

Search for Higgs boson decays in the  
 $H \rightarrow W^+ W^- \rightarrow \ell \nu \ell \nu$  channel via  
vector boson fusion with the  
ATLAS detector at the LHC



Dissertation der Fakultät für Physik  
der  
Ludwig-Maximilians-Universität München

vorgelegt von

**Bonnie Kar Bo Chow**

geboren in Hong Kong

München, den 20.10.2014



Erstgutachter: PD Dr. Johannes Elmsheuser  
Zweitgutachter: Prof. Dr. Wolfgang Dünneweber

Tag der mündlichen Prüfung: 27.11.2014



## Abstract

A search for the evidence of Higgs boson production via vector boson fusion in the  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  decay mode is presented using data recorded with the ATLAS detector from proton-proton collisions at the Large Hadron Collider. The data were taken in 2011 and 2012, at a centre-of-mass energy of  $\sqrt{s} = 7$  TeV and 8 TeV respectively, corresponding to a total integrated luminosity of  $25 \text{ fb}^{-1}$ . A multivariate analysis has been developed using boosted decision trees, in parallel with a cut-based approach which provides a valuable cross-check. For both methods, regions with a large signal presence are identified, with separate control regions used to check the modelling and yields of background processes. The analysis is conducted separately for the 7 TeV and 8 TeV datasets to allow focussed optimisation, before recombining to achieve a final result.

Evidence for the vector boson fusion production in the  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  channel has been established, with an observed significance of 3.3 standard deviations. Assuming the existence of a Standard Model Higgs boson with a mass of 125.36 GeV, the ratio of the measured cross section to that predicted by the Standard Model for vector boson fusion is observed to be consistent with unity:

$$\mu_{obs} = 1.28^{+0.44}_{-0.40} (stat.)^{+0.29}_{-0.20} (syst.)$$

As the vector boson fusion production mode is sensitive to couplings of the Standard Model Higgs boson to the  $W$  and  $Z$  vector bosons, a value for the bosonic coupling has been observed to be consistent with the Standard Model prediction:

$$\kappa_V = 1.04^{+0.10}_{-0.11}$$



## Zusammenfassung

In dieser Arbeit wird die Analyse der Higgs-Boson-Produktion durch Vektor-Boson-Fusion im  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  Zerfallskanal vorgestellt. Die der Untersuchung zugrunde liegenden Proton-Proton-Kollisionsdaten wurden am ATLAS Detektor des Large Hadron Collider in den Laufzeiten der Jahre 2011 und 2012 gesammelt. Die Daten wurden bei einer Schwerpunktsenergie von  $\sqrt{s} = 7$  TeV bzw. 8 TeV genommen und haben eine integrierte Luminosität von  $25 \text{ fb}^{-1}$ . Ein multivariates Analyseverfahren mit Boosted-Decision-Trees wurde entwickelt, wobei parallel eine bewährte schnittbasierte Herangehensweise als Gegenprobe genutzt wurde. Mit beiden Analyseverfahren wurden Regionen mit hoher Signaleffizienz definiert, sowie unabhängige Kontrollregionen, in denen die Modellierung und Beitragsstärke der Hintergrundprozesse überprüft werden konnten. Um spezifische Parameteroptimierungen zu ermöglichen, wurden die 7 TeV- und 8 TeV-Datensätze zunächst separat analysiert, bevor das Gesamtergebnis aus beiden Analysen zusammengeführt wurde.

Higgs-Produktion durch Vektor-Boson-Fusion konnte im Zerfallskanal  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  mit einer beobachteten Signifikanz von 3.3 Standardabweichungen nachgewiesen werden. Unter Annahme eines Standardmodell-Higgs-Bosons mit einer Masse von 125.36 GeV konnte ein mit der Standardmodellvorhersage für Vektor-Boson-Fusion konsistenter Wirkungsquerschnitt gemessen werden. Das Verhältnis des experimentell bestimmten Wertes zur Standardmodellvorhersage beträgt:

$$\mu_{obs} = 1.28^{+0.44}_{-0.40} (stat.)^{+0.29}_{-0.20} (syst.)$$

Aufgrund des Zusammenhangs des Vektor-Boson-Fusion-Produktionsmodus mit der Kopplung des Standardmodell-Higgs-Boson zu den W- und Z-Vektor-Bosonen konnte die Bosonenkopplung im Einklang mit der Standardmodellvorhersage mit folgendem Wert gemessen werden:

$$\kappa_V = 1.04^{+0.10}_{-0.11}$$



# Contents

<b>Introduction</b>	<b>1</b>
<b>1 Theoretical Principles</b>	<b>1</b>
1.1 The Standard Model of Particle Physics . . . . .	1
1.2 Electroweak Interaction Theory and Electroweak Symmetry Breaking . . . . .	3
1.3 Quantum Chromodynamics . . . . .	5
1.4 Higgs Boson Production and Decay . . . . .	8
1.4.1 Discovery of a Higgs-like Boson . . . . .	12
<b>2 Experimental Setup</b>	<b>13</b>
2.1 CERN Accelerator Complex and the LHC . . . . .	13
2.2 The ATLAS Detector . . . . .	14
2.2.1 Inner Detector . . . . .	15
2.2.2 Electromagnetic Calorimeter . . . . .	18
2.2.3 Hadronic Calorimeter . . . . .	19
2.2.4 Muon Spectrometer . . . . .	20
2.2.5 Luminosity Determination . . . . .	21
2.2.6 Data-acquisition and Trigger System . . . . .	22
<b>3 Signal and Background Processes</b>	<b>27</b>
3.1 Signature of Vector Boson Fusion $H \rightarrow WW^{(*)} \rightarrow l\nu l\nu$ . . . . .	27
3.1.1 Associated Production . . . . .	29
3.2 Background Processes . . . . .	29
3.2.1 Gluon-gluon Fusion . . . . .	29
3.2.2 Top quark production . . . . .	30
3.2.3 Diboson production . . . . .	31
3.2.4 Drell-Yan production . . . . .	32
3.2.5 $W$ +jets and QCD Multijet processes . . . . .	33
<b>4 Object Reconstruction</b>	<b>35</b>
4.1 Leptons . . . . .	35
4.1.1 Electrons . . . . .	36
4.1.2 Muons . . . . .	37
4.2 Jets . . . . .	39
4.2.1 $b$ -Jets . . . . .	40

4.3	Missing Transverse Energy . . . . .	41
4.4	Overlap Removal . . . . .	43
<b>5</b>	<b>Boosted Decision Trees</b>	<b>45</b>
5.1	Training and Classifying . . . . .	46
5.1.1	Quantifying the Separation . . . . .	46
5.1.2	Input Variables . . . . .	46
5.1.3	Stop Criteria . . . . .	47
5.1.4	Boosting . . . . .	47
5.1.5	Bagging . . . . .	48
5.2	Evaluation of the Training . . . . .	49
5.2.1	Overtraining . . . . .	49
<b>6</b>	<b>Data and Simulation</b>	<b>51</b>
6.1	Data . . . . .	51
6.1.1	Event Cleaning . . . . .	51
6.1.2	Triggers . . . . .	52
6.2	Monte Carlo Simulation . . . . .	55
6.2.1	Pile-up Reweighting . . . . .	56
<b>7</b>	<b>Observables and Event Selection</b>	<b>63</b>
7.1	Observables . . . . .	63
7.2	Event Preselection . . . . .	66
7.3	Event Selection . . . . .	66
7.3.1	Blinding . . . . .	69
7.3.2	BDT event selection and training . . . . .	69
7.3.3	Cut-based selection . . . . .	74
7.3.4	Modifications for the 7 TeV Analyses . . . . .	78
<b>8</b>	<b>Background Estimations</b>	<b>83</b>
8.1	Calculation of Normalisation, Correction and Extrapolation Factors . . . . .	83
8.2	Sources of Theoretical Uncertainties . . . . .	84
8.3	Top background . . . . .	86
8.4	$WW$ production . . . . .	90
8.5	$Z/\gamma^* \rightarrow ee + \mu\mu$ background . . . . .	93
8.5.1	Non-closure Concerns with the ABCD Method . . . . .	101
8.6	$Z/\gamma^* \rightarrow \tau\tau$ background . . . . .	101
8.7	$W$ +jets and QCD Multijet Backgrounds . . . . .	104
8.8	Modifications for the 7 TeV analyses . . . . .	109
<b>9</b>	<b>Statistical Treatment</b>	<b>111</b>
9.1	Fit Regions . . . . .	111
9.2	Profile Likelihood Procedure . . . . .	112
9.3	$CL_s$ Method . . . . .	115
9.3.1	Test Statistic, p-values and Significance . . . . .	115

9.4	Combination of Results . . . . .	116
<b>10</b>	<b>Systematic Uncertainties</b>	<b>117</b>
10.1	Leptons . . . . .	117
10.1.1	Lepton Efficiency, Energy Scale and Resolution . . . . .	118
10.2	Jets . . . . .	118
10.2.1	Jet Energy Scale and Resolution . . . . .	118
10.2.2	$b$ -jet Identification . . . . .	119
10.2.3	Jet Vertex Fraction . . . . .	119
10.2.4	Pile-up . . . . .	119
10.3	Missing Transverse Energy . . . . .	119
10.4	Luminosity . . . . .	120
10.5	Theoretical Uncertainties . . . . .	120
10.5.1	VBF Higgs Production . . . . .	120
10.5.2	ggF Higgs Background . . . . .	121
10.5.3	Non- $WW$ Diboson Production . . . . .	121
10.6	Final Impact on Results . . . . .	122
<b>11</b>	<b>Results</b>	<b>125</b>
11.1	8 TeV BDT and Cut-based yields . . . . .	125
11.2	7 TeV BDT and Cut-based yields . . . . .	133
11.3	Significance and Interpretation of Results . . . . .	140
11.4	$H \rightarrow WW^{(*)}$ Combination . . . . .	143
<b>12</b>	<b>Summary</b>	<b>145</b>
	<b>Bibliography</b>	<b>155</b>
<b>A</b>	<b>2011 Data Periods</b>	<b>157</b>
<b>B</b>	<b>Cut-based yields for the 7 TeV and 8 TeV analyses</b>	<b>159</b>
	<b>Acknowledgements</b>	<b>163</b>



# Introduction

The concept that all matter is composed of elementary building blocks is not new, and has its roots in the ideas of ancient Greek scientists, who believed that indivisible atoms made up the world as we know it. However, particle physics in a more familiar modern context begins in 1897 with J. J. Thompson's discovery of the electron, the first of the elementary particles to be revealed. Following quickly in 1905, was the interpretation of the photoelectric effect by A. Einstein which led to the recognition of photons as particles in their own right.

Since then, many rapid developments in the 20th century allowed the understanding of particle physics to progress at an astonishing rate. From the prediction of antimatter by P. Dirac in 1928, to the proposal of the quark model by M. Gell-Mann, G. Zweig and Y. Ne'eman, all these accumulated constituents helped to form the modern Standard Model of particle physics. So far, the Standard Model has predicted a wide range of phenomena which have been subsequently confirmed experimentally, and has survived any attempts to find deviations from its predictions.

Within the Standard Model, one of the key ideas is that of electroweak symmetry breaking via the Brout-Englert-Higgs mechanism, resulting in the emergence of a scalar particle referred to as the Higgs boson. As a new particle with the expected properties of a Higgs boson has recently been discovered by the ATLAS and CMS experiments at the Large Hadron Collider located at CERN, it is of vital importance to determine if this is indeed the sought-after particle. No evidence to the contrary has been found yet, and observations of the Higgs boson in all individual decay channels would provide essential confirmations of the Standard Model predictions.

In this thesis, a search for the evidence of Higgs boson production via vector boson fusion in the  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  decay mode is presented. The data used has been collected by the ATLAS detector in 2011 and 2012, in proton-proton collisions with a centre-of-mass energy of 7 TeV and 8 TeV respectively. The thesis begins with the theoretical concepts and motivation for this analysis, and continues with its technical implementation. The results of the search, along with full statistical and systematic uncertainty estimates are presented at the end.



# Chapter 1

## Theoretical Principles

The Standard Model of Particle Physics (SM) provides our current understanding of all known subatomic particle dynamics, mediated by the electromagnetic, weak and strong interactions. The backbone of the SM was developed largely in the 1960s and 1970s, and is constructed from two quantum field theories: electroweak (EW) theory [1, 2, 3] and quantum chromodynamics (QCD) [4, 5]. Since then, it has withstood many rigorous experimental tests and precisely predicted a wide variety of phenomena, which have only served to give further credence to these concepts. This section gives a brief overview of the SM, with special focus on the EW processes and the Brout-Englert-Higgs mechanism [6, 7, 8] by which EW symmetry is broken.

### 1.1 The Standard Model of Particle Physics

Matter in the SM is composed of twelve elementary particles which are the physical manifestations of quantum field excitations [9]. Specifically, these are three generations of eight fermionic fields, where each generation has an up-type quark and down-type quark (in three flavours each), a charged lepton, and a corresponding lepton neutrino. The basic properties of each generation differ only in the masses, such that the mass of a fermion is greater than that of the analogous fermion in a lower generation. Accompanying all the matter particles are their anti-matter counterparts, which are not shown in the summary given in Table 1.1.

Forces are described by twelve bosonic fields: eight coloured gluons ( $g$ ), a massless photon ( $\gamma$ ) and the massive  $W^+$ ,  $W^-$ ,  $Z^0$  bosons which mediate the strong, electromagnetic and weak forces respectively. At higher energies, the electromagnetic and weak forces are unified to form the *electroweak* force. Key properties of these bosons are given in Table 1.2.

Both the EW and QCD quantum field theories are required to be renormalisable and local gauge invariant, in order to remove any unphysical divergences which may arise within the theory. Local gauge invariance is the requirement that the following

	1st generation		2nd generation		3rd generation	
	Flavour	Mass [MeV]	Flavour	Mass [GeV]	Flavour	Mass [GeV]
<b>Quarks</b>	$u$	$2.3^{+0.7}_{-0.5}$	$c$	$1.28 \pm 0.03$	$t$	$173.07 \pm 1.24$
	up		charm		top	
	$d$	$4.8^{+0.5}_{-0.3}$	$s$	$0.10 \pm 0.01$	$b$	$4.18 \pm 0.03$
	down		strange		bottom	
<b>Leptons</b>	$e^-$	0.511	$\mu^-$	0.105	$\tau^-$	1.777
	electron		muon		tau	
	$\nu_e$	$< 2$ eV	$\nu_\mu$	$< 0.2$ MeV	$\nu_\tau$	$< 18.2$ MeV
	electron neutrino		muon neutrino		tau neutrino	

Table 1.1: Elementary fermions of the SM with their masses, sorted by generation. Each up-type and down-type quark occurs in three flavours, referred to as “colours”. [10]

	Interaction	Mass [GeV]
Gluon $g$	Strong	0
Photon $\gamma$	Electromagnetic	0
$W^+, W^-$ bosons	Electroweak	$80.385 \pm 0.015$
$Z^0$ boson	Electroweak	$91.188 \pm 0.002$

Table 1.2: Force mediators of the SM with their masses [10].

transformation leaves the Lagrangian unchanged:

$$\psi \rightarrow \exp(-i\alpha^k(x)\beta^k)\psi \quad (1.1)$$

where  $\alpha^k(x)$  are arbitrary functions of spacetime and  $\beta^k$  are generators of the gauge group operating on the fields  $\psi$ . Noether's theorem implies that each gauge symmetry must be accompanied by a conserved charge [11]. Only massless gauge fields are allowed using this construction, since introducing mass terms causes the theory to be unrenormalisable. Choosing the unitary group  $U(1)$  as a gauge group yields the theory of Quantum Electrodynamics (QED) for the electromagnetic interaction with the conserved electric charge, and choosing the special unitary group  $SU(3)$  leads to QCD for strong interactions, with three conserved 'colour' charges. Both these theories are highly successful in describing their respective effects. However, as the weak interaction is mediated by massive bosons, it is unable to fit directly into this framework. A solution to this problem is provided via the Brout-Englert-Higgs (BEH) mechanism [6, 7, 8], via which gauge bosons can acquire masses when the symmetry of the system is spontaneously broken. Higgs also noted that this mechanism produces a further massive scalar boson, which is commonly referred to as the Higgs boson. Within this scheme, the Lagrangian of the SM is given as:

$$\mathcal{L}_{SM} = \mathcal{L}_{EW} + \mathcal{L}_{Higgs} + \mathcal{L}_{QCD} \quad (1.2)$$

with each contribution to the Lagrangian discussed briefly in the following sections. Further details on the quantum field theory treatment can be found in [9].

## 1.2 Electroweak Interaction Theory and Electroweak Symmetry Breaking

The unified theory of electromagnetic and weak interactions result from a  $SU(2) \times U(1)$  gauge group. Experimentally, it has been shown that only *left-handed* field components participate in the weak interaction generated by  $SU(2)$ , which has a Lagrangian of  $V - A$  form (vector minus axial vector) [9]. The left-handed fermions are thus arranged in doublets with respect to *weak isospin*  $I$  (the charge under  $SU(2)$ ) and the right-handed fermions in weak isospin singlets, which do not transform under  $SU(2)$ . Transformations under  $SU(2)$  involve three massless vector boson fields:  $W_1^\mu$ ,  $W_2^\mu$  and  $W_3^\mu$ . A fourth massless vector boson  $B$  is a consequence of  $U(1)$  transformations with respect to the *weak hypercharge*,  $Y$ . Both left and right-handed field components participate equally in electromagnetic interactions generated by  $U(1)$ . The Gell-Mann-Nishijima formula [12, 13] relates the weak hypercharge, the third component of weak isospin and electric charge of a particle  $Q$  in the following way:

$$Q = I_3 + \frac{Y}{2} \quad (1.3)$$

where  $I_3$  is the third component of the weak isospin. A summary table of fermions and their associated weak isospin and hypercharge are given in Table 1.3. The EW term

in the SM Lagrangian and the associated covariant derivative  $D$  and field tensors are given as follows:

$$\begin{aligned}
\mathcal{L}_{EW} &= \sum_k i\bar{\psi}_k \gamma_\mu D^\mu \psi_k - \frac{1}{4} F_j^{\mu\nu} F_{\mu\nu}^j - \frac{1}{4} B^{\mu\nu} B_{\mu\nu} \\
D_\mu &= \partial_\mu - ig \frac{I}{2} W_\mu - ig' \frac{Y}{2} B_\mu \\
F_{\mu\nu}^j &= \partial_\mu W_\nu^j - \partial_\nu W_\mu^j + g\epsilon_{kl}^j W_\mu^k W_\nu^l \\
B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu
\end{aligned} \tag{1.4}$$

where  $g, g'$  are coupling constants and  $\gamma_\mu$  are the Dirac matrices. The term  $\sum_k i\bar{\psi}_k \gamma_\mu D^\mu \psi_k$  is the kinetic term for the fermions, whereas  $\frac{1}{4} F_j^{\mu\nu} F_{\mu\nu}^j$  and  $\frac{1}{4} B^{\mu\nu} B_{\mu\nu}$  describe the interaction between the  $W$  and  $B$  particles.

	<b>Q</b>	<b>Y</b>	<b>I<sub>3</sub></b>		<b>Q</b>	<b>Y</b>	<b>I<sub>3</sub></b>
$u_L$	2/3	1/3	1/2	$u_R$	2/3	4/3	0
$d_L$	-1/3	1/3	-1/2	$d_R$	-1/3	-2/3	0
$e_L^-$	-1	-1	-1/2	$e_R^-$	-1	-2	0
$\nu_L$	0	-1	1/2	$\nu_R$	0	0	0

Table 1.3: Left and right-handed fermions and their associated weak isospin and hypercharge. As all fermion generations have the same charges, only the 1st generation has been detailed here [10].

As mentioned previously, massless gauge fields are required using this framework. However as massive vectors bosons have been observed, the  $SU(2) \times U(1)$  symmetry is therefore not exact and must be spontaneously broken. This is achieved by using the BEH mechanism, which introduces an additional complex scalar doublet field  $\phi$  and imposes a quartic potential  $V$ . The scalar field breaks the EW symmetry spontaneously and is also transformed in the process, acquiring a vacuum expectation value,  $v$ , at the minimum of the potential. This extra contribution to the SM Lagrangian results in:

$$\mathcal{L}_{Higgs} = |D_\mu \phi|^2 - V(\phi) - \sum_k f_k \bar{\psi}_k \phi \psi_k \tag{1.5}$$

$$V(\phi) = -\mu^2 \phi^\dagger \phi + \frac{\lambda}{2} (\phi^\dagger \phi)^2 \tag{1.6}$$

The term  $\sum_k f_k \bar{\psi}_k \phi \psi_k$  sums over Yukawa contributions which are used to add mass to fermions for every field  $k$ . Since  $\langle \phi \rangle = v$ , these mass terms have mass  $v \cdot f_k$ , where  $f_k$  is a free parameter which can be tuned to the rest mass measurement of the particular fermion. These mass terms are additionally also interaction terms with the Higgs boson with coupling strength proportional to the fermion mass. The parameters  $\mu, \lambda$  are required to be real due to unitarity and are chosen such that  $v$  is non-zero

- thus requiring additionally that  $\lambda > 0$  and  $\mu^2 > 0$ . Consequently, the potential  $V$  has the form as illustrated in Figure 1.1, with a maximum at zero and a minimum at  $\phi^\dagger\phi = \frac{\mu^2}{\lambda} =: \frac{v^2}{2}$ . By expanding  $\phi$  about the  $v$  value at an arbitrary point in the minimum, four real fields are obtained. Three of these form massless Goldstone bosons [14, 15], which are then absorbed into the originally massless gauge fields. These three absorbed degrees of freedom form the third longitudinal polarisation component needed to make such fields massive - thus giving mass to the  $W^+$ ,  $W^-$ ,  $Z^0$  bosons. As a result of the symmetry breaking, the following mass eigenstates are obtained:

$$\begin{aligned}
W_\mu^\pm &= \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) & m_W &= g\frac{v}{2} \\
Z_\mu &= \frac{1}{\sqrt{g^2 + g'^2}}(gW_\mu^3 - g'B_\mu) & m_Z &= \frac{v}{2}\sqrt{g^2 + g'^2} \\
A_\mu &= \frac{1}{\sqrt{g^2 + g'^2}}(g'W_\mu^3 + gB_\mu) & m_A &= 0
\end{aligned} \tag{1.7}$$

The coupling constants define the *weak-mixing* or *Weinberg* angle  $\theta_W$ , and also relate it to the electric charge  $e$ :

$$\begin{aligned}
\theta_W &= \arctan(g/g') \\
e &= g \sin \theta_W
\end{aligned} \tag{1.8}$$

It can be seen from the above formulae that the masses  $m_W$  and  $m_Z$  are connected via  $\theta_W$ . The  $W$  bosons couple to left-handed fields only with a strength proportional to  $g$ . In contrast, the  $Z$  boson has unequal couplings to left and right-handed states. There is a coupling to left-handed fields with a strength also proportional to  $g$ , and additionally a coupling proportional to  $Q$  which acts equally on left and right-handed components. The final mass eigenstate  $A_\mu$  is massless, as the combination  $Q = I_3 + \frac{Y}{2}$  remains unbroken by this mechanism. Since  $Q$  is the electromagnetic charge,  $A_\mu$  is identified as the photon, with the expected coupling to fermions according to their electromagnetic charge. Feynman diagrams of allowed EW interaction vertices involving fermions, or bosons are illustrated in Figure 1.2.

This mechanism also produces an additional massive scalar particle from the remaining field, the Higgs boson, with a mass of  $m_H = \sqrt{2\lambda}v = \sqrt{2}\mu$ . Within the SM,  $m_H$  is a parameter to be measured, since  $\lambda$  does not feature in couplings between the Higgs boson and massive particles. Feynman diagrams of allowed Higgs interaction vertices can be seen in Figure 1.3.

### 1.3 Quantum Chromodynamics

For the theory of QCD, the  $SU(3)$  gauge group is chosen in order to describe the strong interaction between elementary particles. The three charges associated with  $SU(3)$  are referred to as “colour” in the context of QCD, and particles which transform with respect to this factor of the gauge group are said to be “coloured”. Each of the six quark

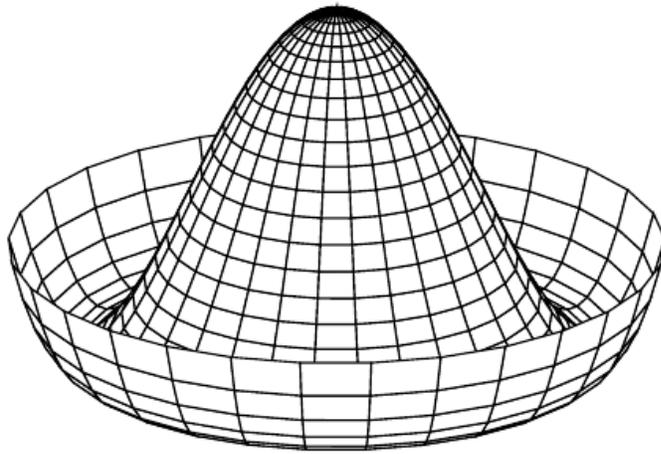


Figure 1.1: The “mexican hat” potential of the Higgs field.

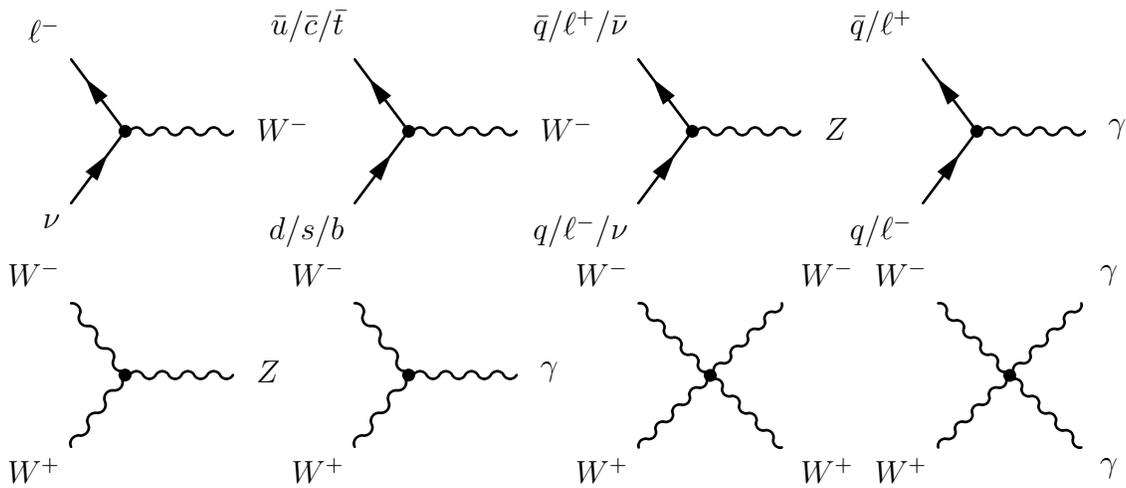


Figure 1.2: Electroweak Feynman diagrams of possible interaction vertices containing fermions, or bosons. Here,  $\ell = e, \mu, \tau$ .

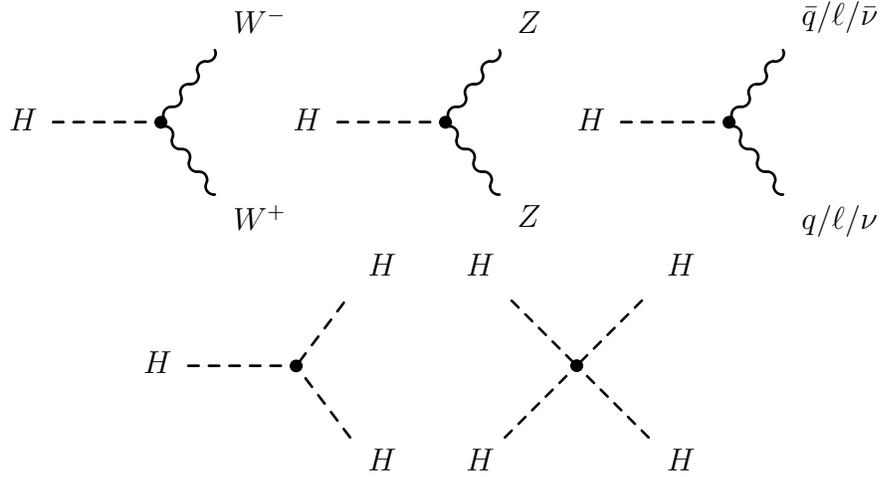


Figure 1.3: Feynman diagrams of possible Higgs boson interaction vertices.

flavours exist in a colour-triplet of three differently coloured fields, which are referred to as “red”, “green” and “blue” for convenience. Gluons exist in eight different colour combinations, where each gluon carries a colour and anti-colour charge. The actual colour choice of a particle is arbitrary as colour doesn’t break the  $SU(3)$  symmetry, and can be rotated around. Contributing to  $\mathcal{L}_{QCD}$  are the eight gluon fields  $G_\mu^\alpha$  which are generated by the  $SU(3)$  symmetry and fermionic quark fields  $q_i$ , together with the strong coupling constant  $g_s$ :

$$\begin{aligned}
 \mathcal{L}_{QCD} &= \sum_j i\bar{q}_j \gamma_\mu (\partial^\mu - ig_s G^\mu) q_j - \frac{1}{2} \text{tr} G_{\mu\nu} G^{\mu\nu} \\
 G_{\mu\nu} &= \partial_\mu G_\nu - \partial_\nu G_\mu - ig_s [G_\mu, G_\nu] \\
 G_\mu &= \sum_{\alpha=1}^8 G_\mu^\alpha \frac{\lambda^\alpha}{2}
 \end{aligned} \tag{1.9}$$

The first term  $\sum_j i\bar{q}_j \gamma_\mu (\partial^\mu - ig_s G^\mu) q_j$ , gives the interaction between quarks and gluons, with the latter term  $\frac{1}{2} \text{tr} G_{\mu\nu} G^{\mu\nu}$  describing gluon-gluon self-interaction. Feynman diagrams illustrating the possible QCD vertices are shown in Figure 1.4.

An important feature of QCD is the phenomenon of *asymptotic freedom* which results from the self-interaction of the gluons. This leads to decreasing  $g_s$  with increasing energies (or shorter distances) or in other words, at high energies, quarks and gluons behave as free particles, allowing the use of perturbation theory and Feynman calculus in this regime. Another property of QCD is *confinement*. This is also due to the self-interaction of the gluons and requires that all naturally occurring particles must be “colourless”, such that either the total amount of each colour is zero, or that there is an equal sum of all colours present. Thus, individual quarks and gluons can never be isolated on large scales, with quarks existing only in combination as mesons

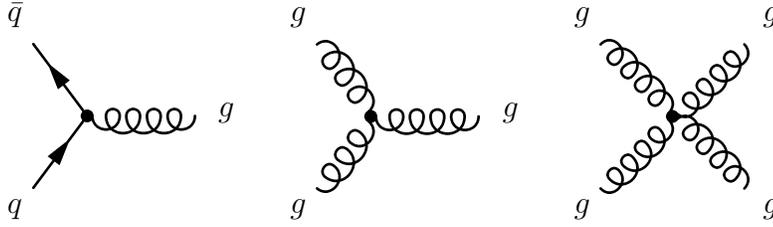


Figure 1.4: Feynman diagrams of allowed QCD vertices

and (anti-)baryons. The further away the quarks or gluons move from each other, the greater the potential energy becomes, until sufficient energy is present to produce a new quark-antiquark pair, which recombine to form other hadrons. Via this hadronisation process, a collimated shower of particles is produced from the initial partons, which is referred to as a *jet*.

## 1.4 Higgs Boson Production and Decay

Within the SM, Higgs production can happen via the following channels:

- ***Gluon-gluon fusion*** - two gluons combine to form a quark loop, which then radiates a Higgs boson. As the Higgs coupling is proportional to the quark mass, the dominant contribution derives from  $t$ -quark loops.
- ***Vector boson fusion*** - two fermions (or anti-fermions) interact and radiate a virtual vector boson each, which collide to form a Higgs boson
- ***Associated production*** - the collision of a fermion and anti-fermion, producing a virtual  $W$  or  $Z$  vector boson which in turn emits a Higgs boson if the vector boson has sufficient energy
- ***Top quark fusion*** - requires two gluons to each decay into a heavy quark-antiquark pair. A quark and antiquark from each pair then combine to produce a Higgs boson.

Feynman diagrams showing these various production modes are illustrated in Figure 1.5. Summary plots of the production modes and their expected cross sections at the Large Hadron Collider (LHC), at a centre-of-mass energy of 7 TeV, and 8 TeV can be seen in Figure 1.6). For instance, given a SM Higgs boson mass  $m_H = 125$  GeV, the expected cross section for VBF is  $\sim 2$  pb at a centre-of-mass energy of 8 TeV. Since the LHC is a proton-proton collider, it is most probable that two gluons collide - thus gluon-gluon fusion (ggF) is expected to be the dominant production mode at the LHC. Vector boson fusion (VBF) is the second greatest process, since the colliding fermions (or anti-fermions) do not need to be of the same flavour. The associated production mode has a much smaller cross section than either ggF or VBF, due to the lower probability of achieving a fermion and anti-fermion collision. Top quark fusion is by far the

smallest production method, since extremely high energy gluons are required, in order for them to be able to emit a heavy quark-antiquark pair each.

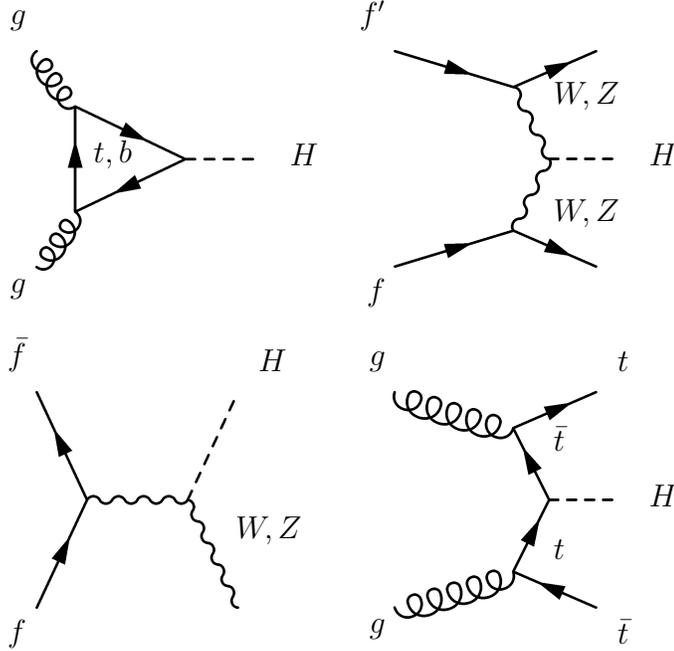


Figure 1.5: Feynman diagram of Higgs production modes

In order to identify the produced Higgs boson however, one has to consider its decay products. For the decay of the Higgs boson, an overview of the possible decay channels can be seen in Figure 1.7, expressed as *branching ratios* which give the probability of such a decay. Viable decay options are listed below:

- Fermion-antifermion pair
- Pair of massive gauge bosons ( $W$  or  $Z$ )
- Pair of massless gauge bosons ( $g$  or  $\gamma$ ) via an intermediate loop of heavy quarks or massive gauge bosons

The branching ratio of the decay of a Higgs boson into a pair of  $W$  bosons is comparatively large over the  $m_H$  range, especially for  $m_H > 2m_W$  where it is the dominant decay channel, making it an attractive search option. Each  $W$  boson can decay either into a quark and anti-quark, or into a charged lepton and the corresponding lepton neutrino. It is however simpler to focus only on the  $H \rightarrow WW^{(*)} \rightarrow l\nu l\nu$  process where both  $W$  bosons decay leptonically, as the dilepton requirement is able to suppress many background processes which contain two jets. Due to the presence of the neutrinos, the mass resolution of this channel is too broad to be able to reconstruct  $m_H$ , although it is sensitive to the presence of a Higgs boson over a range of masses.

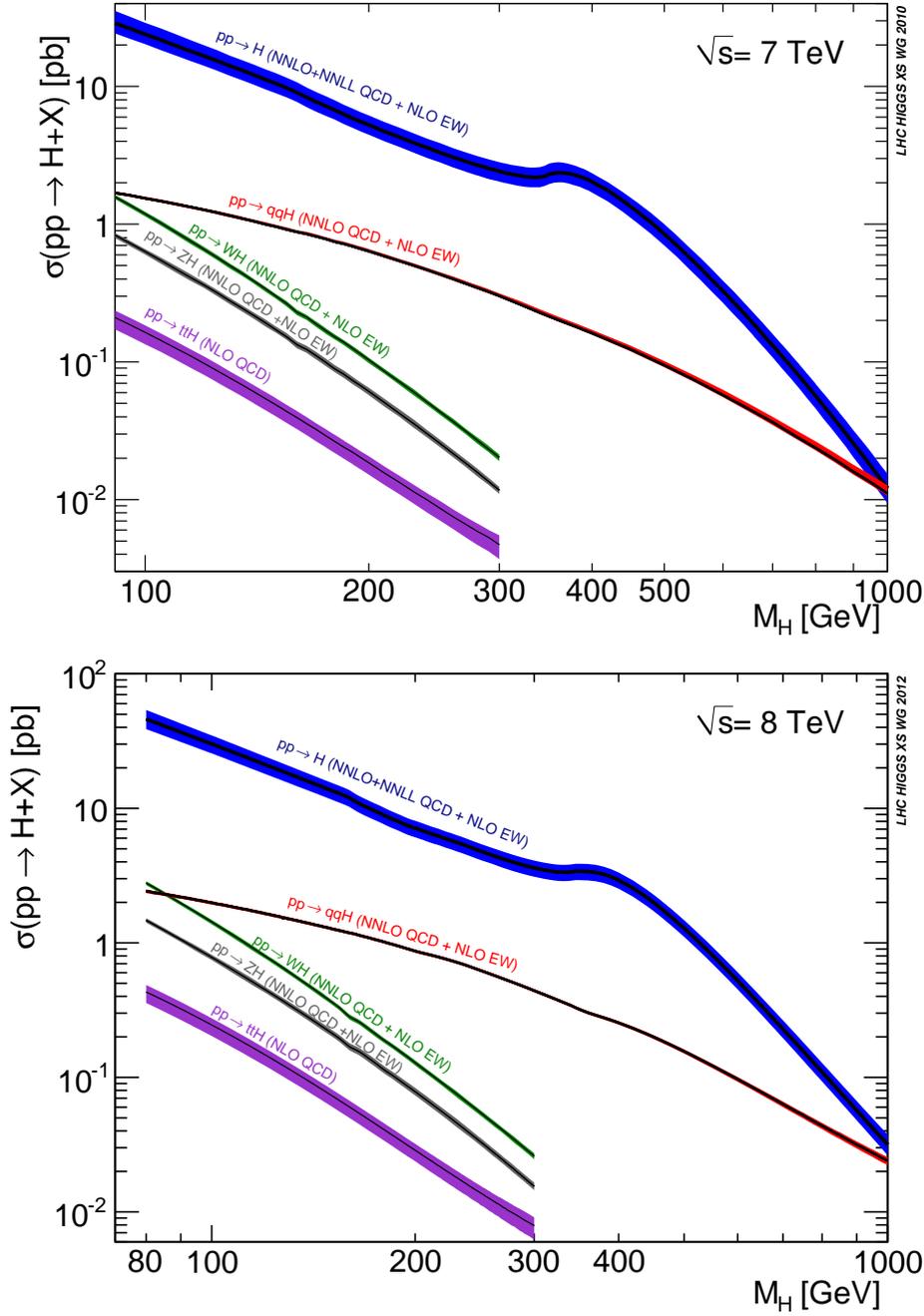


Figure 1.6: SM Higgs boson production cross sections at centre-of-mass energy  $\sqrt{s} = 7$  TeV and  $\sqrt{s} = 8$  TeV as a function of the Higgs mass  $m_H$  [16]. ggF (blue) and VBF (red) can be seen to be the most dominant production modes followed by associated production, separated into production with a  $W$  boson (green) or a  $Z$  boson (grey). It can be seen easily that top quark fusion (purple) processes are an order of magnitude lower. The bands around each line represent an uncertainty of one standard deviation.

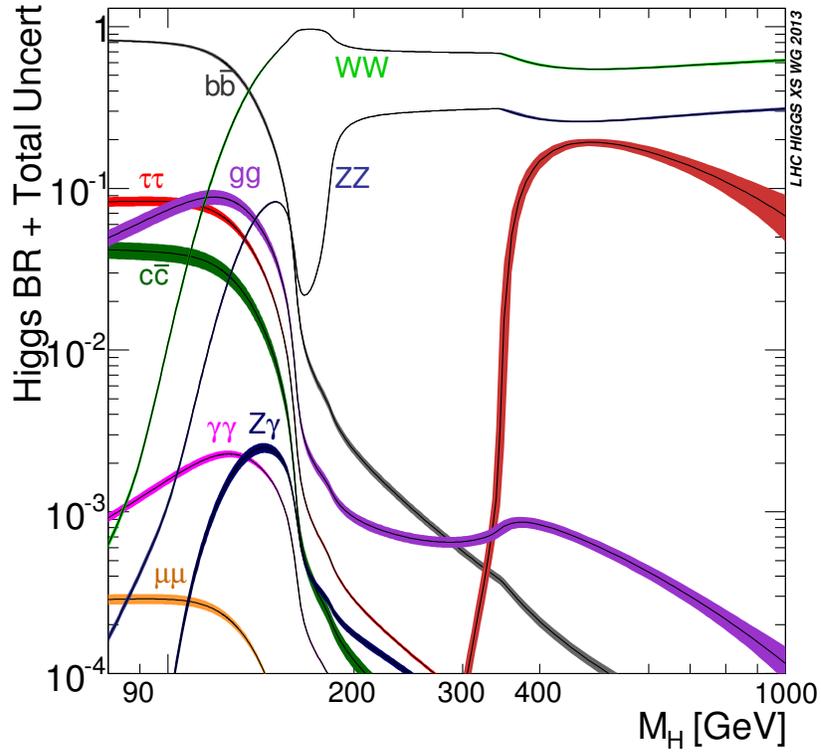


Figure 1.7: SM Higgs boson decay branching ratios as a function of the Higgs boson mass  $m_H$  [16]. The line representing the decay of a Higgs boson into a pair of  $W$  bosons (green) is shown at the top of the plot, becoming dominant for for  $m_H > 2m_W$ . The bands around each line represent an uncertainty of one standard deviation.

The measurement of  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  production via the VBF and associated production processes are sensitive to couplings of the Higgs boson to  $W$  and  $Z$  bosons, while the ggF process probes the couplings to both heavy quarks and  $W$  bosons. Hence within the  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  search channel in ATLAS, further divisions are made according to the production process, allowing the SM coupling predictions to be more thoroughly tested. This thesis concerns only the VBF production mode, although the latest combination of results involving all the current  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  searches will be included at the end.

### 1.4.1 Discovery of a Higgs-like Boson

A new particle with the expected spin and gauge boson couplings of a SM Higgs boson has recently been discovered by the ATLAS [17, 18, 19, 20] and CMS [21] experiments at the LHC, using the  $ZZ \rightarrow 4\ell$ ,  $\gamma\gamma$  and  $WW^*$  decay channels. Measurements of the particle's mass [22, 23] and also of its decays to fermions at the LHC [24, 25] are consistent with the properties of a SM Higgs boson. The most recent measurement of the Higgs boson mass yields  $m_H = 125.36 \pm 0.41$  GeV, corresponding to the central value of the ATLAS measurement in the  $ZZ \rightarrow 4\ell$  and  $\gamma\gamma$  decay modes [26]. Searches for the  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  decay where  $\ell$  is an electron or muon, have provided the first direct exclusions of the SM Higgs boson at a hadron collider [27, 28, 29, 30, 31]. Subsequent  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  measurements determined the couplings and spin of the discovered particle to increasing precision, for instance the spin-0 nature of the Higgs boson [20], with positive parity being strongly preferred. Evaluation of the properties of the discovered particle are crucial in understanding the SM predictions, and key to this is its observed cross section via VBF production and its coupling to the  $W$  and  $Z$  bosons.

# Chapter 2

## Experimental Setup

The Large Hadron Collider (LHC) is a synchrotron-type accelerator and is the world's most powerful hadron collider built to date. Located at the European Organization for Nuclear Research (CERN) at the Swiss-French border, it is also the largest collider in the world, with a diameter of 27km. It allows for the acceleration of two counter-rotating bunched beams of protons, which are provided by the CERN accelerator complex at an injection energy of 450 GeV, up to a design energy of 7TeV. In addition, the LHC is also able to collide beams of heavy ions, which could allow researchers to answer questions about the early Universe and properties of quark-gluon plasma.

The two proton beams are brought to collision at four points in the LHC ring, coinciding with the four installed experiments: ATLAS (A Toroidal LHC ApparatuS), CMS (Compact Muon Solenoid), ALICE (A Large Ion Collider Experiment) and LHCb (LHC beauty). Of these, ATLAS and CMS are general purpose detectors, with various target areas of interest including Higgs physics, Supersymmetry, Exotics and precision (Standard Model) measurements. ALICE and LHCb are both special purpose experiments, designed to study heavy ion physics and  $b$ -hadron physics respectively. This Chapter provides a description of the experimental setup used to obtain the data which has been analysed in this thesis.

### 2.1 CERN Accelerator Complex and the LHC

The accelerator complex is simply a succession of machines, which accelerate particles to increasingly higher energy (an overview of the entire chain can be seen in Figure 2.1). Described in this section are only the areas relevant to producing proton beams.

A bottle of hydrogen gas forms the proton source, with the resulting proton beam then injected consecutively into Linac 2, the Proton Synchrotron Booster, the Proton Synchrotron and then the Super Proton Synchrotron. This sequence of events accelerates the protons to an energy of 50 MeV, 1.4 GeV, 25 GeV and 450 GeV respectively. The last link in this chain is the LHC which brings the protons to their maximum energy. For

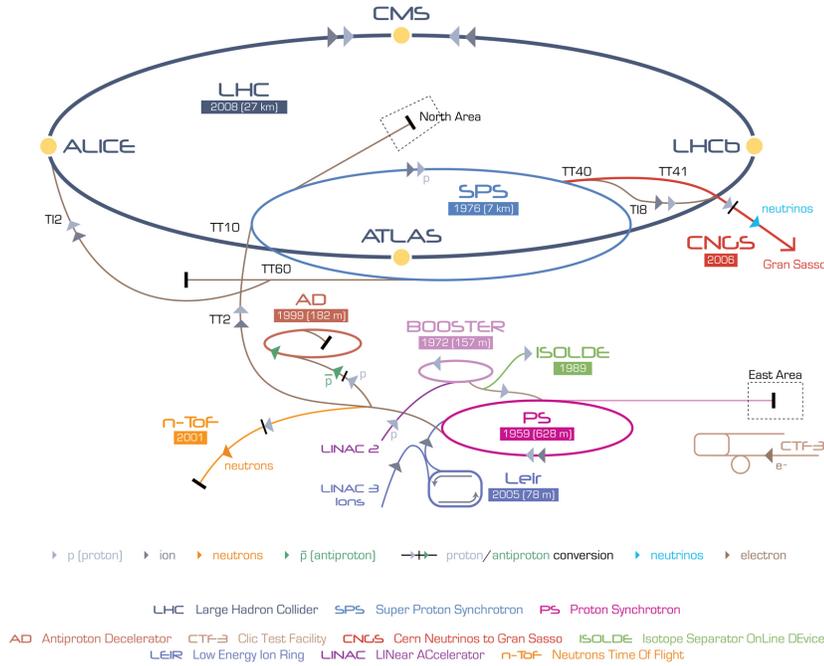


Figure 2.1: Overview schematic of the CERN accelerator complex [32]

data taken during 2012, this was restricted to an energy of 4 TeV, whereas during 2011, it was limited to 3.5 TeV- in other words, a centre-of-mass energy of 8 TeV and 7 TeV respectively. This decision was made following an incident on the 19th of September 2008, which revealed that additional safety measures would be needed before running at the full design energy of 14 TeV. These have largely been implemented already in the current 2013-2014 shutdown phase of the LHC. Once the beams have entered the LHC ring, they are kept in a circular trajectory by 1232 superconducting dipole magnets, while various other superconducting magnets focus the beam. The protons within the beam are grouped in 'bunches', with the possibility to vary the time interval between each bunch, known as the *bunch-spacing*. The bunch-spacing was predominantly 50 ns during the 2011 and also 2012 data-taking periods [33]. Typically, every bunch contains  $\mathcal{O}(10^{11})$  protons. Each intersection of one bunch with another is referred to as a *bunch-crossing*, during which multiple interactions can take place.

## 2.2 The ATLAS Detector

Situated on the LHC ring, the ATLAS detector is a working tribute to many aspects of Physics and Engineering. It consists of a number of subdetectors, namely an inner tracking detector; electromagnetic calorimeter, hadronic calorimeter and a large external muon spectrometer. In addition, three superconducting magnet systems are integrated into the detector to curve any charged particle tracks. These consist of a solenoid around the inner detector, and an air-core barrel and end-cap toroids within the muon spectrometer. An overview can be seen in Figure 2.2 of all the subsystems,

the details of which are described in this section [34, 35].

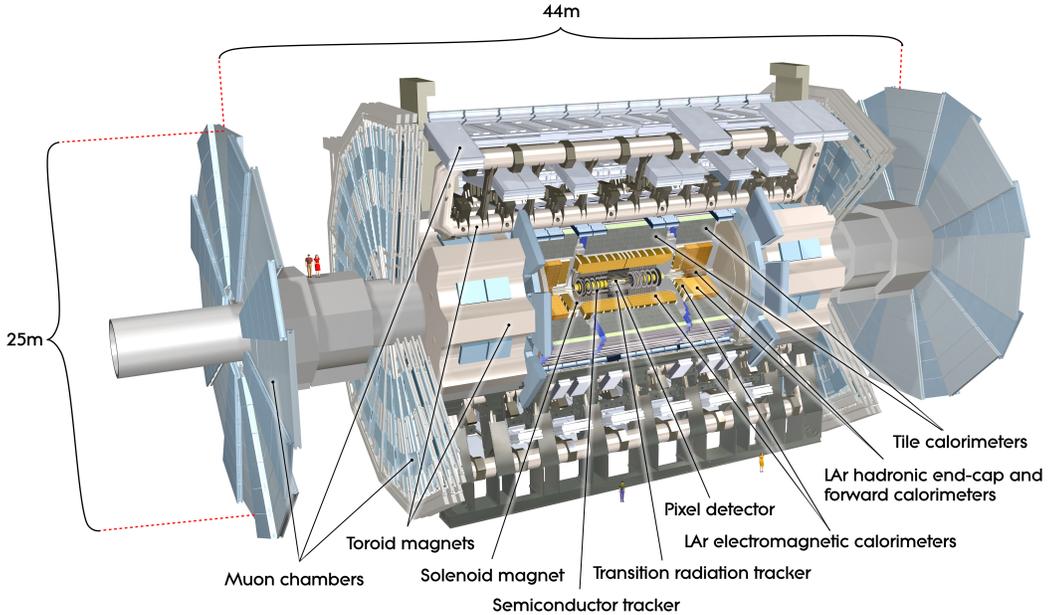


Figure 2.2: ATLAS Detector [36]

A right-handed coordinate system is used for the ATLAS experiment, with the origin located in the centre of the detector, at the nominal interaction point. The  $x$ -axis points from the interaction point to the centre of the LHC ring, the  $y$ -axis upwards, and the  $z$ -axis along the beam line. Cylindrical coordinates are used in the  $x$ - $y$  plane, transverse to the beam direction. The azimuthal angle around the beam line is denoted by  $\phi$ . The pseudorapidity is defined as  $\eta \equiv -\ln \tan \frac{\theta}{2}$ , where  $\theta$  is the polar angle to the beam axis.

### 2.2.1 Inner Detector

The purpose of the inner detector is to reliably provide high resolution positional information, such as particle tracks, impact parameters and vertex information. In order to achieve this, the inner detector is composed of three subsystems: the pixel detector, the semiconductor tracker and the transition radiation tracker, arranged cylindrically around the beam axis. Each of these covers a pseudorapidity range of  $|\eta| < 2.5$  and are optimised to provide different information, depending on its location relative to the interaction point. A superconducting solenoid magnet surrounds the system, providing a 2 T magnetic field. All subsystems are constructed with high irradiation in mind, in order to withstand the harsh conditions during operation, and to maintain a high operational efficiency. An overview can be seen in Figure 2.3 and a cross-section of the inner detector subsystems is illustrated in Figure 2.4, with the relative sizes and radial

positions included. The inner detector has a total length of about 7 m and extends radially to 115 cm, where it is met by the cryostat of the electromagnetic calorimeter [37].

