The Search for a Heavy Higgs Boson in the $H \to WW \to l\nu jj$ Channel with the ATLAS Detector

by

Ossama S. AbouZeid

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

Abstract

The Search for a Heavy Higgs Boson in the $H \to WW \to l\nu jj$ Channel with the ATLAS Detector

Ossama S. AbouZeid
Doctor of Philosophy
Graduate Department of Physics
University of Toronto
2015

A search for a heavy Higgs boson has been performed in the $H \to WW \to \ell\nu jj$ decay channel using the ATLAS detector, with 20.3 fb^{-1} of data, recorded from proton – proton collisions at the Large Hadron Collider at $\sqrt{s} = 8$ TeV. The search is performed in two orthogonal search modes, one targeted at gluon-gluon fusion production (ggF) and the other at vector boson fusion (VBF) production. The data are tested against multiple signal hypothesis, one large width Standard Model-like Higgs, one assuming an electroweak singlet (EWS) extension to the SM Higgs sector, and finally, one model independent Higgs, which is generated with the Narrow Width Approximation – consistent with some parameters of EWS and Two Higgs Doublet Models (2HDM). In all cases, no significant excess in the data is found when compared to the background only (SM) hypothesis. Instead, upper limits on $\sigma \times$ B.R. at the 95% confidence level (using the CL_s method) are presented for the three heavy Higgs models studied.

Dedication

For my family: Ossama, Edith, Anna, and Maria.

Acknowledgements

The first thank-you must of course go to Peter Krieger, who I have thoroughly enjoyed working with, learned a lot from, and who has put up with my continual harassment with grace for too many years.

A special mention to the graduate students and post-docs at the University of Toronto that I have worked with, partied with, and complained with; Travis, Trisha, Bin, Sing, Dominque, Pawan, Rey, Bertrand, Garrin, Joe, Steve, Nikolina, Sami, Behi, Pete, Justin, Nicola, Richard, David, Santiago, Chav, Vince, Miriam, and Aaron.

I must also mention my extended academic family at the Carleton University physics department, some of whom I worked with on ATLAS, and others who I worked with in my undergraduate, and all of whom encouraged and helped me at every chance. Thank-you Bruce Campbell, Manuella Vincter, Alain Bellerive, Cliff Hargrove, and Kevin Graham.

The Liquid Argon group on ATLAS is a group of physicists that I hope I will have the chance to collaborate with again in the future. I enjoyed my time working in the LAr group an immense amount, and the amount I learned while under the direction of the three "chefs" has been irreplaceable. Merci à Isabelle, Emmanuel, et Luis.

The years I have spent in the HSG3 $\ell\nu jj$ group have been a real pleasure, and I owe a lot of that to the team of excellent physicists I worked with. Rikard, Mark, Ben, Mike, Corrinne, Olivier, and Haifeng; thank-you.

The most important thank-you goes to my family. Thank-you for your support.

And to Maria, who made this entire trip worthwhile.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

> "We don't want to conquer the cosmos, we simply want to extend the boundaries of Earth to the frontiers of the cosmos."
>
> Stanislaw Lem, *Solaris*

Much of the press coverage of the Large Hadron Collider (LHC) has been in the context of the search for the Standard Model (SM) Higgs boson. While this overlooks the breadth of the experimental program at the LHC and its associated experiments, it is natural to have a particular focus on the Higgs; the particle has been searched for and studied since its original postulation in the 1960s [1, 2, 3, 4, 5, 6]. With the highest energy proton–proton collisions in the world being produced by the LHC, combined with the already extensive restrictions on the Higgs mass from experiments at LEP (the Large Electron-Positron Collider) [7, 8] and the Tevatron [9], as well as theoretical constraints [10], the Standard Model Higgs – if it existed – was expected to be discovered by the experiments at the LHC.

On July 4^{th}, 2012, the ATLAS and CMS Collaborations announced the discovery of a Higgs boson [11, 12]. The discovery validates decades of work to understand the fundamental building blocks and interactions of our universe. The theory that collectively encompasses these building blocks and their interactions (excluding gravity) is called the Standard Model (for more discussion on the Standard Model, refer to Chapter 2). The presence of the Higgs (or something like it) is needed by the Standard Model in order to provide the spontaneous symmetry breaking (Section 2.2) which generates mass for the fundamental particles. Without it, the Standard Model produces unphysical results, with some cross-sections for electroweak processes (such as WW scattering) breaking unitary (*i.e.*, give a probability of a process occurring > 1) [10]. While a Higgs boson has been found, the precise nature of the Higgs sector is not entirely known, *i.e.*, is the Higgs discovered at 125 GeV *a* Higgs, or *the* Higgs. The discovery, while providing a first glimpse of the Higgs sector, does not indicate whether the SM prediction is the correct one. There is only one Higgs predicted by the SM, with very specific properties, and so ATLAS and CMS experiments have been (and still are) performing measurements to confirm the properties of Higgs.

While precision tests and measurements of the Higgs boson are one way to examine the Higgs sector, a more direct way to prove that this is not the SM Higgs boson, is to discover another (perhaps heavier) Higgs boson, which is part of an extended Higgs sector, not predicted of the SM. Therefore, in parallel

to precision tests of Higgs properties, studies by both CMS and ATLAS Collaborations are underway to search for an additional Higgs.

One such search, looking for a heavy Higgs in the $H \to WW \to \ell\nu jj$ channel, is presented in this thesis. The search was originally a Standard Model Higgs search, but one that was initially somewhat of a black-sheep in the family. The process is most sensitive in the higher range of Higgs masses ($m_H \gtrsim 300$ GeV), while the fit on the Standard Model parameters [7] indicated that the Standard Model prefers a light Higgs. Therefore, before the discovery was made, the search channel was sensitive in a region of Higgs mass "not preferred" by the Standard Model. With the discovery of a relatively low mass Higgs (in good agreement with results of the electroweak fit), the Higgs mass is in a region not accessible to this analysis due to the high number of background events. While the prospects for an observation, or precise property measurement of the 125 GeV Higgs in this search channel are not good, the search for an extended Higgs sector in the high mass region (where this analysis is most sensitive) provides an excellent motivation to study this final state.

While most interesting physics searches and measurements have a theoretical motivation, this channel was personally motivating because the final state itself is *interesting*; it requires a detailed knowledge of the detector and the object reconstruction, as the reconstruction of the final state objects (jets, leptons and neutrinos) relies on information from all portions of the detector. Added to that is the potential for the two jets to become merged, in the case of a very heavy Higgs, and the corresponding need for the use of jet substructure techniques.

1.1 Results from 2011 Dataset

The most recent result published by ATLAS in the $H \to WW \to \ell\nu jj$ channel was in 2012, using data collected during the 2011 pp run [13]. This was before the discovery in July 2012, so that search was for a SM-like Higgs at high mass. Compared to the 2012 dataset, the 2011 dataset was at a lower center of mass energy[1] of $\sqrt{s} = 7$ TeV, and an integrated luminosity of 4.7 fb^{-1}.

The limit from that analysis is presented in Figure 1.1. The limited search range ($300 \leq m_H \leq 600$ GeV) was due to the fact that no treatment was yet in place to account for the lineshape of a heavy Higgs (which will be discussed in Chapter 2), and no mechanism in place for evaluation of the interference of the $H \to WW \to \ell\nu jj$ process with the non-resonant WW continuum background. As a result, there were large uncertainties placed on the theoretical cross-section ($\approx 1.5 \times (m_H/\text{ TeV})^3$), which became prohibitive at large masses ($\gtrsim 30$ % at $m_H = 600$ GeV). As a result, both CMS and ATLAS heavy Higgs searches only set cross-section limits up to $m_H = 600$ GeV [14] using the 2011 data.

1.2 The 2012 Search for a Heavy Higgs Boson

The 2012 LHC proton–proton run brought with it promising changes and exciting developments. In 2012 the LHC was able to quadruple the 2011 pp dataset delivered, and at a higher center of mass energy of $\sqrt{s} = 8$ TeV (See Chapters 3 and 4 for discussions on the LHC and the delivered dataset). Further, the theory community provided more accurate lineshape predictions for a heavy Higgs, and Monte Carlo generators and techniques able to quantify the interference between $H \to WW \to \ell\nu jj$ and the

[1] In both 2011 and 2012, the energy was reduced from the LHC design energy of $\sqrt{s} = 14$ TeV. The reduction of the collision energy was due to catastrophic damage suffered during the initial commissioning of the LHC in 2008.

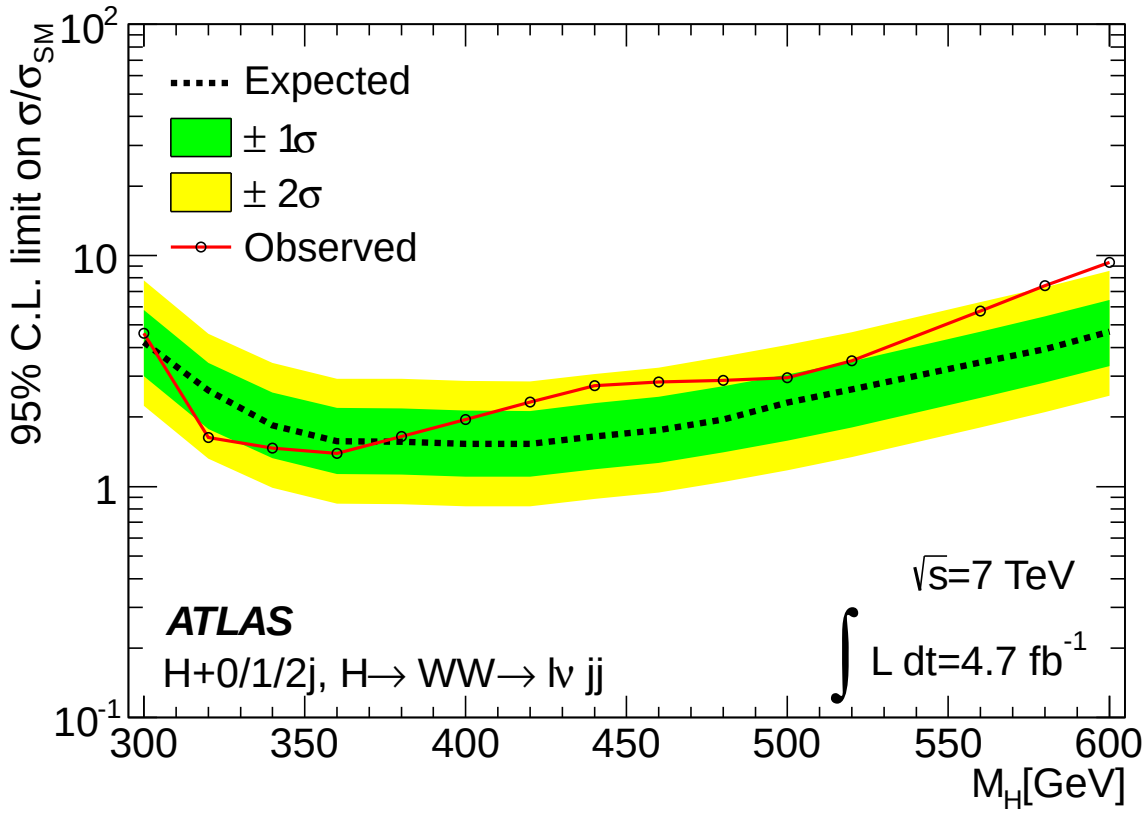


Figure 1.1: Expected and observed upper limits on the Standard Model like Higgs boson $\sigma \times$ BR (at 95% CL) using the 2011 pp dataset in the $H \to WW \to \ell\nu jj$ channel. The upper limit is divided by the SM prediction, σ_{SM}. The dotted black line shows the *expected* upper cross-section limit, assuming the (null) background only hypothesis, with the coloured bands indicating the uncertainty on the expected limits with $\pm 1\sigma$ and $\pm 2\sigma$. The red line shows the *observed* upper limits as determined from data [13].

continuum WW background (Chapter 2), allowing for the search across a broad range of Higgs masses without (prohibitively) large theoretical uncertainties. These factors, in addition to the larger 2012 dataset and experience gained from the 2011 run, result in the 8 TeV 2012 dataset providing a superior experimental environment to search for the production of heavy Higgs-like resonances, as discussed in Chapter 5. Of course, the discovery of what appears to be a SM-like Higgs boson at 125 GeV means that this analysis cannot simply perform a search for an additional Higgs which has properties consistent with SM expectation. Instead, a suite of theories that predict an extended Higgs sector must be examined. These are briefly discussed in the next chapter.

The results presented in this thesis do not close the chapter on the Higgs sector, or even on the $H \to WW \to \ell\nu jj$ search. The LHC Run-II (starting as this is being written) aims to provide much more data, and a higher center of mass energy than was available from LHC Run-I, presenting the possibility to perform an even more sensitive search for an extended Higgs sector. Before the ink has begun to dry, I look forward to LHC Run-II.

Chapter 2

The Standard Model and SM-like Higgs Bosons

"Six years ago they began to get invisible, glass or no glass. Nobody has ever seen the last five I made because no glass is strong enough to make them big enough to be regarded truly as the smallest things ever made."

Flann O'Brien, *The Third Policeman*.

2.1 The Standard Model of Particle Physics

The Standard Model of particle physics [15, 16, 17] is a fundamental theory of matter, and its interactions. It mathematically describes the basic building blocks of the universe and the three fundamental forces of nature through which they interact; the strong force, weak force, and electromagnetic force. A description of gravity is noticeably absent from the Standard Model, although at the relevant energy and length scales, the contributions from gravity are negligible.

The SM consists of twelve fundamental spin-$\frac{1}{2}$ particles (fermions) and four force mediating spin-1 particles (the gauge bosons)[1]. The fermions are what make up everyday matter, while the spin-1 gauge bosons mediate the three fundamental forces/interactions between the particles. Fermions can be divided into two categories; the leptons and the quarks, each consisting of six spin-$\frac{1}{2}$ particles. These are often grouped into three sets of pairs, or *generations*.

For the leptons, the three generations are the electron (e), muon (μ) and tau (τ), and their associated neutrinos (ν_e, ν_μ, ν_τ) grouped into generations, as;

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix} \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix} \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}. \tag{2.1}$$

There are also six distinct types, or *flavours* of quarks: up, down, charm, strange, top and bottom

[1]Spin is a dimensionless quantity when using *natural units* ($\hbar = c = 1$), which are used throughout this thesis.

Table 2.1: Summary of fundamental fermion properties. Associated anti-particles have the same mass, but opposite sign charge. Electric charge is reported in units of the elementary electric charge, e. Masses for the u, d, and s are evaluated in a mass-independent subtraction scheme ($\overline{\text{MS}}$), while the c, and b quark masses are evaluated with "running" masses in the $\overline{\text{MS}}$ scheme. The quoted top quark mass was determined by direct mass measurements [18].

	The Leptons			**The Quarks**		
	Lepton	Mass [MeV]	Electric charge	**Quark**	Mass [MeV]	Electric charge
Gen. I	e	0.511	-1	u	2.3	$+2/3$
	ν_e	$< 2 \times 10^{-3}$	0	d	4.8	$-1/3$
Gen. II	μ	105.66	-1	c	1275	$+2/3$
	ν_μ	$< 2 \times 10^{-3}$	0	s	95	$-1/3$
Gen. III	τ	1776.82	-1	t	173.21×10^3	$+2/3$
	ν_τ	$< 2 \times 10^{-3}$	0	b	4.18×10^3	$-1/3$

(u, d, c, s, t, b). Each quark is either an *up* type or *down* type quark, with the up types having an electric charge of $+2/3$, and the down types having $-1/3$. The quarks, like the leptons, are organized in three generations, each having one *up* type and one *down* type quark:

$$\begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix}. \tag{2.2}$$

Each fermion has an associated anti-particle, which is identical to its partner, but carries opposite charge. The anti-particles are typically differentiated from their partners with a bar, *e.g.*, an up quark, u, and it's anti-particle, $\bar{u}$. A summary of the fermion properties is found in Table 2.1.

The Standard Model also describes four spin-1 bosons (the gauge bosons), which act as the mediators for the three fundamental forces that are described by the Standard Model. They are the massless photon, γ (mediator for the electromagnetic force), the W and Z bosons (mediators of the weak force), and the gluon, g (massless mediator of the strong force). A summary of the gauge boson properties can be found in Table 2.2.

The Standard Model is a $SU_C(3) \times SU_L(2) \times U_Y(1)$ locally gauge invariant theory. The strong interaction, described by QCD (quantum chromodynamics), obeys $SU_C(3)$ transformations ("C" indicating colour, the charge of the strong force), and contains eight gauge bosons (gluons) which mediate the strong force. Weak and electromagnetic interactions exhibit invariance under $SU_L(2) \times U_Y(1)$ transfor-

Table 2.2: Summary of gauge boson properties. Electric charge reported in units of the elementary electric charge, e [18]. The masses for the W and Z bosons are from experimental results, while the gluon and photon masses are from theoretical prediction (experimental results are consistent with a massless gluon and photon).

	Mass [GeV]	Electric charge	Mediator of
γ	0	0	Electromagnetic force
g	0	0	Strong force
$W^\pm$	80.385	± 1	Weak force
Z^0	91.1876	0	Weak force

mations ("Y" for weak hypercharge and "L" indicating that only interactions with left-handed fermions are allowed). The $SU_L(2)$ symmetry has three group generators, correspond to three gauge bosons, while $U_Y(1)$ has a single generator. The four generators of the $SU_L(2) \times U_Y(1)$ group are related to the four physical SM gauge bosons (Z, $W^{\pm}$, γ) which arise from the spontaneous symmetry breaking in the Higgs sector (detailed in the following sections), and accounts for their non-zero mass.

2.2 The Higgs Mechanism

Although the Standard Model has been extremely successful, making experimentally verified predictions, and holding up to precision measurements, there was one piece missing for many years: the Higgs boson. The Higgs boson (and it's associated field) solves the problem that arises from a simple observation; fundamental particles have mass, but the mathematical framework of the SM does not allow the Lagrangian to have explicit mass terms for the gauge bosons. In the case that the fermion and boson fields in the Standard Model are kept massless, the Lagrangian is locally gauge invariant. Adding mass terms for the quarks and leptons when considering only the strong force is trivial, and does not break the local gauge $SU_C(3)$ invariance which describes the strong force [10]. However, attempting to add a mass term for the electroweak bosons of the form $\frac{1}{2}M_V^2 W_\mu W^\mu$ breaks the local $SU_L(2) \times U_Y(1)$ gauge symmetry. Adding explicit fermion mass terms results in similar behaviour, with the local gauge symmetry being broken. One must then look for a way to generate the fermion and relevant boson masses without breaking the local $SU_L(2) \times U_Y(1)$ symmetry. The way to do this is through the mechanism of spontaneous symmetry breaking [1, 2, 3, 4, 5, 6].

2.2.1 Spontaneous Symmetry Breaking in a Toy Model

As an illustrative example [10, 19], consider the following simple Lagrangian, which describes a real scalar field ϕ, with potential $V(\phi)$:

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - V(\phi), \quad V(\phi) = \frac{1}{2}\mu^2\phi^2 + \frac{1}{4}\lambda\phi^4, \quad \lambda > 0. \tag{2.3}$$

Note that this Lagrangian has a readily apparent parity symmetry ($\phi \to -\phi$). In the case that the mass term (μ^2) is greater than zero, this is simply the Lagrangian of a spin-0 particle with mass μ. The field value which minimizes the potential is defined as the vacuum expectation value (vev), ϕ_0. With $\mu^2 > 0$, the potential has a minimum at $\phi_0 = \phi = 0$. However, in the case that $\mu^2 < 0$, then the minimum of the potential is *not* at $\phi = 0$ and the vacuum expectation for the field is instead

$$\phi_0 = \pm\sqrt{-\mu^2/\lambda} \equiv \pm v, \tag{2.4}$$

which results in two separate, but degenerate, non-zero vacuum expectation values. While the Lagrangian remains unchanged, and still has an explicit reflection symmetry, there is no longer reflection symmetry about the field's non-zero ground state, ϕ_0.

In order to understand the consequence of the non-zero ground state, one of the degenerate vacuum states is arbitrarily chosen ($v = +\sqrt{-\mu^2/\lambda}$), and a shifted field, σ, is defined in relation to this ground state:

$$\sigma \equiv \phi - \phi_0 = \phi - v. \tag{2.5}$$

By construction, this new field σ has a vacuum expectation value, $\sigma_0 = 0$. Rewriting the original Lagrangian in Equation 2.3 in terms of the new shifted field yields:

$$\mathcal{L} = \frac{1}{2}\partial_\mu\sigma\partial^\mu\sigma - (-\mu^2)\sigma^2 - \sqrt{-\mu^2\lambda}\sigma^3 - \frac{\lambda}{4}\sigma^4 + \text{c}, \tag{2.6}$$

where c represents a collection of constant terms. The result is a scalar field with a mass of $m = \sqrt{-2\mu^2}$. However, the original reflection symmetry that was present in the Lagrangian found in Equation 2.3 has been lost. In choosing one of the two degenerate states, the intrinsic reflection symmetry has been *spontaneously* broken.

2.2.2 Spontaneous Symmetry Breaking and Goldstone Bosons

Now that the basic principle of spontaneous symmetry breaking has been introduced, whereby a shifted field was defined in relation to a non-zero vev, consider the more complicated case of a scalar field doublet $\boldsymbol{\phi} = \binom{\phi_1}{\phi_2}$, with a Lagrangian written as:

$$\mathcal{L} = \frac{1}{2}\left[\partial_\mu\phi_1\partial^\mu\phi_1 + \partial_\mu\phi_2\partial^\mu\phi_2\right] - V(\phi_1^2 + \phi_2^2), \tag{2.7}$$

$$V(\phi_1^2 + \phi_2^2) = \frac{1}{2}\mu^2(\phi_1^2 + \phi_2^2) + \frac{1}{4}\lambda(\phi_1^2 + \phi_2^2)^2. \tag{2.7a}$$

Like the case in Section 2.2.1 there is an explicit symmetry; in this case it is a continuous $SO(2)$ rotational symmetry. Once again, if $\mu^2 > 0$, the symmetry is unbroken and the Lagrangian simply describes two scalar fields, both with mass μ^2. In the case of $\mu^2 < 0$ the original $SO(2)$ symmetry is broken, and the vev can be defined as

$$\boldsymbol{\phi}_0^2 = -\mu^2/\lambda \equiv v. \tag{2.8}$$

In this case, the degenerate vacuum state forms a circle of radius v in the $\phi_1 \times \phi_2$ plane. While the choice of final vacuum state is completely arbitrary (the $SO(2)$ invariant Lagrangian demands it), for simplicity the vacuum state $\boldsymbol{\phi}_0 = \binom{v}{0}$ is chosen.

A new set of shifted fields can be defined, analogous to the shift done in the previous section (Equation 2.5), so that the vev of the new field is zero.

$$\boldsymbol{\sigma} \equiv \boldsymbol{\phi} - \boldsymbol{\phi}_o = \begin{pmatrix}\eta\\ \xi\end{pmatrix} \tag{2.9}$$

Substituting this shifted field into the original Lagrangian in Equation 2.7:

$$\mathcal{L} = \frac{1}{2}\left[\partial_\mu\eta\partial^\mu\eta + 2\mu^2\eta^2\right] + \frac{1}{2}\left[\partial_\mu\xi\partial^\mu\xi\right] + \text{c}. \tag{2.10}$$

This Lagrangian describes two particles: η, which has a mass of $\sqrt{-2\mu^2}$ and ξ, which is massless. The fact that a single massless boson is left after the symmetry breaking is a direct result of Goldstone's theorem [20, 21, 22], which states that one massless boson will appear for every continuous symmetry that is broken.

2.2.3 Spontaneous Symmetry Breaking in the Standard Model

The process of defining a shifted field to account for a non-zero vev can be used when considering a locally gauge invariant Lagrangian (something analogous to the Standard Model). The following is an example of the Abelian Higgs model [10, 19, 23].

A U(1) invariant Lagrangian that couples a charged complex scalar, ϕ, to electromagnetism can be written as follows:

$$\mathcal{L} = D_\mu \phi^* D^\mu \phi - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} - V(\phi) \tag{2.11}$$

$$V(\phi) = -\mu^2(\phi^*\phi) - \lambda(\phi^*\phi)^2, \quad \lambda > 0. \tag{2.11a}$$

Here, D_μ is the covariant derivative. As in the previous subsection, there are two cases, depending on the sign of μ^2. The case where $\mu^2 > 0$ is uninteresting as there is no symmetry breaking, and so the $\mu^2 < 0$ case is examined, which will lead to a broken symmetry. With $\mu^2 < 0$, the vev of the field ϕ occurs at

$$\phi_0 = \sqrt{\frac{-\mu^2}{2\lambda}} \equiv v/\sqrt{2}. \tag{2.12}$$

Shifting the field by defining a new field with a vanishing vacuum expectation value, where $\phi' = \phi - \phi_0$, and introducing two new fields (η and ζ), the original complex scalar field can be written as;

$$\phi = \frac{e^{i\zeta/v}(v+\eta)}{\sqrt{2}} \approx \frac{v + \eta + i\zeta}{\sqrt{2}}, \tag{2.13}$$

which can be used in order to rewrite the original Lagrangian (Equation 2.11) in terms of these shifted fields:

$$\mathcal{L} = \frac{1}{2}\left[\partial_\mu \eta \partial^\mu \eta + 2\mu^2 \eta^2\right] + \frac{1}{2}\left[\partial^\mu \zeta \partial_\mu \zeta\right] - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} + qv A_\mu \partial^\mu \zeta + \frac{q^2 v^2}{2} A_\mu A^\mu + \mathcal{L}_{int}. \tag{2.14}$$

In the above equation, $\mathcal{L}_{int}$ contains the interaction terms (which are not of interest here). In this Lagrangian, the term $qvA_\mu\partial^\mu\zeta$ is bilinear and unphysical; this implies a transformation can be performed in order to get the Lagrangian in a form in which the particle content is more readily apparent. Performing the gauge transformation $A_\mu \to A'_\mu = A_\mu - \frac{1}{qv}\partial_\mu \zeta$, eliminates all of the bilinear terms, resulting in the following:

$$\mathcal{L} = \frac{1}{2}\left[\partial_m \eta \partial^\mu \eta + 2\mu^2 \eta\right] - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{q^2 v^2}{2} A'_\mu A'^\mu + \text{c.} \tag{2.15}$$

In this form, the particle content can be seen; a massive photon A'_μ, and massive boson, η, resulting from the shifted field. Note that the massless Goldstone boson, ζ, is missing from this form of the Lagrangian. In popular parlance, the ζ field is said to have been *eaten* by the symmetry breaking and gauge transformation, resulting in the massive A'_μ.

Note that no degrees of freedom have been lost. Originally there were 2 degrees of freedom from the massless A_μ, and two from the complex scalar field. In the end, there are 3+1 degrees of freedom left: one from the massive scalar η, and the other three from A'_μ, which now has a longitudinal polarization

state.

***Finally*, the Standard Model**

The same process can be applied in the case of the Standard Model, with $SU_L(2) \times U_Y(1)$ symmetry. Although the calculation is more involved, the result is the same; the Goldstone bosons are *eaten* to give an extra degree of freedom (mass/longitudinal polarization) to each of the $W^{\pm}$, Z^0 gauge bosons. In order to provide the required degrees of freedom for the massive gauge bosons, the Higgs field cannot be a simple scalar field as in the above example, but rather a complex doublet, with four degrees of freedom. In the end, the three weak gauge bosons become massive, using three degrees of freedom, leaving a single degree of freedom representing a massive scalar boson: the Higgs.

While the mass terms of the gauge bosons fall out of the shifted Lagrangian, the question of fermion masses still remains; although they are not present in the final Lagrangian, they can be be given explicit mass terms through the $SU_L(2) \times U_Y(1)$ invariant Yukawa couplings of the form:

$$\mathcal{L}_{fermion} = \frac{gv}{\sqrt{2}} \left(\bar{\psi}_L \psi_R + \text{h.c.} \right) + \dots \tag{2.16}$$

In the above expression, g is the Yukawa coupling, which is proportional to the fermion mass. While the coupling of the Higgs to fermions is proportional to the fermion mass, the coupling of the Higgs to the massive vector bosons is proportional to the boson mass squared ($g \propto m_V^2$).

Throughout the discussion of the Higgs and electroweak symmetry breaking, there has been one feature of the Higgs boson that has not been discussed: its mass. The reason behind this is simple; there is no prediction in the Standard Model for the mass of the Higgs. Searches for the Higgs boson therefore have to cover a large range of possible mass hypotheses, covering the range in m_H that has not been excluded by experiments at LEP and the Tevatron, or on theoretical grounds. Once the value of the Higgs mass has been determined, the Standard Model has exact predictions for other related properties, such as the Higgs width, and the branching ratios for all possible Higgs decays. Of particular interest in this thesis is impact of the large width for a heavy (SM-like) Higgs, which is discussed in the next section.

2.3 Heavy Higgs Bosons

Unstable particle resonances are generally modelled in high energy physics as having a Breit-Wigner (BW) lineshape. The BW Higgs lineshape can be written as;

$$\text{BW}(s) = \frac{1}{\pi} \frac{\mu_H \Gamma_H}{\left(s - \mu_H^2\right)^2 + \left(\mu_H \Gamma_H\right)^2}, \tag{2.17}$$

where s is the Higgs virtuality, and μ_H its mass. The BW modelling works well up until the unstable particle's width (Γ) becomes comparable to its mass [14]. A general rule of thumb is that once the ratio Γ/m exceeds ~10%, the lineshape can no longer naively be taken as a Breit-Wigner.

Figure 2.1 shows that at relatively small Higgs masses, the width is very small, and the approximation works well (take for example $m_H = 125$ GeV where the Higgs width is 4.07×10^{-3} GeV [25], which is $\ll 10\%$ of the mass). At higher masses, this approximation breaks down; at 400 GeV the width is 29.2 GeV (~7.3%), and at 450 GeV the width/mass ratio is greater than 10% ($\Gamma = 46.8$ GeV) [25].

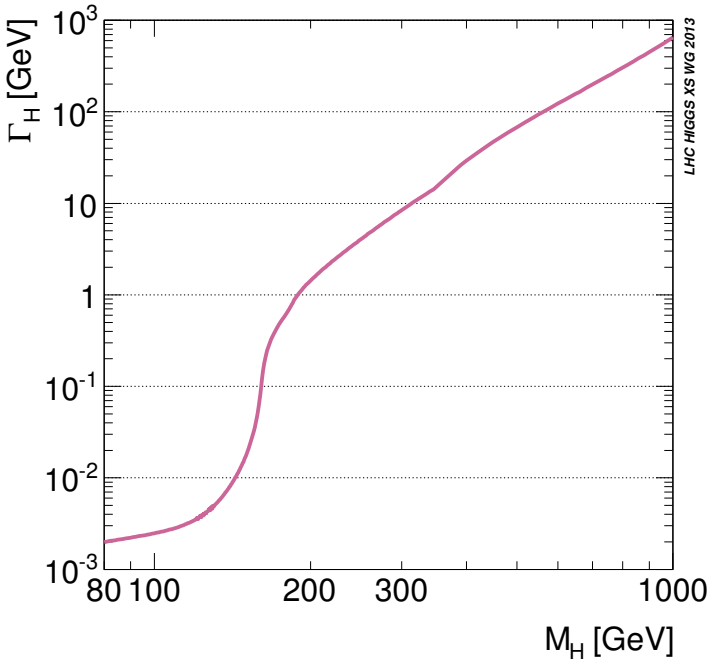


Figure 2.1: Higgs width as a function of SM Higgs mass [24].

Therefore, for $m_H > 400$ GeV, the approximate BW lineshape must be abandoned. Its replacement, the complex pole scheme, should be used instead [26]. The cross-section for the Higgs (going from some initial state i to some final state f), as described by the CPS is given as follows:

$$\sigma_{i\to H\to f}(s) = \frac{1}{\pi}\sigma_{i\to H}\frac{s^2}{|s-s_H|^2}\frac{\Gamma_H^{tot}(\sqrt{s})}{\sqrt{s}}\mathrm{BR}_{H\to f}, \tag{2.18}$$

$$\Gamma_H^{tot}(\sqrt{s}) = \sum_f \Gamma_{H\to f}(\sqrt{s}), \tag{2.18a}$$

with s being the Higgs virtuality and s_H the Higgs complex pole. The complex pole of the Higgs (and of unstable particles in general) can be written as

$$s_H = \mu^2 - i\mu\gamma, \tag{2.19}$$

where the parameter μ is the particle's mass, and γ is its width.

2.4 Higgs Production at the Large Hadron Collider

Assuming a Standard Model Higgs boson, for any given m_H the cross-sections for the production mechanisms and the branching ratios (BR) for the decays are well defined and are shown in Figures 2.2 and 2.3, respectively. In the Higgs mass range $80 \leq m_H \leq 1000$ GeV the dominant production mechanisms at the LHC are gluon-gluon fusion (ggF), vector boson fusion (VBF) and associated production (VH). Feynman diagrams for these three main LHC production processes are shown in Figure 2.4.

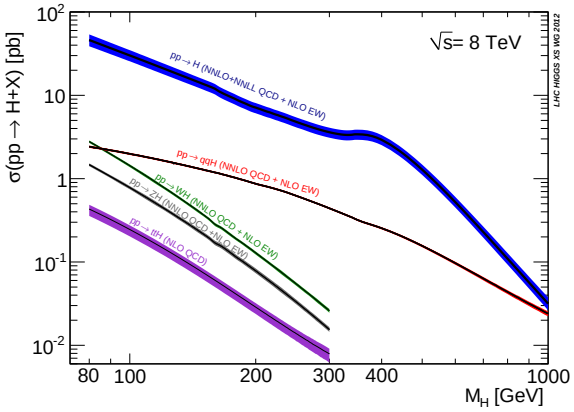


Figure 2.2: Higgs production cross-sections at $\sqrt{s} = 8$ TeV [24].

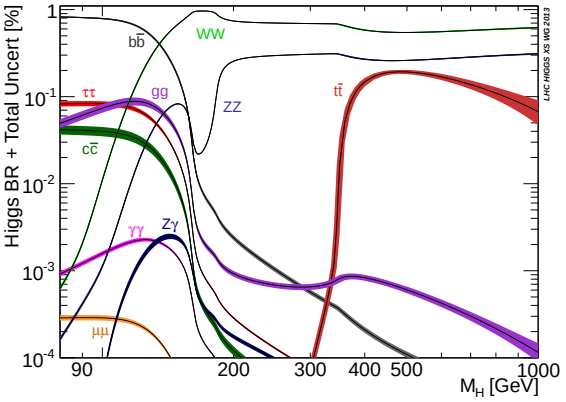


Figure 2.3: Higgs decay branching ratios broken up by final state channel as a function of Higgs mass [14].

Of these three main contributors to Higgs production, the two dominant production mechanisms at $\sqrt{s} = 8$ TeV are ggF and VBF. While the ggF production mode has a larger cross-section than VBF throughout the mass range (by an order of magnitude at low m_H), the difference in their cross-sections decreases until at $m_H \sim 1000$ GeV the two production rates are similar. At lower masses, associated production becomes more comparatively important, and exceeds the VBF cross-sections for an extremely light Higgs ($m_H \sim 80$ GeV).

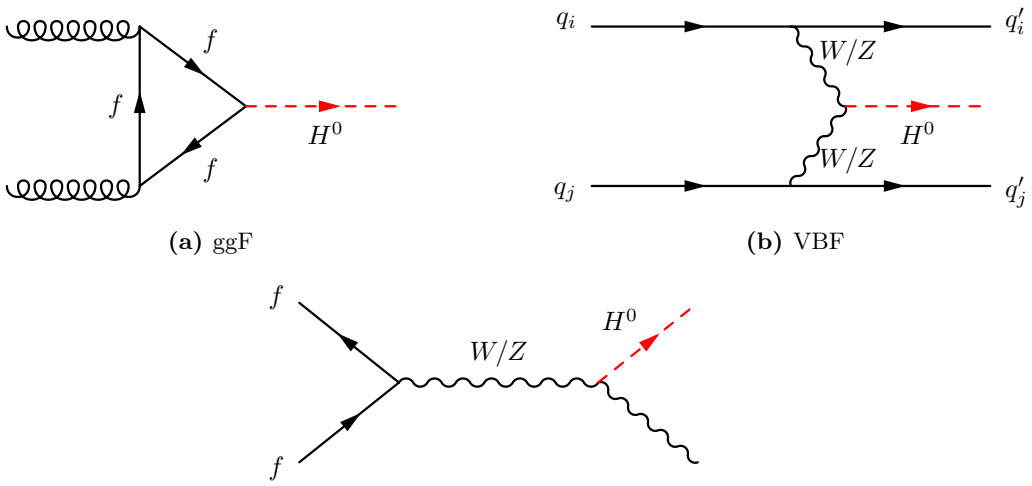
As the Higgs couples to mass, the largest branching ratio for a given Higgs mass will be the heaviest kinematically accessible (and dynamically allowed) final state. Therefore, at low masses ($m_H \leq 150$ GeV) decays of the Higgs are dominated by $H \to b\bar{b}$ decays. However, due to the purity of the signal[1], the two most promising channels for low mass discovery are $H \to \gamma\gamma$ and $H \to ZZ^{(*)}$. At higher mass, the $H \to VV$ decays are preferred, with the highest branching ratio being in the $H \to WW$ channel (which is a factor of 2 higher than the other VV final state, ZZ, due to the production of identical particles in the final state).

In the summer of 2012, the ATLAS and CMS Collaborations both announced the observation of a Higgs-like boson at $m_H = 125$ GeV [11, 12]. Subsequent measurements have shown that this newly discovered particle has properties consistent with the Higgs boson predicted by the SM. Spin-parity measurements, production and coupling measurements by CMS and ATLAS have found no significant deviation from SM predictions for a SM Higgs of this mass [27, 28, 29, 30, 31, 32].

2.5 Higgs Sector Beyond the Standard Model

Although there is no evidence yet for a deviation of the properties of the observed state from the SM Higgs boson predictions, the possibility that the observed Higgs is not exactly as predicted by the SM is not ruled out. Instead, the Higgs already observed could be part of an extended Higgs sector, as postulated by some theories that are *beyond the Standard Model* (BSM). As mentioned earlier in this chapter, the Higgs mechanism is needed in order to preserve local gauge invariance when introducing mass terms for the fermions and massive gauge bosons. The addition of Higgs mediated diagrams in

[1] Purity being the relative number of signal and background events in the final state.



(a) ggF **(b)** VBF **(c)** VH

Figure 2.4: Feynman diagrams of dominant Higgs production modes at the LHC: gluon-gluon fusion (a), vector boson fusion (b), and associated production (c).

W/Z scattering is also needed in order to preserve unitary at high center-of-mass energies. An open question is whether the Higgs discovered at 125 GeV is solely responsible for cancelling out the divergent process, and preserving unitarity.

In the case that there are one or more undiscovered Higgs bosons that are part of a BSM Higgs sector, there are two common benchmark models [14] which predict an extended Higgs sector. In the first, the couplings to the gauge bosons and fermions are scaled from the SM prediction by a common factor (as in the electroweak singlet model, Section 2.5.1). In the second, the couplings to the gauge bosons and fermions are scaled by separate factors, as in Two Higgs Doublet Models (2HDM, Section 2.5.2). The EWS and 2HDM benchmarks are briefly explained in the following sections, along with a description of the Narrow Width Approximation, which can be used to probe the parameter space of these models, or apply model-independent constraints on other BSM Higgs models.

2.5.1 Electroweak Singlet (EWS)

The EWS model adds a real singlet field to the original Higgs complex doublet [14]. The end result is two neutral CP-even scalars, h and H. Convention dictates that the more massive of the two is H. While not predicted in the original Higgs theory, the second scalar, H, is referred to in this thesis as an additional Higgs boson, or a Higgs-like boson as it has similar properties as h.

The couplings of each of the two scalars to the vector bosons and fermions is scaled by a common factor relative to the SM prediction by κ and κ' (for h and H, respectively). To preserve unitary for longitudinal W/Z gauge boson scattering, the following sum rule is required:

$$\kappa^2 + \kappa'^2 = 1. \tag{2.20}$$

This results in the cross-section and width of the lighter Higgs being scaled by the coupling strength κ^2, with branching ratios remaining the same as for the SM Higgs.

$$\sigma_h = \kappa^2 \times \sigma_{h,\mathrm{SM}} \tag{2.21}$$

$$\Gamma_h = \kappa^2 \times \Gamma_{h,\mathrm{SM}} \tag{2.22}$$

$$\mathrm{BR}_{h\to f} = \mathrm{BR}_{\mathrm{SM},h\to f} \tag{2.23}$$

For the heavier H state, alternate decay modes, such as $H \to hh$, may be possible (depending on the mass of H). Therefore, the scaling of the SM cross-sections and widths is not as simple as in the light case. Instead, the scaling is a function of κ' and $\mathrm{BR}_{\mathrm{new}}$, which is the branching ratio of H into decay modes not allowed in the SM case:

$$\sigma_H = \kappa'^2 \times \sigma_{H,\mathrm{SM}}, \tag{2.24}$$

$$\Gamma_H = \frac{\kappa'^2}{1-\mathrm{BR}_{\mathrm{new}}} \times \Gamma_{H,\mathrm{SM}}, \tag{2.25}$$

$$\mathrm{BR}_{H\to f} = (1-\mathrm{BR}_{\mathrm{new}}) \times \mathrm{BR}_{\mathrm{SM},H\to f}. \tag{2.26}$$

A careful reader will notice that κ^2 plays the same role in the EWS model as the SM-normalized signal strength μ ($\mu \equiv \sigma_{obs}/\sigma_{\mathrm{SM}}$). Therefore, measurements of μ in the SM context imply a constraint on κ'^2 and any possible extra Higgs bosons ($\kappa'^2 = 1-\mu_h$). From Equations 2.24–2.26 the SM relative signal strength for H, μ', can be written as:

$$\mu' = \frac{\sigma' \times \mathrm{BR}'}{\sigma_{\mathrm{SM}} \times \mathrm{BR}_{\mathrm{SM}}} = \kappa'^2(1-\mathrm{BR}_{\mathrm{new}}). \tag{2.27}$$

Figure 2.5 summarizes this information by showing how the signal strength and the width of the heavy boson vary in the $\mathrm{BR}_{\mathrm{new}} \times \kappa'^2$ plane.

The already discovered Higgs shows that μ is close to one, implying that κ^2 (in the context of an EWS extension to the Higgs sector) is also close to one. Examining Figure 2.5 this would indicate that any new heavy Higgs-like resonance would almost certainly be narrower than Γ_{SM} (unless $\mathrm{BR}_{\mathrm{new}}$ is significant), and is most likely quite narrow.

2.5.2 Two Higgs Doublet Models

2HDM models [14, 33] are minimal extensions to the Standard Model that have an extended Higgs sector in the form of a second electroweak doublet. While itself a benchmark model, some well known extensions to the Standard Model have a two Higgs doublet Higgs sector – supersymmetry (SUSY) [34] being the most notable example. A Higgs sector with a two Higgs doublet is also promising as it is able to generate CP violation and baryon asymmetry at a larger scale than the SM – large enough to match cosmological observations [33].

With the addition of the second doublet, there are now 5 particles in the Higgs sector: h, H, A and $H^{\pm}$. The two bosons h and H are the more familiar neutral, CP-even scalar bosons[1], while A is a neutral, CP-odd, pseudo-scalar. The two remaining scalars, $H^{\pm}$, are CP-even, charged Higgs bosons. There are two further parameters in the 2HDM models; α, which is the mixing angle between the two

[1] Like the EWS convention, h is the less massive of the two.

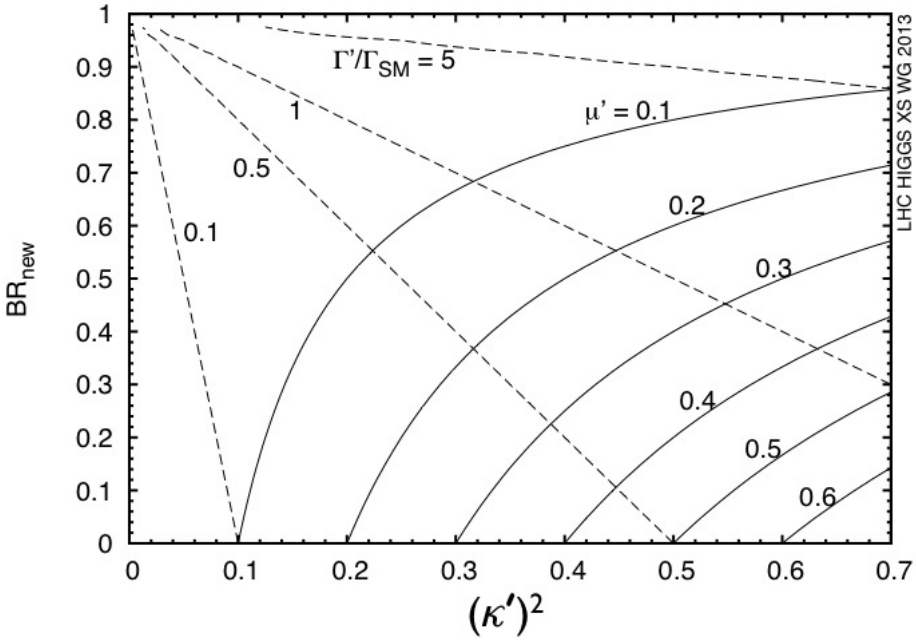


Figure 2.5: Possible values of μ' and Γ' in the $\mathrm{BR}_{\mathrm{new}} \times \kappa'^2$ plane for an extended Higgs sector with an EWS Higgs [14].