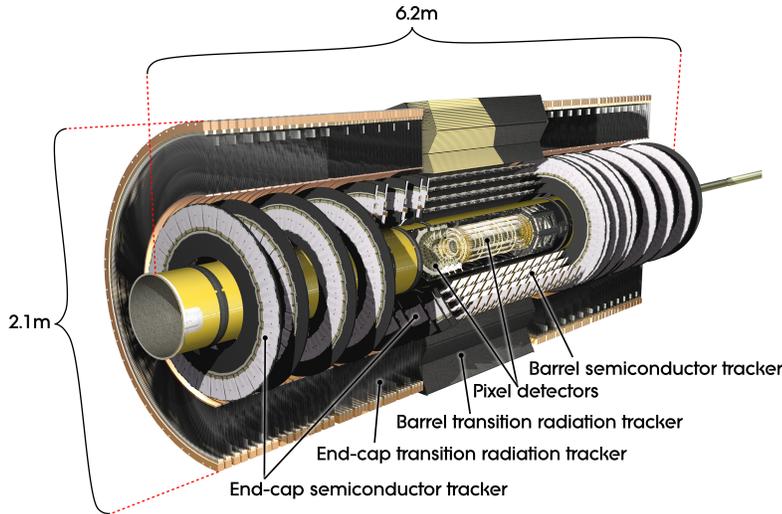


Figure 2.3: ATLAS Inner Detector [36]

## Pixel Detector

The pixel detector is designed to provide measurements of extremely high granularity and precision, as close to the interaction point as possible. It enables the reconstruction of primary vertices even in the presence of multiple interactions, and of secondary interaction vertices which are crucial in order to identify jets originating from long-lived particles such as  $b$ -hadrons. The detector consists of three 'barrel' layers of silicon sensors, arranged concentrically at radii of 5, 9 and 12 cm from the beam axis, and six 'end-cap' discs of sensors, located at  $\pm 50$ ,  $\pm 58$  and  $\pm 65$  cm along the  $z$ -axis. Typically, each track crosses three layers of pixels due to this arrangement. The sensor layers and discs together contain 1456 and 288 pixel modules respectively, with each module comprising 46080 pixel elements. In total, the pixel detector consists of around 80 million pixels, covering an area of  $1.7 \text{ m}^2$ . The pixel detector is capable of a  $\sim 99.9\%$  efficiency with a spatial resolution of  $10 \mu\text{m}$  in the transverse direction and  $115 \mu\text{m}$  along the beam axis [38].

## Semiconductor Tracker

Arranged around the pixel detector, are four nested barrel layers of silicon microstrip sensors which are located at radii of approximately 30, 37, 44 and 51 cm from the  $z$ -axis. In addition, there are also two end-cap modules, containing nine sensor discs each, situated in a range  $\pm(85 - 273)$  cm from the interaction point along the beam axis. The layout of the detector ensures that charged particle tracks pass through at

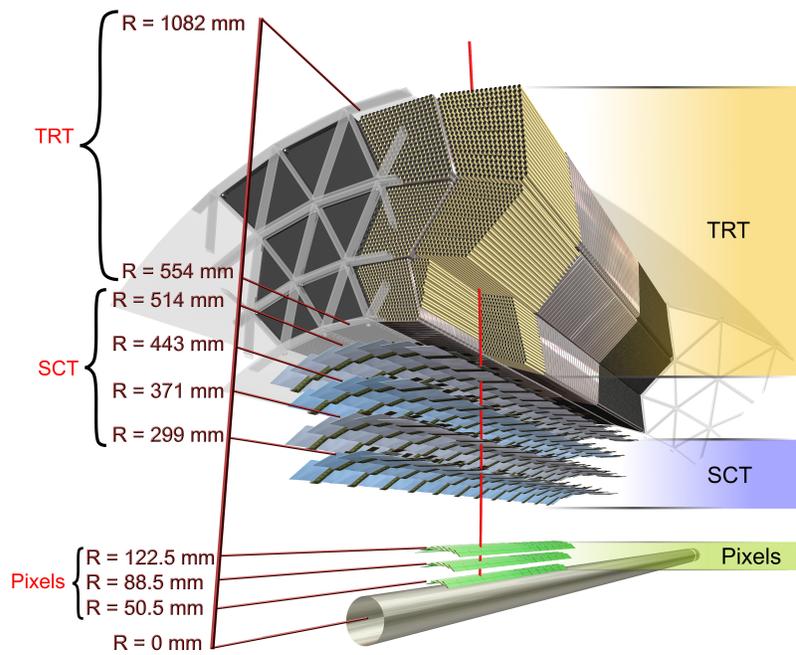


Figure 2.4: Particle tracking in the Inner Detector [36]

least four layers everywhere in the acceptance region. Sensor modules on the barrel and end-caps have a similar construction, with the main difference being that the end-cap modules are tapered (instead of rectangular) in shape so as to provide complete coverage. The semiconductor tracker (SCT) consists of 2122 and 1976 modules in the barrel and the end-caps - giving rise to  $61 \text{ m}^2$  of active silicon sensors. The greater active surface area of the SCT was one of the motivations in using strips of sensors here, rather than the more expensive pixel sensors. It has maintained an extremely high working efficiency throughout the data-taking periods in 2010-13, with  $\sim 99\%$  of the channels remaining operational. The SCT provides a spatial resolution of  $17\mu\text{m}$  in the transverse direction and  $580\mu\text{m}$  along the  $z$ -axis. Tracks are distinguishable if separated by more than  $200\mu\text{m}$  [39, 40].

### Transition Radiation Tracker

Surrounding the semiconductor tracker is the transition radiation tracker (TRT), which provides tracking and electron identification information [41]. The transition radiation measured assists in distinguishing electrons and pions, since such radiation is sensitive to the energy-mass ratio  $E/m$ . It makes use of thin-walled proportional drift tube detectors which are arranged in a barrel layer with two end-caps, in a similar fashion as the other two subdetectors which make up the inner detector. Such tubes were chosen as they are easily arranged in modules and can also be integrated into a Xenon based gas mixture with which the tracker is operated. The TRT provides continuous tracking in the acceptance range and comprises 298304 tubes, with the tubes arranged such that particles cross 35-40 tubes in the barrel region. The spatial resolution achieved by the tubes is  $130\mu\text{m}$ .

### 2.2.2 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) is a finely segmented lead/liquid-argon (LAr) detector [42, 43], where the lead electrodes are arranged with the accordion geometry shown in Figure 2.5. Together with LAr as the active material and in the presence of an electric field, ionization produced by electromagnetically interacting particles as they traverse the detector can be measured. From such particle showers, the energy deposited by the incoming particle can be determined.

The calorimeter is divided into a barrel section which covers a range of  $|\eta| < 1.475$  and two end-caps which cover  $1.375 < |\eta| < 3.2$ . Within the transition region between the barrel and the end-caps, particle identification is unreliable and thus particles entering this region are rejected for analysis. The barrel component is divided into three layers with increasingly rough granularities as the radial distance from the  $z$ -axis increases. The innermost layer has a granularity of  $\Delta\eta \times \Delta\phi = 0.003 \times 0.1$ , allowing  $\gamma/\pi_0$  and  $e/\pi$  discrimination, while the second and third layers have  $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$  and  $\Delta\eta \times \Delta\phi = 0.05 \times 0.025$  respectively. In combination, the hits from each layer allow the reconstruction of the electromagnetic shower shape, as well

as the position. The ECAL resolution<sup>1</sup> in both the barrel and the end-caps is given by  $\frac{\Delta E}{E} = \frac{10\%}{\sqrt{E[\text{GeV}]}} \oplus \frac{170 \text{ MeV}}{E[\text{MeV}]} \oplus 0.7\%$  [43]. The entire LAr calorimeter is enveloped by a cryostat to impose a temperature uniformity, as any fluctuations impact greatly the energy measurement.

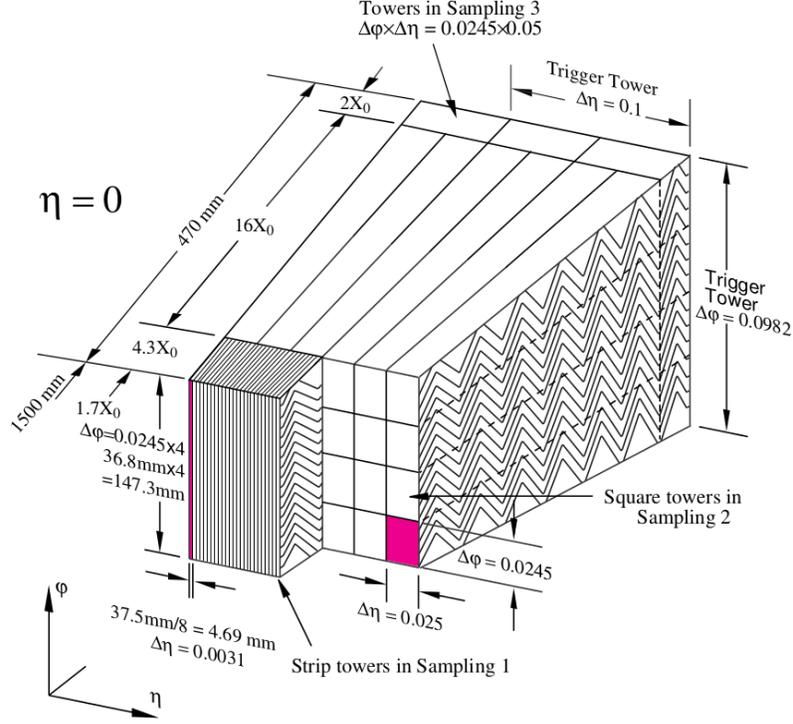


Figure 2.5: Sketch of structure within ECAL [42]

### 2.2.3 Hadronic Calorimeter

As the name suggests, the hadronic calorimeter (HCAL) [35, 42] measures the energy of particles which interact via the strong force, primarily that of hadrons. It consists of a number of highly segmented sections which are optimised to the varying radiation and detection requirements over the large pseudorapidity range. Due to the large active barrel area of the calorimeter, the material choice was made largely with cost in mind.

For  $|\eta| < 1.7$ , a tile calorimeter is used, with three sets of barrel sections, each containing three layers. Plastic scintillator tiles are used as the active material and steel as the absorber. Strongly interacting particles produce light on interaction with

<sup>1</sup>The calorimeter energy resolution is given as  $\frac{\Delta E}{E} = \frac{a}{\sqrt{E[\text{GeV}]}} \oplus \frac{b}{E[\text{MeV}]} \oplus c$ , where  $a$  represents the stochastic term,  $b$  the electronic noise and  $c$  is a constant which includes effects of detector instabilities and any miscalibration.

the scintillator, the amount of which is related to the energy deposited. The light is then read out by wavelength shifting fibres into photo-multiplier tubes for detection. The innermost layer of each barrel section has a granularity of  $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ , whereas the second and third layers have both  $\Delta\eta \times \Delta\phi = 0.2 \times 0.1$ . The resolution of the tile calorimeter is measured to be  $\frac{\Delta E}{E} = \frac{50\%}{\sqrt{E[\text{GeV}]}} \oplus 3\%$ , with negligible electronic noise [44, 45].

Two additional sets of LAr end-cap hadronic calorimeters are used to cover higher pseudorapidity regions. These are known as the hadronic end-cap calorimeter (HEC) and the forward calorimeter (FCAL), and they extend from  $1.5 < |\eta| < 3.2$  and  $3.1 < |\eta| < 4.9$  respectively. In contrast to the lead absorbers within the ECAL, copper is used as the absorbing material in the HEC with the pseudorapidity range  $1.5 < |\eta| < 2.5$  having  $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ , and  $2.5 < |\eta| < 3.2$  having  $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ . A combination of copper and tungsten absorbers are used in the FCAL, with a granularity  $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ . The resolution of the end-cap hadronic calorimeters is measured to be  $\frac{\Delta E}{E} = \frac{100\%}{\sqrt{E[\text{GeV}]}} \oplus 10\%$ , again with negligible electronic noise [34]. Conveniently, the HEC and FCAL are contained within the same cryostat as that of the ECAL end-cap discs. An overview of the entire calorimeter system can be seen in Figure 2.6.

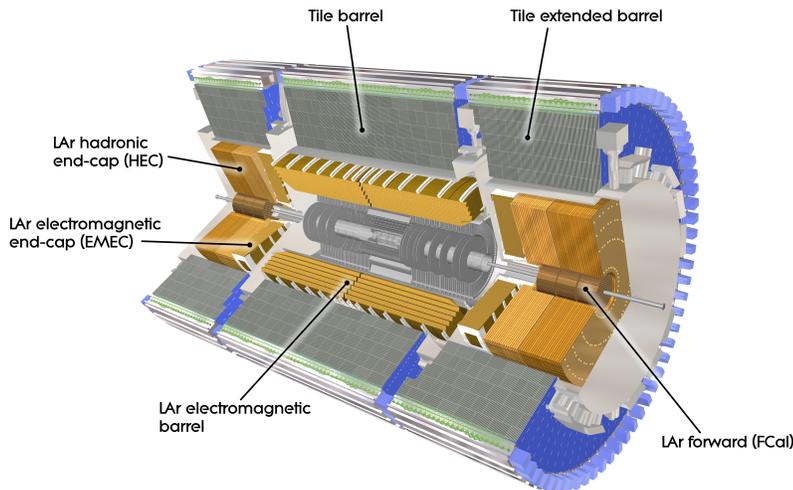


Figure 2.6: Overview of electromagnetic and hadronic calorimeters [36]

## 2.2.4 Muon Spectrometer

Due to their lack of strong interactions and relatively large mass (thus suppressing bremsstrahlung), muons are usually able to escape through all the inner detector and the calorimeter subdetectors. The muon spectrometer (MS) is designed to trigger on

and provide tracking for such muons using a variety of components [35, 46]. The muon tracks are deflected using a system of superconducting magnets, which enable their momentum to be determined from the track curvature. It is designed to have a combined  $p_T$  resolution of 10% for tracks with energies of 1 TeV [34] and to cope with any ageing effects induced by the harsh radiation conditions - all components were tested to withstand at least five times the expected radiation levels. A cut-away overview of the complete MS system is illustrated in Figure 2.7.

Within the barrel region are three concentrically arranged layers of monitored drift tube (MDT) chambers and resistive plate chambers (RPC) at radii of about 5, 7.5 and 10 m from the  $z$ -axis. These cover a pseudorapidity range  $|\eta| < 1$  and provide quick triggering signals, and precision tracking information with a resolution of  $80\mu\text{m}$  and 10 mm respectively, along the beam direction [34].

To round off the trigger system, three layers of thin gap chambers (TGC) are installed in the end-caps of the MS, which cover the pseudorapidity range  $1.05 < |\eta| < 2.7$ . Besides the muon triggers provided, the trigger system also enables the identification of bunch crossings during runtime. Additional MDT layers are also installed in the end-caps to provide tracking at larger distances and pseudorapidities. Nearer to the interaction point and beam line, for  $2 < |\eta| < 2.7$ , tracking is provided by the cathode strip chambers (CSC) which are better able to endure the high rate of charged particles here. The resolution possible using the CSC is measured to be  $60\mu\text{m}$  per CSC layer [46].

Three superconducting air-core toroid magnets are used to provide magnetic deflection - one larger barrel magnet integrated into the MDT and RPC system ( $|\eta| < 1.1$ ), and two smaller end-cap pieces located between the CSC and the other outer end-cap subdetectors ( $1.1 < |\eta| < 2.7$ ). The combined magnet configuration delivers a magnetic field which is largely orthogonal to the muon trajectories.

## 2.2.5 Luminosity Determination

In the context of proton beam collisions, the luminosity  $L$  describes the number density of proton encounters per beam cross-sectional area per time. The rate of a process is  $f = \sigma \times L$ , where the cross-section  $\sigma$  of a process is the size of the effective area that a proton needs to hit in order for this process to occur. As the ATLAS experiment is run over a long period of time, the quantity of interest is the *integrated luminosity*  $\int L dt$  which can then be multiplied by a cross-section to yield the number of events expected in the entire time period for a particular process. In terms of the accelerator parameters, the luminosity is determined in ATLAS [47] using the following equation:

$$L = \frac{n_b f_r n_1 n_2}{2\pi \Sigma_x \Sigma_y} \quad (2.1)$$

where  $n_b$  is the number of bunch pairs colliding per revolution,  $f_r$  the known revolution frequency of the LHC,  $n_1$  and  $n_2$  the number of protons in beam 1 and 2 respectively,

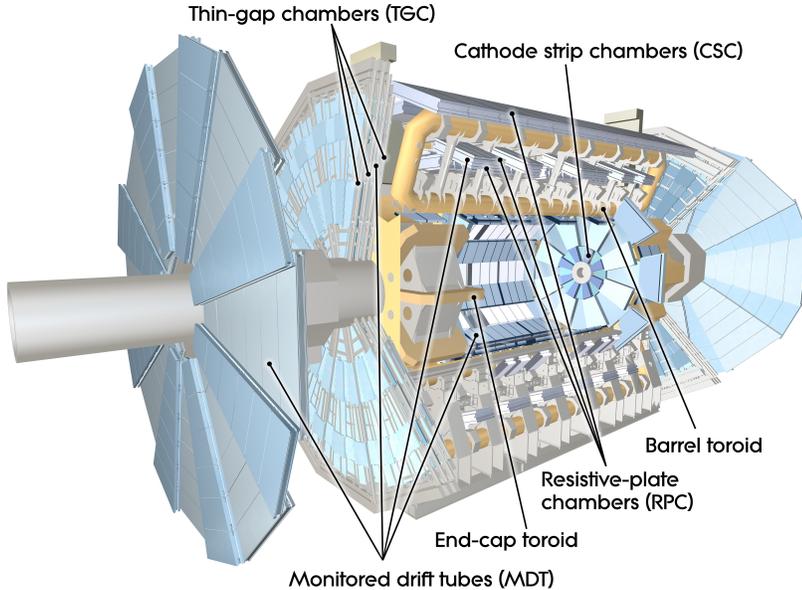


Figure 2.7: Illustration of the muon spectrometer system [36].

and  $\Sigma_x$  and  $\Sigma_y$  the Gaussian widths of the bunches in the directions perpendicular to that of the beam.

Measurements of  $\Sigma_x$  and  $\Sigma_y$  are conducted using *van der Meer scans*, in which the event rate is measured whilst one of the two beams is displaced stepwise with respect to the other in the  $x$  ( $y$ ) direction [48, 49].  $n_b$ , and  $n_1 n_2$  (the bunch population product), are determined from an external analysis of beam currents in the LHC.

The total integrated luminosity delivered to the ATLAS experiment in 2012 was  $22.8 \text{ fb}^{-1}$ , with  $20.3 \text{ fb}^{-1}$  delivered under stable beam conditions. The corresponding values for data taken during 2011, are  $5.5 \text{ fb}^{-1}$ , and  $4.5 \text{ fb}^{-1}$  respectively [50]. A summary of these total integrated luminosities during the 2011 and 2012 data-taking periods is shown in Figure 2.8.

## 2.2.6 Data-acquisition and Trigger System

Due to the high rate of proton-proton collisions and the amount of information associated with each event<sup>2</sup>, it is simply not feasible to record all the events which occur. For instance during data-taking in 2012, the proton-proton collision frequency was 20 MHz, with 21 interactions occurring simultaneously on average [50]. The mean number of interactions per crossing for both data-taking periods to date is illustrated in Figure 2.9.

---

<sup>2</sup>An *event* refers to a triggered bunch crossing as recorded by the detector.

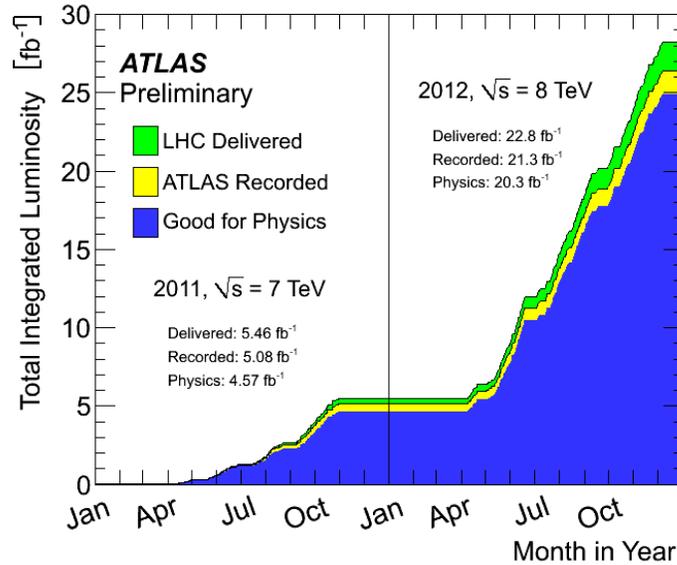


Figure 2.8: Cumulative luminosity versus time delivered to (green), recorded by ATLAS (yellow), and certified to be good quality data taken during stable beam conditions (blue). Shown for proton-proton collision data taken during 2011 with a centre-of-mass energy  $\sqrt{s} = 7$  TeV, and during 2012 with  $\sqrt{s} = 8$  TeV. [50].

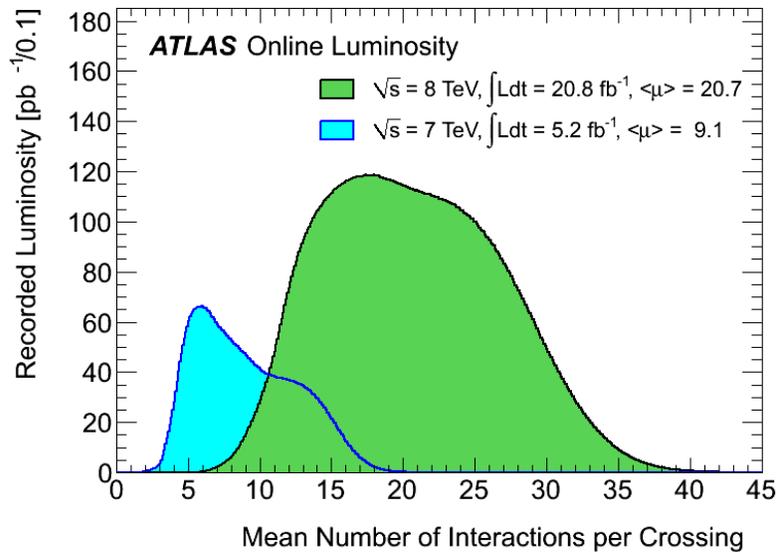


Figure 2.9: Mean number of interactions per crossing  $\mu$ , with respect to the recorded luminosity. Shown for proton-proton collision data taken during 2011 with a centre-of-mass energy  $\sqrt{s} = 7$  TeV, and during 2012 with  $\sqrt{s} = 8$  TeV. The integrated luminosities and  $\langle\mu\rangle$  are given in the Figure [50].

ATLAS employs a trigger and data-acquisition (DAQ) system based on three levels of online event selection to reduce this to a manageable rate, whilst retaining events which contain interesting physics at high efficiency. Figure 2.10 gives a simplified overview of the trigger/DAQ system, which is described in the following subsection in greater detail.

The level-1 (L1) trigger makes an initial hardware-based selection using reduced information from various subdetectors, thus bringing the event rate to 75 kHz. For instance, events with high transverse momentum ( $p_T$ ) muons can be identified using the RPC and TGC systems in the MS, or those containing electrons/photons/hadronic decays selected for with the ECAL and HCAL. Other such candidate objects include taus which decay hadronically, as well as large total  $p_T$  and missing energies, the latter of which implies the existence of neutrinos or other undetectable particles. The L1 trigger decision is based on combinations of object signatures in coincidence or veto, and is highly flexible in its choice of selection. Due to the short bunch-crossing intervals, the L1 trigger has to quickly identify which objects belong together in an event, and also to decide whether it is kept and passed on to the next level.

At the next stage, the level-2 (L2) trigger makes use of 'region-of-interest' (RoI) information from the L1 trigger. This incorporates positional information of the candidate objects, with full precision and granularity available if needed. Additional  $p_T$  and tracking requirements can be placed on candidates to reduce the rate. For example, candidate electrons should have a high  $p_T$  track in the inner detector which is distinct from other tracks, with additional discriminating power from the transition radiation signature. In this way, the rate is brought down to roughly 3.5 kHz, depending on the object requirements.

The last level of triggers is the event filter (EF), which employs offline algorithms and methods, calibrated to the detector conditions during runtime. The rejection power of the EF (compared to L2) derives from more refined algorithms, some of which are not feasible at L2 due to the processing time. Both the L2 and EF are software-based and are collectively referred to as the High Level Trigger (HLT). The rate is reduced now to approximately 500 Hz and this final selection of events are then written to mass storage for offline analysis.

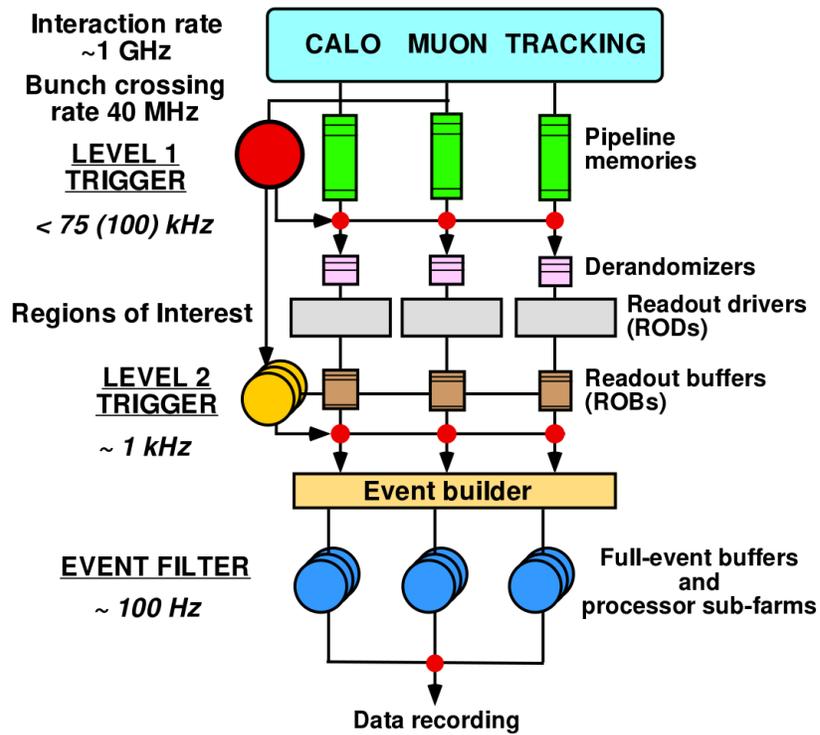


Figure 2.10: Diagram representing the ATLAS trigger/DAQ system [35].



# Chapter 3

## Signal and Background Processes

In order to proceed with a measurement, one must first understand the processes involved in order to take advantage of their properties. Here in this Chapter, the details of the VBF  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  signal are discussed, along with the background processes which will be encountered.

### 3.1 Signature of Vector Boson Fusion

$$H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$$

Although VBF is expected to be the second dominant Higgs production mode at the LHC, it has a very specific kinematic signature which can be exploited in searches. The two interacting quarks will be seen as two highly energetic jets at high rapidities, in other words, relatively close to the beam axis. The vector bosons ( $W$  or  $Z$ ) then fuse to form a Higgs boson, which decays with a relatively large branching ratio to a pair of excited  $W$  bosons.

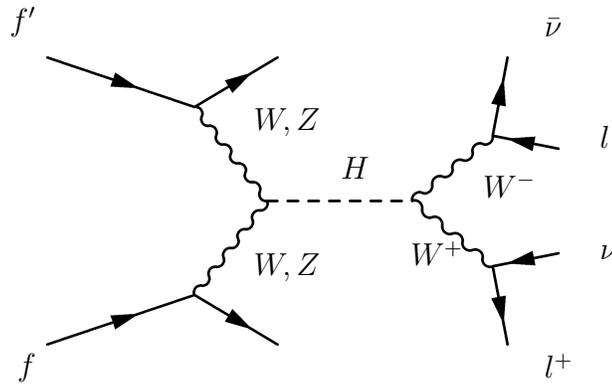


Figure 3.1: Feynman diagram of Higgs production via VBF, decaying into a pair of excited  $W$  bosons which decay leptonically in turn

$W$  bosons decay leptonically with equal branching ratios to electrons, muons and taus, and their corresponding neutrinos. By focussing only on the leptonic decays of the  $W$  bosons, many background processes can be rejected as they do not contain the two required leptons in the final state. The neutrinos which arise from the leptonic  $W$  decays cannot be detected directly and appear as missing energies in the detector - their presence is inferred from a measured momentum imbalance in the transverse plane of the detector. As taus are unstable and decay into lighter particles before detection is feasible in the ATLAS detector, the selection is made only for electrons and muons. Just over a third of taus produced can decay further into lighter leptons and their corresponding neutrinos, providing final states which contain also two leptons and missing energies - these situations are included implicitly in the event selection. The Feynman diagram of the nominal VBF signal process can be seen in Figure 3.1.

Due to the assumption that the Higgs is a spin-0 particle, and the  $V - A$  structure of the weak interaction, the charged leptons tend to be emitted in the same direction, with the two neutrinos produced in the opposite direction to balance the dilepton system. This alignment of the leptons originating from the  $H \rightarrow WW^{(*)}$  decay can be seen in Figure 3.2. The charged leptons and neutrinos are produced within the rapidity range spanned by the two energetic jets, with little additional hadronic activity expected here as the mediating weak bosons do not exchange colour.

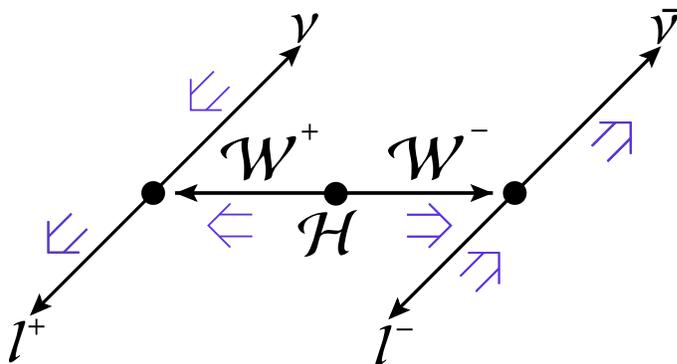


Figure 3.2: Illustration of the  $H \rightarrow WW^{(*)} \rightarrow l\nu l\nu$  decay, showing the directions of particle motion with the thin arrows (black), and the spin projections with the double arrows (purple).

The large numbers of correlations between the various final state particles can be best exploited using multivariate techniques such as Boosted Decision Trees (BDT). A traditional “cut-based” analysis, in which requirements are imposed on observables to obtain a signal-enriched region, was also developed. In order to corroborate the BDT results, the cut-based analysis is used independently alongside the BDT analysis as a cross-check.

### 3.1.1 Associated Production

The measurement of  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  production via the VBF and associated production processes are sensitive to couplings of the Higgs boson to  $W$  and  $Z$  bosons and are combined together for such coupling measurements. Although this thesis concerns only the VBF production mode, with all optimisation procedures targeting this signal, the inclusion of associated Higgs production was made to check the sensitivity of the analysis for this process too. The topologies of the associated Higgs production modes  $WH$  and  $ZH$  are very different to that of VBF and contain also extra leptons and jets from the weak boson decays. It is thus essentially negligible, also due to its lower production cross section, but is incorporated for completeness. The corresponding Feynman diagram for this production mode can be seen in Figure 3.3.

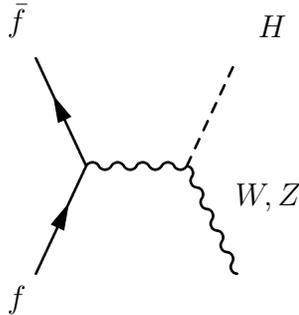


Figure 3.3: Feynman diagram of Higgs production via associated production.

## 3.2 Background Processes

Despite the fact that VBF  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  has a distinctive signature, there is a sizable and diverse background consisting of ggF Higgs production, top quark production, diboson production,  $Z/\gamma^* \rightarrow \ell\ell$ ,  $W$ +jets, and QCD multijet processes. Within the analysis, the “same-flavour” and “different-flavour” final states are considered, in other words, having either  $ee + \mu\mu$  or  $e\mu + \mu e$  as the final state leptons. For the  $e\mu + \mu e$  channel, the “first” lepton is that with the higher transverse energy. The splitting of the analysis is done in order to target backgrounds such as  $Z/\gamma^* \rightarrow \ell\ell$  which are lepton-flavour sensitive. The  $e\mu + \mu e$  channel carries the bulk of the sensitivity as it can be better differentiated from the background processes.

### 3.2.1 Gluon-gluon Fusion

Although the ggF Higgs process produces predominantly Higgs bosons with less than two jets, higher jet multiplicity final states are possible. Such events arise from higher order ggF processes, in which additional radiation is generated from incoming gluons. Clearly if this Higgs boson then decays leptonically via a pair of  $W$  bosons, the signature is extremely similar to that of VBF. However, the jets which originate in this

way have a different topology to those of the VBF production and will generally be of lower energy, making it easy to distinguish from VBF produced Higgs bosons. The Feynman diagram for ggF can be seen in Figure 3.4.

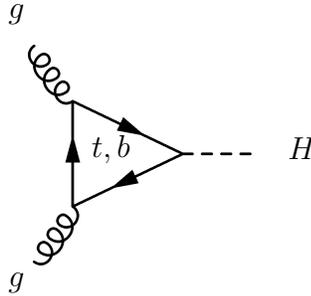


Figure 3.4: Feynman diagram of Higgs production via ggF.

### 3.2.2 Top quark production

Top quarks are principally pair produced in hadronic collisions via either gluon fusion, or the annihilation of a quark and antiquark. These background processes are referred to collectively as  $t\bar{t}$  backgrounds. Single top quarks can also be produced via the weak interaction. In both cases, further jets can be produced due to gluon radiation from incoming partons. As the top quark decays almost exclusively to a  $W$  boson and  $b$ -quark, both  $t\bar{t}$  and single-top production are considered backgrounds as the  $b$ -jets (jets originating from  $b$ -quarks) from top decays mimic the jets seen in VBF production.  $t\bar{t}$  processes in particular are a problem due to the two  $b$ -jets produced, with a much higher production cross section than that of VBF Higgs production. Top backgrounds can be effectively reduced by identifying  $b$ -jets by their kinematic properties and rejecting events in which they are present. Possible  $t\bar{t}$  and single-top diagrams are shown in Figures 3.5 and 3.6.

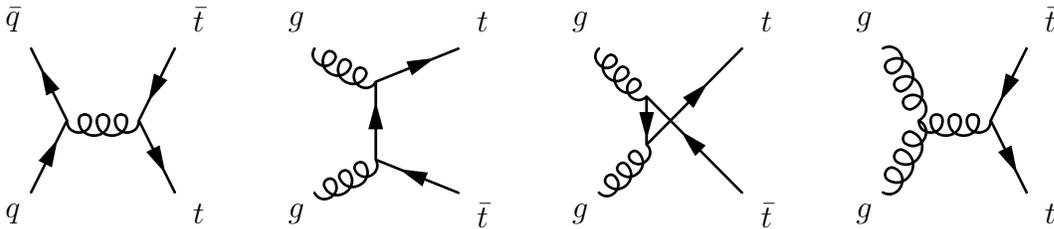


Figure 3.5: Pair production of top quarks via quark-antiquark annihilation, or gluon fusion.

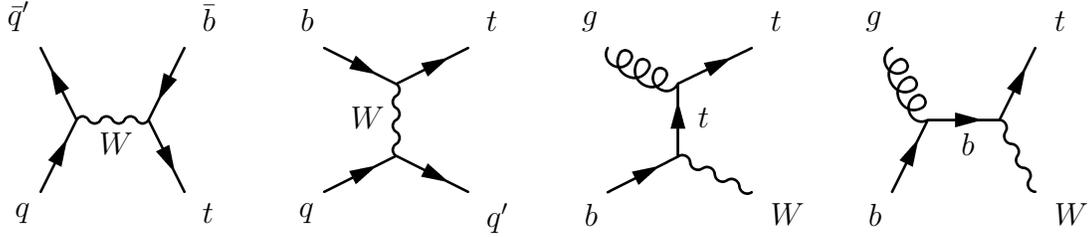


Figure 3.6: Feynman diagrams of singly-produced top quarks

### 3.2.3 Diboson production

Of the diboson production processes, the most challenging background is the pair production of  $W$  bosons where the  $W$  bosons decay leptonically. It has exactly the same final state, with similar kinematics, although the opening angle between the charged leptons can be used to characterise the process - the  $WW$  process tends to have a larger separation between the charged leptons than those from the VBF Higgs signal. The invariant mass of the charged leptons tends also to be greater for the  $WW$  process, due to the relatively low Higgs boson mass ( $m_H < 2m_W$ ).  $WW$  can occur from quark-antiquark initial states or gluon-gluon initial states as shown in Figure 3.7.

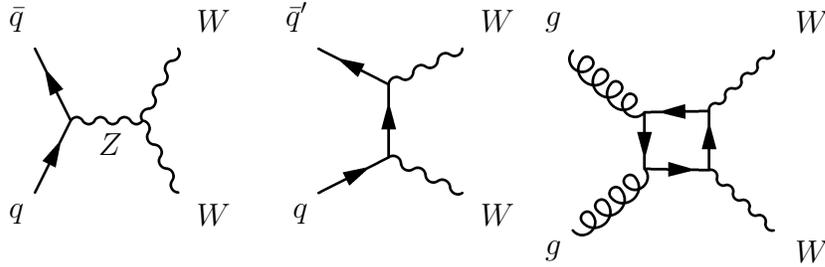


Figure 3.7: Pair-produced  $W$  bosons from quark-antiquark or gluon-gluon initial states.

Other diboson processes include the production of  $WZ$ ,  $ZZ$ ,  $Z\gamma$ ,  $W\gamma$  and  $W\gamma^*$ , in which the bosons decay leptonically. All of these however have a much smaller cross section than that of  $WW$  pair production, and are illustrated in Figure 3.8. The  $W\gamma$  and  $Z\gamma$  backgrounds enter the signal region if the photon converts into an  $e^+e^-$  pair within the detector material. In a similar fashion, the  $W\gamma^*$  becomes a background when the virtual photon produces a charged lepton pair. Additional jets in diboson production processes can be generated radiatively from incoming partons.

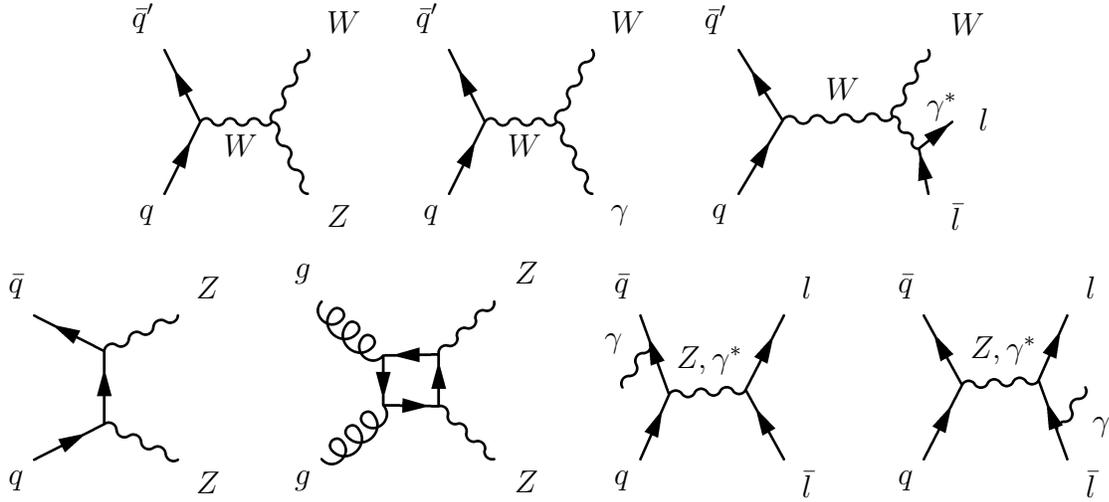


Figure 3.8: Feynman diagrams of non- $WW$  diboson backgrounds.  $WZ$ ,  $W\gamma$  and  $W\gamma^*$  processes are shown in the first row, with  $ZZ$  and  $Z\gamma$  processes in the second.

### 3.2.4 Drell-Yan production

The Drell-Yan (DY) background or  $Z/\gamma^* \rightarrow \ell\ell$  process is an energetic quark-antiquark annihilation producing a  $Z$  boson or a virtual photon, with subsequent decay into a pair of oppositely charged leptons. Extra jets are radiated from the high energy incoming quarks or antiquarks. In all cases as the leptons are of the same flavour,  $Z/\gamma^* \rightarrow \ell\ell$  is largely a problem of the  $ee + \mu\mu$  channel. A distinction is made between  $Z/\gamma^* \rightarrow ee + \mu\mu$  and  $Z/\gamma^* \rightarrow \tau\tau$  processes, as only the latter has “true” neutrinos resulting from the decay of the tau into lighter leptons. However, missing energies can still be present due to neutrinos originating from heavy quark decays (from the associated jets), and also detector resolution effects. The DY process is presented in Figure 3.9.

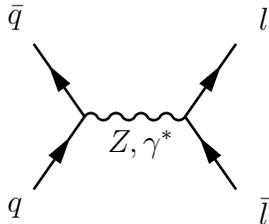


Figure 3.9: The  $Z/\gamma^* \rightarrow \ell\ell$  process with subsequent decay into a pair of oppositely charged leptons.

### 3.2.5 $W$ +jets and QCD Multijet processes

$W$ +jets processes encompass all those in which a  $W$  boson, which decays further leptonically, is produced in association with one or more jets. These processes constitute backgrounds when a jet, or an object radiated from a jet is misidentified as a prompt lepton, thus appearing to satisfy the lepton multiplicity requirement. This category also includes the production of non-prompt leptons, arising from the decay of hadrons containing a heavy quark. The QCD background refers to multijet production, in which two of the jets are misidentified as the required prompt leptons. As the rate of misidentification is low and the fact that the presence of large missing transverse energies is unlikely, both these backgrounds have small cross sections compared to the other processes presented.



# Chapter 4

## Object Reconstruction

The  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  analysis uses a common set of “physics objects”, which must be defined and reconstructed using detector information, for all its sub-analyses in order to facilitate later combination of results. This Chapter details the objects used for the  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  analysis - namely electrons, muons, jets and missing transverse energy. The procedure applied when these objects overlap with one another in the detector are also described. Differences between the reconstruction criteria between the 7 TeV and 8 TeV analyses are detailed in addition where appropriate.

### 4.1 Leptons

Based on the information from all the differing sub-detector systems in ATLAS, it is possible to classify a set of tracks, or energy deposits as an electron or muon. One of the main issues with accurate lepton reconstruction however, is the production of non-prompt leptons, for instance via decays within hadronic jets, or via a photon conversion in the detector material. A number of methods can be used to suppress such leptons, such as requiring that the leptons are *isolated*. This isolation is defined using information from both the calorimeter and tracking systems. The calorimeter isolation is based on the sum of the transverse energies of the surrounding energy deposits within a given cone in the  $\eta - \phi$  space with respect to the lepton, where cone sizes are given by  $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ . Track isolation is based similarly on the sum of the transverse momenta from all tracks within a given cone size around the lepton, excluding the lepton track itself. Typically, the isolation criteria are expressed as a percentage of the lepton transverse energy  $E_T$  or transverse momentum  $p_T$ .

The primary vertex of an event is selected to be the vertex with the largest value of  $\sum(p_T)^2$ , where the sum is over all the tracks associated to that particular vertex. To further suppress backgrounds from non-prompt leptons, cuts on impact parameters are applied to ensure that the leptons originate from the primary vertex. These include restrictions on the significance of the transverse impact parameter, defined as the ratio of the measured transverse impact parameter  $d_0$  to its estimated uncertainty  $\sigma_{d_0}$ , and the longitudinal impact parameter  $z_0$ . These impact parameters are defined with re-

spect to the primary vertex. Further reconstruction details particular to electrons and muons are discussed below.

### 4.1.1 Electrons

Within the ATLAS detector, electron reconstruction originates from energy deposits in the ECAL, which are then matched with reconstructed charged particle tracks in the inner detector. Electron identification (ID) is limited to the range  $|\eta| < 2.47$  due to the inner detector acceptance, excluding the  $1.37 < |\eta| < 1.52$  region as this is the transition region between the end of the barrel calorimeter and the start of the endcap. The  $E_T$  of an electron is computed from the calorimeter-based energy and the track direction at the interaction point.

Differing cuts on calorimeter and tracking variables, such as the shower shape, hadronic leakage<sup>1</sup> or track-cluster matching, are then applied to sort the objects - into groups of “Loose”, “Medium” and “Tight” and so on - according to their increasing quality [51]. Requirements on the track quality, based on the number of hits in various subdetector systems, are also imposed on the electron candidates. Such ID conditions are useful in helping to distinguish true electron candidates from fake electrons which can arise for example, from jets. The ID quality is improved at successive levels by tightening the cut values of the variables which make up the ID criteria or by adding more discriminant variables. Due to the high number of interactions per bunch crossing, the ID criteria were reoptimised for data taken during 2011, resulting in the new “Loose++”, “Medium++” and “Tight++” selections with improved electron selection efficiencies.

Variations to the quality criteria exist in order to allow greater flexibility in analysis optimisation - one of these is the use of a MVA likelihood technique, denoted by a suffix “LH” which stands for the “likelihood”. Enhanced background rejection is possible with MVA techniques due to the combined evaluation of several properties during the selection decision. The LH ID working points use largely the same discriminant variables as the corresponding cut-based ID method, and are designed to correspond roughly to each other - for instance, the “Very Tight Likelihood” (VTLH) ID matches approximately the electron efficiency of the “Tight” cut-based selection, except with a better rejection of light-flavour jets and photon conversions.

For electrons with  $10 < E_T < 25$  GeV, the VTLH ID is found to be optimal as it provides better background rejection, especially against  $W$ +jets processes, than other criteria without sacrificing signal efficiency. However for electrons with  $E_T > 25$  GeV, such misidentification backgrounds are less important and hence the “Medium++” ID is sufficient. In addition to the “Medium++” ID, electron candidates are rejected if

---

<sup>1</sup>Deposits in the HCAL are considered too, as energetic leptons can leave deposits in the HCAL, resulting in *hadronic leakage*. This is measured in a region of  $0.2 \times 0.2$  in  $\eta \times \phi$  space around the particle direction behind the ECAL cluster corresponding to the particle.

their associated track is part of a conversion vertex or does not have a hit in the innermost layer of the pixel detector (the B-layer), in order to reduce electrons from photon conversions. The ID criteria used for 7 TeV electrons is the “Tight++” cut-based criteria, over the entire  $E_T$  range.

The electron isolation for the 8 TeV analysis is also dependent on the electron  $E_T$ . For electrons, the calorimeter isolation requirement is based on  $E_{T,cone30}$ , defined as the sum of the cluster transverse energies, of surrounding energy deposits in the ECAL and HCAL within a cone of  $\Delta R = 0.3$  in the  $\eta - \phi$  space around the electron. The cells within  $0.125 \times 0.175$  in  $\eta \times \phi$  around the centre of the electron object are excluded. The track isolation term is defined by  $p_{T,cone30}$ , which is the scalar sum of all tracks with  $p_T > 400$  MeV inside a cone of  $\Delta R = 0.3$  with respect to the electron, excluding the electron track itself. Tracks considered in the sum must originate from the primary vertex associated to the electron track. For electrons with  $E_T < 15$ , the track isolation cone size is increased to  $\Delta R = 0.4$ . These energy terms are divided by the electron  $E_T$  ( $p_T$  for the 7 TeV analysis) to form the isolation requirements. Electron isolation is tighter for the 7 TeV selection due to a less efficient estimation of the misidentification backgrounds, namely  $W$ +jets and QCD multijet processes.

Impact parameter requirements are imposed on the electrons, with the 8 TeV values being more stringent in order to further suppress non-prompt leptons, due to the higher number of simultaneous proton-proton interactions (known as *pile-up*) conditions experienced at a higher centre-of-mass energy. The total electron selection is summarised in Table 4.1.

	$E_T$ [ GeV ]	Electron ID	Calo. isolation	Track isolation	Impact param.
8 TeV	10-15	VTLH	$E_{T,cone30}/E_T < 0.20$	$p_{T,cone40}/E_T < 0.06$	$ d_0 /\sigma_{d_0} < 3.0,$ $z_0 \sin \theta < 0.4$ mm
	15-20		$E_{T,cone30}/E_T < 0.24$	$p_{T,cone30}/E_T < 0.08$	
	20-25	Medium++ with “CBL”	$E_{T,cone30}/E_T < 0.28$	$p_{T,cone30}/E_T < 0.10$	
	> 25				
7 TeV	–	Tight++	$E_{T,cone30}/p_T < 0.12$	$p_{T,cone40}/p_T < 0.06$	$ d_0 /\sigma_{d_0} < 10, z_0 < 1$ mm

Table 4.1: Electron selection as a function of  $E_T$ . “CBL” refers to the conversion vertex and B-layer hit requirements. The 7 TeV electron selection is defined over the entire  $E_T$  range.

## 4.1.2 Muons

Muons are reconstructed using the STACO algorithm [52], which begins by extrapolating tracks from the outermost layer of the muon spectrometer (MS) back towards the interaction point and primary vertex. The MS track is required to have a track

segment in all three layers of the MS. This information is then combined with tracks from the inner detector, with requirements on the number of hits in various subdetector systems, to ensure that the reconstructed object is not a result of a cosmic muon track, or false signals [53]. Specifically, the sum of pixel detector hits and dead pixel sensors traversed by the track must be greater than zero, the sum of SCT hits and dead SCT sensors crossed by the track must be greater than four, and the number of missing hits in a working sensor crossed by the track is required to be less than three. In addition, if the track is within the acceptance of the TRT, a successful extension to the MS track must exist. Only muons which have  $|\eta| \leq 2.5$  are included, due to the loss in track quality at higher  $\eta$  in the MS. The  $p_T$  of a muon is computed from the track curvature and position within the MS.

As for the electrons, impact parameter and isolation cuts are also placed on the muons to reject muons arising from heavy flavour decays within jets, and are optimised using the same procedure as that of the electrons. The track isolation terms  $p_{T,cone30}$  (and  $p_{T,cone40}$ ) are defined identically to those of the electrons. However, the calorimeter isolation discriminant, based on  $E_{T,cone30}$ , is defined instead for muon candidates as the sum of the ECAL and HCAL calorimeter cells  $E_T$  above a noise threshold ( $\sim 3$  GeV) [54] inside a cone of  $\Delta R = 0.3$  in the  $\eta - \phi$  space around the muon. The muon isolation terms are then divided by the muon  $p_T$  to form the isolation requirement. For the 7 TeV analysis, the isolation requirements are tighter for lower  $p_T$  muons, and is given as either a fixed value, or a function of the muon  $p_T$ , whichever is the smaller. The idea behind such isolation conditions was to target non-prompt muons, which are typically of lower energy, whilst retaining a high level of acceptance for true muons. The muon selection is dependent on the muon  $p_T$  and is summarised in Table 4.2.

	$p_T$ [ GeV ]	Calo. isolation	Track isolation	Impact param.
8 TeV	10-15	$E_{T,cone30}/p_T < 0.06$	$p_{T,cone40}/p_T < 0.06$	$ d_0 /\sigma_{d_0} < 3.0,$ $z_0 \sin \theta < 1.0$
	15-20	$E_{T,cone30}/p_T < 0.12$	$p_{T,cone30}/p_T < 0.08$	
	20-25	$E_{T,cone30}/p_T < 0.18$	$p_{T,cone30}/p_T < 0.12$	mm
	> 25	$E_{T,cone30}/p_T < 0.28$		
7 TeV	–	$E_{T,cone30}/p_T < \min(0.20, 0.0125p_T - 0.14)$	$p_{T,cone30}/p_T < \min(0.15, 0.011p_T - 0.12)$	$ d_0 /\sigma_{d_0} < 3.0,$ $z_0 \sin \theta < 1.0$ mm

Table 4.2: Muon selection as a function of  $p_T$ .  $E_{T,cone30}$  and  $p_{T,cone30}$  refer respectively to the  $E_T$ , or  $p_T$  present in a cone of  $\Delta R = 0.3$  in the  $\eta - \phi$  space around the muon. The 7 TeV muon definition is valid over the entire  $p_T$  range.

## 4.2 Jets

Hadronic jets are a crucial part of the object definition, as the  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  analysis relies heavily on the jet multiplicity to distinguish the VBF Higgs production mode from that of ggF. Jet objects are reconstructed using the anti- $k_T$  sequential clustering algorithm [55], chosen as the default jet construction in ATLAS as it produces stable, cone-like jets (circular in  $\eta$ - $\phi$  space). It is relatively insensitive to pile-up, and underlying event processes compared to other jet algorithms. In addition, it is infrared and collinear safe, meaning that the reconstructed jet is affected by neither soft gluon radiation nor collinear parton splitting. Three-dimensional calorimeter clusters [56, 57], which are groups of calorimeter cells containing significant energy deposits from both the ECAL and HCAL, are used as input to the anti- $k_T$  algorithm. The clusters are calibrated to the energy scale of electromagnetic showers, and also corrected for any pile-up contributions and the position of the primary vertex. The distance parameter of jets used in the  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  analysis is  $R = 0.4$ , corresponding roughly to a cone size of  $\Delta R = 0.4$  in  $\eta - \phi$  space. Subsequent calibration of the jets to the hadronic energy scale is then performed, using  $p_T$  and  $\eta$ -dependent correction factors calculated from MC simulation (for the 2011 dataset) and from data (2012 dataset).

A jet cleaning method [58, 59] is applied to the reconstructed jets, which discards events with jet candidates from false signals, due to cosmic-ray showers, unstable beam conditions or subdetector issues, if the jet has  $p_T > 20$  GeV. Such “bad-jet” candidates, as well as distorting the jet multiplicity, also result in an unreliable measurement of missing energies. For the purposes of classifying an event in terms of jet multiplicity, the jets also have further  $p_T$ ,  $\eta$  requirements, where the increased  $p_T$  threshold in the higher  $|\eta|$  region suppresses jets resulting from pile-up. Additionally a cut on the jet vertex fraction (JVF) is applied, which is designed to reduce the number of jets originating from pile-up vertices. The JVF gives the fraction of total  $p_T$  associated with the tracks of a particular jet within a cone size of  $\Delta R = 0.4$ , to that of the total  $p_T$  associated with the primary interaction vertex. Jets with  $p_T < 50$  GeV within the inner detector acceptance of  $|\eta| < 2.4$  are required to have  $|JVF| > 0.5$ . An absolute value for JVF is imposed, as a value of  $-1$  could be achieved if no tracks are associated to the jet.