CP-even neutral Higgs states, and $\tan\beta$, which is the ratio of the vacuum expectation values of the two doublets. Therefore, in total a 2HDM has six free parameters: m_h, m_H, m_A, $m_{H^\pm}$, $\tan\beta$, and α.

Unlike the single common scaling of couplings in the EWS case, in 2HDM the scalar Higgs' (h and H) couplings to the vector bosons and fermions are scaled by two separate factors: $\kappa_V^{(\prime)}$ for the vector boson couplings, and $\kappa_f^{(\prime)}$ for the fermionic couplings. Like the case of the EWS models, the scaling of the couplings to the vector bosons is required to satisfy a sum rule in order to maintain unitarity.

2.5.3 Narrow Width Approximation

The Narrow Width Approximation (NWA, also known as the Zero Width Approximation, ZWA) assumes that the Higgs width is negligibly small; much smaller than both the Higgs mass, and the experimental mass resolution. As discussed in the case of the EWS, depending on the values of κ^2, and $\mathrm{BR}_{\mathrm{new}}$, second, more massive Higgs boson may be quite narrow, and for small values of μ' might be negligibly narrow (compared to m_H). This is not an observation that is unique to the EWS case; in some regions of the 2HDM phase space the heavier neutral Higgs, H, also has a negligibly small width. For these cases, it is safe to use the NWA scheme, where the Higgs is generated on-shell, as a very narrow BW.

The search for a Higgs boson with a NWA lineshape is not intended to be a direct search in the extreme of the EWS parameter space, or to apply direct limits to the 2HDM parameter space (although that could be done). Rather, the limits provide model-independent restrictions, which can then be extended to any exotic Higgs theory which predicts a very narrow resonance.

2.6 Interference Effects

Another factor that becomes increasingly important at higher masses (and therefore larger widths[1]) is the quantum mechanical interference between the Higgs production $pp \to H \to WW$ and the related non-resonant background, in this case $pp \to WW$. The interference between these two processes can not only change the overall cross-section of Higgs production, but can also alter the kinematic distributions and lineshape. These effects are non-negligible once the Higgs boson mass is greater than ∼400 GeV [14, 35, 36, 37], assuming the SM relationship between the Higgs mass and width.

When speaking of interference effects, it is useful to parameterize the overall $pp \to WW$ invariant mass distribution into three separate components; signal ($\mathcal{S}$), background ($\mathcal{B}$) and interference ($\mathcal{I}$). The signal component refers to the portion of the WW spectrum coming solely from Higgs production and decay (with no interference taken into account). The background component is the non-resonant $pp \to WW$ component that would be present *without* the presence of an additional heavy Higgs. Finally, the interference component, which is the difference between the combined $\mathcal{SBI}$ and $\mathcal{S}+\mathcal{B}$ distributions. Note that the interference term does not have to be positive, as the interference can be constructive, or destructive (depending on $m_{WW}(m_H)$). The total cross-section for the $pp \to WW$ process can then be written as;

$$\sigma_{tot,\ pp\to WW} = \sigma_{\mathcal{S},\ pp\to WW} + \sigma_{\mathcal{B},\ pp\to WW} + \sigma_{\mathcal{I},\ pp\to WW}. \tag{2.28}$$

As the interference spectrum is not present if the signal itself is not present, the $\mathcal{S}+\mathcal{I}$ spectrum can be treated as the de-facto signal when performing a search.

Existing Monte Carlo event generators for Higgs production do not include the interference terms by default. Therefore, the change in the lineshape and kinematics due to the interference must be accounted for afterwards. There are now multiple Monte Carlo generators that can be used to extract the interference spectrum for the $pp(\to H) \to WW$ process. Monte Carlo generators such as Phantom [38], aMC@NLO [39] and VBFNLO [40, 41, 42] are capable of determining the interference effects from the inclusion or exclusion of a Higgs diagram for VBF WW production, while MCFM [43] is able to account for the interference in ggF production [36]. Further discussion of the procedure used to account for interference effects is discussed in Chapter 5.

2.6.1 Uncertainties on Interference Effects

In the case of interference in the ggF production mode, the background and interference components are only known to LO (leading order), while the signal is known to NNLO (next to next leading order). While one can treat the LO interference term as a correction to the NNLO signal, there is no guarantee that the NLO (next to leading order) and NNLO corrections to the interference are small, since higher-order corrections to $\mathcal{B}$ are unknown.

The two options are to either keep the interference as a LO correction to the signal evaluated at NNLO (*additive*), or to include a large K-factor to have an effective NNLO interference term (*multiplicative*). The following is a treatment discussed in Refs. [35, 14], which details both solutions, and proposes a central *intermediate* value. The *intermediate* value provides the nominal estimate of $\sigma_{\mathcal{S}+\mathcal{I}}$, and the *additive* and *multiplicative* values are used to define the associated theoretical uncertainty.

[1] This assumes an SM-like relationship between the width and the mass. For lineshapes, such as the NWA, where the width is negligible, the interference effects are similarly negligible.

First, consider the observable distribution, D, which in this case is the invariant mass spectrum,

$$D = \frac{d\sigma}{dx}, \qquad x = m_{WW}. \tag{2.29}$$

Additive and *multiplicative* corrections to the $\mathcal{S}+\mathcal{I}$ distribution can be defined, as found in Equations 2.30 and 2.31.

$$\left.\frac{d\sigma^{\text{NNLO}}_{\text{eff}}}{dx} = \left(\frac{d\sigma^{\text{NNLO}}}{dx}\right)_{\mathcal{S}} + \left(\frac{d\sigma^{\text{LO}}}{dx}\right)_{\mathcal{B}} + \left(\frac{d\sigma^{\text{LO}}}{dx}\right)_{\mathcal{I}} \right\} \textit{Additive} \tag{2.30}$$

$$\left.\frac{d\sigma^{\text{NNLO}}_{\text{eff}}}{dx} = K_D\left[\left(\frac{d\sigma^{\text{LO}}}{dx}\right)_{\mathcal{S}} + \left(\frac{d\sigma^{\text{LO}}}{dx}\right)_{\mathcal{I}}\right] + \left(\frac{d\sigma^{\text{LO}}}{dx}\right)_{\mathcal{B}} \right\} \textit{Multiplicative} \tag{2.31}$$

$$K_D = \left(\frac{d\sigma^{\text{NNLO}}/dx}{d\sigma^{\text{LO}}/dx}\right)_{\mathcal{S}} \tag{2.31a}$$

The additive form leaves the interference as a LO correction to the signal, while the multiplicative option uses a (large) K-factor to attempt to correct for the NNLO effects.

One can then define the *intermediate* correction (Equation 2.32) which is to be used as the central value of the interference. The intermediate value provides a *best guess* value for the $\mathcal{S}+\mathcal{I}$ cross-section, where the signal component has a multiplicative factor assigned in order to match NNLO prediction, and the interference spectrum is scaled by a K-factor which only includes gluon initiated diagrams.

$$\left.\frac{d\sigma^{\text{NNLO}}_{\text{eff}}}{dx} = K_D\left(\frac{d\sigma^{\text{LO}}}{dx}\right)_{\mathcal{S}} + \sqrt{K_D^{gg}}\left(\frac{d\sigma^{\text{LO}}}{dx}\right)_{\mathcal{I}} + \left(\frac{d\sigma^{\text{LO}}}{dx}\right)_{\mathcal{B}} \right\} \textit{Intermediate} \tag{2.32}$$

$$K_D = K_D^{gg} + K_D^{\text{rest}} \tag{2.32a}$$

$$K_D^{gg} = \frac{\left(d\sigma^{\text{NNLO}}/dx\right)_{gg\to H(g)\to WW(g)}}{\left(d\sigma^{\text{LO}}/dx\right)_{gg\to H\to WW}} \tag{2.32b}$$

With the intermediate option giving the central value, the multiplicative and additive options form a band, which can then be used as the theoretical uncertainty on the normalization of the ggF interference spectrum.

In the case of the uncertainty for the VBF interference, the difference between the LO and NNLO interference is expected to be small [44, 45, 46, 47, 48]. Therefore, the only uncertainty applied is on the modelling of the interference spectrum. The uncertainty is evaluated by comparing the $WW_{\mathcal{S}+\mathcal{I}}$ invariant mass spectrum obtained from two separate methods, and assigning the difference between the two as a systematic uncertainty. Repolo[1] is used to reweight the existing signal samples to obtain the $\mathcal{S}+\mathcal{I}$ lineshape and is then compared to the WW invariant mass lineshape obtained with a VBFNLO generated $\mathcal{SBI}-\mathcal{B}$ invariant mass spectrum. The effect of varied renormalization and factorization scales on the interference spectrum is used as an additional theoretical systematic uncertainty. A detailed discussion of the obtained interference weights, and the associated uncertainties is found in Chapter 5.

[1] **Re**weighting **Po**hweg at **L**eading **O**rder, part of the VBFNLO framework

Chapter 3

The ATLAS Detector at the LHC

"If only I had thought of a Kodak! I could have flashed that glimpse of the Under-world in a second, and examined it at leisure."

H.G. Wells, *The Time Machine*

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [49] is a superconducting particle accelerator, the largest of its kind, and currently producing the highest energy proton–proton (pp) and heavy ion collisions in the world. Its home is CERN (*Conseil Européenne pour la Recherche Nucléaire*), located just outside Geneva, Switzerland. The collider sits in a 26.7 km long circular tunnel originally constructed for LEP (the Large Electron-Positron Collider) and its associated experiments.

The collisions produced by the LHC are observed by four large scale experiments situated at different interaction points around the LHC ring. They are ALICE [50], ATLAS [51], LHCb [52] and CMS [53]. Most of the experiments are primarily interested in proton – proton collisions, except for ALICE, which was built solely to study lead nuclei collisions from the LHC.

The LHC, while the largest and highest energy collider at CERN, is not the only accelerator. Multiple other accelerators, boosters and storage rings support CERN's experimental program, and protons that eventually make it into the main LHC ring are first fed through a network of accelerators in order to bring them up to the appropriate energy and proton bunch density. The full CERN accelerator complex is illustrated in Figure 3.1.

When running in pp mode, the LHC is designed to produce collisions at a center of mass energy of $\sqrt{s} = 14$ TeV and an instantaneous luminosity of up to $\mathcal{L}_{\text{inst}} = 10^{34}$ cm^{-2}s^{-1}. The instantaneous luminosity is what determines the observed rate of physical processes, and is related to the beam parameters as follows:

$$\mathcal{L} = \frac{f_r n_b n_1 n_2}{2\pi \Sigma_x \Sigma_y}, \tag{3.1}$$

where f_r is the machine revolution frequency, n_b is the number of colliding bunches, $n_{1,2}$ are the number of protons in each colliding bunch, and $\Sigma_{x,y}$ are the convoluted transverse beam profiles [54]. At design

instantaneous luminosity, the LHC will be able to deliver a maximum projected integrated luminosity of $80 - 120$ fb^{-1} per year [55].

While the original design energy of the LHC is 14 TeV, during the initial start-up and commissioning of the machine in 2008, a major technical incident forced the LHC to shutdown for a year in order to complete repairs. Once back online, the center of mass energy had to be lower than the design center of mass energy: $\sqrt{s} = 7$ TeV in 2011, and $\sqrt{s} = 8$ TeV in 2012. In May of 2015 the LHC produced pp collisions at a world record center of mass energy of 13 TeV.

3.2 ATLAS Detector Concept

ATLAS is a $\sim 4\pi$ hermetic detector and consists of three main components – the inner detector (ID), calorimeter, and muon spectrometer (MS) – all of which consisting of multiple subdetectors. Figure 3.2 shows the full ATLAS detector, with components and subdetectors labelled.

The ATLAS coordinate system is right-handed with the positive z-axis pointing in the direction of counter-clockwise beam rotation (as viewed from above). The $x - y$ plane is then perpendicular to the beam axis with the positive x-axis pointing towards the center of the LHC ring, and the positive y-axis pointing upwards. Due to the nature of hadron colliders, it is often useful to use transverse quantities, such as the transverse momentum, p_{T}, which is the momentum in the $x - y$ (transverse) plane.

Because of the approximate cylindrical symmetry of the detector, it is useful to specify the coordinates of detector components and particle trajectories in the form of polar coordinates. The angle θ is the polar angle as measured from the positive z-axis, while the azimuthal angle, ϕ, is measured as the angle about the beam-pipe (with the $\phi = 0$ sitting on the positive x-axis, $\phi = \arctan\left[\frac{x}{y}\right]$). The polar coordinate can be used to construct another commonly used coordinate, pseudorapidity, given by

$$\eta = -\ln\left[\tan\left(\theta/2\right)\right]. \tag{3.2}$$

Pseudorapidity is preferred over the polar angle coordinate, as in the massless limit it approaches the Lorentz invariant *rapidity* of a particle, defined as;

$$y = \frac{1}{2}\ln\left(\frac{E + p_{\mathrm{z}}}{E - p_{\mathrm{z}}}\right). \tag{3.3}$$

Finally, of interest for discussion of issues such as object (electrons, jets, etc.) reconstruction, and object overlap removal it is useful to define the angular separation of two objects in the detector:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \tag{3.4}$$

3.3 Physics Goals and Detector Performance

ATLAS is considered a general purpose detector, meaning the detector is able to make precision measurements of a large number of physics processes, each with distinct topologies and final states. The design of the ATLAS detector took into account the performance needed to observe a wide range of processes, from QCD and super-symmetry, to electroweak and other SM processes. Observation and precision measurement of the Higgs boson (see Chapter 2) across a variety of final states was also a major design requirement. In order to fulfill these goals, ATLAS is required to have excellent charged

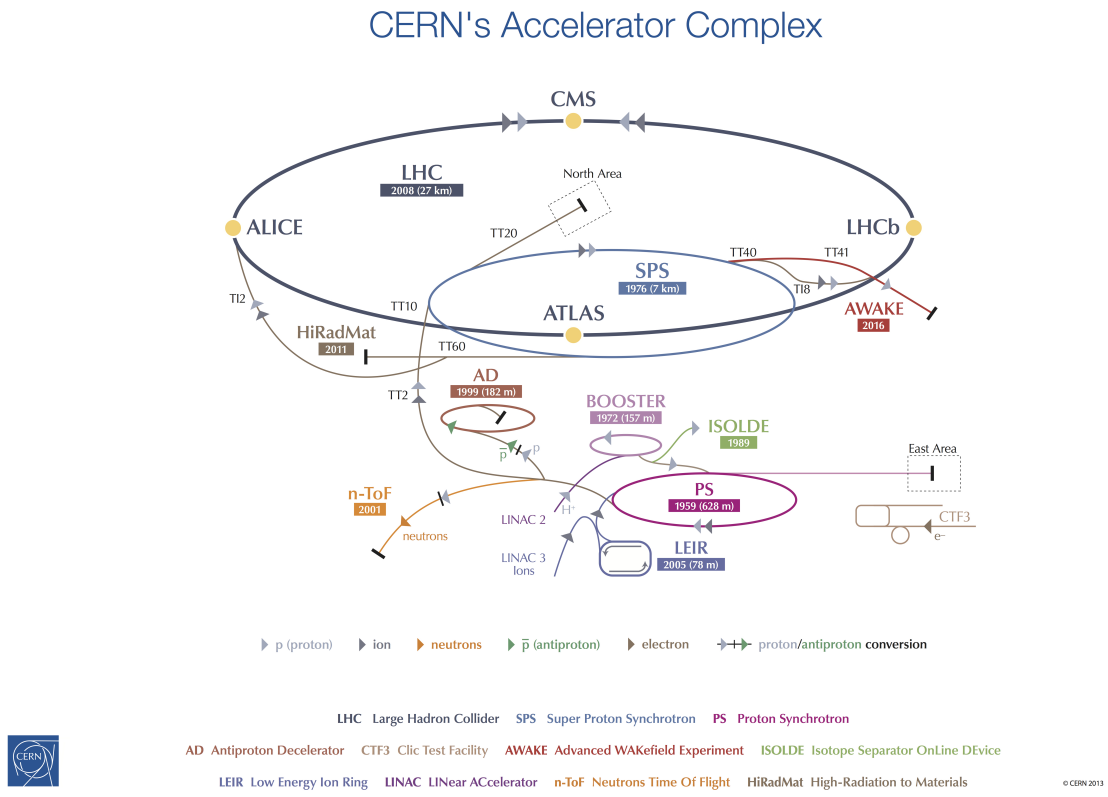


Figure 3.1: The CERN accelerator complex [56]

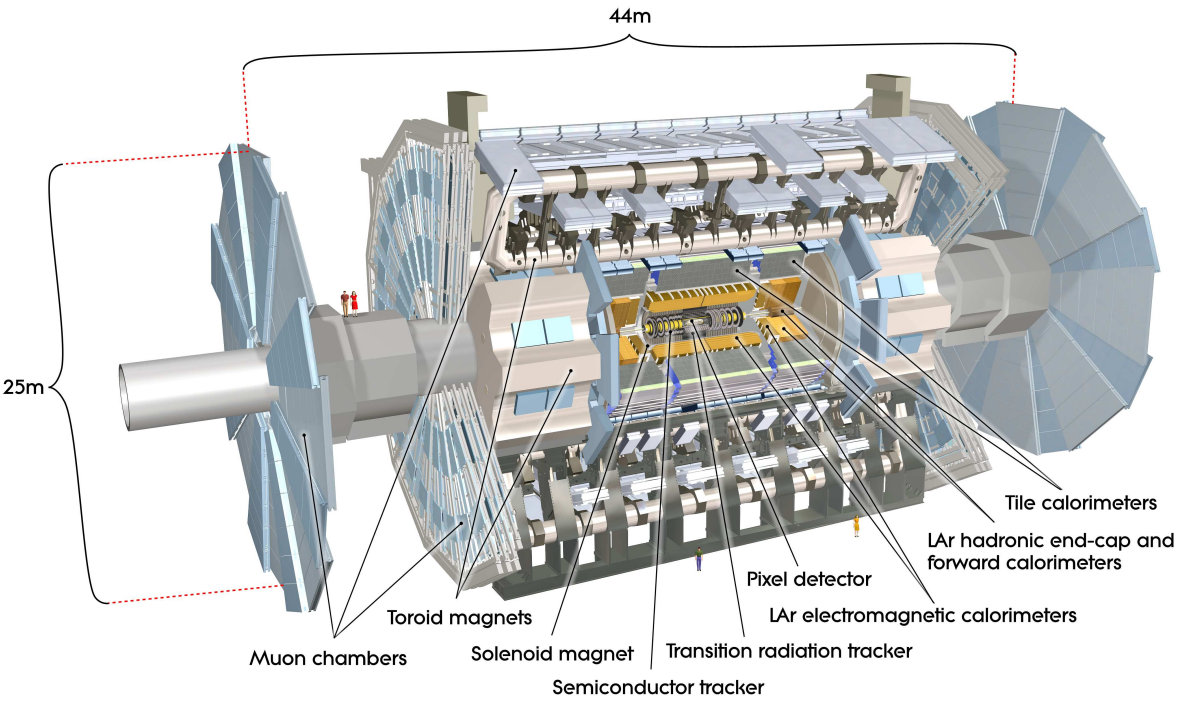


Figure 3.2: The ATLAS detector [51]

Table 3.1: Design performance of the ATLAS systems [51].

Detector Component	Resolution	Coverage	
		Measurement	Trigger
Tracking	$\sigma_{p_T}/p_T = 0.05\%\ p_T \oplus 1\%$	$\lvert\eta\rvert < 2.5$	–
EM Calorimetry	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$	$\lvert\eta\rvert < 3.2$	$\lvert\eta\rvert < 2.5$
Hadronic Calorimetry			
Barrel and end-cap	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	$\lvert\eta\rvert < 3.2$	$\lvert\eta\rvert < 3.2$
Forward	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$	$3.1 < \lvert\eta\rvert < 4.9$	$3.1 < \lvert\eta\rvert < 4.9$
Muon Spectrometer	$\sigma_{p_T}/p_T = 10\%$ @ $p_T = 1$ TeV	$\lvert\eta\rvert < 2.7$	$\lvert\eta\rvert < 2.4$

particle tracking, identification and charge measurements, precise calorimetry for both hadronic and EM showers, efficient triggering methods, precise muon transverse momentum measurement, all over a large range of pseudorapidity [57]. The expected performance of the ATLAS subsystems are summarized in Table 3.1.

Additionally, due to the large instantaneous luminosity delivered by the LHC machine, the electronic and mechanical components of ATLAS are required to be radiation hard, to prevent unacceptable degradation throughout the lifetime of the experiment.

The detector systems of ATLAS are described below. As the final state in this analysis makes particular use of the calorimeters for the object reconstruction, an emphasis is placed on the ATLAS calorimeter systems.

3.4 Inner Detector

The ATLAS Inner Detector (ID) is made up of three trackers close to the interaction point, along with the Beam Conditions Monitor (BCM) [58] which acts as a beam monitor and luminosity detector. The ID measures 5.3 m long and 2.5 m in diameter and has instrumented coverage for tracking up to $|\eta| = 2.5$. The three trackers (listed in order of increasing radial distance from the beam pipe) are the Pixel Detector (Pixel) [59], Semiconductor Tracker (SCT) [60], and Transition Radiation Tracker (TRT) [61]. The entire structure of the inner detector is shown in Figure 3.3. Figure 3.4 displays the radial positioning of the three trackers' central (barrel) components in relation to the LHC beampipe. Figure 3.5 displays the end-cap region of the trackers showing the positioning of the modules along the z-axis. The entire ID volume is located in a 2 T axial magnetic field (further discussion of the ATLAS magnet system is found in Section 3.7), which allows the determination of charged particle polarity and transverse momentum.

The ID is meant to provide charged particle track and momentum reconstruction, and identification of primary vertices for each event/bunch crossing. Due to the large number of interactions per bunch crossing (a result of the large proton density in each proton bunch), approximately 1000 particles will traverse the ID per bunch crossing at design luminosity. In order to be able to resolve each of the tracks, the ID needs to have superb granularity, and fast response.

Both the Pixel and SCT are semiconductor-based detectors, and so to limit the amount of noise (as a result of leakage current, which will grow with increasing integrated luminosity), the silicon detectors

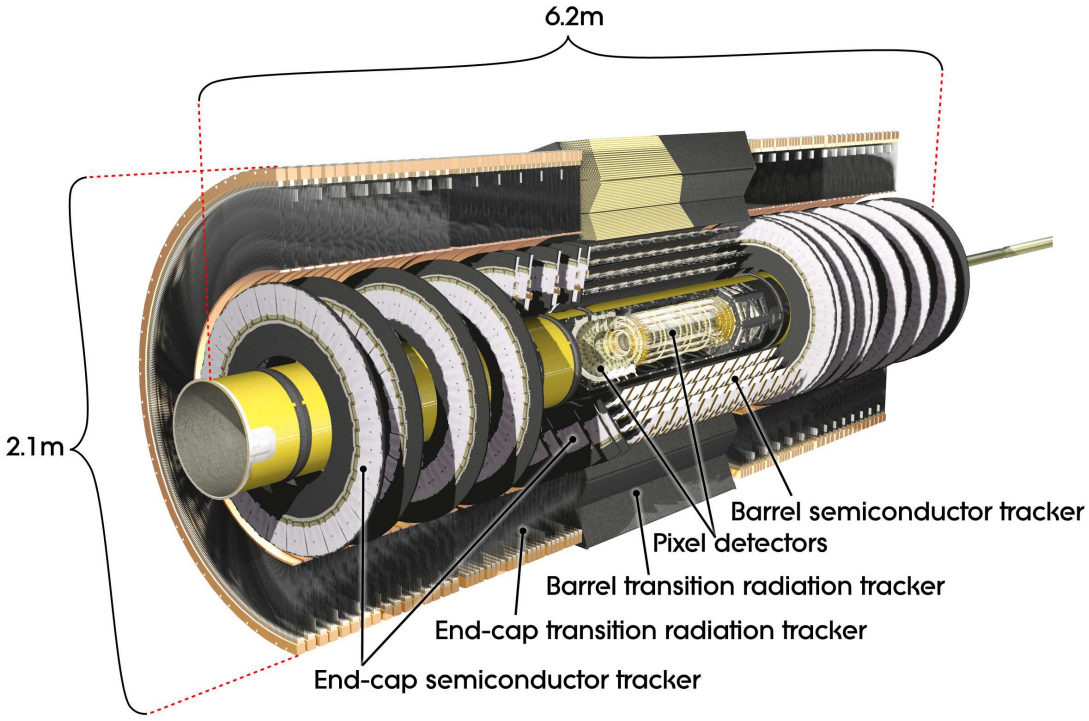


Figure 3.3: Cut-away view of inner detector trackers, with both the barrel and end-cap components of each sub-system (Pixel, Semiconductor Tracker and Transition Radiation Tracker) visible and labelled. While not shown in the figure, the Beam Conditions Monitor is located in the end-cap section, close to the beam pipe, and is hidden from view by the other subdetectors [51].

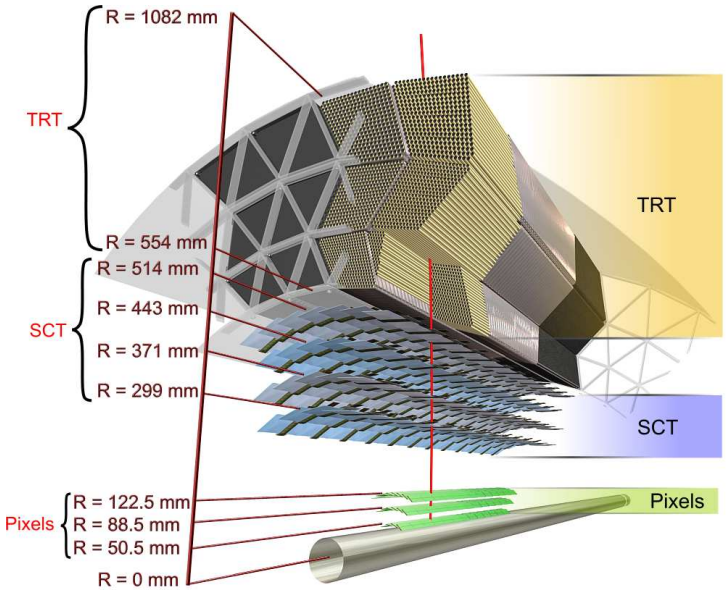


Figure 3.4: Cut-away view of the barrel section of the inner detector trackers. Superposed in red is a charged particle track going through the three trackers making up the ID [51].

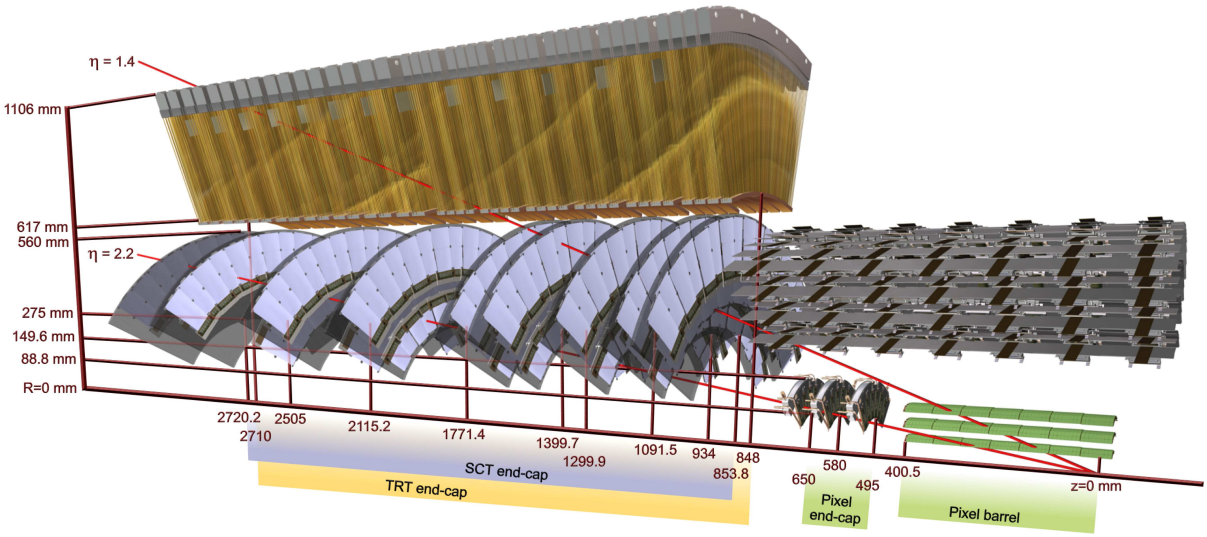


Figure 3.5: Cut-away view of the end-cap section of the inner detector trackers [51].

Table 3.2: Precision of the Inner Detector Trackers [51].

Item		Intrinsic Precision [μm]
Pixel		
Barrel	3 cylindrical layers	10 $(R-\phi)$ 115 (z)
End-cap	2×3 disks	10 $(R-\phi)$ 115 (R)
SCT		
Barrel	4 cylindrical layers	17 $(R-\phi)$ 580 (z)
End-cap	2×9 disks	17 $(R-\phi)$ 580 (R)
TRT		
Barrel	73 straw planes	130 $(R-\phi)$
End-cap	160 straw planes	130 $(R-\phi)$

must be cooled, and are operated between $-5°C$ and $-10°C$. The TRT on the other hand, is designed to operate at room temperature.

The spatial precision of the individual modules that make up each ID subdetector are listed in Table 3.2. These values represent the intrinsic resolving capabilities of a single particle traversing a single plane of modules, and do not take into account the improved precision and resolution obtained by fitting a particle trajectory through the multiple layers of sensors in the tracker.

3.4.1 Pixel Detector

The Pixel tracker is the subdetector component that is located closest to the interaction point and has the finest granularity in $R-\phi$ coordinates, with more than 80.4 million readout channels. The excellent granularity of the Pixel detector means that it is able to reconstruct particle trajectories with high precision, and is able to reconstruct track vertices with high precision – an essential quality for the tagging and identification of b-jets (discussed in Section 4.6.2).

The Pixel detector is made up of three concentric cylinders of modules that make up the barrel section, and six disks (three on either side of the interaction point), aligned perpendicular to the z-axis,

in the end-cap region. Therefore, for each charged particle traversing the ID, the Pixel detector will nominally provide three space point measurements. Figure 3.5 clearly shows the barrel and end-cap components of the Pixel detector.

The sensors [62] making up the Pixel detector are all identical, each containing 47,232 pixels. The individual pixels are 50 μm in the azimuthal coordinate and 400 μm in the z direction (R direction for the end-cap sensors). The tracking precision of each pixel sensors (which has multiple individual readout pixels) is 10 μm $\times$ 115 μm in $(R-\phi)\times z$ for the barrel, and $(R-\phi)\times R$ for the end-caps.

3.4.2 Semiconductor Tracker

The SCT is laid out similarly to the Pixel detector (Figure 3.5) with a concentric cylindrical geometry for the central barrel section, and disks aligned perpendicular to the beamline for the end-caps. Like the Pixel detector, the semiconductor tracker is a high granularity detector, with over 6.3 million readout channels. In the barrel region, charged particles traverse eight SCT strip layers, mounted in small-angle stereo pairs (40 mrad angle), which results in four space point measurements. In the end-cap, there are nine disks of sensors per side. The SCT is made from 8448 identical p-in-n rectangular sensors for the barrel and 6944 p-in-n wedge shaped sensors in the end-caps. The sensors are each 6.4 cm long and have a strip pitch[1] of 80 μm. Each layer of SCT modules consists of two sensors mounted mounted back to back. In the barrel (end-cap) one sensor is aligned with readout strips pointing along the beam direction (pointing out radially), and the second sensor is rotated by 40 mrad with respect to the first. This arrangement results in a position precision of 17 μm $\times$ 580 μm (in $(R-\phi)\times z$) in the barrel and 17 μm $\times$ 580 μm (in $(R-\phi)\times R$) in the end-cap disks.

3.4.3 Transition Radiation Tracker

Surrounding the Pixel and SCT is the TRT. While the Pixel and SCT focus on making a low number of space point measurements with high precision/granularity, the TRT makes many measurements over a larger track length at much lower precision. A typical track will have 7 space point measurements in the Pixel and SCT (three in Pixel, four in SCT) but have 36 in the TRT. The lower granularity of the TRT is also evident in the number of readout channels that it contains: 351,000, versus the millions in the SCT and Pixel. The decreased spatial resolution/granularity of the TRT straw tubes is offset by the larger number of space point measurements, and the fact that these measurements are taken over a larger track length than the other detectors in the ID. In addition to reduced granularity, the TRT also has a more limited η coverage, only covering $|\eta| \leq 2.0$. The TRT is made from a matrix of 4 mm diameter straw drift tubes filled with a gaseous $Xe/CO_2/O_2$ mixture at $5-10$ mbar gauge pressure; at the center of each straw is a gold-plated tungsten anode readout wire. The straws are constructed from a multilayer of Kapton, aluminum and carbon-loaded polyimide, which are then reinforced with carbon fibers to enhance the mechanical stability. The main gas in the mixture is xenon (70%), which acts as the active medium and was chosen because of its high efficiency for capturing transition radiation photons, radiation hardness and linear response. The CO_2 (27%) is added as a natural quencher, which results in a constant electron drift velocity, and limits their lateral diffusion. Finally, O_2 (3%) is added to stabilize the amplification of the signal.

[1] The sensors in the end-cap do not have a constant pitch as a result of the irregular wedge-shaped geometry.

In the barrel region 52,544 straws are laid out parallel to the beam axis, and are 144 cm in length. The readout wires at the center of each barrel TRT straw tube are split into two halves at $\eta = 0$, meaning that there is no z position reconstructed from the TRT hits (other than which side of the barrel the charged particle traversed). As such, the TRT barrel is only able to provide information on the $R - \phi$ coordinate. In this coordinate, the TRT has an intrinsic precision of 130 μm for each straw hit.

In the end-cap region, the TRT straws are shorter (37 cm) and are splayed out radially in a fan-like pattern. Each of the end-caps contains 122,880 straws. Like in the barrel section, each TRT disk is only able to provide the single $R - \phi$ coordinate.

In the barrel, polypropylene fibers are placed between the straw planes, and in the end-caps polypropylene foils are woven between the straws. These act as periodic radiators for the creation of transition radiation photons, which are emitted whenever a relativistic charged particle traverses between two materials with different dielectric constants [63]. As charged particles travel through the ATLAS TRT, there is a chance to emit transition radiation photons each time a polypropylene fiber/foil is crossed. The probability for emitting a transition radiation photon is dependant on the Lorentz factor, γ, of the charged particle. As a result, high energy electrons will emit much more transition radiation than hadronic objects (such as charged pions), and muons. This allows the TRT to aid in particle identification in addition to providing momentum measurements for charged particles.

3.4.4 Beam Conditions Monitor

The BCM is a small detector originally designed and tasked with detecting dangerous beam conditions that have the ability to damage elements of the ATLAS detector. One such dangerous scenario would be if the LHC beam were to hit the collimators that sit on either side of the ATLAS detector, causing a large shower of energetic particles to spray into the detector. In such a case, the BCM is able to order an abort sequence which will dump the LHC beam.

There are two BCM detector stations located on either side of the ATLAS interaction point, at $z = \pm 1.9$ m ($|\eta| \approx 4.2$), and only 20 mm away from the beam pipe. To withstand the intense radiation at this position in ATLAS, the BCM uses a set of radiation hard synthetic diamond sensors. These sensors also have a very fast response time, which allows the BCM to discriminate between LHC pp collision products, and protons from the LHC beam, based on the Δt measured between the two opposite BCM stations [51]. Each station consists of four of these sensors (measuring 1×1 cm^2), separated azimuthally by $\phi = \pi/2$.

Although not initially intended for the purpose, the BCM now functions as one of the main luminosity detectors in ATLAS [64].

3.5 Calorimeters

The final state of the search presented in this thesis, while dependant on all sub-systems of the ATLAS detector, relies most heavily on the ATLAS calorimetry, which provides measurements of the electron and jet energies, as well as the missing transverse momentum (and provides supporting information for muon reconstruction and identification). Therefore, before describing in detail the ATLAS calorimeters, some basic aspects of calorimetry will first be discussed [65].

3.5.1 Calorimetry

A calorimeter can be said to be a homogeneous calorimeter or a sampling calorimeter. Homogeneous calorimeters are calorimeters where the entire volume is constructed from the active material, and energy deposited in any part of the calorimeter will be read out and measured. Conversely, sampling calorimeters are made up of *active* and *passive* layers (also called absorbers). The passive material is uninstrumented, and so the energy read out is only a *sample* of the full energy deposit. An important parameter of a sampling calorimeter is the *sampling fraction*, which is defined in ATLAS as the ratio of the measured energy of an electron to the electron's total energy. This definition of the sampling fraction is used as it defines the sampling fraction with respect to the *EM scale*. The EM scale is the energy deposited by a particle via the electromagnetic force (and therefore, electromagnetic showers), and is the scale at which the ATLAS calorimeters measure energy deposits. Electromagnetic showers, and the EM scale are briefly discussed below.

Electromagnetic Calorimetry

When EM particles, such as electrons and photons, transverse through a calorimeter they lose energy via the electromagnetic force, creating showers of photons, electrons, and positrons in the process. This also applies to charged hadronic particles, such as the charged pion, although a large portion of their energy is deposited through other, non-EM processes (this is discussed further below). The energy from the EM showers may be deposited via radiative losses (bremsstrahlung at high energies), ionization, Čerenkov radiation, δ-rays, etc. For high energy electrons, the dominant mechanism for energy loss is the Coulomb interaction, through the emission of bremsstrahlung radiation. It is only at lower electron energies that ionization losses become dominant. The energy at which the ionization losses are equal to the radiative losses is called the *critical energy*, ϵ_c. For electrons, the critical energy is quite low (<100 MeV), but scales as a function of $(m/m_e)^2$ for other massive charged particles.

Two important parameters for understanding the development of EM showers are the radiation length (X_0) and the Molière radius (ρ_M), which parameterize the longitudinal and lateral development of the shower, respectively. The radiation length is defined as the average distance over which an electron loses $(1 - e^{-1})$ of its energy ($\approx 63.2\%$) via bremsstrahlung radiation. The Molière radius is defined as

$$\rho_M = m_e c^2 \sqrt{4\pi/\alpha} \times \frac{X_0}{\epsilon_c} = 21.2 \ \text{MeV} \times \frac{X_0}{\epsilon_c}, \tag{3.5}$$

where α is the fine-structure constant. On average, approximately 90% of the energy deposited in an EM shower is contained in a cylinder of radius ρ_M.

Photons, being neutral, cannot radiate or ionize the calorimeter medium like electrons. Their main source of energy loss at high energies is through pair production and Compton scattering. At lower energies (on the order of 100 keV), the photoelectric effect and Rayleigh scattering become dominant. The probabilities for these processes are governed by properties of the calorimeter material, and the number of radiation lengths traversed. It can be shown that for very high energy photons, the cross-section for photon interactions in a given material (with radiation length X_0) is given by;

$$\sigma(E_\gamma \to \infty) = \frac{7}{9} \frac{A}{N_A X_0}, \tag{3.6}$$

where A is the atomic weight, and N_A is Avagadro's number. From Equation 3.6 it can be seen that

the mean free path for a high energy photon is $\frac{9}{7}X_0$.

Hadronic Calorimetry

While charged hadrons will lose energy due to ionization, much like in the case of an EM object, the largest source of energy loss is through strong interactions with atomic nuclei in the calorimeter (for neutral hadrons, this is the only method of energy loss). An interesting consequence is that some fraction of the hadronic energy deposits are unmeasureable by the calorimeter systems. This lost energy may escape detection in the form of nuclear binding energy and/or the creation of particles in the showering process such as neutrons (which may not interact with the calorimeter again), or neutrinos and high energy muons (which only interact very weakly with the detector). Energy measurements that do not take into account the missing hadronic energy, and rely only on direct measurement of the EM objects in hadronic showers are said to be measured on the EM scale.

Since the main mechanism for hadronic energy loss is intrinsically stochastic, a useful parameter for hadronic containment is the interaction length, λ_{int}, which represents the mean free path of a hadron traversing the calorimeter.

Energy Resolution in Sampling Calorimetry

The fractional resolution of the energy measurement of a calorimeter is commonly parameterized in the following way (components added in quadrature) [66]:

$$\frac{\sigma_E}{E} = \frac{A}{\sqrt{E}} \oplus \frac{B}{E} \oplus C. \tag{3.7}$$

Here, A is called the *stochastic term*, B is the *noise term* and C is called the *constant term*, with the dominant component of the equation depending on the overall energy deposit, E.

The stochastic term comes from statistical fluctuations in the shower development. Because of the inhomogeneous design of sampling calorimeters, this term can be rather large (at least when compared to homogeneous calorimeters). For sampling calorimeters, this term is related to the number of samplings, and the overall sampling fraction of the calorimeter.

At low $\mathcal{L}_{\text{inst}}$ the noise term is dominated by noise in the readout electronics. However, at the high luminosity environment of the LHC, the noise term is dominated by pile-up (see Section 4.1.1). The noise term can be reduced through electronic means via signal shaping and filtering.

Non-uniformities in the calorimeters that affect the readout energy, or gain through the readout chain, contribute to C. These non-uniformities can come in the form of inhomogeneous radiation damage in the calorimeter, channel-to-channel gain or response differences, temperature gradients in the active material, dead material, and cracks. As the deposited energy grows, the impact that the constant term makes on detector energy resolution becomes dominant.

3.5.2 ATLAS Calorimetry

The ATLAS calorimeters (shown in Figure 3.6) must provide good containment for electromagnetic and hadronic showers, and must also limit punch-through into the muon system, making calorimeter depth an important design consideration. The total thickness of the EM calorimeter is about 22 X_0 in the barrel and 24 X_0 in the end-caps. For capturing hadronic showers, the 9.7 interaction lengths

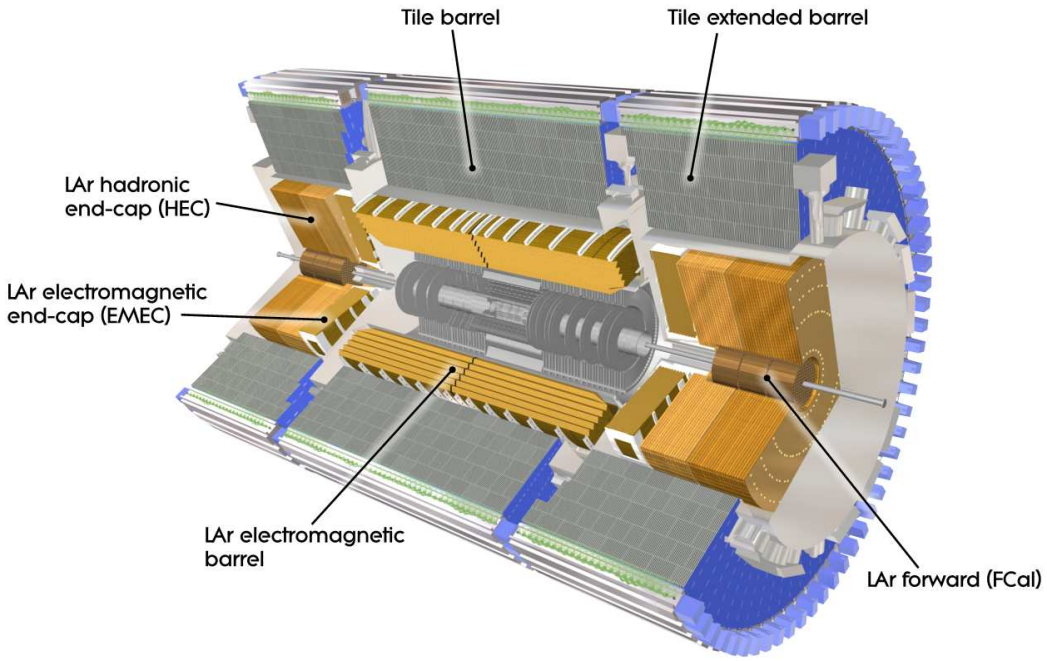


Figure 3.6: Cut-away view of ATLAS calorimeters [51].

Table 3.3: Specifications of the Calorimeter System [51].

Detector Sub-System	Absorber	Active Medium	Coverage
Barrel EM	Lead	LAr	$0 < \|\eta\| < 1.475$
Barrel Hadronic	Steel	Scintillating tiles	$0 < \|\eta\| < 1.7$
End-cap EM	Lead	LAr	$1.375 < \|\eta\| < 3.2$
End-cap Hadronic	Copper	LAr	$1.5 < \|\eta\| < 3.2$
Forward EM	Copper	LAr	$3.1 < \|\eta\| < 4.9$
Forward Hadronic	Tungsten	LAr	

in the barrel, and 10 λ_{int} in the end-caps, (including contributions from both the EM and hadronic calorimeters) are adequate to provide good resolution for high energy jets, and limit punch-through [67]. The total thickness, including 1.3 λ_{int} from material upstream of the calorimeter, is 11 λ_{int} at $\eta = 0$.

The ATLAS calorimeters, like the ID, can be separated into two distinct regions: the barrel, and the end-caps. The barrel is situated in the central region of the ATLAS detector, while the end-caps cover the two forward regions to ensure as close to 4π solid angle coverage as possible. In both the end-cap and the barrel region, the EM calorimeters are Liquid Argon (LAr) based sampling calorimeters. Liquid argon was chosen because of its intrinsic properties: radiation hardness, linear response, low cost and uniformity [67]. The hadronic end-cap calorimeter is also LAr based. However, the hadronic calorimeter in the barrel is a sampling calorimeter, with scintillator tiles forming the active region and steel plates as the absorber [68].

The passive and active materials, as well as the coverage for each of the calorimeters is given in Table 3.3 while Figure 3.7 shows the relative number of interaction lengths as a function of $|\eta|$, including material in front of the calorimeters. As argon is gaseous under standard temperature and pressure, the

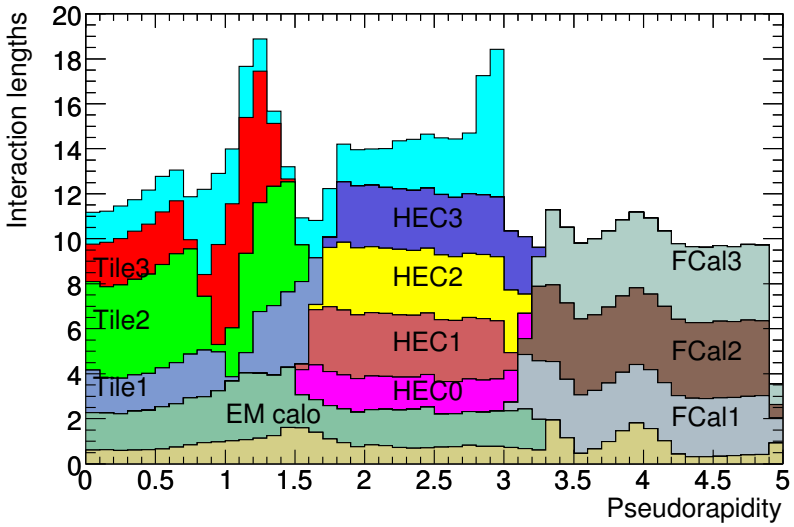


Figure 3.7: Interaction lengths as a function of $|\eta|$ for the ATLAS calorimeters. Also shown is the material in front of the first calorimeter layer (light brown), and uninstrumented material in front of the Muon Spectrometer (light blue). These materials include the LHC beampipe, cryostat/vacuum walls, structural support and the ID [51].

LAr calorimeters are enclosed in cryostats and cooled to an operating temperature of about 88.5 K. The EM barrel is housed inside the same cryostat as the solenoid magnet, while the end-cap and forward calorimeters are enclosed in common end-cap cryostats (one on each side).

3.5.3 Tile Calorimeter

The Tile Calorimeter is responsible for the hadronic calorimetry in the barrel region and is made of three sections: one barrel section placed between two extended barrels. Figure 3.6 shows how the barrel region of the calorimeter is laid out with respect to the extended barrel and the end-cap components. The barrel covers $|\eta| < 1.0$, while the two extended barrel regions cover the range $0.8 < |\eta| < 1.7$.

The Tile Calorimeter is split azimuthally into 64 wedge shaped modules; besides the small cracks between these modules, the Tile Calorimeter is uniform in ϕ. A schematic of one of these wedges can be seen in Figure 3.8. Each wedge consists of steel absorber and plastic scintillator as the active medium, and extends from an inner radius of 2.28 m to 4.25 m. Additionally, the calorimeter is segmented into three concentric readout layers of 1.5, 4.1 and 1.8 λ_{int} in the barrel, and 1.5, 2.6 and 3.3 λ_{int} in the extended barrel sections. The first two layers have a granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$, while the rear-most layer has a granularity of $\Delta\eta \times \Delta\phi = 0.2 \times 0.1$.

The scintillators are read out on both ends by wavelength shifting fibers to two separate photomultiplier tubes (PMTs). Alumininisation of the ends of the fibers opposite the PMTs increases the light output as less light is lost on the dead end of the fiber. These fibers are grouped together to define readout channels, which are arranged to be pseudo-projective, pointing towards the interaction point.

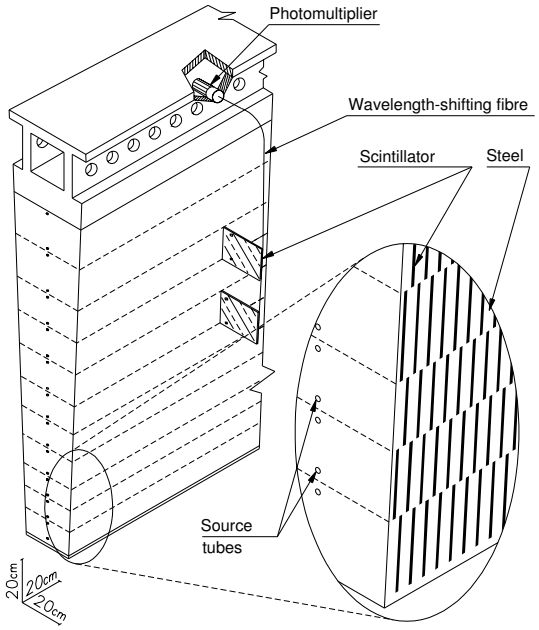


Figure 3.8: Single wedge shaped module from the Tile Calorimeter [69].

3.5.4 Liquid Argon Calorimeters

The Liquid Argon (LAr) Calorimeters are a system of four separate sampling calorimeters. Although their structure and physical layout are quite different, they all use the same active medium of liquid argon, operate under a similar principle, and share a common readout and calibration system. The four LAr Calorimeters are the Electromagnetic Barrel Calorimeter, Electromagnetic End-cap Calorimeter, the Hadronic End-cap Calorimeter and the Forward Calorimeter; each of these subdetectors is described in more detail below.

Electromagnetic Barrel Calorimeter

The layout of the electromagnetic barrel calorimeter (EMB) is unique because of the way the absorber plates are shaped; to avoid azimuthal cracks, the lead absorber plates are accordion shaped. Figure 3.9 shows the EMB under construction with the accordion absorber structure clearly visible. To maintain structural rigidity in the calorimeter, the lead plates are sandwiched between two 0.2 mm layers of stainless steel. The EMB has a minimum depth of 22 X_0, but increases in thickness as seen by a particle originating from the interaction point, with increasing $|\eta|$. At $|\eta| = 0.8$, the thickness of the lead absorber changes from 1.53 mm to 1.13 mm in order to limit the decrease in sampling fraction, which negatively impacts the energy resolution. To maintain the same absorber thickness, and therefore constant LAr gap of 4.5 mm, epoxy is coated onto the thinner lead plates. Although the absorber thickness is decreased, due to the large angle of incidence of particles originating from the interaction point, the distance traversed through the calorimeter will be larger, resulting in an increased number of radiation lengths. At $|\eta| = 0.8$ the thickness is 30 X_0, while at the edge of the barrel calorimeter, the thickness is 33 X_0. The increase in the number of radiation lengths is solely due to trajectory through the calorimeter for particles originating from the interaction point, as the lead absorber thickness is constant between $0.8 \leq |\eta| \leq 1.475$.

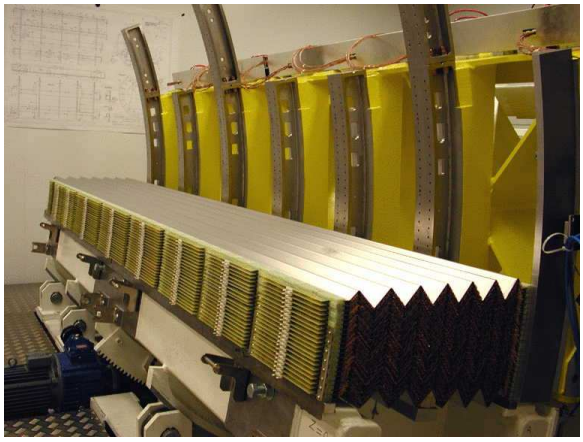


Figure 3.9: Photo of single EMB module under construction, with accordion shaped absorber on display [51].

Figure 3.10: Photo of EMEC wheel under construction. Both inner and outer wheels are visible [51].

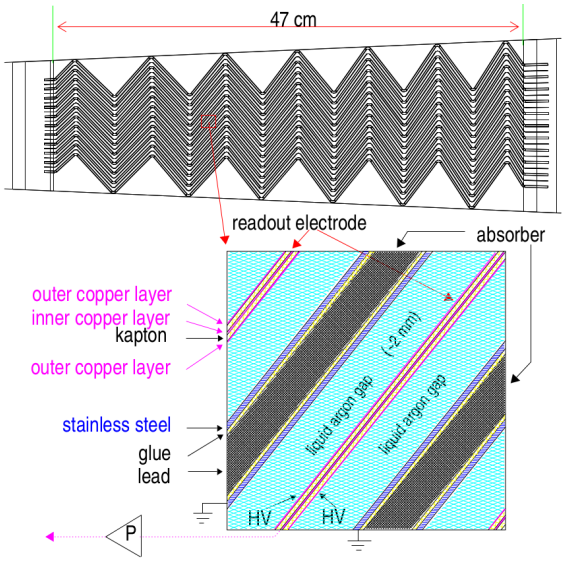


Figure 3.11: Diagram of a radial segment of the LAr EM Calorimeter, with an expanded view of the gap structure and central anode placement [70].

Readout electrodes are placed in the middle of the LAr gaps, insulated and positioned between the absorber plates by a honeycomb array of spacers. This splits the 4.5 mm LAr gap into two 2.1 mm drift regions as seen in Figure 3.11.

The EMB consists of two separate half barrels that are separated at $z = 0$ by a 4 mm gap. Each half barrel is composed of 1024 accordion-shaped absorber plates and is made up of 16 azimuthal modules (one module under construction seen in Figure 3.9). As illustrated in Figure 3.12, the EMB is segmented and read out in three separate layers by depth (labelled front, middle, and back). The middle and back layers are readout from the rear of the EMB, while the front section is readout from the front, closer to the interaction point. The granularity of the EMB (indicated in Figure 3.12) also decreases as a function of depth, the front layer providing a finely segmented description of the lateral shower profile, the middle layer providing the bulk of the energy measurement, and the rear layer which is meant to capture the tail end of EM showers.

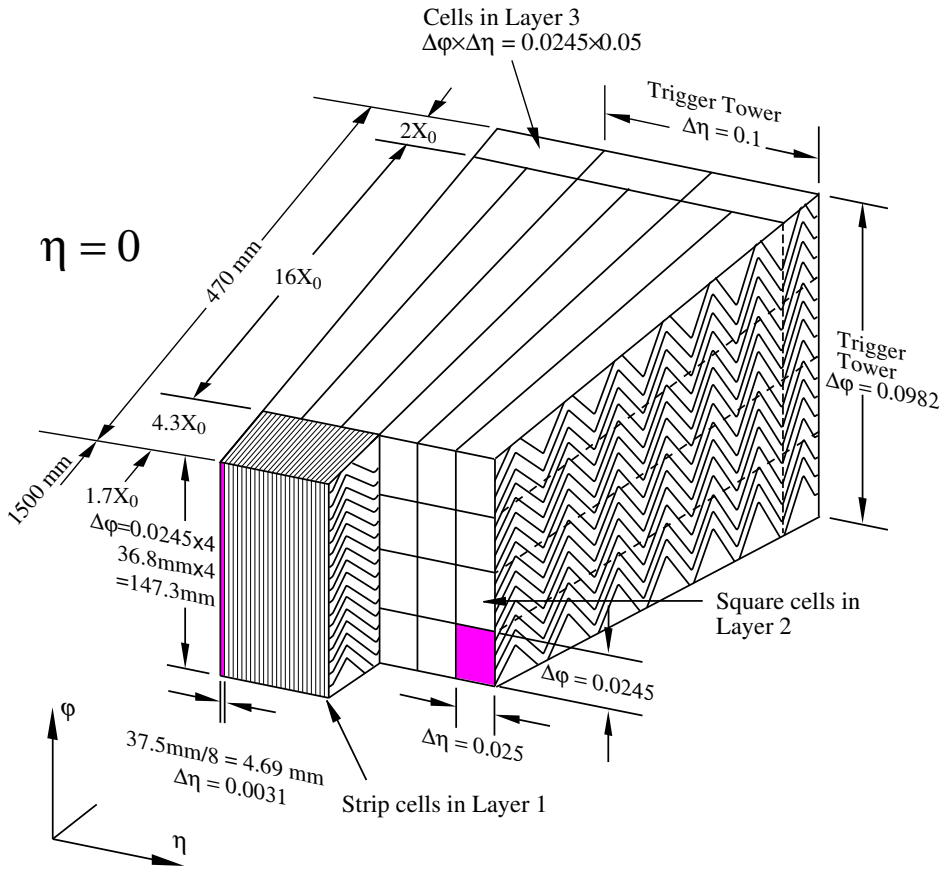


Figure 3.12: Cut-away sketch of EMB displaying $\eta - \phi$ readout granularity in all three layers [67].

Electromagnetic End-cap Calorimeter

All of the end-cap electromagnetic (EM) calorimeters use liquid argon as the active medium, and so to minimize the amount of un-instrumented material between the calorimeters and the interaction point all of the calorimeters are housed in the same cryostat. Each of the two end-cap cryostats (one on either side of ATLAS) contain the EM end-cap (EMEC), hadronic end-cap (HEC) and forward calorimeters (FCal), as shown in Figure 3.13.

Each EMEC is segmented into two coaxial disks, or wheels, with an accordion absorber geometry, similar to the EMB (see Figure 3.10). The orientation of the accordion structure is rotated 90° with respect to that of the EMB (compare Figure 3.10 with Figure 3.9) to ensure full azimuthal coverage. The outer wheel of the EMEC covers the region $1.375 < |\eta| < 2.5$, while the inner wheel covers $2.5 < |\eta| < 3.2$. The minimum number of radiation lengths in the EMEC is 24 X_0, but as there tends to be more energetic particles at high $|\eta|$, the thickness increases to 36 X_0 as the calorimeter extends to its maximum depth at $|\eta| = 3.2$. This is accomplished by changing the thickness of the absorber plates from the inner to the outer wheels. In the outer wheels the absorber is 1.7 mm thick, while the absorbers are 2.2 mm thick in the inner wheel.

Each of the EMEC wheels is composed of eight wedge shaped detectors with accordion shaped absorbers making up the entire wheel shaped detector. Because of the accordion shaped absorbers, this does not introduce azimuthal gaps in the detector. Like the EMB detector, the outer wheel is read out from both sides of the detector. However, as the inner wheel covers higher $|\eta|$ values (and therefore is subject to higher radiation levels) the readout must be done from the back of the detector.

Similar to the EMB, the EMEC outer wheel is divided into three longitudinal readout sections labelled the front, middle and back (relating to their respective distance from the interaction region) with the granularity ($\Delta\eta \times \Delta\phi$) for each layer decreasing with depth. The inner wheel is only segmented into two sections in z, and has coarser granularity than the rest of the EMEC. The electrodes for the EMEC are also similarly structured, with the readout anode positioned in the middle of the LAr gap, creating two drift regions.

LAr Presampler

Ideally, the tracking components in the inner detector of ATLAS have no effect on the energy of the outgoing particles as they deposit negligible energy into them. The materials that make up the inner detector are specifically chosen to minimize the energy deposited, while maintaining the desired tracking performance. However, since quantum mechanical cross-sections are involved, there is always the possibility that a particle can interact with the inner detector material and may shower prematurely. The ID is not the only material upstream of the calorimeters; there are also structural support systems, vacuum walls, cryostats and a solenoid magnet present, which can result in a particle showering before entering the calorimeters. It is therefore important to have a system in place that can detect these premature showers, so that it is known that the particle in question may have lost some portion of its energy before reaching the calorimeter. The presamplers are located directly in front of the EMB, as well as portions of the end-cap calorimeters, with the goal of detecting these preshowers.

The presamplers are LAr based devices and are incorporated into the existing barrel and end-cap cryostats. The barrel presampler covers the entire η range of the EMB while the end-cap presampler covers the range of $1.5 < |\eta| < 1.8$.

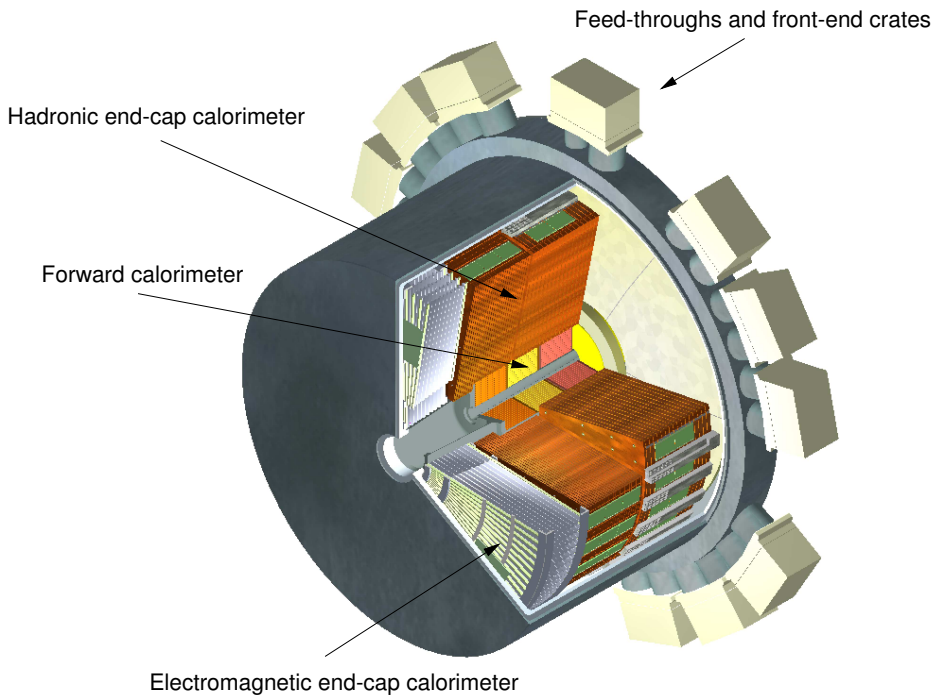


Figure 3.13: Layout and cut-away view of the LAr end-cap cryostat [67].

The barrel presampler has a depth of 11 mm and is azimuthally segmented into 64 identical sectors. Each of these sectors has dimensions of 1.52×0.2 ($\Delta\eta \times \Delta\phi$), and covers the length of an entire half-barrel. In order to make the η-granularity of the presampler constant ($\Delta\eta = 0.2$) as a function of $|\eta|$, each sector is composed of 8 modules which increase in length for higher pseudorapidity ($|\eta|$). The exception is the barrel module that sits on the end of the sector, at the highest pseudorapidity, which has a coverage of $\Delta\eta = 0.12$. The end-cap presampler is constructed from 32 identical azimuthal modules, and has a readout granularity of 0.025×0.1 ($\Delta\eta \times \Delta\phi$).

Hadronic End-cap Calorimeter

Like the EMEC, the HEC consists of two segmented wheels in each end-cap together covering the region $1.5 < |\eta| < 3.2$, located one behind the other in z (a front wheel: HEC1, and back wheel: HEC2). Each of these wheels is made up of 32 wedge shaped modules, and is segmented longitudinally into two sections (HEC1 and HEC2). A drawing of one of these wedge shaped modules is shown in Figure 3.14.

The HEC modules differ from the other LAr subdetectors mentioned so far in two key aspects. First, the absorbers are flat-plates, and second, the absorber is made from copper instead of lead. In the HEC1 module, there are twenty-four 25 mm Cu plates, while the rear HEC module contains sixteen 50 mm thick plates, resulting in a total depth of approximately 10 λ_{int}. In both modules the absorber plates are separated by a 8.5 mm LAr gap, with three electrodes equally spaced in each (similar to the EMB and EMEC), which defines four drift regions of width 1.8 mm. The granularity of the readout cells is pseudorapidity dependent, with a granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ for $|\eta| < 2.5$, and 0.2×0.2 otherwise.

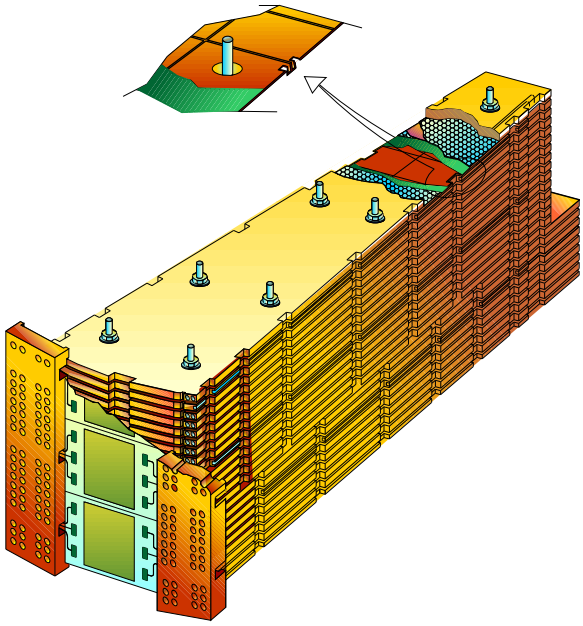


Figure 3.14: Cut-away schematic of single HEC module [67].

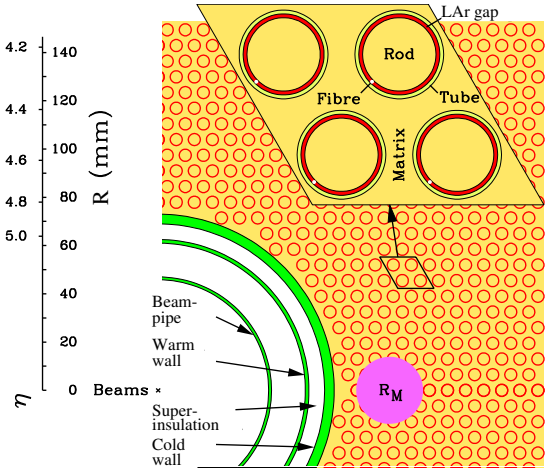


Figure 3.15: Schematic of FCal1 absorber/tube structure. A circle with radius equal to the Molière radius is superposed in purple [67].

Forward Calorimeter

The Forward Calorimeter (FCal) [71] is a combined hadronic and electromagnetic calorimeter. It is cylindrical in shape and is composed of three modules, segmented in z, called FCal1, FCal2 and FCal3 (Figure 3.16). The FCal1 is considered an EM calorimeter while the other two modules are intended for hadronic shower measurements and containment. There is a fourth cylindrical section behind FCal3 that is uninstrumented and made of brass; it acts to further shield the forward muon spectrometers from any energetic showers that leak through the back of the FCal. Because of its location close to the beam pipe, the FCal has to endure much higher flux of particles than the other detector systems positioned at lower $|\eta|$. Because of this larger flux of particles, a typical argon gap of 2 mm is too large. If the FCal had 2 mm LAr gaps there would be a large amount of ion buildup in the gap which would cause a charge screening effect, and limit the detector's response at high luminosity. As a result, the FCal is made with very small gaps.

The FCal signal is readout by cylindrical electrodes (a hollow tube as the cathode, and a rod as the anode), oriented parallel to the beamline. The anode rods fit into hollow tubes, inserted into the main FCal absorber, and are concentrically positioned and electrically isolated from the tube by a PEEK[1] fiber wound around the electrode rods. Figure 3.15 shows a schematic view of the absorber/rod/tube structure, while Figure 3.17 details the FCal LAr gap.

Electrodes in the FCal are not read-out individually, but are summed into *tube groups*. The number of electrodes in a tube group (along with other FCal properties) is listed in table 3.4. Each tube group is summed with three neighbouring groups to form one FCal readout channel. To avoid awkward readout cell shapes, tube groups that sit on the very inner and outer edges of the FCal modules are not summed into groups of four. These readout channels are called *unsummed* channels (as opposed to *summed*

[1] A radiation hard polymer, short for polyetheretherketone.

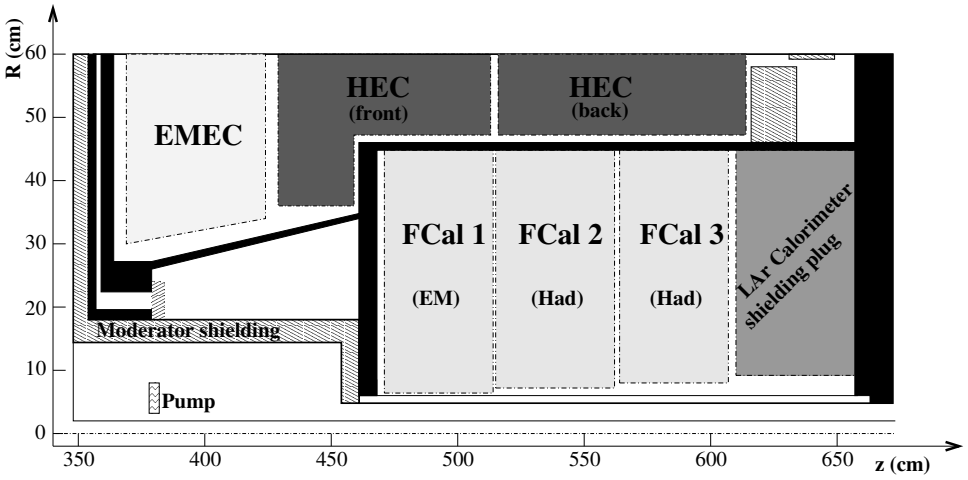


Figure 3.16: Cross-section drawing of the three FCal modules (and shielding brass plug) in the end-cap cryostat. Positioning of other end-cap detectors in relation to FCal also shown [51].

Table 3.4: Properties of the three FCal modules [71].

	FCal1	FCal2	FCal3
Absorber Material	Copper	Tungsten	Tungsten
Electrode Material	Copper	Tungsten	Tungsten
LAr Gap Dimension (mm)	0.269	0.376	0.508
Depth (λ_{int})	2.7	3.7	3.6
Sampling Fraction (%)	1.61	1.33	1.70
Grouped Electrodes	4	6	9
Readout Channels (per side)	1008	500	254

channels), and consist of only a single tube group. This carries the additional advantage that the inner regions of the FCal have a finer readout granularity.

Instead of copper, the absorber material for FCal2 and FCal3 is tungsten. Tungsten was chosen to maximize the number of interaction lengths and limit the transverse size of hadronic showers in the FCal. However, tungsten is a very brittle material and thus difficult to fabricate. Therefore, the main portion of the absorber is not made out of solid tungsten but instead a matrix of tungsten slugs that fit around the copper tubes. Figure 3.18 shows how these slugs fit around the tubes to form a solid matrix.

While the three FCal modules differ in construction materials, gap size, and *tube group* formation, they all share a common readout (along with the rest of the LAr Calorimeter), and calibration infrastructure, and are housed in a common support tube, inside the end-cap cryostats.

3.5.5 LAr Calorimeter Signal and Processing

There are two stages to the readout of the LAr Calorimeter signals, one step occurring directly on the detector itself (done by the *front-end* electronics [72]), and the second (by the *back-end* electronics [73]) which is done in a separate, but nearby, underground service cavern.

The front-end electronics that sit on the ATLAS detector (but outside the LAr cryostats) are responsible for receiving the raw signal from the calorimeter, performing analogue processing, digitizing

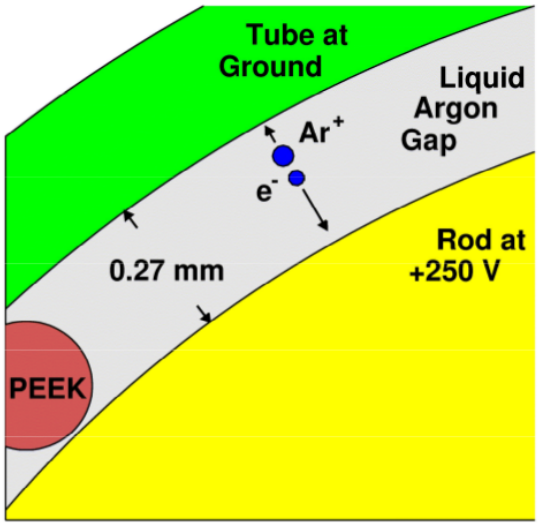


Figure 3.17: View of LAr gap in FCal1 module [71].

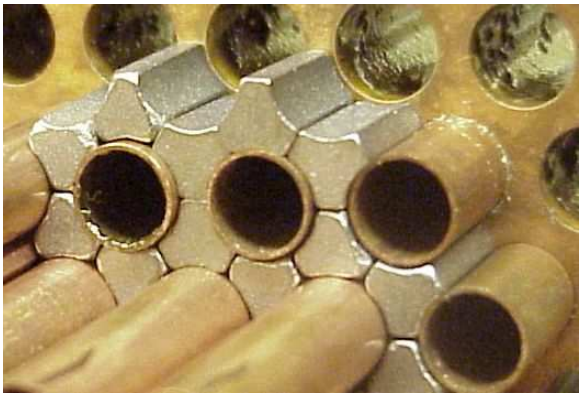


Figure 3.18: Arrangement of tungsten slugs around copper tubes in an FCal mock up [71].

the resulting shaped pulse and then transmitting the resulting signal to the back-end electronics (if an L1 trigger accept is received).

The initial pulse shape from the calorimeter is triangular shaped, with a sharp rise and a slowly falling tail from the drift of the ionization electrons to the electrodes. When the signal arrives at the front-end boards (FEB, part of the front-end electronics) [74], the signal is amplified and an analogue filter is applied which transforms the original triangular pulse shape into a bipolar pulse with the same amplitude but zero net area. Both the triangular pulse shape and the resulting filtered shape are shown together in Figure 3.19. The integration time of the shaping is optimized in order to minimize the overall noise, which is expanded upon below. After the analogue shaping, the pulse is sampled and digitized at the LHC clock frequency of 40 MHz. During typical operating conditions, only 5 of these samples are sent to the back-end electronics. During special runs, or calibrations more samples may be kept, although a lower trigger rate is needed.

The LAr back-end electronics are located in an adjacent underground services cavern, which is shielded from the main experimental cavern. Once the amplified, shaped and digitized samples have been transferred from the FEBs they arrive at the back-end *readout drivers* (RODs). The RODs are responsible for applying the optimal filtering coefficients [75], and performing data integrity checks. The optimal filtering coefficients applied to the samples allows the extraction of the energy in the channel (measured in ADC counts, and then converted to GeV), and the timing of the signal, while minimizing the effects of noise.

Due to the long intrinsic drift time for the calorimeters (∼450 ns for most of the calorimeters at design operating voltage) compared to the time between collisions (50 ns), the readout of multiple collisions and energy deposits will overlap. This feature, known as out-of-time *pile-up*, introduces the dominant source of noise in the calorimeters at design luminosity (discussed further below), and introduces complications when reconstructing objects, which are discussed further in Chapter 4. There is a second component to pile-up called *in-time* pile-up which refers to the actual number of interactions in a specific bunch-crossing, which is also discussed in Chapter 4.

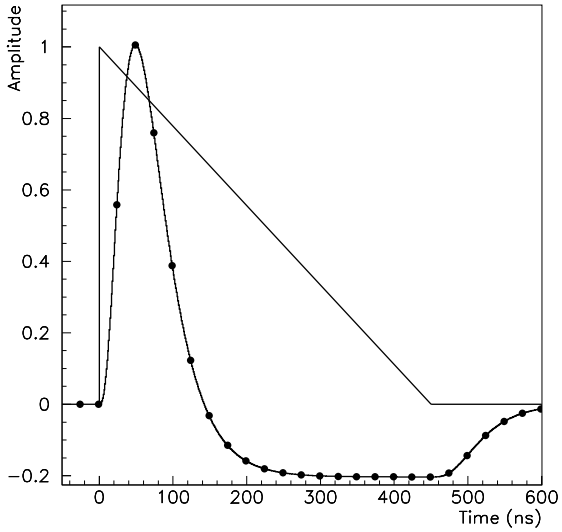


Figure 3.19: Diagram of LAr Calorimeter pulse shape showing current amplitude as a function of time for both the raw triangular pulse, and the bipolar shaped pulse. The solid points on the shaped pulse indicate the 25 ns sampling that is performed [51].

Pulse Shaping and Noise

There are two sources of noise which are of importance when discussing the total noise in the LAr Calorimeters [67]. The first source is the thermal noise generated by the readout electronics, and the second is the *pile-up* noise, which is proportional to the instantaneous luminosity, and was the dominant contribution during the 2012 *pp* run. The pile-up noise is a direct result of the multiple calorimeter waveforms which overlap as a result of the short colliding bunch separation. The bipolar pulse shaping was chosen in order to minimize the effect of these overlaps, with the average deposit of energy for minimum bias events summed over many events equal to zero. However, this cancellation is not perfect, and while the average reconstructed energy will peak at zero, the distribution will have some non-zero width; this width represents the pile-up noise.

The relative contributions of these sources of noise are determined by the time constant used in the bipolar shaping on the FEBs. The electronics noise increases as the integration time decreases, while the pile-up noise will decrease as the integration time decreases (as this will permit a fewer number of colliding bunches to impact the waveform). The optimal time constant for the signal shaping can be calculated by evaluating the pile-up noise and electronics noise as a function of integration time. The optimal integration time is the value which gives the lowest overall noise, which will be instantaneous luminosity dependant [67].

3.6 Muon Spectrometer

The Muon Spectrometer (MS) is made up of the monitored drift tubes (MDT), cathode strip chambers (CSC), resistive plate chambers (RPC) and thin gap chambers (TGC) [51, 76]. A drawing of the MS, with the subdetectors labelled is shown in Figure 3.20. Each of these subdetectors is used for slightly different purposes, and cover different regions in $|\eta|$. As a group, they provide muon reconstruction up to $|\eta| = 2.7$ and triggering up to $|\eta| = 2.4$. The main properties of the MS subdetectors can be found in Table 3.5.

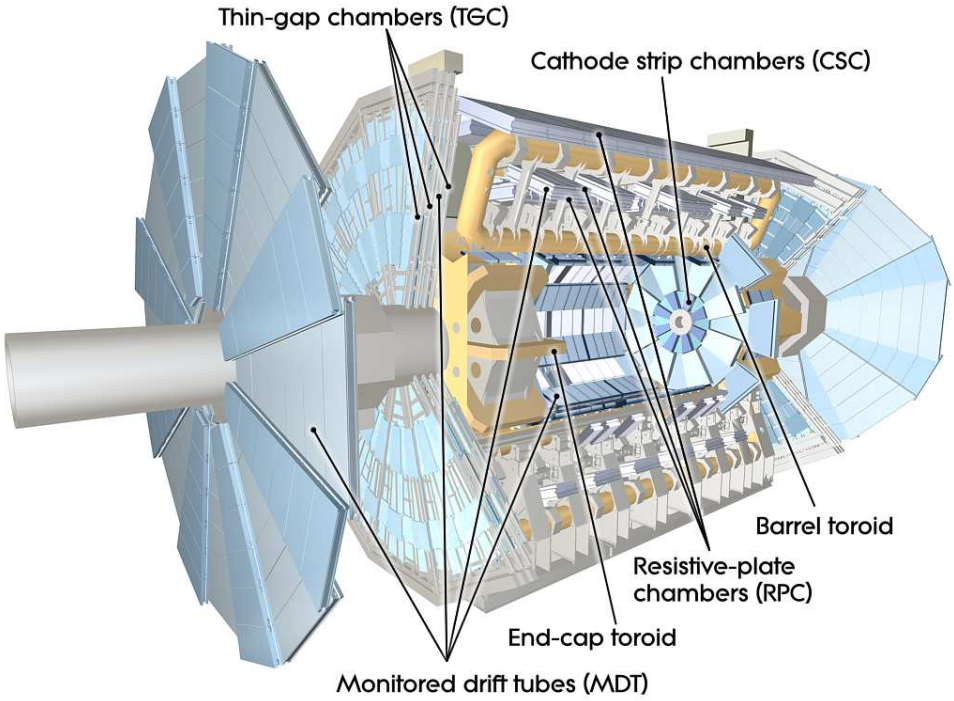


Figure 3.20: The ATLAS Muon Spectrometers [51].

Much like the ID, the MS is a tracker, but is located on the very outside of the ATLAS detector[1]. The MS functions by collecting a series of space points, which can then be used to reconstruct the muon's trajectory. The momentum and polarity can also be reconstructed based on the bending of the track through the magnetic field. This magnetic field is supplied by air-core toroidal magnets in both the end-caps and the barrel which provide an azimuthal magnetic field and bending in the η coordinate. The magnet systems are described in further detail in Section 3.7.

For the central region of the ATLAS detector, the muon detectors are arranged about the z-axis in 3 cylindrical layers (at $R \approx 5$, 7.5 and 10 m). In the end-cap region of ATLAS, there are four instrumented tracking stations per side, mounted perpendicular to the z-axis in a disk geometry (at $|z| \approx 7.4$, 10.8, 14 and 21.5 m).

The RPCs, along with the TGCs, act as part of the triggering system (see Section 3.8 for more details), and therefore focus on providing fast response, as opposed to fine spatial granularity. The RPCs and TGCs cover the range $|\eta| < 2.4$, with the RPCs being responsible for coverage in the barrel region ($|\eta| < 1.05$), and the TGCs for the end-caps ($1.05 < |\eta| < 2.4$). While the RPCs and TGCs are not designed for precision position measurements, they are able to supplement the information from the precision detectors. The MDTs are only able to reconstruct the muon trajectory in the bending coordinate, and therefore, the secondary (non-bending) coordinate is solely provided by the triggering chambers.

[1]Only muons are typically capable of fully traversing the ID and calorimeters.

Table 3.5: Summary of properties of the ATLAS muon spectrometer [51].

Monitored drift tubes (MDT)	
Coverage	$\|\eta\| < 2.7$ (innermost layer: $\|\eta\| < 2.0$)
Number of chambers	1150
Number of channels	354000
Function	Precision tracking
Cathode strip chambers (CSC)	
Coverage	$2.0 < \|\eta\| < 2.7$
Number of chambers	32
Number of channels	31000
Function	Precision tracking
Resistive plate chambers (RPC)	
Coverage	$\|\eta\| < 1.05$
Number of chambers	606
Number of channels	373000
Function	Triggering, secondary coordinate
Thin gap chambers (TGC)	
Coverage	$1.05 < \|\eta\| < 2.7$ (2.4 for triggering)
Number of chambers	3588
Number of channels	318000
Function	Triggering, secondary coordinate

3.6.1 Monitored Drift Tubes

The MDT chambers provide coverage over most of the MS ($|\eta| < 2.7$) and are responsible for precision measurements of the muon trajectory. Each chamber is made up of 3 cm diameter aluminum drift tubes, with a central anode readout wire, which are arranged into two distinct sub-layers. Each of these sub-layers contains three to four layers of drift tubes. The length of the chambers/tubes is between 0.9 m and 6.2 m depending on the chamber's position within ATLAS.

The individual drift tubes are filled with a Ar/CO_2 gas mixture at 3 bar absolute pressure. Each tube has a position resolution better than 100 μm, and each MDT chamber has a resolution of 35 μm in the bending coordinate.

The MDT chambers are present in all three layers of the barrel MS, and the three outer layers of the end-cap. The inner-most end-cap station uses CSCs to cope with the higher particle fluence present in that region of ATLAS.

3.6.2 Cathode Strip Chambers

The cathode strip chambers are responsible for precision tracking solely in the inner-most end-cap station between $2 < |\eta| < 2.7$. Here, the particle fluence is significantly higher than in the barrel or outer end-cap stations. As a result of the higher occupancy, the MDTs cannot be used (MDTs are only able to function up to a rate of 150 Hz/cm^2), and the CSCs are employed (which can withstand up to 1000 Hz/cm^2).

The CSCs are multi-wire proportional counters (MWPC) with 2 orthogonal cathode planes (one on each side of the anode plane), and are thus able to measure two spatial coordinates (as opposed to

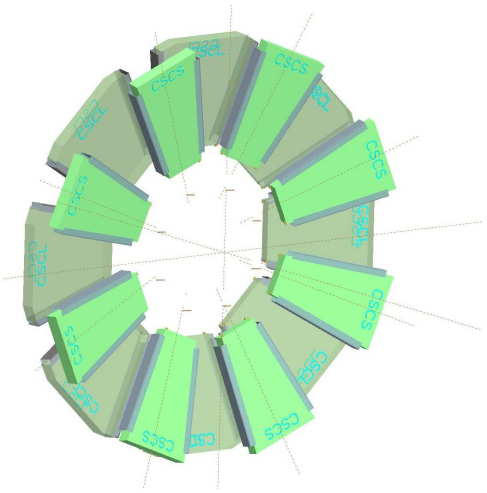


Figure 3.21: Layout of CSC chambers that make up the muon small wheel. Large and small chambers are visible in dark and light green, respectively [51].

the MDT). The CSC chambers are arranged azimuthally into two disks (making up one wheel), each containing eight chambers. One disk contains only large chambers, the other, only small chambers. The geometry of the CSC disks is shown in Figure 3.21. Each CSC chamber is made up of four CSC planes, allowing four space point measurements for every incident track.

The resolution of the CSC is 60 μm per CSC plane in the bending coordinate, and 5 mm in the non-bending coordinate.

3.6.3 Resistive Plate Chambers

Like the MDTs in the barrel, the RPCs are positioned in three concentric cylindrical layers, providing full azimuthal coverage about the beam axis. Figure 3.22 shows how the three layers are positioned in relation to the three MDT layers. The spacing of the chambers allows effective triggering for both high and low p_{T} muons.

The RPCs are simple avalanche detectors: one gas unit consists of parallel electrode-plates (separated by a 2 mm gap) filled with a $C_2H_2F_2$ (94.7%), Iso $- C_4H_{10}$ (5%), SF_6 (0.3%) gas mixture.

Each layer of RPC chambers contains two independent detector layers, and measures the muon trajectory in both the η and ϕ coordinates through readout strips with pitches varying between $25 - 35$ mm. Although primarily meant for triggering, the RPCs also supplement the MDT position information, which only provides the coordinate in the magnetic field bending plane.

3.6.4 Thin Gap Chambers

The TGCs provide triggering in the muon end-cap region as well as a complementary measurement of the space points for the end-cap MDTs.

The TGCs are MWPCs that have a characteristically small gap between the anode wire and cathode (1.4 mm vs 1.8 mm between adjacent anode wires). The cathode wires are read out in groups and can provide the radial coordinate, while the azimuthal coordinate is provided by radial readout strips coupled to the cathode. The number of anode wires that are grouped is $|\eta|$ dependant in order to optimize granularity.

The TGCs are stationed on the inner MS station, and the outer-middle station, with two layers of TGCs on the former, and seven on the latter.

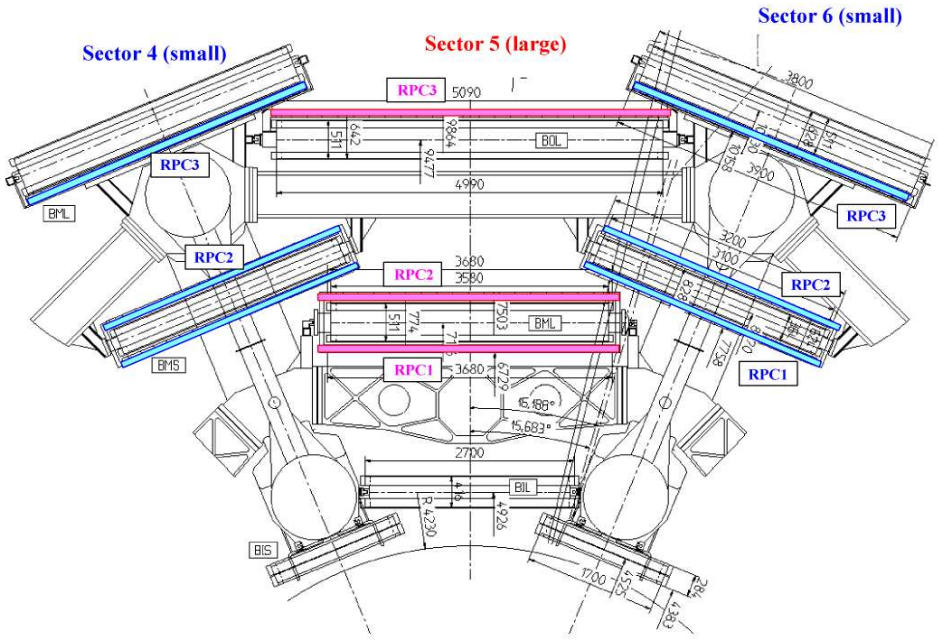


Figure 3.22: Placement of RPCs (blue and pink) in relation to barrel MDTs. The outer-most MDT layer has a single RPC placed on the inner surface of the chamber, while the central MDT chambers have RPCs on either side. The there are no RPCs stationed on the inner-most MDT layer [51].

3.7 Magnet Systems

ATLAS features four large scale liquid He cooled superconducting magnet systems in order to provide bending power for the inner detector and muon spectrometer [51]. One solenoid magnet [77] provides a magnetic field in inner detector, while three air-core toroid magnet systems (two on either end-cap and one for the barrel region) provide for the muon spectrometer. A computer generated image of the complete magnet system can be seen in Figure 3.23.