For the VBF analyses, the two highest  $p_T$  jets which pass the selection criteria are referred to as the *tagging jets*, and are the only ones used for any “dijet” observable. A summary of the jet selection can be found in Table 4.3. The 7 TeV analysis uses exactly the same jet selection, although as pile-up conditions were less severe during the 2011 data-taking period, the requirement on the JVF  $|JVF| > 0.75$ , can be stricter without a loss in jet acceptance.

$ \eta  < 2.4$	$ \eta  \geq 2.4$
Jet cleaning	
$ \eta  < 4.5$	
$p_T > 25$ GeV	$p_T > 30$ GeV
$ JVF  > 0.5$ if $p_T < 50$ GeV	–

Table 4.3: Jet selection as a function of  $p_T$  and  $\eta$  for the nominal jets used in the 8 TeV analysis. The 7 TeV analysis used the same jet selection, except with a  $|JVF| > 0.75$  cut. The jet cleaning method is described in Refs. [58, 59]

Additional jets which are not included when defining an event by jet multiplicity have relaxed requirements - so called “sub-threshold jets”, as they fall below the standard  $p_T$  lower limit of 25 GeV. Such jets with  $p_T > 20$  GeV feature in the analysis for  $b$ -jet identification (described below) and also in construction of one of the VBF observables, namely the Central Jet Veto, which is discussed in Chapter 7.

### 4.2.1 $b$ -Jets

Jets originating from  $b$ -hadrons are referred to as “ $b$ -jets”, and distinguished from light quark jets or gluon jets using “ $b$ -tagging” algorithms. Such algorithms exploit the relatively long lifetime of  $b$ -hadrons, which travel a measurable distance from the primary vertex before further decay, resulting in a secondary vertex which is somewhat displaced. Track information of decaying particles from the secondary vertex is also useful in determining the likelihood that a specific jet is classified as a  $b$ -jet. Due to the use of tracking information, the acceptance range of  $b$ -tagging is limited to  $|\eta| < 2.5$ , within the range covered by the inner detector. For the  $H \rightarrow WW^{(*)} \rightarrow l\nu l\nu$  analysis,  $b$ -jet requirements are used to suppress the significant top quark production backgrounds present.

The “MV1” multivariate technique is used [60], which is based on a neural network using the following three tagging algorithms as input: “IP3D”, “JetFitterCombNN” and “SV1” [61]. These use respectively the impact parameter information, typical topology of various hadron decays, and the possible reconstruction of a secondary vertex. The efficiency with which the algorithm identifies  $b$ -jets is measured [62] in a high-statistics data sample of  $t\bar{t}$  pair candidates with two leptons in the final state. A working point of the MV1 algorithm which is 85% efficient in tagging  $b$ -jets is used, with a  $p_T$  threshold of 20 GeV. At this working point, the probability of misidentifying a light quark jet as a  $b$ -jet is 10.3% [63]. The usage of sub-threshold  $b$ -jets maximises the ability to identify  $b$ -jets for this analysis, as the algorithm can then make use of a larger range of jet information. The identification of  $b$ -jets is restricted to the acceptance of the inner detector,  $|\eta| < 2.5$ .

### 4.3 Missing Transverse Energy

As the two neutrinos in the  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  final state are not directly detectable in the ATLAS detector, the momentum imbalance in the transverse plane is used to infer their presence, since the initial transverse momentum is zero. This missing transverse energy is defined as the negative vector  $p_T$  sum of reconstructed objects, such as leptons, jets and photons, and detector deposits not associated with these objects. It is an important discriminating observable as many background processes are characterised by it, and for those processes which should not have any neutrinos in the final state, its presence points to detector effects or mismeasurement of this quantity. There exist several methods of calculating this observable depending on the subdetector information used. Missing transverse energy reconstructed from the calorimeter is simply referred to as  $\cancel{E}_T$ , which is used in the analysis in order to make use of the large rapidity acceptance of the calorimeter and its sensitivity to neutral particles. The transverse momenta of identified objects (leptons, jets and photons) determined using the calorimeter system, and unidentified calorimeter deposits are included in the calculation of  $\cancel{E}_T$ .

For this analysis a modified version of missing energy is used in addition,  $\cancel{E}_T^{jet,Track}$ , which uses transverse momenta calculated from the reconstructed tracks, but with those from jet objects replaced with calorimeter-based momenta in order to use the full energy of the jets.  $\cancel{E}_T^{jet,Track}$  has been shown to give a better energy resolution of  $\mathcal{O}(10\%)$ , as it has reduced pile-up dependence due to the use of tracking information [64, 65]. Only tracks which originate from the primary vertex are included in the calculation, and for those which are not identified as leptons or jets, a further requirement of  $p_T > 0.5$  GeV is imposed to suppress detector noise. They are both defined as follows:

$$\cancel{E}_T = - \sum_{i, Calo} p_T^{\rightarrow i} \quad (4.1)$$

$$\cancel{E}_T^{jet,Track} = - \sum_{i, TRKS} p_T^{\rightarrow i} + \sum_{j, jets} \left( p_T^{\rightarrow j, TRK} - p_T^{\rightarrow j, Calo} \right) \quad (4.2)$$

where ‘‘Calo’’ and ‘‘TRKS’’ refer respectively to the calorimeter objects or tracks, the sum over  $i$  includes all identified particles and any detector energy measurements, and the sum over  $j$  refers to all reconstructed jets. Use of both missing transverse energy definitions allows for a better overall background rejection, and the inclusion of neutral particles originating from jets, which are not detected by the tracker. An example of the improved resolution is shown in Figure 4.1 over two other missing energy definitions, by reconstructing the mass of the di-tau system  $m_{\tau\tau}$  from  $Z/\gamma^* \rightarrow \tau\tau$  decays. It can be seen that the  $m_{\tau\tau}$  distribution for  $\cancel{E}_T^{jet,Track}$  is more strongly peaked, and in good agreement with the data.

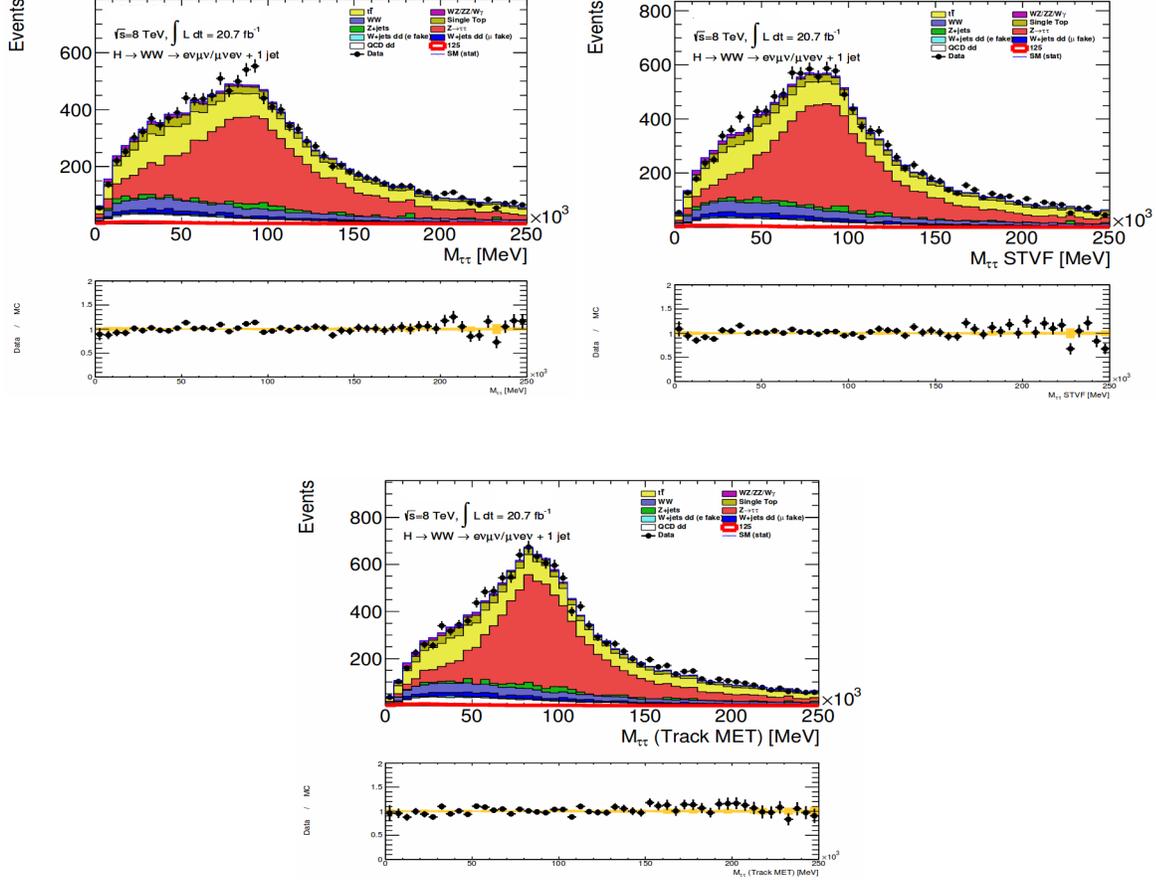


Figure 4.1: Distributions of  $m_{\tau\tau}$  with different missing transverse energy definitions in the  $e\mu + \mu e$  flavour channel. The upper left and right plots are calculated using  $\cancel{E}_T^{RefFinal}$  and  $\cancel{E}_T^{STVF}$  respectively, both of which are previous definitions considered in the  $H \rightarrow WW^{(*)} \rightarrow l\nu l\nu$  analysis. The lower distribution uses  $\cancel{E}_T^{jet,Track}$  which can be seen to have the strongest peak and thus the best resolution. The ratio plots below the  $m_{\tau\tau}$  distributions show the data/MC ratio, with the yellow band representing only the statistical error [64].

## 4.4 Overlap Removal

To avoid mismeasurement or false signals arising from the subdetectors, a defined set of steps is taken when two leptons, or a lepton and a jet, are in close proximity to one another. Any electron candidates with tracks extending to the MS are rejected, as are those which are separated from a muon by  $\Delta R < 0.1$ . Such cases are indicative of a muon which has undergone bremsstrahlung within the inner detector or calorimeter. Also due to bremsstrahlung, an electron can produce another electron candidate and thus, for two electrons separated by  $\Delta R < 0.1$ , the more energetic electron is retained. A highly energetic electron is always reconstructed as a jet, therefore if an electron is within  $\Delta R < 0.3$  of a jet, the jet is removed. In contrast, for muons in the close vicinity of a jet, the muon candidate is instead removed as it could be a non-prompt muon from a heavy quark decay. Only objects which have survived the entire cleaning and the overlap removal enter into the analysis. The overlap removal procedure is summarised in Table 4.4:

Objects	$\Delta R$	Procedure
$\mu - e$	$< 0.1$	Keep $\mu$ , remove $e$
$\mu - e$	$< 0.05$	Remove entire event
$e - e$	$< 0.1$	Keep higher $p_T$ $e$
$e - jet$	$< 0.3$	Keep $e$ , remove jet
$\mu - jet$	$< 0.3$	Keep jet, remove $\mu$

Table 4.4: Summary of overlap removal procedure. As a reminder,  $\Delta R$  is defined in the  $\eta - \phi$  metric as  $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$



# Chapter 5

## Boosted Decision Trees

Multivariate analysis (MVA) techniques refer to any method used to analyse data that involves more than one input variable. Boosted decision trees (BDT) are an example of such techniques and are a popular choice due to the simplicity of the method and speed of the algorithm [66]. A decision tree is simply a sequence of binary splits of the data as shown in Figure 5.1, forming a tree-structure, which eventually allows one to classify an event as either signal or background. This Chapter discusses the various aspects which a BDT is comprised of, with focus on the procedures used in the VBF  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  analysis.

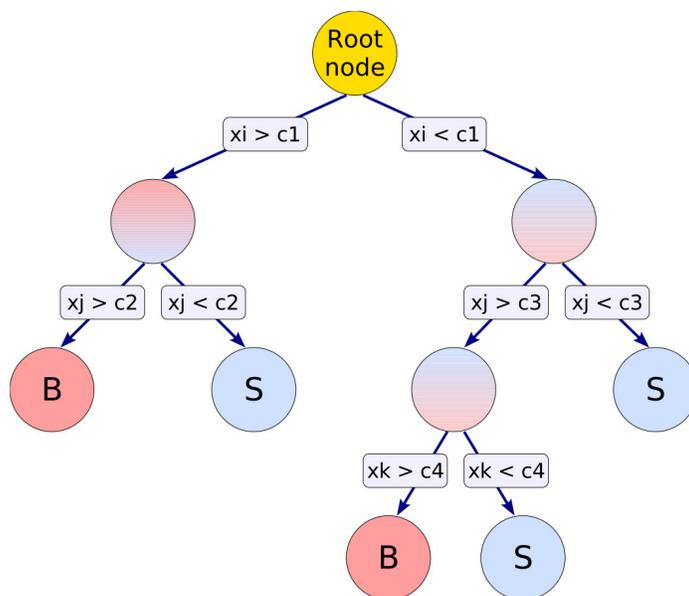


Figure 5.1: Schematic view of a decision tree, split using discriminating variables  $x$  at cut values  $c$ . The leaf nodes at the bottom are labelled as signal-like “S” or background-like “B” depending on the relative fraction of events which end up in the respective nodes [67].

## 5.1 Training and Classifying

To *train* such a tree, a set of known signal and background testing events are provided to the tree algorithm which then splits the input data at each tree node, according to a particular input observable and its value, in order to maximise the separation between signal and background events. This process continues until some pre-defined stop criteria are reached, leaving a number of final nodes (called “leaves”). Each leaf node is classed as signal-like or background-like depending on the majority of events which are present. Useful input observables are those for which the shape of the signal distribution is markedly different to that of the background.

The trained tree can then be applied to fresh sets of data so as to *classify* these events accordingly. The events which the tree classifies should be independent to those used to train the tree, to avoid overtraining on features particular to the training set. If simulated events are classified, it is possible to check the validity of the training procedure as it is known which of the inputs are signal or background. For the BDT, the output of the classification is given in terms of a *BDT score* ranging for example from  $[-1, 1]$ , where a score of  $-1$  means that an event is completely background-like, and  $+1$  is for a completely signal-like event. In effect, the BDT produces a new event observable, which can be simply cut on, or used for a shape analysis. For this analysis, the BDT implementation within the TMVA [67] package was used.

### 5.1.1 Quantifying the Separation

Various separation criteria can be used to evaluate the performance of a specific variable and its cut value at the tree node. They are symmetric with respect to the signal and background classes, as obviously a cut which singles out background events is equally valuable as one which selects for signal events. For all the variations possible within TMVA, there is a maximum if the samples are fully mixed, which descends to zero if the sample consists of purely signal or background events. Studies have been conducted [67] which show no significant difference in performance between the separation criteria. The benchmark used for this analysis was the *Gini Index* which is defined by  $p \cdot (1 - p)$ , where the purity  $p$  of a node is given by the ratio of signal events, to the total events in that node. To determine the variable and cut value which optimises the separation, the algorithm scans the cut values over the variable range with a granularity which can also be adjusted.

### 5.1.2 Input Variables

Observables from the experiment which show good separation between signal and background processes are good inputs to the algorithm. For instance the difference in rapidity between the two tagging jets  $\Delta Y_{jj}$  as illustrated in Figure 5.2.

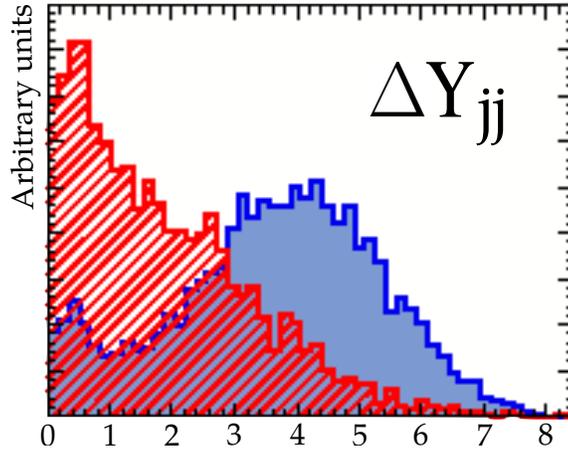


Figure 5.2: The difference in rapidity between the two tagging jets  $\Delta Y_{jj}$  is a good example of a well separated variable. The  $\Delta Y_{jj}$  shape for VBF produced Higgs signal (blue) and other backgrounds (red) is normalised to an arbitrary value to allow the shape difference to be better shown.

### 5.1.3 Stop Criteria

Node splitting terminates once a pre-defined set of conditions is fulfilled. For instance, once the minimum number (or fraction) of events contained in that node is reached; or the maximum number of nodes allowed; or the maximum depth of the tree has been attained.

### 5.1.4 Boosting

Single decision trees can be unstable with respect to statistical fluctuations in the training set. To overcome this, several trees can be consecutively trained using the same training set and by *boosting* events along the way. Simply put, boosting is the reweighting of events which were misclassified previously - the following tree then works harder to correctly identify these events. Each event is then classified according to the average over the outputs from all trees. “Gradient boosting” is one such boosting algorithm, which is used for this analysis. Details can be found in Ref. [67], from which the following description is based.

A model response  $F(\mathbf{x})$  can be defined as a weighted sum of parameterised base functions  $f(x; \alpha_n)$ :

$$F(\mathbf{x}; P) = \sum_{n=0}^N \beta_n f(x; \alpha_n); P \in \{\beta_n; \alpha_n\}_0^N \quad (5.1)$$

where  $P$  denotes the parameters to be adjusted and  $\mathbf{x}$  are the input variables. Each base function corresponds to a decision tree. The boosting procedure adjusts  $P$  such that the deviation between  $F(\mathbf{x})$  and the true value  $y$  obtained using the training

sample is minimised. This deviation is measured by a *loss function*  $L(F, y)$ , which fully determines the boosting method. Gradient boosting uses the binomial log-likelihood loss function:

$$L(F, y) = \ln(1 + \exp^{-2F(\mathbf{x})y}) \quad (5.2)$$

When faced with noisy data, such a loss function performs better than an exponential loss function, which is not robust if outliers are present. The current gradient of the loss function is calculated and then a regression tree<sup>1</sup> formed, whose leaf values are adjusted to match the mean value of the gradient in each region defined by the tree structure. Iterating this procedure thus produces the set of decision trees which minimises  $L$ .

Related to the boosting method, the number of trees to be grown can also be controlled. Generally, having too few trees could result in a poorly performing classifier which has not been able to make full use of the information available in the training set, whilst being also more vulnerable to fluctuations in the training data. In contrast, if the number of trees is increased, after a certain point the classifier performance ceases to improve whilst the algorithm simply consumes more computational power.

If gradient boosting is used, a “shrinkage” parameter can also be given which defines the learning rate of the boosting algorithm. This controls the weight of the individual decision trees, and can improve the robustness of the classifier if it is lowered. A small shrinkage (0.1-0.3) requires more trees to be grown, but can enhance the accuracy of the classifier.

### 5.1.5 Bagging

“Bagging” is a resampling technique [68] in which a classifier such as a decision tree, is repeatedly trained using randomly resampled events from the training set, such that the combined classifier represents the average of the individual classifiers. It is however not technically a boosting technique, as it does not enhance the classifier performance. Rather, it effectively smears over statistical representations of the training events and therefore stabilises the classifier response. In the TMVA framework, if bagging is used then it is specified by the ‘bagging fraction’, which gives the relative size of the bagged event sample compared to the original training set.

Resampling includes the possibility of replacement, and is equivalent to viewing the training set as a representation of the probability density distribution of the parent sample.

---

<sup>1</sup>A regression tree is similar to a decision tree, except that the final leaf nodes represent a specific value of the target variable, as opposed to sorting the events into categories. The target variable is that which the regression function is trying to estimate, for example, the price of a car given the properties it has. Within the same analogy, a decision tree would use the properties of the object to determine whether it is really a car, instead of a horse.

## 5.2 Evaluation of the Training

A number of checks are performed after the training, to ensure its validity and evaluate the performance. In this analysis, the statistical significance defined as  $S/\sqrt{B}$  is used as a figure of merit, where  $S$  and  $B$  represent the number of signal and background events respectively. The most basic method is to simply apply the trained BDT onto an independent set of events to see how well it performs - if there is a significant decrease in the performance, this could be a sign of *overtraining*. Comparisons between different sets of input variables and BDT parameters can also be conducted, by comparing the significance achieved - in the context of this analysis, the BDT optimisations performed are described in Chapter 7.

### 5.2.1 Overtraining

Overtraining of a BDT can occur for instance, if its parameters are adjusted such that it is trained to recognise well the current training sample leaving it vulnerable to fluctuations within this set. Essentially, the BDT has insufficient degrees of freedom as too many of the algorithm parameters have been adjusted to too few of the training data points. This leads to a seemingly large increase in the classifier performance with the training sample, but when tested on an independent sample, shows a sizeable drop-off. It is straightforward however to detect any overtraining and also easy to counteract it. Specifically for BDT training, the simplest method is to limit either directly the number of nodes present in the tree, or to restrict the tree depth to a lower value. The TMVA package provides also a numerical quantification of overtraining in the form of a Kolmogorov-Smirnov test, between the normalised BDT score distributions of the training and testing samples. This test quantifies the difference in shape between two distributions and is used as a rough guide when checking for overtraining.



# Chapter 6

## Data and Simulation

### 6.1 Data

Data used for this analysis were taken during 2011 at a centre-of-mass energy  $\sqrt{s} = 7$  TeV and also during 2012 at  $\sqrt{s} = 8$  TeV. Not all the data collected are usable due to the detector conditions at runtime, and only data from a “good runs list” (GRL) [69] are included for analysis. Such a GRL is an indicator of data quality and ensures stable beam conditions, and that all required subdetector and trigger systems were operating correctly, for all the data it contains. The recorded data are grouped into periods corresponding to the detector requirements during that time. A list of the dates corresponding to each data period can be found in the Appendix. An integrated luminosity  $\int L dt = 4.5 \text{ fb}^{-1}$  and  $\int L dt = 20.3 \text{ fb}^{-1}$  from the 7 TeV and 8 TeV datasets respectively are used in this analysis. The luminosity is evaluated from data included in a GRL using the standard ATLAS Luminosity Tool [70], which has a uncertainty of  $\pm 1.8\%$  for the 7 TeV data and  $\pm 2.8\%$  for 8 TeV data, from a preliminary calibration of the luminosity scale derived from van der Meer scans performed in November 2012 [47]. The following Sections detail the selection of data used in this analysis, including the event cleaning applied to the datasets and the triggers required.

#### 6.1.1 Event Cleaning

Apart from the application of a GRL, a number of other event conditions need also to be met [59]:

- **LAr Error Flag** - this flags a certain data-taking time period as being unsuitable for analysis, due to irregular noise bursts experienced by the LAr ECAL. These are automatically recorded by the detector, and defines a time interval of one second after such a noise burst [71].
- **Missing Energy Cleaning** - for accurate reconstruction, events containing at least one “bad” jet, as defined in Section 4.2, are removed [58].

- **LAr Hole Cleaning** - this refers to the rejection of data events containing a jet with  $p_T > 25$  GeV which intersects a faulty area of the LAr ECAL during the 2011 data-taking period between E-H. This procedure is outlined by the ATLAS Jet-Etmiss performance group [58] and is a result of the malfunctioning of four front end boards in the ECAL, producing a “hole” in  $\eta \times \phi$  space of  $(0.0 < \eta < 1.45) \times (-0.78847 < \phi < -0.59213)$ . In MC simulation, this method is also applied to the percentage of events corresponding to the same integrated luminosity affected by this fault.

## 6.1.2 Triggers

This analysis uses data events selected with triggers that require the presence of a single lepton or two leptons (dilepton). As it is more likely that an event contains a single lepton rather than multiple leptons, the lepton identification requirements belonging to the single-lepton triggers are more restrictive, with higher  $p_T$  thresholds than that of the dilepton triggers, in order to reduce the rate of event acceptance. The specific software EF<sup>1</sup> triggers used in the 2011 and 2012 data collection are summarised in Table 6.1 and 6.2. The triggers are denoted by their  $p_T$  threshold and also the quality criteria required.

Period	ee channel	$\mu\mu$ channel	$e\mu + \mu e$ channel
B - I	e20_medium	mu18_MG	e20_medium    mu18_MG
J	e20_medium	mu18_MG_medium	e20_medium    mu18_MG_medium
K	e22_medium	mu18_MG_medium	e22_medium    mu18_MG_medium
L - M	e22vh_medium	mu18_MG_medium	e22vh_medium    mu18_MG_medium

Table 6.1: EF triggers used for the 2011 dataset according to the data periods. Only single-lepton triggers were used during the 2011 data-taking. For the  $e\mu + \mu e$  channel, the triggers are combined in a logical OR condition. “MG” refers to the MuGirl algorithm used to identify muons in the ATLAS detector, which had better low  $p_T$  ( $< 50$  GeV) performance at the time of data-taking [72]

Due to the increase in luminosity between the 2011 and 2012 data-taking period, the level of pile-up and trigger rate increased considerably, provoking a change in the trigger definitions used. This is reflected in the changing of the lepton triggers used over the 2011 period. The triggers used for each time period during 2011 are the lowest *unprescaled*<sup>2</sup> triggers, which have the lowest  $p_T$  threshold possible whilst still being

<sup>1</sup>As described in Section 2.2.6, “EF” refers to the Event Filter trigger level, which makes up the final stage of the trigger system before events are written to mass storage

<sup>2</sup>Prescaling with a value  $n$  refers to the procedure of accepting only, one in  $n$  firings of the trigger, reducing the effective recorded luminosity for that trigger by  $1/n$ .

ee channel	$\mu\mu$ channel	$e\mu + \mu e$ channel
e24vhi_medium    e60_medium    2e12Tvh_loose    2e12Tvh_loose1_L2StarB	mu24i_tight    mu36_tight    mu18_tight_mu8_EFFS	e24vhi_medium    e60_medium    mu24i_tight    mu36_tight    e12Tvh_medium_mu8

Table 6.2: EF triggers used for the 2012 dataset, combined in a logical OR condition.

recorded in full, in order to cope with the high event rate. Apart from the first muon trigger, all the single-lepton triggers used during the 2011 runs satisfy the “Medium” cut-based quality criteria. In order to reduce further the high rate of electron triggers, the energy thresholds in the ECAL were increased slightly (by 1 or 2 GeV) in specific  $\eta$  regions with minimal trigger efficiency losses. The increased rate of electron triggers due to highly energetic hadronic jets was also suppressed by requiring that the energy deposit in the HCAL belonging to the electron candidate is less than 1 GeV. Together, these two additional conditions are indicated by the suffix “vh” in the electron trigger [73]. “MG” refers to the MuGirl algorithm used to identify muons in the ATLAS detector, which had better low  $p_T$  ( $< 50$  GeV) performance at the time of data-taking [72]

A combination of single and dilepton triggers were used for the 2012 data collection to compensate for a higher centre-of-mass energy and pile-up conditions, and additionally to allow the reduction of the lepton  $p_T$  thresholds to gain acceptance. No changes were necessary to the trigger choice throughout 2012 and hence no further differentiation between data periods is required here. “T” and “i” in the trigger name refer respectively to further  $E_T$  and track isolation cuts applied to the trigger [74]. Alternative ID criteria for electrons were introduced after the 2011 data-taking period, such as “L2StarB” [75], which is used in conjunction with the standard L2 triggers for the 2012 analyses. This is an alternative L2 trigger with different criteria, which shows up to 20% improvement at low  $p_T$  (below  $\sim 24$  GeV). An example can be seen in Figure 6.1 of the comparison between the standard quality criteria and the L2StarB criteria for the e7\_medium and e12Tvh\_loose electron triggers. For muons, the Event Filter Full Scan (EFFS) [76] trigger definition was added, which searches for muons at the EF level, independently of any L1 or L2 seed. In this way, the number of muons accepted can be increased.

The efficiencies of the lepton triggers have been measured using the *tag-and-probe* method with a data sample of  $Z/\gamma^* \rightarrow ee + \mu\mu$  candidates, by the ATLAS ElectronGamma Combined Performance Group [51, 77] and the ATLAS Muon Combined Performance Group [78, 79]. The single-lepton trigger efficiency of leptons satisfying the analysis selection criteria is approximately 90% for electrons, and 90% for muons with  $|\eta| > 1.05$  (70% for muons with  $|\eta| < 1.05$ ). Differences between the trigger efficiencies in data and Monte Carlo (MC) simulation were observed, leading to the calculation of  $p_T$  and  $\eta$ -dependent scale factors with which the MC simulation is cor-

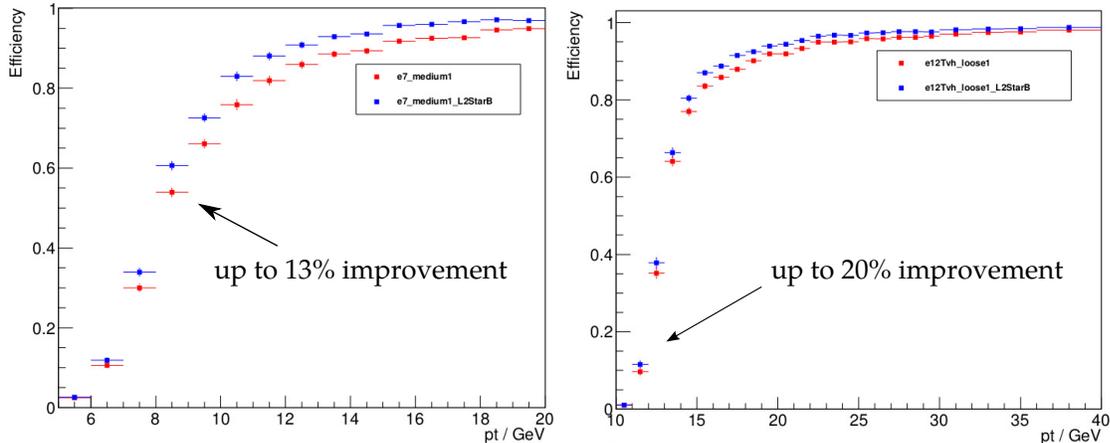


Figure 6.1: Comparison of the trigger efficiencies of the e7\_medium (left) and e12Tvh\_loose (right) electron triggers, when using the standard (red) or L2StarB (blue) quality criteria. With the use of the L2StarB criteria, e7\_medium and e12Tvh\_loose show up to 13% and 20% improvement at low  $p_T$  respectively.

rected to take into account real detector effects.

## Tag-and-Probe Method

Lepton efficiency measurements are conducted using the tag-and-probe method, by which *probe* leptons are selected using other information in the event which makes up the *tag* requirements. Depending on the efficiency measurement, the probe lepton is usually loosely (with respect to the quality and selection criteria) selected, with the tag requirement being at least a well-reconstructed lepton. For trigger efficiency evaluation in particular, the tag lepton is further required to have fired the trigger under investigation.  $Z$  boson or  $J/\psi$  decays are often used for efficiency evaluation and thus a further tag requirement that the combined invariant mass of the tag and probe leptons corresponds to that of the  $Z$  boson or  $J/\psi$  is imposed. The efficiency  $\epsilon$  from the sample of probe leptons are calculated from data in the following way:

$$\epsilon = \frac{N_{\text{matched probes}}}{N_{\text{probes}}} \quad (6.1)$$

where  $N_{\text{matched probes}}$  are the subset of probe leptons matching the property under investigation, and  $N_{\text{probes}}$  are the total number of probe leptons selected.  $p_T$ ,  $\eta$ ,  $\phi$ -dependent efficiencies are derived for these efficiencies in order to cover all possible analysis selection of lepton objects. The resolution of the efficiencies with respect to each observable is limited largely by the number of available probes, but also by the accuracy of any background subtraction techniques used. For instance, by subtracting background processes based on MC simulation, or by defining control regions as needed.

The entire procedure is then repeated using MC simulation, to obtain MC-predicted efficiencies which can then be compared to the data-measured efficiencies in order to

derive sets of scale factors with corresponding systematic uncertainties. An efficiency scale factor is defined as the ratio of the data to MC efficiency, which is then applied to the MC simulation so that real detector effects can be taken into account in the MC. These efficiencies and scale factors are provided by the ATLAS Combined Performance Groups [51, 77, 78, 79].

## 6.2 Monte Carlo Simulation

Dedicated Monte Carlo (MC) simulations of the signal and background processes are produced so as to allow comparison with the collected data. For the majority of the processes, separate programs are used to generate the hard<sup>3</sup> scattering process with predictions from perturbation theory matrix elements, and to model the subsequent parton showering<sup>4</sup> (PS) and hadronisation. The various MC generators used for each signal and background process are summarised in Table 6.3, along with their cross sections.

All the 7 TeV MC samples have been produced during the 2011 ATLAS MC production campaign “MC11c” [80], whereas the 8 TeV samples are predominantly from the “MC12a” part of the 2012 ATLAS MC production run, with a handful of samples produced during “MC12b” [81, 82]. Events are filtered during generation where necessary, allowing an increase in sample size for a particular process. The number of simulated events are scaled to the measured integrated luminosity for 7 TeV and 8 TeV each, using the predicted cross section of the particular process. For instance, if there are  $N_{MC}$  events present in the MC sample with effective luminosity  $L_{MC}$ , which corresponds to a data sample with luminosity  $L_{data}$ , then the background yield  $B$  is given as  $B = (L_{data}/L_{MC}) \cdot N_{MC}$ . As the number of generated events is large, the statistical uncertainty on each MC prediction is given by the standard deviation  $\sqrt{N_{MC}}$  of a Gaussian distribution, such that the statistical uncertainty on the background yield  $\delta = (L_{data}/L_{MC}) \cdot \sqrt{N_{MC}}$ .

For most of the processes, the acceptances and efficiencies of the ATLAS detector are then calculated using a detailed detector model with the Geant4 program [83]. This is referred to as the *FullSim* simulation and takes also into account a realistic treatment of the pile-up conditions during data-taking. A handful of processes use instead the Atl-Fast-II [84] simulation to provide the detector response, which contains parameterisations of the energy profiles in the calorimeter, thus reducing the simulation time by one order of magnitude. This makes it possible to generate a larger MC sample within the limited computing resources and still allows for the use of all standard ATLAS reconstruction code as for FullSim samples.

---

<sup>3</sup>“Hard” and “soft” refer to processes with high and low momentum transfer respectively.

<sup>4</sup>Parton showers refer to cascades of radiation produced by coloured partons, which radiate gluons. As the gluons themselves are also charged, further radiation can therefore be emitted, forming a “parton shower”

The CT10 parton distribution function (PDF) set [85] is mainly used for processes generated with POWHEG [86]. CTEQ6LI [87] tends to be used for samples produced with ALPGEN, AcerMC, PYTHIA 6 and PYTHIA 8 [88, 89, 90, 91]. Depending on whether the SHERPA sample belongs to the 8 TeV or 7 TeV set, either CT10 or CTEQ6LI is used. The  $Z/\gamma^* \rightarrow \ell\ell$  ALPGEN samples are reweighted to the MRSTM-Cal set [92], to improve the modelling of the lepton pseudorapidity distributions [93]. Detailed overviews of the 7 TeV and 8 TeV MC simulation used are shown in Table 6.4 and 6.5. Leptonic decays of  $W$  and  $Z$  bosons are always assumed for the MC samples considered here.

### 6.2.1 Pile-up Reweighting

Pile-up can arise from simultaneous interactions in the detector, as well as overlapping signals due to interactions occurring in other nearby bunch crossings. Pile-up interactions are modelled with PYTHIA 8, and the ATLAS detector response is simulated using either Geant4 or a Geant4-based calorimeter simulation [94]. As the MC production is not synchronised in time with the data-taking periods, the actual pile-up conditions may not be reflected in the simulation. Hence, it is necessary to apply a reweighting to the MC simulation, using the ATLAS pile-up reweighting tool [95]. The average number of interactions per bunch crossing  $\mu$  was 9 and 21 during the 2011 and 2012 data periods respectively. For the 2011 samples an additional  $\mu$  reweighting is not required, as the MC prediction was tuned to match the data [96], however for the 2012 MC samples,  $\mu$  is reweighted down by a factor of 0.9 [63]. The improvement on  $\mu$  due to the rescaling is presented in Figure 6.2 and shows the correction of the data to MC prediction ratio to a flat factor close to unity.

### Higgs processes

The Higgs production modes VBF, ggF and the associated Higgs production channels  $WH/ZH$  are considered to decay only in the  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  mode, with final states featuring two charged leptons. Of these leptons, only  $\ell = e, \mu$ , including small contributions from leptonic  $\tau$  decays which are taken into account. All Higgs samples are generated with a mass  $m_H = 125$  GeV. The branching fraction for this decay channel as a function of the Higgs mass  $m_H$ , is obtained from the HDECAY [97] program with an associated uncertainty of 4.5% for  $m_H = 125$  GeV. Other small production processes such as the Higgs production via top or bottom quark fusion are neglected here as their contribution is negligible.

For the VBF process, the total cross section is computed using an approximated QCD next-to-next-to-leading order (NNLO) computation with the VBF@NNLO [98] program, with EW corrections evaluated at leading order (LO) with the HAWK program [99]. The POWHEG generator is used to simulate the VBF Higgs production.

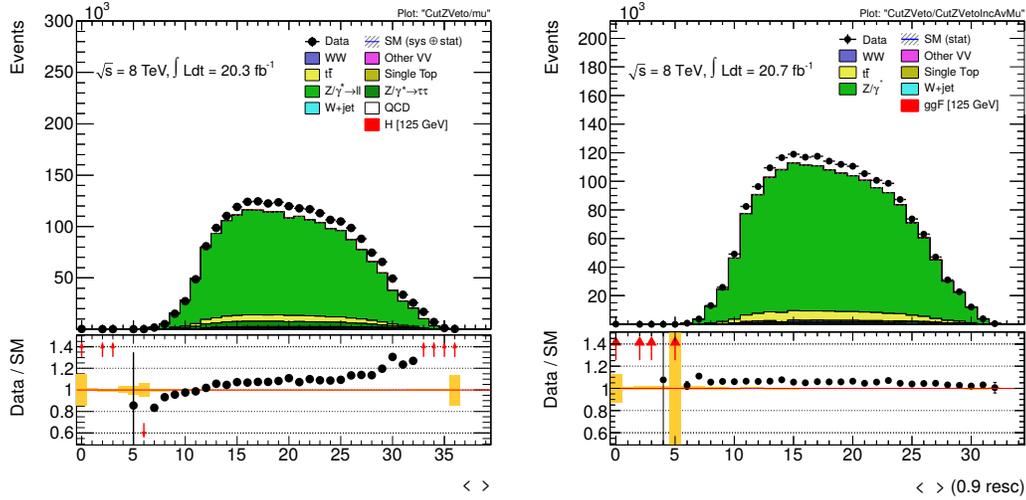


Figure 6.2: Comparison of the  $\mu$  distribution before (left) and after (right) application of the  $\mu$  rescaling for the 2012 MC production. The yellow band represents only the statistical error [96].

The total cross section for ggF Higgs production is calculated to NNLO in QCD [100] with soft gluon radiation resummations at next-to-next-to-leading log (NNLL) [101]. Next-to-leading order (NLO) EW corrections are also applied [102] and together, all these computations make up the total inclusive ggF cross section [103]. The POWHEG generator is used to simulate ggF Higgs processes, with showering performed using PYTHIA 8. The  $p_T$  spectrum of ggF Higgs events containing two or more jets are reweighted to that predicted by the NLO POWHEG-MINLO [104] simulation of Higgs production for better jet multiplicity agreement - further details can be found in Ref. [105].

The cross sections of the  $WH/ZH$  processes are evaluated up to NLO EW corrections [106], and NNLO QCD corrections [107, 108] using PYTHIA 6 for the 7 TeV dataset, and PYTHIA 8 for the 8 TeV dataset.

## Top quark production

$t\bar{t}$  forms the bulk of the top quark background and is modelled using POWHEG. The 8 TeV  $t\bar{t}$  sample, was produced using a dilepton-filter - in other words, only events containing two leptons in the final state were included, removing other processes - to increase the sample size. Such a filter was not considered necessary for the 7 TeV  $t\bar{t}$  sample. The single-top (st) production channels, namely  $s$ -channel,  $t$ -channel and  $Wt$ , are also modelled using POWHEG.

## Diboson production

For the VBF analysis, SHERPA is used to model all  $WW$  production with jets, instead of POWHEG which models the process poorly if more than one jet is present. The  $WW$  processes are divided into two sets: those production diagrams which only contain EW vertices, and those which have a QCD vertex present.

POWHEG is used to generate  $W\gamma^*$  with dilepton invariant mass  $m_{\ell\ell} > 7$  GeV, whereas SHERPA is used to simulate the range  $m_{\ell\ell} < 7$  GeV, as POWHEG is unable to model such a low dilepton invariant mass range. For (non-EW)  $WZ$  processes, POWHEG is used, with events having invariant mass of the  $Z$  or excited photon  $m_{Z/\gamma} < 7$  TeV removed in order to avoid overlap with the SHERPA  $W\gamma^*$  sample. The (non-EW)  $WZ$  processes are required to have at least two charged leptons present with  $p_T > 5$  GeV and  $|\eta| < 2.7$ .  $ZZ$  processes are modelled using POWHEG, where all events have a dilepton invariant mass cut of  $m_{\ell\ell} > 4$  GeV. EW  $WZ$  and EW  $ZZ$  processes (those containing only EW vertices) are generated instead with SHERPA.

The  $W\gamma$  sample is modelled using ALPGEN, with kinematic criteria applied in the generation of the different processes. For  $W(\rightarrow \ell\nu)\gamma$  events, the photon must have  $p_T > 8$  GeV with a separation from the charged lepton of  $\Delta R > 0.25$ .  $W(\rightarrow \ell\nu)\gamma(\rightarrow \ell'\ell')$  events must have at least two charged leptons with  $p_T > 5$  GeV, and either  $|\eta| < 3$  if the two leading leptons are electrons or muons, or  $|\eta| < 5$  if they are taus.  $Z\gamma$  processes are also simulated using ALPGEN, with the photon  $p_T > 7$  TeV.

## $W$ +jets and QCD Multijet Processes

These are not generated with MC, but modelled instead with a data-driven technique described in Section 8.7. Both  $W$ +jets and multijet backgrounds arise from the misidentification of jets or non-prompt leptons (*fake leptons*) as true leptons - hence a background estimation carried out using data is more appropriate here.

### $Z/\gamma^* \rightarrow \ell\ell$

The DY process is generated using ALPGEN, with a lower limit on the dilepton invariant mass of  $m_{\ell\ell} > 10$  GeV. To enhance the sample size, the generation was performed with a dilepton filter (at least one charged lepton with  $p_T > 20$  GeV and at least two charged leptons with  $p_T > 7$  GeV, all with  $|\eta| < 3$ ). Specifically with the VBF analysis in mind, additional higher jet multiplicity samples were generated using ALPGEN with a “VBF filter” which imposes the following jet requirements: at least two jets with  $p_T > 15$  GeV and  $|\eta| < 5$ , dijet invariant mass  $m_{jj} > 200$  GeV and  $\Delta\eta_{jj} > 2$ . Additionally, EW  $Z$ +jets processes in which no QCD vertices exist are modelled using SHERPA.

Process	Generator	$\sigma \cdot \text{Br}(8 \text{ TeV})$ (pb)	$\sigma \cdot \text{Br}(7 \text{ TeV})$ (pb)
ggF $H \rightarrow WW^{(*)}$	POWHEG +PYTHIA 8	0.435	0.341
VBF $H \rightarrow WW^{(*)}$	POWHEG +PYTHIA 8	$36 \cdot 10^{-3}$	$28 \cdot 10^{-3}$
$WH/ZH H \rightarrow WW^{(*)}$	PYTHIA 8 (PYTHIA 6)	$25 \cdot 10^{-3}$	$21 \cdot 10^{-3}$
QCD $WW + 2$ jets	SHERPA	0.568	-
EW $WW + 2$ jets	SHERPA	0.039	0.027
$t\bar{t}$ dileptonic	POWHEG +PYTHIA 6	26.6	18.6
$tW/tb$ leptonic	POWHEG +PYTHIA 6	4.17	3.15
$tqb$ leptonic	AcerMC+PYTHIA 6	28.4	20.7
inclusive $W$	ALPGEN +HERWIG [109]	$37 \cdot 10^3$	$31 \cdot 10^3$
inclusive $Z/\gamma^*(m_{\ell\ell} \geq 10) \text{ GeV}$	ALPGEN	$16.5 \cdot 10^3$	$14.9 \cdot 10^3$
EW $Z/\gamma^*$	SHERPA	5.36 (inc. t-ch)	2.26
$W(Z/\gamma^*)$	POWHEG +PYTHIA 8	12.7	10.8
$W(Z/\gamma^*)(m_{Z/\gamma^*} < 7 \text{ GeV})$	SHERPA	12.2	10.6
EW $WZ + 2$ jets	SHERPA	$13 \cdot 10^{-3}$	$8.5 \cdot 10^{-3}$
EW $ZZ + 2$ jets ( $4\ell, 2\ell 2\nu$ )	SHERPA	$73 \cdot 10^{-5}(12 \cdot 10^{-4})$	$53 \cdot 10^{-5}(8.8 \cdot 10^{-4})$
$W\gamma$	ALPGEN +HERWIG	369	313
$Z\gamma(p_T^\gamma > 7 \text{ GeV})$	SHERPA	163	-

Table 6.3: MC generators used to model the signal and background processes for the VBF analysis, and corresponding cross sections for the 7 TeV and 8 TeV samples. In the case of the signal processes, these are given for  $m_H = 125 \text{ GeV}$ . Leptonic decays of  $W/Z$  bosons are always assumed, and the quoted cross sections include the branching ratios and are summed over all three lepton flavours.

Process	Generator	PDF	PS	Simulation
$ggF: H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$	POWHEG	CT10	PYTHIA 8	FullSim
$VBF: H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$	POWHEG	CT10	PYTHIA 8	FullSim
$WH: H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$	PYTHIA 6	CTEQ6L1	PYTHIA 6	FullSim
$ZH: H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$	PYTHIA 6	CTEQ6L1	PYTHIA 6	FullSim
$WW \rightarrow \ell\nu\ell\nu + jj$ (EW $WW$ )	SHERPA	CT10	SHERPA	FullSim
$WW \rightarrow \ell\nu\ell\nu + \text{jets}$ (QCD $WW$ )	SHERPA	CT10	SHERPA	FullSim
$qq \rightarrow ZZ \rightarrow 4\ell(2\ell 2\nu)$ (QCD $ZZ$ )	POWHEG	CTEQ6L1	PYTHIA 6	FullSim
$ZZ \rightarrow \ell\nu\ell\nu + jj$ (EW $ZZ$ )	SHERPA	CT10	SHERPA	FullSim
$ZZ \rightarrow 4\ell + jj$ (EW $ZZ$ )	SHERPA	CT10	SHERPA	FullSim
$qq \rightarrow WZ$ (QCD $WZ$ )	POWHEG	CTEQ6L1	PYTHIA 6	FullSim
$WZ \rightarrow 3\nu + jj$ (EW $WZ$ )	SHERPA	CT10	SHERPA	FullSim
$W\gamma$	ALPGEN	CTEQ6L1	HERWIG	FullSim
$W\gamma^*$	SHERPA	CT10	SHERPA	FullSim
$Z \rightarrow ee$	ALPGEN	CTEQ6L1*	HERWIG	FullSim
$Z \rightarrow ee$ VBF Filter	ALPGEN	CTEQ6L1*	HERWIG	FullSim
$Z \rightarrow \mu\mu$	ALPGEN	CTEQ6L1*	HERWIG	FullSim
$Z \rightarrow \mu\mu$ VBF Filter	ALPGEN	CTEQ6L1*	HERWIG	FullSim
$Z \rightarrow \tau\tau$	ALPGEN	CTEQ6L1*	HERWIG	FullSim
$Z \rightarrow \tau\tau$ VBF Filter	ALPGEN	CTEQ6L1*	HERWIG	FullSim
$Z \rightarrow ee + bb$	ALPGEN	CTEQ6L1*	HERWIG	FullSim
$Z \rightarrow \mu\mu + bb$	ALPGEN	CTEQ6L1*	HERWIG	FullSim
$Z \rightarrow \tau\tau + bb$	ALPGEN	CTEQ6L1*	HERWIG	FullSim
$Z \rightarrow ee + \text{jets}$ (EW)	SHERPA	CTEQ6L1	SHERPA	FullSim
$Z \rightarrow \mu\mu + \text{jets}$ (EW)	SHERPA	CTEQ6L1	SHERPA	FullSim
$Z \rightarrow \tau\tau + \text{jets}$ (EW)	SHERPA	CTEQ6L1	SHERPA	FullSim
$DY \rightarrow ee$	ALPGEN	CTEQ6L1*	HERWIG	FullSim
$DY \rightarrow \mu\mu$	ALPGEN	CTEQ6L1*	HERWIG	FullSim
$DY \rightarrow \tau\tau$	ALPGEN	CTEQ6L1*	HERWIG	FullSim
$DY \rightarrow ee + \text{jets}$ (EW)	SHERPA	CTEQ6L1	SHERPA	FullSim
$DY \rightarrow \mu\mu + \text{jets}$ (EW)	SHERPA	CTEQ6L1	SHERPA	FullSim
$DY \rightarrow \tau\tau + \text{jets}$ (EW)	SHERPA	CTEQ6L1	SHERPA	FullSim
$W \rightarrow e\nu$	ALPGEN	CTEQ6L1	HERWIG	FullSim
$W \rightarrow \mu\nu$	ALPGEN	CTEQ6L1	HERWIG	FullSim
$W \rightarrow \tau\nu$	ALPGEN	CTEQ6L1	HERWIG	FullSim
$W \rightarrow \ell\nu + bb$	ALPGEN	CTEQ6L1	HERWIG	FullSim
$W \rightarrow \ell\nu + cc$	ALPGEN	CTEQ6L1	HERWIG	FullSim
$W \rightarrow \ell\nu + c$	ALPGEN	CTEQ6L1	HERWIG	FullSim
$t\bar{t}$ (no-all had)	Powheg	CT10	PYTHIA 6	FullSim
st $t$ -channel(leptonic)	AcerMC	CTEQ6L1	PYTHIA 6	AF2
st $s$ -channel(leptonic)	POWHEG	CT10	PYTHIA 6	AF2
st $Wt$ -channel(dilepton)	POWHEG	CT10	PYTHIA 6	FullSim

Table 6.4: Summary table for all MC samples used in the 7 TeV VBF  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  analyses. PDFs marked with “\*” indicates samples which are reweighted to the MRST PDF set.

MC12 Version	Process	Generator	PDF	PS	Simulation
12a	$ggF: H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$	POWHEG	CT10	PYTHIA 8	FullSim
12a	$VBF: H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$	POWHEG	CT10	PYTHIA 8	FullSim
12a	$WH: H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$	PYTHIA 8	CTEQ6L1	PYTHIA 8	FullSim
12a	$ZH: H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$	PYTHIA 8	CTEQ6L1	PYTHIA 8	FullSim
12a	$WW \rightarrow \ell\nu\ell\nu + jj$ (EW $WW$ )	SHERPA	CT10	SHERPA	FullSim
12a	$WW \rightarrow \ell\nu\ell\nu + \text{jets}$ (QCD $WW$ )	SHERPA	CT10	SHERPA	FullSim
12a	$gg \rightarrow ZZ \rightarrow 4\ell$ (QCD $ZZ$ )	gg2ZZ	CT10	HERWIG	FullSim
12a	$qq \rightarrow ZZ \rightarrow 4\ell(2\ell 2\nu)$ (QCD $ZZ$ )	POWHEG	CT10	PYTHIA 8	FullSim
12a	$ZZ \rightarrow \ell\nu\ell\nu + jj$ (EW $ZZ$ )	SHERPA	CT10	SHERPA	FullSim
12a	$ZZ \rightarrow 4\ell + jj$ (EW $ZZ$ )	SHERPA	CT10	SHERPA	FullSim
12a	$ZZ \rightarrow 2\ell 2\nu + \text{jets}$	SHERPA	CT10	SHERPA	FullSim
12a	$qq \rightarrow WZ \rightarrow 3\ell\nu$ (QCD $WZ$ )	POWHEG	CT10	PYTHIA 8	FullSim
12a	$WZ \rightarrow 3\ell\nu + jj$ (EW $WZ$ )	SHERPA	CT10	SHERPA	FullSim
12a	$W\gamma$	ALPGEN	CTEQ6L1	HERWIG	FullSim
12b	$W\gamma^* \rightarrow \ell\nu\ell\ell$	SHERPA	CT10	SHERPA	FullSim
12a	$Z \rightarrow ee$	ALPGEN	CTEQ6L1*	HERWIG	FullSim
12a	$Z \rightarrow ee$ VBF Filter	ALPGEN	CTEQ6L1*	HERWIG	FullSim
12a	$Z \rightarrow \mu\mu$	ALPGEN	CTEQ6L1*	HERWIG	FullSim
12a	$Z \rightarrow \mu\mu$ VBF Filter	ALPGEN	CTEQ6L1*	HERWIG	FullSim
12a	$Z \rightarrow \tau\tau$	ALPGEN	CTEQ6L1*	HERWIG	FullSim
12a	$Z \rightarrow \tau\tau$ VBF Filter	ALPGEN	CTEQ6L1*	HERWIG	FullSim
12a	$Z \rightarrow ee$	ALPGEN	CTEQ6L1*	PYTHIA 6	FullSim
12a	$Z \rightarrow \mu\mu$	ALPGEN	CTEQ6L1*	PYTHIA 6	FullSim
12a	$Z \rightarrow \tau\tau$	ALPGEN	CTEQ6L1*	PYTHIA 6	FullSim
12a	$Z \rightarrow ee + bb, cc$	ALPGEN	CTEQ6L1*	HERWIG	FullSim
12a	$Z \rightarrow \mu\mu + bb, cc$	ALPGEN	CTEQ6L1*	HERWIG	FullSim
12a	$Z \rightarrow \tau\tau + bb, cc$	ALPGEN	CTEQ6L1*	HERWIG	FullSim
12a	$Z \rightarrow ee + bb, cc$	ALPGEN	CTEQ6L1*	PYTHIA 6	FullSim
12a	$Z \rightarrow \mu\mu + bb, cc$	ALPGEN	CTEQ6L1*	PYTHIA 6	FullSim
12a	$Z \rightarrow \tau\tau + bb, cc$	ALPGEN	CTEQ6L1*	PYTHIA 6	FullSim
12a	$Z \rightarrow ee + \text{jets}$ (EW)	SHERPA	CT10	SHERPA	FullSim
12a	$Z \rightarrow \mu\mu + \text{jets}$ (EW)	SHERPA	CT10	SHERPA	FullSim
12a	$Z \rightarrow \tau\tau + \text{jet}$ (EW)	SHERPA	CT10	SHERPA	FullSim
12a	$Z \rightarrow ee + \gamma$	SHERPA	CT10	SHERPA	FullSim
12a	$Z \rightarrow \mu\mu + \gamma$	SHERPA	CT10	SHERPA	AFII
12a	$DY \rightarrow ee$	ALPGEN	CTEQ6L1*	HERWIG	FullSim
12b	$DY \rightarrow ee$ VBF Filter	ALPGEN	CTEQ6L1*	HERWIG	FullSim
12a	$DY \rightarrow \mu\mu$	ALPGEN	CTEQ6L1*	HERWIG	FullSim
12b	$DY \rightarrow \mu\mu$ VBF Filter	ALPGEN	CTEQ6L1*	HERWIG	FullSim
12a	$DY \rightarrow \tau\tau$	ALPGEN	CTEQ6L1*	HERWIG	FullSim
12a	$DY \rightarrow ee$	ALPGEN	CTEQ6L1*	PYTHIA 6	FullSim
12a	$DY \rightarrow \mu\mu$	ALPGEN	CTEQ6L1*	PYTHIA 6	FullSim
12a	$DY \rightarrow \tau\tau$	ALPGEN	CTEQ6L1*	PYTHIA 6	FullSim
12a	$DY \rightarrow ee + bb, cc$	ALPGEN	CTEQ6L1*	PYTHIA 6	FullSim
12a	$DY \rightarrow \mu\mu + bb, cc$	ALPGEN	CTEQ6L1*	PYTHIA 6	FullSim
12a	$DY \rightarrow \tau\tau + bb, cc$	ALPGEN	CTEQ6L1*	PYTHIA 6	FullSim
12b	$DY \rightarrow ee$ (EW)	SHERPA	CT10	SHERPA	FullSim
12b	$DY \rightarrow \mu\mu$ (EW)	SHERPA	CT10	SHERPA	FullSim
12b	$DY \rightarrow \tau\tau$ (EW)	SHERPA	CT10	SHERPA	FullSim
12a	$W \rightarrow e\nu$	ALPGEN	CTEQ6L1	HERWIG	FullSim
12a	$W \rightarrow \mu\nu$	ALPGEN	CTEQ6L1	HERWIG	FullSim
12a	$W \rightarrow \tau\nu$	ALPGEN	CTEQ6L1	HERWIG	FullSim
12a	$W \rightarrow e\nu$	ALPGEN	CTEQ6L1	PYTHIA 6	FullSim
12a	$W \rightarrow \mu\nu$	ALPGEN	CTEQ6L1	PYTHIA 6	FullSim

– continued on following page

– continued from previous page

MC12 Version	Process	Generator	PDF	PS	Simulation
12a	$W \rightarrow \tau\nu$	ALPGEN	CTEQ6L1	PYTHIA 6	FullSim
12a	$W \rightarrow \ell\nu + bb$	ALPGEN	CTEQ6L1	HERWIG	FullSim
12a	$W \rightarrow \ell\nu + cc$	ALPGEN	CTEQ6L1	HERWIG	FullSim
12a	$W \rightarrow \ell\nu + c$	ALPGEN	CTEQ6L1	HERWIG	FullSim
12a	$W \rightarrow \ell\nu + bb$	ALPGEN	CTEQ6L1	PYTHIA 6	FullSim
12a	$W \rightarrow \ell\nu + cc$	ALPGEN	CTEQ6L1	PYTHIA 6	FullSim
12a	$W \rightarrow \ell\nu + c$	ALPGEN	CTEQ6L1	PYTHIA 6	FullSim
12b	$t\bar{t}$ (dilepton)	Powheg	CT10	PYTHIA 6	AF2
12a	st $t$ -channel(leptonic)	AcerMC	CTEQ6L1	PYTHIA 6	AF2
12a	st $s$ -channel(leptonic)	POWHEG	CT10	PYTHIA 6	AF2
12a	st $Wt$ -channel(dilepton)	POWHEG	CT10	PYTHIA 6	AF2
12a	$ggF : H \rightarrow \tau\tau \rightarrow \ell\ell/\ell h$	POWHEG	CT10	PYTHIA 8	FullSim
12a	$VBF : H \rightarrow \tau\tau \rightarrow \ell\ell/\ell h$	POWHEG	CT10	PYTHIA 8	FullSim

Table 6.5: Summary table for all MC samples used in the 8 TeV VBF  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  analyses. PDFs marked with “\*” here indicates samples which are reweighted to the MRST PDF set.

# Chapter 7

## Observables and Event Selection

For the  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  analysis, a common preselection is applied to all data and MC simulation events before dividing further by exclusive jet multiplicity cuts, in order to separate out the ggF and VBF Higgs production modes. Within each analysis, the  $ee + \mu\mu$  and  $e\mu + \mu e$  final states are considered separately in order to better target the background processes. The event selection applied to each is very similar, with alterations where necessary to discriminate against specific backgrounds. This Chapter describes first the observables used in the VBF  $H \rightarrow WW^{(*)}$  analysis, before moving on to the common preselection used. A detailed event selection for both the VBF BDT and cut-based analyses is then given, with a separate section listing modifications made for the 7 TeV analyses.

### 7.1 Observables

A number of observables are used within the analyses, designed to either reject certain reducible backgrounds, or to exploit the Higgs decay topology over the irreducible  $WW$  background. Where mentioned, “leading” and “subleading” refer to the highest  $p_T$  object of that kind, and the object with the next highest  $p_T$ . A brief description and definition of the variables used in the BDT and cut-based analyses is given below:

- $\Delta\phi_{\ell\ell}$ :  
Azimuthal angle between the two leptons produced in the hard scattering interaction. Due to the spin zero Higgs boson in the initial state and the  $V - A$  structure of the weak interaction, the leptons produced in the  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  decay tend to be more collimated than those produced, for example from the irreducible  $WW$  background.
- $m_{\ell\ell}$ :  
The invariant mass of the two leptons originating from the leptonic decay of the  $W$  bosons. For the VBF Higgs signal,  $m_{\ell\ell}$  tends towards low values for a low Higgs mass.

- $\cancel{E}_T$  and  $\cancel{E}_T^{\text{jet,Track}}$ :

Two missing transverse energy definitions used in the VBF analyses, calculated from various transverse momenta sums. They are defined as:

$$\cancel{E}_T = - \sum_i p_T^{\rightarrow i, Calo} \quad (7.1)$$

$$\cancel{E}_T^{\text{jet,Track}} = - \sum_i p_T^{\rightarrow i, Track} + \sum_{j, jets} \left( p_T^{\rightarrow j, Track} - p_T^{\rightarrow j, Calo} \right) \quad (7.2)$$

where ‘‘Calo’’ and ‘‘Track’’ refer respectively to the calorimeter objects or tracks included in the calculation, the sum over  $i$  includes all identified particles and any detector energy measurements, and the sum over  $j$  refers to all reconstructed jets.  $\cancel{E}_T^{\text{jet,Track}}$  improves the missing transverse energy by  $\mathcal{O}(10\%)$  and the use of both missing transverse energy definitions together allows for a better overall background rejection. This has been discussed already in greater detail in Section 4.3.

- $m_T^{\text{jet,Track}}$ :

Transverse mass of the system, defined as follows:

$$m_T^{\text{jet,Track}} = \sqrt{(E_T^{\ell\ell} + \cancel{E}_T^{\text{jet,Track}})^2 - |\mathbf{p}_{T,\ell\ell} + \mathbf{E}_T^{\text{jet,Track}}|^2}, \quad (7.3)$$

where  $p_{T,\ell\ell}$  is the transverse momentum of the dilepton system and the transverse energy of the dilepton system  $E_T^{\ell\ell} = \sqrt{|\mathbf{p}_{T,\ell\ell}|^2 + m_{\ell\ell}^2}$ .  $\mathbf{p}_{T,\ell\ell}$  and  $\mathbf{E}_T^{\text{jet,Track}}$  are the vector sums of their respective quantities. The  $\cancel{E}_T^{\text{jet,Track}}$  variable is used here in the definition to improve the  $m_T$  resolution.

- $|\mathbf{m}_{\ell\ell} - \mathbf{m}_Z| > 15 \text{ GeV}$  (**Z-mass Veto**):

The Z-Veto requires the  $m_{\ell\ell}$  to fall outside of a 15 GeV mass window around the mass of the Z-boson,  $m_Z$ . This cut is applied at preselection level in the  $ee + \mu\mu$  channel and removes events in which the leptons have their origin in  $Z/\gamma^* \rightarrow \ell\ell$  decays.

- $m_{\tau\tau} < 66.1876 \text{ GeV}$  (**Z/ $\gamma^*$ /H  $\rightarrow \tau\tau$ -Veto**):

Similar to the Z-Veto, the Z/ $\gamma^*$ /H  $\rightarrow \tau\tau$ -Veto requires the mass of the di-tau system,  $m_{\tau\tau}$ , to be 25 GeV less than the mass of the Z-boson,  $m_Z$ . To calculate  $m_{\tau\tau}$ , the collinear approximation method [110] is used, which assumes that the charged leptons observed are the decay products of a pair of  $\tau$  leptons, and that they are collinearly emitted with the neutrinos. As the neutrinos are the only source of missing transverse energy in the event, the energy fractions carried by the neutrinos can be determined. A value for  $m_{\tau\tau}$  can then be calculated if the energy fractions are physical. Consistent with the rest of the analysis, the  $\cancel{E}_T^{\text{jet,Track}}$  is used within the calculation. This cut is applied to suppress Z/ $\gamma^* \rightarrow \tau\tau$  decays, and also H  $\rightarrow \tau\tau$  decays.

- **$N_{jets}$ :**  
The number of jets,  $N_{jets}$ , as defined by the anti- $k_T$  algorithm with the selection criteria described in Section 4.2.
- **$N_{b-jet}=0$  (b-Jet Veto):**  
The number of  $b$ -jets,  $N_{b-jet}$ , in an event are required to be exactly zero, in order to reduce top quark production backgrounds. The MV1 algorithm is used at an efficiency working point of 85%, as described in Section 4.2.1.
- **$\mathbf{p}_T^{tot,jet,Track}$ :**  
The total transverse momentum  $p_T^{tot,jet,Track}$  is calculated using the vector sum of the jet-corrected track-based missing transverse energy  $\cancel{E}_T^{jet,Track}$ . It is defined as follows:

$$p_T^{tot,jet,Track} = |\mathbf{p}_T^{lead\ lep} + \mathbf{p}_T^{sublead\ lep} + \cancel{E}_T^{jet,Track} + \sum_{jet} \mathbf{p}_T^{jet}| \quad (7.4)$$

Here,  $\mathbf{p}_T^{leadlep}$ ,  $\mathbf{p}_T^{subleadlep}$  and  $\mathbf{p}_T^{jets}$  are the transverse momentum vectors of the leading lepton, sub-leading lepton and jets present in the event. The sum over jets here includes only those which have passed our jet definition as defined above in Section 4.2. This variable is helpful in suppressing events with appreciable soft gluon radiation that recoils against the dilepton+dijet system, with no high  $p_T$  jets.

- **$\Delta Y_{jj}$ :**  
Difference in rapidity of the tagging jets. The VBF signal is characterised by a large rapidity separation of the two tagging jets.
- **$m_{jj}$ :**  
The tagging jets are expected to be highly energetic and thus, a large invariant mass of the tagging jets  $m_{jj}$  is required.
- **(Central Jet Veto) CJV :**  
This veto removes any events if they contain good reconstructed jets within the pseudorapidity range spanned by the tagging jets, with  $p_T > 20$  GeV. From the VBF topology, little hadronic activity is expected between the tagging jets, and thus this observable is extremely helpful in removing unwanted backgrounds. Although additional uncertainties arise from the requirements on further sub-threshold jets, the gain in signal sensitivity outweighs the effect of the extra uncertainties.
- **(Outside Lepton Veto) OLV:**  
In a similar vein to the CJV, the two charged leptons are required to have pseudorapidities within the pseudorapidity range spanned by the tagging jets, as the leptonic Higgs decay products are expected here.

- $\eta_{lep}$  centrality:

This continuous observable quantifies the exact locations of the leptons with respect to the two tagging jets in the  $\eta$ -plane and is used as an input variable to the BDT analysis.  $\eta_{lep}$  centrality is an extension to the OLV idea and is given by:

$$\begin{aligned}
 OLV_{lead\ lep} &= 2 \cdot \left| \frac{\eta_{lead\ lep} - \bar{\eta}}{\eta_{lead\ jet} - \eta_{sublead\ jet}} \right| \\
 OLV_{sublead\ lep} &= 2 \cdot \left| \frac{\eta_{sublead\ lep} - \bar{\eta}}{\eta_{lead\ jet} - \eta_{sublead\ jet}} \right| \\
 \eta_{lep}\ centrality &= OLV_{lead\ lep} + OLV_{sublead\ lep} \tag{7.5}
 \end{aligned}$$

where  $\bar{\eta} = (\eta_{lead\ jet} + \eta_{sublead\ jet})/2$  is the average pseudorapidity  $\eta$  of the two tagging jets. For each lepton,

$$OLV \begin{cases} = 0 & : \text{lepton is exactly in the centre of the dijet pseudorapidity gap} \\ < 1 & : \text{lepton lies within the dijet pseudorapidity gap} \\ > 1 & : \text{lepton is outside the dijet pseudorapidity gap} \end{cases} \tag{7.6}$$

- $\sum_{\ell,j} M_{\ell j}$ :

Sum of the invariant masses of the four possible lepton-tagging jet pairs, which tends towards higher values for the VBF Higgs signal compared to the other backgrounds. This occurs due to the fact that in the VBF topology the tagging jets are typically more forward, whereas the leptons tend to be produced more centrally, resulting in larger  $\sum_{\ell,j} M_{\ell j}$  values.

## 7.2 Event Preselection

The common preselection is a minimal collection of cuts, after which jet multiplicity cuts and further analysis requirements are imposed. A summary of the preselection cuts is given in Table 7.1, and the plots in Figures 7.1 and 7.2 present the understanding of the reconstructed objects and observables at this stage in the analysis, for both the 7 TeV and 8 TeV analyses. It can be seen that the MC simulation shows an acceptable agreement with the data at this level. No other corrections apart from the reweighting due to pile-up and average number of interactions per bunch crossing is applied here.

## 7.3 Event Selection

Both the BDT and cut-based analyses use largely the same set of observables, with the BDT using a subset as input variables to the training. After the common preselection cuts, the two analyses begin to diverge and will thus be described separately in the following subsections. An overview of the final observable usage for both BDT and cut-based analyses is presented in Table 7.3.

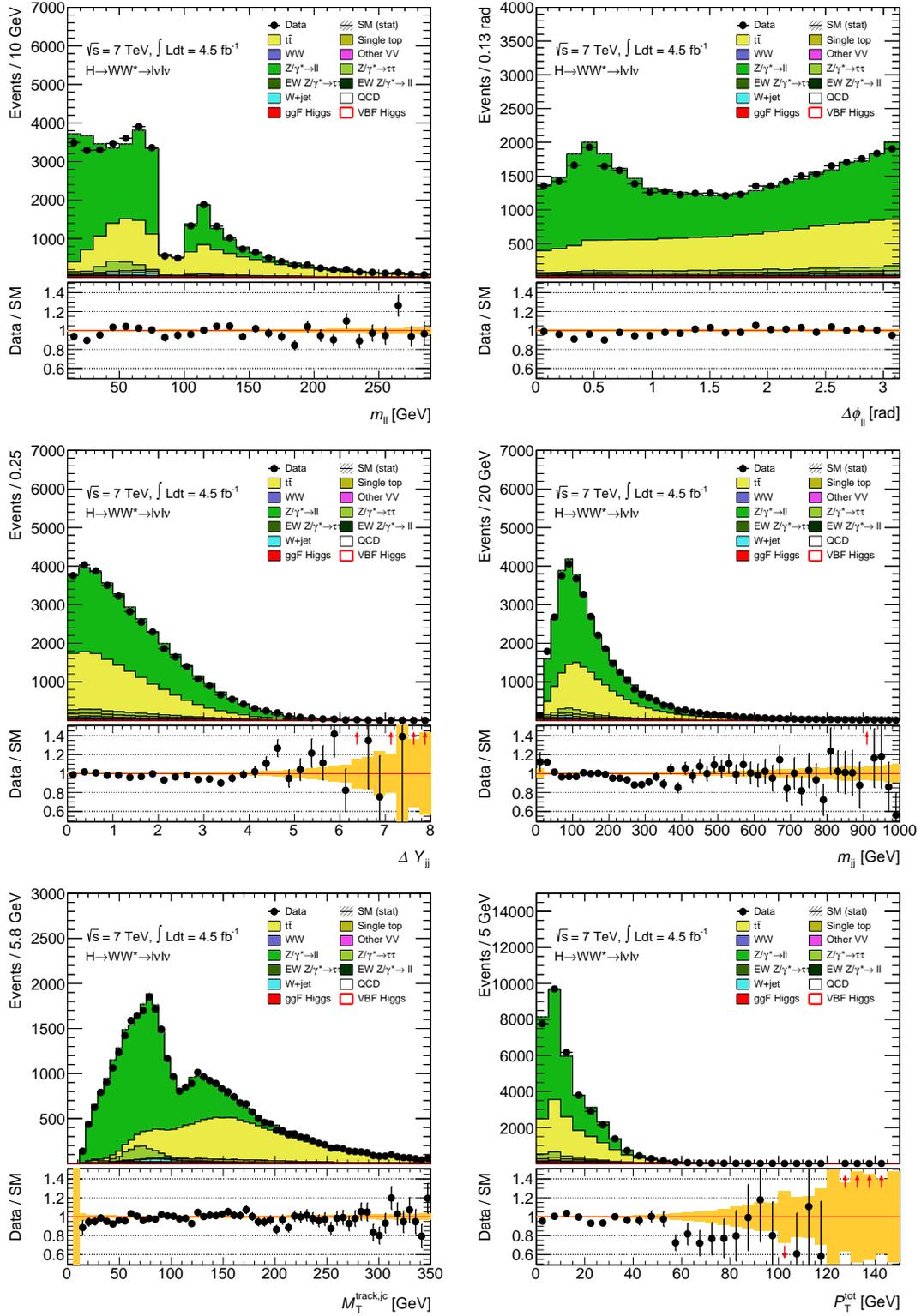


Figure 7.1: Distributions of  $m_{\ell\ell}$ ,  $\Delta\phi_{\ell\ell}$ ,  $\Delta Y_{jj}$ ,  $m_{jj}$ ,  $m_T^{\text{jet,Track}}$ ,  $p_T^{\text{tot,jc,Track}}$ ,  $\cancel{E}_T$ ,  $\cancel{E}_T^{\text{jet,Track}}$ ,  $\eta_{lep}$  centrality and  $\sum_{\ell,j} M_{\ell j}$  at the preselection stage, for all lepton channels combined in the 7 TeV dataset. The yellow uncertainty band shown represents only the statistical MC error in the data/MC ratio plots below the main distributions.

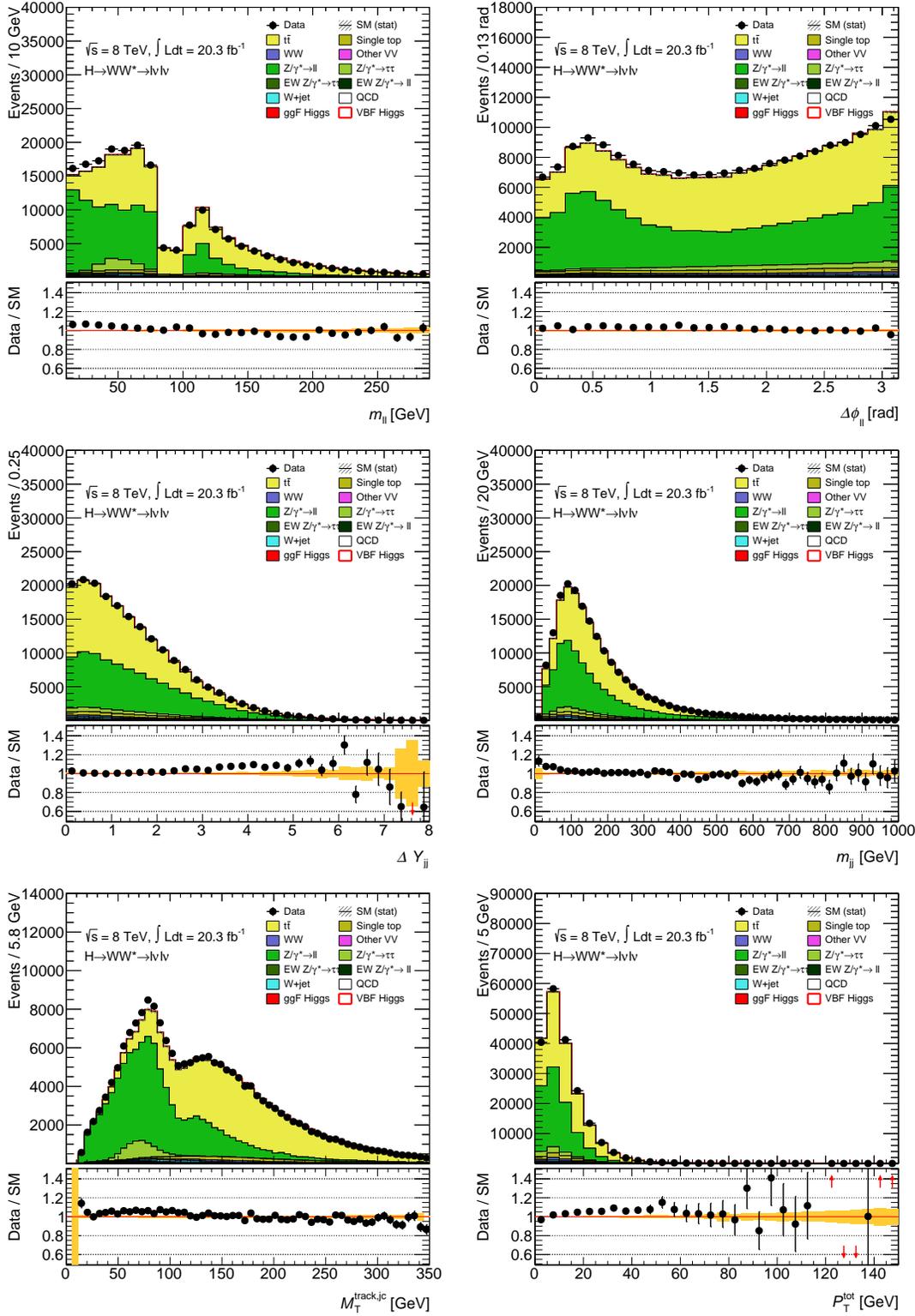


Figure 7.2: Distributions of  $m_{\ell\ell}$ ,  $\Delta\phi_{\ell\ell}$ ,  $\Delta Y_{jj}$ ,  $m_{jj}$ ,  $m_T^{\text{jet},\text{Track}}$ ,  $p_T^{\text{tot},\text{jet},\text{Track}}$ ,  $\cancel{E}_T$ ,  $\cancel{E}_T^{\text{jet},\text{Track}}$ ,  $\eta_{\text{lep}}$  centrality and  $\sum_{\ell,j} M_{\ell j}$  at the preselection stage, for all lepton channels combined in the 8 TeV dataset. The yellow uncertainty band shown represents only the statistical MC error in the data/MC ratio plots below the main distributions.

$e\mu + \mu e$ channel	$ee + \mu\mu$ channel
Exactly two oppositely charged leptons	
$p_T^{\text{lead lep}} > 22 \text{ GeV}, p_T^{\text{sublead lep}} > 10 \text{ GeV}$	
$m_{\ell\ell} > 10 \text{ GeV}$	$m_{\ell\ell} > 12 \text{ GeV}$
–	$ m_{\ell\ell} - m_Z  > 15 \text{ GeV}$ ( $Z$ -mass veto)

Table 7.1: Total preselection as applied to all analyses, depending on the final state leptons.

### 7.3.1 Blinding

By carrying out the analyses using the full range of data collected, there is a risk of introducing a bias into the results. Certain assumptions have been made prior to the execution of the analysis, for example that the SM Higgs mass  $m_H \approx 125 \text{ GeV}$ , or about the VBF Higgs topology and thus where the signal-sensitive region should occur. Given the finite amount of data available, the event selection and any subsequent analysis optimisation could be based on the individual properties of the data used, as opposed to a selection based on general properties expected from the theoretical predictions.

To avoid this bias, these analyses are carried out and optimised *blindly* at first. This refers to the practice of excluding the region in data, in which the Higgs signal is expected to be found. Only after the analyses have been shown to be robust and all procedures fixed, are the signal-like data events included. All parameter or variable optimisations carried out in this thesis are blinded. The use of data in background-dominated *control regions* in order to check the modelling of background processes and to calculate correction factors where appropriate, is allowed.

### 7.3.2 BDT event selection and training

In this analysis, the BDT is trained using both  $e\mu + \mu e$  and  $ee + \mu\mu$  channels combined, as the majority of backgrounds are insensitive to lepton flavour. Attempts at using dedicated BDTs for each channel yielded no significant gain - by having one combined BDT training for both lepton channels however, the analysis profits from the full statistical power of the training samples. The preselection for the BDT training consists of the common preselection cuts described in Section 7.2, except with the change that  $p_T^{\text{sublead lep}} > 15 \text{ GeV}$ . In other words, the training is performed without “low- $p_T$ ” events - those in which the subleading lepton  $p_T$  ranges from  $10 < p_T^{\text{sublead lep}} < 15 \text{ GeV}$ . The inclusion of such low- $p_T$  events in the training brought only marginal improvements in expected sensitivity and would have required the recomputation of many theoretical

uncertainties. In addition the following cuts are used:  $N_{jets} \geq 2$ ,  $b$ -veto,  $\cancel{E}_T > 45$  GeV,  $\cancel{E}_T^{jet,Track} > 40$  GeV and  $m_{\ell\ell} < 75$  GeV; the latter three cuts being for the  $ee + \mu\mu$  channel only, to further suppress  $Z/\gamma^* \rightarrow \ell\ell$  background. The Toolkit for Multivariate Analysis with ROOT (TMVA) [67] package was used to implement the BDT. Of the physics processes described in Section 6.2, the following MC samples are used in the training:

- **Signal:** VBF  $H \rightarrow WW^{(*)}$
- **Background:** ggF  $H \rightarrow WW^{(*)}$ ,  $t\bar{t}$ , single top,  $Z/\gamma^* \rightarrow \ell\ell$ , diboson backgrounds

Associated Higgs production channels ( $WH/ZH$ ) are not included in the training as their contribution is negligible and furthermore, their inclusion causes the BDT algorithm to become confused about defining a signal-like phase space. This can be seen in Figure 7.3, where some small overtraining at low BDT scores can be seen. The data-driven  $W$ +jets and QCD multijet backgrounds are also left out as their background estimation method can occasionally produce negatively weighted events. These can not be treated properly in the BDT training if the gradient boosting method is used. However as  $W$ +jets and multijet processes make up less than 5% of the total background in the signal-sensitive region, the impact of leaving these two processes out of the training is expected to be minimal [96].

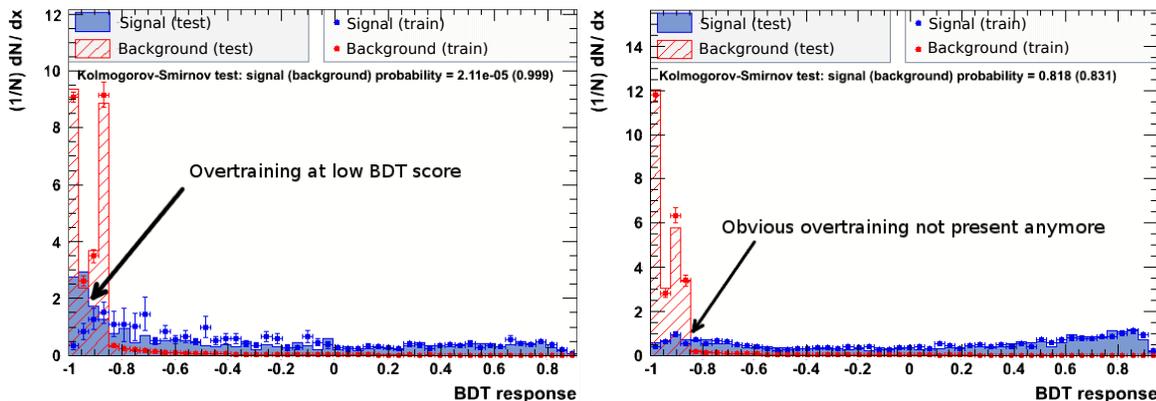


Figure 7.3: Distributions of BDT response with  $WH$  and  $ZH$  processes included (left) and without (right). The signal processes are shown in blue, with the background processes shown in red. The points show the training sample, superimposed onto the solid histograms which represent the testing sample.

To determine the optimal TMVA parameters, a *grid scan* was conducted using  $S/\sqrt{B}^1$  as a figure of merit. This entailed the discrete variation of the  $N$  TMVA

<sup>1</sup>For a number of signal events  $S$  and background events  $B$ , the Gaussian approximation for the expected Poisson significance is  $S/\sqrt{B}$ .

parameters, thus forming an  $N$ -dimensional grid with which the ideal set of parameter values could be identified. Beginning at the high BDT score (most signal-like) boundary, the  $S/\sqrt{B}$  is calculated for a region increasing in size, in steps of BDT score, for each grid point in order to compare the performance. The final parameters chosen are summarised in Table 7.2.

TMVA parameter	Value
Boosting algorithm	Gradient
Maximum tree depth	5
Number of trees	1000
Minimum number of events required per mode	1000
Use bagging in gradient boosting	true
Bagging fraction	0.25
Shrinkage (algorithm learning rate)	0.125

Table 7.2: TMVA parameters used in the analysis.

Eight training variables were input to the BDT:  $\Delta\phi_{\ell\ell}$ ,  $m_{\ell\ell}$ ,  $\Delta Y_{jj}$ ,  $m_{jj}$ ,  $p_T^{tot,jet,Track}$ ,  $m_T$ ,  $\eta_{lep}$  *centrality* and  $\sum_{\ell,j} M_{\ell j}$ . The distributions of these variables can be seen in Figures 7.4 and 7.5 for the 8 TeV BDT analysis, illustrating the difference in shape between that of the signal and background processes. This indicates that the input variables are good discriminators between the signal and backgrounds. The eight variables used were chosen using a “ $N - 1$  minimal loss variable pruning” procedure, using the expected significance, modified to include the statistical uncertainty on the background rate, as a figure of merit:

- Begin with a BDT trained on a large collection of  $N$  potentially useful discriminating variables
- Remove one variable at a time and evaluate the BDT output performance of the  $(N - 1)$ -variable BDT based on the expected significance  $S/\sqrt{B + \Delta B}$ , where  $\Delta B$  represents the statistical uncertainty on the background rate
- As for the parameter optimisation, calculate  $S/\sqrt{B + \Delta B}$  in steps of BDT score, in order to compare performance
- Use the best performing  $(N - 1)$ -variable BDT as the new benchmark
- Iterate procedure until the BDT performance drops significantly, around  $\mathcal{O}(10\%)$

The available data and MC samples are divided randomly into two, such that a BDT can be trained on each half, and then applied using TMVA to the other half of the events to avoid overtraining. Both halves of the data and MC samples are then

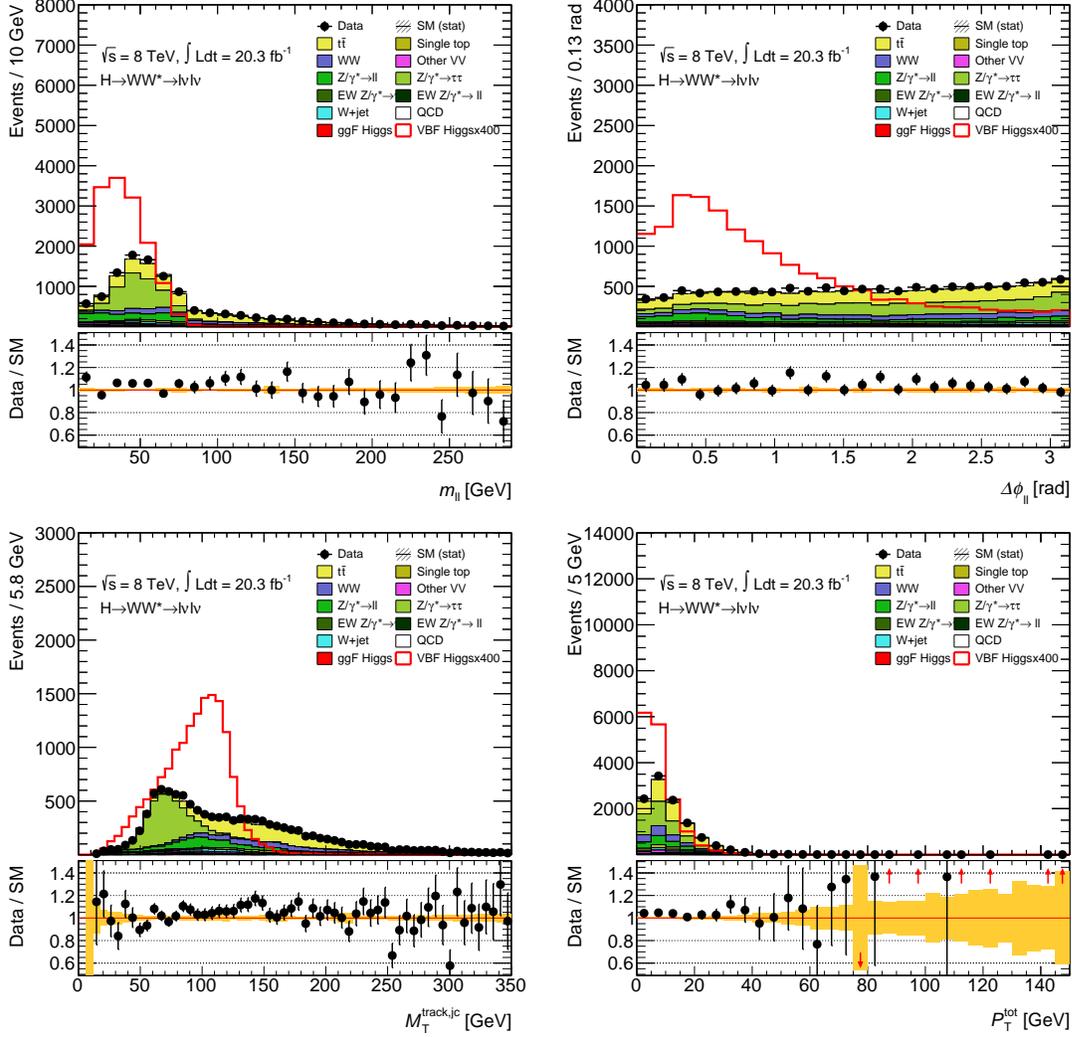


Figure 7.4: Distributions of the input variables which enter the BDT training,  $m_{\ell\ell}$ ,  $\Delta\phi_{\ell\ell}$ ,  $m_T$  and  $p_T^{\text{tot},jet,Track}$  after the BDT training selection. The VBF signal is superimposed in red, scaled by a factor of 400 to show the difference in shape between it and the background processes. Shown for all lepton channels combined in the 8 TeV dataset. The yellow uncertainty band shown represents only the statistical MC error.

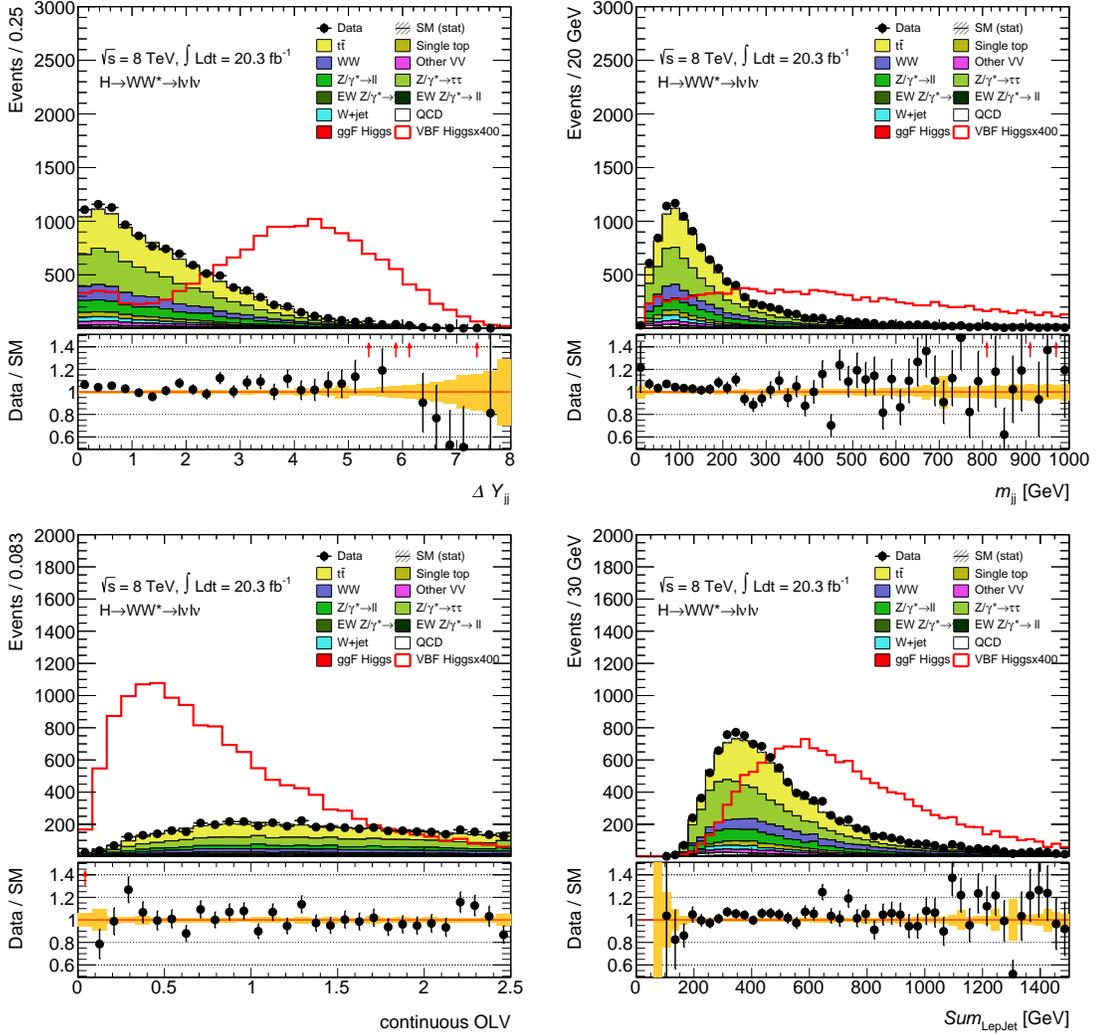


Figure 7.5: Distributions of the input variables which enter the BDT training,  $\Delta Y_{jj}$ ,  $m_{jj}$ ,  $\eta_{lep}$  centrality and  $\sum_{\ell,j} M_{\ell j}$ , after the BDT training selection. The VBF signal is superimposed in red, scaled by a factor of 400 to show the difference in shape between it and the background processes. Shown for all lepton channels combined in the 8 TeV dataset. The yellow uncertainty band shown represents only the statistical MC error.

rejoined together. The training is applied to all events with  $p_T^{\text{sublead lep}} > 10$  GeV (consistent with the common preselection). Additional CJV, OLV and  $Z/\gamma^* \rightarrow \tau\tau$  veto cuts are imposed to further enrich the signal region (SR).

The distribution of the BDT score is used in the final fits of the analysis, with a binning optimised for the maximal sensitivity to the VBF Higgs process to enable the full separation power of the BDT method to be exploited. Whilst the primary object is to enhance the sensitivity, care is taken to avoid optimising on statistical fluctuations and also the depletion of background processes from regions, which then cannot obtain a reliable background estimate and the associated systematic errors. Beginning at the high BDT score (most signal-like) boundary, the expected Poisson significance is calculated for a region, increasing in size, in steps of BDT score of size 0.02. Once a significance maximum is reached, and each background process is confirmed to be present, a boundary is set and the next iteration then begins at the new boundary.

In this way, the resulting binning for the BDT analysis is achieved. The low BDT score region,  $-1 < \text{BDT} < -0.48$  is not considered as it contains the majority of the background ( $> 90\%$ ) and has negligible signal presence. The remaining BDT output region is divided into three bins and define the signal fit region, which is *blinded* for data events during the initial stage of the analysis:  $-0.48 < \text{BDT} < 0.3$ ,  $0.3 < \text{BDT} < 0.78$  and  $0.78 < \text{BDT} < 1.0$ . Throughout the rest of this thesis, the three BDT signal bins will be denoted by the lower bound of each region as follows:  $[-0.48, 0.3, 0.78]$ . The mechanics of the actual fit are discussed in greater detail in Chapter 9. The same BDT trained with 8 TeV MC simulation is used both in the 8 TeV analysis and the 7 TeV re-analysis, due to the lower statistical power of the 7 TeV MC samples and to facilitate the combination of results, as many theoretical uncertainties can be backported. A summary of the BDT binning used in the BDT analyses can be seen in Table 7.4.

### 7.3.3 Cut-based selection

A traditional “cut-based” analysis, in which requirements are imposed on observables to obtain a signal-enriched region, was also developed which is also used independently alongside the BDT analysis as a cross-check. Some of the cuts which feature in the BDT analysis such as the  $Z/\gamma^* \rightarrow \tau\tau$  veto and CJV were designed originally for the cut-based analysis, but were found to improve the sensitivity of the BDT analysis as well. The cut-based analysis is especially useful for investigating the background estimation techniques and associated systematic errors induced, which are described in Chapter 8. A pictorial representation of the cuts used, along with their impact on the signal and background yields in the  $e\mu + \mu e$  and  $ee + \mu\mu$  flavour channels can be seen in Figure 7.6

After the common preselection, the cuts are imposed in the order described in Table 7.3. Although physically the cut order has no overall effect, this sequence of cuts has been chosen in order to ensure that enough events remain in order to estimate the

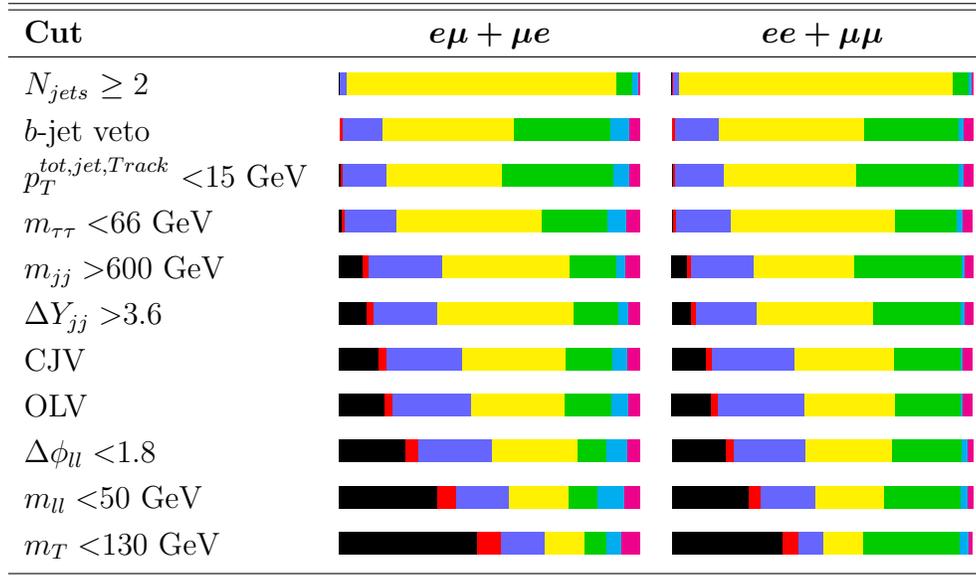


Figure 7.6: Representation of the 8 TeV cut-based analysis, split into the  $e\mu + \mu e$  (left) and  $ee + \mu\mu$  (right) flavour channels. The total event yield is represented within each bar, with the percentage composition of each process shown at the various cut stages. The processes shown from left to right are VBF (black), ggF (red),  $WW$  production (blue), top backgrounds (yellow),  $Z/\gamma^* \rightarrow \ell\ell$  (green),  $W$ +jets and QCD multijet (cyan) and lastly non- $WW$  diboson processes (purple).

backgrounds properly in each region. The blinding is implemented for the cut-based analysis by removing all data events which pass the event selection after the common preselection. Although the blinding criteria are not fully efficient, the remaining Higgs signal was estimated to be less than 1.7% of the SM prediction [96] and thus negligible.

For the cut-based analysis, the “steepest descent method” was used to optimise the cut-values for each observable, using the expected significance as a figure of merit. This method converges towards local minima/maxima and is described in greater detail in Ref. [96]. Optimisation for the  $m_{\ell\ell}$ ,  $\Delta\phi_{\ell\ell}$ ,  $p_T^{\text{tot},jet,Track}$ ,  $m_{jj}$  and  $\Delta Y_{jj}$  cuts was performed in the  $e\mu + \mu e$  channel, as it has dominant sensitivity to the VBF signal - these values were then adopted by the  $ee + \mu\mu$  channel. In contrast, the optimisation for the  $\cancel{E}_T$  and  $\cancel{E}_T^{\text{jet,Track}}$  cuts was conducted in the  $ee + \mu\mu$  channel, as they are applied only here. As for the BDT procedure, ggF-produced  $H \rightarrow WW^{(*)}$  is considered as a background during the optimisation. All other cuts are necessary by definition due to the desired VBF signal ( $N_{jets}$ ,  $b$ -veto) or exist as a veto (CJV, OLV,  $Z/\gamma^* \rightarrow \tau\tau$  veto).

Within the cut-based analysis, the  $m_T$  distribution is used in the statistical fit procedure as it has good discrimination between signal and background yields. To select the most signal-sensitive region for the cut-based analysis, the bin boundaries were optimised using the relative expected significance<sup>2</sup> as a figure of merit, with a method which parametrised the signal and background  $m_T$  shapes with continuous functions [96]. The ensuing bin boundaries in the  $m_T$  fit are denoted by their lower bounds [0, 80, 130] GeV, where the upper bound of the third bin is set at  $\infty$ . As shown for instance in Figure 7.5, the VBF Higgs signal and the background processes have different shapes as a function of  $m_{jj}$ , where the background tails off faster than the VBF signal and thus, the SR in the  $e\mu + \mu e$  channel is further subdivided into two bins in  $m_{jj}$  with the partition at 1 TeV. This can be seen in greater detail in Figure 7.7. No division in  $m_{jj}$  for the  $ee + \mu\mu$  channel is implemented due to low statistics. The  $m_{jj}$ -split brings an additional improvement of  $\sim 7\%$  to the expected significance for the cut-based analysis [96]. The exact  $m_T$  fit procedure is discussed in Chapter 9.

---

<sup>2</sup>The relative expected significance is defined as the ratio  $Z/Z_{reference}$ , where  $Z$  is the expected significance, and  $Z_{reference}$  the expected significance from the unbinned fit.

	Observable	Cut-based value	BDT usage
Common preselection	lepton $p_T$	$p_T^{\text{lead lep}} > 22 \text{ GeV}$ , $p_T^{\text{sublead lep}} > 10 \text{ GeV}$	
	$m_{\ell\ell}$	$> 10 \text{ GeV}$ for $e\mu + \mu e$ $> 12 \text{ GeV}$ for $ee + \mu\mu$	
	Lepton charge	opposite	
	Z-Veto	$ m_{\ell\ell} - m_Z  > 15 \text{ GeV}$ , for $ee + \mu\mu$ only	
VBF preselection	$N_{jets}$	$\geq 2$	
Background rejection	$\cancel{E}_T (e\mu + \mu e)$	no cut	
	$\cancel{E}_T (ee + \mu\mu \text{ only})$	$\cancel{E}_T^{\text{jet,Track}} > 50 \text{ GeV}$ , $\cancel{E}_T > 55 \text{ GeV}$	$\cancel{E}_T^{\text{jet,Track}} > 40 \text{ GeV}$ , $\cancel{E}_T > 45 \text{ GeV}$
	b-Veto	$N_{b-jet}=0$ , MV1 85%	
	$p_T^{\text{tot,jet,Track}}$	$< 15 \text{ GeV}$	INPUT 1
	$Z/\gamma^* \rightarrow \tau\tau$ -Veto	$m_{\tau\tau} < 66 \text{ GeV}$	
VBF topology	$m_{jj}$	$> 600 \text{ GeV}$ , split at 1 TeV	INPUT 2
	$\Delta Y_{jj}$	$> 3.6$	INPUT 3
	CJV	$< 20 \text{ GeV}$	
	OLV	used	
$H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$ topology	$\Delta\phi_{\ell\ell}$	$< 1.8 (p_T^{\text{sublead lep}} > 15 \text{ GeV})$ , $< 2.8$ for low- $p_T$ events	INPUT 4
	$m_{\ell\ell}$	$< 50 \text{ GeV}$	INPUT 5
	$m_T$	$< 130 \text{ GeV}$ , with $m_T > 80 \text{ GeV}$ cut for low- $p_T$ events	INPUT 6
BDT-specific	$\sum_{\ell,j} M_{\ell j}$	-	INPUT 7
	$\eta_{lep} \text{ centrality}$	-	INPUT 8

Table 7.3: Summary of cuts used in the VBF BDT and cut-based analyses, including the common preselection. The cut order shown here is important only for the cut-based analysis, to assure enough events remain for reliable background estimations. BDT training variables are denoted as “INPUT” and the exact training procedure can be found in Section 7.3.2. Low- $p_T$  events are those for which  $10 < p_T^{\text{sublead lep}} < 15 \text{ GeV}$ .

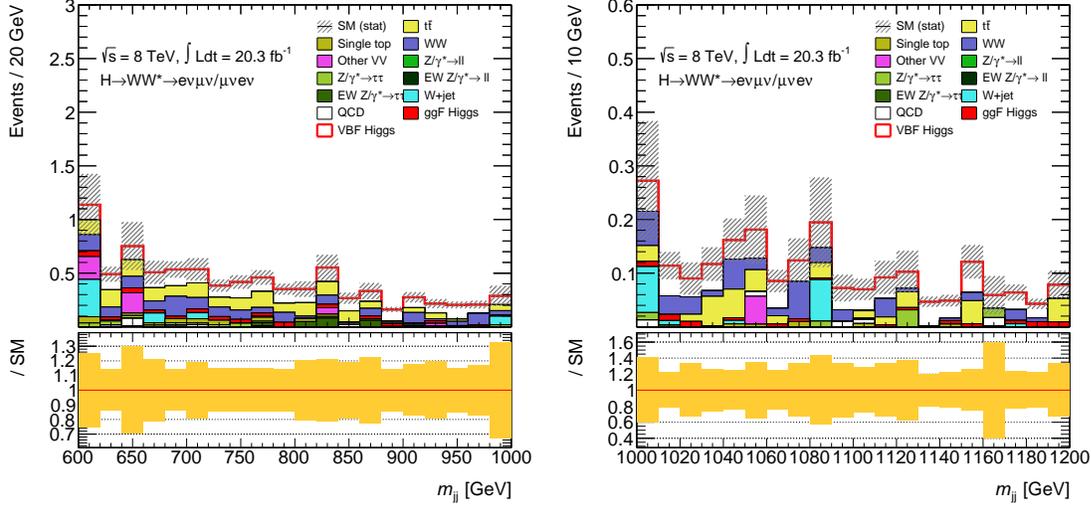


Figure 7.7: Distributions of  $m_{jj}$ , shown at the level of the  $m_T$  cut, which is the last cut applied before the fit is performed. Only the 8 TeV  $e\mu + \mu e$  flavour channel is shown here, to illustrate the motivation for splitting this fit region further according to  $m_{jj}$ . As this is part of the signal region, the data is blinded here. The yellow uncertainty band shown represents only the statistical MC error.

### 7.3.4 Modifications for the 7 TeV Analyses

The analysis of the 7 TeV data sample follows very closely the selection used for the 8 TeV sample - the majority of the differences are found in the object definition, as newer triggers and identification algorithms were developed since the 2011 7 TeV data-taking runs. Similar procedures between the 7 TeV and 8 TeV analyses allow for a more streamlined combination of results. For the 7 TeV BDT analysis, a two-bin  $[-0.48, 0.3, 1.0]$  fit to the BDT response is performed instead in the  $e\mu + \mu e$  channel, with a simple single-bin  $[-0.48, 1.0]$  calculation of the significance used in the  $ee + \mu\mu$  channel due to lower data event yields. Binning choices for the 7 TeV BDT analysis are summarised in Table 7.4. The distributions of the input variables to the BDT algorithm for the 7 TeV analysis can be seen in Figures 7.8 and 7.9.