3.7.1 Solenoid Magnet

The solenoid measures 5.8 m long, is 0.10 m thick, with an inner diameter of 2.46 m and is designed to provide a 2 T axial magnetic field. In order to not greatly impact the resolution of the calorimeter systems sitting directly behind it, the solenoid was designed to contribute only $\sim$0.66 X_0 of material. This is in part achieved by the solenoid and LAr calorimeters sharing a common vacuum wall and the use of an aluminum stabilized superconductor, which allows the construction of a high field magnet with low weight/thickness [78]. There are 222 windings/m of NbTi/Cu superconducting cable, cooled to 4.5 K, supported by an aluminium structure. This support structure also serves to help cool the superconducting windings.

3.7.2 Toroid Magnets

Each of the three toroid magnets consists of eight coils symmetrically arranged in ϕ about the z axis. In contrast to the solenoid magnet, the toroid magnets have a magnetic field that points in the azimuthal

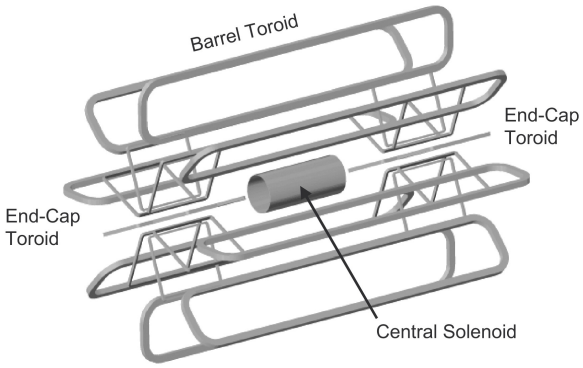


Figure 3.23: Layout of ATLAS magnet system. All three system components are shown: solenoid, barrel toroid and end-cap toroids [77].

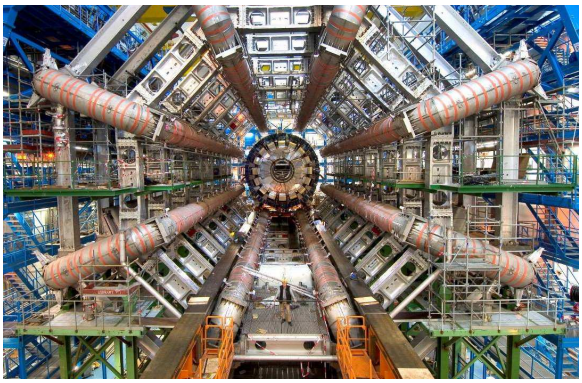


Figure 3.24: Photo of ATLAS during construction with barrel toroid magnets on prominent display [51].

direction. Like the solenoid magnet, the toroid magnets are constructed from NbTi/Cu superconducting cable windings. To ensure good magnetic field coverage between the barrel and end-cap toroids, the end-cap toroid coils are offset by 22.5° in ϕ with respect to the barrel toroid coils [51]. There is some overlap between the magnetic fields of the two toroid magnet systems in the range $1.4 < \eta < 1.6$. Figure 3.24 shows the exposed barrel toroid coils during construction of ATLAS. The toroid magnets provide a ≈ 0.5 T field in the barrel region and a ≈ 1.0 T field in the end-caps.

The coils for the barrel toroid are each contained in their own stainless steel cryogenic compartment (clearly visible in Figure 3.24), while each end-cap toroid has all of its eight coils in a common cryogenic compartment cooled to 4.6 K. The overall size of the installed barrel toroid system is 25.3 m in length, with an outer diameter of 20.1 m (inner diameter is 9.4 m).

For $|\eta| < 1.4$ the bending power is achieved solely through the use of the barrel toroid while the for the range $1.6 < |\eta| < 2.7$ the bending comes from the end-cap toroid magnets. In the intermediate range, $1.4 < |\eta| < 1.6$ (transition region), the deflection is provided by the magnetic field from both magnets.

3.8 Trigger and Data Acquisition

At LHC design luminosity, $\mathcal{L}_{\text{inst}} = 10^{34}$ cm^{-2}s^{-1}, the approximate collision rate inside ATLAS is 40 MHz. However, the recording rate, based on the current technology is only 400 Hz. Therefore, ATLAS uses a sophisticated trigger and data acquisition system (TDAQ), in order to select the events recorded, based on the physics contained in the events.

The ATLAS trigger is a three level system. The Level-1 (L1) trigger is custom hardware based, and is responsible for reducing the data rate to 75 kHz. The events that pass the L1 trigger are then sent to the Level-2 trigger (L2) system and then, if they pass L2, to the event filter (EF). The latter two stages of the trigger, EF and L2, are collectively known as the high-level trigger, or HLT. These two trigger stages further bring the rate down from 75 kHz to the final write-to-disk rate of about 400 Hz. Each trigger level refines the information from the preceding level, and can apply stricter selection criteria.

The L1 trigger, selects events that have high p_{T} objects in the calorimeters or the MS, as well as events that contain large transverse energy, or a large imbalance in the transverse momentum. To do

this, it uses information from the calorimeters, RPCs, and TGCs, with lower granularity than the final readout. The L1 trigger has approximately 2.5 μs to make a trigger decision before it either passes the event on to the HLT, or discards it. The L1 trigger also defines *regions of interest* (RoI), which are the regions of the detector (in $\eta \times \phi$) in which any of the L1 trigger thresholds have been surpassed; these RoI are then passed on to the HLT.

If the L1 trigger decision is to retain the event, the digitized data from each readout channel are transferred to the detector subsystem's readout drivers (RODs) for formatting, while the RoI is passed to the L2 trigger. The L2 trigger takes information from the RoI identified by the L1 trigger and applies stricter selection criteria than L1 in order to further reduce the rate ($\sim$40 ms decision time). The L2 trigger, while only looking at a fraction of the detector (determined by the ROIs), is able to use the full detector granularity in order to make a decision.

Finally, the EF fully reconstructs the event using algorithms similar to those used in the offline event reconstruction, taking all detector systems into account, and using the full detector granularity. The EF then selects events at the final rate of approximately 400 Hz ($\sim$4 s processing time).

Assuming that the event passes all of the three trigger levels, the data are then permanently stored.

Chapter 4

Dataset and Analysis Objects

> "When asked "What do we need to learn this for?" any high-school teacher can confidently answer that, regardless of the subject, the knowledge will come in handy once the student hits middle age and starts working crossword puzzles in order to stave off the terrible loneliness."
>
> David Sedaris, *Me Talk Pretty One Day*

This chapter discusses the *pp* dataset and the reconstructed objects used in the analysis (which is presented in Chapter 5). The object definition and selection described in this chapter only form the basis for definitions and requirements used in the analysis, with the further (more stringent) requirements discussed in Section 5.5.1. As mentioned in Chapter 1, one of the features of this analysis channel is the large variety of reconstructed objects – electrons, muons, missing energy and two different types of reconstructed jets – which demands a detailed understanding of all the detector subsystems.

4.1 Dataset

The dataset used is the entire 2012 *pp* dataset delivered by the LHC, and recorded by ATLAS. This amounts to 20.3 fb^{-1} of data designated as *good for physics*[1]. The plot showing the LHC delivered and ATLAS recorded integrated luminosity as a function of day in 2012, with a *good for physics* overlay, is shown in Figure 4.1.

4.1.1 Pile-up

In the press, the LHC receives praise for being the highest energy collider in the world, but this overlooks another feature that makes it a great "discovery machine": the high instantaneous luminosity. In order to achieve the high luminosity at the LHC in 2012, the time between the colliding bunches in ATLAS is a very short 50 ns, and the number of *pp* interactions for each bunch crossing is large.

[1] What determines if data is *good for physics* is described in Section 4.1.2.

The features and complications (from the standpoint of both the detector instrumentation and object reconstruction) that arise from the short time separation between bunch crossings, and the result of the multiple interactions per bunch crossing are collectively referred to as *pile-up*. The component of pile-up that originates from the short time between bunch crossings is referred to as *out-of-time pile-up*, while *in-time pile-up* is related to the number of *pp* interactions per bunch crossing. For the 2012 *pp* dataset, the average number of interactions per bunch crossing, $\langle\mu\rangle$, was 20.7 (compared to 9.1 in the 2011 *pp* dataset). Figure 4.2 shows the profile of in-time pile-up over 2012 (along with 2011, for reference).

Both in-time and out-of-time pile-up present experimental difficulties that impact the detector response, and object reconstruction. The out-of-time component heavily affects detectors which have a relatively long response time (on the order of the bunch spacing, or longer), such as the LAr Calorimeters where the pulse shape is ≈ 600 ns long. A result of the slow response is that the long pulse shape can be deformed by energy deposited by preceding and subsequent collisions, as discussed in Section 3.5.5. A result of this feature of the calorimeters is that the energy read out will be dependant on both the average number of collisions per bunch crossing, and the actual number of collisions in the event. In the case of in-time pile-up, each hard-scatter event is accompanied by, on average, 20 other softer *pp* collisions. The additional collisions produce additional tracks in the ID, and showers in the calorimeter, which negatively impacts the reconstruction of objects which originate from the *pp* hard-scatter under investigation.

4.1.2 Good Runs List

As is visible in Figure 4.1, not all of the delivered luminosity was recorded by ATLAS, or used for physics analysis. There are many possible faults in ATLAS that can prevent data from being recorded or used for physics analysis. The fault can be a result of large-scale hardware problems in one of the detector sub-systems, to transient problems with the data acquisition and trigger. There are also a subset of detector systems that are only operated after the LHC has declared stable beam conditions (this turn on is not instant, resulting in the loss of some delivered luminosity). Any of these reasons can prevent ATLAS from recording data from a particular sub-system, or being able to record data at all. However, even if the data are recorded successfully, there are a number of smaller problems that may prevent the data from being used in a physics analysis. Issues such as the magnets not delivering the correct magnetic field, part of a sub-detector being temporarily off, or very noisy readout cells can prevent data collected from being declared *good for physics* for the duration of the fault.

Each event recorded by ATLAS is associated with a collection known as a luminosity block, or *lumiblock*, each containing events recorded during a 60 s consecutive interval (there are some cases where the lumiblock is incremented, regardless of how much time has elapsed). The lumiblocks for which there are no unrecoverable faults in any of the detector sub-systems are *good for physics* and are kept documented in a *good runs list* (GRL). Due to the fact that this analysis uses a wide array of the detector subsystems to make the measurement, only lumiblocks with all detector sub-systems operating without an unrecoverable fault are used.

Table 4.1 breaks down the percentage of data recorded that is *good for physics* by sub-system component. The total percentage of recorded data that is declared as all good for physics is 95.5%.

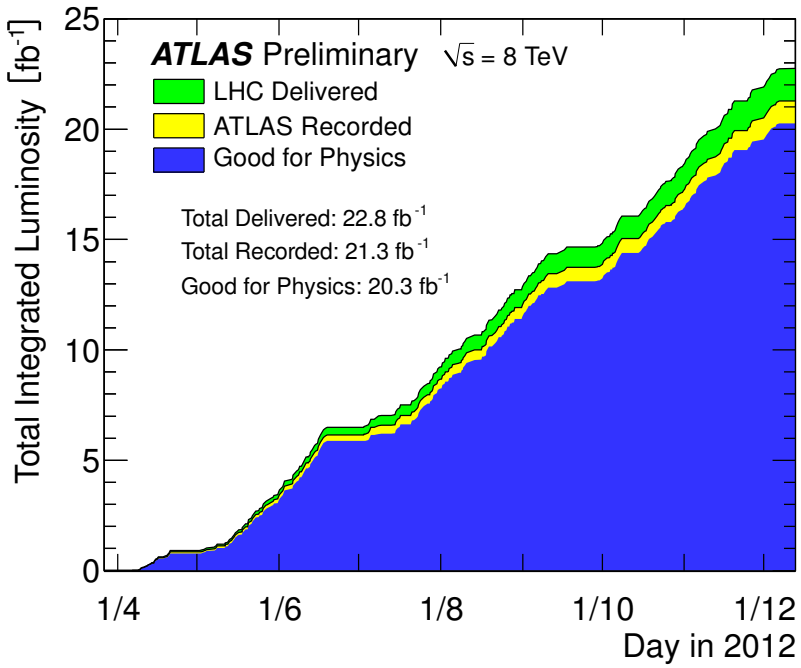


Figure 4.1: Luminosity delivered, recorded, and *good for physics* (*pp*), by day, in 2012 (ATLAS specific) [79].

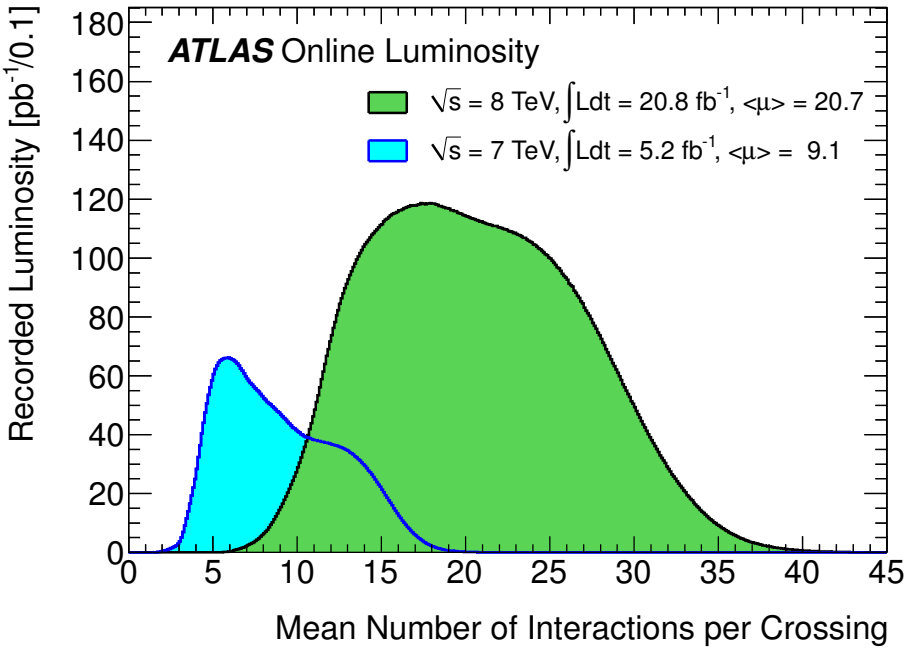


Figure 4.2: Average number of interactions per bunch crossing (pile-up, μ) for the 2011 and 2012 run. Although only data collected during the 2012 run, the 2011 pile-up conditions are also displayed as a reference [79].

Table 4.1: Percentage of data recorded that is classified as *all good for physics*, broken down by detector sub-system in 2012 [80].

Inner Detector			**Calorimeters**		**Muon Spectrometer**				**Magnet Systems**	
Pixel	SCT	TRT	LAr	Tile	MDT	RPC	CSC	TGC	Solenoid	Toroids
99.9	99.1	99.8	99.1	99.6	99.6	99.8	100	99.6	99.8	99.5
All good for physics: 95.5%										

4.2 Triggers

The analysis presented in this thesis uses single lepton triggers in order to select events with $W \to \ell\nu$ decays (see Section 3.8 for a discussion of the ATLAS trigger and DAQ systems). Single lepton triggers are used to trigger on the leptonically decaying W instead of jet triggers (triggering on the hadronic W), because of their relative high efficiency (even down to low p_{T}) and the lower rate of lepton backgrounds. This is in contrast to the jet based trigger, which would accept events from one of the many "jet-y" backgrounds possible. The triggers used are summarized in Table 4.2, and described further below.

Table 4.2: Triggers used for the analysis.

Triggered Object	**Trigger Name**
Electron:	`e24vhi` || `e60`
Muon:	`mu24i` || `mu36`

Some of the single lepton triggers demand that there is track isolation in the ID about the triggering object. To measure the degree of track isolation, the scalar sum of the transverse momentum ($\sum p_{\mathrm{T}}$) for all tracks within a cone of radius R, centered on (but excluding) the object, forms the quantity p_{T}^{iso}. The ratio of p_{T}^{iso} to the triggering object's p_{T} (or E_{T}) then provides a measure of how isolated the object is. Only tracks with $p_{\mathrm{T}} > 1$ GeV are included in the calculation of p_{T}^{iso}.

The specific definition and description of each of the four triggers used is provided below.

- `e24vhi`: An electron trigger with an E_{T} threshold of 24 GeV, and *medium* electron ID requirements (discussed in Section 4.4). The `vhi` in the name refers to further requirements of the trigger:
  - `v`: The trigger compensates at the L1 stage for the different amount of uninstrumented material in front of the calorimeters, which is not uniform in η, by adjusting the triggering threshold as a function of η. The threshold changes by ≤ 2 GeV over the full η acceptance of the trigger.
  - `h`: Requires the triggering candidate's energy deposit in the hadronic calorimeter (behind the EM calorimeter energy deposit) to be below a threshold[1].
  - `i`: A track isolation requirement of $p_{\mathrm{T}}^{iso}/E_{\mathrm{T}} < 0.1$ within a cone of radius $R = 0.2$, centered on the candidate object, is applied.

[1] Even energetic electrons should deposit the large majority of their energy in the EM calorimeter alone.

- `e60:` An electron trigger, with an E_{T} threshold of 60 GeV, and *medium* identification requirements. Although not indicated in the name, this trigger also applies the `v` and `h` requirements of the lower threshold electron trigger, described above, but not the isolation requirement.

- `mu24i:` A muon trigger, with p_{T} threshold of 24 GeV. Like the electron triggers, the `i` denotes that a track isolation requirement is applied to the candidate object. For muon triggers, the isolation required is $p_{\mathrm{T}}^{iso}/p_{\mathrm{T}} < 0.12$ within a cone of radius $R = 0.2$.

- `mu36:` A muon trigger, with a p_{T} threshold of 36 GeV. No isolation requirement is applied.

Events are kept for analysis if any of the triggers selects the event, which maintains triggering sensitivity throughout a large kinematic range.

To measure the trigger efficiency for low p_{T} muons ($p_{\mathrm{T}} < 10$ GeV), the tag-and-probe method is used on J/ψ meson di-muon decays, while the intermediate p_{T} range ($10 \leq p_{\mathrm{T}} \leq 100$ GeV) utilizes tag-and-probe on $Z \to \mu\mu$ events. Above 100 GeV, $W \to \mu\nu$ decays triggered by the ATLAS missing transverse momentum triggers are used to evaluate the trigger performance [81].

In the tag-and-probe method using offline reconstructed di-muon events, one of the reconstructed muons triggers the event (the tag), while the other muon serves as the probe, allowing measurement of the trigger efficiency at the three trigger stages (L1, L2, EF). The probe muon can be associated (with a ΔR matching requirement) to an *online* muon, identified by the trigger software. The fraction of the reconstructed probe muons that overlap with a muon identified by the trigger defines the trigger efficiency. The same procedure is used to determine the efficiency of the electron triggers, using $Z \to ee$ events.

4.2.1 Trigger Scale Factors

The Monte Carlo samples used to model the signal and background in this analysis do not perfectly model the efficiency of the triggers used. As a result, an ATLAS software package applies scale factors to MC events in order to reproduce the observed behaviour of the trigger. These scale factors are used for all MC datasets analyzed in this thesis.

4.3 Clustering Algorithms in ATLAS

The first step in reconstructing an object in the calorimeter – whether it be a photon, electron or jet – is defining clusters of calorimeter cells (or, just *clusters*). There are two types of clusters used by ATLAS and in this analysis. The first type of clusters are generated with a sliding-window algorithm, and have a fixed size and shape[1], and rely on supporting information from the trackers to reconstruct electrons, photons and converted photons. The second type, topological clusters, have a variable cluster size and shape, and are used in the first step of jet reconstruction, as well as for electron and photon reconstruction outside the acceptance of the ATLAS trackers. For this analysis, the sliding-window clusters are used to reconstruct electrons, while the jets are reconstructed from topological clusters.

[1] The size/shape of the cluster depends on the type of object being reconstructed, but is identical for all objects of the same type.

4.3.1 Sliding-Window Clustering

To generate the sliding-window clusters [82] used for electron reconstruction, the EM calorimeters[1] are divided into cells of size $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$. The energy in all longitudinal layers of the calorimeter is summed into one common energy per cell, forming a *tower* and an associated *tower* energy, in a process called *tower building*.

From the constructed towers, a sliding-window algorithm takes over, attempting to find seeds for a cluster. The seed finding algorithm examines all possible groups of towers in a fixed size window, until it finds a group that gives a local maximum in the transverse energy sum (ΣE_{T}) of the towers in the window. If the group forming a maximum satisfies $E_{\mathrm{T}} > E_{\mathrm{T}}^{\mathrm{thres.}} = 2.5$ GeV, the group of towers in the window forms a seed for a cluster, with the position of the seed being layer dependant and defined by an energy-weighted average of the cells in each calorimeter layer. The dimensions of the window used to find the seeds is defined by an integer number of adjacent towers along the η and ϕ coordinates, $N_{\eta}^{cluster} \times N_{\phi}^{cluster} = 3 \times 5$.

The final cluster is constructed from a group of towers ($N_{\eta}^{cluster} \times N_{\phi}^{cluster}$) centered on the seed, with the size of the cluster dependant on whether the seed is in the barrel or end-caps. In the EMB, the cluster size is $N_{\eta}^{cluster} \times N_{\phi}^{cluster} = 3 \times 7$, while for the EMEC it is 5×5. The cluster size is relatively large in ϕ (the solenoid magnet bending coordinate) for clusters in the EMB in order to capture any bremsstrahlung, or other emissions radiated as the electron traverses the material in front of the calorimeters. In the end-caps, the bending effect from the solenoid's magnet field is smaller, therefore, the cluster size does not have to be as large in ϕ. Instead, the size of the cluster in η is larger to account for the smaller cell size at higher pseudorapidity [82].

4.3.2 Topological Clustering

Topological clusters [82], unlike clusters created with the fixed-window algorithm, do not have a fixed size or shape. They are formed by grouping together nearby cells that have a significant energy deposit (in relation to the noise of the cell).

There are three signal to noise thresholds defined for topological clusters – t_{seed}, $t_{neighbour}$, and t_{cell} – which satisfy the relation $t_{cell} \leq t_{neighbour} \leq t_{seed}$ (in practice $t_{neighbour} < t_{seed}$). Clusters are seeded by cells that have a signal to noise ratio, t, greater than t_{seed}, with the signal being defined as the absolute energy of a cell, $|E|$. This definition of the signal is used in order to symmetrize the contribution from clusters that are created purely from noise in the calorimeters.

Starting with one of the identified seed cells, the proto-cluster is built outward iteratively, adding all cells neighbouring the proto-cluster that satisfy $t \geq t_{neighbour}$ (in both the EM and hadronic calorimeters). Once there are no more cells that can be added in this way, all the cells bordering the proto-cluster, with $t \geq t_{cell}$ are added to the cluster. This final step is done only once, **not** iteratively.

The topological clusters in ATLAS used for this analysis are *Had 420* clusters, which are intended to reconstruct hadronic objects, which have a larger more diffuse shower shape compared to EM objects. The numbers in the name denote the value of t_{seed}, $t_{neighbour}$, and t_{cell} thresholds, respectively. A cut of $t_{cell} = 0$ indicates that the cell will be added to the cluster regardless of the signal to noise ratio, which results in **all** cells neighbouring the proto-cluster to be included in the final step.

[1] This algorithm is only used to reconstruct electrons and photons, which typically only deposit energy in the EM calorimeters.

Cluster Splitting

As a result of the large number of interactions per bunch crossing (Section 4.1.1, Figure 4.2), there is a large probability that two distinct, but nearby energy deposits will be merged into a single cluster as a result of unrelated energy deposits between the cluster seeds (this effect grows with increasing $|\eta|$). Therefore, after the topological clusters are constructed, they are split by identifying *local maximum* cells within the cluster that meet the following requirements:

- $E > 500$ MeV
- Energy greater than any neighbouring cell
- The cell has at least four neighbouring cells which are part of the same cluster.

If a cluster contains more than one such local maximum, it is split so that each resulting cluster has at most one local maximum. Clusters with only a single local maximum (or none) are untouched by the splitting algorithm. Cells in regions of the calorimeter where large energy deposits are unlikely, such as the LAr presampler, may be excluded from being defined as local maximum in order to minimize the number of fake clusters.

4.4 Electron Reconstruction and Identification

This section explains how central electrons are reconstructed. Forward electrons ($|\eta| > 2.47$) outside the acceptance of the ID are not used in this analysis.

After the EM cluster seeds have been defined, as discussed in Section 4.3.1, the process of attempting to reconstruct the cluster seed as an electron begins [83, 84].

The collection of cells making up the cluster seed is required to pass modest shower shape requirements which place cuts on the maximum fraction of energy allowed in each of the EM calorimeter layers. There are also loose requirements on the transverse shower shape and hadronic leakage. The cluster seed is also required to have at least one associated track, with $p_{\mathrm{T}} > 1$ GeV, close to its center (in η and ϕ coordinates) when extrapolated to the middle layer of the EM calorimeter [84].

In the case that a cluster seed passes these selection criteria, it is regarded as an electron candidate, and an electron cluster is built around it (described in Section 4.3.1). The four-momentum of the electron is calculated using energy of the cluster (from the calorimeters), and the trajectory of the track (given by the ID).

The reconstruction efficiency of electrons in ATLAS is required to be high; for electrons with at least $E_{\mathrm{T}} = 15$ GeV, the efficiency is 99% as measured by Monte Carlo in the central regions of the detector [84].

After the reconstruction of the electrons is finished, the electron candidate must then pass a suite of identification criteria [84]. These are a set of cut-based selections with three working points: *loose*, *medium*, and *tight*. As the name implies, the *tight* working point applies the most stringent selection (and yields the largest background rejection power), while *loose* applies the loosest (and provides the worst background rejection). The working points are not orthogonal as electron candidates that pass any of the tighter identification requirements also pass all of the looser requirements. As the selection gets more stringent, cuts that were applied at the looser operating points are tightened, and new cuts are introduced. The working points used in this analysis correspond to *medium* and *tight* selection

criteria. The identification requirements use a wide range of discriminating variables including track quality, object isolation, track matching, vertexing, hadronic leakage and longitudinal and transverse shower shapes.

Further selection on the electrons are required before they are accepted as electrons in the analysis. This is discussed in more detail in Section 5.5.1. More specific details on the reconstruction and identification of electrons in ATLAS during Run I are found in Refs. [84, 85].

Scale factors are applied to the Monte Carlo in order to correctly model the electron reconstruction and identification efficiency in data (in bins of E_{T} and η).

4.5 Muon Reconstruction

Like the electron reconstruction, the muon reconstruction and identification must be quite efficient. Over most of the p_{T} range $5 \leq p_{\mathrm{T}} \leq 100$ GeV, the efficiency is greater than 99% [86] over the full MS acceptance ($|\eta| \leq 2.7$).

The reconstruction of the muons uses information from both the inner detector and the dedicated muon spectrometer[1], with supporting information from the calorimeters. So-called *combined muons* have the highest purity, and are the ones most often used in ATLAS physics analyses, including this one.

To construct combined muons, tracks are reconstructed in the MS by first constructing track segments in the individual muon chambers. These tracks are extended and matched with track segments in the other muon chambers to form a complete muon track. These muon tracks are then combined with matching tracks in the ID, which gives an independent measurement of the muon trajectory, to form a complete combined muon track. The matching of the ID muon tracks and the MS muon tracks is done on the basis of a χ^2 test. The final reconstruction of the muon trajectory is done in this analysis via the *Staco* algorithm which performs a statistical combination of the ID track and MS track in order to give the final muon trajectory [87].

Like the case of the trigger scale factors and electron reconstruction scale factors, scale factors are applied to MC events with reconstructed muons to ensure that the MC properly models the data. The scale factors are calculated by measuring the efficiencies using the tag-and-probe method in $Z \to \mu\mu$ events, and are applied in bins of muon ϕ and η.

4.6 Jet Algorithms and Reconstruction

Unlike the reconstruction of electrons and muons, which attempt to reconstruct individual physical objects in the calorimeters/trackers, jets are not discrete objects. Instead, they are made up of a cascade of hadronic and EM particles (a *shower*) resulting from the hadronization of the coloured object(s) created at the interaction point. The result is a more diffuse deposit of energy in the calorimeters when compared to the energy deposits of electrons and photons. In the case of this analysis, we are primarily interested in jets coming from the hadronization of quarks resulting from the decay of a W-boson.

[1] Muons can be reconstructed using information from the MS alone; these are referred to as *standalone muons*, and are not used in this analysis.

4.6.1 Jet Algorithms

As energy deposits from jets are more diffuse than those from EM objects, and have a more variable shower shape in the calorimeters, the algorithms used to reconstruct them must be chosen carefully. Of particular note for jets in ATLAS, is the large amount of unrelated energy that is deposited in the calorimeters as a result of pile-up. Therefore, having an algorithm that is not overly sensitive to soft contributions is crucial.

The most commonly used jet algorithms are k_t [88, 89], Cambridge/Aachen (C/A) [90], and anti-k_t [91]. All three combine the four-momenta from close-by final state jet constituents to form the final jet. In ATLAS analyses, the jet constituents are most commonly *Had 420* topological clusters (although tracks from the ID may also be used), while for MC simulations, the final state particles are used to create *truth* jets. For the remainder of this section, clusters will be used as the jet constituents.

All jet algorithms mentioned begin by first calculating the distance parameters for all clusters/pairs of clusters, denoted by indices i and j;

$$d_{ij} = \min\left(k_{ti}^{2p}, k_{tj}^{2p}\right)\frac{\Delta_{ij}^2}{R^2}, \tag{4.1}$$

$$d_{iB} = k_{ti}^{2p}. \tag{4.2}$$

Here, k_t is a cluster's transverse momentum, $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, and R is the radius parameter of the jet. The radius parameter determines the physical size of the jet, and ensures that the center of each jet is separated from any other jet's center by at least R. When reconstructing jets in ATLAS, the rapidity, y, is replaced by the pseudorapidity, η, in the definition of Δ_{ij}^2 (which thereby becomes the ΔR defined in Equation 3.4). The value of p in Equation 4.1 determines the relative power of the transverse momentum term versus the topological term (Δ_{ij}^2/R^2), as discussed further below.

For all clusters, the quantities d_{ij} and d_{iB} are computed, and the smallest entry is identified. If the smallest entry is the distance parameter between two clusters, d_{ij}, the four-momenta of the two clusters are summed to form a *proto-jet*, and the two constituents that make up the proto-jet are removed from further consideration. The quantities d_{ij} and d_{iB} for all pairs of clusters/proto-jets is recalculated, and the process is repeated. Alternatively, if the smallest quantity is d_{iB}, then the proto-jet, i, is considered a jet, and removed from the jet finding algorithm, and the next smallest distance parameter is identified. This process continues until all clusters/proto-jets have been accepted as jets.

The three jet algorithms mentioned above differ only in their value of p, with $p = 1$ for k_t, -1 for anti-k_t, and 0 for the C/A algorithm, resulting in different clustering behaviour. The anti-k_t algorithm produces conical jets, whose shape is not altered greatly by the presence of small energy deposits, unlike the k_t algorithm, which produces rather irregularly shaped jets, seeded by softer energy deposits. This makes anti-k_t jets ideal for high pile-up environments such as ATLAS. The C/A clustering on the other hand is based purely on the topological structure of the clusters, with no weight given to the transverse momentum of the jet constituents. This produces irregularly shaped jets, much like the k_t algorithm jets.

A variety of R parameters can be used in each jet algorithm, however, the most common in ATLAS are $R = 0.4$ jets constructed with the anti-k_t algorithm. Also used in this analysis is a large-R C/A jet[1] with $R = 1.2$. The reasons behind this are discussed in Section 5.5.1.

[1] By convention, any jet with $R \geq 1$ is referred to as a large-R jet (and sometimes, a fat jet).

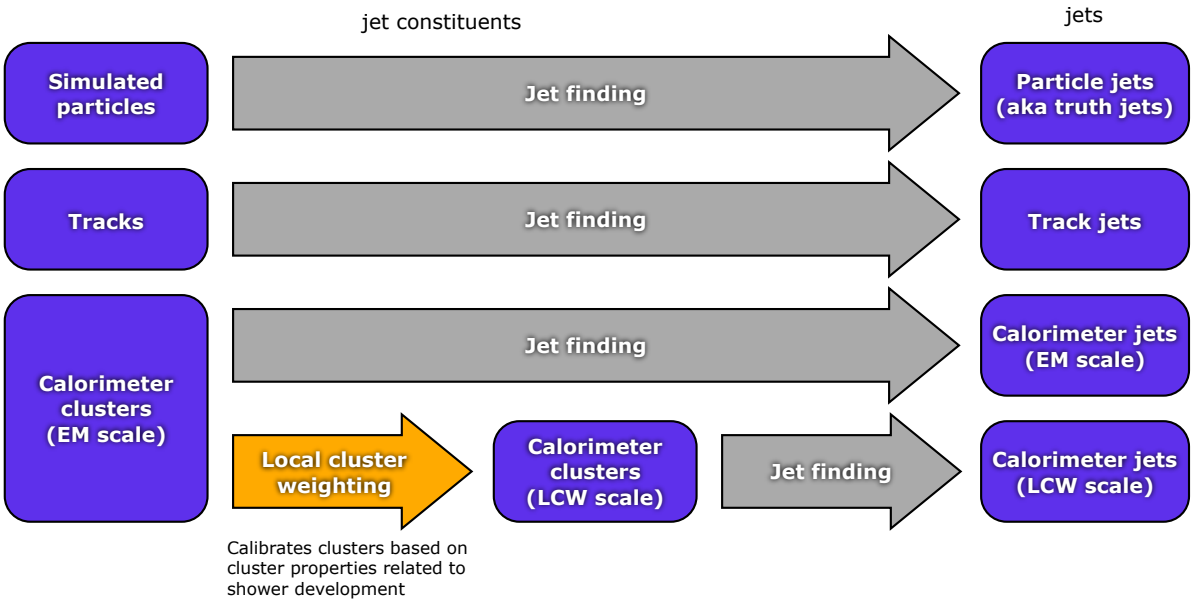


Figure 4.3: ATLAS jet reconstruction [92].

4.6.2 Reconstructing Jets in ATLAS

ATLAS jets start with *Had 420* topological clusters (Section 4.3.2), which are constructed from cells with energies determined at the EM scale. Before jets are constructed, the clusters are calibrated with the Local Cell Weighting calibration (LCW), which first classifies the cluster as either "EM-like" or "hadron-like", and then weights the cluster energy to account for the lower calorimeter response to hadronic showers [92, 93, 94]. The LCW calibration is derived from studies of single charged and neutral pions and their associated clusters in a GEANT4 [95, 96] Monte Carlo simulation of the ATLAS detector[1].

Although not used in this analysis, ATLAS is also able to reconstruct jets at the EM scale (where no LCW is done), and track jets. In the case of truth jets, the relevant jet algorithm is applied to the final state particles after showering in a Monte Carlo program. The steps to reconstruct jets for these different jets types are found in Figure 4.3.

Jet Calibration

After a jet is constructed from LCW topological clusters, the ATLAS jet calibration is performed. This is a sequential process which accounts for the effects of pile-up and the hadronic response of the calorimeters (which is not completely corrected for in the LCW procedure), and is outlined below.

- **Jet Origin Correction:** The jet's origin is corrected to be located at the primary vertex of the interaction (instead of the center of ATLAS, which is assumed in the jet making step). The jet's four-vector is adjusted so that the jet points back to the primary vertex, with its energy unchanged.
- **Pile-up Correction:** Energy deposits inside the jet area that are a result of pile-up, if unaccounted for, cause the jet's reconstructed energy to be dependant on the pile-up (both in-time and out-of-time pile-up). To reduce this effect, the median p_{T} density from pile-up is evaluated for

[1] More information about the modelling of ATLAS in MC events can be found in Section 5.3.1.

each event. For each jet individually, the pile-up p_{T} density is then scaled by the jet's area, and subtracted from the jet energy [97, 98].

The median p_{T} density, ρ, is calculated as the median p_{T} density for all jets i;

$$\rho = \mathrm{Median}\left[\frac{p_{\mathrm{T},i}^{jet}}{A_i^{jet}}\right], \tag{4.3}$$

where A_i^{jet} is the jet's area. The jets used in this definition of ρ are k_t jets with radius parameter $R = 0.4$, reconstructed with topological clusters in the central region of the detector ($|\eta| \leq 2.0$). k_t jets are used instead of the standard anti-k_t jets due to their sensitivity to the soft energy deposits characteristic of pile-up

- **Residual pile-up correction:** After the jet area pile-up correction is applied, there remains a residual dependence on the number of reconstructed primary vertices (N_{PV}) and the average number of collisions per bunch crossing ($\langle\mu\rangle$) seen in the reconstructed jet energy. A correction is applied, which parameterizes the residual energy dependence on the variables N_{PV} and $\langle\mu\rangle$ independently.

  The final corrected jet energy, which combines all corrections done to account for pile-up, takes the form:

$$p_{\mathrm{T}}{}^{jet,corr} = p_{\mathrm{T}}{}^{jet} - \left(\rho \times A^{jet}\right) - \left(\alpha \times (N_{PV} - 1)\right) - \left(\beta \times \langle\mu\rangle\right), \tag{4.4}$$

  with α and β being the residual pile-up correction coefficients (which are dependant on the jet algorithm and area).

- **Absolute Jet Energy Scale:** The jet energy is calibrated to the jet energy scale with constants extracted from Monte Carlo studies comparing the energy response of reconstructed jets to truth jets. Both the energy and direction of the jet are adjusted. The calibration coefficients are parameterized as a function of the jet p_{T} and η.

- **Residual In-situ Calibration:** This is final step of the calibration which is *only applied to data*. The residual in-situ calibration compares data and MC in-situ calibrations, by looking at the p_{T} balance of jets from reference objects (for example, dijets from Z-boson decays). The constants derived are of the form,

$$\frac{\left(p_T^{probe}/p_T^{ref}\right)_{Data}}{\left(p_T^{probe}/p_T^{ref}\right)_{MC}},$$

  and vary as a function of jet p_{T} and η.

Flavour Tagging

The ability to flavour tag reconstructed jets is useful to both highlight a particular signal, or reject unwanted backgrounds. In particular, jets originating from the hadronization of b quarks can be tagged using a process called b-tagging.

In ATLAS, b-tagging is done with a multivariate neural net algorithm, MV1 [99, 100, 101]. The MV1 method uses a combination of likelihood based b-tag algorithms as its input. These other algorithms

search for topological differences between b and light jets, such as displaced vertices and the weak decay topologies of b quark decays.

The MV1 algorithm can be used at a number of different operating points, each of which select b-jets with a particular efficiency and associated purity. The operating point used in this analysis corresponds to a b-tagging efficiency of 85%, and a purity of ∼77%, with the contamination being dominated by mistagged charm quarks. This provides an excellent mechanism to reject top quark backgrounds, which are one of the dominant backgrounds in this analysis (see Section 5.3.4). The b-tagging algorithms rely on data from the ID, and therefore flavour tagging of jets is only possible for those within the acceptance of the trackers.

Flavour tagging is used to reject the high amount of top quark background in this analysis. The details of the flavour tagging requirements used will be discussed in Section 5.5.1.

As with the trigger scale factors, and lepton efficiency scale factors, there are associated data-driven b-tag scale factors which are applied to the MC in order to provide good agreement with data.

4.7 Missing Transverse Momentum

Missing transverse momentum ($E_{\mathrm{T}}^{\mathrm{miss}}$, and sometimes MET) is defined as the transverse momentum imbalance as measured by energy deposited in the detector, and is calculated from the energy deposited in the calorimeters, and momentum measurements from the muon spectrometers. The missing transverse momentum is used in order to reconstruct the vector sum of the transverse momentum/energy of particles that do not interact at all with the trackers or calorimeters, such as neutrinos.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated in separate x and y components [102, 103];

$$E_x^{miss} = E_x^{miss,e} + E_x^{miss,\gamma} + E_x^{miss,\tau} + E_x^{miss,jets} + E_x^{miss,soft\,terms} + E_x^{miss,\mu}, \tag{4.5}$$

$$E_y^{miss} = E_y^{miss,e} + E_y^{miss,\gamma} + E_y^{miss,\tau} + E_y^{miss,jets} + E_y^{miss,soft\,terms} + E_y^{miss,\mu}. \tag{4.6}$$

Each of these terms is calculated as the negative sum of the energy of all identified and *calibrated* objects, projected onto the relevant axis, as in Equations 4.7 and 4.8. The exceptions are the $E_{x(y)}^{miss,soft\,terms}$ terms, which are not derived from identified objects, but from LCW topological clusters and tracks not associated with any reconstructed object.

$$E_x^{miss,\,term} = -\sum_{i=0}^{N^{cell}} E_i \sin\theta_i \cos\phi_i, \tag{4.7}$$

$$E_y^{miss,\,term} = -\sum_{i=0}^{N^{cell}} E_i \sin\theta_i \sin\phi_i. \tag{4.8}$$

Once the two components of the missing transverse momentum are evaluated, the magnitude of the total missing transverse momentum is defined as $E_{\mathrm{T}}^{\mathrm{miss}} = \sqrt{(E_x^{miss})^2 + (E_y^{miss})^2}$, with azimuthal angle $\phi^{\mathrm{miss}} = \arctan\left(E_y^{miss}/E_x^{miss}\right)$.

A requirement on the missing transverse momentum is used in this analysis to reject backgrounds with low (or fake) $E_{\mathrm{T}}^{\mathrm{miss}}$, and to reconstruct the transverse momentum of the neutrino from the leptonically decaying W. The requirement on an event's $E_{\mathrm{T}}^{\mathrm{miss}}$ will be discussed in Section 5.5.3.

Chapter 5

Searching for a Heavy Higgs-like Boson

> "We are both busy people, so let's cut the small talk."
>
> David Mitchell, *number9dream*

This chapter provides a complete description of the search for a heavy Higgs boson in the $H \to WW \to \ell\nu jj$ channel. The discussion begins with an overview of the search final state, and modelling of the expected signal and background processes. As two of the signal models considered (EWS and SM-like) predict a Higgs with a non-negligible width, the interference between the search process, $H \to WW \to \ell\nu jj$, and the WW continuum background must also be treated, and will be discussed in Section 5.4 (motivation for the interference reweighting can be found in Section 2.6). Starting in Section 5.5 the final object selection requirements are detailed, along with the first event selection criteria. To ensure high sensitivity for the full range of Higgs masses considered, these event *cuts* are then tightened as a function of the Higgs mass hypothesis, as will be discussed in Section 5.9. The requirements outlined in the preselection define the *signal regions* in the analysis; the selection that defines the signal regions is slightly modified in order to provide *control regions* which are used to check the modelling of the dominant backgrounds of the analysis (see Section 5.6), and are also used in order to estimate the data driven multijet background, which is discussed in Section 5.8. Before results can be shown, the modelling of the MC is checked in the control regions, and in the signal region (which is defined by preselection cuts), and is discussed in Sections 5.10 and 5.11, along with expected signal and background event yields. Finally, the systematic uncertainties of the search are presented in Section 5.12, and the limit setting infrastructure and final results are presented in Sections 5.13 and 5.14.

5.1 Search Final State

The analysis discussed in this thesis searches for a heavy Higgs boson decaying into two W-bosons in the semi-leptonic final state[1]. The analysis searches for events with a single isolated lepton (vetoing events

[1] One W decays into a lepton – lepton neutrino pair, and the other hadronically, into quarks.

with more than one lepton), missing transverse momentum (representing the transverse momentum of the neutrino, which cannot be directly reconstructed by ATLAS), and jets from the decay of the hadronically decaying W. At low masses, the two quarks from the hadronic W decay are well enough separated to be treated as distinct reconstructed jets using a narrow cone anti-k_t jet algorithm. However, for high m_H ($m_H \gtrsim 700$ GeV), both W bosons will be boosted, resulting in energy deposits from the decay products being too close together to be separately reconstructed using narrow cone jets. Instead, a wide cone C/A algorithm (with a filtering algorithm applied), is used to capture the hadronization of the two quarks into a single reconstructed object. The basic reconstruction of all of these final state objects has already been discussed in Chapter 4. However, further selections are applied to the reconstructed objects, as will be outlined in Section 5.5.1.

5.2 Analysis Strategy

The experimental landscape has changed since the last Higgs search in the $H \to WW \to \ell\nu jj$ channel (discussed in Chapter 1), as a result of the discovery of the Higgs at 125 GeV. The same strategy of applying cross-section limits to a Higgs with Standard Model properties is no longer relevant, and therefore Higgs models beyond the Standard Model must be considered. While there are some concrete predictions for extended Higgs sectors outside the Standard Model (in theories such as supersymmetry), the focus here is on three pseudo-model-independent interpretations; a Higgs produced with a large width, as modelled by an SM-like lineshape (CPS lineshape at high mass) [26], a Higgs produced with a very narrow width (NWA) and an EWS Higgs with an unknown width (a direct result of an unknown κ') [14].

Therefore, there are three separate signal hypothesis being tested (not counting the range of widths present in the case of the EWS search). First is the SM-like Higgs signal, modelled with the CPS and BW lineshapes, probed in the region $300 \leq m_H \leq 1000$ GeV. The EWS model is probed in the same mass range, for widths ranging from $0.2{-}1.0{\times}\Gamma_{SM}$ in increments of $0.2{\times}\Gamma_{SM}$, with a Breit-Wigner lineshape[1]. This corresponds to a range of $0.2 \leq \kappa'^2 \leq 1.0$ for $\mathrm{BR}_{\mathrm{new}} = 0$. Finally, the model-independent case of a Higgs produced with a NWA lineshape is examined in the range $300 \leq m_H \leq 2000$ GeV. For each of these interpretations no deviation from the SM background hypotheses were found, and therefore upper limits on $\sigma \times \mathrm{BR}$ at the 95% confidence level are given. In the case of the EWS limits, $\mathrm{BR}_{\mathrm{new}}$ is assumed to be 0.