For the 7 TeV cut-based analysis, a few cut values are relaxed back to those originating from the optimisation conducted in Ref. [111], so that reasonable event yields were retained for all background processes in order to estimate their contribution in the SR, and the associated systematic errors. These differences are listed in Table 7.5, alongside the 8 TeV values. The final SR for the 7 TeV cut-based analysis has a lower event yield than that of the 8 TeV analysis, and thus is not divided further into two  $m_{jj}$  regions, but kept as a single region after the  $m_T$  cut. No fit to the  $m_T$  distribution is performed due to the lower statistics and thus a simple significance calculation is performed in the  $ee + \mu\mu$  and  $e\mu + \mu e$  SRs separately after the cut on  $m_T < 130$  GeV.

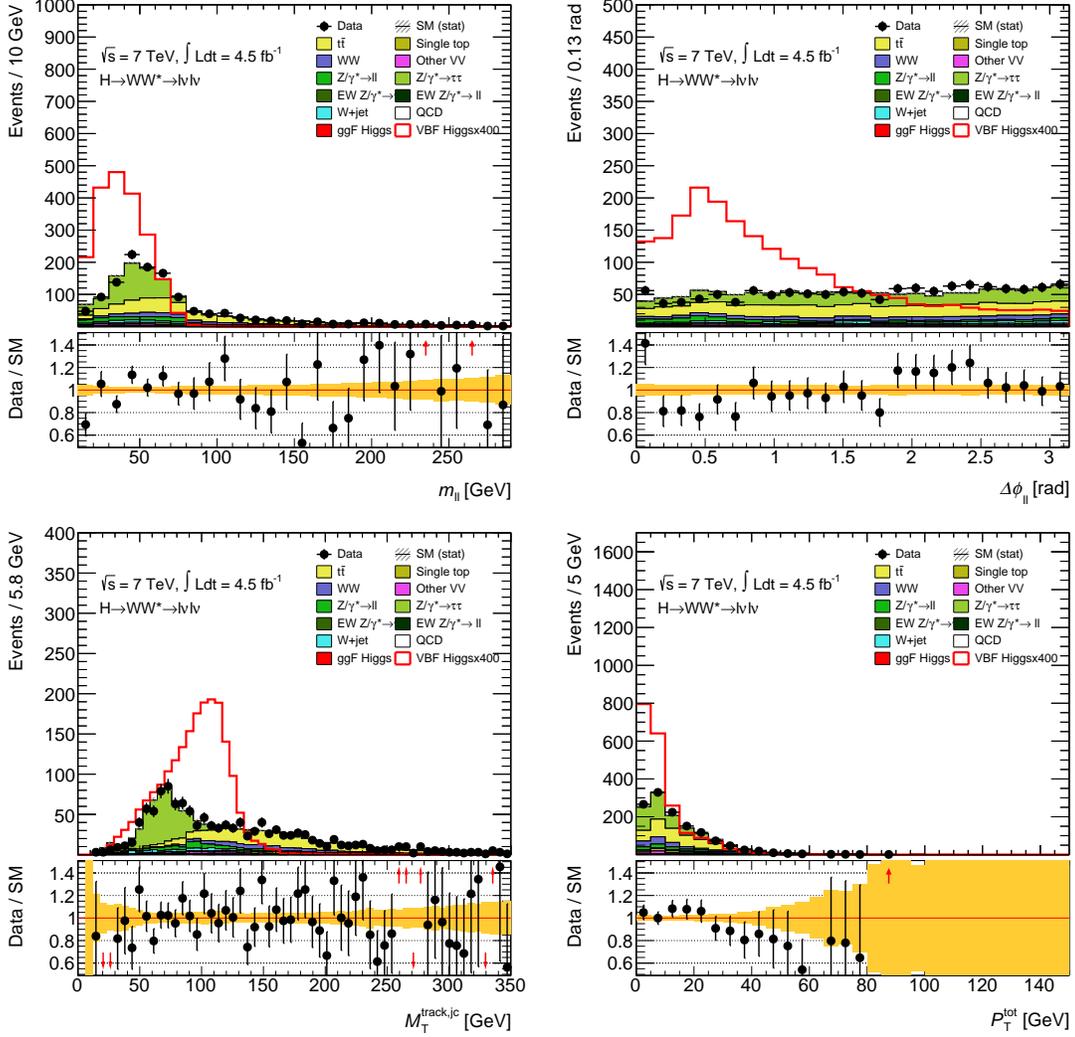


Figure 7.8: Distributions of the input variables which enter the BDT training,  $m_{\ell\ell}$ ,  $\Delta\phi_{\ell\ell}$ ,  $m_T$  and  $p_T^{tot,jet,Track}$  after the BDT training selection. The VBF signal is superimposed in red, scaled by a factor of 400 to show the difference in shape between it and the background processes. Shown for all lepton channels combined in the 7 TeV dataset. The yellow uncertainty band shown represents only the statistical MC error.

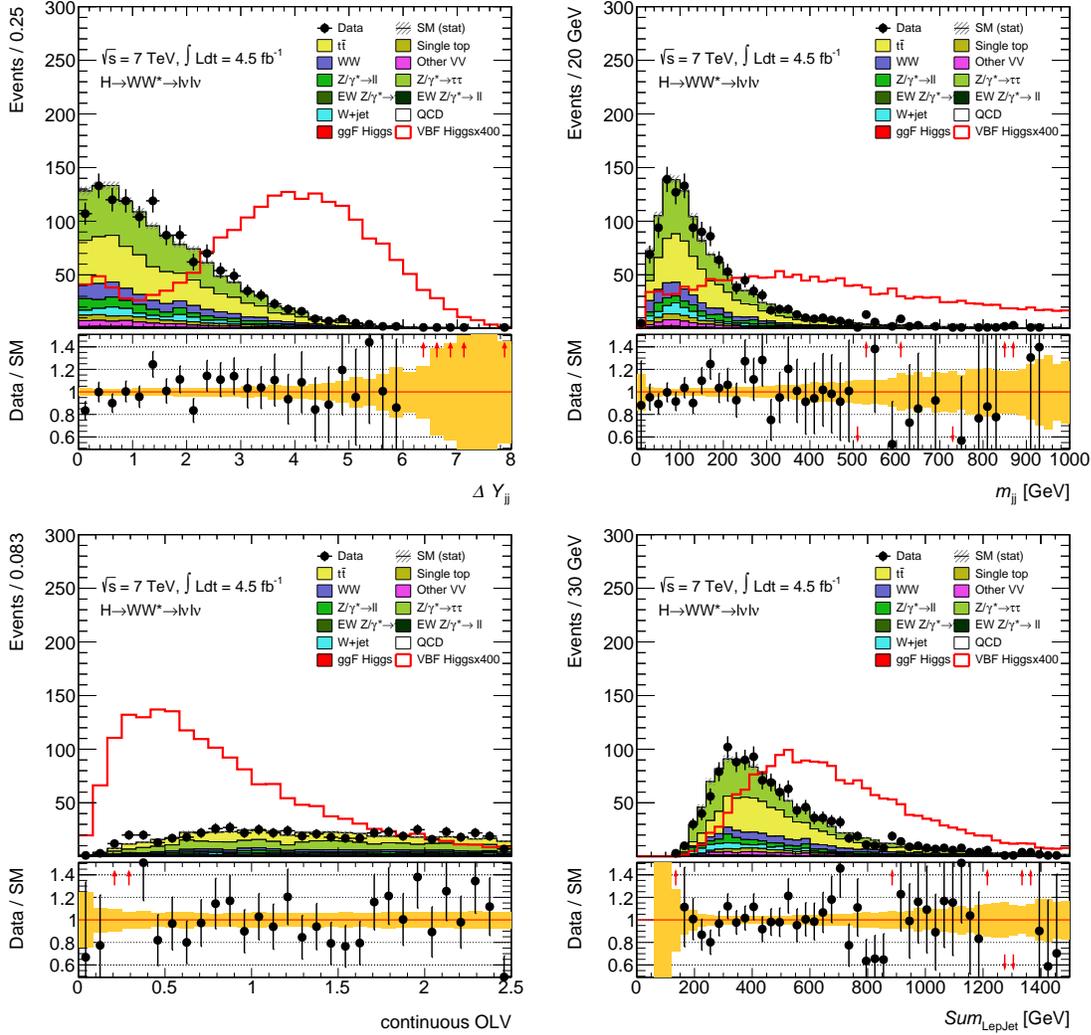


Figure 7.9: Distributions of the input variables which enter the BDT training,  $\Delta Y_{jj}$ ,  $m_{jj}$ ,  $\eta_{lep}$  centrality and  $\sum_{\ell,j} M_{\ell j}$ , after the BDT training selection. The VBF signal is superimposed in red, scaled by a factor of 400 to show the difference in shape between it and the background processes. Shown for all lepton channels combined in the 7 TeV dataset. The yellow uncertainty band shown represents only the statistical MC error.

BDT score	7 TeV		8 TeV
	$e\mu + \mu e$	$ee + \mu\mu$	
$-1 < \text{BDT} < -0.48$	Background bin		
$-0.48 < \text{BDT} < 0.3$	FR1		FR1
$0.3 < \text{BDT} < 0.78$		only one FR	FR2
$0.78 < \text{BDT} < 1.0$	FR2		FR3

Table 7.4: Summary of the different binning used for the 7 TeV and 8 TeV BDT analyses. “FR” represents the fit region used by the analysis.

7 TeV	8 TeV
$\cancel{E}_T > 45 \text{ GeV}$	$\cancel{E}_T > 55 \text{ GeV}$
–	$\cancel{E}_T^{\text{jet,Track}} > 50 \text{ GeV}$
$m_{jj} > 500 \text{ GeV}$	$m_{jj} > 600 \text{ GeV}$
$\Delta Y_{jj} > 2.8$	$\Delta Y_{jj} > 3.6$

Table 7.5: Deviations in the 7 TeV cut-based analysis selection to that of the 8 TeV analysis.



# Chapter 8

## Background Estimations

Different methods for estimating the background processes are utilised by the analysis depending on the MC capabilities, the ability to define a pure control sample and also on the importance of the background. For non- $WW$  diboson backgrounds ( $WZ$ ,  $ZZ$ ,  $Z\gamma$ ,  $W\gamma$  and  $W\gamma^*$ ) and the ggF Higgs process, the estimations rely purely on the MC simulation. In contrast, for all top and  $Z/\gamma^* \rightarrow \ell\ell$  backgrounds, the simulation is normalised to the measured data in control regions (CR) constructed specifically for each background, so that any mismodelling can be accounted for. Each CR is ideally defined to be of high purity - without contamination from other backgrounds or the signal process - and to be orthogonal to that of the signal region (SR). Although it is not feasible to define a CR for the  $WW$  background, a validation region is selected to check the modelling of this important, irreducible background. Data-driven techniques are more suited to the estimation of  $W$ +jets and QCD multijet processes, as these originate from the misidentification of other objects as leptons.

Background estimation techniques in both the BDT and cut-based analyses are given here in this Chapter, with modifications and results for the 7 TeV analyses included at the end. An overview of all the corrections applied to the background yields for the 8 TeV analyses are given in Tables 8.1 and 8.2. Systematic uncertainty calculations associated to the background estimations have been listed here as appropriate. Other systematic uncertainties due to experimental sources are discussed later in Chapter 10, along with the uncertainties related to those backgrounds which have not been described here.

### 8.1 Calculation of Normalisation, Correction and Extrapolation Factors

In general, unless otherwise specified, the procedure to calculate a normalisation factor (NF) from a CR is given as:

$$\text{NF} = \frac{\text{Data} - (\text{Total Bkg.} - B_{CR})}{B_{CR}} \quad (8.1)$$

where ‘‘Data’’ refers to the number of measured data events, ‘‘Total Bkg.’’ to the total number of expected background events from MC simulation (including yields of backgrounds using data-driven estimates) and  $B_{CR}$  to the number of MC events of the background in question, all in the defined CR. The background estimate in the SR is thus normalised to the actual data and is given as:  $NF \times B_{SR}$ , where  $B_{SR}$  is the MC prediction in the SR. The calculated NFs are always quoted with their statistical uncertainties only, unless otherwise specified. Where appropriate for the BDT analyses, the NFs are given per signal-like bin used for the final fit, where the three BDT signal bins are denoted by the lower bound of each bin as follows: [-0.48, 0.3, 0.78].

A correction factor (CF) can also be determined from a CR, which takes into account the difference between the data and MC prediction when applying a set of cuts. It is defined as follows:

$$CF = \frac{N_{\text{data,cuts}}/N_{\text{data}}}{N_{\text{MC,cuts}}/N_{\text{MC}}} \quad (8.2)$$

where  $N_{\text{data,cuts}}$  is the number of data events in the CR after a particular set of cuts have been applied, and  $N_{\text{data}}$  the number of data events without these cuts applied (and likewise for the  $N_{\text{MC,cuts}}$  and  $N_{\text{MC}}$  terms). The CF is used in the cut-based analyses in conjunction to correct for cuts which were not included in the CR of the NF, due to a lack of remaining background events, but do feature in the SR definition.

Another use of the CR is to calculate an extrapolation factor  $\alpha$ , defined as the ratio of MC yields between the SR and CR. The estimate  $B_{SR}^{est}$  of the expected background under consideration in the SR can be written as:

$$B_{SR}^{est} = B_{SR} \cdot \underbrace{N_{CR}/B_{CR}}_{NF} = N_{CR} \cdot \underbrace{B_{SR}/B_{CR}}_{\alpha} \quad (8.3)$$

where the observed yield in the CR is  $N_{CR} = \text{Data} - (\text{Total Bkg.} - B_{CR})$ . The theoretical uncertainties associated with the direct background estimation from MC can be replaced by the combination of two smaller uncertainties - namely the statistical uncertainty on  $N_{CR}$  and the theoretical uncertainty on  $\alpha$ . For the  $Z/\gamma^* \rightarrow ee + \mu\mu$  and misidentification backgrounds ( $W$ +jets and QCD multijet) which use *data-driven* techniques however, the theoretical uncertainties are estimated using other means, as such an extrapolation factor cannot be defined.

## 8.2 Sources of Theoretical Uncertainties

The theoretical uncertainty on a background estimate can be estimated by studying the variation produced on the background yield or the extrapolation factor  $\alpha$  when one of the theoretical assumptions are changed. Standard sources of theoretical uncertainties considered are:

	Applied to (for Background)	NF $\pm$ stat.
$NF_{ABCD,i}^{BDT}$	Each $ee + \mu\mu$ SR bin ( $Z/\gamma^* \rightarrow ee + \mu\mu$ )	[1.02 $\pm$ 0.16, 0.90 $\pm$ 0.28, 0.90 $\pm$ 0.28]
Top NF	Each SR bin (Top)	[1.58 $\pm$ 0.15, 0.95 $\pm$ 0.31, 0.95 $\pm$ 0.31]
$Z/\gamma^* \rightarrow \tau\tau$ NF	SR ( $Z/\gamma^* \rightarrow \tau\tau$ )	1.23 $\pm$ 0.32

Table 8.1: Normalisation factors used in the 8 TeV VBF BDT analysis, including the regions to which they are applied. Where no lepton flavour channel is specified, this refers to both  $ee + \mu\mu$  and  $e\mu + \mu e$  channels combined.

	Applied to (for Background)	NF $\pm$ stat.
$NF_{ABCD}^{cut-based}$	–	0.71 $\pm$ 0.20
$CF_{Z/DY}$	–	1.34 $\pm$ 0.46
$NF_{ZDY}^{cut-based}$	$ee + \mu\mu$ SR ( $Z/\gamma^* \rightarrow ee + \mu\mu$ )	0.95 $\pm$ 0.50
Top NF	SR (Top)	1.04 $\pm$ 0.19
$NF_{\tau\tau}$	–	0.97 $\pm$ 0.04
$CF_{\tau\tau} (ee + \mu\mu)$	–	1.26 $\pm$ 0.05
$CF_{\tau\tau} (e\mu + \mu e, m_{jj} < 1 \text{ TeV})$	–	1.24 $\pm$ 0.06
$CF_{\tau\tau} (e\mu + \mu e, m_{jj} > 1 \text{ TeV})$	–	1.34 $\pm$ 0.10
$Z/\gamma^* \rightarrow \tau\tau$ NF ( $ee + \mu\mu$ )	$ee + \mu\mu$ SR ( $Z/\gamma^* \rightarrow \tau\tau$ )	1.22 $\pm$ 0.06
$Z/\gamma^* \rightarrow \tau\tau$ NF ( $e\mu + \mu e, m_{jj} < 1 \text{ TeV}$ )	$e\mu + \mu e, m_{jj} < 1 \text{ TeV}$ SR ( $Z/\gamma^* \rightarrow \tau\tau$ )	1.20 $\pm$ 0.07
$Z/\gamma^* \rightarrow \tau\tau$ NF ( $e\mu + \mu e, m_{jj} > 1 \text{ TeV}$ )	$e\mu + \mu e, m_{jj} > 1 \text{ TeV}$ SR ( $Z/\gamma^* \rightarrow \tau\tau$ )	1.30 $\pm$ 0.11

Table 8.2: Normalisation factors used in the 8 TeV VBF cut-based analysis, including the regions to which they are applied. Where no lepton flavour channel is specified, this refers to both  $ee + \mu\mu$  and  $e\mu + \mu e$  channels combined.

- **QCD scale:**

Missing higher order terms in the calculation of a process can affect the kinematics, and thus the value of  $\alpha$ . Such uncertainties are obtained by varying the QCD factorisation and renormalisation scales, which regulate the infra-red and ultraviolet cutoff points within the the QCD theory. These scales are varied by a factor of two in both directions, and the resulting  $\alpha$  values are then compared.

- **Parton distribution function:**

This estimates the impact of the choice of parton distribution function (PDF) sets on the process yields, and is achieved by simply changing the PDF set.

- **Parton showering:**

The effect of the choice of parton shower (PS) model can be studied by comparing the obtained  $\alpha$  values using a fixed event generator, but with differing parton shower generators. This also includes any effects resulting from the *underlying event*, which refers to all the additional radiation and hadronisation processes produced, excluding the hard scattering process.

- **Generator modelling:**

Using different matrix element level generators, together with the same PS generator, allows an estimate of the uncertainty due to the choice of event generator.

Systematic uncertainties arising from experimental sources are discussed later in Chapter 10, along with the uncertainties related to those backgrounds which have no estimation procedure.

## 8.3 Top background

NFs derived from the top CR are used to correct the top background in the SR as described in Section 8.1. For both 8 TeV BDT and cut-based analyses a combined top CR ( $ee + \mu\mu$  and  $e\mu + \mu e$ ) is used, with the implicit assumption that any mismodelling present is independent of lepton flavour. Cut selections for the 8 TeV top CRs can be seen in Table 8.3. In the BDT analysis, the cuts defining the top CR are exactly the same as those for the SR, except with the  $b$ -veto replaced by requiring exactly  $N_{b-jet} = 1$ . Using a CR composed of events containing exactly one  $b$ -jet instead of an inclusive region in  $b$ -jet multiplicity, brings the quark flavour composition of the tagging jets closer to that of the SR. Distribution of the BDT input variables in the top CR selection can be seen in Figure 8.1, before any cut on the BDT response is made.

The BDT top NFs are calculated for each of the three bins in the final BDT score distribution entering the fit, as the phase space associated with lower BDT outputs is significantly different from that of the higher BDT score regions. However due to a lack of data events, the two highest BDT output bins (in other words, the two most signal-like bins) share a combined top NF. The ensuing top NFs are thus  $[1.58 \pm 0.15, 0.95 \pm 0.31, 0.95 \pm 0.31]$  in each of the three BDT signal-like bins. The larger top

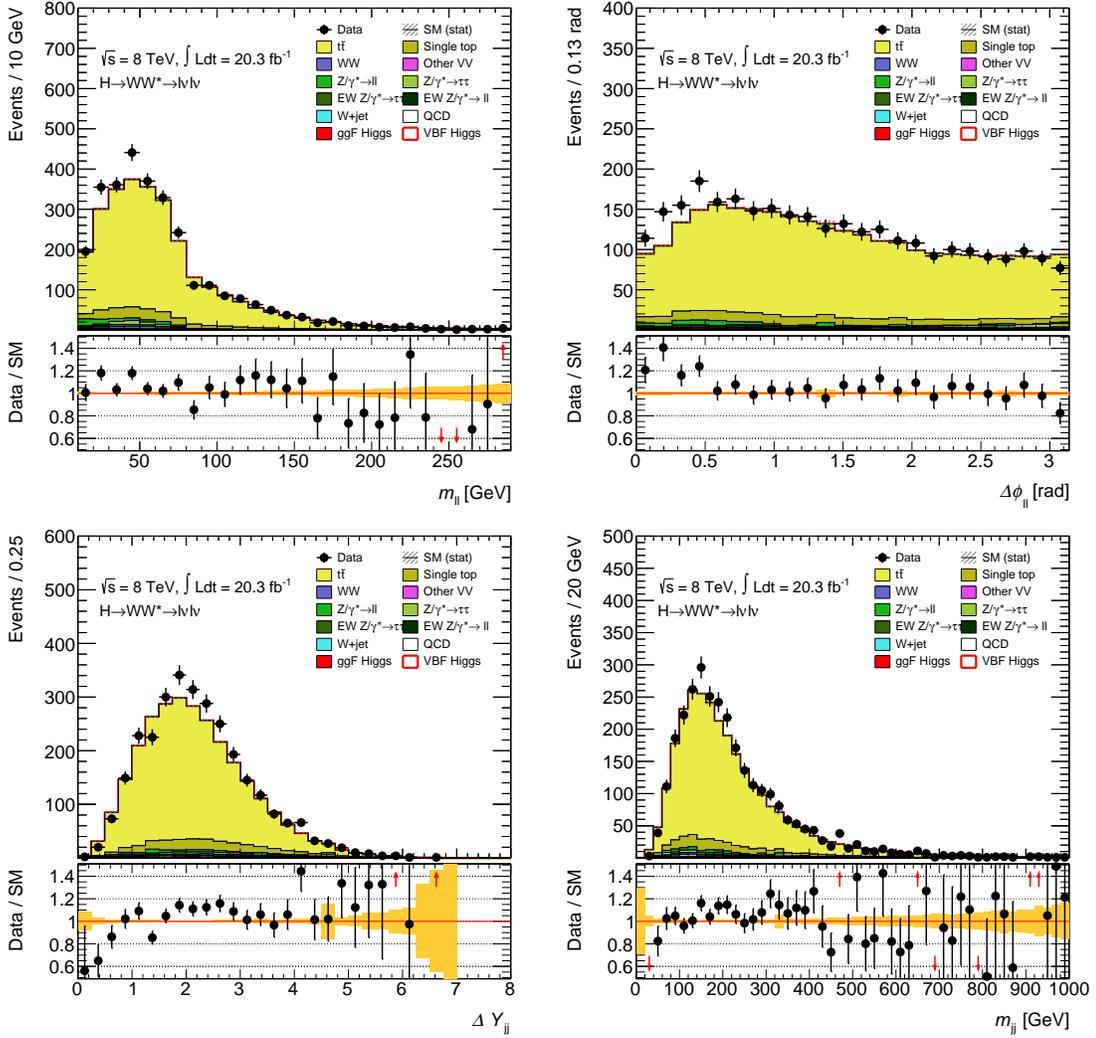


Figure 8.1:  $\Delta\phi_{\ell\ell}$ ,  $m_{\ell\ell}$ ,  $\Delta Y_{jj}$  and  $m_{jj}$  distributions within the BDT top CR. Shown for all lepton channels combined in the 8 TeV dataset, with no cut made on the BDT response here yet. The yellow uncertainty band shown represents only the statistical MC error in the data/MC ratio plots.

BDT usage	Cut-based usage
Common preselection	
$N_{jets} \geq 2$	
$\cancel{E}_T^{jet,Track} > 40 \text{ GeV},$ $\cancel{E}_T > 45 \text{ GeV } ee + \mu\mu \text{ only}$	$\cancel{E}_T^{jet,Track} > 50 \text{ GeV},$ $\cancel{E}_T > 55 \text{ GeV } ee + \mu\mu \text{ only}$
$N_{b-jet}=1, \text{ MV1 } 85\%$	
–	$p_T^{tot,jet,Track} < 15 \text{ GeV}$
$m_{\tau\tau} < 66 \text{ GeV}$	
–	$m_{jj} > 600 \text{ GeV}$
–	$\Delta Y_{jj} > 3.6$
CJV	
OLV	
BDT score	–

Table 8.3: Top CR definition for the 8 TeV BDT and cut-based analyses. The common preselection refers to that defined in Table 7.1. The “BDT score” cut refers to the separate calculation of the two top NFs, corresponding to the BDT response binning.

NF in the  $-0.48 < \text{BDT} < 0.3$  region has been studied and is due mostly to an excess at low  $\Delta\phi_{\ell\ell}$  in the top CR. Alternative top MC samples produced with ALPGEN and MC@NLO were used to check this region, with MC@NLO showing good modelling of  $\Delta\phi_{\ell\ell}$ . Given that the discrepancy from unity is covered if all theoretical (including a MC generator comparison) and detector systematics are included, it was concluded that this is not worrisome, especially as this concerns the least sensitive of the three BDT signal bins.

For the cut-based analysis, exactly  $N_{b-jet} = 1$  is also required in place of the  $b$ -veto. The cut-based top CR selection is also identical to that of the cut-based SR, except with the requirements on  $m_{\ell\ell}$  and  $\Delta\phi_{\ell\ell}$  removed -  $m_{\ell\ell}$  and  $\Delta\phi_{\ell\ell}$  are not included, as they are well modelled variables and thus are assumed not to be the cause of any mis-modelling. Their absence allows more events to be retained in the CR, thus reducing the statistical error on the NF derivation. Distributions of the key variables in the 8 TeV cut-based top CR can be seen in Figure 8.2. The possibility of using separate top NFs to accommodate the two cut-based  $e\mu + \mu e$  fit regions ( $600 \text{ GeV} < m_{jj} < 1 \text{ TeV}$ , and  $m_{jj} > 1 \text{ TeV}$ ) were investigated, but as both NFs derived were statistically consistent with each other as well as unity, one NF of  $1.04 \pm 0.19$  integrated over the full  $m_{jj}$  range was used instead [112]. The cut-based top NF matches well with that determined for the two most signal-like bins in the BDT analysis.

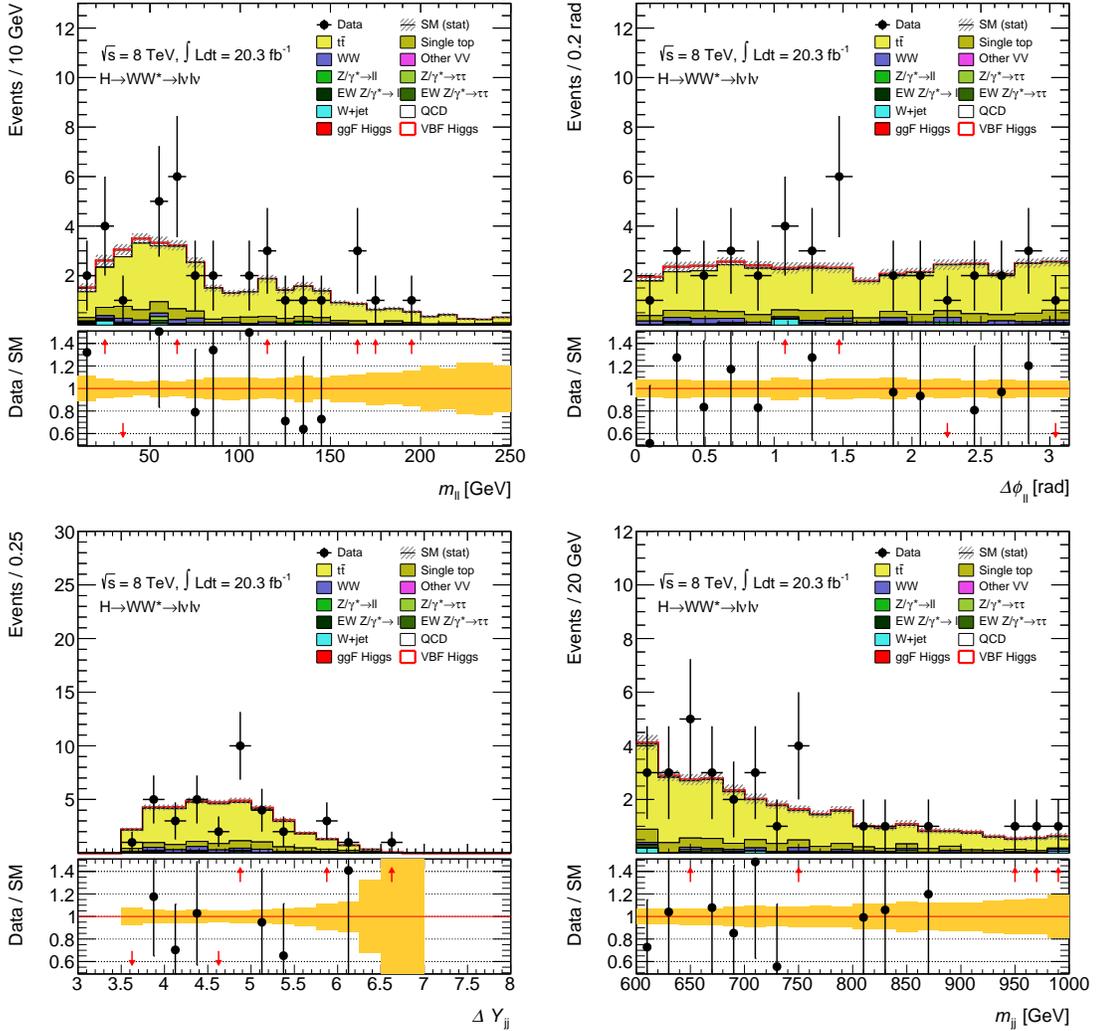


Figure 8.2: Distributions of  $\Delta\phi_{\ell\ell}$ ,  $m_{\ell\ell}$ ,  $\Delta Y_{jj}$  and  $m_{jj}$  within the cut-based top CR from which the top NF is directly determined. Shown for all lepton channels combined in the 8 TeV dataset. The yellow uncertainty band shown represents only the statistical MC error.

The top background extrapolation factor  $\alpha_{top}$  is defined by the ratio of simulated top MC events in the SR compared to those in the top CR:  $\alpha_{top} = N_{top}^{SR,MC} / N_{top}^{CR,MC}$ , as described in Section 8.1. As the top CR is completely dominated by the  $t\bar{t}$  background, the single-top processes are neglected in the calculation of theoretical uncertainties.  $\alpha_{top}$  is determined for each bin of the BDT output as is done for the NF calculation, in order to incorporate shape systematics where appropriate. Correspondingly, one overall  $\alpha_{top}$  is calculated for the cut-based analysis.

The total theoretical uncertainty is dominated by the matrix element component,

which is calculated by comparing the  $t\bar{t}$  prediction of the POWHEG, MC@NLO and ALPGEN generators, with all PS calculated using HERWIG. Comparing MC@NLO and ALPGEN produces the most discrepant values of  $\alpha_{top}$ , which is then taken as the systematic matrix element uncertainty in each bin, [10%, 12%, 21%]. Likewise for the cut-based cross-check, the difference in yields produced using MC@NLO and ALPGEN were the largest, resulting in an overall 22% modelling uncertainty, which agrees well with the 21% modelling uncertainty in the most signal-like BDT bin. The QCD scale makes up the second largest source of theoretical uncertainty and is found to be less than 10% both for the  $\alpha_{top}$  in each BDT signal-like bin and in the cut-based top CR.

PDF uncertainties are obtained by comparing the CT10 and NNPDF [113] PDF sets, and PS uncertainties by observing the difference between a POWHEG  $t\bar{t}$  MC sample showered with PYTHIA, and with HERWIG. For the BDT analysis, both PDF and PS uncertainties are within the statistical error and hence are not taken into account in the final fitting procedure. In the cut-based analysis, the PS uncertainty is also negligible for the same reason but has a PDF uncertainty which is comparable with the statistical error, of  $\mathcal{O}(10\%)$ . The difference here between the BDT and cut-based PDF uncertainties originate from the different treatment, described further in Ref. [112]. Combining all theoretical uncertainties in quadrature, total values of [14%, 16%, 23%] and 26% are achieved for the top background estimate in the 8 TeV BDT and cut-based analyses respectively.

## 8.4 $WW$ production

In the VBF analyses, no  $WW$  CR can be established due to large contamination from  $t\bar{t}$  production in  $WW$ -enriched regions - hence the background estimation relies completely on the SHERPA MC prediction. Instead, a validation region (VR) is defined with a reasonably pure sample of  $WW$  events produced with two associated jets, in order to check the modelling of the  $WW$  background.

The VR uses the lepton  $p_T$  cuts from the common preselection, along with the following:  $m_{\ell\ell} > 10$  GeV,  $N_{jets} \geq 2$  and  $N_{b-jet} = 0$ . A  $m_T^{jet,Track} > 100$  GeV cut is applied to further enhance the  $WW$  contribution. Lastly, the  $m_{T2}$  variable [114], defined using the lepton and jet momenta, and the  $\cancel{E}_T$  present, is used with a value  $m_{T2} > 160$  GeV to bring the purity of the  $WW$  VR to 60%. This variable has been used previously in other ATLAS analyses with good results, to select out  $t\bar{t}$  events decaying into two leptons [115]. The top background present in this VR is normalised in a separate CR defined by the same criteria except with the  $N_{b-jet}$  requirement inverted, using an NF as described in Section 8.1. The ratio between the observed and expected number of  $WW$  events in this region is  $1.13 \pm 0.19$ , which is consistent with unity given the statistical error. Figure 8.3 shows the  $m_{\ell\ell}$  and  $\Delta\phi_{\ell\ell}$  distributions in the VR, with a data/MC ratio consistent with unity mostly.

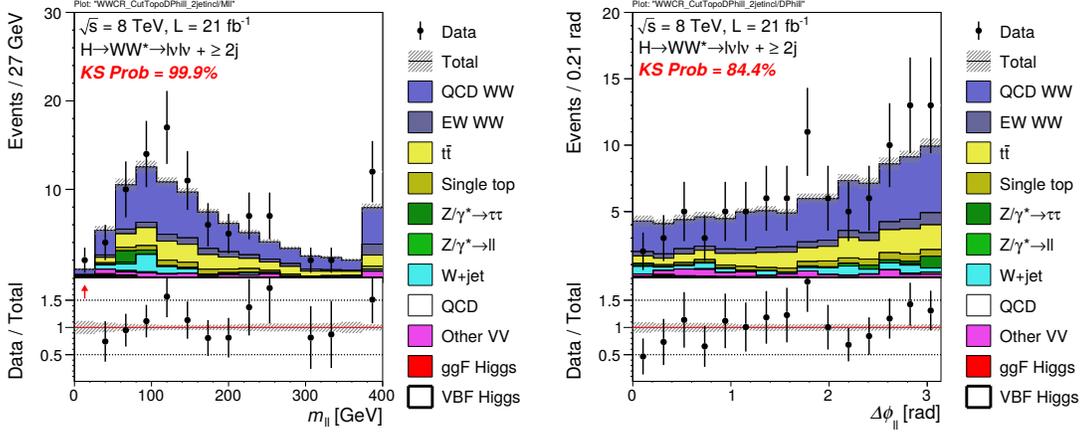


Figure 8.3: The  $m_{\ell\ell}$  and  $\Delta\phi_{\ell\ell}$  distributions in the  $WW$  VR, with only statistical uncertainties shown [93].

As described in Section 6.2.1, the  $WW$  background is made up of two MC samples, one containing processes with EW vertices only, and those containing a QCD vertex. Although the total cross section of the QCD  $WW$  processes are roughly a factor of twenty larger than those of the EW  $WW$ , both contributions are comparable in the most signal-like BDT bins (corresponding approximately to the VBF cut-based SR of large  $m_{jj}$  and  $\Delta Y_{jj}$ ), and thus have separate uncertainty calculations. Additional QCD and EW  $WW$  samples are generated using the multi-purpose MC generator MADGRAPH [116] in order to derive these various uncertainties, in the BDT SR using the same cuts as defined in Table 7.3. Just as for the top background, the BDT analysis calculates bin-by-bin theoretical uncertainties in order to take any shape effects into account. More extensive details on this topic can be found in Ref. [112]

QCD  $WW$  theoretical uncertainties are evaluated for generator modelling by comparing SHERPA and MADGRAPH predictions, for QCD scale by varying the renormalization and factorisation scales of additional MADGRAPH samples, and for PDF choice by comparing the yields obtained with CT10 and MSTW2008 [117] PDF sets. The uncertainty from the generator modelling is [14%, 7%, 12%] in each signal-like BDT output bin, and includes also PS uncertainties as the two samples have different PS schemes - SHERPA performs parton showering internally, whereas the MADGRAPH sample was showered using PYTHIA 6. Although the QCD scale variation induces only small BDT score shape differences relative to the statistical uncertainties, there is an overall yield difference of 27% in the BDT fit region for which a flat uncertainty across all three signal-like BDT bins is assigned. An additional uncertainty of 4% is assigned across all bins in the fit region due to the PDF choice. The EW  $WW$  sample is treated in the same way, leading to generator uncertainties of [16%, 12%, 10%], a flat uncertainty for QCD scale of 10%, and an extra 3% for the PDF comparison.

Another source of uncertainty on the  $WW$  production cross section arises from interference between the EW and QCD  $WW$  processes, as well as that between the VBF Higgs production and EW  $WW$ . To study such effects, combined MADGRAPH samples of QCD+EW  $WW$  and EW  $WW$ +VBF Higgs were generated. The interference uncertainties were then derived by comparing the cross sections of these combined samples to those of the 'normal' separated samples, leading to uncertainties of 2% and 1.2% for the EW-QCD  $WW$ , and the EW  $WW$ -VBF Higgs interferences respectively, the latter of which is applied only to the EW  $WW$  sample.

The  $WW$  prediction in the cut-based analysis is evaluated using the same procedure using the relevant cuts listed in Table 7.3. As for the treatment of the top background, the possibility of assigning different uncertainties was examined for the two cut-based  $e\mu + \mu e$  fit regions. No statistically significant discrepancies were observed and hence overall uncertainties were applied.

Solid agreement between the BDT and cut-based values are generally observed. For the QCD  $WW$  cut-based procedure, modelling uncertainties of 8% and symmetrised QCD scale uncertainties of 34% were derived. No PDF uncertainty was calculated for the cut-based QCD  $WW$  sample, as it was expected to be small given the corresponding BDT result, and would anyway be dominated by the QCD scale uncertainty. For the EW  $WW$  sample, the modelling and PDF uncertainties were each found to be 3% only. An example distribution from the EW  $WW$  modelling uncertainty determination can be seen in Figure 8.4. The QCD scale uncertainty for EW  $WW$  was determined in the  $ee$ ,  $\mu\mu$  and  $e\mu + \mu e$  channels separately, in order to check if any differences were observed, leading to values of 2-3% depending on the lepton channel. Interference uncertainties between the EW and QCD processes add 2-4% to each EW and QCD  $WW$  sample in addition, again depending on the lepton channel. The interference between the EW  $WW$  process and the VBF Higgs signal contributes a further 0.4% to the EW  $WW$  uncertainty.

In summary for the BDT analysis, the total theoretical uncertainties combined in quadrature are [31%, 28%, 30%] and [19%, 16%, 14%] for the QCD and EW  $WW$  samples, for each of the three signal-like bins. Corresponding values for the cut-based analysis are 37% and 6%, when averaged also over the different lepton flavour values where present.

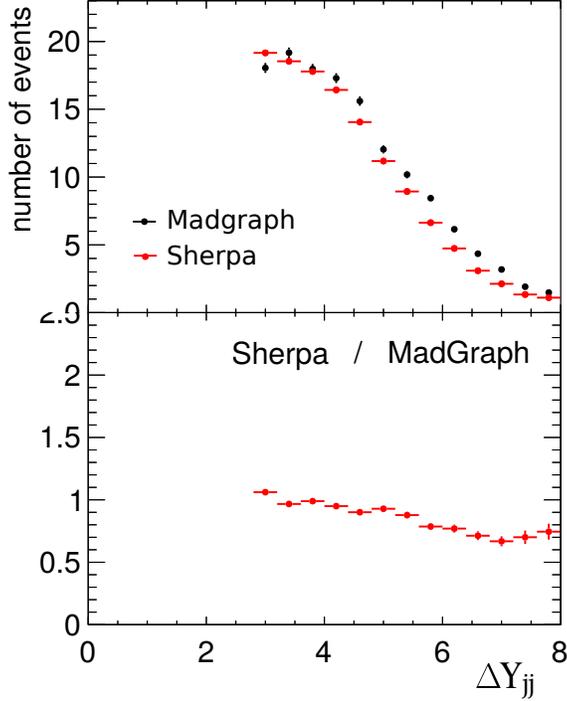


Figure 8.4: Distribution of  $\Delta Y_{jj}$  used in the calculation of the EW  $WW$  modelling uncertainty, with a comparison between SHERPA and MADGRAPH. Good agreement between the two generators can be seen. Only statistical uncertainties are shown here.

## 8.5 $Z/\gamma^* \rightarrow ee + \mu\mu$ background

It is known that  $Z/\gamma^* \rightarrow ee + \mu\mu$  processes show some slight mismodelling within the MC simulation for high missing transverse energies, which is part of the SR selection. In this case, a data-driven estimation is preferable. As  $Z/\gamma^* \rightarrow ee + \mu\mu$  is only dominant in the  $ee + \mu\mu$  channel, the estimation for this background is corrected for only in the SR of the  $ee + \mu\mu$  channel. The  $Z/\gamma^* \rightarrow ee + \mu\mu$  BDT response shape is taken from the data in a low- $\cancel{E}_T$   $Z$  background CR (referred to as “Region B”), corrected by subtracting the non- $Z/\gamma^* \rightarrow ee + \mu\mu$  background contributions from MC simulations. This shape is normalised using the  $\cancel{E}_T$  cut efficiency calculated using data within a 15 GeV mass window around the  $Z$  boson mass  $m_Z$  (Regions C and D). An overview of the various regions can be seen in Table 8.4, and the cuts prior to allocation of regions are given in Table 8.5. The  $Z/\gamma^* \rightarrow ee + \mu\mu$  estimate is estimated for each BDT score bin  $i$ :

$$N^{\text{SR},i} = N_{data}^{\text{B},i} \cdot \frac{N_{data}^{\text{C}}}{N_{data}^{\text{D}}} \cdot f_{\text{non-closure}} \quad (8.4)$$

where  $N_{data}$  is the number of data events in the region, with the non- $Z/\gamma^* \rightarrow ee + \mu\mu$  background subtracted, and  $f_{\text{non-closure}}$  is the non-closure factor which represents the

discrepancy in  $\cancel{E}_T$  cut efficiencies between the low and high  $m_{\ell\ell}$  regions (Regions A+B, and Regions C+D). Note that the terms  $N_{data}^C$  and  $N_{data}^D$  are independent of the bin  $i$  in question and are taken inclusively from the entire BDT score range to reduce statistical error, as the BDT score and  $\cancel{E}_T$  are shown to be uncorrelated. The non-closure is defined by:

$$f_{\text{non-closure}} = \frac{N_{MC}^A/N_{MC}^B}{N_{MC}^C/N_{MC}^D} \quad (8.5)$$

where  $N_{MC}$  is the number of  $Z/\gamma^* \rightarrow ee + \mu\mu$  MC events in each of the ABCD regions. The BDT NF in each bin  $i$  is thus given by:

$$NF_{ABCD,i}^{BDT} = N^{\text{SR},i}/N_{MC,i}^A \quad (8.6)$$

<b>Region A (SR)</b>  $\cancel{E}_T > 45 \text{ GeV}$ $m_{\ell\ell} < 75 \text{ GeV}$	<b>Region C</b>  $\cancel{E}_T > 45 \text{ GeV}$ $ m_{\ell\ell} - m_Z  < 15 \text{ GeV}$
<b>Region B</b>  $25 < \cancel{E}_T < 45 \text{ GeV}$ $m_{\ell\ell} < 75 \text{ GeV}$	<b>Region D</b>  $25 < \cancel{E}_T < 45 \text{ GeV}$ $ m_{\ell\ell} - m_Z  < 15 \text{ GeV}$

Table 8.4: Summary of the regions used for the  $Z/\gamma^* \rightarrow ee + \mu\mu$  estimation technique used in the  $ee + \mu\mu$  channel of the VBF BDT analysis.

Figure 8.5 shows Regions A, B, C and D, before further division into bins of BDT score. Due to a lack of data events in Region B for the two highest BDT score bins, there is a shared NF for these bins. The resulting NFs are  $[1.02 \pm 0.16, 0.90 \pm 0.28, 0.90 \pm 0.28]$  for each of the three BDT response regions.

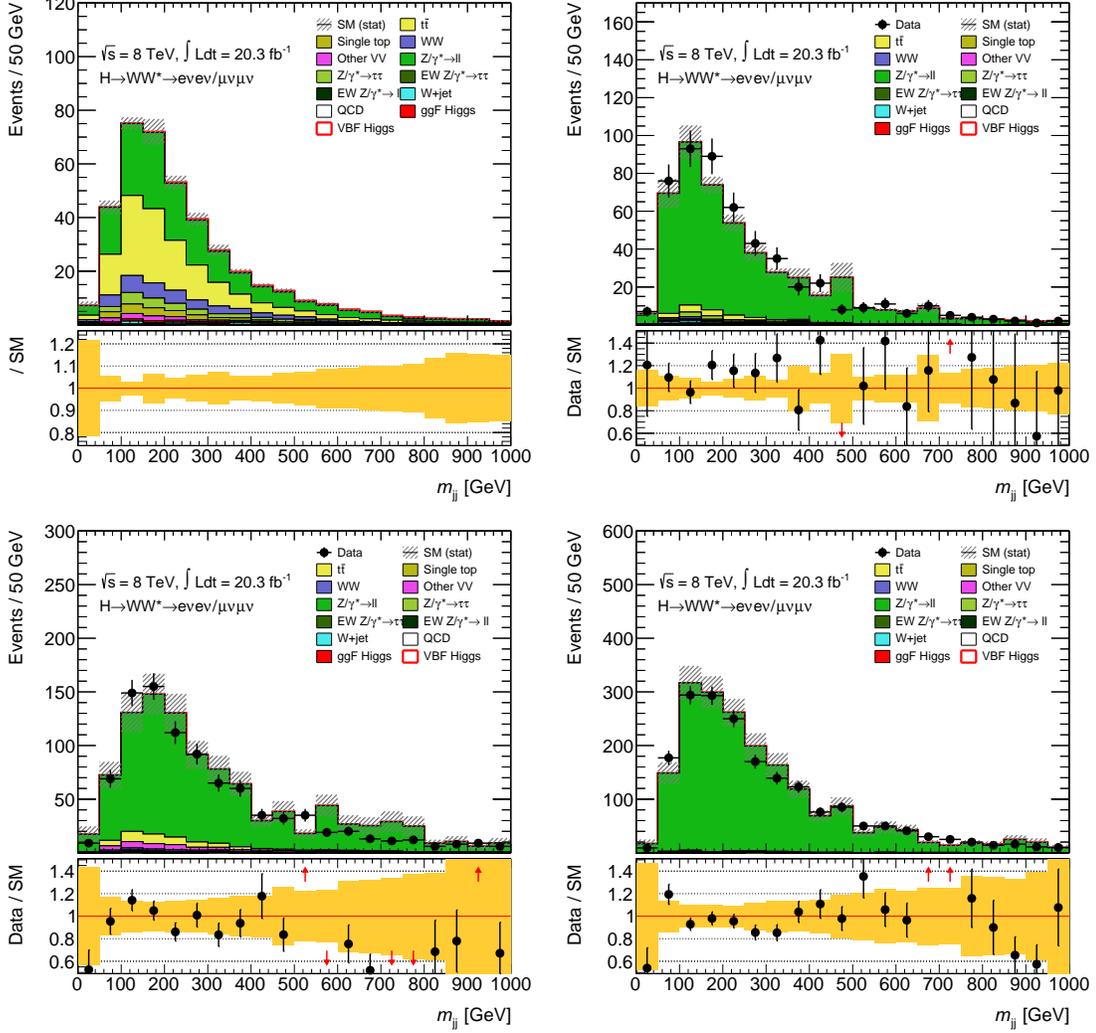


Figure 8.5:  $m_{jj}$  distributions for each of the Regions A, B, C and D before cutting on the BDT score to enable the corresponding NF calculation for each fit region. Region A is the SR and hence has no data points due to the blinding criteria. The statistical error only is represented by the yellow band in the data/MC ratio plots below the main distribution.

BDT usage	Cut-based usage
Common preselection without $Z$ -mass veto	
$N_{jets} \geq 2$	
$N_{b-jet}=1, MV1 \text{ 85\%}$	
$ee + \mu\mu$ only	–
$\cancel{E}_T^{jet,Track} > 40 \text{ GeV}$	$\cancel{E}_T^{jet,Track} > 50 \text{ GeV}$
$m_{\tau\tau} < 66 \text{ GeV}$	
–	$p_T^{tot,jet,Track} < 15 \text{ GeV}$
–	$m_{jj} > 600 \text{ GeV}$
CJV	–
OLV	–

Table 8.5: Selection used before dividing into the “ABCD” regions, for the 8 TeV BDT and cut-based analyses. The common preselection refers to that defined in Table 7.1.

Two assumptions are implicit for this estimation procedure:

1. BDT response and  $\cancel{E}_T$  are uncorrelated for the  $Z/\gamma^* \rightarrow ee + \mu\mu$  background
2.  $\cancel{E}_T$  cut efficiency is uncorrelated with  $m_{\ell\ell}$

Both these assumptions are based on empirical results from MC simulation and therefore there is an uncertainty in how accurately the MC simulation deals with such assumptions. The first assumption allows for the shape of the BDT output for  $Z/\gamma^* \rightarrow ee + \mu\mu$  to be extracted from Region B, and also for the same  $N_{data}^C/N_{data}^D$  factor to be used across all bins. This is illustrated in Figure 8.6, in which the BDT score shape is shown for the low  $\cancel{E}_T$  (Regions B+D) and high  $\cancel{E}_T$  (Regions A+C), normalised arbitrarily to unity. It can be seen that in all of the three BDT fit regions, that the shape is extremely similar and within the associated statistical uncertainty.

The negligible correlation between the BDT score and  $\cancel{E}_T$  for  $Z/\gamma^* \rightarrow ee + \mu\mu$  can be explained by the fact that there should be no “true” missing energy in  $Z/\gamma^* \rightarrow ee + \mu\mu$  events - this process produces no neutrinos. Any apparent  $\cancel{E}_T$  seen here results from mismeasurement or fluctuations in the calorimeter response. Inputs to the BDT algorithm which are related directly to the missing energy, such as  $m_T^{jet,Track}$  and  $p_T^{tot,jet,Track}$ , are defined using jet-corrected  $\cancel{E}_T^{jet,Track}$  which uses tracking rather than calorimeter information. Hence it is evident that the BDT output and  $\cancel{E}_T$  are largely uncorrelated. This dependence is further weakened in the BDT response by the presence of more strongly discriminating input variables such as  $m_{jj}$  and  $\Delta Y_{jj}$ . In order to assign a theoretical uncertainty to this assumption, the shape discrepancy in each of the three

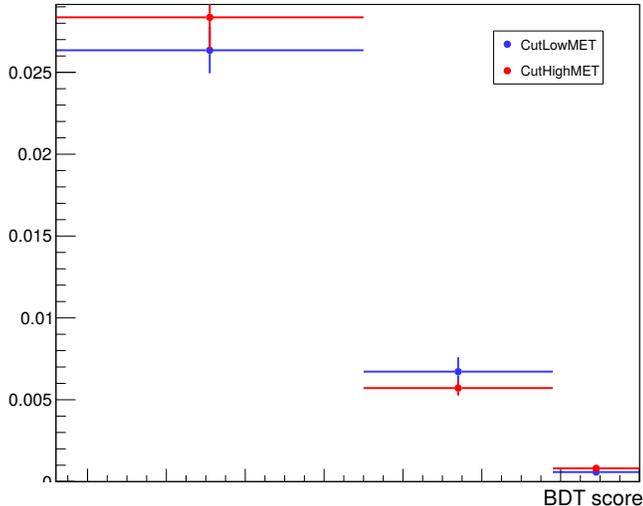


Figure 8.6: BDT response shape for a low  $\cancel{E}_T$  (Regions B+D) and high  $\cancel{E}_T$  (Regions A+C) region, to show the non-correlation of the BDT score and  $\cancel{E}_T$  for the  $Z/\gamma^* \rightarrow ee + \mu\mu$  background.

BDT score bins for  $Z/\gamma^* \rightarrow ee + \mu\mu$  events between the SR and Region B is calculated. Two samples with different showering were used - ALPGEN with PYTHIA, and ALPGEN with HERWIG- and the largest discrepancy in each bin taken as the uncertainty. As a result, the theoretical uncertainties due to any  $\cancel{E}_T$  dependence in the BDT response is [4%, 10%, 60%] in each of the three BDT signal-like bins.

At leading order, the assumption that  $m_{\ell\ell}$  and  $\cancel{E}_T$  are uncorrelated is valid as  $Z/\gamma^* \rightarrow ee + \mu\mu$  produces no true neutrinos and therefore no true missing energy. However if the hadronic recoil<sup>1</sup> is also taken into account, it can be seen that a larger  $m_{\ell\ell}$  results from a higher momentum transfer and thus increased initial-state radiation. As the resolution in the calorimeter is related roughly to  $\sqrt{E}$ , where  $E$  is the energy of the object, the additional hadronic activity induces a poorer resolution in the measurement of  $\cancel{E}_T$ . Due to this, there exists a small correlation between  $m_{\ell\ell}$  and  $\cancel{E}_T$  which results in the non-closure of this method.  $f_{\text{non-closure}}$  is measured to be  $0.83 \pm 0.22$  and is already incorporated into the background estimate. The deviation of  $f_{\text{non-closure}}$  from unity, 17%, is taken as a background estimation systematic across all of the BDT score bins. The  $\cancel{E}_T$  cut efficiency is measured also from  $Z/\gamma^* \rightarrow ee + \mu\mu$  MC simulation and has a value  $0.47 \pm 0.03$ , consistent with that measured in data of  $0.43 \pm 0.02$ . The statistical uncertainty on the data-derived value is propagated also to the fit for all the BDT output bins.