For the SM-like and EWS Higgs, which have non-negligible widths, the interference with the WW continuum background must be taken into account, as described in Section 2.6. The experimental treatment of this is discussed in Section 5.4.

The SM-like Higgs search, and the EWS search terminate at 1 TeV. For the SM-like case, this is done due to the fact that the CPS lineshape was calculated with SM constraints, and thus is only valid up to ∼1 TeV [26]. The EWS has the same search range due to the fact that the Monte Carlo samples are derived from the SM-like samples, through a reweighting method described in Section 5.3.3. In the case of the NWA Higgs search, which has dedicated MC samples, and is not confined by any SM interpretation, the upper range of the search can be set arbitrarily high. For the NWA search, samples

[1] The BW lineshape was chosen as a benchmark for the EWS after discussion with theorists in the LHC Cross-section Working group, as the simplest candidate for theoretical interpretation. Note that the EWS model makes no concrete prediction on the lineshape of the extra Higgs, and so the *simplest* lineshape was chosen.

were generated up to $m_H = 2000$ GeV, however limits are only set up to $m_H = 1200$ GeV due to a loss of sensitivity at higher mass.

5.3 Signal and Background Modelling

A series of MC generated samples are created in order to model the expected signal and background yields and topologies. All backgrounds in the analysis, except for the QCD multijet background, are modelled using MC simulations. A series of control regions (CR) which contain an enriched sample of a particular background are defined; these CRs allow the simulation predictions to be compared to the data, without biasing the analysis. The CRs are also used in order to form a data driven estimate of the cross-section for the dominant Top and W+jets backgrounds. More discussion on the various control regions and their definitions is provided in Section 5.6.

5.3.1 Modelling the ATLAS Detector

The ATLAS simulation infrastructure makes use of the Geant4 simulation framework to completely model the ATLAS detector, and predict its response from Monte Carlo simulation [95, 96, 104]. This modelling is done in three separate steps. In the first, the parton-parton hard-scatter process is simulated, unstable particles created in the hard-scatter are decayed (such as the Higgs, or massive gauge bosons), and any coloured final state particles undergo parton showering and hadronization. At this stage of the simulation, all the event information (including all particles and four-momenta), as calculated by the MC generator, are recorded as *truth* information, which is present in the final output. Next, the interactions between the final state particles with the detector itself are modelled. Finally, the interactions that occur between the sensitive portions of the ATLAS detector and the final state particles are digitized and converted into currents and voltages for simulated readout.

The output of the simulation chain is stored in the same format as ATLAS data originating from collisions in the LHC. As a result, output derived from either MC or collision data can be processed by the same reconstruction and analysis tools.

The computing resources required to simulate events with the full ATLAS Geant4 simulation are incredibly large. In some cases, generating MC samples with sufficiently high statistics would be impossible if run through the full simulation (fullsim) in the conventional way (keeping in mind the large number of different MC datasets needed for the wealth of physics processes, and associated backgrounds, interesting to all ATLAS analyzers). In order to be able to speed up the simulation process, which allows more events to be simulated with the same amount of CPU time, a faster simulation, ATLFAST-II (abbreviated as AFII, or AF2), has been developed [104]. Instead of simulating the full detector response to the final state particles, the AFII simulation uses the full Geant4 simulation of the inner detector and muon spectrometer, while making use of a simplified model of the calorimeters. As opposed to fullsim, which directly models the interactions of final state particles with the calorimeter, the AFII simulation parameterizes the longitudinal and transverse energy profiles of showers in the calorimeter based on extensive fullsim MC samples of photons and charged pions. These parameterizations are used to model how single particles deposit energy in the calorimeters, with the fullsim photon showers modelling shower development for photons and electrons, and the charged pions modelling hadronic showers. This provides a reduction in the CPU time required to process an event by an order of magnitude.

The broad kinematic range considered in this analysis means that high statistics samples are needed to sufficiently cover the full range of reconstructed $m_{\ell\nu jj}$. This can be quite difficult at high mass due to the falling exponential shape of the background (and is even more difficult in the VBF category, due to the stricter selection cuts, outlined in Section 5.5.3). Therefore, the availability of AFII samples is of great importance. When possible, AFII samples are used in this analysis in order to take advantage of the higher statistics they usually provide[1]. Whenever possible, these have been validated against their fullsim equivalents to ensure the modelling is consistent.

5.3.2 Monte Carlo Generators

There are multiple MC generators used to model the signal and background processes for this search. These can be split into two main types: general purpose generators, and matrix element generators [18]. General purpose generators are able to fully model high-energy collisions, both at small distance scales (hard scattering processes, and the decays of intermediate states), and larger scales to model the parton showering of the final state particles, emission of initial and final state radiation (ISR/FSR), and the remnants from multiple soft pp interactions (the underlying event). Examples of general purpose generators used are Herwig [105] (with the Jimmy [106] libraries used for the hadronization and parton showering), Sherpa [107, 108, 109, 110], and Pythia [111, 112].

The matrix element (ME) generators simulate only the hard scatter process, and therefore must rely on an interface with another Monte Carlo generator to simulate the parton showering process and emission of ISR/FSR. For samples used in this analysis, the MC generator used to shower ME generated samples is one of the general-purpose generators. The matrix element generators are more accurate for hard, large-angle emissions, when compared to the general purpose generators. Examples of the ME based MC programs used in the analysis are Alpgen [113] and AcerMC [114]. The final class of Monte Carlo generators are the NLO generators, which take into account NLO QCD processes. Again, the events generated through this process must be interfaced with a multi-purpose generator to model the parton showers (although some NLO generators have a parton shower model built in). An example of one of these such generators[2] is Powheg Box [115, 116, 117] (often just Powheg), which is used to model the signal in this analysis.

Regardless of the type of MC generator used, all must take into account the structure of the proton in the collisions, which impacts the cross-section of the underlying hard-scatter process, and the emission of ISR before the hard-scatter. Independent groups produce the parton distribution functions (PDFs) which describe the parton constituents of the protons. The parton distribution functions with the form, $f\left(x, Q^2\right)$, give the probability to find a particular parton carrying momentum x (expressed as a fraction of the proton's total momentum), when probed with a momentum transfer of Q^2. There are two sets of PDFs used in this analysis, CTEQ6L1 [118], a leading order PDF set used by the leading order generators, and CT10 [119], which is a NLO PDF set used by the NLO generators.

Monte Carlo generators often must be provided with the renormalization and factorization scales (μ_F and μ_R, respectively) to be used when simulating a process. The renormalization scale is the energy scale at which the generator calculates the strong coupling constant α_s, while the factorization scale determines the scale at which the MC generator queries the PDF set.

[1] Some AFII samples were not generated in order to provide a higher statistics, but rather to generate the samples required on a short timescale.

[2] Sherpa, a general purpose generator, also includes a NLO ME generator for simulating the hard-scatter.

Monte Carlo Tunes

The modelling of the parton showers and underlying event is quite sensitive to small changes to input parameters given to the MC generators (for example, the minimum p_{T} of scattered partons). These parameters can be *tuned* in order to match the observed data. Many of the MC generators responsible used in the analysis are custom tuned by the collaboration to match ATLAS observations (the ATLAS underlying event tunes, AUET [104]). Pythia has a set of Perugia tunes, which are derived by the Pythia authors, meant to model LHC data. The P2011C Perugia tunes [120] are used for samples showered with Pythia.

5.3.3 Signal Modelling

The two signal hypotheses that are directly modelled in Monte Carlo are the SM-like and NWA Higgs; the EWS signal samples are constructed through a reweighting of the SM-like samples as described in the next section.

The modelling of the SM-like signal is done with two different lineshapes. Below $m_H = 400$ GeV, samples are generated with a BW lineshape. Above $m_H = 400$ GeV, as the width becomes greater than 10% of the mass, and therefore the CPS lineshape must be used instead (see Section 2.3 for this discussion). All samples are generated with Powheg Box in 20 GeV m_H increments for $300 \leq m_H \leq 600$ GeV, and 50 GeV increments otherwise ($600 < m_H \leq 1000$ GeV), and subsequently showered with Pythia8 [112]. Samples are generated only including the signal contribution; the effects of interference are evaluated separately, as described in Section 5.4. A summary of the SM-like samples generated is displayed in Table 5.1.

The NWA samples are also generated with Powheg Box + Pythia8, with a BW lineshape and Higgs width of 4.07 MeV in m_H increments of 50 GeV for $m_H \leq 1500$ GeV. Above $m_H = 1500$ GeV the increments are 100 GeV. The NWA samples used are summarized in Table 5.2.

EWS Signal Samples

In order to model the variable width BW lineshape used to probe the EWS parameter space, the SM-like samples are reweighted to have a BW lineshape and width in the range $0.2 - 1.0 \times \Gamma_{SM}$. For $m_H < 400$ GeV samples, which are already generated with the correct lineshape, the reweighting only adjusts the lineshape to have a different width (in the case that $\Gamma' = \Gamma_{\mathrm{SM}}$, no reweighting is applied). For the high mass samples where the CPS lineshape is used, the reweighting changes both the width and the WW invariant mass lineshape.

The expected lineshape for the EWS samples is computed from high-statistics Powheg Box samples generated for each mass point and width used in the analysis. The m_{WW} spectrum of the SM-like samples is then reweighted to match the BW lineshape with a map that relates the true invariant WW mass to the lineshape weight needed.

5.3.4 Background Samples

The main backgrounds to the $H \to WW \to \ell\nu jj$ signal process are from W+jets and Top events, with smaller components coming from diboson, Drell-Yan, Z+jets and QCD multijet processes. Except in the case of the multijet background, all background kinematics are determined directly from MC simulation.

Table 5.1: SM-like signal samples used in the analysis. All mass points in the SM-like search are listed, with information on the lineshape, MC generators (and tunes), PDF sets, and detector simulation used. All signal samples are generated without the WW interference component.

m_H [GeV]	Generator (Tune)	PDF Set	AFII/Fullsim
	ggF		
300 (Breit-Wigner)	Powheg + Pythia8 (AU2)	CT10	Fullsim
320 (Breit-Wigner)	Powheg + Pythia8 (AU2)	CT10	Fullsim
340 (Breit-Wigner)	Powheg + Pythia8 (AU2)	CT10	Fullsim
360 (Breit-Wigner)	Powheg + Pythia8 (AU2)	CT10	Fullsim
380 (Breit-Wigner)	Powheg + Pythia8 (AU2)	CT10	Fullsim
400 (CPS)	Powheg + Pythia8 (AU2)	CT10	Fullsim
420 (CPS)	Powheg + Pythia8 (AU2)	CT10	Fullsim
440 (CPS)	Powheg + Pythia8 (AU2)	CT10	Fullsim
460 (CPS)	Powheg + Pythia8 (AU2)	CT10	Fullsim
480 (CPS)	Powheg + Pythia8 (AU2)	CT10	Fullsim
500 (CPS)	Powheg + Pythia8 (AU2)	CT10	Fullsim
520 (CPS)	Powheg + Pythia8 (AU2)	CT10	Fullsim
540 (CPS)	Powheg + Pythia8 (AU2)	CT10	Fullsim
560 (CPS)	Powheg + Pythia8 (AU2)	CT10	Fullsim
580 (CPS)	Powheg + Pythia8 (AU2)	CT10	Fullsim
600 (CPS)	Powheg + Pythia8 (AU2)	CT10	Fullsim
650 (CPS)	Powheg + Pythia8 (AU2)	CT10	Fullsim
700 (CPS)	Powheg + Pythia8 (AU2)	CT10	Fullsim
750 (CPS)	Powheg + Pythia8 (AU2)	CT10	Fullsim
800 (CPS)	Powheg + Pythia8 (AU2)	CT10	Fullsim
850 (CPS)	Powheg + Pythia8 (AU2)	CT10	Fullsim
900 (CPS)	Powheg + Pythia8 (AU2)	CT10	Fullsim
950 (CPS)	Powheg + Pythia8 (AU2)	CT10	Fullsim
1000 (CPS)	Powheg + Pythia8 (AU2)	CT10	Fullsim
	VBF		
300 (Breit-Wigner)	Powheg + Pythia8 (AU2)	CT10	Fullsim
320 (Breit-Wigner)	Powheg + Pythia8 (AU2)	CT10	Fullsim
340 (Breit-Wigner)	Powheg + Pythia8 (AU2)	CT10	Fullsim
360 (Breit-Wigner)	Powheg + Pythia8 (AU2)	CT10	Fullsim
380 (Breit-Wigner)	Powheg + Pythia8 (AU2)	CT10	Fullsim
400 (CPS)	Powheg + Pythia8 (AU2)	CT10	AFII
420 (CPS)	Powheg + Pythia8 (AU2)	CT10	AFII
440 (CPS)	Powheg + Pythia8 (AU2)	CT10	AFII
460 (CPS)	Powheg + Pythia8 (AU2)	CT10	AFII
480 (CPS)	Powheg + Pythia8 (AU2)	CT10	AFII
500 (CPS)	Powheg + Pythia8 (AU2)	CT10	AFII
520 (CPS)	Powheg + Pythia8 (AU2)	CT10	AFII
540 (CPS)	Powheg + Pythia8 (AU2)	CT10	AFII
560 (CPS)	Powheg + Pythia8 (AU2)	CT10	AFII
580 (CPS)	Powheg + Pythia8 (AU2)	CT10	AFII
600 (CPS)	Powheg + Pythia8 (AU2)	CT10	AFII
650 (CPS)	Powheg + Pythia8 (AU2)	CT10	AFII
700 (CPS)	Powheg + Pythia8 (AU2)	CT10	AFII
750 (CPS)	Powheg + Pythia8 (AU2)	CT10	AFII
800 (CPS)	Powheg + Pythia8 (AU2)	CT10	AFII
850 (CPS)	Powheg + Pythia8 (AU2)	CT10	AFII
900 (CPS)	Powheg + Pythia8 (AU2)	CT10	AFII
950 (CPS)	Powheg + Pythia8 (AU2)	CT10	AFII
1000 (CPS)	Powheg + Pythia8 (AU2)	CT10	AFII

Table 5.2: NWA signal samples used in the analysis. All mass points in the SM-like search are listed, with information on the lineshape, MC generators (and tunes), PDF sets, and detector simulation used.

m_H [GeV]	Generator (Tune)	PDF Set	AFII/Fullsim
	ggF		
300	Powheg + Pythia8 (AU2)	CT10	Fullsim
350	Powheg + Pythia8 (AU2)	CT10	AFII
400	Powheg + Pythia8 (AU2)	CT10	Fullsim
450	Powheg + Pythia8 (AU2)	CT10	AFII
500	Powheg + Pythia8 (AU2)	CT10	Fullsim
550	Powheg + Pythia8 (AU2)	CT10	AFII
600	Powheg + Pythia8 (AU2)	CT10	Fullsim
650	Powheg + Pythia8 (AU2)	CT10	AFII
700	Powheg + Pythia8 (AU2)	CT10	Fullsim
750	Powheg + Pythia8 (AU2)	CT10	AFII
800	Powheg + Pythia8 (AU2)	CT10	Fullsim
850	Powheg + Pythia8 (AU2)	CT10	AFII
900	Powheg + Pythia8 (AU2)	CT10	Fullsim
950	Powheg + Pythia8 (AU2)	CT10	AFII
1000	Powheg + Pythia8 (AU2)	CT10	Fullsim
1050	Powheg + Pythia8 (AU2)	CT10	AFII
1100	Powheg + Pythia8 (AU2)	CT10	AFII
1150	Powheg + Pythia8 (AU2)	CT10	AFII
1200	Powheg + Pythia8 (AU2)	CT10	AFII
1250	Powheg + Pythia8 (AU2)	CT10	AFII
1300	Powheg + Pythia8 (AU2)	CT10	AFII
1350	Powheg + Pythia8 (AU2)	CT10	AFII
1400	Powheg + Pythia8 (AU2)	CT10	AFII
1450	Powheg + Pythia8 (AU2)	CT10	AFII
1500	Powheg + Pythia8 (AU2)	CT10	AFII
1600	Powheg + Pythia8 (AU2)	CT10	AFII
1700	Powheg + Pythia8 (AU2)	CT10	AFII
1800	Powheg + Pythia8 (AU2)	CT10	AFII
1900	Powheg + Pythia8 (AU2)	CT10	AFII
2000	Powheg + Pythia8 (AU2)	CT10	AFII
	VBF		
300	Powheg + Pythia8 (AU2)	CT10	Fullsim
350	Powheg + Pythia8 (AU2)	CT10	AFII
400	Powheg + Pythia8 (AU2)	CT10	Fullsim
450	Powheg + Pythia8 (AU2)	CT10	AFII
500	Powheg + Pythia8 (AU2)	CT10	Fullsim
550	Powheg + Pythia8 (AU2)	CT10	AFII
600	Powheg + Pythia8 (AU2)	CT10	Fullsim
650	Powheg + Pythia8 (AU2)	CT10	AFII
700	Powheg + Pythia8 (AU2)	CT10	Fullsim
750	Powheg + Pythia8 (AU2)	CT10	AFII
800	Powheg + Pythia8 (AU2)	CT10	Fullsim
850	Powheg + Pythia8 (AU2)	CT10	AFII
900	Powheg + Pythia8 (AU2)	CT10	Fullsim
950	Powheg + Pythia8 (AU2)	CT10	AFII
1000	Powheg + Pythia8 (AU2)	CT10	Fullsim
1050	Powheg + Pythia8 (AU2)	CT10	AFII
1100	Powheg + Pythia8 (AU2)	CT10	AFII
1150	Powheg + Pythia8 (AU2)	CT10	AFII
1200	Powheg + Pythia8 (AU2)	CT10	AFII
1250	Powheg + Pythia8 (AU2)	CT10	AFII
1300	Powheg + Pythia8 (AU2)	CT10	AFII
1350	Powheg + Pythia8 (AU2)	CT10	AFII
1400	Powheg + Pythia8 (AU2)	CT10	AFII
1450	Powheg + Pythia8 (AU2)	CT10	AFII
1500	Powheg + Pythia8 (AU2)	CT10	AFII
1600	Powheg + Pythia8 (AU2)	CT10	AFII
1700	Powheg + Pythia8 (AU2)	CT10	AFII
1800	Powheg + Pythia8 (AU2)	CT10	AFII
1900	Powheg + Pythia8 (AU2)	CT10	AFII
2000	Powheg + Pythia8 (AU2)	CT10	AFII

The background processes considered in the analysis are described below, with detail regarding the MC samples used.

- **W+jets:** Production of leptonically decaying W, with associated jets. This background dominates over the full kinematic region studied. The modelling of the W+jets background is checked with data, using a dedicated control region (Section 5.6); while kinematics of the background are modelled in MC, the overall normalization of the W+jets background is determined by a comparison to data yields in the signal and control regions in the final fit. As a result of the high mass region being considered in this analysis, gaining sufficient statistics in the high mass tails for this background is quite difficult, since the majority of W+jets events have low p_{T}^{W}. In order to generate sufficient statistics in the tails, without the need to generate a large number of events that would preferentially populate the low mass region, Sherpa samples are generated in five bins of p_{T}^{W} ($40-70$, $70-140$, $140-280$, $280-500$, $500+$), and an inclusive sample for $p_{\mathrm{T}}^{W} < 40$ GeV. When available, AFII samples are used due to the larger statistics available.

- **Top:** Single top and $t\bar{t}$ pairs. Along with W+jets, top backgrounds are dominant processes in the $H \to WW \to \ell\nu jj$ analysis. Like the W+jets background, the Top processes are modelled by MC, with the normalization determined from the final fit. A control region with an enriched sample of Top is defined in order to examine the MC Top modelling, and constrain the overall Top event yield. In order to have sufficient statistics in the high-mass region, the Top samples are all generated with the AFII simulation.

- **Z+jets:** Production of a leptonically decaying Z-boson, with associated jets. Both the kinematics and overall normalization are determined by MC. Like the case for the W+jets background, in order to have sufficient statistics in the high mass region Sherpa, p_{T} binned samples are used.

- **Diboson/Drell-Yan:** Background from production of $WW/WZ/ZZ$ and $Z/\gamma^{*} \to \ell\ell$ ($\ell = e, \mu$). Modelling and yields are from MC.

- **Multijet:** QCD multijet events where a jet fakes an electron or muon. While having a high cross-section at the LHC, these events are heavily suppressed at high mass, and play only a minor role. Due to the difficulty in modelling the QCD multijet processes in MC, a data-driven technique is used, as described in Section 5.8.

The simulation samples used for the Top, diboson, W+jets, Z+jets and Drell-Yan processes are summarized in Tables 5.3–5.5.

5.4 Interference Reweighting

For broad resonances above $m_H = 400$ GeV, the interference between the $H \to WW \to \ell\nu jj$ process and the $WW \to \ell\nu jj$ continuum, non-resonant background must be taken into account, as was discussed in Section 2.6. In this analysis, separate methods are used to account for the interference in ggF and VBF production modes.

The process of generating, showering and modelling the ATLAS response in MC is extremely time consuming, especially if a high number of samples, or statistics are needed (see Section 5.3.1)[1]. If the

[1] Also, the reweighting procedure was new at the time, and took some number of iterations, and collaboration with theorists, MC authors and CMS colleagues, in order to get reasonable and verified results.

Table 5.3: Top, Diboson and Drell-Yan Monte Carlo samples used in the analysis, with information on the process, MC generator (and tune), cross-section, and detector simulation type.

Process	Generator (Tune)	PDF Set	$\sigma\times$ BR [pb] (@ $\sqrt{s} = 8$ TeV)	AFII/Fullsim
Single top (t-channel)	AcerMC + Pythia6 (P2011C)	CTEQ6L1	28.44	AFII
Single top (s-channel)	Powheg + Pythia6 (P2011C)	CTEQ6L1	1.818	AFII
$t\bar{t}$	Powheg + Pythia6 (P2011C)	CTEQ6L1	252.89	AFII
Wt (inclusive)	Powheg + Pythia6 (P2011C)	CTEQ6L1	22.37	AFII
WW	Herwig (AUET2)	CTEQ6L1	32.501	Fullsim
WZ	Herwig (AUET2)	CTEQ6L1	12.009	Fullsim
ZZ	Herwig (AUET2)	CTEQ6L1	4.6914	Fullsim
$Z/\gamma^* \to ee$ ($p_{\mathrm{T}} > 10$ GeV)	Sherpa	CT10	32.26	Fullsim
$Z/\gamma* \to \mu\mu$ ($p_{\mathrm{T}} > 10$ GeV)	Sherpa	CT10	32.317	Fullsim

output of the reweighted Repolo + Powheg event files were then used as inputs in the full ATLAS simulation chain, there would be a significant burden placed on the simulation infrastructure. This is especially true for the case of the EWS search, where samples for each width being studied would need to be put through the entire simulation chain. Therefore, the interference is taken into account *after* the MC samples have been fully processes through the ATLAS simulation chain. For both the ggF and VBF interference weights, a map that relates the true invariant WW mass to the weight to be applied, $m_{WW,truth} \to w_{\mathcal{I}}$, is constructed. In the subsequent processing of the MC signal samples in the analysis chain, the MC truth information for each event is used as a key to find the appropriate weight.

The result of the interference is that the overall lineshape is broadened, with destructive interference occurring for $m_{WW} > m_H$ and constructive interference for $m_{WW} < m_H$. As the Higgs mass increases, this effect becomes more and more pronounced.

5.4.1 ggF Interference Reweighting

The ggF interference weights were calculated with MCFM v6.3 together with the cross-section reweighting scheme outlined in Section 2.6. MCFM was used to model the shape of the signal + interference curves. This new $\mathcal{S}+\mathcal{I}$ lineshape was then scaled to the appropriate cross-section using Equation 2.32. The final interference weights, which contain the modified shape and normalization information, are stored in a lookup table which maps the weights as a function of m_{WW}.

In order to properly account for the uncertainty on the scaling of the interference cross-section from LO $\to$ N(N)LO, the additive and multiplicative cross-sections are calculated and used to define the uncertainty on the method (see Equations 2.30 and 2.31).

5.4.2 VBF Interference Reweighting

The VBF interference reweighting is performed using Repolo. As input, Repolo reads the Les Houches event file (LHEF[1]) [121] from Powheg, before showering/hadronization, and assigns a weight to each event in order to transform the invariant mass spectrum from $(d\sigma/dm_{WW})_{\mathcal{S}} \to (d\sigma/dm_{WW})_{\mathcal{S}+\mathcal{I}}$. The weight is calculated by evaluating the matrix elements for the background ($\mathcal{M}_{\mathcal{B}}$), signal ($\mathcal{M}_{\mathcal{S}}$) and the combined spectrum ($\mathcal{M}_{\mathcal{SBI}}$, including interference). The interference weight is calculated by Repolo as,

[1]LHEF is pseudo-human readable text file which stores truth particle information for MC events. It is an open standard and allows different MC programs and algorithms to interface with each other easily.

Table 5.4: W+jets Monte Carlo samples used to model background used in the analysis, with information on the process, MC generator (and tune), cross-section, and detector simulation type. In addition to the W+jets samples being binned in p_{T}^{W}, the samples are also filtered by quark content, for use by the ATLAS Top working group.

Process	Generator (Tune)	PDF Set	$\sigma\times$ BR [pb] (@ $\sqrt{s}=8$ TeV)	AFII/Fullsim
W$\to e\nu$	Sherpa	CT10	11866.0	Fullsim
W$\to e\nu$, $40<p_{\mathrm{T}}<70$, b-jet filter	Sherpa	CT10	652.82	AFII
W$\to e\nu$, $40<p_{\mathrm{T}}<70$,c-jet filter, b-jet veto	Sherpa	CT10	652.83	AFII
W$\to e\nu$, $40<p_{\mathrm{T}}<70$, c-jet veto, b-jet veto	Sherpa	CT10	653.16	AFII
W$\to e\nu$, $70<p_{\mathrm{T}}<140$, b-jet filter	Sherpa	CT10	250.55	AFII
W$\to e\nu$, $70<p_{\mathrm{T}}<140$, c-jet filter, b-jet veto	Sherpa	CT10	250.71	AFII
W$\to e\nu$, $70<p_{\mathrm{T}}<140$, c-jet veto, b-jet veto	Sherpa	CT10	250.43	AFII
W$\to e\nu$, $140<p_{\mathrm{T}}<280$, b-jet filter	Sherpa	CT10	31.155	AFII
W$\to e\nu$, $140<p_{\mathrm{T}}<280$, c-jet filter, b-jet veto	Sherpa	CT10	31.189	AFII
W$\to e\nu$, $140<p_{\mathrm{T}}<280$, c-jet veto, b-jet veto	Sherpa	CT10	31.112	AFII
W$\to e\nu$, $280<p_{\mathrm{T}}<500$, b-jet filter	Sherpa	CT10	1.8413	Fullsim
W$\to e\nu$, $280<p_{\mathrm{T}}<500$, c-jet filter, b-jet veto	Sherpa	CT10	1.8370	Fullsim
W$\to e\nu$, $280<p_{\mathrm{T}}<500$, c-jet veto, b-jet veto	Sherpa	CT10	1.8426	Fullsim
W$\to e\nu$, $p_{\mathrm{T}}>500$, b-jet filter	Sherpa	CT10	0.10188	Fullsim
W$\to e\nu$, $p_{\mathrm{T}}>500$, c-jet filter, b-jet veto	Sherpa	CT10	0.10101	Fullsim
W$\to e\nu$, $p_{\mathrm{T}}>500$, c-jet veto, b-jet veto	Sherpa	CT10	0.10093	Fullsim
W$\to \mu\nu$	Sherpa	CT10	11867.0	Fullsim
W$\to \mu\nu$, $40<p_{\mathrm{T}}<70$, b-jet filter	Sherpa	CT10	652.73	AFII
W$\to \mu\nu$, $40<p_{\mathrm{T}}<70$, c-jet filter, b-jet veto	Sherpa	CT10	653.14	AFII
W$\to \mu\nu$, $40<p_{\mathrm{T}}<70$, c-jet veto, b-jet veto	Sherpa	CT10	653.06	AFII
W$\to \mu\nu$, $70<p_{\mathrm{T}}<140$, b-jet filter	Sherpa	CT10	250.55	AFII
W$\to \mu\nu$, $70<p_{\mathrm{T}}<140$, c-jet filter, b-jet veto	Sherpa	CT10	250.57	AFII
W$\to \mu\nu$, $70<p_{\mathrm{T}}<140$, c-jet veto, b-jet veto	Sherpa	CT10	250.77	AFII
W$\to \mu\nu$, $140<p_{\mathrm{T}}<280$, b-jet filter	Sherpa	CT10	31.164	AFII
W$\to \mu\nu$, $140<p_{\mathrm{T}}<280$, c-jet filter, b-jet veto	Sherpa	CT10	31.165	AFII
W$\to \mu\nu$, $140<p_{\mathrm{T}}<280$, c-jet veto, b-jet veto	Sherpa	CT10	31.173	AFII
W$\to \mu\nu$, $280<p_{\mathrm{T}}<500$, b-jet filter	Sherpa	CT10	1.8380	Fullsim
W$\to \mu\nu$, $280<p_{\mathrm{T}}<500$, c-jet filter, b-jet veto	Sherpa	CT10	1.8395	Fullsim
W$\to \mu\nu$, $280<p_{\mathrm{T}}<500$, c-jet veto, b-jet veto	Sherpa	CT10	1.8433	Fullsim
W$\to \mu\nu$, $p_{\mathrm{T}}>500$, b-jet filter	Sherpa	CT10	0.10163	Fullsim
W$\to \mu\nu$, $p_{\mathrm{T}}>500$, c-jet filter, b-jet veto	Sherpa	CT10	0.10210	Fullsim
W$\to \mu\nu$, $p_{\mathrm{T}}>500$, c-jet veto, b-jet veto	Sherpa	CT10	0.10186	Fullsim
W$\to \tau\nu$	Sherpa	CT10	11858.0	Fullsim
W$\to \tau\nu$, $40<p_{\mathrm{T}}<70$, b-jet filter	Sherpa	CT10	652.84	AFII
W$\to \tau\nu$, $40<p_{\mathrm{T}}<70$, c-jet filter, b-jet veto	Sherpa	CT10	652.58	AFII
W$\to \tau\nu$, $40<p_{\mathrm{T}}<70$, c-jet veto, b-jet veto	Sherpa	CT10	652.99	AFII
W$\to \tau\nu$, $70<p_{\mathrm{T}}<140$, b-jet filter	Sherpa	CT10	250.57	Fullsim
W$\to \tau\nu$, $70<p_{\mathrm{T}}<140$, c-jet filter, b-jet veto	Sherpa	CT10	250.61	Fullsim
W$\to \tau\nu$, $70<p_{\mathrm{T}}<140$, c-jet veto, b-jet veto	Sherpa	CT10	250.60	Fullsim
W$\to \tau\nu$, $140<p_{\mathrm{T}}<280$, b-jet filter	Sherpa	CT10	31.162	Fullsim
W$\to \tau\nu$, $140<p_{\mathrm{T}}<280$, c-jet filter, b-jet veto	Sherpa	CT10	31.151	Fullsim
W$\to \tau\nu$, $140<p_{\mathrm{T}}<280$, c-jet veto, b-jet veto	Sherpa	CT10	31.176	Fullsim
W$\to \tau\nu$, $280<p_{\mathrm{T}}<500$, b-jet filter	Sherpa	CT10	1.8362	Fullsim
W$\to \tau\nu$, $280<p_{\mathrm{T}}<500$, c-jet filter, b-jet veto	Sherpa	CT10	1.8395	Fullsim
W$\to \tau\nu$, $280<p_{\mathrm{T}}<500$, c-jet veto, b-jet veto	Sherpa	CT10	1.8368	Fullsim
W$\to \tau\nu$, $p_{\mathrm{T}}>500$, b-jet filter	Sherpa	CT10	0.10208	Fullsim
W$\to \tau\nu$, $p_{\mathrm{T}}>500$, c-jet filter, b-jet veto	Sherpa	CT10	0.10139	Fullsim
W$\to \tau\nu$, $p_{\mathrm{T}}>500$, c-jet veto, b-jet veto	Sherpa	CT10	0.10201	Fullsim
VBF W$\to e\nu$	Sherpa	CT10	4.2114	Fullsim
VBF W$\to \mu\nu$	Sherpa	CT10	4.2128	Fullsim
Wγ	Alpgen + Jimmy (AUET2)	CTEQ6L1	229.88	Fullsim
W$\gamma+1$ parton, (Lepton/Photon Filter)	Alpgen + Jimmy (AUET2)	CTEQ6L1	59.518	Fullsim
W$\gamma+2$ partons, (Lepton/Photon Filter)	Alpgen + Jimmy (AUET2)	CTEQ6L1	21.39	Fullsim
W$\gamma+3$ partons, (Lepton/Photon Filter)	Alpgen + Jimmy (AUET2)	CTEQ6L1	7.1203	Fullsim
W$\gamma+4$ partons	Alpgen + Jimmy (AUET2)	CTEQ6L1	2.1224	Fullsim
W$\gamma+5$ partons	Alpgen + Jimmy (AUET2)	CTEQ6L1	0.46612	Fullsim

Table 5.5: Z + jets Monte Carlo samples used to model background used in the analysis, with information on the process, MC generator (and tune), cross-section, and detector simulation type. Samples are binned in p_{T}^{W} and also filtered by quark content, as mentioned in the caption for Table 5.4.

Process	Generator (Tune)	PDF Set	$\sigma\times$ BR [pb] (@ $\sqrt{s} = 8$ TeV)	AFII/Fullsim
Z$\to ee$	Sherpa	CT10	1207.8	Fullsim
Z$\to ee$, $40 < p_{\mathrm{T}} < 70$, b-jet filter	Sherpa	CT10	70.48500	AFII
Z$\to ee$, $40 < p_{\mathrm{T}} < 70$, c-jet filter, b-jet veto	Sherpa	CT10	70.53000	AFII
Z$\to ee$, $40 < p_{\mathrm{T}} < 70$, c-jet veto, b-jet veto	Sherpa	CT10	70.43100	AFII
Z$\to ee$, $70 < p_{\mathrm{T}} < 140$, b-jet filter	Sherpa	CT10	29.49400	AFII
Z$\to ee$, $70 < p_{\mathrm{T}} < 140$, c-jet filter, b-jet veto	Sherpa	CT10	29.48700	AFII
Z$\to ee$, $70 < p_{\mathrm{T}} < 140$, c-jet veto, b-jet veto	Sherpa	CT10	29.49100	AFII
Z$\to ee$, $140 < p_{\mathrm{T}} < 280$, b-jet filter	Sherpa	CT10	3.987700	AFII
Z$\to ee$, $140 < p_{\mathrm{T}} < 280$, c-jet filter, b-jet veto	Sherpa	CT10	3.981100	AFII
Z$\to ee$, $140 < p_{\mathrm{T}} < 280$, c-jet veto, b-jet veto	Sherpa	CT10	3.989000	AFII
Z$\to ee$, $280 < p_{\mathrm{T}} < 500$, b-jet filter	Sherpa	CT10	0.241600	Fullsim
Z$\to ee$, $280 < p_{\mathrm{T}} < 500$, c-jet filter, b-jet veto	Sherpa	CT10	0.241280	Fullsim
Z$\to ee$, $280 < p_{\mathrm{T}} < 500$, c-jet veto, b-jet veto	Sherpa	CT10	0.241580	Fullsim
Z$\to ee$, $p_{\mathrm{T}} > 500$, b-jet filter	Sherpa	CT10	0.013235	Fullsim
Z$\to ee$, $p_{\mathrm{T}} > 500$,c-jet filter, b-jet veto	Sherpa	CT10	0.013454	Fullsim
Z$\to ee$, $p_{\mathrm{T}} > 500$, c-jet veto, b-jet veto	Sherpa	CT10	0.013307	Fullsim
Z$\to \mu\mu$	Sherpa	CT10	1207.7	Fullsim
Z$\to \mu\mu$, $40 < p_{\mathrm{T}} < 70$, b-jet filter	Sherpa	CT10	70.48600	AFII
Z$\to \mu\mu$, $40 < p_{\mathrm{T}} < 70$, c-jet filter, b-jet veto	Sherpa	CT10	70.46900	AFII
Z$\to \mu\mu$, $40 < p_{\mathrm{T}} < 70$, c-jet veto, b-jet veto	Sherpa	CT10	70.53400	AFII
Z$\to \mu\mu$, $70 < p_{\mathrm{T}} < 140$, b-jet filter	Sherpa	CT10	29.49100	AFII
Z$\to \mu\mu$, $70 < p_{\mathrm{T}} < 140$, c-jet filter, b-jet veto	Sherpa	CT10	29.44700	AFII
Z$\to \mu\mu$, $70 < p_{\mathrm{T}} < 140$, c-jet veto, b-jet veto	Sherpa	CT10	29.52100	AFII
Z$\to \mu\mu$, $140 < p_{\mathrm{T}} < 280$, b-jet filter	Sherpa	CT10	3.984200	AFII
Z$\to \mu\mu$, $140 < p_{\mathrm{T}} < 280$, c-jet filter, b-jet veto	Sherpa	CT10	3.991100	AFII
Z$\to \mu\mu$, $140 < p_{\mathrm{T}} < 280$, c-jet veto, b-jet veto	Sherpa	CT10	3.984100	AFII
Z$\to \mu\mu$, $280 < p_{\mathrm{T}} < 500$, b-jet filter	Sherpa	CT10	0.242190	Fullsim
Z$\to \mu\mu$, $280 < p_{\mathrm{T}} < 500$, c-jet filter,b-jet veto	Sherpa	CT10	0.241690	Fullsim
Z$\to \mu\mu$, $280 < p_{\mathrm{T}} < 500$, c-jet veto, b-jet veto	Sherpa	CT10	0.242720	Fullsim
Z$\to \mu\mu$, $p_{\mathrm{T}} > 500$, b-jet filter	Sherpa	CT10	0.013161	AFII
Z$\to \mu\mu$, $p_{\mathrm{T}} > 500$, c-jet filter, b-jet veto	Sherpa	CT10	0.013480	AFII
Z$\to \mu\mu$, $p_{\mathrm{T}} > 500$, c-jet veto, b-jet veto	Sherpa	CT10	0.013264	AFII
Z$\to \tau\tau$	Sherpa	CT10	1206.9	Fullsim
Z$\to \tau\tau$, $40 < p_{\mathrm{T}} < 70$, b-jet filter	Sherpa	CT10	70.44100	AFII
Z$\to \tau\tau$, $40 < p_{\mathrm{T}} < 70$, c-jet filter, b-jet veto	Sherpa	CT10	70.53800	AFII
Z$\to \tau\tau$, $40 < p_{\mathrm{T}} < 70$, c-jet veto, b-jet veto	Sherpa	CT10	70.52800	AFII
Z$\to \tau\tau$, $70 < p_{\mathrm{T}} < 140$, b-jet filter	Sherpa	CT10	29.48900	Fullsim
Z$\to \tau\tau$, $70 < p_{\mathrm{T}} < 140$, c-jet filter, b-jet veto	Sherpa	CT10	29.49900	Fullsim
Z$\to \tau\tau$, $70 < p_{\mathrm{T}} < 140$, c-jet veto, b-jet veto	Sherpa	CT10	29.49400	Fullsim
Z$\to \tau\tau$, $140 < p_{\mathrm{T}} < 280$, b-jet filter	Sherpa	CT10	3.988700	Fullsim
Z$\to \tau\tau$, $140 < p_{\mathrm{T}} < 280$, c-jet filter, b-jet veto	Sherpa	CT10	3.988000	Fullsim
Z$\to \tau\tau$, $140 < p_{\mathrm{T}} < 280$, c-jet veto, b-jet veto	Sherpa	CT10	3.987100	Fullsim
Z$\to \tau\tau$, $280 < p_{\mathrm{T}} < 500$, b-jet filter	Sherpa	CT10	0.241900	Fullsim
Z$\to \tau\tau$, $280 < p_{\mathrm{T}} < 500$, c-jet filter, b-jet veto	Sherpa	CT10	0.241020	Fullsim
Z$\to \tau\tau$, $280 < p_{\mathrm{T}} < 500$, c-jet veto, b-jet veto	Sherpa	CT10	0.241470	Fullsim
Z$\to \tau\tau$, $p_{\mathrm{T}} > 500$, b-jet filter	Sherpa	CT10	0.013285	Fullsim
Z$\to \tau\tau$, $p_{\mathrm{T}} > 500$, c-jet filter, b-jet veto	Sherpa	CT10	0.013308	Fullsim
Z$\to \tau\tau$, $p_{\mathrm{T}} > 500$, c-jet veto, b-jet veto	Sherpa	CT10	0.013284	Fullsim

$$w_{\mathcal{I}} = \frac{{\mathcal{M}_{\mathcal{SBI}}}^2 - {\mathcal{M}_{\mathcal{B}}}^2}{{\mathcal{M}_{\mathcal{S}}}^2}. \tag{5.1}$$

A new LHEF is created, which contains the same events as before, but with the appropriate interference weights applied to each event. The lineshapes from these difference components (as well as the final *effective* signal) are shown in Figure 5.1, for four values of m_H.

Especially at lower m_{WW} ($m_{WW} \lesssim {}^{m_H}/_2$), the constructive interference results in these weights becoming extremely large. This poses the problem of having large statistical variations in adjacent m_{WW} bins (a direct result of the limited number of events in the distribution to reweight, and ${\mathcal{M}_{\mathcal{S}}}^2 \ll {\mathcal{M}_{\mathcal{SBI}}}^2 - {\mathcal{M}_{\mathcal{B}}}^2$). The observed statistical fluctuations also occur at very large m_{WW}, where the $\mathcal{S}$ spectrum has only a small contribution.

In order to mitigate the effects of the spuriously large $w_{\mathcal{I}}$ that occur mostly in the low m_{WW} region, an iterative reweighting procedure is applied, which attempts to smooth out the distribution through averaging. The iterative reweighting algorithm, developed to run on top of Repolo, runs Repolo over all events a single time, after which the weighted Les Houches file is parsed, and the weights are examined. A threshold for the weights is constructed ($w_{\mathcal{I}}^{\text{thres.}}$), and if the interference weight calculated for the event is above this threshold, the event is classified as having a large event weight (weights not classified as *large* are in this context considered *small*). The Les Houches event file is then split into two separate files, one containing only events with large weights, and the other containing only small weights. The small weights file are regarded as final, and complete. However, the LHEF file containing the larger weights is then parsed through the Repolo machinery 20 more times (all iterations are given unique random seeds to start the evaluation) to ensure that the large weights present are not a result of the low statistics of the distribution being sampled (thereby attempting to smooth out the statistical jitter). Once the recalculations are complete, the interference weights obtained from both sources are merged.

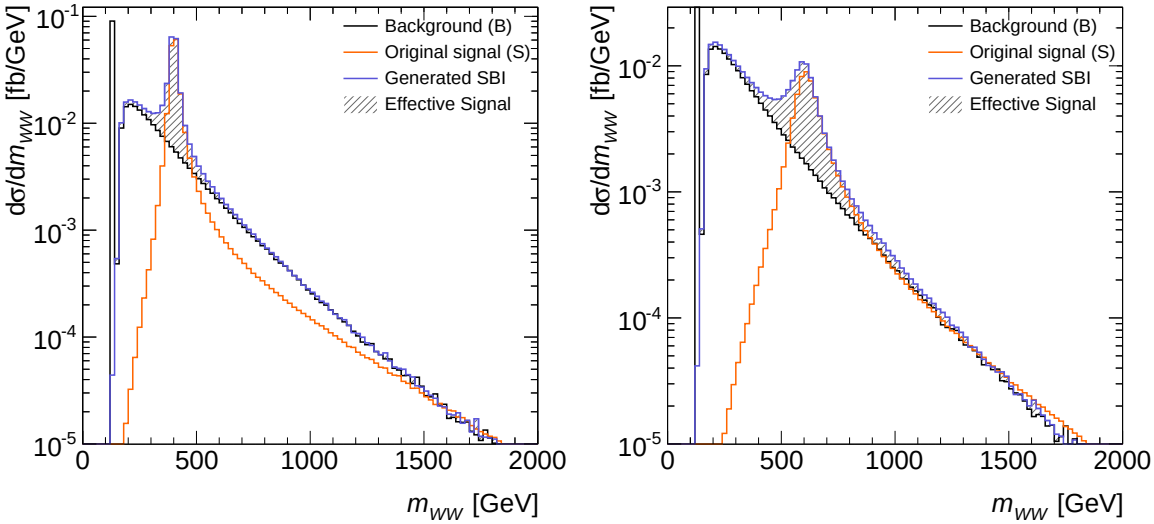
The threshold was empirically determined, and set to a low enough value, ensuring that the majority of low m_{WW} weights were caught ($w_{\mathcal{I}}^{\text{thres.}} = 15 \times ({}^{m_H}/_{m_{WW}})^3$). This results in a interference spectrum with a minimal amount of statistical jitter.

Following this process, the weights are then used to construct a map that assigns a weight based on the event's WW invariant mass in bins of width $\mathrm{d}m_{WW} = 2$ GeV.