Further theoretical uncertainties are evaluated for the two regions,  $-0.48 < \text{BDT} < 0.3$  and  $0.3 < \text{BDT} < 1.0$ , in this background estimation procedure. In each case, the

<sup>1</sup>In first order  $Z/\gamma^* \rightarrow \ell\ell$  diagrams, this is essentially the remaining parts of the colliding protons and any initial state radiation.

relative expected background in each of the three BDT fit regions was compared to its nominal value. To take into account uncertainties arising from QCD scale, PS and PDF variations, additional MC samples were produced using SHERPA. The uncertainties were derived by taking the discrepancy in the yield of these extra samples to that of the base sample. Combining these three source of uncertainties leads to a 11% uncertainty in the highest BDT score bin.

The cut-based analysis uses the same approach, with a slightly different selection for the CR shown in Table 8.6, and the use of one overall NF instead of a bin-by-bin procedure. The motivation for the cut-based ABCD selection is driven by the need to correct for the cuts used and the cut values corresponding to the cut-based SR selection are used here. Each of the Regions A, B, C and D used in this background estimate can be seen in Figure 8.7. The yield is given by:

$$N^{\text{SR}} = N_{\text{data}}^{\text{B}} \cdot \frac{N_{\text{data}}^{\text{C}}}{N_{\text{data}}^{\text{D}}} \cdot f_{\text{non-closure}} \quad (8.7)$$

where  $N_{\text{data}}$  is the number of data events in each region, with the non- $Z/\gamma^* \rightarrow ee + \mu\mu$  background subtracted and  $f_{\text{non-closure}} = \frac{N_{MC}^{\text{A}}/N_{MC}^{\text{B}}}{N_{MC}^{\text{C}}/N_{MC}^{\text{D}}}$ . As for Eq. 8.6, the cut-based NF is given similarly by:

$$NF_{\text{ABCD}}^{\text{cut-based}} = N^{\text{SR}}/N_{MC}^{\text{A}} \quad (8.8)$$

<b>Region A (SR)</b>  $\cancel{E}_T > 55 \text{ GeV}$ $m_{\ell\ell} < 50 \text{ GeV}$	<b>Region C</b>  $\cancel{E}_T > 55 \text{ GeV}$ $ m_{\ell\ell} - m_Z  < 15 \text{ GeV}$
<b>Region B</b>  $20 < \cancel{E}_T < 55 \text{ GeV}$ $m_{\ell\ell} < 50 \text{ GeV}$	<b>Region D</b>  $20 < \cancel{E}_T < 55 \text{ GeV}$ $ m_{\ell\ell} - m_Z  < 15 \text{ GeV}$

Table 8.6: Summary of the regions used for the  $Z/\gamma^* \rightarrow ee + \mu\mu$  estimation technique used in the SF channel of the VBF cut-based analysis.

Additional cuts applied before the CR is split into the four regions can be seen in Table 8.5. The lower  $\cancel{E}_T$  bound of 20 GeV in the cut-based ABCD method is motivated by increased proximity to the SR and higher statistical power. The cut-based ABCD NF is calculated to be  $0.71 \pm 0.20$  and as for the BDT approach, already has  $f_{\text{non-closure}}$  incorporated into the estimate which takes into account the limitations of

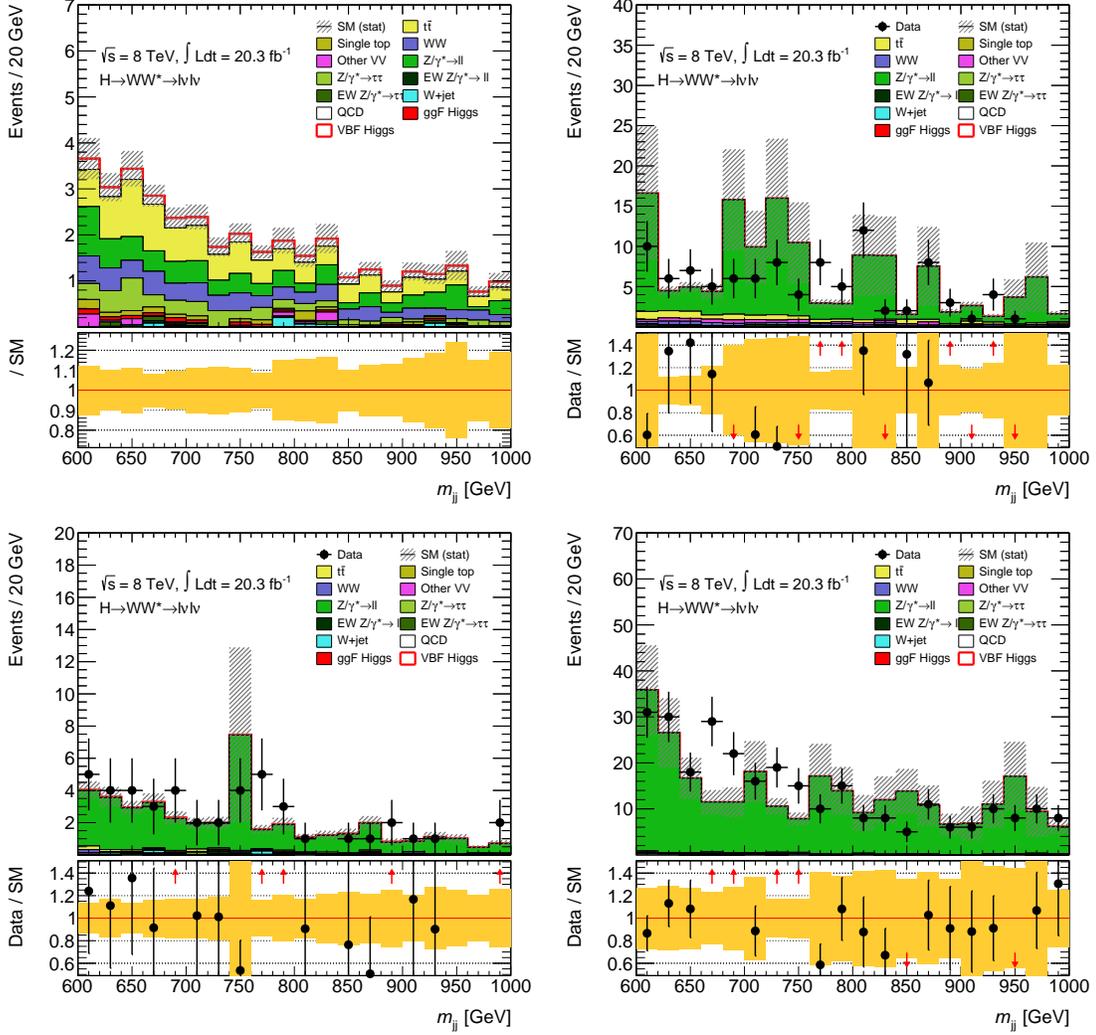


Figure 8.7:  $m_{jj}$  distributions for each of the Regions A, B, C and D from which  $N_{ABCD}^{cut-based}$  is calculated for the 8 TeV cut-based analysis. As Region A is the SR, it has no data points due to the blinding criteria. The statistical error is represented by the yellow band in the data/MC ratio plots below the main distribution.

the decorrelation of  $\cancel{E}_T$  and  $m_{\ell\ell}$ . An additional correction factor  $CF_{Z/DY}$  is required in order to correct for the effects induced by the topological VBF cuts not included in the ABCD selection. The correction is made separately as a CR which consists of all the VBF cuts would be extremely depleted of events, as these cuts are designed to be very restrictive for background processes. The CF is calculated from the region described in Table 8.7 and has a value  $CF_{Z/DY} = 1.34 \pm 0.46$ . Given that  $CF_{Z/DY}$  carries such a high statistical uncertainty, no additional theoretical uncertainties were assigned to it.

The final factor used in the cut-based analysis estimation of  $Z/\gamma^* \rightarrow ee + \mu\mu$  is therefore  $NF_{ZDY}^{cut-based} = NF_{ABCD}^{cut-based} \times CF_{Z/DY} = 0.95 \pm 0.50$  and is applied to the  $ee + \mu\mu$  flavour channel at the  $m_{\ell\ell}$  cut stage. This agrees well with those derived for the BDT analysis within statistical error. No calculation of separate NFs is required for the two SRs split in  $m_{jj}$ , as the split occurs only for the  $e\mu + \mu e$  channel which does not require an  $Z/\gamma^* \rightarrow ee + \mu\mu$  estimation. The deviation of  $f_{\text{non-closure}}$  from unity is taken as the only systematic uncertainty in the  $Z/\gamma^* \rightarrow ee + \mu\mu$  background estimation for the cut-based analysis, which being 29% is expected to dominate over any calculated theoretical uncertainties.

---



---

Common preselection without $Z$ -mass veto
$N_{jets} \geq 2$
$20 < \cancel{E}_T < 55 \text{ GeV}$
$\cancel{E}_T^{jet,Track} > 50 \text{ GeV}$
$b$ -veto
$p_T^{tot,jet,Track} < 15 \text{ GeV}$
$m_{\tau\tau} < 66 \text{ GeV}$
$\Delta Y_{jj} > 3.6$
CJV
OLV
$\Delta\phi_{\ell\ell} < 1.8$ ( $< 2.8$ for low- $p_T$ events)
$m_{\ell\ell} < 50 \text{ GeV}$

---



---

Table 8.7: Definition of the  $Z/\gamma^* \rightarrow ee + \mu\mu$  CR used to derive the  $CF_{Z/DY}$ , as part of the cut-based  $Z/\gamma^* \rightarrow ee + \mu\mu$  background estimation. The common preselection refers to that described in Table 7.1. Low- $p_T$  events are those for which  $10 < p_T^{\text{sublead lep}} < 15 \text{ GeV}$ .

### 8.5.1 Non-closure Concerns with the ABCD Method

In order to check that the ABCD approach used was sound, the  $f_{\text{non-closure}}$  factors were calculated in 8 TeV data and MC simulation within a low  $\cancel{E}_T$  region chosen for its high purity in  $Z/\gamma^* \rightarrow \ell\ell$  events. Three subdivisions in the low  $\cancel{E}_T$  region are made, illustrated in Figure 8.8, in order to calculate two non-closure factors. For instance, one non-closure factor is calculated using Region 1 as the low  $\cancel{E}_T$  region (B+D with respect to the naming used in Tables 8.4) and 8.6 and Region 2 as the higher  $\cancel{E}_T$  region (A+C). The other factor is determined using Region 2 and Region 3 in the same way. The selection used for this extra study is the same as that described in Table 8.5 for the cut-based analysis.

The  $f_{\text{non-closure}}$  factors obtained are consistent with both unity, and each other within statistical uncertainties, and can be seen in Figure 8.9. Further studies were also performed with various cuts to be sure that this consistency was maintained, thus showing that any non-closure present is taken into account already through the statistical error applied to the  $Z/\gamma^* \rightarrow ee + \mu\mu$  background estimate.

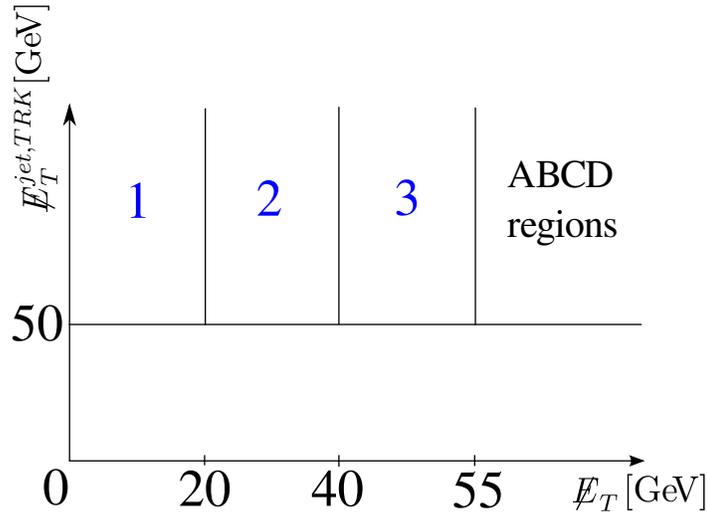


Figure 8.8: Schematic of the low  $\cancel{E}_T$  regions used to calculate  $f_{\text{non-closure}}$  factors for the non-closure comparisons between data and MC. The nominal ABCD regions used for the actual derivation of NFs is also shown for reference.

## 8.6 $Z/\gamma^* \rightarrow \tau\tau$ background

For this background, both BDT and cut-based analyses define a  $Z/\gamma^* \rightarrow \tau\tau$  CR in order to estimate the background contribution. The cut definitions of the  $Z/\gamma^* \rightarrow \tau\tau$  CR can be found in Table 8.8. The requirements on  $m_{\ell\ell}$  and  $m_{\tau\tau}$  increase the  $Z/\gamma^* \rightarrow \tau\tau$  purity of the region, while for the BDT analysis, the cut on BDT score corresponds

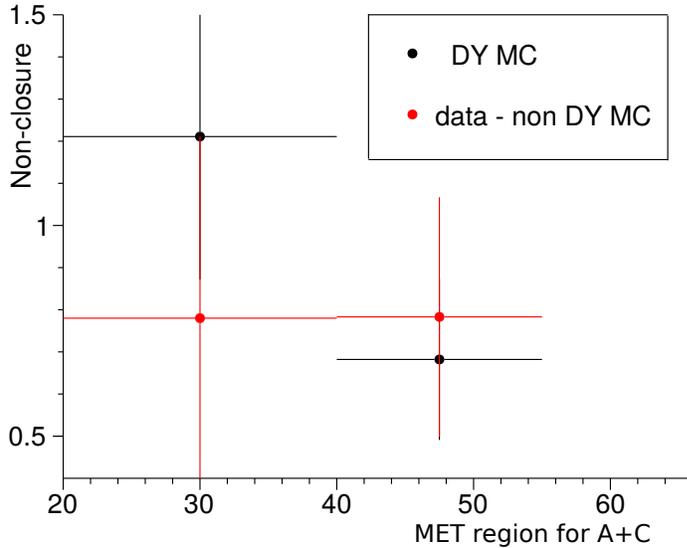


Figure 8.9: Non-closure tests performed to check the assumption that  $\cancel{E}_T$  and  $m_{\ell\ell}$  are uncorrelated. Both data and MC derived  $f_{\text{non-closure}}$  factors are consistent with each and unity, within the statistical errors shown.

to that of the fit region. Some example distributions of the BDT  $Z/\gamma^* \rightarrow \tau\tau$  CR can be seen in Figure 8.10 from which a single NF is derived  $NF_{\tau\tau}^{BDT} = 1.23 \pm 0.32$ , which is consistent with unity given the large statistical error. This is applied to the  $Z/\gamma^* \rightarrow \tau\tau$  background in the SR for both lepton flavour channels. Despite the large statistical uncertainty, the  $Z/\gamma^* \rightarrow \tau\tau$  background is an extremely small contribution in the BDT fit region, and the expected significance is shown to be unchanged even with the application of this NF. This statistical uncertainty is included in each bin of the BDT distribution, and given its size, no further theoretical uncertainties are assigned.

In the cut-based analysis, the  $Z/\gamma^* \rightarrow \tau\tau$  background is treated in a slightly different way. The overall correction is the product of a normalisation factor  $NF_{\tau\tau}$  estimated from the  $Z/\gamma^* \rightarrow \tau\tau$  CR to correct for any  $\tau$  lepton mismodelling, and a correction factor  $CF_{\tau\tau}$  calculated from the data/MC cut efficiencies, which corrects for mismodelling induced by the VBF topology cuts not included in the  $Z/\gamma^* \rightarrow \tau\tau$  CR definition. The top background contamination in the cut-based  $Z/\gamma^* \rightarrow \tau\tau$  CR is  $\sim 6\%$  and thus a separate top NF is defined to correct the top background estimate here. This is derived from the top CR (summarised in Table 8.3) but using only the cuts up to and including the  $p_T^{\text{tot},jet,Track}$  cut so as to match the  $Z/\gamma^* \rightarrow \tau\tau$  CR definition, and has a value  $1.05 \pm 0.01$ . The  $Z/\gamma^* \rightarrow \tau\tau$  CR is shown in Figure 8.11 and the resulting normalisation factor from it is  $NF_{\tau\tau} = 0.97 \pm 0.04$ .

As the VBF topology cuts are based on jet observables, they have no correlation to lepton flavour and thus it can be assumed that they are lepton flavour insensitive. This

BDT	Cut-based
Common preselection	
$N_{jets} \geq 2$	
$b$ -veto	
$ m_{\tau\tau} - m_Z  < 25 \text{ GeV}$	
–	$p_T^{tot,jet,Track} < 15 \text{ GeV}$
$m_{\ell\ell} < 80 \text{ GeV}$ for $e\mu + \mu e$	$m_{\ell\ell} < 80 \text{ GeV}$
$m_{\ell\ell} < 75 \text{ GeV}$ for $ee + \mu\mu$	
BDT score $> -0.48$	–

Table 8.8: Definition of  $Z/\gamma^* \rightarrow \tau\tau$  CR for BDT and cut-based analyses. The common preselection refers to that described in Table 7.1.

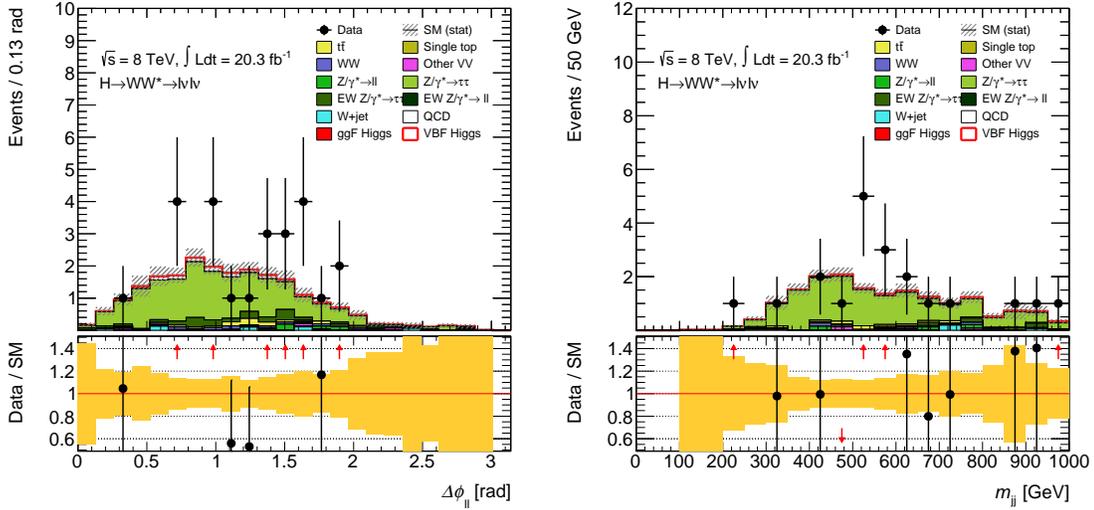


Figure 8.10:  $\Delta\phi_{\ell\ell}$  (left) and  $m_{jj}$  (right) distributions, for the  $Z/\gamma^* \rightarrow \tau\tau$  CR events in the 8 TeV BDT analysis. The yellow band in the data/MC ratio plots represent only the statistical error. This region is poorly populated with events, leading to a large statistical uncertainty on the  $Z/\gamma^* \rightarrow \tau\tau$  NF.

allows the use of a  $ee + \mu\mu$  channel only  $Z/\gamma^* \rightarrow ee + \mu\mu$  CR as defined in Table 8.9 and shown in Figure 8.12, for the  $CF_{\tau\tau}$  derivation as it benefits from higher event yields. To justify this choice, a comparison showing the shapes of  $Z/\gamma^* \rightarrow \tau\tau$  events in the  $e\mu + \mu e$  SR, and  $Z/\gamma^* \rightarrow ee + \mu\mu$  events in the  $ee + \mu\mu$  channel  $Z/\gamma^* \rightarrow ee + \mu\mu$  CR can be seen in Figure 8.13, with reasonable agreement with the statistical uncertainties. Due to the difference in fit regions in the cut-based analysis between the  $ee + \mu\mu$  and the  $e\mu + \mu e$  channels, a different  $CF_{\tau\tau}$  is calculated for each SR. Referring to Table 8.9, the  $ee + \mu\mu$  SR uses a  $CF_{\tau\tau} = 1.26 \pm 0.05$  estimated at the  $\Delta\phi_{\ell\ell}$  cut-stage. This region is then split into a low and high  $m_{jj}$  region with the division at 1 TeV, which are each used to calculate a  $CF_{\tau\tau}$  for the  $e\mu + \mu e$  channel. These are  $CF_{\tau\tau} = 1.24 \pm 0.06$  and  $1.34 \pm 0.10$  for the lower and higher  $m_{jj}$  fit regions in the  $e\mu + \mu e$  channel respectively. The overall correction factor for each region is obtained by multiplying the  $NF_{\tau\tau}$  by the corresponding  $CF_{\tau\tau}$ . The final correction factors all agree with that used by the BDT analysis, within statistical error. As with the 8 TeV BDT analysis, no further sources of uncertainty are considered for the cut-based  $Z/\gamma^* \rightarrow \tau\tau$  estimation as the statistical component outweighs any theoretical component. This is justified by an estimate of the systematic errors associated with the  $Z/\gamma^* \rightarrow \tau\tau$  background estimation, performed as described in Ref.[96].

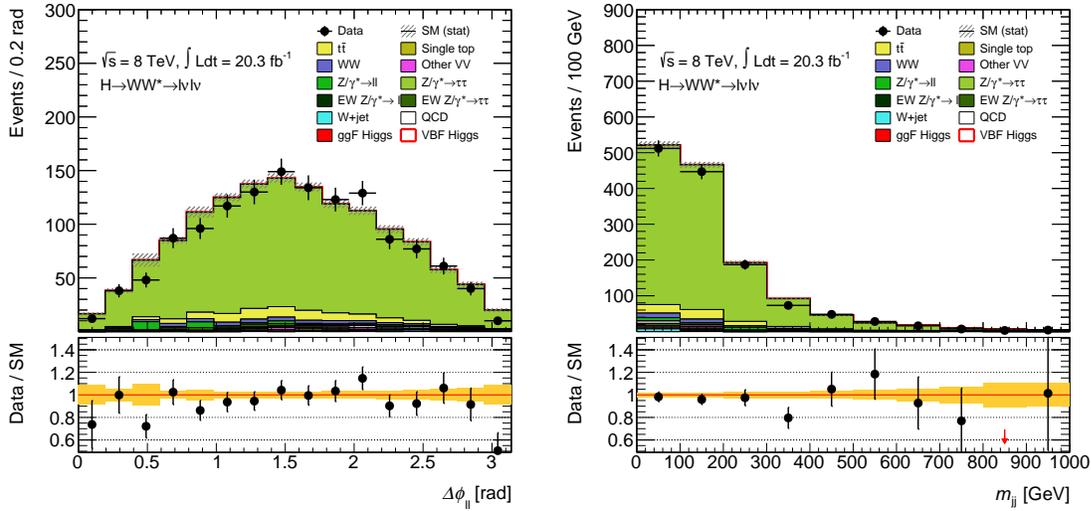


Figure 8.11: Distributions of  $\Delta\phi_{\ell\ell}$  (left) and  $m_{jj}$  (right) from the  $Z/\gamma^* \rightarrow \tau\tau$  CR in the 8 TeV cut-based analysis, from which  $NF_{\tau\tau}$  is calculated. The yellow band in the data/MC ratio plots represent only the statistical error.

## 8.7 $W$ +jets and QCD Multijet Backgrounds

Both  $W$ +jets and QCD multijet processes are especially minor backgrounds for the  $H \rightarrow WW^{(*)} \rightarrow l\nu l\nu$  signal process, in the VBF SR (for both BDT and cut-based

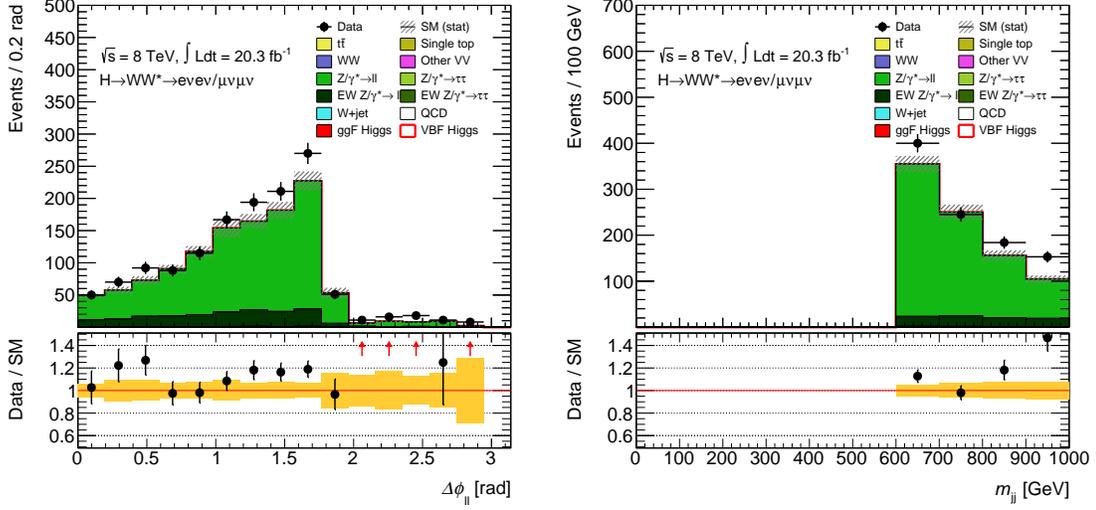


Figure 8.12: Distributions of  $\Delta\phi_{\ell\ell}$  (left) and  $m_{jj}$  (right) from the  $ee + \mu\mu$  channel  $Z/\gamma^* \rightarrow ee + \mu\mu$  CR in the 8 TeV cut-based analysis, from which  $CF_{\tau\tau}$  is calculated. These are shown at the  $\Delta\phi_{\ell\ell}$  cut-stage which is that at which the  $ee + \mu\mu$  channel  $CF_{\tau\tau}$  is derived. The yellow band in the data/MC ratio plots represent only the statistical error.

---



---

$ee + \mu\mu$  events only

Common preselection without  $Z$ -mass veto

$$N_{jets} \geq 2$$

$b$ -veto

$$p_T^{tot,jet,Track} < 15 \text{ GeV}$$

$$m_{jj} > 600 \text{ GeV}$$

$$\Delta Y_{jj} > 3.6$$

CJV

OLV

$$\Delta\phi_{\ell\ell} < 1.8 \text{ (} < 2.8 \text{ for low-}p_T \text{ events)}$$

---



---


$$600 \text{ GeV} < m_{jj} < 1 \text{ TeV, and } m_{jj} > 1 \text{ TeV}$$


---



---

Table 8.9: Definition of the  $ee + \mu\mu$  channel  $Z/\gamma^* \rightarrow ee + \mu\mu$  CR used to derive the  $CF_{\tau\tau}$ , as part of the cut-based  $Z/\gamma^* \rightarrow \tau\tau$  background estimation. The common preselection refers to that described in Table 7.1. Low- $p_T$  events are those for which  $10 < p_T^{\text{sublead lep}} < 15 \text{ GeV}$ .

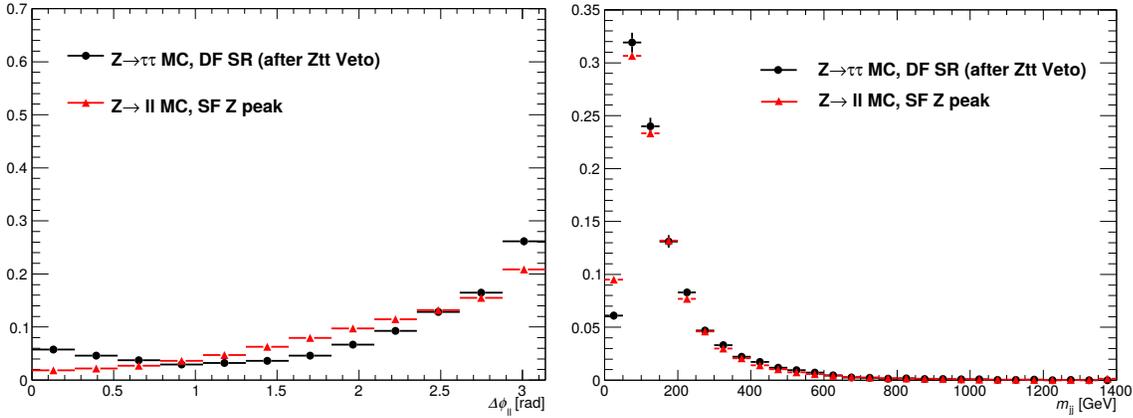


Figure 8.13: Comparison of the  $\Delta\phi_{\ell\ell}$  (left) and  $m_{jj}$  (right) distributions, for the  $Z/\gamma^* \rightarrow \tau\tau$  events in the  $e\mu + \mu e$  SR (black) and  $Z/\gamma^* \rightarrow ee + \mu\mu$  (red). Both distributions are normalised to unity so that the good shape agreement can be better seen [96].

analyses). As these backgrounds arise from the misidentification of jets or non-prompt leptons (*fake leptons*) as true leptons, a background estimation from data is better suited here. The MC prediction is not well modelled due to the simulated detector response, and so a data-driven technique is used to better account for real detector effects. Separate estimations are carried out for  $W$ +jets and QCD multijet processes, using the same method, and also for the BDT and cut-based analyses due to the different SR selection. The estimation procedure is described here for a general SR.

CRs are defined in which at least one of the leptons fulfils only a looser lepton definition meaning that some of the requirements in the more stringent lepton quality criteria as described in Section 4.1.1 and 4.1.2 are relaxed. Such leptons are referred to as “anti-identified” leptons, as opposed to the “identified” leptons which are selected for analysis. The selection for the anti-identified leptons are given in Table 8.10. Thus, regions which are enriched in fake leptons are obtained, which are orthogonal to the SR. To extrapolate from these regions to the SR, a *fake factor*  $f_\ell$  is used, defined as the ratio of identified to anti-identified leptons:

$$f_\ell = \frac{N_{\text{identified}}}{N_{\text{anti-identified}}}, \text{ where } \ell = e \text{ or } \mu \quad (8.9)$$

$f_\ell$  is measured as a function of both lepton  $p_T$  and pseudorapidity, for events with oppositely charged lepton pairs  $e^+e^-$  or  $\mu^+\mu^-$  present, as the SR selection requires also two oppositely signed leptons. A data region enriched in  $Z$ +jets processes (those in which a  $Z$  boson is produced in association with one or more jets) and dijet processes is used to calculate  $f_\ell$ .

For the  $W$ +jets predictions, two  $W$ +jets CRs are defined for the  $ee + \mu\mu$  and  $e\mu + \mu e$  channel separately, with the same selection as that of the SR. The difference is

however that there is one identified and one anti-identified lepton required (instead of two identified analysis leptons). The number of events in this  $W$ +jets CR  $N_{id+anti-id}$ , is used to estimate the amount  $W$ +jets background in the SR:

$$N_{W+jets} = f_\ell \times N_{id+anti-id} \quad (8.10)$$

Similarly for the QCD process, a CR is defined with the same selection as that of the SR (and again for both flavour channels separately) but using two anti-identified leptons. The number of events in this QCD CR,  $N_{anti-id+anti-id}$ , estimates the number of QCD events in the SR:

$$N_{QCD} = f_\ell^2 \times N_{anti-id+anti-id} \quad (8.11)$$

where the use of the  $f_\ell^2$  factor results from having two fake leptons present in such processes. In both cases, contamination from non- $W$ +jets, or non-QCD multijet processes is subtracted using MC simulation. QCD multijet processes are present also in the  $W$ +jets CR, as either of the jets in a dijet event could be misidentified as a lepton. However, as this contamination makes up only a few percent of the total  $W$ +jets background, it is taken as negligible [93]. Figure 8.14 illustrates the fake factors measured in data for electrons and muons.

<b>Electron</b>	<b>Muon</b>
Same $p_T$ and $\eta$ range as an identified electron	$p_T > 15 \text{ GeV}$ , $ \eta  < 2.5$
Same impact parameter requirements	$d_0$ requirements removed
$E_{T,cone30}/E_T < 0.30$	$\Delta z_0 \sin \theta < 1.0 \text{ mm}$
$p_{T,cone40}/E_T < 0.16$	$E_{T,cone30}/p_T < 0.30$
No CBL	Track isolation cuts removed
Fails Medium++	
Not an identified electron	Not an identified muon

Table 8.10: Definition of the anti-identified electron and muon.  $E_T$  or  $p_T$ -dependent criteria are given as a range. ‘‘CBL’’ refers to the conversion flag and b-layer hit requirements. The identified electron and muon definitions can be found in Table 4.1 and 4.2.

The total theoretical and experimental (due to the data-driven estimate) uncertainties associated with the extrapolation factor  $f_\ell$ , vary as a function of the anti-identified lepton  $p_T$ : between 29% and 61% for anti-identified electrons, and 25% to 46% for

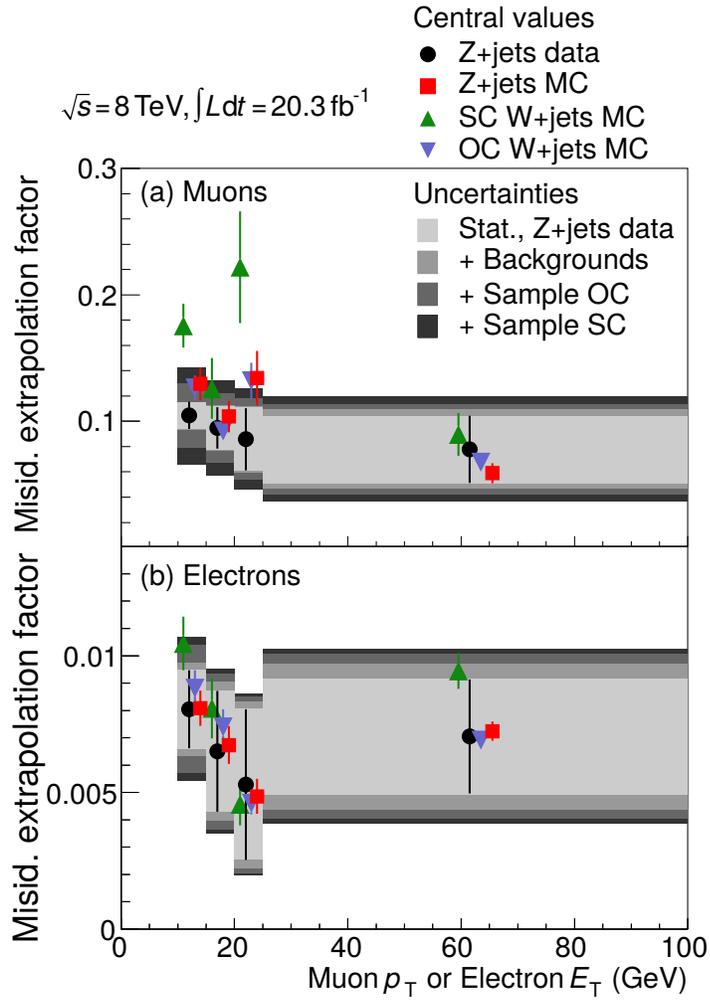


Figure 8.14: Fake factors for muons (top) and electrons (bottom) as a function of the anti-identified lepton  $p_T$  [96].

anti-identified muons. Sources of uncertainty accounted for here are the effect of pile-up on  $f_\ell$  due to changing run conditions, the impact of real lepton contamination from  $W$  or  $Z$  bosons in the CR, and the sample dependence. The latter term encompasses the differences in jet kinematics and quark flavour fractions between the sample used to determine  $f_\ell$ , and the CR to which  $f_\ell$  is applied to in order to obtain a  $W$ +jets or QCD multijet estimate. This is evaluated using MC predictions of dijets,  $Z$ +jets and  $W$ +jets processes, and completely dominates the total uncertainty.

## 8.8 Modifications for the 7 TeV analyses

The 7 TeV BDT and cut-based analyses follow the same procedures used to estimate the backgrounds here. Only minor changes are required, such as the replacing of cut values for the 7 TeV cut-based analysis in the definitions of CRs where they differ from that of the 8 TeV analysis, listed previously in Table 7.5. As the 7 TeV dataset suffers from lower event yields, the final fit definitions have been altered as described in Section 7.3.4. This affects only the calculation of NFs, in that only one overall NF is calculated each for top,  $Z/\gamma^* \rightarrow ee + \mu\mu$  and  $Z/\gamma^* \rightarrow \tau\tau$  backgrounds. The NFs calculated for the 7 TeV BDT and cut-based analyses are given in Tables 8.11 and 8.12. Theoretical uncertainties on any extrapolation factors are adopted directly by the 7 TeV analyses, while experimental sources are unique and are accounted for separately - this will be described in Chapter 10. In all cases where the binning differs, the convention followed is to take the larger of the uncertainties. For instance in the 7 TeV BDT analysis, the  $ee + \mu\mu$  channel contains a single merged bin instead of three separate regions and thus, the largest of the uncertainties calculated for each of the three bins is used.

	<b>Applied to (for Background)</b>	<b>NF <math>\pm</math> stat.</b>
$NF_{ABCD,i}^{BDT}$	$ee + \mu\mu$ SR ( $Z/\gamma^* \rightarrow ee + \mu\mu$ )	$1.460 \pm 0.814$
Top NF	SR (Top)	$0.816 \pm 0.294$
$Z/\gamma^* \rightarrow \tau\tau$ NF	SR ( $Z/\gamma^* \rightarrow \tau\tau$ )	$1.520 \pm 0.911$

Table 8.11: Normalisation factors used in the 7 TeV VBF BDT analysis, including the regions to which they are applied. Where no lepton flavour channel is specified, this refers to both  $ee + \mu\mu$  and  $e\mu + \mu e$  channels combined.

	Applied to (for Background)	NF $\pm$ stat.
$NF_{ABCD}^{cut-based}$	–	$1.12 \pm 0.23$
$CF_{Z/DY}$	–	$1.15 \pm 0.18$
$NF_{ZDY}^{cut-based}$	$ee + \mu\mu$ SR ( $Z/\gamma^* \rightarrow ee + \mu\mu$ )	$1.29 \pm 0.33$
Top NF	SR (Top)	$1.33 \pm 0.44$
$NF_{\tau\tau}$	–	$1.04 \pm 0.12$
$CF_{\tau\tau}$	–	$1.44 \pm 0.08$
$Z/\gamma^* \rightarrow \tau\tau$ NF	SR ( $Z/\gamma^* \rightarrow \tau\tau$ )	$1.49 \pm 0.19$

Table 8.12: Normalisation factors used in the 7 TeV VBF cut-based analysis, including the regions to which they are applied. Where no lepton flavour channel is specified, this refers to both  $ee + \mu\mu$  and  $e\mu + \mu e$  channels combined.

# Chapter 9

## Statistical Treatment

In order to evaluate the results of the VBF analyses a profile likelihood [118] procedure - which is defined to model, or fit the yields of all the processes considered - is used to fit the available data. Instead of using the actual yields in the fit, an equivalent parametrisation constructed from *strength parameters* is used. The signal strength parameter  $\mu$  is defined as the ratio of the measured signal yield to the expected (SM) yield - by definition,  $\mu = 1$  would mean a signal yield identical to the theoretical yield expected from the SM, and  $\mu = 0$  would correspond to no observed signal in the data. The goal is to extract an observed signal strength  $\mu_{obs}$  to characterise the results of the analyses, and to allow comparison between them. A limit can then be placed on  $\mu_{obs}$  using a modified frequentist method known as  $CL_s$  [119].

The BDT response distribution and the  $m_T$  distribution is used in the fit for the BDT and the 8 TeV cut-based analyses respectively. The 7 TeV cut-based analysis uses instead a simpler approach due to a lack of data events, which involves simply counting the number of events in the SR rather than fitting a distribution. This Chapter describes the fit regions and the entire statistical procedure.

### 9.1 Fit Regions

A fit is performed separately for the 7 TeV and 8 TeV analyses over data samples defined by the SR listed in Table 9.1, and profiled CRs which are given in Table 9.2 (non-profiled CRs are also included here for completeness). In a statistics context, *profiling* refers to the expression of a parameter in terms of other parameters of interest, thereby reducing the number of independent parameters. The profiled CRs calculate the normalisation of the background that it represents, whereas non-profiled CRs do not have explicit terms in the fit. Bin boundary optimisations and details in the fit region were described previously in Sections 7.3.2 and 7.3.3.

For the baseline BDT analyses, the BDT response is used to fit the data samples. As the signal purity increases with BDT score, the bin widths decrease in order to maximise the sensitivity. As a reminder, only the regions  $-0.48 < \text{BDT} < 0.3$ ,  $0.3 < \text{BDT} < 0.78$

and  $0.78 < \text{BDT} < 1.0$  contribute to the fit, with the bins represented by their lower boundaries as  $[-0.48, 0.3, 0.78]$ . The cross-check cut-based analyses use a fit on the  $m_T$  distribution with the following three bins, denoted by their lower boundaries  $[0, 80, 130]$  GeV. The 8 TeV cut-based analysis is further divided into two  $m_{jj}$  regions with a split at 1 TeV, represented again by the respective lower bounds as  $[600, 1000]$  GeV. Due to a lack of statistics in the 7 TeV dataset, the bins are merged such that reasonable event yields remain in all regions. The 7 TeV BDT analysis uses a two-bin fit region for the  $e\mu + \mu e$  channel, and a single-bin region for the  $ee + \mu\mu$  channel. The 7 TeV cut-based analysis uses a single  $m_T$  bin for both flavour channels, with a simple significance calculation performed instead of a fit.

Dataset		Channel	Fit variable	Bin boundaries	
8 TeV	BDT	$e\mu + \mu e$	BDT score	$[-0.48, 0.3, 0.78]$	
	BDT	$ee + \mu\mu$	BDT score	$[-0.48, 0.3, 0.78]$	
	CB	$e\mu + \mu e$	$m_T$	$[0, 80, 130]$ GeV	$m_{jj} < 1$ TeV
	CB	$ee + \mu\mu$	$m_T$	$[0, 80, 130]$ GeV	$m_{jj} > 1$ TeV
7 TeV	BDT	$e\mu + \mu e$	BDT score	$[-0.48, 0.3]$	
	BDT	$ee + \mu\mu$	BDT score	$[-0.48]$	
	CB	$e\mu + \mu e$	$m_T$	$[0, 130]$ GeV	
	CB	$ee + \mu\mu$	$m_T$	$[0, 130]$ GeV	

Table 9.1: Fit region definitions used in the likelihood, for each signal region category. “CB” here refers to the cut-based cross-check analyses

## 9.2 Profile Likelihood Procedure

A *likelihood function* is a function of the parameters of a particular statistical model. In general terms, the likelihood of a set of parameter values with given outcomes is equal to the probability of those observed outcomes given those parameter values. A *profile likelihood function* is a likelihood function in which certain parameters are represented as functions of other parameters so as to reduce the total number of independent parameters present in the expression. The likelihood function is represented as  $\mathcal{L}(\mu, \boldsymbol{\theta} | \mathbf{N})$ , which is a function of the signal strength, and a set of nuisance parameters  $\boldsymbol{\theta} = \{\theta_a, \theta_b, \dots\}$  given a set of event yields  $\mathbf{N} = \{N_a, N_b, \dots\}$ .  $\mathcal{L}$  is maximised with respect to its arguments and evaluated to obtain  $\mu_{obs}$ .

CR	Profiled?	Flavour
Top	•	$ee + \mu\mu$ & $e\mu + \mu e$
$Z/\gamma^* \rightarrow ee + \mu\mu$	×	$ee + \mu\mu$ only for BDT $ee + \mu\mu$ & $e\mu + \mu e$ for CB
$Z/\gamma^* \rightarrow \tau\tau$	×	$ee + \mu\mu$ & $e\mu + \mu e$
Misid.	×	Resp.

Table 9.2: Summary of all background CRs, where “misid.” refers to  $W$ +jets and QCD multijet processes, and “CB” to the cut-based analysis. The top CR is the only profiled CR (•) with the non-profiled regions marked by ×. The sample for the top and also  $Z/\gamma^* \rightarrow \tau\tau$  CRs are used for both lepton-flavour SRs; the  $Z/\gamma^* \rightarrow ee + \mu\mu$  sample treatment varies depending on the BDT/CB analysis; and the misid. use samples corresponding to the lepton flavour of the SR (Resp.).

$\mathcal{L}$  is given by the product of four probability distribution functions,  $\mathcal{L} = f_1 \cdot f_2 \cdot g_3 \cdot f_4$ , where:

- $f_1$ , a Poisson function for the statistics of a signal region  $i$  and bin  $b$  of the fit distribution used, with an observed yield  $N_{ib}$
- $f_2$ , another Poisson function for the statistics of profiled control regions  $l$ , with an observed yield  $N_l$ , for a given background process  $k$
- $g_3$ , a Gaussian function to constrain the systematic uncertainties which affect the expected signal and background yields
- $f_4$ , a Poisson function to take into account the finite statistics of a sample used for the MC prediction

Statistical uncertainties enter explicitly into  $f_1$ ,  $f_2$  and  $f_4$  only. The terms  $f_1$  and  $f_2$  also incorporate the random error associated with the value predicted by the MC simulation, for instance, for a background yield estimate  $B$ , which has a random error  $\sqrt{B}$  associated. The exact definition of  $\mathcal{L}$  is given in Equation 9.1. Each of the four terms will now be described further, with a more detailed treatment given in [105, 96].

$$\begin{aligned}
f_1 &= \prod_{i,b} f(N_{ib}|\mu \cdot S_{ib} \cdot \prod_r \nu_{br}(\theta_r) + \sum_k \beta_k \cdot B_{kib} \cdot \prod_s \nu_{bs}(\theta_s)) \\
f_2 &= \prod_l f(N_l | \sum_k \beta_k \cdot B_{kl}) \\
g_3 &= \prod_t g(\vartheta_t | \theta_t) \\
f_4 &= \prod_k f(\xi_k | \zeta_k \cdot \theta_k)
\end{aligned} \tag{9.1}$$

Most simply, a Poisson function  $f$  for the probability of observing  $N$  events given  $\lambda$  expected events is  $f(N|\lambda) = e^{-\lambda}\lambda^N/N!$ . The first Poisson function  $f_1$  has an expected value which consists of the signal event yield  $S$  multiplied by the signal strength  $\mu$  and the sum of the background contribution  $\sum_k \beta_k B_k$  in a given signal region. The background yields  $B_k$  are those obtained either from estimation techniques as described in Chapter 8, or purely from MC simulation, with any normalisation factors  $\beta_k$  appropriate to that background. Both signal and background yields are scaled by response functions  $\nu$  which parametrise the effect on the yields of the systematic uncertainty  $\theta$ . The systematic uncertainties enter the fit as nuisance parameters in  $\mathcal{L}$ , and as  $\theta$  represents different systematic sources, the correlations between each of these is taken into account in the fit as well.

For  $f_2$ , the product runs over all profiled CRs  $l$ , and has an expected yield  $\sum_k \beta_k B_{kl}$ . The normalisation factors are the same as those appearing in  $f_1$ .

$g_3$  is made up of Gaussian terms which are of the form  $g(\vartheta|\theta) = e^{-(\vartheta-\theta)^2/2}/\sqrt{2\pi}$ . The  $\vartheta$  is the nominal value of a quantity, which has an associated nuisance parameter  $\theta$ . Together, these terms represent the systematic uncertainties which impact the signal and background yields, which will be described in Chapter 10.

The last Poisson term  $f_4$  takes into account the error due to the finite number of generated MC events for each background process  $k$ . The number of actual MC events used is given by  $\xi$ , with the expected value  $\zeta \cdot \theta$ , where  $\theta$  is a nuisance parameter.  $\zeta$  represents the square of the ratio of the background yield estimate  $B$  to its statistical uncertainty<sup>1</sup>  $\delta$ :  $\zeta = (B/\delta)^2$ . The likelihood  $\mathcal{L}$  is then maximised with respect to its arguments  $\mu$  and  $\theta$ , and evaluated at  $\vartheta = 0$  and  $\xi = \zeta$ , in order to determine  $\mu_{obs}$ .

---

<sup>1</sup>As discussed previously in Section 6.2, if there are  $N_{MC}$  events present in the MC sample with effective luminosity  $L_{MC}$ , which corresponds to a data sample with luminosity  $L_{data}$ , then the background yield  $B$  is given as  $B = (L_{data}/L_{MC}) \cdot N_{MC}$ . As the number of generated events is large, the statistical uncertainty on each MC prediction is given by the standard deviation  $\sqrt{N_{MC}}$  of a Gaussian distribution, such that the statistical uncertainty on the background yield is  $\delta = (L_{data}/L_{MC}) \cdot \sqrt{N_{MC}}$ .

## 9.3 CL<sub>s</sub> Method

The CL<sub>s</sub> method [119] (from confidence levels) is a method used to set upper limits - also referred to as *exclusion limits* - on model parameters, and is used to set limits on many experimental results at the LHC. It is used here to compute the 95% confidence level (CL) exclusions and the  $p_0$  value using the distribution of a test statistic  $q_\mu$ . The limit calculation can be interpreted as a hypothesis test, where the null hypothesis (background-only) assumes the absence of a signal, and the alternative hypothesis assumes its existence. In this case, the assumption of the alternative hypothesis is that of a SM Higgs boson with  $m_H = 125.36 \pm 0.41$  GeV, corresponding to the central value of the ATLAS measurement in the  $ZZ \rightarrow 4\ell$  and  $\gamma\gamma$  decay modes [26].

### 9.3.1 Test Statistic, p-values and Significance

A test statistic is a way of quantifying a property of a sample, which in this case is an indicator of how signal or background-like the observed result is, so that a hypothesis test can be performed. In order to derive a limit on  $\mu_{obs}$ , a test statistic  $q_\mu$  is defined, which is a function of  $\mu$ :

$$q_\mu = -2 \ln \frac{\mathcal{L}(\mu, \boldsymbol{\theta})}{\mathcal{L}_{max}} \Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_\mu} \quad (9.2)$$

where  $\mathcal{L}_{max}$  is the maximum possible value of  $\mathcal{L}$  (and hence is not dependent on  $\mu$  or  $\boldsymbol{\theta}$ ), and  $\hat{\boldsymbol{\theta}}_\mu$  are the  $\boldsymbol{\theta}$  values which maximise  $\mathcal{L}$  for a given  $\mu$ . The value of  $\mu$  which maximises  $\mathcal{L}$  is  $\hat{\mu}$ . The range of  $\hat{\mu}$  is restricted to  $0 \leq \hat{\mu} \leq \mu$  in order for the results to remain physical for the limit calculation. In general however, the value of  $\mu$  is allowed to be negative in order to account also for downward fluctuations.  $\frac{\mathcal{L}(\mu, \boldsymbol{\theta})}{\mathcal{L}_{max}}$  is known as the profile likelihood ratio and takes on values between 0 and 1 by construction, where values close to 1 imply good agreement between the observed result and the hypothesised  $\mu$ . For the null hypothesis,  $\mu = 0$  by definition, and hence the corresponding test statistic achieved in this case is  $q_0$ .

Using a test statistic, a probability can be calculated which expresses the strength of the evidence against the null hypothesis in favour of the alternative, known as a *p-value*. The desired p-value is that which gives the probability that, were there no SM Higgs boson in existence, a statistical fluctuation that could explain the observed data would occur. Mathematically, this is given by the  $p_0$  value, defined as the probability of obtaining a value of  $q_0$  equal to or larger than the observed value under the null hypothesis. Here,  $\hat{\mu}$  and therefore  $q_0$  also, are allowed to be negative so that downward fluctuations can be quantified as well. Both  $p_0$  and  $q_\mu$  depend implicitly on the assumed value of the SM  $m_H$  as this changes the expected yield.

In particle physics, for the purposes of excluding a hypothesis, a threshold p-value of 0.05 is considered, otherwise known as a 95% confidence level. The p-value is also

converted into an equivalent significance  $Z_0$ , which expresses the result in terms of the number of standard deviations,  $\sigma$ . Further details can be found in Ref. [118].

## 9.4 Combination of Results

In order to form a full picture of the VBF  $H \rightarrow WW^{(*)} \rightarrow l\nu l\nu$  channel, the final results from the 7 TeV and 8 TeV analyses are combined. Within the combination of results it is important to take into account the correlated effects of systematic uncertainties across signal and background processes, and lepton flavour channels, as well as between the two data-taking periods. The correlation scheme used follows closely the recommendations by the ATLAS Combined Performance Groups and the LHC Higgs Cross Section Working Group [120, 121, 122].

# Chapter 10

## Systematic Uncertainties

It is important in this analysis to estimate any possible systematic uncertainties which are those arising for example, from the limited knowledge of the reconstructed physics objects, or the precision of the luminosity measurement. Further sources of systematic uncertainties emerge from the background estimation methods described in Chapter 8 and the theoretical uncertainties associated with the signal and background processes. In the context of uncertainties, differences between the BDT and cut-based analyses arise only from the background estimation procedures. Estimates for all the additional systematic uncertainties for the  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  VBF analysis are presented here.

In general, the 7 TeV analyses adopt the same uncertainties as for the 8 TeV analyses, with the exception of those from experimental sources for instance, detector, luminosity, pile-up and trigger uncertainties, which are calculated using the detector conditions during the 2011 data-taking period. In cases where the binning differs, the convention followed is to take the larger of the theoretical uncertainties. For instance in the 7 TeV BDT analysis, the  $ee + \mu\mu$  channel contains a single merged bin instead of three separate regions and thus, the largest of the uncertainties calculated for each of the three bins is used.

For each experimental systematic uncertainty source, such as the lepton energy scale or jet energy resolution, the uncertainty response is varied by one standard deviation in both directions. The resulting change in the lepton or jet energy is then propagated to the relevant uncertainty calculation and their impact determined by observing the changes in the final event yields. This procedure is explained in greater detail in Ref.[96] and does not apply to systematic uncertainties arising from background estimations, theoretical assumptions or the luminosity measurement.

### 10.1 Leptons

The simulation of both electrons and muons in MC are limited by the level of detail available in the Geant4 detector simulation, as well as possible mismodelling in within the simulation. The intrinsic resolution of the various detectors as well as uncertainty

on the magnetic field can also be a source of systematic uncertainties for the lepton momentum and energy measurements. Additional systematic uncertainties arise from the uncertainty on the efficiency of the lepton reconstruction, identification and triggers used. All such resolutions and efficiencies are determined from data.

### 10.1.1 Lepton Efficiency, Energy Scale and Resolution

The lepton efficiency uncertainties refer to those originating from the lepton reconstruction and identification procedures, and are determined by the ATLAS performance groups [51, 77, 78, 79]. As for the trigger efficiencies described in Section 6.1.2, the calculation of such uncertainties are carried out with the tag-and-probe method, using a data sample of  $Z/\gamma^* \rightarrow ee + \mu\mu$  candidates. Lepton energy scale and resolution uncertainties originate from the calibration of the inner detector, ECAL and MS subdetector systems using test beams [123, 124]. As a cross check, further in-situ calibrations with  $Z$  boson and  $J/\psi$  decays are used to extract a resolution measurement for both the lepton energy scale and resolution by fitting the data to the peak of the corresponding resonance. From these results, a further scale and resolution correction can be obtained, as well as uncertainty estimates from both sources. In the context of the VBF analysis, the lepton efficiency uncertainties are all of  $\mathcal{O}(1\%)$ , for both the 2011 and 2012 data-taking periods.

## 10.2 Jets

Jet systematic uncertainties are crucial for the VBF analysis, as it relies heavily on the identification of jets and their properties. The dominant sources of uncertainty appear from the jet energy measurements, and the identification of  $b$ -jets. These jet uncertainties are common to the 7 TeV and 8 TeV analyses as the jet selection used is the same, with the exception of the JVF cut which is found to have a negligible uncertainty contribution.

### 10.2.1 Jet Energy Scale and Resolution

Jet energy scale (JES) and jet energy resolution (JER) uncertainties together cover any systematics arising from the jet calibration and resolution in the detector. JES and JER corrections are applied to simulated jets in MC, to account for data-MC differences as for the lepton efficiency scale factors, and are provided by the ATLAS JetMET Combined Performance group. As for the lepton energy scale, the JES is determined from a combination of test beam calibrations and in-situ measurements, such as tag and probe measurements [122]. JES systematics are calculated as a function of jet  $p_T$  and  $\eta$ , where the JES uncertainty tends to be higher for lower jet  $p_T$  and higher jet  $\eta$  values, as expected due to the detector capacity. For the VBF analyses, the JES uncertainties enter most strongly into the VBF Higgs signal yield, and the top quark production backgrounds, with maximum effect 4.9% and 6.0% respectively, depending on the  $p_T$  and  $\eta$  of the jet [63]. JER systematics are an additional uncorrelated source, determined

also from in-situ measurements [125, 126]. The impact of the JER uncertainties range from 2% to 40%, with the largest uncertainty associated with jets around the 20 GeV threshold of the jet selection. Translating these values into their effects on the analysis, an overall combined uncertainty of 5.4% and 2.7% is added to the VBF Higgs signal yield, and that of the cumulative background respectively for the 8 TeV BDT analysis.

### 10.2.2 *b*-jet Identification

The efficiency with which the MV1 algorithm identifies *b*-jets is measured in a high-statistics data sample of  $t\bar{t}$  pair candidates with two leptons in the final state [62]. This method is based on a likelihood fit to the data, and is also combined with another calibration method using samples containing muons reconstructed in the vicinity of a jet, in order to obtain the highest precision possible [127]. Uncertainties related to *b*-jet identification and the reconstruction of *c*-jets as *b*-jets are both  $p_T$  dependent and range from 0.01% to 7.8%, and 6% to 14% respectively. Additional uncertainties due to the misidentification of light jets as *b*-jets are  $p_T$  and  $\eta$  dependent, and have an impact between 9% and 19%. In the context of the total background yield in the VBF analysis however, the efficiency and misidentification uncertainties add only  $\sim 2\%$  each for the 7 TeV and 8 TeV analyses.

### 10.2.3 Jet Vertex Fraction

Contributions towards the systematic uncertainties could enter via the use of the jet vertex fraction during the jet selection. Such effects were studied using recommendations from the ATLAS JetMET Combined Performance group [128] and found to be negligible for all regions.

### 10.2.4 Pile-up

As the level of pile-up increased significantly between the 2011 and 2012 data-taking periods, a detailed investigation was conducted into the related pile-up uncertainties [96]. These originate from the choice of the pile-up reweighting factor, and within the jet energy calibration uncertainties. An additional source of pile-up systematics arise from event migration between jet multiplicity bins, which can occur if a pile-up produced jet has sufficient energy to be reconstructed as an analysis jet. The total impact of pile-up systematics for the 8 TeV analyses on the VBF Higgs signal yield is 1.7%, and 0.8% for the total background yield. For the 7 TeV analyses, the effect of pile-up is smaller due to the lower centre-of-mass energy and is thus considered negligible.

## 10.3 Missing Transverse Energy

Missing transverse energies are calculated from other high  $p_T$  reconstructed objects, such as jets and leptons, which already have systematic uncertainties assigned to them

as described in the sections above. However, the remaining uncertainty introduces itself via pileup, sub-threshold jets, and the measurement of soft (low  $p_T$ ) calorimeter deposits not associated with any analysis objects. Both the VBF BDT and cut-based analyses use  $\cancel{E}_T$  together with  $\cancel{E}_T^{jet,Track}$ , for which separate systematic uncertainties are assigned, by varying the scale and resolution at which these low  $p_T$  components are measured. For  $\cancel{E}_T$ , an overall uncertainty of 17% has been derived for events with  $\cancel{E}_T > 45$  GeV, whereas the effect induced by using  $\cancel{E}_T^{jet,Track}$  was found to be negligible as this variable uses mainly tracker information [96]. In total, missing transverse energies contribute 1.2% and 1.6% to uncertainties for the 8 TeV BDT analysis on the VBF Higgs signal yield and the total background yield respectively. Uncertainty values determined for the 7 TeV BDT analysis are smaller due to reduced pile-up effects.

## 10.4 Luminosity

From the van der Meer scans used to calculate the luminosity, the uncertainty is 1.8% for  $4.5 \text{ fb}^{-1}$  of data taken in 2011 at 7 TeV, and 2.8% for  $20.3 \text{ fb}^{-1}$  of data taken in 2012 at 8 TeV [47]. The dominant source of uncertainty here is the calibration uncertainty from the van der Meer scan.

## 10.5 Theoretical Uncertainties

For the VBF Higgs signal process, the theoretical uncertainties on the expected cross section in the SR must be taken into account. Similarly, the uncertainties on the yield are needed for a number of backgrounds such as the non- $WW$  diboson processes ( $WZ$ ,  $ZZ$ ,  $Z\gamma$ ,  $W\gamma$  and  $W\gamma^*$ ) and the ggF Higgs production, which are normalised directly from the theoretical MC prediction. The other background processes however are normalised using data, and thus don't require uncertainties associated with the cross sections, but rather on the extrapolation between the relevant CR and SR. As these have already been discussed in Chapter 8, they are only included in the overall summary in Table 10.1 for completeness. As for the other estimated backgrounds, the sources of uncertainty considered are: QCD scale, PDF choice, PS choice and also the generator modelling uncertainties. Additional uncertainties relevant to a particular process are considered in some cases. This section summarises the impact of theoretical uncertainties for both signal and background processes, with a detailed discussion of the Higgs production cross section uncertainties and those of the non- $WW$  diboson backgrounds. All the theoretical uncertainties on any extrapolation factors are adopted directly by the corresponding BDT or cut-based 7 TeV analyses.

### 10.5.1 VBF Higgs Production

Uncertainties on the VBF Higgs production cross section are calculated separately for the BDT and cut-based analyses, as the final fit region varies slightly in the region of

variable phase-space covered. Variations of the QCD renormalisation and factorisation scales are found to have negligible effect on the VBF Higgs yield in the BDT analysis, although it adds an uncertainty of 3.6% for that of the cut-based analysis. The choice of PDF, evaluated using the CT10 and NNPDF sets, contributes roughly 2% to each BDT signal-like bin, which translates into an overall uncertainty of 4%, agreeing well with the corresponding cut-based value, which is also 4%. PS uncertainties determined by studying the difference between POWHEG showered with HERWIG or PYTHIA 8, provide values of [ $< 1\%$ , 5.2%, 14%] and 8% for the BDT and cut-based analyses respectively, which also agrees well between the analyses if the average of the BDT values are taken. Modelling uncertainties are obtained by comparing the POWHEG and aMC@NLO [116] MC predictions, both showered using HERWIG, which give 4.2% for the BDT analysis and 2.4% for the cut-based. A further source of uncertainty was considered for the VBF Higgs process, namely the effect of the QCD scale in combination with the use of the CJV in the event selection. The effect of the CJV is taken into account here<sup>1</sup>, as it impacts the jet multiplicity in the event selection, which in turn is affected by the QCD scale. This brings an additional 3% and 4% for the BDT and cut-based analyses.

## 10.5.2 ggF Higgs Background

Higgs production via ggF (with two or more jets) is an important, although minor, background for the VBF analysis, as any contamination arising from this process must be estimated. All other standard sources of uncertainties are found to be negligible except those from PS which adds up to 15% when comparisons between POWHEG +HERWIG and POWHEG +PYTHIA 8 are considered, and from the impact of the QCD scale on the CJV acceptance, as for the VBF Higgs process. This combined jet multiplicity and veto uncertainty is evaluated using the Stewart-Tackmann method [129] which treats the inclusive ggF Higgs+2 jets and Higgs+3 jets cross sections as uncorrelated quantities. The uncertainties originating from this source are [30%, 30%, 56%] and a symmetrised value of 32% for the BDT and cut-based analyses respectively. Further details on the ggF theoretical uncertainty estimations can be found in Ref. [112].

## 10.5.3 Non- $WW$ Diboson Production

Other than the  $WW$  background, the other diboson processes namely  $WZ$ ,  $ZZ$ ,  $Z\gamma$ ,  $W\gamma$  and  $W\gamma^*$  have no background estimation as their contribution to the total background yield is relatively small. Their yields in the SR for both the BDT and cut-based analyses are simply normalised to that of the MC prediction from the various generators, scaled to the measured integrated luminosity for 7 TeV and 8 TeV each.

Theoretical uncertainties on the  $W\gamma$  process cross section are calculated using MCFM [130] following the procedure described in the Ref. [112], which assigns a 100%

---

<sup>1</sup>Essentially, the CJV defines regions which include at least two or three jets

uncertainty for both the BDT and cut-based analyses. In addition, an uncertainty of 3.1% is applied due to the choice of PDF set used. For the  $W\gamma^*$  and  $WZ$  backgrounds, MCFM is also used to calculate the theoretical uncertainties, resulting in an overall 26% error on the cross sections. Other non- $WW$  diboson production backgrounds are not considered due to their small component in the total background yield.

## 10.6 Final Impact on Results

In order to check the impact of all the experimental and theoretical uncertainties, these are entered into the fit procedure as nuisance parameters in the likelihood function, and their impact on the total signal and cumulative background yields are calculated. The dominant uncertainties for the signal yield are related to the theoretical Higgs process uncertainties, and that of the jet energy scale and resolution. In contrast, for the total background yield, the uncertainties associated with the background estimation methods have the largest effect. A summary of the experimental and theoretical uncertainties on the total signal and background yields for the 7 TeV and 8 TeV BDT analyses are shown in Table 10.1. The corresponding values for the cut-based cross-check analyses are not shown as the experimental systematics are identical due to the common object reconstruction, and the remaining systematics dependent on the background estimates have been shown to have largely good agreement between them.

Statistical uncertainties are also relevant for those backgrounds which are estimated using data, namely for top quark production processes,  $Z/\gamma^* \rightarrow ee + \mu\mu$  and also  $Z/\gamma^* \rightarrow \tau\tau$ . The statistical uncertainties due to the data samples used in the derivation of  $W$ +jets and QCD multijet background extrapolation factors are listed instead as experimental uncertainties in Table 10.1. Background contamination in CRs cause anti-correlations between various background processes such that the uncertainty on the total background is smaller than the sum in quadrature from its components, known as *cross-talk*. Other systematic uncertainties also contain correlations between them, for instance in the case of JER and JES. A correlation scheme is used by the analysis which follows closely the recommendations by the ATLAS Combined Performance Groups and the LHC Higgs Cross Section Working Group [120, 121, 122] The final uncertainties on the total signal and background yields after the fitting procedure has been performed is shown in Tables 10.2 and 10.3 for the 7 TeV and 8 TeV BDT analyses. Due to cross-talk and correlation effects, these numbers are not directly comparable with those described earlier in Chapter 8.

<b>Systematic source</b>	<b>7 TeV</b>	<b>8 TeV</b>
VBF $H$ , total cross section	3.0	2.9
VBF $H$ theoretical model	3.8	5.5
ggF $H$ , $N_{jets} \geq 2$ cross section	6.7	6.9
ggF $H$ , $N_{jets} \geq 3$ cross section	2.9	3.1
ggF $H$ , total cross section	1.9	2.0
ggF $H$ theoretical	3.8	4.0
$H \rightarrow WW^{(*)}$ branching fraction	4.3	
Integrated luminosity	1.8	2.8
Jet energy scale & reso.	5.2	5.4
Missing transverse energy scale & resolution	0.2	1.2
Trigger efficiency	0.2	0.4
Electron id., iso., reco. eff.	2.3	1.0
Muon id., iso., reco. eff.	1.1	0.9
Pile-up model	-	1.7

<b>Systematic source</b>	<b>7 TeV</b>	<b>8 TeV</b>
$WW$ theoretical model	4.2	3.0
$VV$ theoretical model	2.1	0.5
Top estimate	3.1	3.0
$Z/\gamma^* \rightarrow \tau\tau$ estimate	11	1.6
$Z/\gamma^* \rightarrow ee + \mu\mu$ estimate	9.2	4.8
$W$ +jets estimate	1.7	1.3
QCD Multijet estimate	0.4	0.9
Integrated luminosity	0.4	
Jet energy scale & reso.	7.5	2.7
Missing transverse energy scale & resolution	0.1	1.6
$b$ -tagging efficiency	1.3	2.0
Light- and $c$ -jet mistag	1.3	2.0
Trigger efficiency	0.2	-
Electron id., iso., reco. eff.	1.1	0.3
Muon id., iso., reco. eff.	0.7	0.2
Pile-up model	-	0.8

Table 10.1: Sources of uncertainty (in %) on the signal yield and the cumulative background yields. Entries marked with a dash show that the corresponding uncertainties either do not apply or are less than 0.1% and therefore neglected. The values are given for the 7 TeV and 8 TeV BDT analysis only, as the corresponding values for the cut-based analyses are very similar if not identical.  $VV$  refers to the non- $WW$  diboson backgrounds.

<b>Sample</b>	<b>Total error</b>	<b>Stat. error</b>	<b>Expt. syst. err.</b>	<b>Theo. syst. err.</b>
$N_{sig}$	12	-	6.2	11
$N_{bkg}$	23	15	16	5.6
$N_{WW}$	31	-	11	29
$N_{top}$	46	41	11	13
$N_{misid}$	29	-	29	4.9
$N_{VV}$	42	-	17	38
$N_{Z/\gamma^* \rightarrow \tau\tau}$	68	60	32	2.6
$N_{Z/\gamma^* \rightarrow ee+\mu\mu}$	57	45	35	-

Table 10.2: Composition of uncertainty (in %) on the total signal, total background, and individual background yields in the signal region for the 7 TeV BDT analysis. The total uncertainty (total) is decomposed into three different components: statistical (stat.), experimental (expt.) and theoretical (theo.). Entries marked with a dash show that the uncertainties either do not apply or are less than 0.1%.

<b>Sample</b>	<b>Total error</b>	<b>Stat. error</b>	<b>Expt. syst. err.</b>	<b>Theo. syst. err.</b>
$N_{sig}$	13	-	6.8	12
$N_{bkg}$	9.2	4.7	6.4	4.5
$N_{WW}$	32	-	14	28
$N_{top}$	15	9.5	7.6	8.5
$N_{misid}$	22	-	12	19
$N_{VV}$	20	-	12	15
$N_{Z/\gamma^* \rightarrow \tau\tau}$	40	25	31	2.9
$N_{Z/\gamma^* \rightarrow ee+\mu\mu}$	18	11	15	-

Table 10.3: Composition of uncertainty (in %) on the total signal, total background, and individual background yields in the signal region for the 8 TeV BDT analysis. The total uncertainty (total) is decomposed into three different components: statistical (stat.), experimental (expt.) and theoretical (theo.). Entries marked with a dash show that the uncertainties either do not apply or are less than 0.1%.

# Chapter 11

## Results

This Chapter presents the yields and distributions of the described analyses for the 7 TeV and 8 TeV data-taking periods, followed by an interpretation of the observed results. Signal and background yields for all processes will be shown at each stage of the selection described in Chapter 7. So that these results can be related to in a broader context, a brief discussion of the latest combined  $H \rightarrow WW^{(*)}$  analysis is also included at the end.

### 11.1 8 TeV BDT and Cut-based yields

A summary of total signal and background yields, for the BDT and cut-based analysis are shown in Tables 11.1 and 11.2. Further cutflows for the BDT analysis are shown in Tables 11.3 and 11.4 divided into each lepton flavour channel. The corresponding cut-based tables are included for completeness in the Appendix. Although the  $VH$  production mode is included as signal, it makes up  $\mathcal{O}(1\%)$  of the expected signal and therefore is negligible in all final signal regions. Only statistical uncertainties are included in these Tables, as the correct inclusion of systematic uncertainties can only be performed within the fit, due to correlation effects.

Distributions of key variables used in the BDT analysis are shown in Figures 11.1 and 11.2 for the  $e\mu + \mu e$  and  $ee + \mu\mu$  flavour channels respectively. The BDT response distributions which enter the fitting procedure can be seen in Figure 11.3, with the equivalent  $m_T$  distributions for the cut-based analysis in Figure 11.4. From these BDT response and  $m_T$  distributions, it can be seen that the  $e\mu + \mu e$  channel in both cases carries the bulk of the sensitivity to the signal. The expected  $Z_0$ , calculated as described in Chapter 9, for the BDT analysis is  $2.1\sigma$  and  $1.2\sigma$  for the  $e\mu + \mu e$  and  $ee + \mu\mu$  channels respectively, and  $2.4\sigma$  if combined. The corresponding numbers for the cut-based analysis are  $1.8\sigma$  and  $0.8\sigma$ , and  $2.0\sigma$  if combined. Roughly, the improvement based on the expected  $Z_0$  of the BDT analysis over that of the traditional cut-based one is a gain of 17% in the  $e\mu + \mu e$  channel and 50% in the  $ee + \mu\mu$  channel.

Cuts	VBF+VH	Total Bkg.	Data	Data/MC
$\cancel{E}_T > 45$ GeV	92.94±1.30	103368.21±327.25	107093	1.04
$N_{jets} \geq 2$	90.71±1.23	85167.63±89.43	89071	1.04
$b$ -Veto	65.86±0.99	15588.16±68.69	16067	1.03
$\cancel{E}_T^{jet,Track} > 40$ GeV	63.19±0.98	10764.79±38.09	11185	1.03
CJV	52.80±0.89	8598.06±34.80	8966	1.04
OLV	29.81±0.43	1965.41±14.22	1928	0.97
$m_{\tau\tau} < 66.1876$	25.06±0.37	1140.98±9.91	1187	1.02
BDT > -0.48	18.07±0.19	99.80±2.89	130	1.10
-0.48 < BDT < 0.3	6.50±0.12	91.89±2.83	90	0.91
0.3 < BDT < 0.78	6.71±0.11	16.78±0.81	28	1.19
BDT > 0.78	4.86±0.09	2.87±0.25	12	1.55

Table 11.1: Summary of total expected signal and background, and data yields at each selection stage after the common preselection, for all the lepton flavour channels combined within 8 TeV BDT analysis. Only statistical uncertainties are given here. The last column shows the ratio of data to total MC simulation.

Cuts	VBF+VH	Total Bkg.	Data	Data/MC
$\cancel{E}_T > 55$ GeV				
$\cancel{E}_T^{jet,Track} > 50$ GeV	147.50±1.86	165617.47±126.85	171907	1.04
$N_{jets} \geq 2$	81.78±1.21	88383.04±55.58	88383	1.00
$b$ -Veto	59.54±0.99	10823.18±36.54	10919	1.01
$p_T^{tot,jet,Track} < 15$ GeV	50.69±0.88	7819.69±31.95	7957	1.02
$m_{\tau\tau} < 66.1876$	41.52±0.76	4290.20±17.48	4473	1.04
$m_{jj} > 600$ GeV	12.02±0.15	169.68±4.65	212	1.25
$\Delta Y_{jj} > 3.6$	11.46±0.14	130.23±4.50	160	1.23
CJV	9.58±0.13	65.95±1.57	89	1.35
OLV	9.31±0.13	54.95±1.46	77	1.40
$\Delta\phi_{\ell\ell} < 1.8$	8.60±0.13	32.95±1.09	49	1.49
$m_{\ell\ell} < 50$ GeV	7.57±0.12	17.45±0.86	28	1.60
$m_T < 130$ GeV				
$m_T > 80$ GeV for low $p_T$	6.90±0.11	9.23±0.62	20	2.17
$m_{jj} < 1$ TeV	3.40±0.08	6.45±0.53	11	1.17
$m_{jj} > 1$ TeV	3.50±0.08	2.87±0.32	9	3.13

Table 11.2: Summary of total expected signal and background, and data yields at each selection stage after the common preselection, for all the lepton flavour channels combined within 8 TeV cut-based analysis. Only statistical uncertainties are given here. The last column shows the ratio of data to total MC simulation.

Cuts	VBF+VH	Total Bkg.	Data	Data/MC
$N_{jets} \geq 2$	57.42±1.03	59017.25±38.64	61434	1.04
$b$ -Veto	41.95±0.86	7556.34±23.46	7818	1.03
CJV	35.13±0.79	6086.82±21.08	6313	1.03
OLV	19.33±0.37	1368.88±10.04	1264	0.91
$m_{\tau\tau} < 66.1876$	16.51±0.32	732.87±6.33	718	0.96
BDT > -0.48	11.64±0.15	45.26±1.41	57	1.00
-0.48 < BDT < 0.3	4.28±0.10	43.28±1.42	37	0.78
0.3 < BDT < 0.78	4.23±0.09	7.98±0.49	14	1.15
BDT > 0.78	3.12±0.07	1.55±0.16	6	1.28

Table 11.3: Summary of total expected signal and background, and data yields at each selection stage after the common preselection, for the  $e\mu+\mu e$  channel within the 8 TeV BDT analysis. Only statistical uncertainties are given here. The last column shows the ratio of data to total MC simulation.

Cuts	VBF+VH	Total Bkg.	Data	Data/MC
$\cancel{E}_T > 45$ GeV	35.53±0.80	44350.96±118.52	45659	1.03
$N_{jets} \geq 2$	33.29±0.68	26150.38±80.65	27637	1.06
$b$ -Veto	23.91±0.49	8031.82±64.56	8249	1.02
$\cancel{E}_T^{jet,Track} > 40$ GeV	21.23±0.46	3208.44±30.01	3367	1.04
CJV	17.67±0.40	2511.24±27.69	2653	1.05
OLV	10.48±0.20	596.53±10.07	664	1.09
$m_{\tau\tau} < 66.1876$	8.55±0.19	408.11±7.62	469	1.13
BDT > -0.48	6.44±0.11	54.06±2.48	73	1.21
-0.48 < BDT < 0.3	2.22±0.07	48.61±2.45	53	1.04
0.3 < BDT < 0.78	2.48±0.07	8.81±0.65	14	1.24
BDT > 0.78	1.74±0.06	1.32±0.20	6	1.96

Table 11.4: Summary of total expected signal and background, and data yields at each selection stage after the common preselection, for the  $ee+\mu\mu$  channel within the 8 TeV BDT analysis. Only statistical uncertainties are given here. The last column shows the ratio of data to total MC simulation.

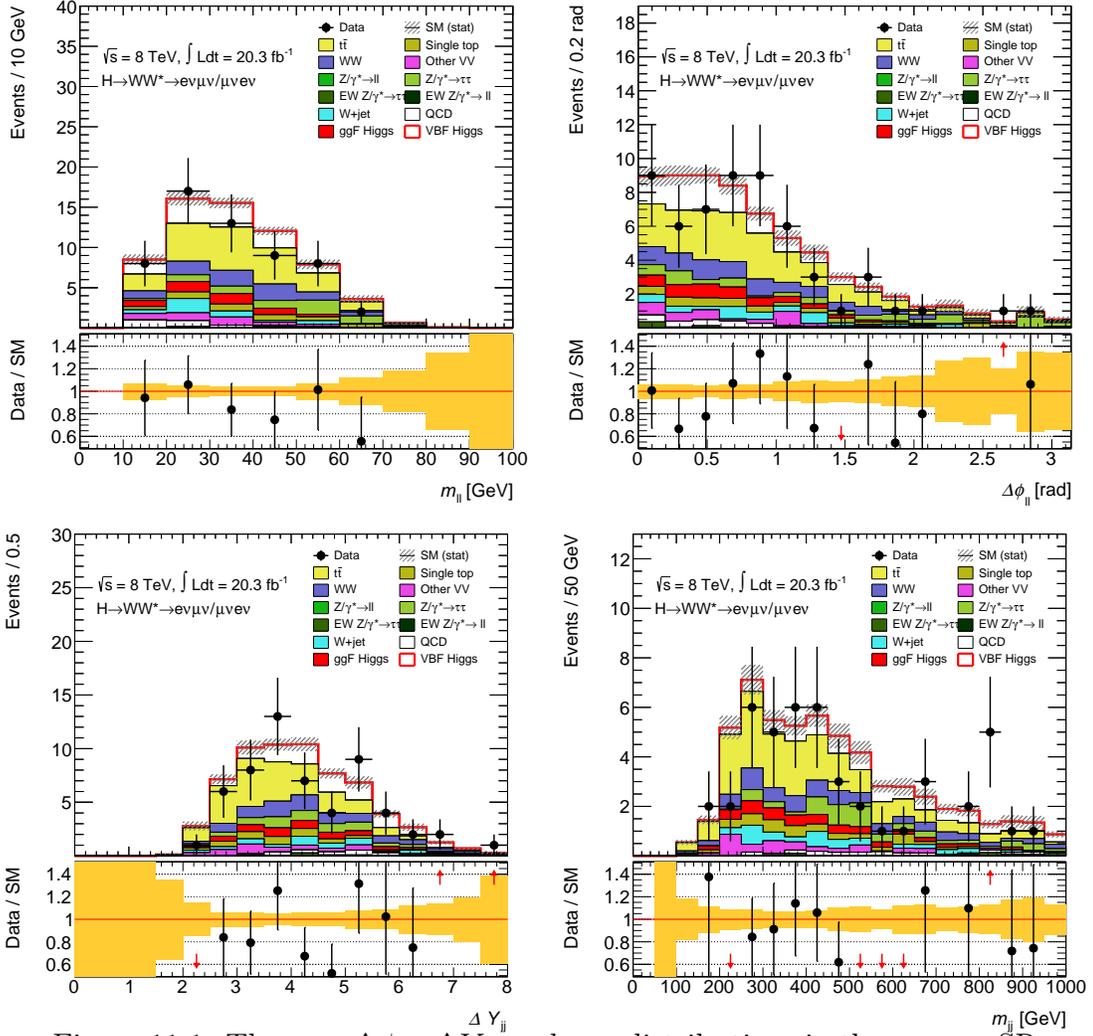


Figure 11.1: The  $m_{\ell\ell}$ ,  $\Delta\phi_{\ell\ell}$ ,  $\Delta Y_{jj}$  and  $m_{jj}$  distributions in the  $e\mu + \mu e$  SR region for the 8 TeV BDT analysis. Only statistical errors are considered here.

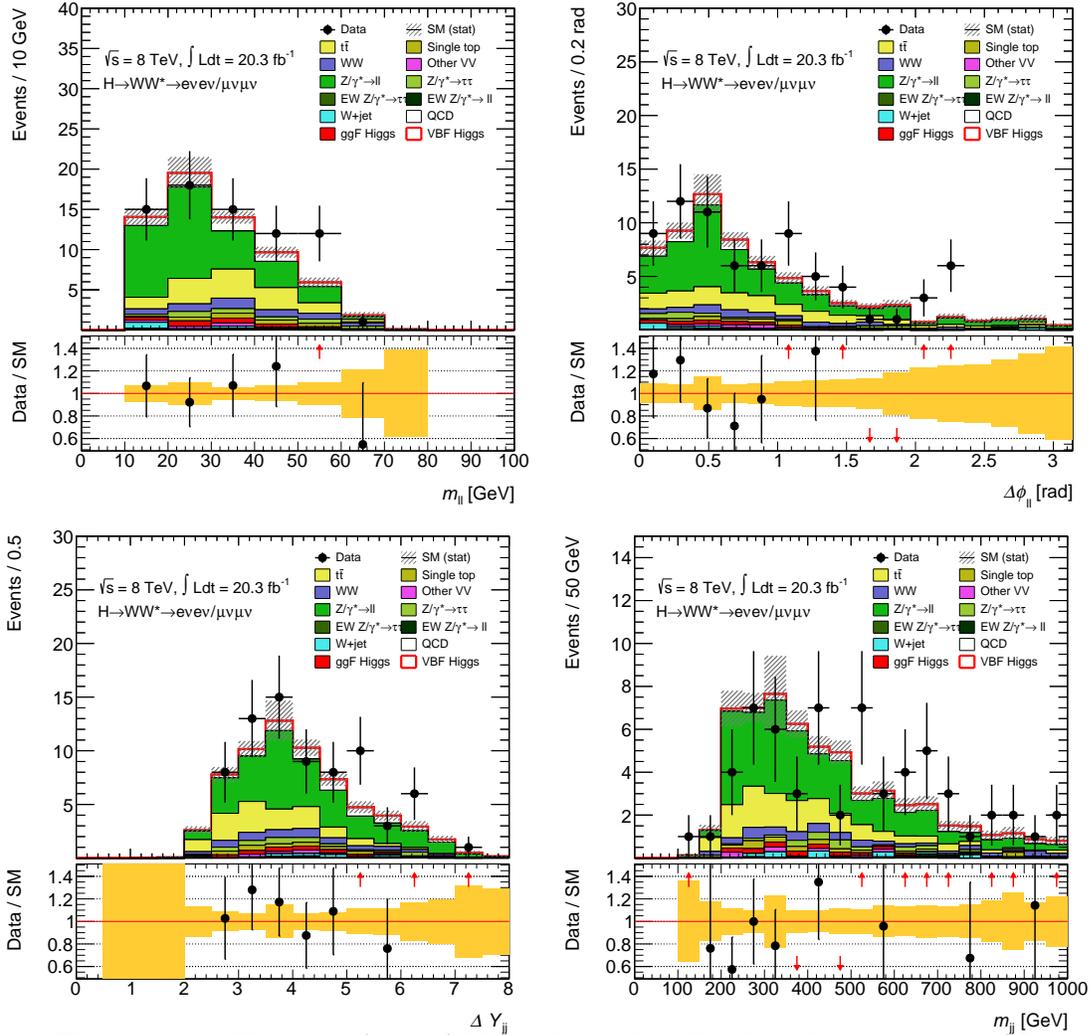


Figure 11.2: The  $m_{\ell\ell}$ ,  $\Delta\phi_{\ell\ell}$ ,  $\Delta Y_{jj}$  and  $m_{jj}$  distributions in the  $ee + \mu\mu$  SR for the 8 TeV BDT analysis. Only statistical errors are considered here.

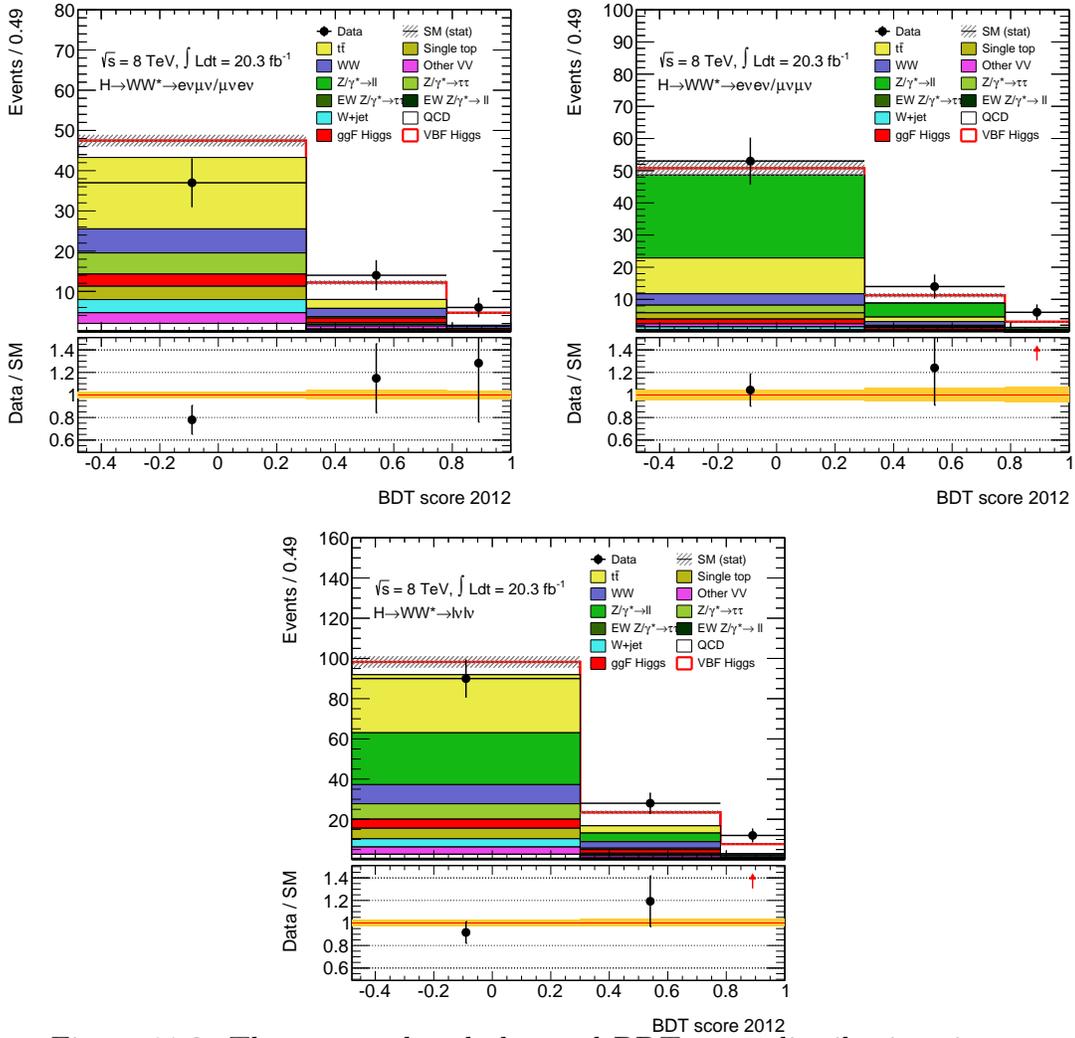


Figure 11.3: The expected and observed BDT score distributions in our final region for the 8 TeV VBF BDT analysis, for the  $e\mu + \mu e$ ,  $ee + \mu\mu$  and combined channels. A three-bin fit to the BDT response is performed in the  $e\mu + \mu e$  and  $ee + \mu\mu$  flavour channel each. The yellow band includes only statistical errors here.

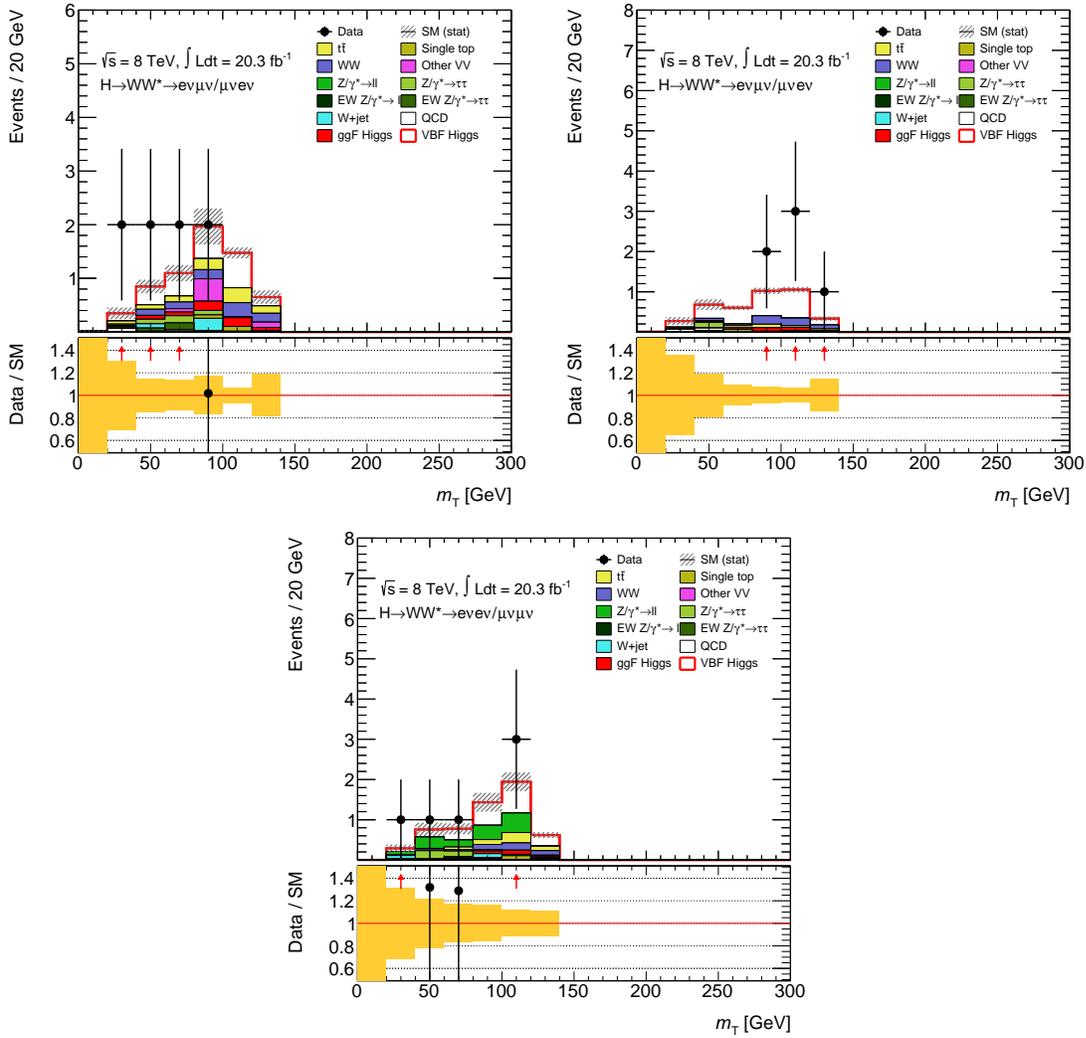


Figure 11.4: Distributions of  $m_T$  for the  $e\mu + \mu e$  channel with  $m_{jj} < 1$  TeV (top left) and  $m_{jj} > 1$  TeV (top right), and the  $ee + \mu\mu$  channel (bottom) where there is no further split. The yellow uncertainty band shown represents only the statistical MC error.

## 11.2 7 TeV BDT and Cut-based yields

The corresponding summary results for the 7 TeV dataset for the BDT and cut-based analysis are shown in Table 11.5 and 11.6. Further cutflows for the BDT analysis are shown in Tables 11.7 and 11.8 divided into each lepton flavour channel. The corresponding cut-based tables are included for completeness in the Appendix. Again, the  $VH$  production mode is included as signal although it is negligible in the SR and only statistical uncertainties are shown in these Tables, as the correct inclusion of systematic uncertainties can only be performed within the fit, due to correlation effects.

Distributions of key variables used in the BDT analysis are shown in Figures 11.5 and 11.6 for the  $e\mu + \mu e$  and  $ee + \mu\mu$  flavour channels respectively. The BDT response distributions which enter the fitting procedure can be seen in Figure 11.7, with the equivalent  $m_T$  distributions for the cut-based analysis in Figure 11.8. The expected  $Z_0$ , calculated as described in Chapter 9, for the BDT analysis is  $0.79\sigma$  and  $0.30\sigma$  for the  $e\mu + \mu e$  and  $ee + \mu\mu$  channels respectively, and  $0.84\sigma$  combined. The corresponding numbers for the cut-based analysis are  $0.72\sigma$  and  $0.38\sigma$ , and  $0.82\sigma$  when combined. Despite the lower significance of the 7 TeV analyses on their own, their combination with the 8 TeV results is extremely important.

Cuts	VBF+VH	Total Bkg.	Data	Data/MC
$\cancel{E}_T > 45$ GeV	$10.90 \pm 0.11$	$12583.70 \pm 27.05$	13098	1.04
$N_{jets} \geq 2$	$10.85 \pm 0.11$	$10387.23 \pm 24.90$	10875	1.05
$b$ -Veto	$7.61 \pm 0.08$	$1532.25 \pm 13.38$	1526	0.99
$\cancel{E}_T^{jet,Track} > 40$ GeV	$7.30 \pm 0.08$	$1269.50 \pm 11.22$	1277	1.00
CJV	$6.49 \pm 0.08$	$1058.14 \pm 10.50$	1060	1.00
OLV	$3.86 \pm 0.04$	$256.88 \pm 5.65$	250	0.96
$m_{\tau\tau} < 66.1876$	$3.25 \pm 0.04$	$137.02 \pm 3.69$	139	0.99
BDT $> -0.48$	$2.30 \pm 0.03$	$9.41 \pm 0.88$	9	0.77
$-0.48 < \text{BDT} < 0.3$	$0.85 \pm 0.02$	$7.79 \pm 0.84$	7	0.81
BDT $> 0.3$	$1.46 \pm 0.02$	$1.62 \pm 0.25$	2	0.65

Table 11.5: Summary of total expected signal and background, and data yields at each selection stage after the common preselection, for all the lepton flavour channels combined within 7 TeV BDT analysis. Only statistical uncertainties are given here. The last column shows the ratio of data to total MC simulation.

Cuts	VBF+VH	Total Bkg.	Data	Data/MC
$\cancel{E}_T > 45$ GeV	20.72±0.16	31818.96±80.40	32785	1.03
$N_{jets} \geq 2$	10.89±0.11	12884.22±27.55	13098	1.02
$b$ -Veto	7.64±0.09	1790.22±14.23	1809	1.01
$p_T^{tot,jet,Track} < 15$ GeV	6.12±0.07	1144.41±11.71	1154	1.01
$m_{\tau\tau} < 66.1876$	5.12±0.07	633.13±8.28	637	1.01
$m_{jj} > 500$ GeV	1.87±0.02	31.49±1.39	28	0.89
$\Delta Y_{jj} > 2.8$	1.85±0.02	22.35±1.07	22	0.98
CJV	1.71±0.02	14.97±0.91	13	0.87
OLV	1.64±0.02	11.75±0.73	11	0.94
$\Delta\phi_{\ell\ell} < 1.8$	1.47±0.02	6.82±0.54	8	1.17
$m_{\ell\ell} < 50$ GeV	1.28±0.02	3.15±0.37	6	1.91
$m_T < 130$ GeV	1.17±0.02	1.72±0.27	4	2.33

Table 11.6: Summary of total expected signal and background, and data yields at each selection stage after the common preselection, for all the lepton flavour channels combined within 7 TeV cut-based analysis. Only statistical uncertainties are given here. The last column shows the ratio of data to total MC simulation.

Cuts	VBF+VH	Total Bkg.	Data	Data/MC
$N_{jets} \geq 2$	6.65±0.09	7670±20.13	8042	1.05
$b$ Veto	4.69±0.07	948.28±9.33	949	1.00
CJV	4.18±0.06	796.10±8.78	799	1.00
OLV	2.50±0.04	195.84±4.89	194	0.98
$m_{\tau\tau} < 66.1876$	2.13±0.03	99.78±3.12	100	0.98
BDT > -0.48	1.48±0.02	5.08±0.54	6	0.91
-0.48 < BDT < 0.3	0.56±0.01	4.12±0.52	6	1.28
BDT > 0.3	0.93±0.02	0.96±0.15	0	–

Table 11.7: Summary of total expected signal and background, and data yields at each selection stage after the common preselection, for the  $e\mu+\mu e$  channel within the 7 TeV BDT analysis. Only statistical uncertainties are given here. The last column shows the ratio of data to total MC simulation.

Cuts	VBF+VH	Total Bkg.	Data	Data/MC
$\cancel{E}_T > 45$ GeV	4.24±0.07	4913.50±18.07	5056	1.03
$N_{jets} \geq 2$	4.19±0.07	2717.03±14.65	2833	1.04
$b$ -Veto	2.91±0.05	583.97±9.58	577	0.98
$\cancel{E}_T^{jet,Track} > 40$ GeV	2.32±0.05	321.22±6.22	328	1.01
CJV	2.32±0.05	262.03±5.77	261	0.99
OLV	1.37±0.03	61.05±2.83	56	0.90
$m_{\tau\tau} < 66.1876$	1.12±0.02	37.24±1.96	39	1.02
BDT > -0.48	0.82±0.02	4.33±0.70	3	0.58
-0.48 < BDT < 0.3	0.29±0.01	3.67±0.67	1	0.25
BDT > 0.3	0.53±0.01	0.66±0.20	2	1.68

Table 11.8: Summary of total expected signal and background, and data yields at each selection stage after the common preselection, for the  $ee+\mu\mu$  channel within the 7 TeV BDT analysis. Only statistical uncertainties are given here. The last column shows the ratio of data to total MC simulation.

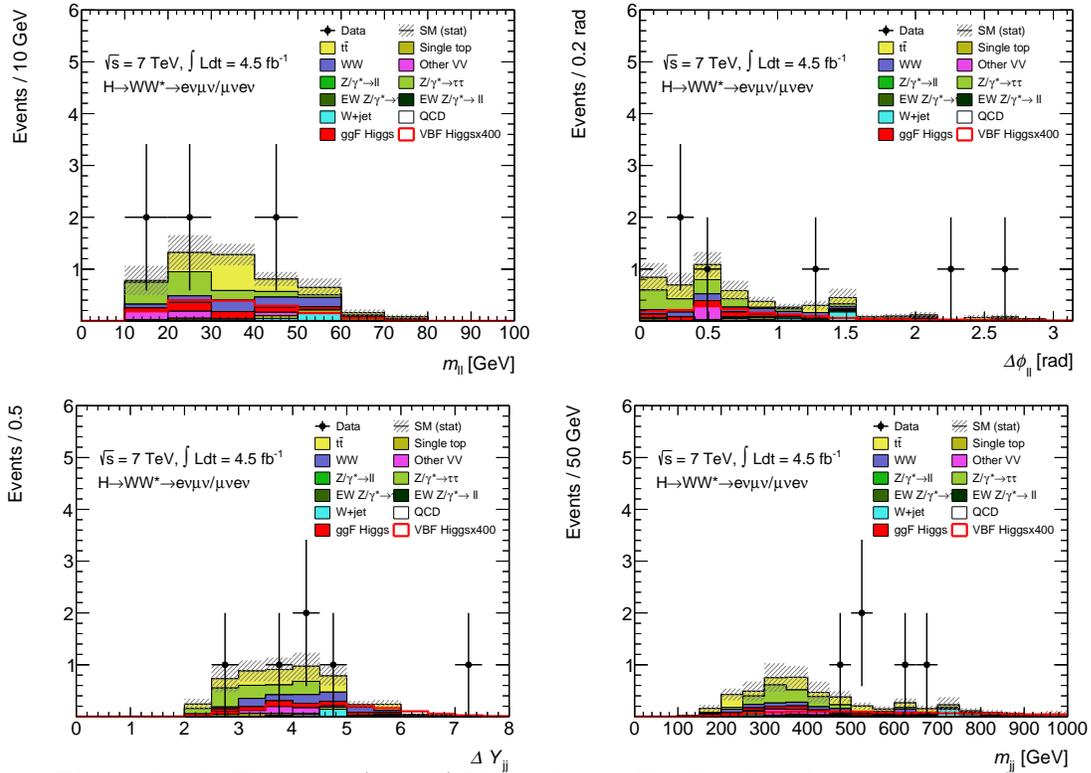


Figure 11.5: The  $m_{\ell\ell}$ ,  $\Delta\phi_{\ell\ell}$ ,  $\Delta Y_{jj}$  and  $m_{jj}$  distributions in the  $e\mu + \mu e$  SR region for the 7 TeV BDT analysis. Only statistical errors are considered here.

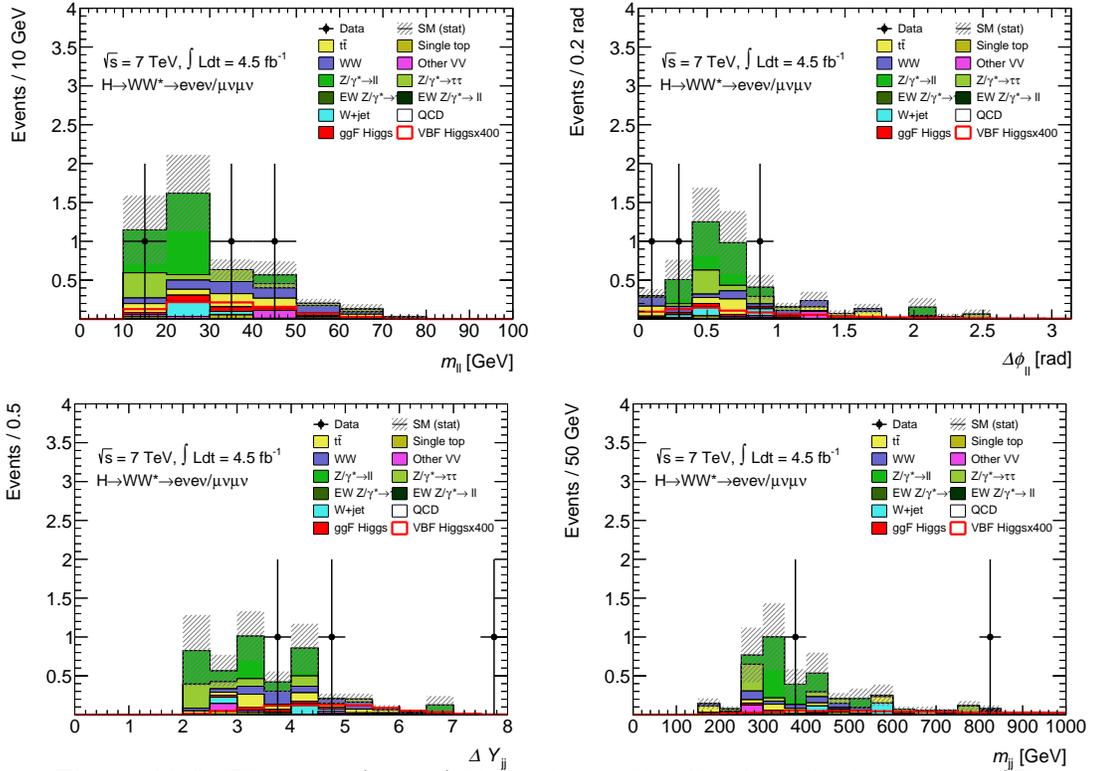


Figure 11.6: The  $m_{\ell\ell}$ ,  $\Delta\phi_{\ell\ell}$ ,  $\Delta Y_{jj}$  and  $m_{jj}$  distributions in the  $ee + \mu\mu$  SR for the 7 TeV BDT analysis. Only statistical errors are considered here.

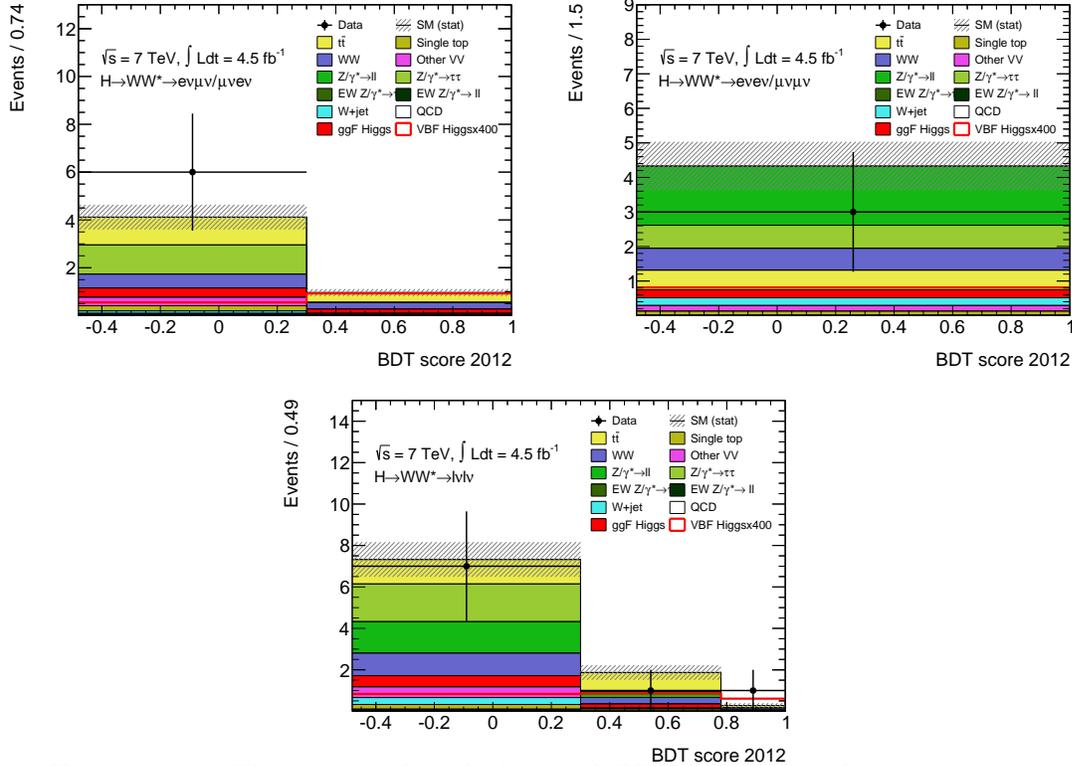


Figure 11.7: The expected and observed BDT score distributions in our final region for the 7 TeV VBF BDT analysis, for the  $e\mu + \mu e$ ,  $ee + \mu\mu$  and combined channels. For the  $e\mu + \mu e$  channel we perform a two-bin fit, whereas for  $ee + \mu\mu$  we perform a one-bin fit. The yellow band includes only statistical errors here.