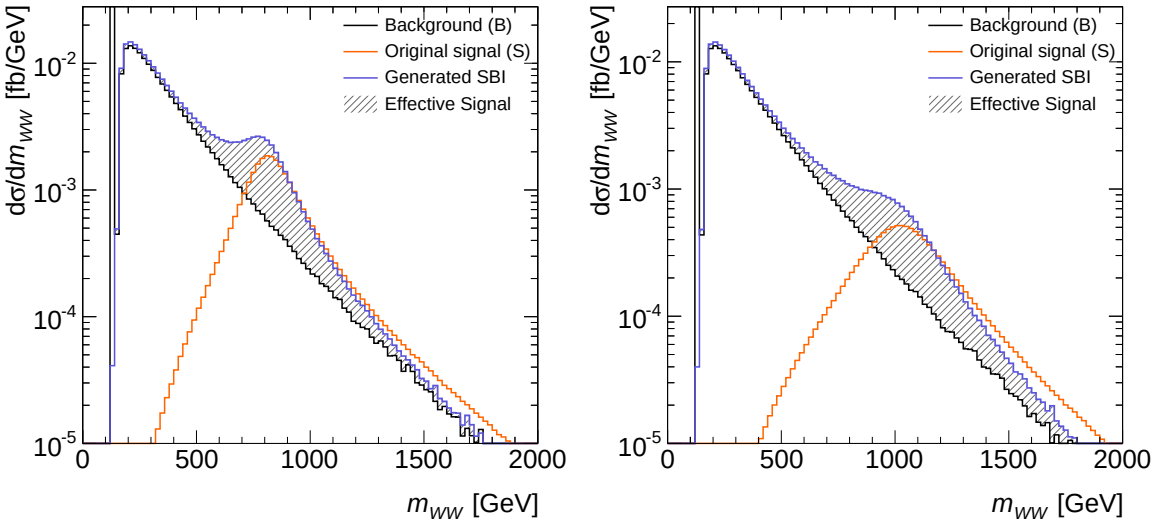
This method ensures that the interference weight spectrum is largely absent of large statistical fluctuations (the exceptions are for the case where m_{WW} is much larger or much smaller than m_H). However, the iterative procedure is not able to completely remove all statistical jitter for events with very low, or very high m_{WW}. This is due to the small number of events in these regions in the original $\mathcal{S}$-only Powheg samples. After consultations with the Repolo authors, it was decided to reweight events in the region $\frac{1}{2}m_H \leq m_{WW} \leq \frac{3}{2}m_H$ normally, but events that fall on the low (high) side of this interval are then reweighted by the same factor as an events with $m_{WW} = \frac{1}{2}m_H$ ($\frac{3}{2}m_H$). The effect of the reweighting on the signal spectrum is shown in Figure 5.2.

5.4.3 Interference with Various Higgs Widths and Lineshapes (EWS)

To study the effect of the interference in the context of the EWS search the interference weights (in both ggF and VBF production modes) were extracted for a series of Higgs widths between $0.2 - 1.0 \times \Gamma_{SM}$. For both production mechanisms, the interference spectra extracted when considering BW and CPS

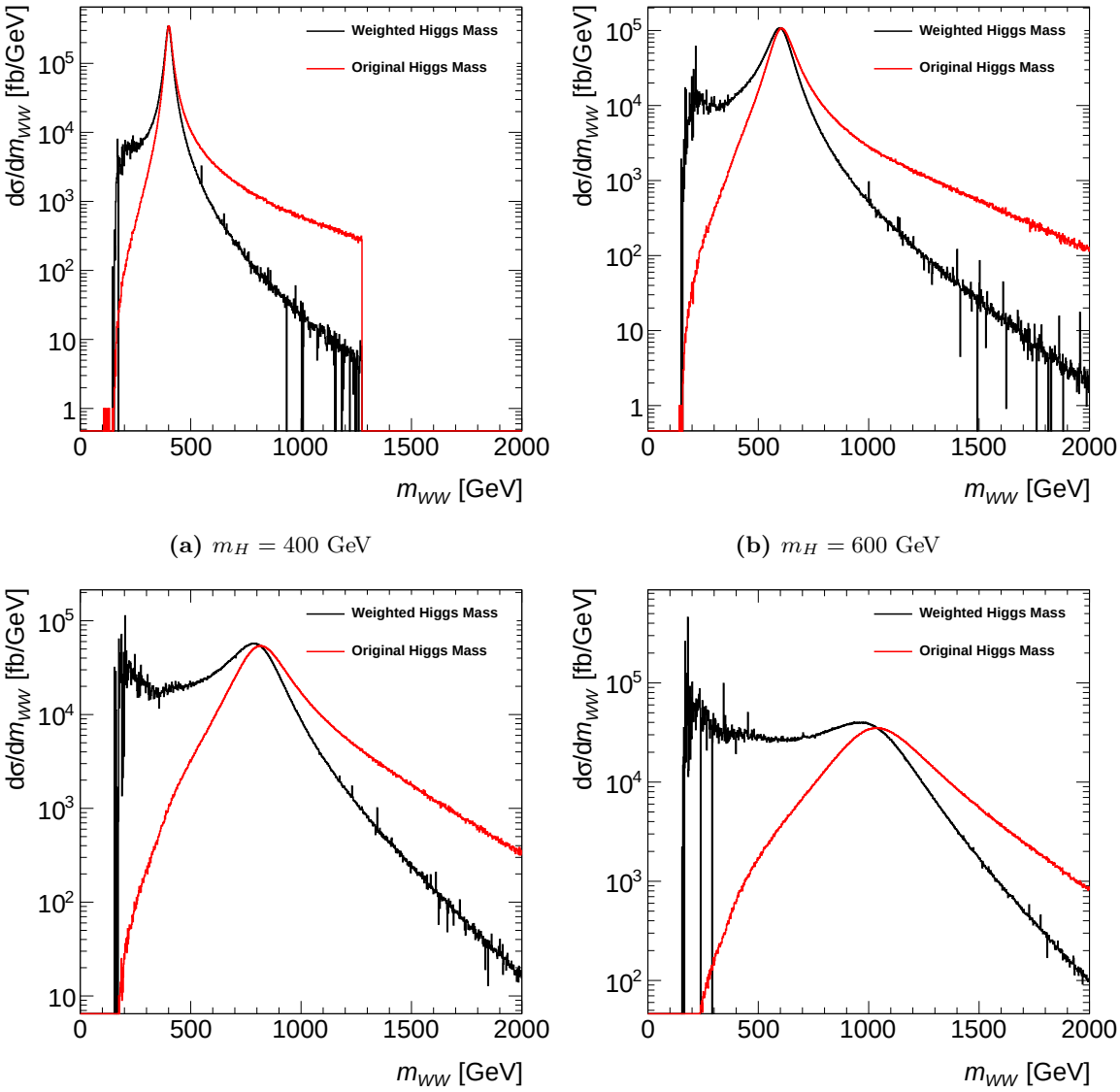


(a) $m_H = 400$ GeV **(b)** $m_H = 600$ GeV



(c) $m_H = 800$ GeV **(d)** $m_H = 1000$ GeV

Figure 5.1: Invariant WW mass spectrum showing the original SM-like signal (signal only), non-resonant background, and the total spectrum when interference is taken into account with Repolo. Also shown is the *effective signal* which is the the $\mathcal{S} + \mathcal{I}$ contribution.



(a) $m_H = 400$ GeV

(b) $m_H = 600$ GeV

(c) $m_H = 800$ GeV

(d) $m_H = 1000$ GeV

Figure 5.2: WW invariant mass spectrum showing the signal shape before and after reweighting with Repolo, in order to account for the interference with the WW continuum background.

lineshapes were found to be the same.

However, in the case of the ggF interference, the interference weights were found to be different for the various widths considered. As a result, separate ggF interference weights were generated for each Higgs mass and width being considered. When transforming the CPS lineshape at $\Gamma = \Gamma_{SM}$, two separate weights need to be applied. The first transforms the CPS lineshape to the BW, with the specified intermediate width. The second then applies the relevant interference weight in order to account for the interference.

5.5 Event Preselection and Classification

The preselection of events in this analysis is a three step process. In the first step, the final, and more stringent object reconstruction requirements are applied (see Section 5.5.1). Second, all events are separated into categories, depending on the object content of the event. In total, there are eight possible categories based on the lepton flavour and charge, the W reconstruction method, and the Higgs production mechanism. The categorization of events will be discussed in Section 5.5.2. The final step of the preselection applies an event based selection, which accepts events based on the topology of the full ensemble of reconstructed objects, and is dependant on the event category determined in the second step.

5.5.1 Object Definition

The objects used in this analysis are electrons, muons, jets, and missing transverse momentum; an initial discussion of how these objects are reconstructed, identified and defined was provided in Chapter 4. A summary of the additional object definition criteria applied in this analysis is found in Table 5.6, with further discussion below.

Primary Vertex

All events in the analysis must have a well defined primary vertex (PV). Cuts on the object selection are applied with respect to the identified PV. Primary vertices are required to have at least three associated tracks that each have $p_{\mathrm{T}} \geq 400$ MeV. If more than one primary vertex is identified, the PV that has the largest sum p_{T}^2 (the sum of each track's p_{T}^2, for all tracks originating from the vertex) is assigned as the event's PV, and the source of the hard-scatter.

Electrons

Electrons are required to pass the *tight* identification selection criteria, have $p_{\mathrm{T}} > 15$ GeV, and are only considered in the region $|\eta| < 2.47$ where the sliding-window clusters in the EM calorimeters can be matched with tracks from the inner detector. Electrons in the small transition region between the EM barrel and end-cap calorimeters ($1.37 \leq |\eta| \leq 1.52$) are rejected.

Electrons are required to pass a calorimeter isolation cut. The variable used is $topoEtConeCor30$ which is constructed by summing the E_{T} for all positive energy topoclusters around the electron inside a $R = 0.3$ cone. The energy of the electron itself, and the estimated contributions from pile-up are subtracted. The ratio of $topoEtConeCor30$ to the electron E_{T} then provides an estimate of the electron calorimeter isolation. The cut applied depends on the electron E_{T}, as shown in Table 5.6.

Table 5.6: Summary of the object definition requirements for leptons and jets. Discussion and definitions of the quantities are found in the text.

Object	Parameter	Requirement
Muons		
	Algorithm:	Staco Combined
	$\lvert z_0 \cdot \sin\theta \rvert$	< 1.0 mm
	$\lvert d_0/\sigma_{d_0} \rvert$	< 3
	Calorimeter Isolation:	$EtConeCor30/E_{\mathrm{T}}$
		< 0.12 $(15 \leq p_{\mathrm{T}} < 20)$
		< 0.18 $(20 \leq p_{\mathrm{T}} < 25)$
		< 0.30 $(p_{\mathrm{T}} \geq 25)$
	Track Isolation:	$PtCone30/E_{\mathrm{T}}$
		< 0.08 $(15 \leq p_{\mathrm{T}} < 20)$
		< 0.12 $(p_{\mathrm{T}} \geq 20)$
	p_{T}	> 15 GeV
	$\lvert\eta\rvert$	< 2.4
Electrons		
	ID:	tight
	$\lvert z_0 \cdot \sin\theta \rvert$	< 0.4 mm
	$\lvert d_0/\sigma_{d_0} \rvert$	< 3
	Calorimeter Isolation:	$topoEtConeCor30/E_{\mathrm{T}}$
		< 0.24 $(15 \leq E_{\mathrm{T}} < 20)$
		< 0.28 $(E_{\mathrm{T}} \geq 20)$
	Track Isolation:	$PtCone30/E_{\mathrm{T}}$
		< 0.08 $(15 \leq E_{\mathrm{T}} < 20)$
		< 0.10 $(E_{\mathrm{T}} \geq 20)$
	p_{T}	> 15 GeV
	$\lvert\eta\rvert$	< 2.47
anti-k_t Jets		
	Radius Parameter:	$R = 0.4$
	p_{T}	> 30 GeV
	$\lvert\eta\rvert$	< 4.5
	JVF	> 0.5
C/A MDF Jets		
	Radius Parameter:	$R = 1.2$
	p_{T}	> 100 GeV
	$\lvert\eta\rvert$	< 1.2
	μ_{frac}	0.67
	y_{cut}	0.09

A track-based isolation requirement is also constructed from the ratio between $PtCone30$ and the E_{T} of the electron, where the quantity $PtCone30$ is analogous to the $topoEtConeCor30$ variable, as it sums the transverse momentum of tracks within a cone of $R = 0.3$, centered on the electron (excluding the contribution from the electron). These tracks must point back to the same primary vertex as the electron track, and have $p_{\mathrm{T}} > 400$ MeV [84]. The cut applied is also E_{T} dependant as summarized in Table 5.6.

Finally, cuts on the electron transverse and longitudinal impact parameters with respect to the primary vertex are applied, in order to reduce the contamination from leptons originating from the decay of heavy flavour quarks. The *transverse impact significance* (the ratio of the transverse impact parameter and it's error), d_0/σ_{d_0} is required to be less than 3. Meanwhile, the longitudinal impact parameter is required to satisfy $|z_0 \cdot \sin\theta| < 0.4$ mm.

Muons

Muons are reconstructed with the *Staco* combined muon algorithm (Section 4.5) which combines information from the ID and MS in order to reconstruct the trajectory and momentum of the muon. To be considered in the analysis, muons must have a $p_{\mathrm{T}} > 15$ GeV and be within the Muon Spectrometer's trigger acceptance of $|\eta| < 2.4$.

As in the case of electrons, muons are required to pass calorimeter and track isolation criteria. The track based isolation variable is the same as is used in the case of electrons, and the cuts applied are also E_{T} dependent. In the case of the calorimeter isolation, the variable used is $EtConeCor30$, which is similar to $topoEtConeCor30$, but sums the raw cell energies inside a cone of $R = 0.3$ instead of using topocluster energies.

Muon candidates are also required to pass cuts on the longitudinal impact parameter and transverse impact parameter significance, to reduce the number of non-prompt muons. The cuts are found in Table 5.6.

Anti-k_t Jets

Jets constructed with the anti-k_t algorithm with a radius parameter $R = 0.4$ are the most commonly used jet reconstruction algorithm in ATLAS. This analysis uses two different types of jet algorithms, and so to separate these two types, we refer to these anti-k_t jets as *resolved* jets.

These jets are only considered up to $|\eta| = 4.5$ and are required to be separated from any reconstructed electrons or muons ($\Delta R > 0.3$). All jets must pass loose selection criteria that distinguish them from energy deposits caused by instrumental effects, cosmic showers or showers generated by beam backgrounds.

There is also a requirement on a quantity known as the jet vertex fraction (JVF) of all resolved jets [98], which suppresses jets not originating from the hard-scatter vertex (the event's PV). The jet vertex fraction is defined as the ratio of the p_{T} of all jet tracks pointing back to a particular primary vertex over the p_{T} of all jet tracks pointing back to any primary vertex (as in Equation 5.2).

$$\mathrm{JVF}(\mathrm{jet_i}, \mathrm{PV_j}) = \frac{\sum_{\mathrm{k}} p_{\mathrm{T}}(track_{\mathrm{k}}^{\mathrm{jet}}, \mathrm{PV_j})}{\sum_{\mathrm{n}} \sum_{\mathrm{l}} p_{\mathrm{T}}(track_{\mathrm{l}}^{\mathrm{jet}}, \mathrm{PV_n})} \tag{5.2}$$

Jets within the ID acceptance of $|\eta| < 2.4$ are required to have JVF > 0.5 with respect to the event's hard-scatter vertex ($\mathrm{PV_j}$ in the above expression). Jets outside the ID acceptance do not have a JVF

requirement.

Resolved jets within $|\eta| < 2.4$ are also flavour tagged with the MV1 (Section 4.6.2) b-tagging algorithm, resulting in a b-tagging efficiency of 85%. It is this tagging of b-jets that allows a significant reduction of Top backgrounds. While the ID covers the region $|\eta| \leq 2.5$, the b-tagged jet acceptance is limited to $|\eta| < 2.4$ because of reduced efficiency of the algorithm when operated at the edge of the ID acceptance.

Large-R Jets and Jet Grooming

The role of jets in this analysis is to first, identify the hadronic showers of the two quarks which originated from the decay of the hadronically decaying W, and second, to identify the recoiling partons in the case of VBF Higgs production. In the low mass region ($m_H \lesssim 600$ GeV) the default resolved jets are able to reconstruct the two W jets without issue. However, when the Higgs is quite heavy, the W decay products (being much lighter than the Higgs) will be significantly boosted. This affects the opening angle of the decay products of the W. For the decay products of a two-body decay of a particle with some p_{T} and mass, m, the opening angle can be approximated as;

$$\Delta R \approx \frac{2m}{p_{\mathrm{T}}}. \tag{5.3}$$

If the opening angle between the jets is $\Delta R \leq 0.4$, the resulting jets will not be resolved by the anti-k_t algorithm with the radius parameter used.

Inserting values into Equation 5.3, the two quarks from the hadronic W decay will typically be within $\Delta R = 0.4$ of each other when $p_{\mathrm{T}}^W \geq 400$ GeV, which means that when the Higgs mass exceeds $m_H \sim 800$ GeV, the two jets will not be individually resolved as anti-k_t jets, and a different strategy must be employed. The calculation assumes that the two W's gain approximately the same transverse momentum in the decay. Further, since the width of the Higgs can be quite large (depending on the model being considered), the effect can be seen at lower m_H. The effect can be plainly seen in Figure 5.3, where the ΔR distribution of the resolved jets reconstructing the W are shown to become less separated as the Higgs mass increases. At Higgs masses above 700 GeV, the distribution is severely truncated at $\Delta R = 0.4$, which is the limit of the resolving power of the $R = 0.4$ jets. This has serious implications for the reconstruction efficiency for signal processes, which will be severely degraded if anti-k_t jets with $R = 0.4$ are used as the sole method to reconstruct the hadronically decaying W.

An easy solution is to expand the radius parameter of the jet, and expect the jet to capture both quark hadronizations, instead of having two discrete *resolved* jets. Expanding the jet size has consequences in ATLAS, where the pile-up conditions are severe; larger jets have a higher probability of unrelated energy (pile-up) being deposited in the jet active area, biasing the energy measurement of the hadronically decaying W. This leads to the use of so-called *jet substructure techniques*, which remove soft, pile-up related contributions from the jet, and examine the physical structure of the proto-jets, and clusters that make up the jet (which can be used as a further background discriminant). In order to study the proto-jet constituents, and remove the soft contributions, the jet must usually be *declustered*, *i.e.*, the clustering history of each jet (*i.e.*, which proto-jets were added to the jet in which order) is examined in reverse.

The large-R jet (also referred to here as a *merged* jet) algorithm used in this analysis is a C/A jet, with a radius parameter of $R = 1.2$ (the C/A jet algorithm is discussed in Section 4.6.1). A grooming

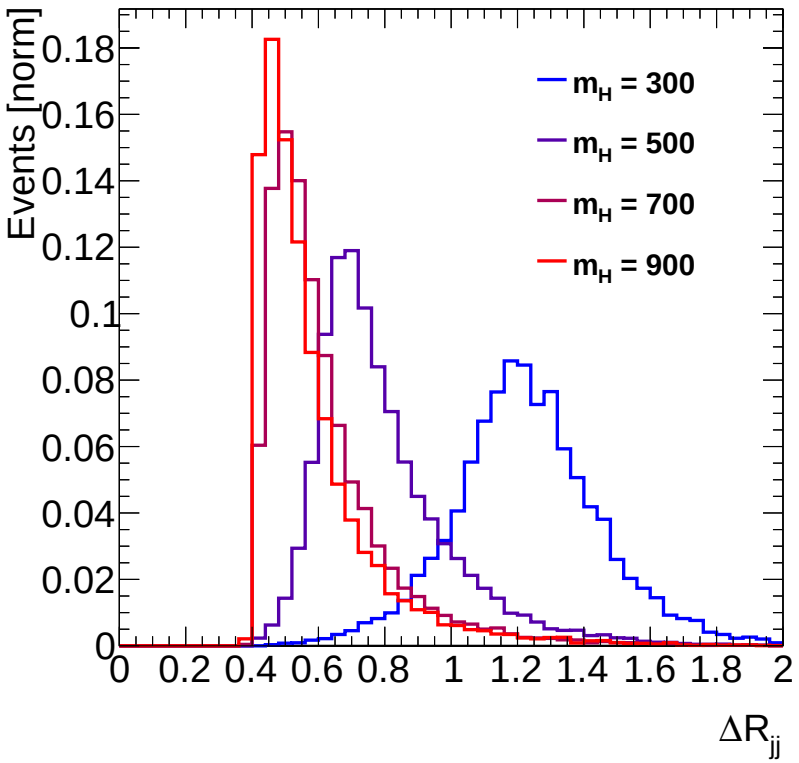
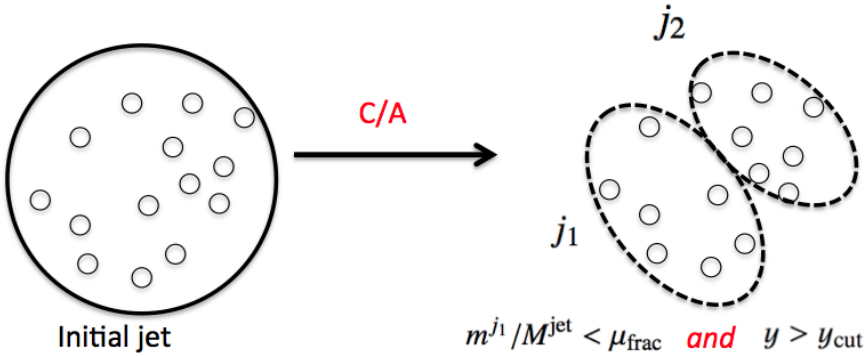
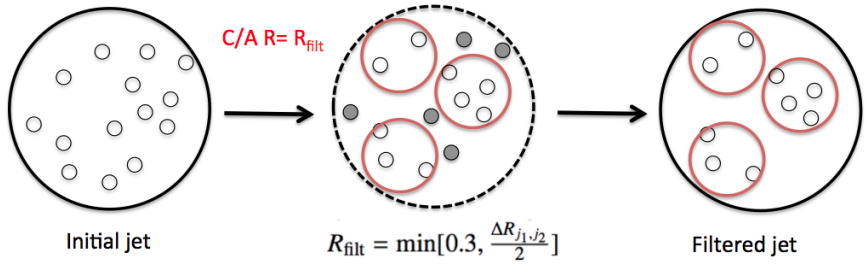


Figure 5.3: ΔR of the resolved W-jets after the preselection stage for different values of m_H (NWA signal model). Unlike the analysis preselection (described in Section 5.5.3), only resolved jets were used.



(a)



(b)

Figure 5.4: Pictorial representation of the two step mass drop filter algorithm. In (a) the C/A jet is declustered, and the two proto-jets must fulfill mass drop (μ_{frac}) and symmetry (y_{cut}) requirements. In (b) the filtering step is illustrated, where the C/A jet constituents are again built into jets, with a new radius parameter R_{filt}. All but the three hardest proto-jets are discarded [122].

technique called a mass drop filter (MDF) is applied to these jets [122], which serves a two-fold purpose. First, it almost completely eliminates the dependence of the jet energy and mass on pile-up conditions, and second, it provides a level of background rejection against events where the jet does not originate from the decay of a boosted object.

The MDF is applied in a two step procedure. First, the jet must first pass the mass drop and symmetry criteria, which requires that the jet has the topological signature of a two body decay. Second, the jet is reclustered, removing all but the highest p_{T} sub-jets in an attempt to select only the energy deposits originating from the W decay products. This two step process is explained in more detail below, and illustrated in Figure 5.4 [122].

Mass Drop and Symmetry Criteria First, the last clustering step of the C/A algorithm is undone. From Equation 4.1, with $p = 0$, the C/A algorithm starts by constructing proto-jets from nearby clusters, and in subsequent iterations combines proto-jets with larger separation. Therefore, the last step in the clustering will combine two well separated proto-jets, as long as they are within the radius parameter set in the algorithm. Undoing this last step then splits the jet into two sub-jets ($j1$, $j2$), which have a large ΔR separation (Figure 5.4a). The mass drop criteria then requires that there is a significant mass drop from the mass of the original jet, m_J, compared to the mass of the heavier of the two sub-jets

(here, m_{j1} is arbitrarily chosen to be larger than m_{j2}). The ratio of these two masses must be less than some pre-defined value μ_{frac}:

$$\frac{m_{j1}}{m_J} < \mu_{frac}. \tag{5.4}$$

The symmetry requirement demands that the two sub-jets are somewhat symmetric, *i.e.* they both contribute a reasonable amount of p_{T} to the jet. The symmetry requirement is:

$$\frac{\min\left[(p_{\mathrm{T}}{}^{j1})^2, (p_{\mathrm{T}}{}^{j2})^2\right]}{m_J^2} \times \Delta R_{j1,j2}^2 > y_{cut}, \tag{5.5}$$

with y_{cut} representing the degree of symmetry required by the specific MDF algorithm, which can be tuned.

In the case that either of the two requirements are not met, the jet is discarded, and not considered further.

Filtering Once the C/A jet passes the above mass drop and symmetry requirements, the jet is filtered in order to remove all but the hardest energy deposits (Figure 5.4b). To do so, the C/A jet algorithm is run once again on the topoclusters/particles making up the the original jet. For this re-clustering, the radius parameter is changed to $R = R_{filt}$ where R_{filt} is given by

$$R_{filt} = \min\left[0.3, \frac{\Delta R_{j1,j2}}{2}\right]. \tag{5.6}$$

The resulting subjets are then filtered, meaning that all of the sub-jets outside of the three highest p_{T} sub-jets are removed from the jet. Three jets are kept instead of two to allow for the capture of gluon radiation from one of the decay quarks.

In this analysis, the values of the cuts used for the large-R C/A MDF *merged* jets are $\mu_{frac} = 0.67$ and $y_{cut} = 0.09$ (see Table 5.6). This is a somewhat standard configuration for MDF C/A jets in ATLAS, with the cuts originally tuned for tagging $H \to b\bar{b}$ decays [123].

For a merged jet to be considered it must also have $p_{\mathrm{T}} > 100$ GeV and be within $|\eta| < 1.2$. The mass of the final groomed jet is taken to be the mass of the summed (massless) clusters making up the final jet,

$$m_J = \left[\left(\sum_i^{N_{clusters}} E_i\right)^2 - \left(\sum_i^{N_{clusters}} \vec{\mathbf{p}}_i\right)^2\right]^{1/2}. \tag{5.7}$$

5.5.2 Event Classification

Events are separated into different categories depending on their topologies and object content. The obvious categorization is the separation of event by charge and flavour (e^+, e^-, μ^+, μ^-). The separation of lepton charge and flavour (which propagates to the final fit) allows for a more sensitive final result, as the expected background – and therefore sensitivity – in each of the charge/flavour channels is different. These events are all treated identically in the preselection. The other categories, which result in differences in the preselection requirements, are the hypothesized production mechanism (ggF vs

VBF) and whether an event reconstructs the hadronically decaying W with two resolved jets, or a single large-R jet (a *resolved* event vs a *merged* event). These two categories are discussed further below.

Production Categories

The categories that represent the two production mechanisms are treated orthogonally. The selection for VBF-like events searches for the two hard, separated jets that are a result of the hadronization of the recoiling partons. These jets are referred to as *tag jets* (as they serve to *tag* the VBF process). The event is declared to fall into the VBF category if it contains two resolved jets that satisfy all of the following three conditions.

- $\mathbf{m_{tagjets} > 600\ GeV}$
  Requires that the two tag jets are energetic and well separated. The mass of the tag jets is taken as the mass of the two tag jet's summed four-vectors.
- $\mathbf{p_T^{tagjet_1} > 40\ GeV}$
  At least one of the two tag jets must have an appreciable transverse momentum.
- $\mathbf{|\Delta y_{tagjet_1,tagjet_2}| > 3}$
  There must be a relatively large rapidity gap between the two tag jets.

Resolved or Merged Category

The method of W reconstruction needs to be chosen, either by two resolved jets, or a single merged jet, on an event by event basis. For each event, both solutions are attempted, and if both solution are possible, the best solution (explained below) is chosen.

Resolved Solutions For a resolved jet to be considered as one of the W-jets it must *not* be identified as one of the VBF tag jets, and must be fairly central in the detector ($|\eta| < 2.4$). All possible pairs of jets are examined, and the pair that gives the best agreement with the W mass is chosen as the resolved solution. A small exception is the case when more than one pair of resolved jets fall close to the W mass (within the so-called W-mass window, $65 \leq m_{jj} \leq 96$). In this case, the pair that has the largest scalar sum p_{T} is chosen, instead of the pair closest to the pole mass. This is done to avoid losing signal events as a result of the p_{T} requirement on the W-jets, that takes place in a later stage of the preselection.

Merged Solutions The merged jet solutions are more straightforward. The merged solution is simply the large-R jet that has the mass with the best agreement with the W mass.

For every event, both solutions are considered, and the solution that gives the best agreement with the W-mass is chosen. Events using resolved dijets to reconstruct the W are classified as *resolved* events, while those using large-R jets with MDF are classified as *merged* events.

5.5.3 Event-level Preselection

This concludes the description of the cuts applied to the reconstructed objects, and the different event categories (which in the case of the merged/resolved and ggF/VBF categories will affect the preselection cuts that are applied). The rest of the preselection (this section) describes requirements on the events,

Table 5.7: Summary of event preselection cuts for $H \to WW \to \ell\nu jj$ analysis. Cuts that are only applied to events with a resolved dijet pair reconstructing the hadronic W are termed "RE" (resolved events), while those only applying to events using a large-R jet to resolve the W are termed "ME" (merged events). All other cuts are applied to both types of events. A description of the cuts not explicitly mentioned in the text is included at the bottom of the table.

	ggF Selection	**VBF Selection**
Basic topology		
	1 isolated charged lepton, $p_{\mathrm{T}} > 25$ GeV	
	veto events with additional leptons ($p_{\mathrm{T}} > 15$ GeV)	
RE:	≥ 2 jets + ≥ 0 groomed jets	≥ 4 jets + ≥ 0 groomed jets
ME:	≥ 0 jets + ≥ 1 groomed jets	≥ 2 jets + ≥ 1 groomed jets
Missing transverse momentum	$E_{\mathrm{T}}^{\mathrm{miss}} \geq 60$ GeV	
VBF topology		
	No two resolved jets satisfying:	Two resolved jets satisfying:
	$m_{\mathrm{tagjets}} > 600$ GeV	
	$p_{\mathrm{T}}^{\mathrm{tagjet}_1} > 40$ GeV	
	$\lvert\Delta y(\mathrm{tagjet}_1, \mathrm{tagjet}_2)\rvert > 3.0$	
Decay topology		
	$\Delta\phi(\ell, \mathrm{E}_{\mathrm{T}}^{\mathrm{miss}}) < 2.5$	
RE:	$p_{\mathrm{T}}^{W-\mathrm{jet}_1} > 60$ GeV	
	$\Delta\phi_{jj}^{W} < 2.5$	
	$\Delta\phi(W\text{−jet}_{1,2}, \ell) > 1.0$	
	$\Delta\phi(W\text{−jet}_{1,2}, E_{\mathrm{T}}^{\mathrm{miss}}) > 1.0$	
ME:	$\Delta\phi(W\text{−jet}, \ell) > 1.0$	
	$\Delta\phi(W\text{−jet}, E_{\mathrm{T}}^{\mathrm{miss}}) > 1.0$	
b-tagging		
RE:	*veto* events with b-tagged (non-W) jets	
	veto events with both W-jets b-tagged	
ME:	*veto* events with any b-tagged jet that is $\Delta R \geq 0.4$ from the W groomed jet	
W-mass window		
RE:	$65 \leq m_{jj} \leq 96$ GeV	
ME:	$65 \leq m_{J} \leq 96$ GeV	

$\Delta\phi(\ell, \mathrm{E}_{\mathrm{T}}^{\mathrm{miss}})$	Azimuthal angle between the lepton and neutrino
$p_{\mathrm{T}}^{W-\mathrm{jet}_1}$	Transverse momentum of the resolved W-jet with the largest p_{T}
$\Delta\phi_{jj}^{W}$	Azimuthal separation between the two resolved W-jets
$\Delta\phi(W\text{−jet}_{1,2}, \ell)$	Azimuthal separation between both W-jets and the single charged lepton
$\Delta\phi(W\text{−jet}_{1,2}, E_{\mathrm{T}}^{\mathrm{miss}})$	Azimuthal separation between both W-jets and the neutrino
$\Delta\phi(W\text{−jet}, \ell)$	Azimuthal separation between the merged W-jet and the single charged lepton
$\Delta\phi(W\text{−jet}, E_{\mathrm{T}}^{\mathrm{miss}})$	Azimuthal separation between the merged W-jet and the neutrino

and the objects previously defined. While the previous sections discussed how to rigorously define objects, (which can reduce backgrounds which involve "fake" reconstructed objects) the following selection cuts are intended to separate the background processes from the signal, based on the final state topologies.

The full set preselection cuts (some of which were briefly mentioned previously) are described in full here, with a summary of the applied requirements shown in Table 5.7.

Primary Vertex Each event is required to have at least one primary vertex, with at least three associated tracks that have p_{T} of at least 400 MeV.

Trigger Every event must pass one of the single lepton triggers outlined in Section 4.2.

Basic topology The event must have *only* one isolated lepton candidate ($p_{\mathrm{T}} > 25$ GeV), and either two central resolved jets to reconstruct the W, or one merged jet. In the case of the VBF production, two additional resolved jets are required, which serve as *tag* jet candidates. As the event can only have one lepton candidate, events with an additional lepton ($p_{\mathrm{T}} > 15$ GeV) are vetoed.

Event Cleaning Throughout the 2012 run, for a variety of reasons, some Tile calorimeter channels were permanently or temporarily masked (the energy readout of the channels is not used). Corrections are normally applied based on the energy of neighbouring channels in the case of masked cells. However, in the case that a large group of neighbouring Tile cells are masked, these corrections were found to either overestimate the energy in the masked cell (in the case of a jet close to the masked module), or underestimate the energy correction (in the case that the jet's core is contained in the masked region). As a result, events with jets that fall into one of these large masked areas are vetoed. Although the problem with masked cell corrections only affects real data, the behaviour is modelled in all Monte Carlo samples, for consistency.

Missing Transverse Momentum In order to select events that have missing transverse momentum from "real" neutrino candidates, the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut is set rather high, to $E_{\mathrm{T}}^{\mathrm{miss}} > 60$ GeV. This greatly reduces the contribution of the multijet background.

VBF topology This cut separates the candidate events into the orthogonal gluon-gluon fusion and vector boson fusion categories by attempting to find two VBF *tag* jets.

Decay topology The event must have the basic topology that would be expected from one leptonic W decay, and one hadronic. Specifically, the lepton and neutrino would be expected to be close together, as would the W-jets (in the case of a resolved event). Angular separation between the W-jet(s) and both the lepton and neutrino (showing angular separation between the two W's) is also required.

***b*-tagging** Events that have a b-tagged resolved jet, that is not one of the W-jets, are vetoed. Events where *both* resolved W-jets are b-tagged are also vetoed; the requirement that both W-jets be b-tagged is done to avoid vetoing signal events where the hadronic W decays into a cs pair (due to the large mistag rate of charm quarks by the tagging algorithm) [101]. These cuts are only relevant for resolved jets which fall within the b-tagging acceptance of $|\eta| < 2.4$. If the event is a merged event, no flavour tagging is applied to the large-R jet. However, if a resolved jet sits close to the merged jet ($\Delta R \leq 0.4$),

the resolved jet is treated as if it where a W-jet, and no b-tagging check is applied. The motivation for this requirement is the same as for the resolved case: to limit the number of signal events discarded due to a mistagged charm quark originating from a W decay.

W-mass window The W-jet(s) mass is required to fall within a ∼15 GeV window around the true W mass ($65 \leq m_{jj/J} \leq 96$). In the case of a merged event, the W mass is just the merged jet's mass. In the case of a resolved event, the mass is taken as the mass of the two jet's summed four-vectors.

Impact of Large-R Jets on Preselection

The inclusion of the large-R groomed jets impacts the analysis and the selection efficiency for signal events in a large way, particularly at high mass. The efficiency gains for the SM-like signal model is illustrated in Figure 5.5, while for the NWA signal model, the relevant figure is Figure 5.6. The fraction of events that use a large-R jet is seen in Figure 5.7 (SM-like) and Figure 5.8 (NWA). As expected, as the Higgs mass grows, the fraction of events that use the large-R jets also increases.

5.6 Background Control Regions

In order to check the modelling of the Monte Carlo generated backgrounds with respect to the data, control regions are defined. The control regions used in this analysis are regions of phase space with an enriched sample of top backgrounds and a region with an enriched sample of W+jets background. A third set of control regions, the multijet control regions (MJCR), serve to provide a data-driven estimate of the multijet background (to be discussed in Section 5.8).

We define control regions that are close in phase space to the already defined signal region (SR, defined by the cuts presented in Section 5.5.3). Usually this is done by simply reversing one of the cuts applied (for example, reversing a b-tagging cut would result in an enriched sample of top events).

The W control region (WCR) has the same cuts and selection as the nominal preselection, up to the final W-mass window cut. The WCR requires the hadronic W mass (m_{jj}/m_J) be within one of two *sidebands* around the nominal mass window (called the upper and lower sideband). The WCR sidebands are defined for the ggF category as follows:

$$52 \leq m_{jj/J} < 65 \text{ GeV} \quad \text{(lower sideband)} \tag{5.8}$$

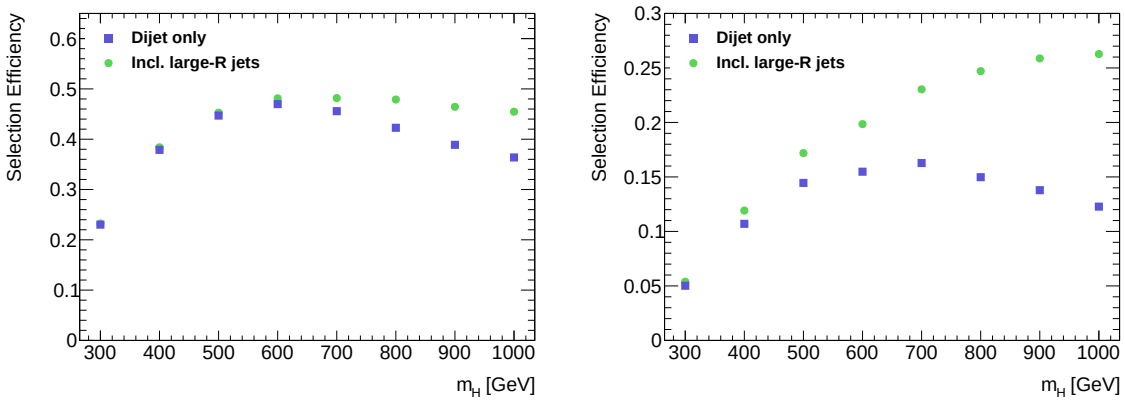
$$96 < m_{jj/J} \leq 126 \text{ GeV} \quad \text{(upper sideband)} \tag{5.9}$$

In the VBF category, the sideband is wider on both the upper and lower side of the m_W window in order to provide sufficient statistics.

$$43 \leq m_{jj/J} < 65 \text{ GeV} \quad \text{(lower sideband)} \tag{5.10}$$

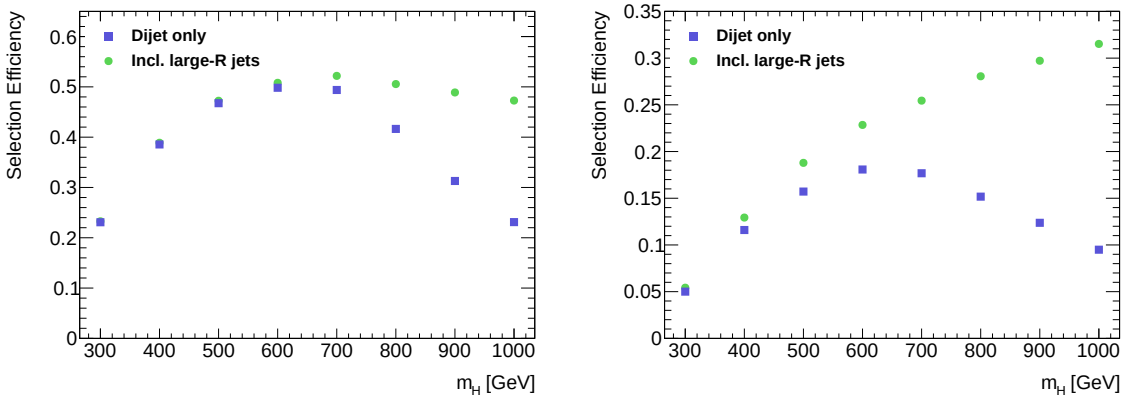
$$96 < m_{jj/J} \leq 200 \text{ GeV} \quad \text{(upper sideband)} \tag{5.11}$$

The top control region (TCR) is not defined by a simple reversal of the b-tagging cuts defined in the preselection. Instead, the veto on doubly tagged W-jets is maintained, and only the non W-jets b-tag cut is reversed.



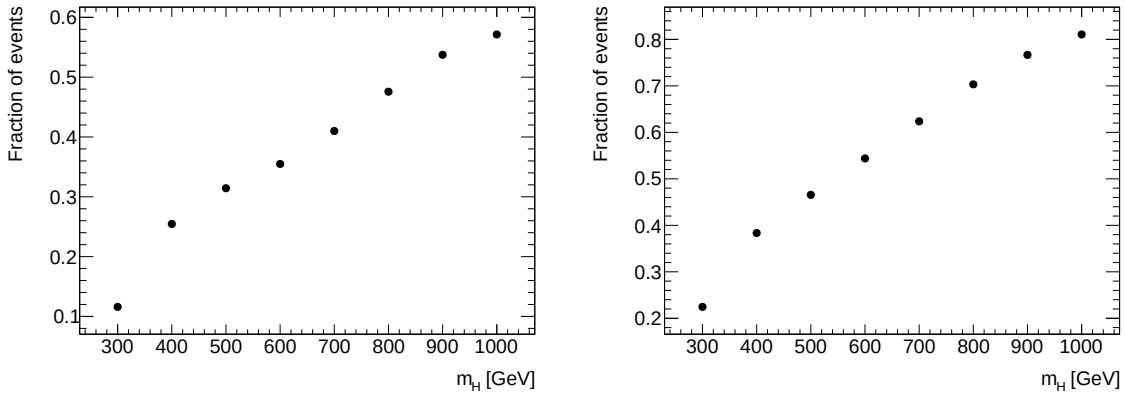
(a) ggF **(b)** VBF

Figure 5.5: Preselection efficiency for signal events (SM-like) for both the ggF (a) and VBF category (b). Shown separately are the efficiencies in the case that only resolved dijet solutions are allowed to reconstruct the hadronic W and the case used in this analysis, where either resolved or large-R jets are able to reconstruct the W.



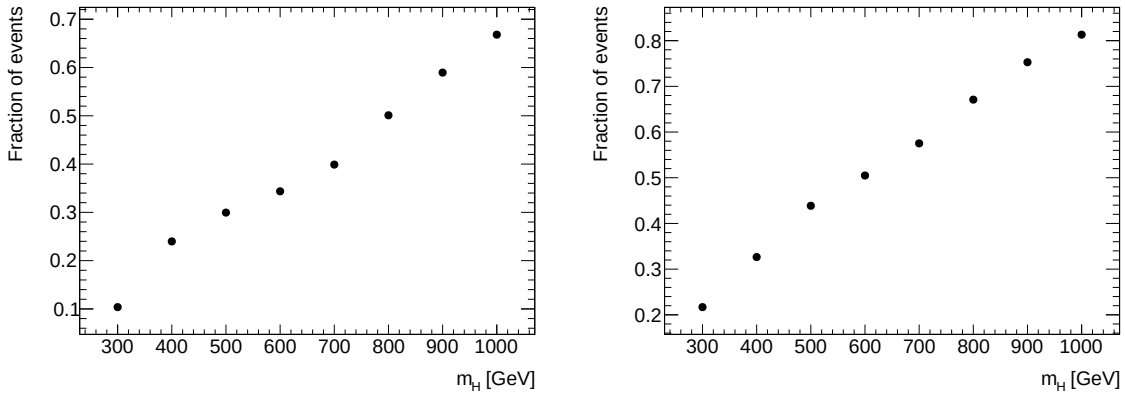
(a) ggF **(b)** VBF

Figure 5.6: Preselection efficiency for signal events (NWA) for both the ggF (a) and VBF category (b). Shown separately are the efficiencies in the case that only resolved dijet solutions are allowed to reconstruct the hadronic W and the case used in this analysis, where either resolved or large-R jets are able to reconstruct the W.



(a) ggF **(b)** VBF

Figure 5.7: Fraction of SM-like signal events using a large-R jet to reconstruct the hadronic W as a function of m_H for the ggF category (a) and VBF category (b).



(a) ggF **(b)** VBF

Figure 5.8: Fraction of NWA signal events using a large-R jet to reconstruct the hadronic W as a function of m_H for the ggF category (a) and VBF category (b).

The final set of CRs are the multijet control regions (MJCR). Events in the MJCR are used to define the multijet templates, which are used to determine the multijet background estimate, as will be described in Section 5.8. Each of the above regions (SR, WCR, TCR) have their own associated MJCR which have the same selection, with the following exceptions: no $E_{\mathrm{T}}^{\mathrm{miss}}$ cut is applied, and the lepton definitions are altered. Electrons are required to pass the looser *medium* identification requirements, but fail the *tight* identification requirements while for muons the cut on the impact parameter significance reversed.