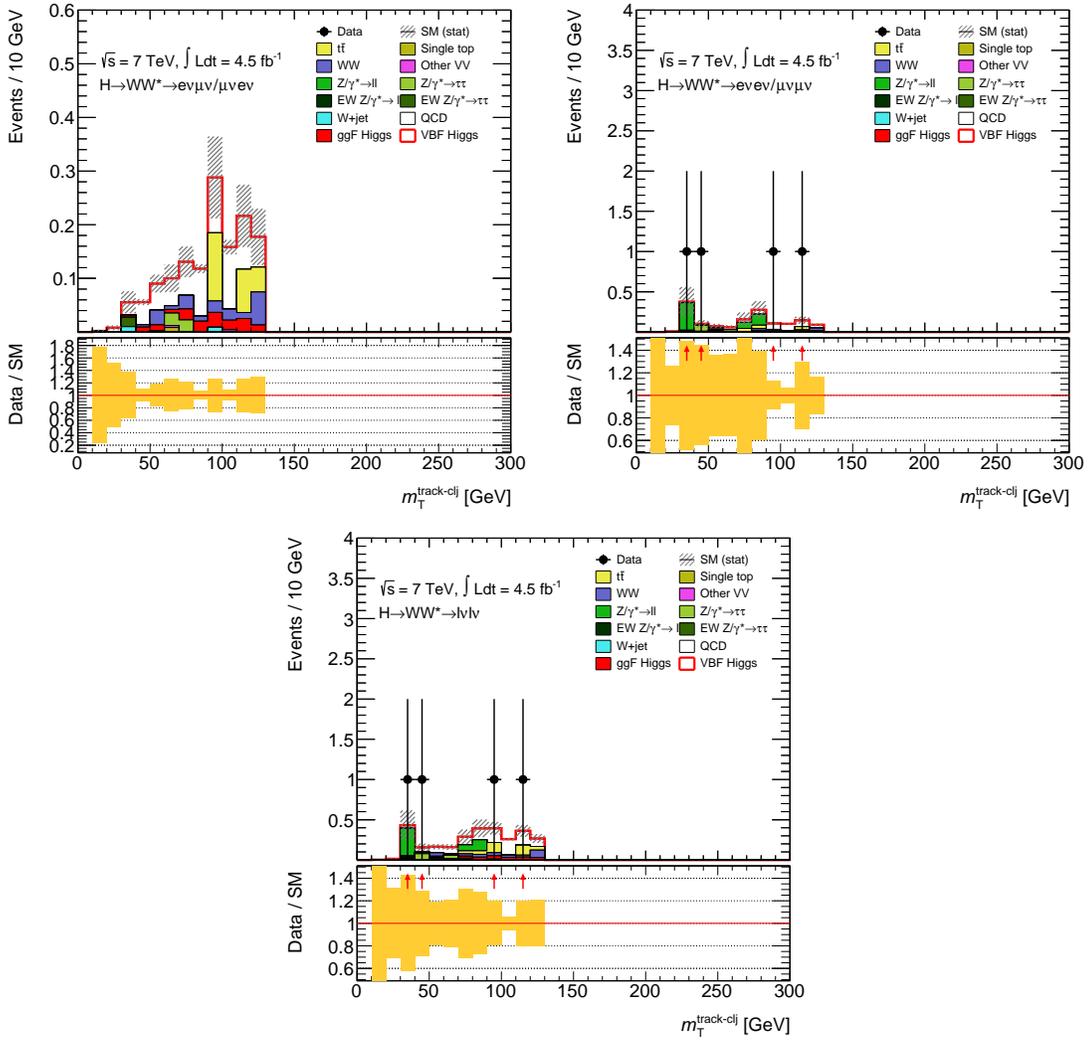


Figure 11.8: Distributions of  $m_T$  for the  $e\mu + \mu e$  (top left) and  $ee + \mu\mu$  (top right) flavour channels for the 7 TeV cut-based analysis, with the combined flavour channel plot is shown below. The yellow uncertainty band shown represents only the statistical MC error.

## 11.3 Significance and Interpretation of Results

As described in Chapter 9, the signal and background rates and shapes are allowed to vary in order to fit to the available data in the signal and control regions, using the profile likelihood procedure. Here, only the 7 TeV and 8 TeV BDT results will be considered further in detail as they contain greater sensitivity than the cut-based analyses. All results are interpreted assuming a SM Higgs boson with mass  $m_H = 125.36$  GeV [26], which is the central value of the ATLAS measurement in the  $ZZ \rightarrow 4\ell$  and  $\gamma\gamma$  Higgs decay modes. The observed  $Z_0$  calculated for the combined  $e\mu + \mu e$  and  $e\mu + \mu e$  8 TeV BDT analysis is  $3.68\sigma$ , with  $-1.24\sigma$  for the 7 TeV BDT analysis. The negative 7 TeV BDT significance value expresses the lack of events in the high BDT score  $e\mu + \mu e$  region, seen easily in Figure 11.7 or 11.8.

In order to establish the observation of the VBF Higgs production mode in the  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  state, the latest results from the complementary 7 TeV and 8 TeV ggF  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  analyses are used. This is done in order to take into account any contamination of the ggF process in the VBF signal region, by simultaneously fitting also the ggF signal regions as described in Ref. [105]. The significance of the VBF production is thus  $3.3\sigma$ , thereby establishing evidence of the VBF production mode in this final state. The likelihood scan is shown in Figure 11.9 as a function of the ratio of ggF and VBF signal strengths  $\mu_{VBF}/\mu_{ggF}$ , which uses the width of the curve to determine the significance of the VBF production signal. Cross-checks were made using the cut-based analysis results and were found to be compatible, giving further strength to the evidence.

The observed signal strength for the VBF production is also determined by simultaneously fitting the VBF and ggF processes, and calculating the  $\mu_{obs}$  for both production categories. The resulting two-dimensional likelihood scan can be seen in Figure 11.11, with  $\mu_{obs} = 1.28^{+0.44}_{-0.40}$  (*stat.*) $^{+0.29}_{-0.20}$  (*syst.*) for the VBF production - well within one standard deviation of the SM prediction.

Measurement of  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  production via the VBF (and associated production) processes are sensitive to couplings of the Higgs boson to  $W$  and  $Z$  bosons, while the ggF process probes the couplings to both heavy quarks and  $W$  bosons. The observed signal strength values for VBF and ggF can be used to test the consistency of the fermionic and bosonic couplings of the SM Higgs boson [131] through a two-dimensional likelihood scan as a function of the fermionic coupling  $\kappa_F$  and the bosonic coupling  $\kappa_V$ . Through the VBF results, a value of  $\kappa_V = 1.04^{+0.10}_{-0.11}$  is achieved [105], which is again consistent with the SM prediction within one standard deviation.

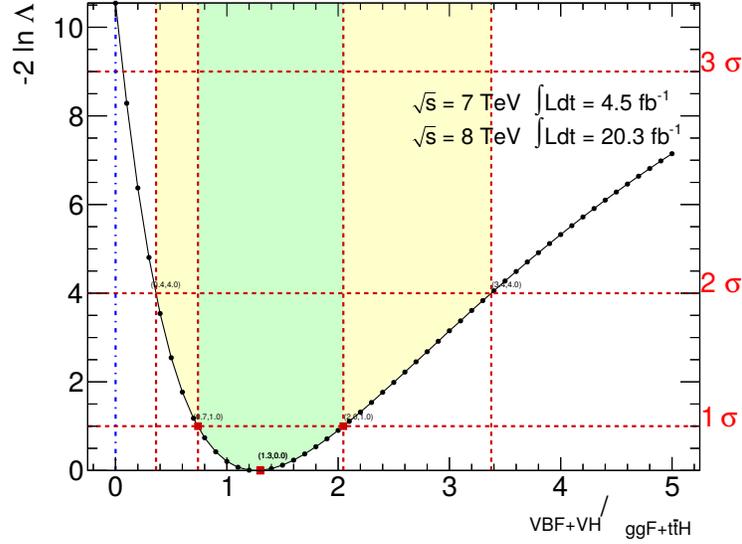


Figure 11.9: The likelihood scan as a function of the signal strength ratio  $\mu_{VBF}/\mu_{ggF}$  [105]. Horizontal lines showing significance of 1, 2 and 3 $\sigma$  are marked in red. The significance of VBF production is determined to be 3.3 $\sigma$ , read off from the plot from the intersection of the curve with the  $\mu_{VBF}/\mu_{ggF} = 0$  line.

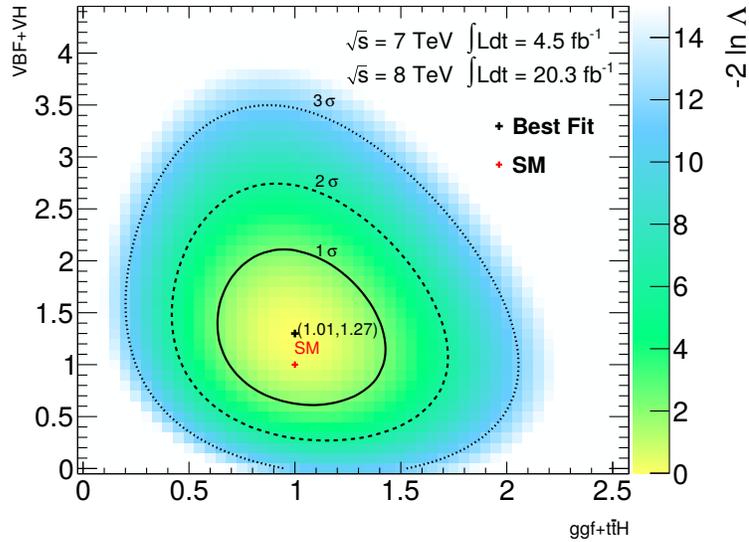


Figure 11.10: The two-dimensional likelihood scan as a function of the ggF and VBF signal strengths  $\mu_{ggF}$  and  $\mu_{VBF}$ . The 1, 2 and 3 $\sigma$  contours around the observed signal strengths are shown, along with the SM prediction in red [105].

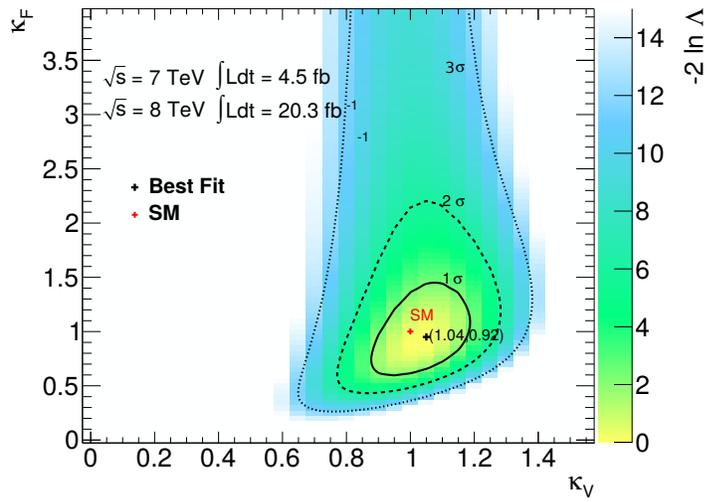


Figure 11.11: The two-dimensional likelihood scan as a function of fermionic and bosonic Higgs boson couplings,  $\kappa_F$  and  $\kappa_V$ . The 1, 2 and  $3\sigma$  contours around the best-fit value are shown, along with the SM prediction in red [105].

## 11.4 $H \rightarrow WW^{(*)}$ Combination

As for the 7 TeV and 8 TeV VBF analyses, all the final results from the  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  7 TeV and 8 TeV analyses (VBF+ggF) using the same data can be combined together to provide a cumulative result of the  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  decay mode. Within the combination of results, it is again important to take into account the correlations between the systematic uncertainties across signal and background processes, lepton flavour channels and data-taking periods [120, 121, 122]. The observation of the  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  decay mode using VBF and ggF analyses is illustrated in Figure 11.12 where the observed and expected  $p_0$  is shown as a function of  $m_H$ . Qualitatively, the observed curve agrees within one standard deviation of the expected curve for a SM Higgs boson of  $m_H = 125.36$  GeV. The corresponding significance values are marked on the right-hand vertical axis, in terms of standard deviations. Given the null (background-only) hypothesis and the alternative hypothesis that a SM Higgs boson with  $m_H = 125.36$  GeV exists, the observed minimum for  $p_0$  is found at  $m_H = 130$  GeV and is equivalent to a significance of  $6.1\sigma$ . This result surpasses the  $5\sigma$  standard used in particle physics to establish discoveries. As the significance of the VBF production is seen to be  $3.3\sigma$ , this production mode does not make up the bulk contribution to the  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  decay significance. However the observation of VBF production is important in establishing the expected properties of the SM Higgs boson due to its sensitivity to the vector boson couplings.

Although these VBF and ggF analyses are optimised for a Higgs boson of mass  $m_H = 125$  GeV, they are sensitive for SM-like Higgs bosons with masses up to 200 GeV and thus can be used to compute exclusion limits using the  $CL_s$  method [119], described briefly in Chapter 9. The exclusion at 95% confidence level is achieved for a SM Higgs boson of mass  $m_H$ , if the value of the signal strength  $\mu = 1$  is excluded at that mass. As by definition  $\mu = 1$  corresponds to the exact prediction of the SM, exclusion is defined at the point at which the results cross this boundary. The results of the exclusion study are shown in Figure 11.13, from which it can be seen that the observed exclusion range is  $132 < m_H < 200$  GeV (with an expected exclusion down to 114 GeV). Again, good agreement between the observed exclusion and that expected from the presence of a SM Higgs boson with  $m_H = 125.26$  GeV is seen. The observed curve is at least  $2\sigma$  away from the expected background-only exclusion for the considered  $m_H$  range due to the poor mass resolution of the  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  mode. The discrepancy between the observed and expected background-only exclusion is therefore not localised at a particular mass point but spread out over the  $m_H$  range.

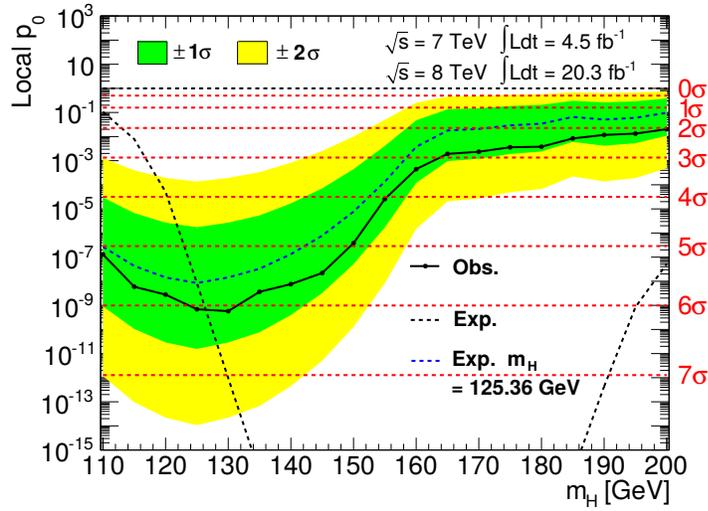


Figure 11.12: The observed (solid) and expected (dotted)  $p_0$  curves as a function of  $m_H$ . The black dotted line shows the expected  $p_0$  as a function of  $m_H$  for a background-only hypothesis, whereas the blue dotted line shows the expected  $p_0$  for a SM Higgs boson with  $m_H = 125.36$  GeV. [105]. The observed significance is  $6.1\sigma$ , with an expected significance of  $5.7\sigma$ .

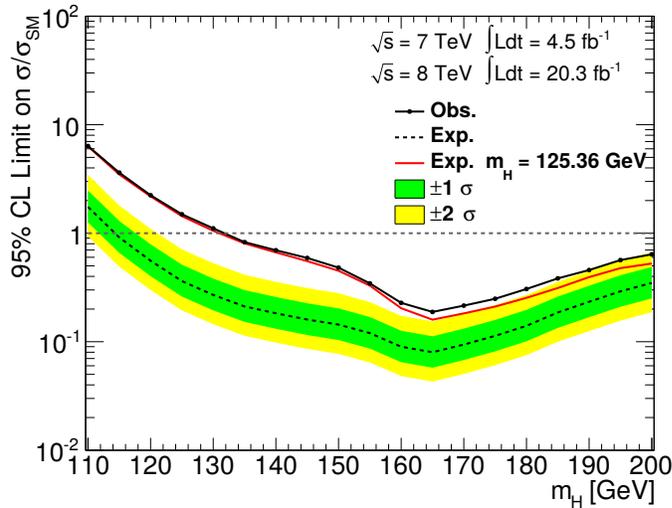


Figure 11.13: Exclusion plot produced using the  $CL_s$  method, showing the observed exclusion (black), the expected exclusion given  $m_H = 125.36$  GeV (red) and the expected background-only exclusion from the analysis (black dotted) with bands showing 1 and  $2\sigma$  deviations from this prediction [105].

# Chapter 12

## Summary

On July 4th 2012, the discovery of a new particle with the expected spin and gauge boson couplings of a Standard Model Higgs boson was announced by the ATLAS and CMS experiments at the Large Hadron Collider at CERN. Since then, all studies so far have found that the properties of this particle are compatible with that predicted by the Standard Model. This thesis presents the search for the evidence of Higgs boson production via vector boson fusion in the  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  decay mode, using data collected by the ATLAS experiment during proton-proton collisions at centre-of-mass energies of 7 TeV and 8 TeV. The data used corresponds to a total integrated luminosity of  $25 \text{ fb}^{-1}$ .

The vector boson fusion signature searched for contains two highly energetic jets, with two leptons and two neutrinos emitted from the  $W$  boson decays within the pseudorapidity range spanned by the jets. Lepton flavour combinations used are either a pair of electrons, a pair of muons, or the case where one electron and one muon are selected for. The two neutrinos are reconstructed as a large amount of missing energy in the ATLAS detector.

A multivariate boosted decision tree analysis is developed, alongside a traditional cut-based analysis which provides a solid cross-check at each stage. The use of a multivariate analysis brings a huge improvement in sensitivity of up to 33% in a particular lepton flavour combination. Regions which are expected to have a large signal presence compared to that of the background are defined, with separate regions used to check the modelling of the background processes. The estimation of the background yields in the signal-like region is crucial to the analyses and was thoroughly studied, especially with the insights gained from the cross-check analysis. Event selections and parameters used for both the multivariate and cut-based approaches have been optimised. The analysis is conducted separately for the 7 TeV and 8 TeV datasets, before recombining to achieve a final result.

Evidence for the vector boson fusion production in the  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  channel has been established, with an observed significance of 3.3 standard deviations. Assuming the existence of a Standard Model Higgs boson with a mass of 125.36 GeV,

the ratio of the measured cross section to that predicted by the Standard Model for vector boson fusion is observed to be consistent with unity:

$$\mu_{obs} = 1.28^{+0.44}_{-0.40} (stat.)^{+0.29}_{-0.20} (syst.)$$

The observation of this production mode is crucial to understanding the properties of the Standard Model Higgs boson. As the vector boson fusion production mode is sensitive to couplings of the Standard Model Higgs boson to the  $W$  and  $Z$  vector bosons, a value for the bosonic coupling has been observed to be consistent with the Standard Model prediction:

$$\kappa_V = 1.04^{+0.10}_{-0.11}$$

The techniques involved in this analysis, especially the use of boosted decision trees, set the scene well for incoming data from the restart of the Large Hadron Collider in 2015. More precision will be achieved due to the larger quantity of data, leading to the reduction of statistical uncertainties which is currently the dominant source of uncertainty for the vector boson fusion measurement. In this way, more stringent examinations of the Standard Model predictions for the Higgs boson can be studied.

Due to its distinctive kinematic signature and as one of the dominant production modes of the Standard Model Higgs boson, vector boson fusion production will continue to play a key role in future coupling measurements and searches to broaden the understanding of particle physics beyond the Standard Model.

# Bibliography

- [1] S.L. Glashow. “Partial Symmetries of Weak Interactions”. In: *Nucl.Phys.* 22 (1961), pp. 579–588.
- [2] A. Salam. “Weak and Electromagnetic Interactions”. In: *Conf.Proc.* C680519 (1968), pp. 367–377.
- [3] S. Weinberg. “A Model of Leptons”. In: *Phys.Rev.Lett.* 19 (1967), pp. 1264–1266.
- [4] D. J. Gross and F. Wilczek. “Ultraviolet Behavior of Nonabelian Gauge Theories”. In: *Phys.Rev.Lett.* 30 (1973), pp. 1343–1346.
- [5] H. D. Politzer. “Reliable Perturbative Results for Strong Interactions?” In: *Phys.Rev.Lett.* 30 (1973), pp. 1346–1349.
- [6] F. Englert and R. Brout. “Broken Symmetry and the Mass of Gauge Vector Mesons”. In: *Phys. Rev. Lett.* 13 (9 1964), 321–323.
- [7] Higgs, P. W. “Broken Symmetries and the Masses of Gauge Bosons”. In: *Phys. Rev. Lett.* 13 (16 1964), 508–509.
- [8] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. “Global Conservation Laws and Massless Particles”. In: *Phys. Rev. Lett.* 13 (20 1964), 585–587.
- [9] M. Peskin and D. Schroeder. *An Introduction to Quantum Field Theory*. Westview Press, Boulder, U.S.A, 1995.
- [10] J. Beringer et al. (Particle Data Group). “2013 Review of Particle Physics”. In: *Phys. Rev. D* 86.010001 (2012).
- [11] E. Noether. *Invariante Variationsprobleme*. Nachrichten der Koeniglichen Gesellschaft der Wissenschaften, 1918.
- [12] T. Nakano and K. Nishijima. “Charge Independence for V-particles”. In: *Prog. Theor. Phys.* 10 (1953), pp. 581–582.
- [13] M. Gell-Mann. “The interpretation of the new particles as displaced charge multiplets”. In: *Nuovo Cim.* 4.S2 (1956), pp. 848–866.
- [14] Y. Nambu. “Quasiparticles and Gauge Invariance in the Theory of Superconductivity”. In: *Phys.Rev.* 117 (1960), pp. 648–663.
- [15] J. Goldstone. “Field Theories with Superconductor Solutions”. In: *Nuovo Cim.* 19 (1961), pp. 154–164.

- [16] The Large Hadron Collider Cross Section Working Group. *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*. Tech. rep. 3. Large Hadron Collider, 2013.
- [17] 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”. In: *Physics Letters B* 716.1 (2012), pp. 1–29. ISSN: 0370-2693.
- [18] G. Aad et al. (ATLAS Collaboration). “Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC”. In: *Physics Letters B* 726.13 (2013), pp. 88–119. ISSN: 0370-2693.
- [19] G. Aad et al. (ATLAS Collaboration). “Corrigendum to Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC [Phys. Lett. B 726 (13) (2013) 88]”. In: *Physics Letters B* 734.0 (2014), pp. 406–. ISSN: 0370-2693.
- [20] G. Aad et al. (ATLAS Collaboration). “Evidence for the spin-0 nature of the Higgs boson using ATLAS data”. In: *Physics Letters B* 726.13 (2013), pp. 120–144. ISSN: 0370-2693.
- [21] S. Chatrchyan et al. (CMS Collaboration). “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”. In: *Physics Letters B* 716.1 (2012), pp. 30–61. ISSN: 0370-2693.
- [22] G. Aad et al. (ATLAS Collaboration). “Measurement of the Higgs boson mass from the  $H \rightarrow \gamma\gamma$  and  $H \rightarrow ZZ^* \rightarrow 4\ell$  channels in  $pp$  collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector”. In: *Phys. Rev. D* 90 (5 2014), 052004.
- [23] “Measurement of the properties of a Higgs boson in the four-lepton final state”. In: *Phys. Rev. D* 89 (9 2014), 092007.
- [24] ATLAS Collaboration. “Evidence for Higgs Boson Decays to the  $\tau^+\tau^-$  Final State with the ATLAS Detector”. In: ATLAS-CONF-2013-108 (2013).
- [25] S. Chatrchyan et al. “Evidence for the direct decay of the 125 GeV Higgs boson to fermions”. In: *Nature Phys.* 10 (2014).
- [26] G. Aad et al. (ATLAS Collaboration). “Measurement of the Higgs boson mass from the  $H \rightarrow \gamma\gamma$  and  $H \rightarrow ZZ^* \rightarrow 4\ell$  channels in  $pp$  collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector”. In: *Phys. Rev. D* 90.052004 (2014).
- [27] T. Aaltonen et al. “Combination of Tevatron searches for the standard model Higgs boson in the  $W^+W^-$  decay mode”. In: *Phys.Rev.Lett.* 104 (2010), p. 061802.
- [28] S. Chatrchyan et al. (CMS Collaboration). “Measurement of  $W^+W^-$  Production and Search for the Higgs Boson in  $pp$  Collisions at  $\sqrt{s} = 7$  TeV”. In: *Phys.Lett.* B699 (2011), pp. 25–47.
- [29] G. Aad et al. (ATLAS Collaboration). “Search for the Higgs boson in the  $H \rightarrow WW(*) \rightarrow \ell\nu\ell\nu$  decay channel in  $pp$  collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector”. In: *Phys.Rev.Lett.* 108 (2012), p. 111802.

- [30] S. Chatrchyan et al. (CMS Collaboration). “Search for the standard model Higgs boson decaying to a  $W$  pair in the fully leptonic final state in  $pp$  collisions at  $\sqrt{s} = 7\text{TeV}$ ”. In: *Phys.Lett.* B710 (2012), pp. 91–113.
- [31] G. Aad et al. (ATLAS Collaboration). “Search for the Standard Model Higgs boson in the  $H \rightarrow WW^0 \rightarrow \ell\nu\ell\nu$  decay mode with  $4.7\text{ fb}^1$  of ATLAS data at  $\sqrt{s} = 7\text{ TeV}$ ”. In: *Physics Letters B* 716.1 (2012), pp. 62–81.
- [32] European Organisation for Nuclear Research. *CERN Home Page*. URL: <http://www.cern.ch>.
- [33] *Status of the LHC*. Vol. 455. 2013.
- [34] G. Aad et al. (ATLAS Collaboration). “The ATLAS Experiment at the CERN Large Hadron Collider”. In: *J. Instrument* 3(08).S08003 (2008).
- [35] ATLAS Collaboration. *ATLAS Detector and Physics Performance: Technical Design Report*. Tech. rep. 1. ATLAS, 1997.
- [36] European Organisation for Nuclear Research. *ATLAS Home Page*. URL: <http://atlas.ch>.
- [37] ATLAS Collaboration. *ATLAS Inner Detector: Technical Design Report*. Tech. rep. 1. ATLAS, 1997.
- [38] G. Aad et al. (ATLAS Collaboration). “ATLAS pixel detector electronics and sensors”. In: *J. Instrument* 3(07).P07007 (2008).
- [39] A. Abdesselam et al. “The integration and engineering of the ATLAS SemiConductor Tracker Barrel”. In: *J. Instrument* 3(10).P10006 (2008).
- [40] J.R. Pater for the ATLAS Collaboration. “The ATLAS SemiConductor Tracker operation and performance”. In: *J. Instrument* 7(04).C04001 (2012).
- [41] ATLAS TRT Collaboration. “The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: design and performance”. In: *J. Instrument* 3(02).P02013 (2008).
- [42] ATLAS Collaboration. *ATLAS Liquid Argon Calorimeter: Technical Design Report*. Tech. rep. ATLAS, 1996.
- [43] ATLAS Collaboration. “Readiness of the ATLAS liquid argon calorimeter for LHC collisions”. In: *Eur. Phys. J.* C70 (2010).
- [44] ATLAS Collaboration. *ATLAS Tile Calorimeter: Technical Design Report*. Tech. rep. ATLAS, 1996.
- [45] ATLAS Collaboration. “Readiness of the ATLAS Tile Calorimeter for LHC collisions”. In: *Eur. Phys. J.* C70 (2010).
- [46] ATLAS Collaboration. *ATLAS Muon Spectrometer: Technical Design Report*. Tech. rep. ATLAS, 1997.
- [47] G. Aad et al. (ATLAS Collaboration). “Improved luminosity determination in  $pp$  collisions at  $\sqrt{s} = 7\text{ TeV}$  using the ATLAS detector at the LHC”. In: *Eur.Phys.J.* C73 (2013).

- [48] S. van der Meer. *Calibration of the effective beam height in the ISR*. Tech. rep. CERN-ISR-PO-68-31. Geneva: CERN, 1968.
- [49] C. Rubbia. *Measurement of the luminosity of  $p\bar{p}$  collider with a generalized van der Meer method*. Tech. rep. CERN- $p\bar{p}$ -Note-38. Geneva: CERN, 1977.
- [50] ATLAS Collaboration. *Public Luminosity Results*. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults>.
- [51] ATLAS Collaboration. “Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data”. In: ATLAS-CONF-2014-032 (2014).
- [52] B. Resende. “Muon identification algorithms in ATLAS”. In: ATL-PHYS-PROC-2009-11 (2009).
- [53] ATLAS Collaboration. “Muon reconstruction performance of the ATLAS detector during Run 1”. To be published in Eur. Phys. J.
- [54] *TrackInCaloTools: a package for measuring muon energy loss and calorimetric isolation in ATLAS*. Vol. 219. 2010.
- [55] M. Cacciari, G. P. Salam, and G. Soyez. “The Anti-k(t) jet clustering algorithm”. In: *JHEP* 0804 (2008), p. 063.
- [56] W Lampl et al. *Calorimeter Clustering Algorithms: Description and Performance*. Tech. rep. ATL-LARG-PUB-2008-002, ATL-COM-LARG-2008-003. Geneva: CERN, 2008.
- [57] G. Aad et al. “Jet energy measurement with the ATLAS detector in proton-proton collisions at  $\sqrt{s} = 7$  TeV”. In: *Eur.Phys.J.* C73 (2013), p. 2304.
- [58] ATLAS JetEtMiss Combined Performance Group. *Recommendations for jet cleaning for data in 2011*. ATLAS Internal Documentation. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/HowToCleanJets2011>.
- [59] 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”. In: ATLAS-CONF-2010-038 (2010).
- [60] ATLAS Collaboration. “Calibration of the performance of  $b$ -tagging for  $c$  and light-flavour jets in the 2012 ATLAS data”. In: ATLAS-CONF-2014-046 (2014).
- [61] ATLAS Collaboration. “Commissioning of the ATLAS high-performance  $b$ -tagging algorithms in the 7 TeV collision data”. In: ATLAS-CONF-2011-102 (2011).
- [62] ATLAS Collaboration. “Calibration of  $b$ -tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment”. In: ATLAS-CONF-2014-004 (2014).
- [63] J. Alison et al. “Object selections in the  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  analysis with  $20.3\text{fb}^{-1}$  of data collected with the ATLAS detector at  $\sqrt{s} = 8$  TeV”. ATLAS Internal Documentation. 2014.

- [64] P. Chang et al. “Track MET definition in the  $H \rightarrow WW^{(*)} \rightarrow l\nu l\nu$  analysis”. ATLAS Internal Documentation. 2014.
- [65] *Measurement of the missing transverse momentum based on tracks in proton-proton collisions at  $\sqrt{s} = 900$  GeV centre-of-mass energy with the ATLAS detector*. Tech. rep. ATLAS-CONF-2010-020. Geneva: CERN, 2010.
- [66] L. Breiman, J. Friedman, C. Stone and R. Olshen. *Classification and regression trees*. CRC press, 1984.
- [67] A. Hoecker et al. *TMVA 4 - Toolkit for Multivariate Data Analysis with ROOT: Users Guide*. URL: <http://tmva.sourceforge.net/>.
- [68] L. Breiman. “Bagging Predictors”. In: *Machine Learning* 24.2 (1996), 123–140. ISSN: 0885-6125.
- [69] ATLAS Collaboration. *Good Runs Lists*. ATLAS Internal Documentation. URL: <https://twiki.cern.ch/twiki/bin/view/Atlas/GoodRunsLists>.
- [70] ATLAS. *ATLAS Luminosity Calculator*. ATLAS Internal Documentation. URL: <https://atlas-lumicalc.cern.ch/>.
- [71] ATLAS Electron-Gamma Combined Performance Group. *LAr Cleaning and Object Quality*. ATLAS Internal Documentation. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/LArCleaningAndObjectQuality>.
- [72] *MuGirl – Muon identification in the ATLAS detector from the inside out*. Vol. 1. 2006.
- [73] ATLAS Collaboration. “Performance of the ATLAS Electron and Photon Trigger in p-p Collisions at  $\sqrt{s} = 7$  TeV in 2011”. In: ATLAS-CONF-2012-048 (2005).
- [74] G. Pasztor, D. Damazio, and F. Monticelli. *Electron and photon trigger performance plots using 2012 data*. Tech. rep. ATL-COM-DAQ-2014-058. Geneva: CERN, 2014.
- [75] P. Bernat. “Architecture and Performance of the Inner Detector Trigger of the ATLAS detector”. ATLAS Internal Documentation. 2012.
- [76] M. Biglietti et al. “Muon Event Filter Software for the ATLAS Experiment at LHC”. In: ATL-DAQ-CONF-2005-008 (2005).
- [77] ATLAS Collaboration. “Performance of the ATLAS Electron and Photon Trigger in p-p Collisions at  $\sqrt{s} = 7$  TeV in 2011”. In: ATLAS-CONF-2012-048 (2012).
- [78] *A measurement of the ATLAS muon reconstruction and trigger efficiency using  $J/\psi$  decays*. Tech. rep. ATLAS-CONF-2011-021. Geneva: CERN, 2011.
- [79] ATLAS Collaboration. *Muon Trigger Public Results*. ATLAS Internal Documentation. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/MuonTriggerPublicResults>.
- [80] ATLAS Collaboration. *MC11c Production Campaign details*. ATLAS Internal Documentation. URL: <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/AtlasProductionGroupMC11c>.

- [81] ATLAS Collaboration. *MC12a Production Campaign details*. ATLAS Internal Documentation. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/AtlasProductionGroupMC12a>.
- [82] ATLAS Collaboration. *MC12b Production Campaign details*. ATLAS Internal Documentation. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/AtlasProductionGroupMC12b>.
- [83] S. Agostinelli et al. “GEANT 4, A Simulation Toolkit”. In: *Nucl. Instrum. Meth.* A506 (2003), p. 250.
- [84] *Fast Simulation for ATLAS: Atlfast-II and ISF*. Vol. 396. 022031. 2012.
- [85] H.-L. Lai et al. “New parton distributions for collider physics”. In: *Phys. Rev. D* 82 (2010).
- [86] P. Nason. “A New method for combining NLO QCD with shower Monte Carlo algorithms”. In: *Journal of High Energy Physics* 0411 (2004).
- [87] J. Pumplin et al. “New generation of parton distributions with uncertainties from global QCD analysis”. In: *Journal of High Energy Physics* 0207 (2002).
- [88] M. L. Mangano et al. “ALPGEN, a generator for hard multi-parton processes in hadronic collisions”. In: *Journal of High Energy Physics* 0307 (2003).
- [89] B. P. Kersevan and E. Richter-Was. “The Monte Carlo event generator AcerMC version 2.0 with interfaces to PYTHIA 6.2 and HERWIG 6.5”. In: *arXiv:hep-ph/0405247* (2004).
- [90] T. Sjostrand, S. Mrenna, and P. Z. Skands. “PYTHIA 6.4 physics and manual”. In: *Journal of High Energy Physics* 0605 (2006).
- [91] T. Sjostrand, S. Mrenna, and P. Z. Skands. “A Brief Introduction to PYTHIA 8.1”. In: *CERN-LCGAPP* 04 (2007).
- [92] A. Sherstnev and R. S. Thorne. “Parton distributions for the LHC”. In: *Eur. Phys. J* C55 (2008).
- [93] J. Alison et al. “Background estimation techniques in the  $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$  analysis with  $20.3\text{fb}^{-1}$  of data collected with the ATLAS detector at  $\sqrt{s} = 8\text{TeV}$ ”. ATLAS Internal Documentation. 2014.
- [94] *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*. ATL-PHYS-PUB-2010-013. 2010.
- [95] ATLAS Collaboration. *ATLAS Pileup Reweighting*. ATLAS Internal Documentation. URL: <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/PileupReweighting>.
- [96] J. Alison et al. “Analysis of  $H \rightarrow WW^{(0)} \rightarrow \ell\nu\ell\nu$  ggF and VBF production modes with  $20\text{fb}^{-1}$  and  $4.5\text{fb}^{-1}$  of data collected with the ATLAS detector at  $\sqrt{s} = 8$  and  $7\text{TeV}$ ”. ATLAS Internal Documentation. 2014.
- [97] A. Djouadi, J. Kalinowski, and M. Spira. “HDECAY: a Program for Higgs Boson Decays in the Standard Model and its Supersymmetric Extension”. In: *Comput.Phys.Commun.* 108 (1998), pp. 56–74.

- [98] P. Bolzoni et al. “Higgs Boson Production via Vector-Boson Fusion at Next-to-Next-to-Leading Order in QCD”. In: *Phys. Rev. Lett.* 105 (1 2010), 011801.
- [99] M. Ciccolini, A. Denner and S. Dittmaier. “Electroweak and QCD corrections to Higgs production via vector-boson fusion at the CERN LHC”. In: *Phys. Rev. D* 77 (1 2008), 013002.
- [100] C. Anastasiou and K. Melnikov. “Higgs boson production at hadron colliders in NNLO QCD”. In: *Nuclear Physics B* 646.12 (2002), pp. 220–256. ISSN: 0550-3213.
- [101] S. Catani et al. “Soft gluon resummation for Higgs boson production at hadron colliders”. In: *JHEP* 0307 (2003), p. 028.
- [102] U. Aglietti et al. “Master integrals for the two-loop light fermion contributions to  $gg \rightarrow H$  and  $H \rightarrow \gamma \gamma$ ”. In: *Phys.Lett.* B600 (2004), pp. 57–64.
- [103] D. de Florian and M. Grazzini. “Higgs production at the LHC: updated cross sections at  $\sqrt{s} = 8$  TeV”. In: *Phys.Lett.* B718 (2012), pp. 117–120.
- [104] K. Hamilton, P. Nason, and G. Zanderighi. “MINLO: multi-scale improved NLO”. In: *Journal of High Energy Physics* 2012.10, 155 (2012).
- [105] ATLAS Collaboration. “Observation and measurement of the Higgs boson decays to  $WW$  with ATLAS at the LHC”. Geneva, 2013.
- [106] S. Dittmaier M. L. Ciccolini and M. Kramer. “Electroweak radiative corrections to associated  $WH$  and  $ZH$  production at hadron colliders”. In: *Phys. Rev.* D68 (2003), p. 073003.
- [107] T. Han and S. Willenbrock. “QCD correction to the  $p p \rightarrow W H$  and  $Z H$  total cross- sections”. In: *Phys. Lett.* B273 (1991), pp. 167–172.
- [108] O. Brein, A. Djouadi, and R. Harlander. “NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders”. In: *Phys. Lett.* B579 (2004), p. 149.
- [109] G. Corcella et al. “HERWIG 6: an event generator for hadron emission reactions with interfering gluons (including super-symmetric processes)”. In: *JHEP* 0101 (2001).
- [110] R. Ellis et al. “Higgs Decay to  $\tau^+ \tau^-$ : A Possible Signature of Intermediate Mass Higgs Bosons at the SSC”. In: *Nucl. Phys. B* 297 (1988).
- [111] O. Arnaez et al. (ATLAS  $H \rightarrow WW$  group). *A  $H$  to  $WW$  reanalysis of the 4.6 of  $\sqrt{s} = 7$  TeV data collected with the ATLAS detector in 2011*. Tech. rep. ATL-COM-PHYS-2013-129. Geneva: CERN, 2013.
- [112] J. Alison et al. “Theoretical studies for the  $H \rightarrow WW$  search and measurement”. ATLAS Internal Documentation. 2014.
- [113] R. Ball et al. “Parton distributions with LHC data”. In: *Nuclear Physics B* 867.2 (2013), pp. 244–289. ISSN: 0550-3213.
- [114] C. Lester and D. Summers. “Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders”. In: *Phys. Lett. B* 463 (1999).

- [115] ATLAS Collaboration. “Top quark mass measurement in the  $e\mu$  channel using the mT2 variable at ATLAS”. In: ATLAS-CONF-2012-082 (2012).
- [116] J. Alwall et al. “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”. In: *JHEP* 1407 (2014).
- [117] A.D. Martin et al. “Parton distributions for the LHC”. In: *European Physical Journal C* 63.2 (2009), 189–285. ISSN: 1434-6044.
- [118] G. Cowan, K. Cranmer, E. Gross and O. Vitells. “Asymptotic formulae for likelihood-based tests of new physics”. In: *Eur. Phys. J. C* 71 (2011).
- [119] A. L. Read. “Presentation of search results: the CLs technique”. In: *J. Phys. G* 28 (2012).
- [120] S. Dittmaier et al. “Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables”. In: (2011).
- [121] *Procedure for the LHC Higgs boson search combination in summer 2011*. Tech. rep. ATL-PHYS-PUB-2011-011. Geneva: CERN, 2011.
- [122] R. Ellis et al. “Jet energy measurement and its systematic uncertainty in proton-proton collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector”. In: *Eur. Phys. J. C* (2014).
- [123] G. Aad et al. “Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data”. In: *Eur.Phys.J. C* 72 (2012), p. 1909.
- [124] *Muon Momentum Resolution in First Pass Reconstruction of pp Collision Data Recorded by ATLAS in 2010*. Tech. rep. ATLAS-CONF-2011-046. Geneva: CERN, 2011.
- [125] G. Aad et al. “Jet energy resolution in proton-proton collisions at  $\sqrt{s} = 7$  TeV recorded in 2010 with the ATLAS detector”. In: *Eur.Phys.J. C* 73 (2013), p. 2306.
- [126] *Jet energy resolution and selection efficiency relative to track jets from in-situ techniques with the ATLAS Detector Using Proton-Proton Collisions at a Center of Mass Energy  $\sqrt{s} = 7$  TeV*. Tech. rep. ATLAS-CONF-2010-054. Geneva: CERN, 2010.
- [127] *Measurement of the b-tag Efficiency in a Sample of Jets Containing Muons with 5 fb<sup>-1</sup> of Data from the ATLAS Detector*. Tech. rep. ATLAS-CONF-2012-043. Geneva: CERN, 2012.
- [128] *Pile-up subtraction and suppression for jets in ATLAS*. Tech. rep. ATLAS-CONF-2013-083. Geneva: CERN, 2013.
- [129] I. W. Stewart and F. J. Tackmann. “Theory Uncertainties for Higgs and Other Searches Using Jet Bins”. In: *Phys.Rev.* D85 (2012), p. 034011.
- [130] J. M. Campbell, R. K. Ellis, and C. Williams. “Vector boson pair production at the LHC”. In: *JHEP* 1107 (2011), p. 018.

- [131] S. Heinemeyer et al. “Handbook of LHC Higgs Cross Sections: 3. Higgs Properties”. In: (2013). Ed. by S Heinemeyer.



# Appendix A

## 2011 Data Periods

Data Period	Date Range
A	March 13th-21st
B	March 21st-24th
C	–
D	April 14th-29th
E	April 30th-May 3rd
F	May 15th-25th
G	May 27th-June 14th
H	June 16th-28th
I	July 13th-29th
J	July 30th-Aug. 4th
K	Aug. 4th-22nd
L	Sept. 7th-Oct. 5th
M	Oct. 6th-30th

Table A.1: Naming of the 2011 data-taking periods, for proton-proton collisions operating at the LHC. No usable data was taken in Period C.



# Appendix B

## Cut-based yields for the 7 TeV and 8 TeV analyses

Cuts	VBF+VH	Total Bkg.	Data	Data/MC
$\cancel{E}_T > 55$ GeV				
$\cancel{E}_T^{jet,Track} > 50$ GeV	107.60±1.60	130259.32±112.94	135734	1.04
$N_{jets} \geq 2$	57.61±1.03	61162.87±41.59	61434	1.00
$b$ -Veto	41.95±0.86	7717.33±25.69	7818	1.01
$p_T^{tot,jet,Track} < 15$ GeV	35.43±0.76	5626.16±21.86	5787	1.03
$m_{\tau\tau} < 66.1876$	29.54±0.67	2970.26±12.50	3129	1.05
$m_{jj} > 600$ GeV	8.24±0.12	99.50±1.88	131	1.32
$\Delta Y_{jj} > 3.6$	7.92±0.12	80.12±1.75	107	1.34
CJV	6.57±0.11	43.00±1.28	58	1.35
OLV	6.40±0.11	36.15±1.20	51	1.41
$\Delta\phi_{\ell\ell} < 1.8$	5.95±0.10	20.87±0.88	34	1.63
$m_{\ell\ell} < 50$ GeV	7.57±0.10	10.60±0.71	19	1.79
$m_T < 130$ GeV				
$m_T > 80$ GeV for low $p_T$	4.72±0.09	5.56±0.47	14	2.52
$m_{jj} < 1$ TeV	2.33±0.06	4.05±0.43	8	1.97
$m_{jj} > 1$ TeV	2.39±0.06	1.57±0.19	6	3.81

Table B.1: Summary of total expected signal and background, and data yields at each selection stage after the common preselection, for the  $e\mu+\mu e$  channel within the 8 TeV cut-based analysis. Only statistical uncertainties are given here. The last column shows the ratio of data to total MC simulation.

Cuts	VBF+VH	Total Bkg.	Data	Data/MC
$\cancel{E}_T > 55$ GeV				
$E_T^{jet,Track} > 50$ GeV	39.89±0.95	35358.15±57.75	36173	1.02
$N_{jets} \geq 2$	24.17±0.62	627220.18±36.88	26949	0.99
$b$ -Veto	17.58±0.49	3105.85±25.99	3101	1.00
$p_T^{tot,jet,Track} < 15$ GeV	15.26±0.45	2193.53±23.30	2170	0.99
$m_{\tau\tau} < 66.1876$	11.99±0.39	1319.94±12.22	1344	1.02
$m_{jj} > 600$ GeV	3.78±0.08	70.18±4.25	81	1.15
$\Delta Y_{jj} > 3.6$	3.54±0.08	50.11±4.15	53	1.06
CJV	3.01±0.08	22.95±0.91	31	1.35
OLV	2.91±0.08	18.80±0.83	26	1.38
$\Delta\phi_{\ell\ell} < 1.8$	2.65±0.07	12.08±0.64	15	1.24
$m_{\ell\ell} < 50$ GeV	2.40±0.07	6.85±0.48	9	1.31
$m_T < 130$ GeV				
$m_T > 80$ GeV for low $p_T$	2.18±0.07	3.66±0.40	6	1.64
$m_{jj} < 1$ TeV	1.07±0.05	2.40±0.31	3	1.25
$m_{jj} > 1$ TeV	1.11±0.04	1.30±0.26	3	2.31

Table B.2: Summary of total expected signal and background, and data yields at each selection stage after the common preselection, for the  $ee+\mu\mu$  channel within the 8 TeV cut-based analysis. Only statistical uncertainties are given here. The last column shows the ratio of data to total MC simulation.

Cuts	VBF+VH	Total Bkg.	Data	Data/MC
$\cancel{E}_T > 45$ GeV	13.05±0.13	20732.77±59.05	21109	1.02
$N_{jets} \geq 2$	6.65±0.09	7861.82±20.55	8042	1.02
$b$ -Veto	4.69±0.07	944.29±9.39	949	1.00
$p_T^{tot,jet,Track} < 15$ GeV	3.76±0.06	602.23±7.71	619	1.03
$m_{\tau\tau} < 66.1876$	3.18±0.05	308.68±5.03	310	1.00
$m_{jj} > 500$ GeV	1.14±0.02	13.32±0.74	13	0.98
$\Delta Y_{jj} > 2.8$	1.12±0.02	10.07±0.55	11	1.09
CJV	1.04±0.02	7.22±0.49	8	1.11
OLV	1.00±0.02	5.93±0.46	6	1.01
$\Delta\phi_{\ell\ell} < 1.8$	0.90±0.02	3.51±0.35	4	1.14
$m_{\ell\ell} < 50$ GeV	0.77±0.02	1.56±0.23	2	1.28
$m_T < 130$ GeV	0.71±0.01	0.70±0.12	0	–

Table B.3: Summary of total expected signal and background, and data yields at each selection stage after the common preselection, for the  $e\mu+\mu e$  channel within the 7 TeV cut-based analysis. Only statistical uncertainties are given here. The last column shows the ratio of data to total MC simulation.

<b>Cuts</b>	<b>VBF+VH</b>	<b>Total Bkg.</b>	<b>Data</b>	<b>Data/MC</b>
$\cancel{E}_T > 45$ GeV	7.67±0.10	11086.19±54.56	11676	1.05
$N_{jets} \geq 2$	4.24±0.07	5022.40±18.34	5056	1.01
$b$ -Veto	2.95±0.05	845.93±10.69	860	1.02
$p_T^{tot,jet,Track} < 15$ GeV	2.36±0.05	542.18±8.81	535	0.99
$m_{\tau\tau} < 66.187$	1.94±0.04	324.45±6.58	327	1.01
$m_{jj} > 500$ GeV	0.74±0.02	18.17±1.18	15	0.83
$\Delta Y_{jj} > 2.8$	0.72±0.02	12.28±0.92	11	0.90
CJV	0.67±0.01	7.75±0.77	5	0.65
OLV	0.64±0.01	5.83±0.56	5	0.86
$\Delta\phi_{\ell\ell} < 1.8$	0.57±0.01	3.31±0.42	4	1.21
$m_{\ell\ell} < 50$ GeV	0.51±0.01	1.59±0.30	4	2.52
$m_T < 130$ GeV	0.47±0.01	1.02±0.24	4	3.91

Table B.4: Summary of total expected signal and background, and data yields at each selection stage after the common preselection, for the  $ee+\mu\mu$  channel within the 7 TeV cut-based analysis. Only statistical uncertainties are given here. The last column shows the ratio of data to total MC simulation.

# Acknowledgements

*“Chance encounters are what keep us going.”  
- Haruki Murakami, Kafka on the Shore*

Firstly, an enormous thank you to Prof. Dr. Dorothee Schaile for providing me with the opportunity to write my thesis and the support during my time here in the group. Her “open-door” policy for all manner of questions is something which I personally appreciate very much. In addition, her generosity has been amazing in allowing a seven month stay at CERN and various summer schools and workshops.

Many thanks to PD Dr. Johannes Elmsheuser for supervising this thesis, including the last-minute proof reading. Thank you for providing useful suggestions and for the help with computing issues during my studies.

I would also like to thank Prof. Dr. Wolfgang Dünneweber for agreeing to write the second review of my thesis, and to Prof. Dr. Gerhard Buchalla and Prof. Dr. Andreas Burkert for participating in my thesis defense.

Thank you very much to Prof. Dr. Otmar Biebel, for some intriguing discussions and his support as the head of the “Particle Physics at the Energy Frontier of New Phenomena” postgraduate program, which supported me for two years.

To all my officemates who have had the (mis)fortune to share an office with me: many thanks go to Christian Meineck, who answers so many of my wumbo questions and is also a great travel companion whether in Geneva or Peru, and Friedrich Hönig, the provider of hazelnuts, Fun Facts and Formal Fridays. To Philippe Calfayan with whom I shared an office at CERN, thank you for all the advice, green ball conversations and shared enjoyment of a few particular songs.

Thank you to all my colleagues in Munich for providing such a fun atmosphere at the Lehrstuhl and also help when needed. All the laughter, table-tennis and the ‘helpful’ German phrases I learnt at the beginning are very dear to me. There are too many to name, but special thanks go to Stefanie Adomeit, Johannes Ebke, Nikolai Hartmann, Florian Rättich, Jonas Will and Josephine Wittkowski.

I would like to thank both Herta Franz and Elke Grimm-Zeidler, for helping me through the mountains of administrative issues and paperwork.

To Angelika, the identical particle with whom I share so many of my thoughts and hopes - thank you for always being there, no matter where in the world I happen to be.

It has been quite a few years since we started at the same school on a quiet hill and we have since then migrated in all directions. How can I compress so many years and experiences into a few sentences? So just simply, thank you to Gib, Frances, Joyce, Eva and Ploy, with whom I grew up with and who continue to support me.

Living in Munich for the past three years has been amazing, largely due to the extended family here. From Carcassone addictions, to ritual weekend dinners - thank you for all the friendships and adventures, and for those to come.

Grateful thanks go towards my parents and brother for supporting me through my studies and their understanding during the most stressful times. My greatest thanks go to Mathias, for his unwavering encouragement and for being a source of strength from the first moment.