5.7 Higgs Mass Reconstruction

For Higgs decays into the $WW \to \ell\nu jj$ final state the Higgs invariant mass m_{WW} (equivalently written as $m_{\ell\nu jj}$ here)[1] can be reconstructed only if the neutrino momentum is known. However, the neutrino momentum is only measured in the transverse plane, determined by the direction and magnitude of the missing transverse momentum. If the invariant mass of the leptonically decaying W is constrained to the W-mass, then the longitudinal component of the neutrino momentum can be solved for analytically.

Recall the definition of the Lorentz invariant mass for a particle with four-momentum p;

$$m^2 \equiv p^2 = E^2 - \vec{\mathbf{p}}^2. \tag{5.12}$$

The invariant mass of the leptonic W is given by mass of the summed W decay products' four vectors. By imposing the leptonic W mass constraint, the neutrino z-component of the momentum, p_{z}^{ν}, can be solved for as follows:

$$p_W^2 = m_W^2 = p_\nu^2 + p_\ell^2 + 2p_\nu \cdot p_\ell \tag{5.13}$$

$$m_W^2 = \underbrace{p_\nu^2 + p_\ell^2}_{\approx 0} + 2p_\nu \cdot p_\ell. \tag{5.14}$$

Expanding Equation 5.14, and solving for p_{z}^{ν} yields:

$$p_{\mathrm{z}}^{\nu} = \frac{\xi p_{\mathrm{z}}^{\ell}}{{E^{\ell}}^2 - {p_{\mathrm{z}}^{\ell}}^2} \pm \left[\underbrace{\left(\frac{\xi p_{\mathrm{z}}^{\ell}}{{E^{\ell}}^2 - {p_{\mathrm{z}}^{\ell}}^2} \right)^2 + \frac{\xi^2 - {E^{\ell}}^2 {p_{\mathrm{T}}^{\nu}}^2}{{E^{\ell}}^2 - {p_{\mathrm{z}}^{\ell}}^2}}_{\equiv \beta} \right]^{1/2}, \tag{5.15}$$

$$\xi \equiv p_{\mathrm{T}}^{\nu} p_{\mathrm{T}}^{\ell} \cos\left(\phi_\ell - \phi_\nu\right) + \frac{m_W^2}{2}.$$

Equation 5.15 is similar to the solutions of the quadratic equation, and therefore the form of the solution (and the number of solutions) depends on the sign of β. In the case that $\beta > 0$, there are two real solutions. If $\beta < 0$, the neutrino momentum has an imaginary component. In this (unphysical) case, the neutrino momentum is taken to be the real part of the solution. Otherwise, the solution that yields the smallest absolute longitudinal momentum $|p_{\mathrm{z}}^{\nu}|$ is chosen. The final expression for the neutrino longitudinal momentum is:

[1] Although written $m_{\ell\nu jj}$, the reconstructed Higgs mass can also be evaluated using a merged jet instead of two resolved jets ($m_{\ell\nu J}$).

$$p_{\mathrm{z}}^{\nu} = \min\left\{\Re\left[\frac{\xi p_{\mathrm{z}}^{\ell}}{E^{\ell\,2} - p_{\mathrm{z}}^{\ell\,2}} \pm \sqrt{\beta}\right]\right\}. \tag{5.16}$$

Picking the smaller $|p_{\mathrm{z}}^{\nu}|$ solution was found to yield a result closest to the true value ∼70% of the time for the low mass Higgs samples. At higher Higgs masses, the performance degrades slightly, to about 60%. This efficiency was found to be independent of the production mechanism, and pile-up conditions (number of reconstructed vertices). The different longitudinal momentum solutions, with the true neutrino momentum overlaid, are shown in Figure 5.9 for several mass points (for a SM-like Higgs produced via ggF). Due to the rather modest efficiency for picking the correct solution[1], the $m_{\ell\nu jj}$ distribution was examined to see whether *incorrect* choices can greatly bias the reconstructed Higgs mass. These distributions are shown for multiple NWA mass points in Figure 5.10 (due to the narrow width, one would expect the NWA mass peak to be the most heavily affected by a mismodelled p_{z}^{ν}). As shown, choosing the *incorrect* p_{z}^{ν} has only a small effect on the reconstructed mass.

Experimental Resolution

The experimental mass resolution is shown in Figure 5.11. This was determined by fitting the peak of the $m_{\ell\nu jj}$ spectrum for NWA signal samples. Since the NWA samples are generated with a negligible width ($\Gamma = 4.07$ MeV), the width of the spectrum is governed purely by the resolution of the reconstruction.

5.8 Multijet Background

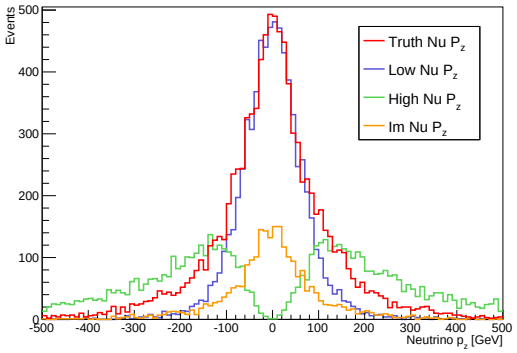
The multijet background dominates in the region of low $m_{\ell\nu jj}$, but falls off sharply at higher values, and is easily suppressed with a cut on the event's $E_{\mathrm{T}}^{\mathrm{miss}}$. The multijet contribution is extremely difficult to model in MC samples, and so a data driven technique, using a template fit to the $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum, is used in the analysis.

The multijet $E_{\mathrm{T}}^{\mathrm{miss}}$ templates are constructed from **data** that falls into the MJCRs. Other backgrounds that are modelled by MC, such as Top and W+jets, may contaminate the MJ templates, and therefore the contribution from these backgrounds is subtracted from the MJ template (this contamination is modelled from MC events which fall into one of the MJCRs). These multijet templates, together with the MC modelled backgrounds is then fit to the SR/WCR/TCR $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum, with all cuts applied except for the one on $E_{\mathrm{T}}^{\mathrm{miss}}$.

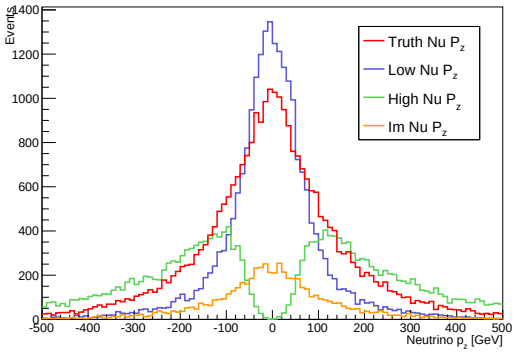
For the WCR and the TCR, the multijet component is estimated using dedicated MJCRs. However, an estimate for the multijet contribution to the signal region cannot use it's associated MJCR, as this would involve fitting to the $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum in the signal region, where there could be a large fraction of signal present. Therefore, the signal region MJ yield is obtained from the $E_{\mathrm{T}}^{\mathrm{miss}}$ fit in the WCR, as the WCR has the selection closest to that of the CR.

This MJ template fit is done separately for the ggF and VBF categories, and for each Higgs mass hypothesis, due to the mass dependant selection. For the multijet estimates for different mass hypothesis, all mass dependant cuts are applied, except the one on $E_{\mathrm{T}}^{\mathrm{miss}}$ (see Section 5.9). The systematic uncertainty on the multijet background normalization is determined by using an alternate multijet template with a different definition for the loosened lepton selection, as will be described in Section 5.12.5.

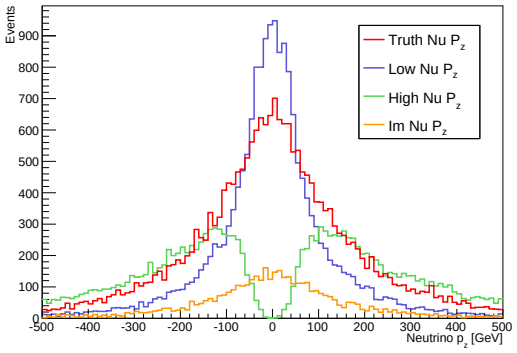
[1] In the case that two real solutions for p_{z}^{ν} are possible the *correct* solution is defined as being the solution closest to the true p_{z}^{ν} (and is considered *incorrect* otherwise).



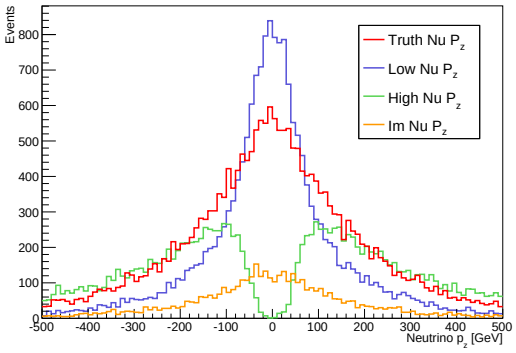
(a) $m_H = 300$ GeV



(b) $m_H = 500$ GeV

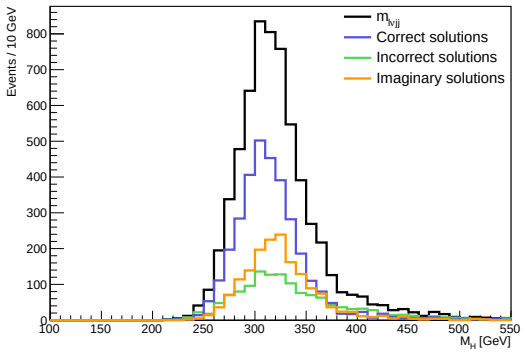


(c) $m_H = 700$ GeV

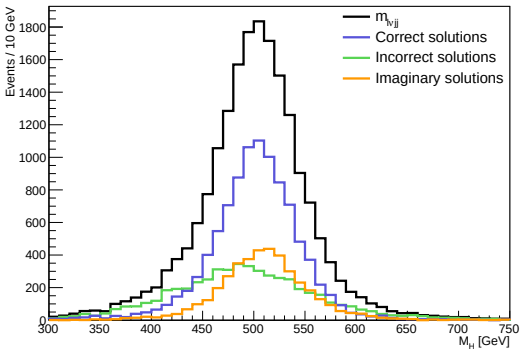


(d) $m_H = 900$ GeV

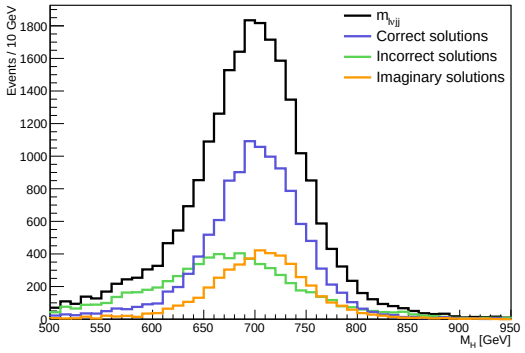
Figure 5.9: The three possible p_z^ν solutions for various m_H points (for a SM-like Higgs produced via ggF) with the true p_z^ν overlaid. In the case that two real solutions are possible, they are both shown on the plot, while in the case of an imaginary component, only the real component is shown. Event yields (which determine the relative normalizations) are after preselection.



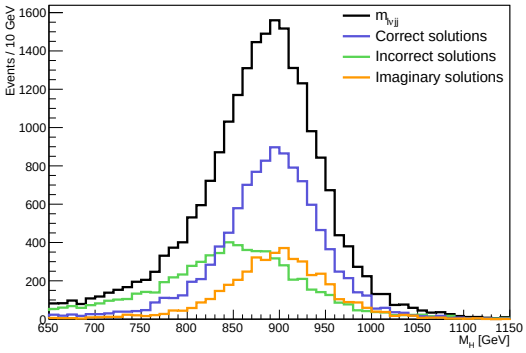
(a) $m_H = 300$ GeV



(b) $m_H = 500$ GeV



(c) $m_H = 700$ GeV



(d) $m_H = 900$ GeV

Figure 5.10: Reconstructed Higgs mass, $m_{\ell\nu jj}$ for events passing the preselection (ggF, NWA). The components making up the the final spectrum are shown separately, depending on whether the solution chosen for the neutrino p_z came from an imaginary solution, a correct solution, or an incorrect solution.

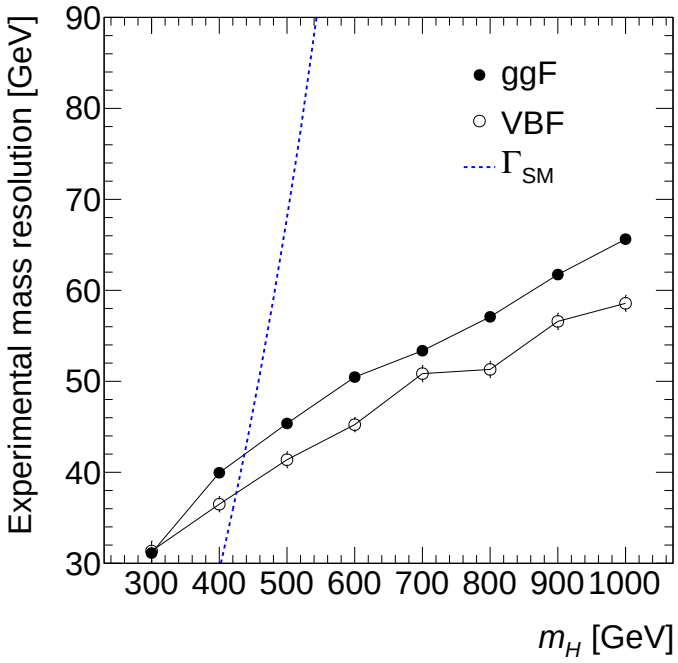


Figure 5.11: The experimental mass resolution (determined by NWA samples) as a function of mass. The blue dashed line shows the width of a SM-like Higgs boson.

As an illustrative example, Figure 5.12 shows the MJ template fits after the preselection are shown for both the ggF and VBF categories in the WCR and TCR (summed over lepton charge and flavour). The remainder of the multijet template fits, are found in Appendix A.

5.9 Mass Dependent Selection Criteria

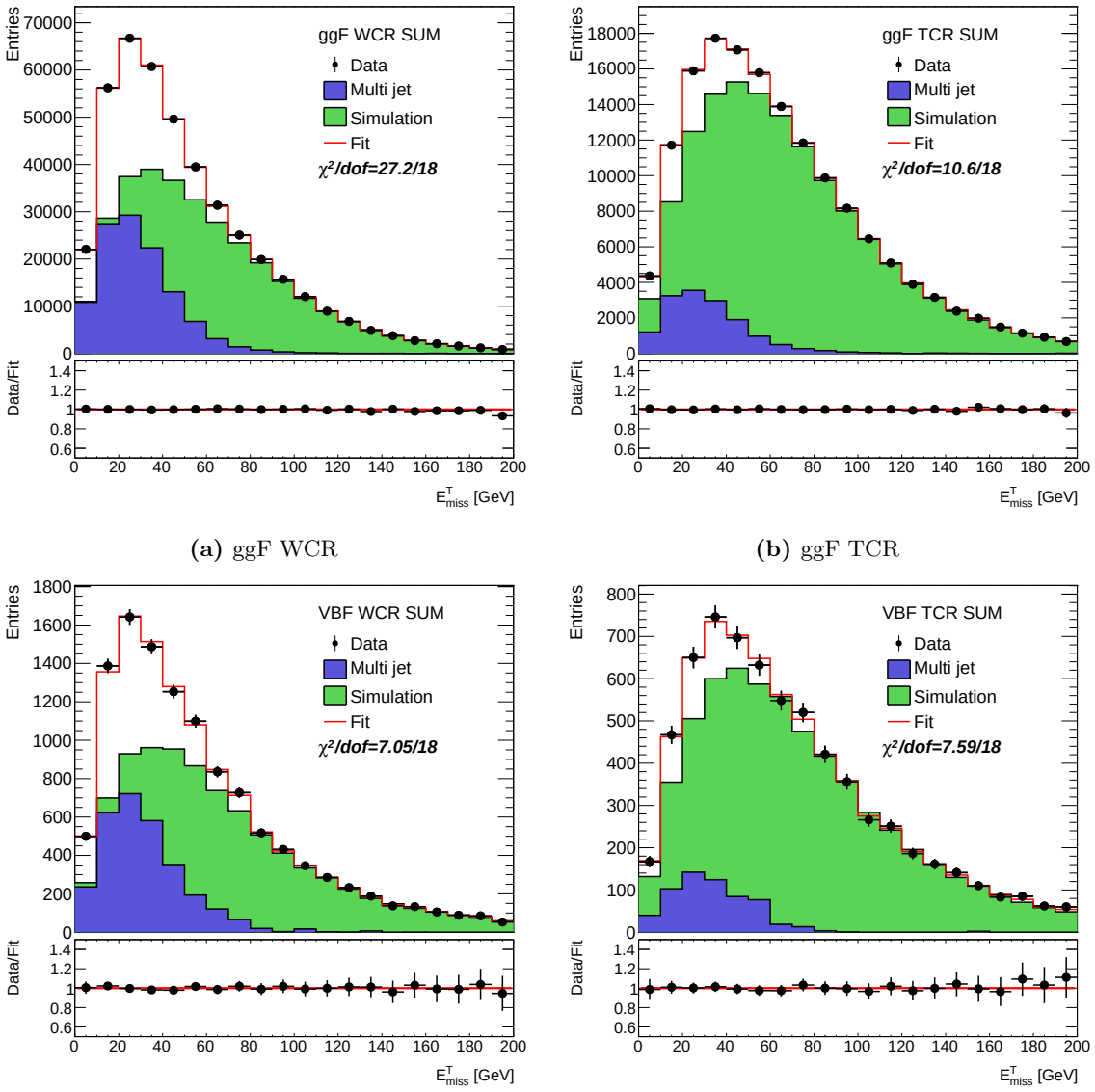
Depending on the mass of the Higgs being searched for, the kinematics of the final state can vary by a large amount. For example, the larger the Higgs mass, the larger the p_{T} of the final state objects. The WW intermediate decay products will also have higher boost with increasing m_H, and so, governed by the relation in Equation 5.3, the W decay products will have a smaller angular separation. In order to maintain high sensitivity throughout the mass range, and take advantage of the extra information that can be used to reject background, selection criteria depending on m_H are employed.

5.9.1 Kinematic Variables

The variables considered for the mass dependant selection differ for the ggF and VBF event categories. The VBF category has a more stringent selection applied at the preselection stage, and as a result, there is a problem of low statistics if all the mass dependant cuts applied in the ggF case are also applied in the case of VBF topologies.

For the ggF category, the cuts applied are on the following kinematic variables:

- $\Delta\phi_{jj}$: The azimuthal angle between the two jets that are tagged in the preselection as W-jets (resolved case only).



(a) ggF WCR **(b)** ggF TCR

(c) VBF WCR **(d)** VBF TCR

Figure 5.12: Summary of multijet background template fits for both ggF and VBF categories in the WCR and TCR. All plots are summed over lepton flavour and charge categories.

- $\Delta\phi_{\ell\nu}$: The azimuthal angle between the neutrino (who's azimuthal angle is determined by the direction of the $E_{\mathrm{T}}^{\mathrm{miss}}$), and the single lepton.
- $p_{\mathrm{T},J}$: Transverse momentum of large-R W-jet (merged case only).
- $p_{\mathrm{T},j1}$: Transverse momentum of the highest-p_{T} resolved W-jet (resolved case only).
- $p_{\mathrm{T},j2}$: Transverse momentum of the other resolved W-jet (resolved case only).
- $p_{\mathrm{T},\ell}$: The lepton p_{T}.
- $E_{\mathrm{T}}^{\mathrm{miss}}$: The magnitude of the missing traverse momentum.

For the VBF category, the variables used are:

- $\Delta\phi_{jj}$: The azimuthal angle between the two jets that are tagged in the preselection as W-jets, as defined above (revolved case only).
- $\Delta\phi_{\ell\nu}$: The azimuthal angle between the neutrino ($E_{\mathrm{T}}^{\mathrm{miss}}$) and the single lepton, as defined above.
- p_{T}-balance: The balance in p_{T} of the final state particles in an event, defined as:

$$p_{\mathrm{T}}-\text{balance} = \frac{\left|(\vec{\mathbf{p}}_\ell + \vec{\mathbf{p}}_\nu + \vec{\mathbf{p}}_{j1} + \vec{\mathbf{p}}_{j2})_{\mathrm{T}}\right|}{p_{\mathrm{T},\ell} + p_{\mathrm{T},\nu} + p_{\mathrm{T},j1} + p_{\mathrm{T},j2}}, \qquad \text{(Resolved case)} \tag{5.17}$$

$$p_{\mathrm{T}}-\text{balance} = \frac{|(\vec{\mathbf{p}}_\ell + \vec{\mathbf{p}}_\nu + \vec{\mathbf{p}}_J)_{\mathrm{T}}|}{p_{\mathrm{T},\ell} + p_{\mathrm{T},\nu} + p_{\mathrm{T},J}}. \qquad \text{(Merged case)} \tag{5.18}$$

The kinematic distributions of the ggF discriminating variables for the signal (for different mass hypotheses) compared to the total combined background after preselection are shown in Figure 5.13. The equivalent plots for the VBF discriminating variables are shown in Figure 5.14. Apparent in both sets of figures is the large difference between the kinematics of the lighter Higgs compared to the heavier Higgs. As the Higgs mass hypothesis increases, the shape differences between the signal and background also grow.

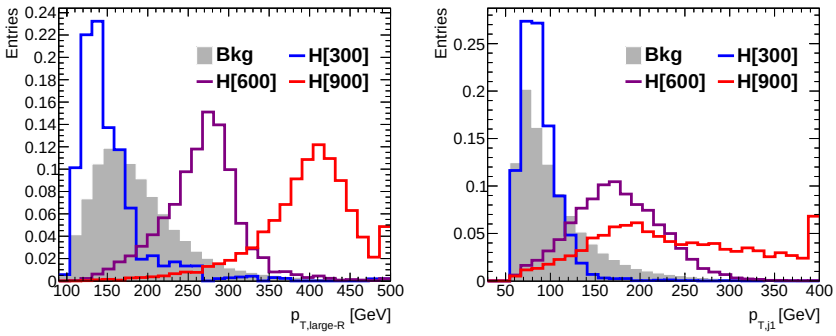
5.9.2 Optimization Procedure for Mass Dependant Selection

The determination of the optimal mass dependant cuts was performed using signal and background events that pass the preselection criteria. Not all signal lineshape/models were considered, instead, the optimization was done only once using the SM-like samples as the signal inputs.

Optimal cuts were determined in the Higgs mass range of $300 \leq m_H \leq 1000$ GeV in 100 GeV increments. For Higgs mass hypothesis that fall between these mass points, the cut value is determined by a linear interpolation between those obtained for the adjacent mass points.

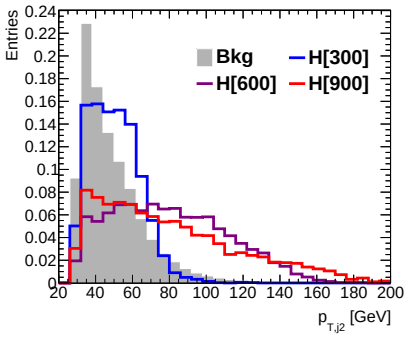
The optimal cuts in the analysis are determined to be those that maximize the estimator of the significance $s/\sqrt{s+b}$, where s is the number of signal events and b is the number of background events. Other significance estimators were also investigated, to see if a change in the definition of the significance had any impact on the optimal cuts [1], but were shown to not alter the optimal cut value. The events that go into the s and b values are not all the signal and background events that pass the pre-selection,

[1] Alternate significance estimators investigated were $s/\sqrt{b}$ and $s/\sqrt{b+\Delta b}$, where $\Delta b = +10\%, +30\%$.

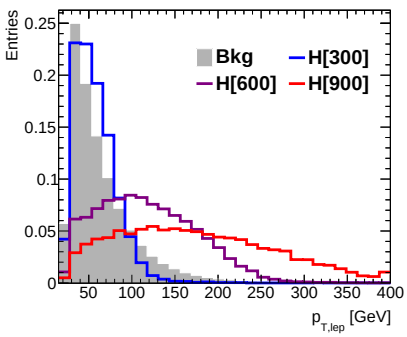


(a) $p_{\mathrm{T},J}$

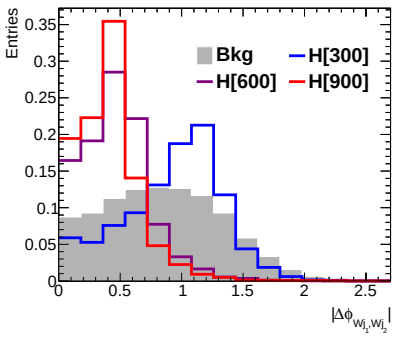
(b) $p_{\mathrm{T},j1}$



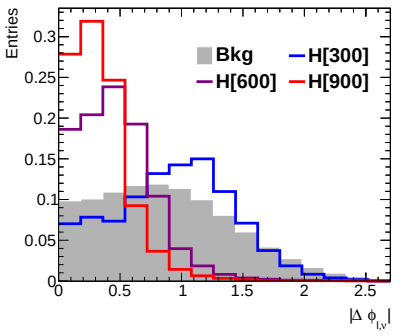
(c) $p_{\mathrm{T},j2}$



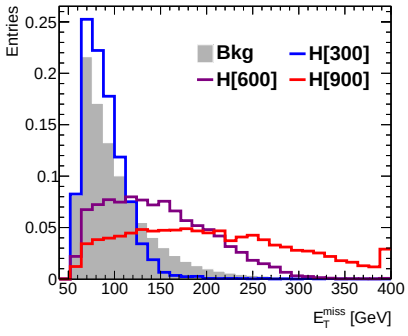
(d) $p_{\mathrm{T},\ell}$



(e) $|\Delta\phi_{jj}|$

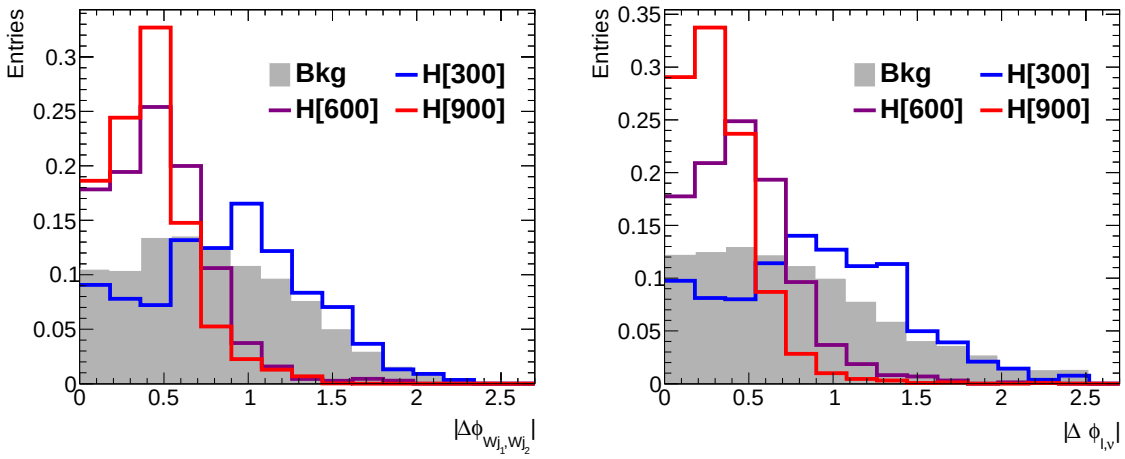


(f) $|\Delta\phi_{\ell\nu}|$

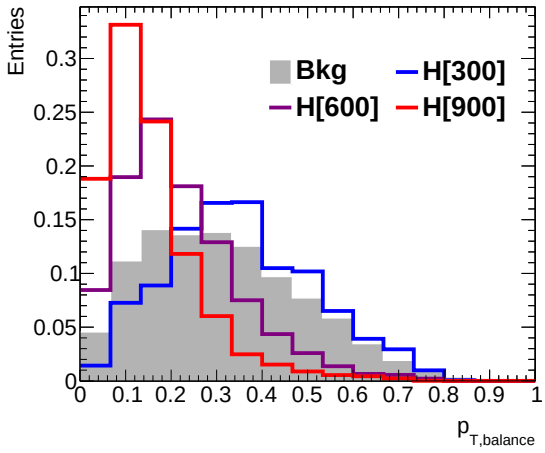


(g) $E_{\mathrm{T}}^{\mathrm{miss}}$

Figure 5.13: Discriminating variables for the ggF mass dependant cuts. Three different SM-like mass hypothesis are shown, along with the total combined background. All distributions are normalized to unit area.



(a) $|\Delta\phi_{jj}|$ **(b)** $|\Delta\phi_{\ell\nu}|$



(c) p_{T}-balance

Figure 5.14: Discriminating variables for the VBF mass dependant cuts. Three different SM-like mass hypothesis are shown, along with the total combined background. All distributions are normalized to unit area.

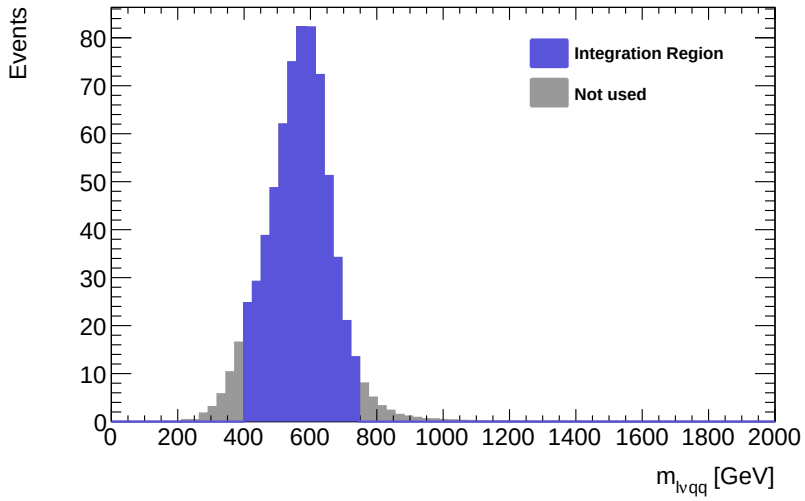


Figure 5.15: Mass dependant optimization integration window for $m_H = 600$ GeV (ggF, SM-like).

but rather, only those that fall into a small range in $m_{\ell\nu jj}$; the optimization range is m_H dependant and is centered about the mean of the reconstructed Higgs mass distribution, and is defined to include 90% of the signal events. An example showing the optimization window for the 600 GeV ggF mass point is seen in Figure 5.15.

Once the optimization window is defined, a scan is done in order to find the cut value with the peak significance for each variable considered. In each scan the cut value is varied between two extremes. For each cut value in the scan the significance is calculated, with the optimal cut value being determined by the cut that yields that largest significance. This procedure is performed for each kinematic variable, and does not account for correlations with the other variables being optimized. An example cut-scan for the variable $|\Delta\phi_{\ell\nu}|$ for the ggF $m_H = 700$ GeV mass point is shown in Figure 5.16; the entire set of cut scan plots for all mass points and variables can be found in Appendix B.

5.9.3 Final Cuts

The optimal cuts determined by the mass dependant selection are quite aggressive, removing a large number of events for both signal and background processes, resulting in limited statistics. While the cuts do maximize the significance, the optimization method only takes into account the final count of the of the signal and background events, and does not take account the sensitivity gained from the shape of the $m_{\ell\nu jj}$ spectrum. As a result, if all mass dependant selection cuts are applied the number of events would be quite limited, diminishing the sensitivity of the final $m_{\ell\nu jj}$ invariant mass shape fit. This is especially true in the case of the ggF category, which has a higher number of kinematic requirements. Therefore, it was decided to systematically loosen the ggF $E_{\mathrm{T}}^{\mathrm{miss}}$ and object p_{T} cuts in order to regain some of the lost events, and regain shape sensitivity for the final fit. These variables were chosen to be loosened because of their comparatively large correlation with the final discriminant $m_{\ell\nu jj}$, and their impact on the final number of events.

To define the loosened cuts, the cut for each mass hypothesis m_H, was changed to the optimal cut value determined for the $(m_H - 100)$ GeV mass point. For $m_H = 300$ GeV, where this is not possible, the cuts default to their preselection values. Figure 5.17 shows the final mass dependant selection cuts

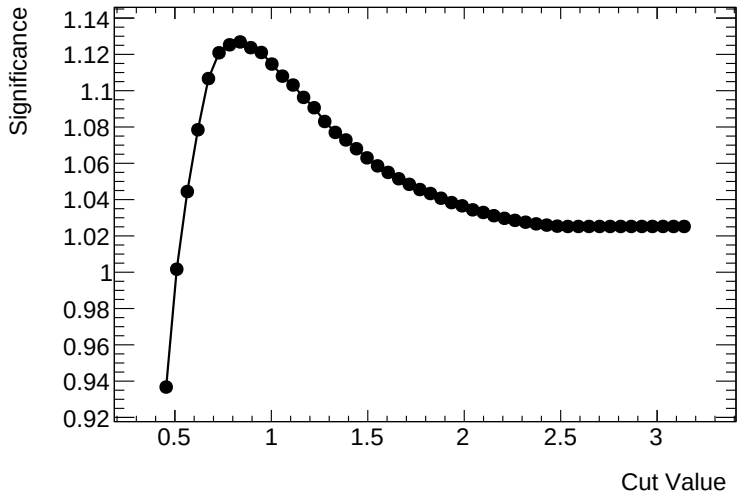
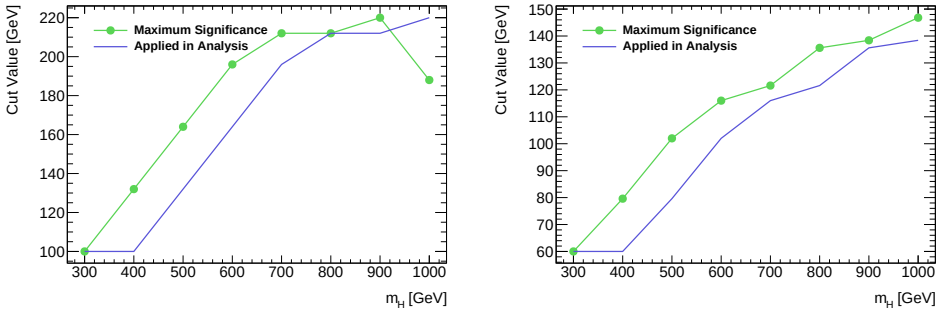


Figure 5.16: Cut scan for $|\Delta\phi_{\ell\nu}|$ (ggF, $m_H = 700$ GeV).

applied for the ggF category. Also shown in the figures are the cut values determined by the optimization procedure to illustrate the loosening of the kinematic requirements.

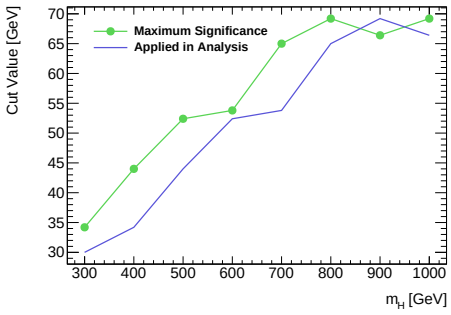
In the VBF category, the issue of reduced event yields in the ggF mass dependant selection was not encountered, and therefore no loosening of cuts was required. However, the gains in significance for the optimized p_{T}-balance cut (see Figure 5.18c) below $m_H = 500$ GeVwere observed to be minimal, with the optimal cut value being determined by small statistical fluctuations. As a result, it was decided to remove this cut for masses below 500 GeV. The final mass dependant cuts used for the VBF category are shown in Figure 5.18, along with those defined by the maximum significance.

In both the ggF and VBF categories, for mass hypotheses with $m_H > 1000$ GeV, the $m_H = 1000$ GeV optimal cuts are used. This only applies to the NWA model, as it is the only signal model in which Higgs mass hypotheses above 1000 GeV are considered.

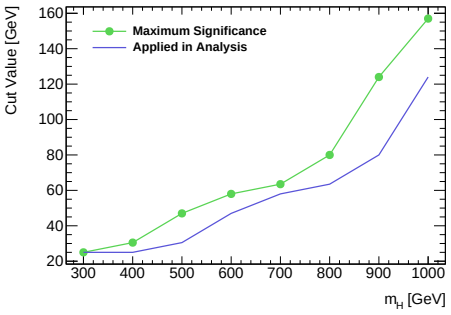


(a) $p_{\mathrm{T},J}$

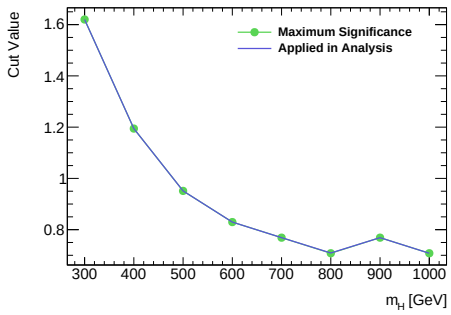
(b) $p_{\mathrm{T},j1}$



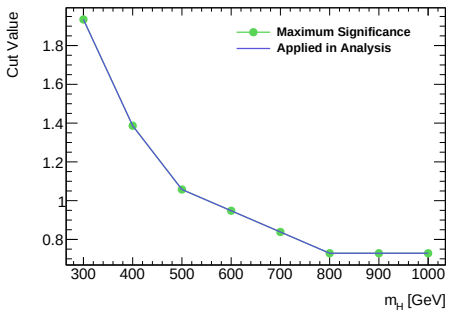
(c) $p_{\mathrm{T},j2}$



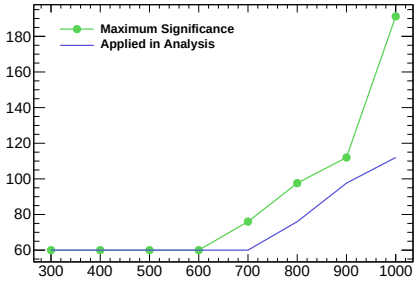
(d) $p_{\mathrm{T},\ell}$



(e) $|\Delta\phi_{jj}|$

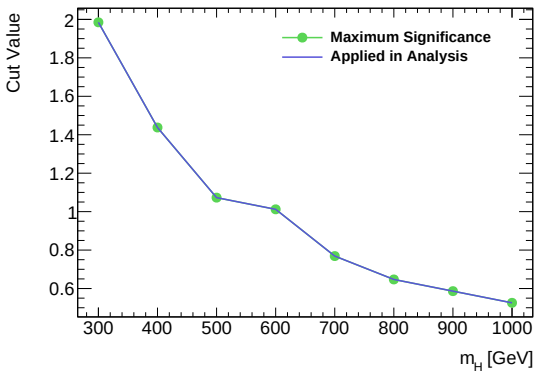


(f) $|\Delta\phi_{\ell\nu}|$

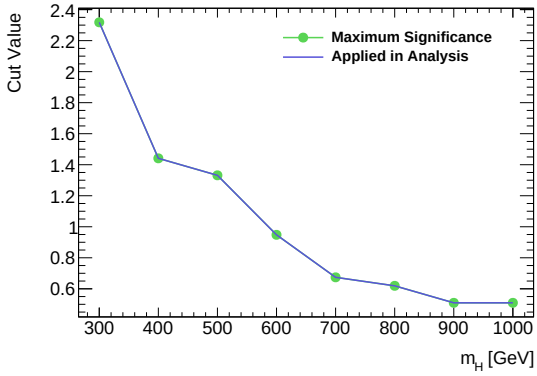


(g) $E_{\mathrm{T}}^{\mathrm{miss}}$

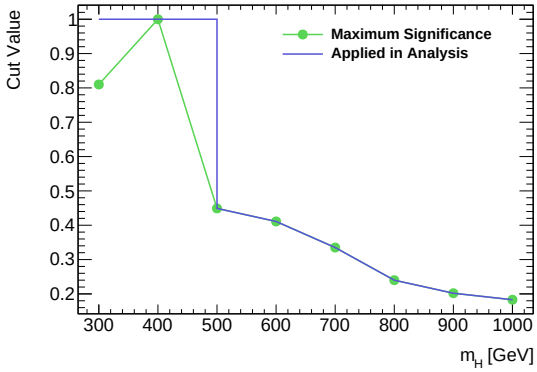
Figure 5.17: Optimal (defined by peak significance) cut values and cut values used in the analysis, as a function of m_H for each of the ggF discriminating variables.



(a) $|\Delta\phi_{jj}|$



(b) $|\Delta\phi_{\ell\nu}|$



(c) p_{T}-balance

Figure 5.18: Optimal (defined by peak significance) cut values and cut values used in the analysis, as a function of m_H for each of the VBF discriminating variables.

5.10 Top and W+jets Modelling

While examining the data modelling in the Top and W control regions, it was observed that for events with large reconstructed $m_{\ell\nu jj}$, the data/MC ratio gets progressively worse with increasing mass. This effect was seen across lepton charge and flavour channels, as well as in both ggF and VBF production categories. For the W+jets modelling this effect was determined to be a result of mismodelling in the Sherpa samples, as no data/MC slope was seen when using an alternate (reduced statistics) Alpgen W+jets sample. Unfortunately, no samples generated with Alpgen had sufficient statistics in the high mass region (especially after the mass dependant selection) to be used.

In the case of the Top samples, this effect had been seen previously by other groups in ATLAS using the same MC samples [124]. Other Higgs groups (looking at Higgs produced in association with top quarks) developed a method to reweight the $t\bar{t}$ transverse momentum, which resolved the data/MC disagreement [125]. The same reweighting procedure is applied in this analysis; however, this was found to be insufficient to fully correct for the effect in this analysis. Therefore, a further correction, described below, was applied.

In order to correct the modelling of the top samples, the preselection and 300 GeV mass dependant selection is applied to the data and MC in the top control region. The $m_{\ell\nu jj}$ spectrum, summed over lepton flavour and charge channels is then made for data and MC with the components of the background coming from the non-Top sources removed from the data by subtracting the corresponding MC background estimates. The remaining data (which should represent only the Top component) and the top MC are plotted together. As is seen in Figure 5.19, a significant slope is present in the ratio. An analytic function is fit to this ratio, which is used to reweight the Top background as a function of $m_{\ell\nu jj}$. The fitting function is chosen to be the lowest order polynomial, still able to fit to the distribution, which in this case is a 1^{st} order polynomial fitted between $200 \leq m_{\ell\nu jj} \leq 1500$ GeV.

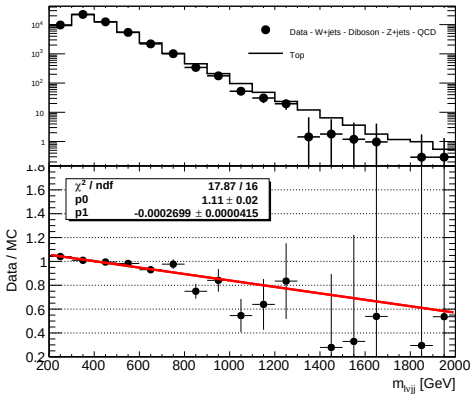
A similar procedure is applied in order to reweight the W+jets samples in the WCR. Once again, a significant slope is seen in the WCR, for both the ggF and VBF categories (see Figure 5.20). Like the case of the top reweighting, a minimal function is found that fits the distribution. For the VBF category, the upper fit range is quite low compared to the ggF category. This is a result of the large statistical error at high $m_{\ell\nu jj}$, resulting in there being a number of analytic functions that are able to fit the data.

Although the slope in the Data/MC ratio is also seen in the p_{T} distributions of the hadronic and leptonic W decay products, it was found that reweighting the distributions with the reconstructed mass was the most effective method to correct the modelling across all other kinematics variables.

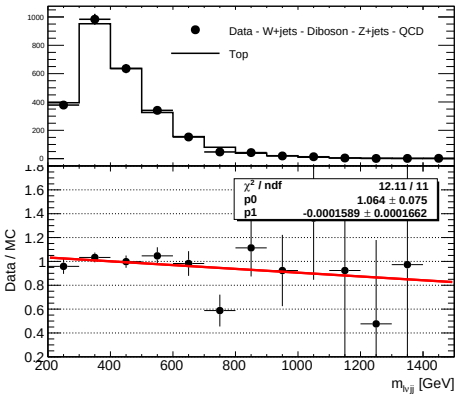
In both cases of the reweighting, events that have an invariant mass above the maximum of the fit range of the correction are corrected by a constant factor, equal to that of an event that sits at the upper extreme of the fit range.

A summary of the fits used to correct the W and Top backgrounds is found in Table 5.8. The systematic uncertainty associated with the reweighting procedure is discussed in Section 5.12.4.

The results of the reweighting (before and after) can be seen in Figures 5.21 (ggF) and 5.22 (VBF) for the TCR, and in Figures 5.23 (ggF) and 5.24 (VBF) for the WCR.

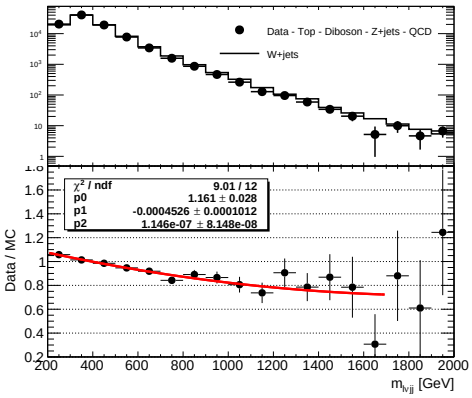


(a) ggF

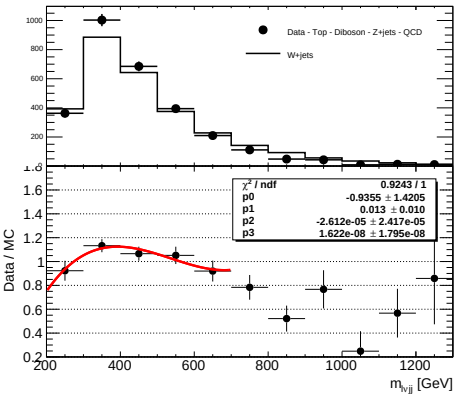


(b) VBF

Figure 5.19: Comparison between data and MC for the reconstructed Higgs mass in the TCR ($m_H =$ 300 GeV selection). All the flavor channels and charge channels are summed together, with the non-top backgrounds subtracted from the data. The top panel of the plots show the data and MC overlaid, while the bottom panel shows the ratio between the data and MC, with the red line indicating the fitted reweighting function.

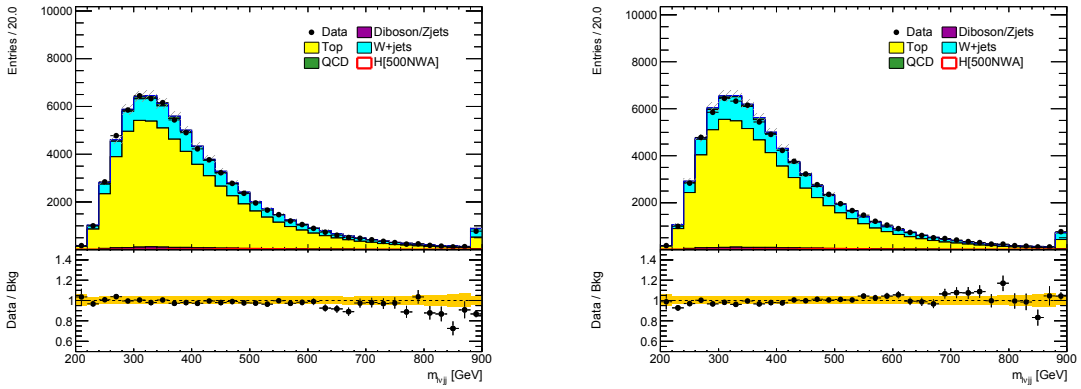


(a) ggF



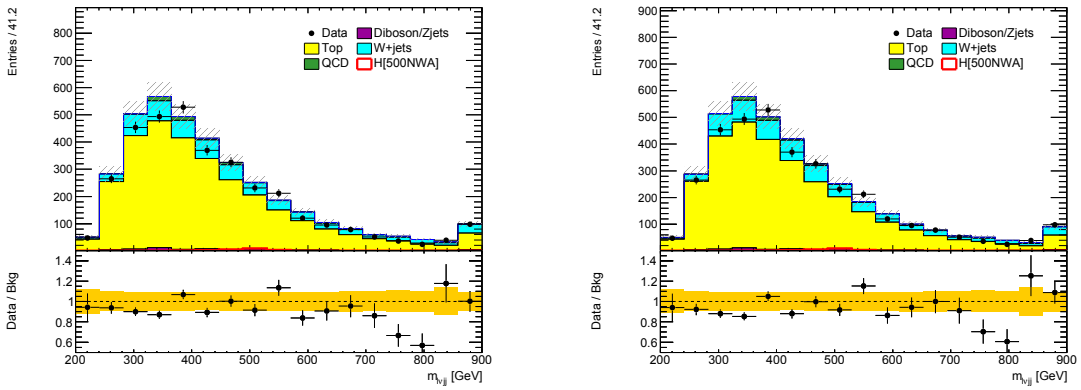
(b) VBF

Figure 5.20: Comparison between data and MC for the reconstructed Higgs mass in the WCR ($m_H =$ 300 GeV selection). All the flavor channels and charge channels are summed together, with the non-W backgrounds subtracted from the data. The top panel of the plots show the data and MC overlaid, while the bottom panel shows the ratio between the data and MC, with the red line indicating the fitted reweighting function.



(a) **(b)**

Figure 5.21: Comparison of data and Monte Carlo simulation in the ggF TCR before (a) and after (b) reweighting.



(a) **(b)**

Figure 5.22: Comparison of data and Monte Carlo simulation in the VBF TCR before (a) and after (b) reweighting.

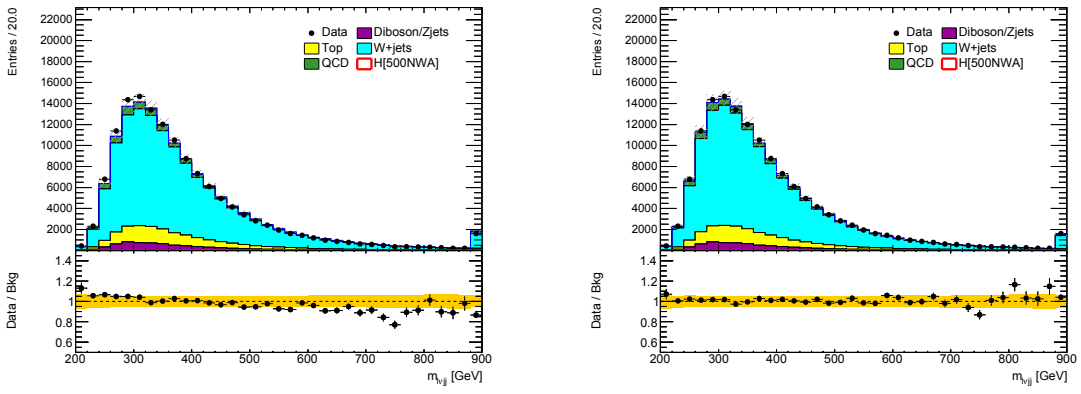


Figure 5.23: Data/MC comparison in ggF WCR before (a) and after (b) reweighting.

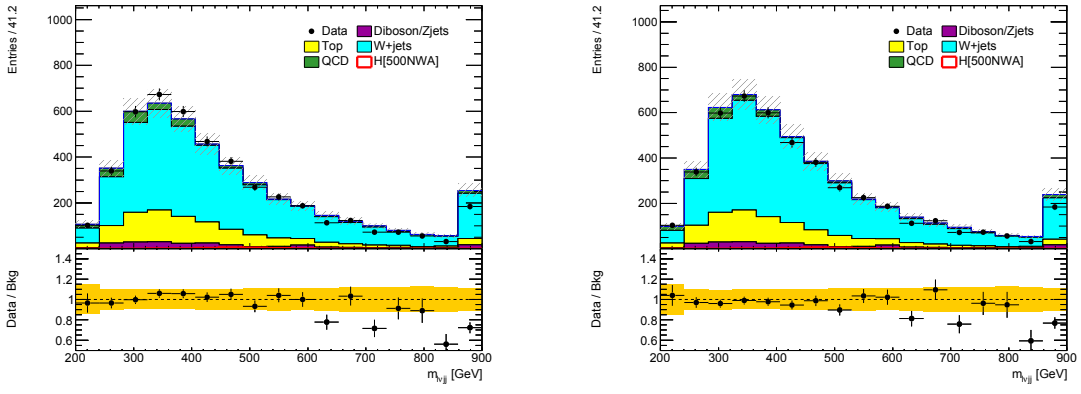


Figure 5.24: Data/MC comparison in VBF WCR before (a) and after (b) reweighting.

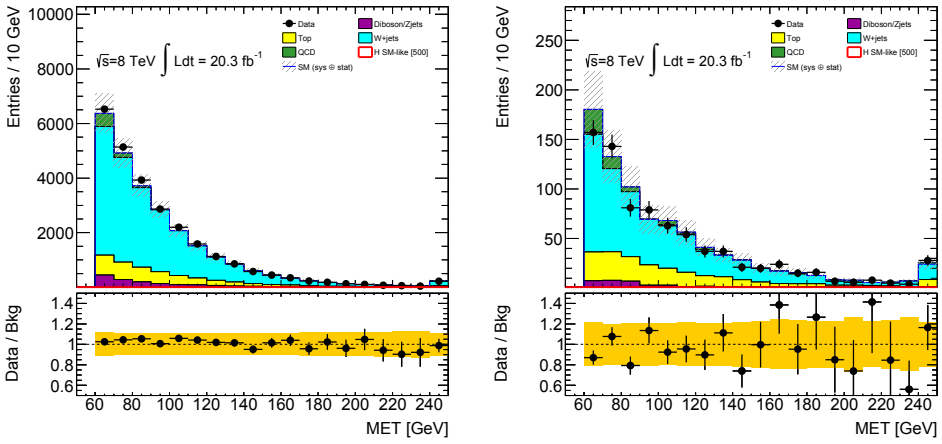
Table 5.8: Fitting functions and ranges for Top and W+jets reweighting.

Process	Production Mode	Fitting Function	Fit Range [GeV]
W+jets	ggF	2^{nd} order polynomial	$200 - 1700$
	VBF	3^{rd} order polynomial	$200 - 700$
Top	ggF VBF	1^{st} order polynomial	$200 - 1500$

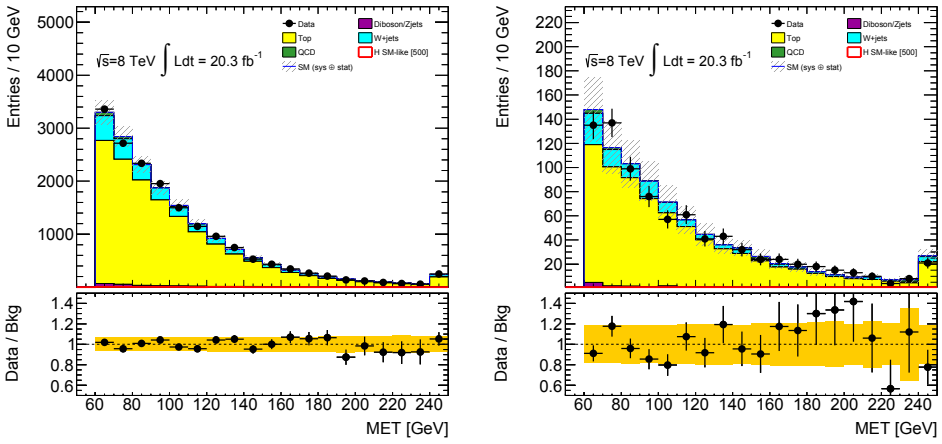
5.11 Event Yields and Modelling in Signal and Control Regions

The MC modelling of the data after preselection, and after W+jet and Top reweighting, is quite good. In Figures 5.25 – 5.30, the data and MC is shown for a selection of kinematic variables ($E_{\mathrm{T}}^{\mathrm{miss}}$, p_{T}^{ℓ}, $\Delta\phi_{\mathrm{W-jets}}$, p_{T}^{H}, $\Delta\eta_{\mathrm{tag\,jets}}$, $m_{\ell\nu jj}$) in the WCR, TCR and SR. All figures are shown for the e^{-} channel, separately for ggF and VBF categories (Figure 5.29, showing $|\Delta\eta_{jj}|$ of the VBF tag jets, only includes the VBF category as the ggF category does not contain VBF tag jets). Modelling for the other charge and lepton flavours are of similar quality. Figure 5.30, which displays the data/MC agreement for $m_{\ell\nu jj}$, is the final fitted discriminant.

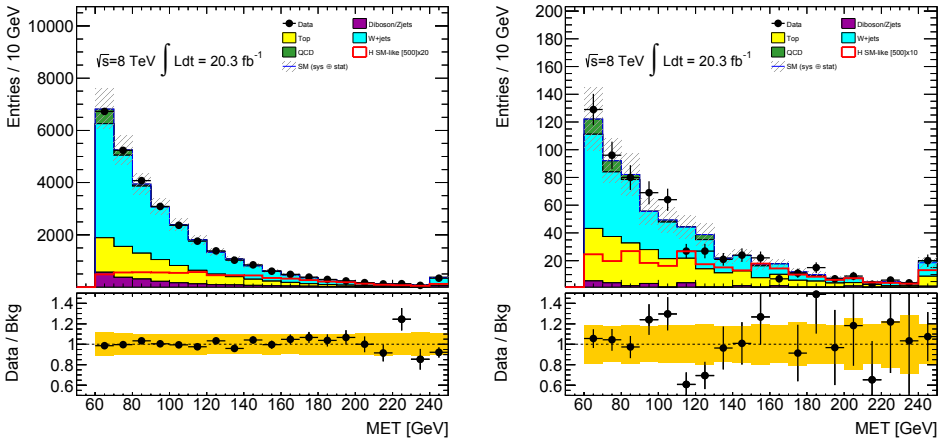
The event yields (summed over lepton charge and flavour) as a function of the mass hypothesis for the SM-like ggF (VBF) samples are listed in Table 5.9 (Table 5.10). The yields for the NWA ggF (VBF) are in Table 5.11 (Table 5.12). As in the case of the above figures, the event yields are determined after the W+jets and Top backgrounds have been reweighted.



(a) ggF WCR **(b)** VBF WCR

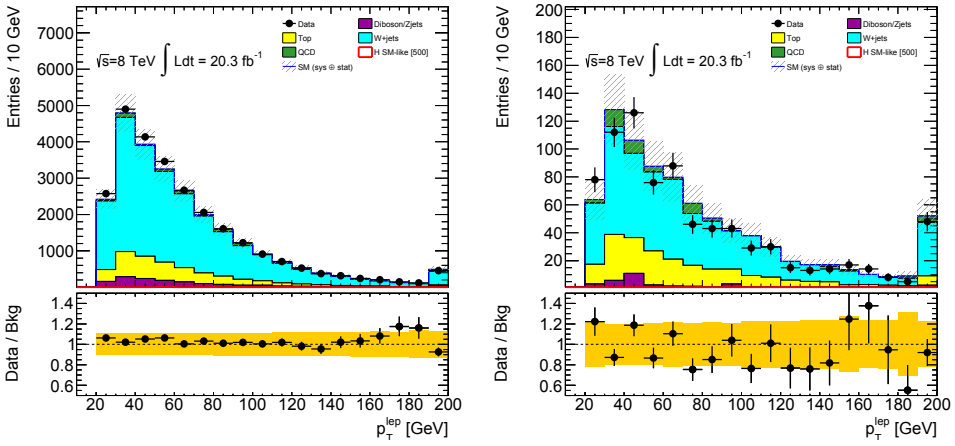


(c) ggF TCR **(d)** VBF TCR

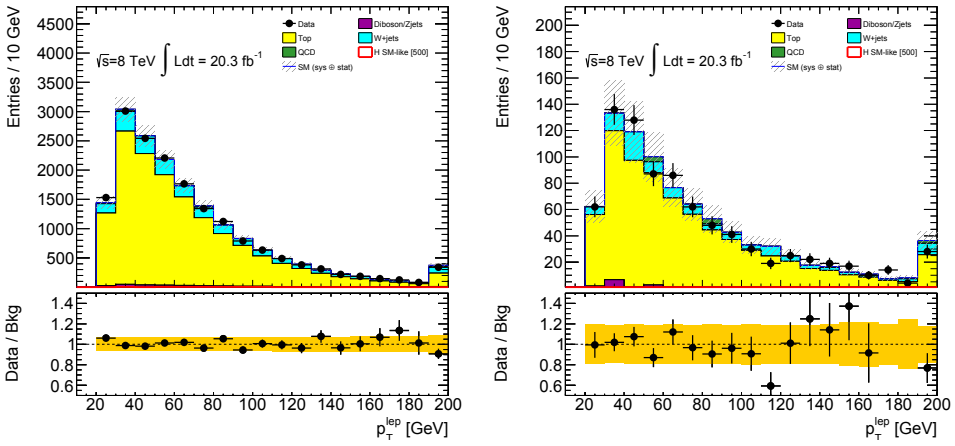


(e) ggF SR **(f)** VBF SR

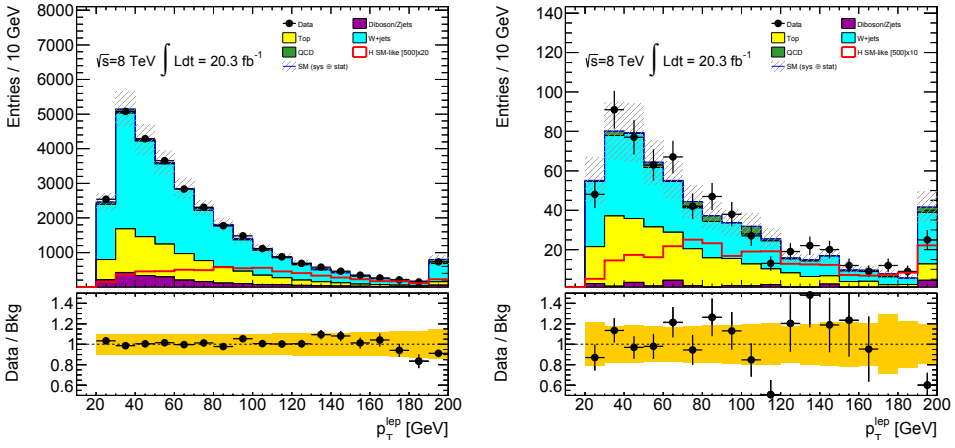
Figure 5.25: $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions for the e^- channel in the WCR, TCR and SR (ggF and VBF categories separated). In each plot, the top panel shows the data and MC (normalization determined by theoretical cross-section) overlaid, with the bottom panels showing the ratio of data over MC. The yellow band in the ratio distributions show the combined statistical and systematic uncertainties.



(a) ggF WCR **(b)** VBF WCR

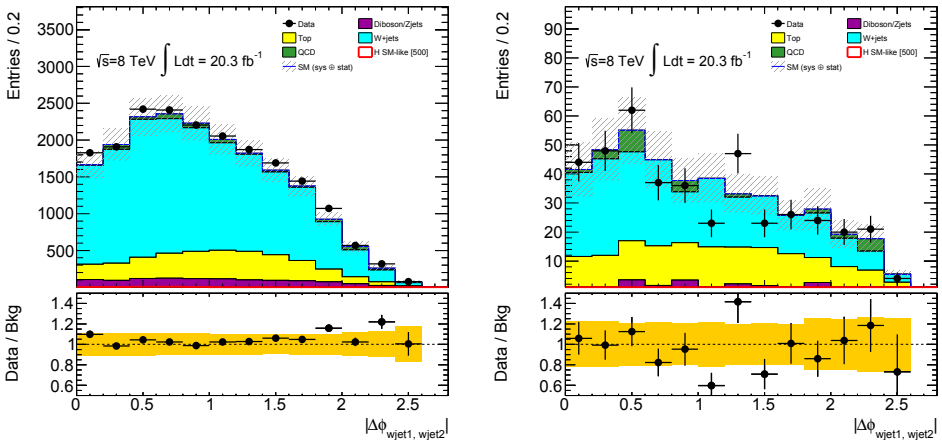


(c) ggF TCR **(d)** VBF TCR

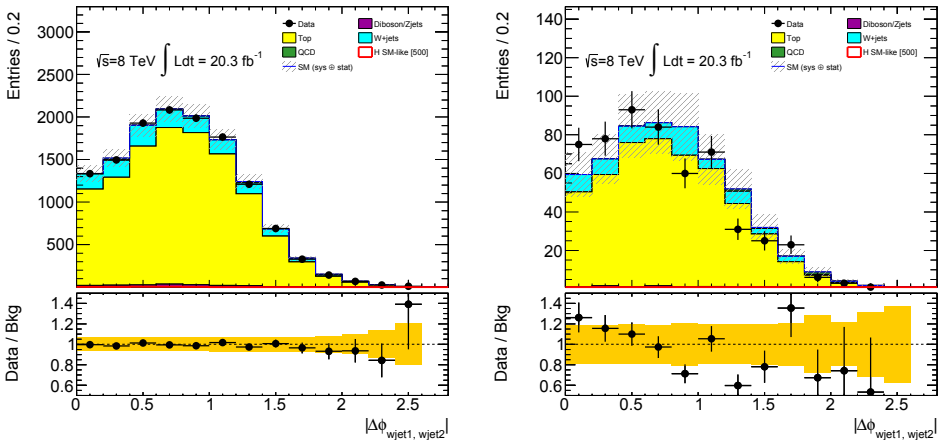


(e) ggF SR **(f)** VBF SR

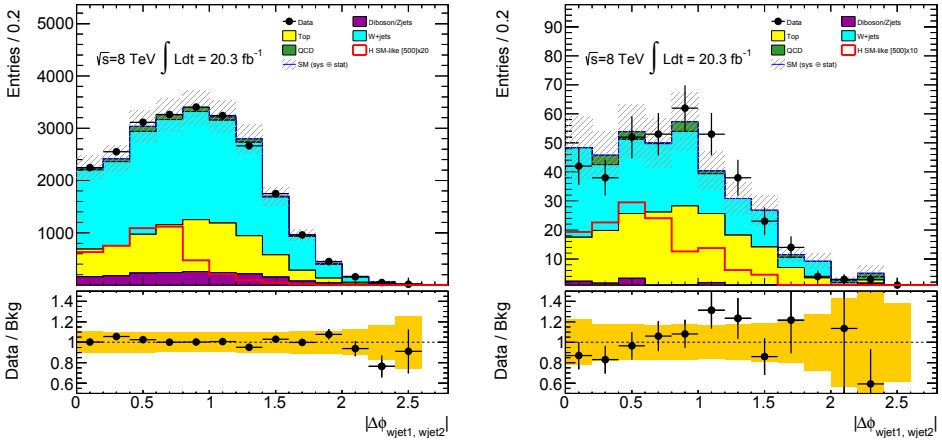
Figure 5.26: Electron p_{T} for the e^- channel in the WCR, TCR and SR (ggF and VBF categories separated). In each plot, the top panel shows the data and MC (normalization determined by theoretical cross-section) overlaid, with the bottom panels showing the ratio of data over MC. The yellow band in the ratio distributions show the combined statistical and systematic uncertainties.



(a) ggF WCR **(b)** VBF WCR

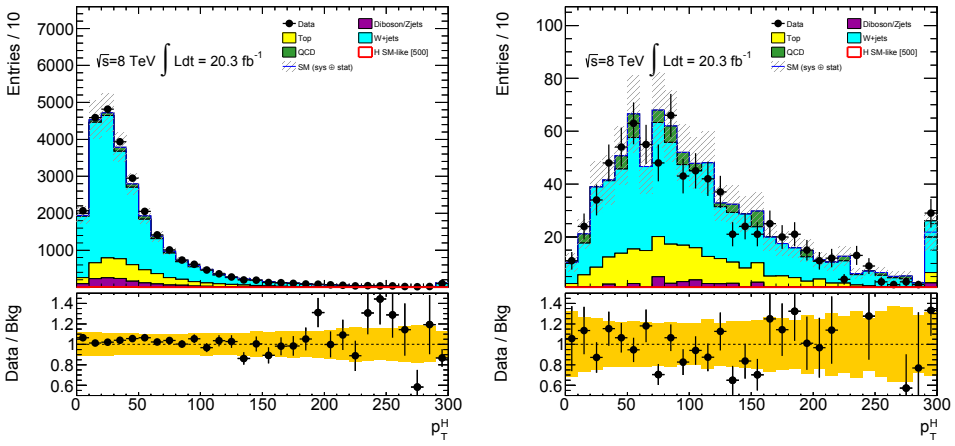


(c) ggF TCR **(d)** VBF TCR

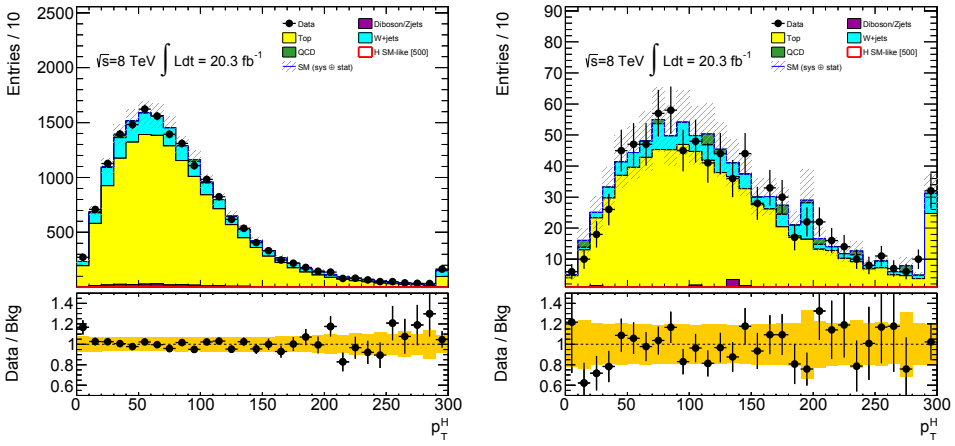


(e) ggF SR **(f)** VBF SR

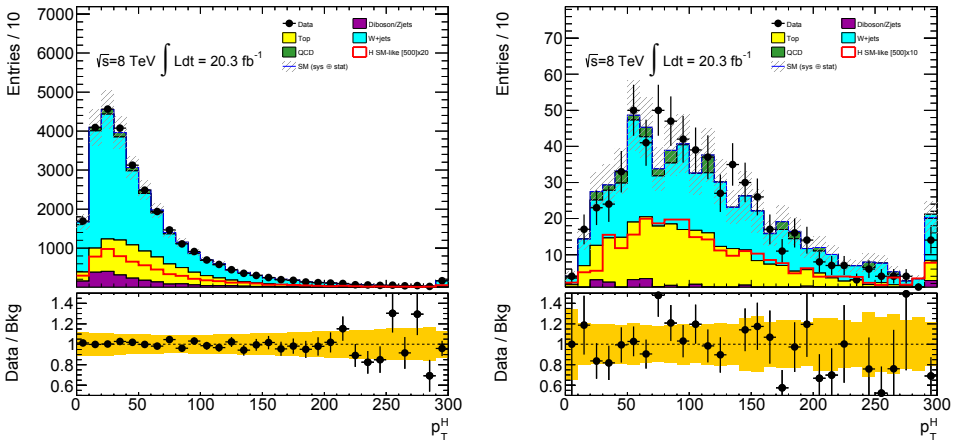
Figure 5.27: $\Delta\phi_{jj}$ of jets from hadronic W for the e^- channel in the WCR, TCR and SR (ggF and VBF categories separated). In each plot, the top panel shows the data and MC (normalization determined by theoretical cross-section) overlaid, with the bottom panels showing the ratio of data over MC. The yellow band in the ratio distributions show the combined statistical and systematic uncertainties.



(a) ggF WCR **(b)** VBF WCR

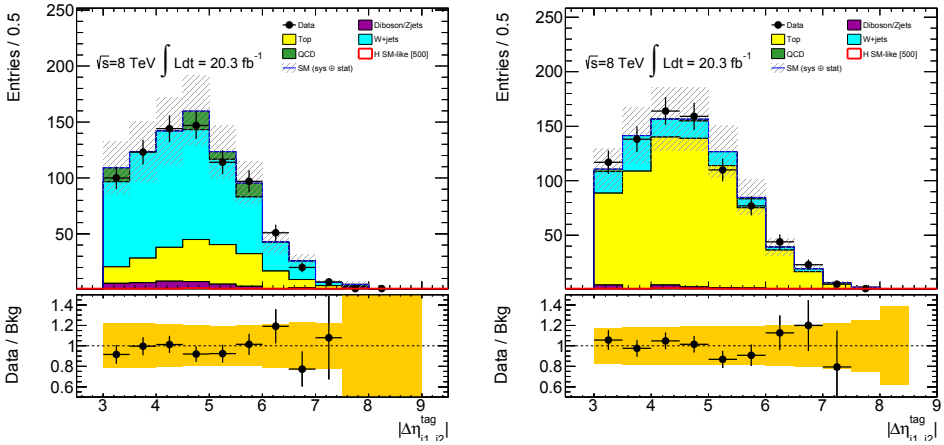


(c) ggF TCR **(d)** VBF TCR

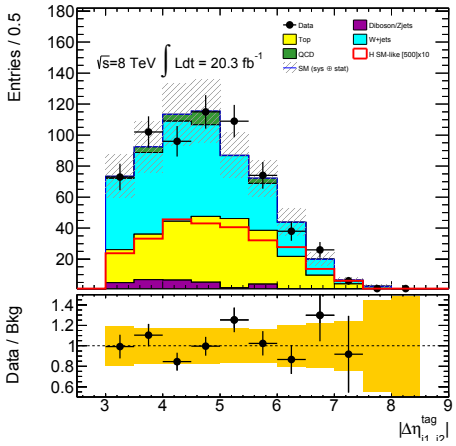


(e) ggF SR **(f)** VBF SR

Figure 5.28: Reconstructed Higgs p_{T} for the e^{-} channel in the WCR, TCR and SR (ggF and VBF categories separated). In each plot, the top panel shows the data and MC (normalization determined by theoretical cross-section) overlaid, with the bottom panels showing the ratio of data over MC. The yellow band in the ratio distributions show the combined statistical and systematic uncertainties.

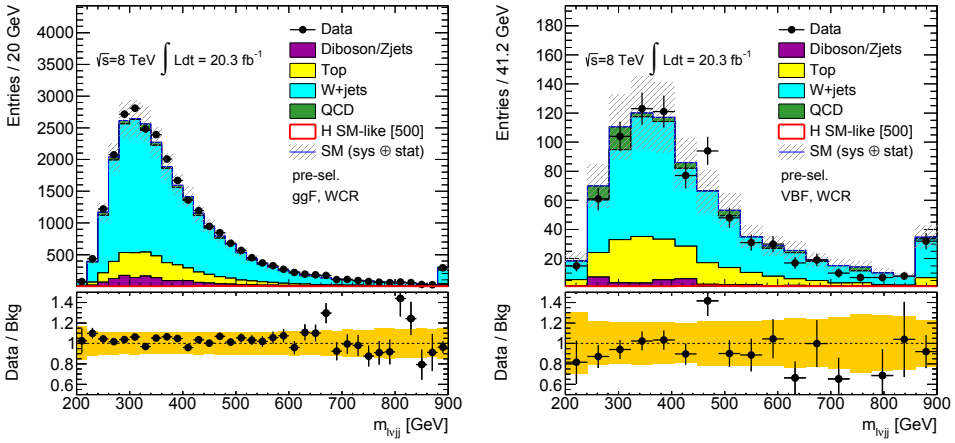


(a) VBF WCR **(b)** VBF TCR



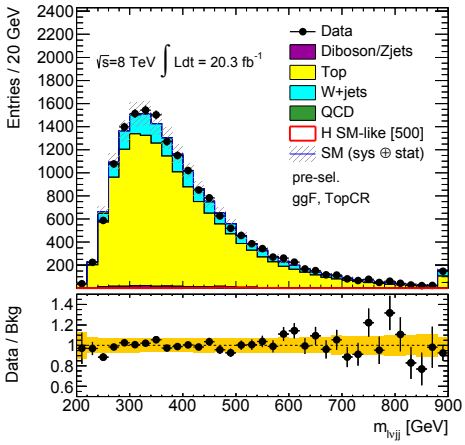
(c) VBF SR

Figure 5.29: Pseudorapidity gap, $|\Delta\eta_{jj}|$, of VBF tag di-jets for the e^- channel in the WCR, TCR and SR (VBF category only). In each plot, the top panel shows the data and MC (normalization determined by theoretical cross-section) overlaid, with the bottom panels showing the ratio of data over MC. The yellow band in the ratio distributions show the combined statistical and systematic uncertainties.

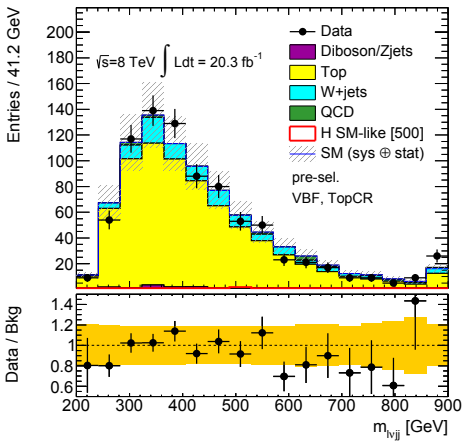


(a) ggF WCR

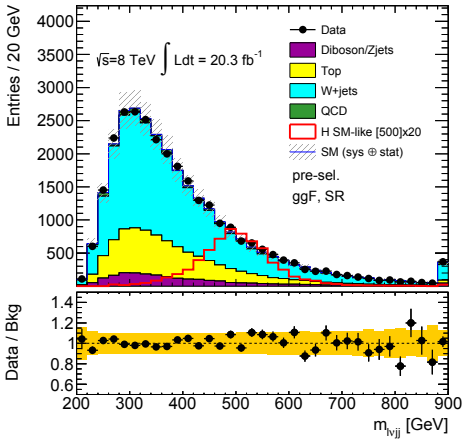
(b) VBF WCR



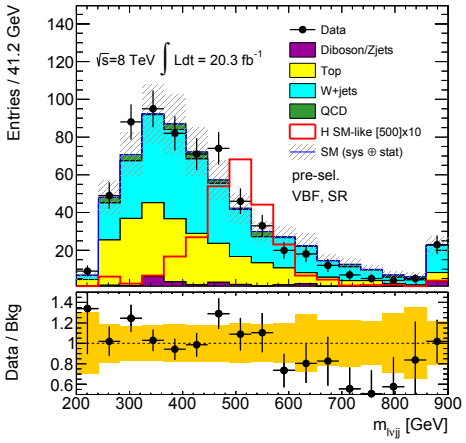
(c) ggF TCR



(d) VBF TCR



(e) ggF SR



(f) VBF SR

Figure 5.30: Reconstructed Higgs mass, $m_{\ell\nu jj}$, for the e^- channel in the WCR, TCR and SR (ggF and VBF categories separated). In each plot, the top panel shows the data and MC (normalization determined by theoretical cross-section) overlaid, with the bottom panels showing the ratio of data over MC. The yellow band in the ratio distributions show the combined statistical and systematic uncertainties.

Table 5.9: Event yields for ggF (SM-like) category. The W+jets, Top and diboson/Z+jets samples are from Monte Carlo simulation with statistical error. The data driven multijet normalization is determined by a template fit, with systematic uncertainties given. Expected signal yields assume SM Higgs production cross-sections and branching ratios.

Region	Signal [ggF]	Signal [VBF]	W+jets	Top	Diboson/Z+jets	Multijet	Data
			$m_H = 300$ GeV				
SR	1344 ± 17	103 ± 2	97010 ± 227	27043 ± 33	9617 ± 63	4457 ± 1752	136310 ± 369
WCR	294 ± 8	26.1 ± 0.8	95447 ± 244	12216 ± 22	6140 ± 54	4769 ± 1871	118414 ± 344
TCR	154 ± 6	16.0 ± 0.7	11613 ± 69	54657 ± 47	1240 ± 26	895 ± 344	67854 ± 260
			$m_H = 400$ GeV				
SR	1815 ± 17	93 ± 2	70902 ± 180	19217 ± 28	7103 ± 54	3189 ± 1245	99176 ± 315
WCR	263 ± 6	13.2 ± 0.5	69255 ± 197	7565 ± 17	4532 ± 48	3808 ± 1470	84809 ± 291
TCR	251 ± 6	15.2 ± 0.6	9061 ± 56	40617 ± 41	1043 ± 25	700 ± 272	51181 ± 226
			$m_H = 500$ GeV				
SR	987 ± 8	65.9 ± 0.9	41943 ± 122	11140 ± 21	4359 ± 43	1622 ± 624	58265 ± 241
WCR	109 ± 3	7.9 ± 0.3	37506 ± 127	3910 ± 12	2522 ± 35	1868 ± 700	45845 ± 214
TCR	134 ± 3	10.4 ± 0.4	5693 ± 38	24184 ± 31	690 ± 20	402 ± 156	30829 ± 176
			$m_H = 600$ GeV				
SR	453 ± 4	45.1 ± 0.7	23059 ± 79	6143 ± 15	2419 ± 31	780 ± 287	32022 ± 179
WCR	43 ± 1	4.6 ± 0.2	18715 ± 76	1990 ± 9	1248 ± 24	716 ± 264	22907 ± 151
TCR	59 ± 1	6.6 ± 0.3	3194 ± 25	13247 ± 23	384 ± 14	206 ± 76	17049 ± 131
			$m_H = 700$ GeV				
SR	219 ± 2	31.2 ± 0.5	14315 ± 56	3774 ± 12	1559 ± 24	522 ± 190	19971 ± 141
WCR	21.4 ± 0.7	3.2 ± 0.2	10893 ± 51	1201 ± 7	737 ± 17	434 ± 167	13388 ± 116
TCR	30.2 ± 0.9	4.2 ± 0.2	2021 ± 19	8147 ± 17	231 ± 12	105 ± 41	10597 ± 103
			$m_H = 800$ GeV				
SR	110.4 ± 1	22.7 ± 0.4	8340 ± 39	2126 ± 9	912 ± 18	144 ± 54	11418 ± 107
WCR	11.6 ± 0.4	2.2 ± 0.1	6019 ± 35	656 ± 5	406 ± 13	162 ± 66	7218 ± 85
TCR	14.6 ± 0.5	2.7 ± 0.1	1206 ± 13	4553 ± 13	131 ± 9	53 ± 25	6014 ± 78
			$m_H = 900$ GeV				
SR	55.7 ± 0.6	16.3 ± 0.3	4942 ± 28	1220 ± 7	557 ± 14	68 ± 28	6635 ± 81
WCR	6.1 ± 0.2	1.8 ± 0.1	3262 ± 23	368 ± 4	223 ± 9	44 ± 17	3943 ± 63
TCR	7.9 ± 0.3	1.8 ± 0.1	736 ± 9	2566 ± 10	76 ± 7	25 ± 10	3420 ± 58
			$m_H = 1000$ GeV				
SR	23.7 ± 0.3	10.2 ± 0.2	1742 ± 14	376 ± 4	207 ± 8	25 ± 11	2243 ± 47
WCR	3.3 ± 0.1	1.24 ± 0.06	1108 ± 11	117 ± 2	83 ± 5	15 ± 7	1317 ± 36
TCR	3.5 ± 0.1	1.10 ± 0.08	287 ± 5	830 ± 5	26 ± 4	16 ± 7	1165 ± 34

Table 5.10: Event yields for VBF (SM-like) category. The W+jets, Top and diboson/Z+jets samples are from Monte Carlo simulation with statistical error. The data driven multijet normalization is determined by a template fit, with systematic uncertainties given. Expected signal yields assume SM Higgs production cross-sections and branching ratios.

Region	Signal [ggF]	Signal [VBF]	W+jets	Top	Diboson/Z+jets	Multijet	Data
			$m_H = 300$ GeV				
SR	25 ± 2	41 ± 1	2011 ± 31	969 ± 6	155 ± 8	131 ± 48	3020 ± 55
WCR	10 ± 2	11.5 ± 0.6	3220 ± 39	771 ± 5	224 ± 10	226 ± 81	4215 ± 65
TCR	8 ± 1	6.9 ± 0.4	618 ± 17	2881 ± 10	61 ± 6	61 ± 23	3401 ± 58
			$m_H = 400$ GeV				
SR	53 ± 3	52 ± 1	1645 ± 26	765 ± 5	132 ± 7	88 ± 35	2480 ± 50
WCR	16 ± 2	10.7 ± 0.5	2546 ± 33	541 ± 4	177 ± 9	195 ± 70	3258 ± 57
TCR	17 ± 2	9.7 ± 0.5	491 ± 13	2297 ± 9	49 ± 5	51 ± 20	2742 ± 52
			$m_H = 500$ GeV				
SR	39 ± 2	43.9 ± 0.7	1103 ± 21	500 ± 4	88 ± 6	59 ± 25	1693 ± 41
WCR	9.2 ± 0.8	6.7 ± 0.3	1739 ± 27	346 ± 4	120 ± 7	127 ± 48	2241 ± 47
TCR	11.3 ± 0.9	6.8 ± 0.3	306 ± 10	1464 ± 7	31 ± 4	8 ± 4	1794 ± 42
			$m_H = 600$ GeV				
SR	24.8 ± 0.9	36.6 ± 0.6	775 ± 15	346 ± 4	67 ± 5	44 ± 20	1196 ± 35
WCR	5.3 ± 0.4	5.0 ± 0.2	1175 ± 20	235 ± 3	84 ± 6	72 ± 27	1548 ± 39
TCR	5.3 ± 0.4	5.1 ± 0.2	210 ± 7	981 ± 6	22 ± 3	8 ± 4	1200 ± 35
			$m_H = 700$ GeV				
SR	12.8 ± 0.5	25.8 ± 0.4	421 ± 11	179 ± 3	41 ± 4	24 ± 10	672 ± 26
WCR	2.7 ± 0.4	3.6 ± 0.2	610 ± 13	121 ± 2	50 ± 5	36 ± 13	806 ± 28
TCR	2.3 ± 0.2	3.2 ± 0.2	113 ± 5	466 ± 4	13 ± 3	7 ± 4	550 ± 23
			$m_H = 800$ GeV				
SR	7.7 ± 0.3	19.6 ± 0.3	260 ± 8	101 ± 2	27 ± 3	9 ± 4	394 ± 20
WCR	1.3 ± 0.1	2.7 ± 0.1	371 ± 10	73 ± 2	34 ± 4	11 ± 4	477 ± 22
TCR	1.3 ± 0.2	1.9 ± 0.1	58 ± 3	234 ± 3	5 ± 1	5 ± 3	292 ± 17
			$m_H = 900$ GeV				
SR	4.1 ± 0.2	14.4 ± 0.3	141 ± 5	55 ± 1	16 ± 2	7 ± 3	223 ± 15
WCR	0.81 ± 0.07	2.33 ± 0.09	208 ± 6	42 ± 1	25 ± 4	12 ± 5	279 ± 17
TCR	0.62 ± 0.06	1.34 ± 0.09	30 ± 2	119 ± 2	2.5 ± 0.9	0.5 ± 0.3	145 ± 12
			$m_H = 1000$ GeV				
SR	2.8 ± 0.1	12.6 ± 0.2	100 ± 4	38 ± 1	12 ± 2	4 ± 2	159 ± 13
WCR	0.60 ± 0.05	2.44 ± 0.09	154 ± 4	33 ± 1	20 ± 3	7 ± 3	215 ± 15
TCR	0.40 ± 0.04	0.82 ± 0.05	21 ± 2	76 ± 2	2.1 ± 0.8	0.3 ± 0.2	98 ± 10

Table 5.11: Event yields for ggF (NWA) category. The W+jets, Top and diboson/Z+jets samples are from Monte Carlo simulation with statistical error. The data driven multijet normalization is determined by a template fit, with systematic uncertainties given. Expected signal yields assume SM Higgs production cross-sections and branching ratios.

Region	Signal [ggF]	Signal [VBF]	W+jets	Top	Diboson/Z+jets	Multijet	Data
			$m_H = 300$ GeV				
SR	1309 ± 21	100 ± 4	97010 ± 227	27043 ± 33	9617 ± 63	4457 ± 1752	136310 ± 369
WCR	282 ± 10	26 ± 2	95447 ± 244	12216 ± 22	6140 ± 54	4769 ± 1871	118414 ± 344
TCR	145 ± 7	19 ± 2	11613 ± 69	54657 ± 47	1240 ± 26	895 ± 344	67854 ± 260
			$m_H = 400$ GeV				
SR	1879 ± 21	100 ± 3	70902 ± 180	19217 ± 28	7103 ± 54	3189 ± 1245	99176 ± 315
WCR	279 ± 8	16 ± 1	69255 ± 197	7565 ± 17	4532 ± 48	3808 ± 1470	84809 ± 291
TCR	258 ± 8	17 ± 1	9061 ± 56	40617 ± 41	1043 ± 25	700 ± 272	51181 ± 226
			$m_H = 500$ GeV				
SR	997 ± 10	71 ± 2	41943 ± 122	11140 ± 21	4359 ± 43	1622 ± 624	58265 ± 241
WCR	105 ± 3	8 ± 1	37506 ± 127	3910 ± 12	2522 ± 35	1868 ± 700	45845 ± 214
TCR	133 ± 4	11 ± 1	5693 ± 38	24184 ± 31	690 ± 20	402 ± 156	30829 ± 176
			$m_H = 600$ GeV				
SR	453 ± 4	50 ± 1	23059 ± 79	6143 ± 15	2419 ± 31	780 ± 287	32022 ± 179
WCR	42 ± 1	3.9 ± 0.3	18715 ± 76	1990 ± 9	1248 ± 24	716 ± 264	22907 ± 151
TCR	61 ± 2	7.1 ± 0.4	3194 ± 25	13247 ± 23	384 ± 14	206 ± 76	17049 ± 131
			$m_H = 700$ GeV				
SR	212 ± 2	34.5 ± 0.8	14315 ± 56	3774 ± 12	1559 ± 24	522 ± 190	19971 ± 141
WCR	20 ± 1	2.8 ± 0.2	10893 ± 51	1201 ± 7	737 ± 17	434 ± 167	13388 ± 116
TCR	27 ± 1	4.3 ± 0.3	2021 ± 19	8147 ± 17	231 ± 12	105 ± 41	10597 ± 103
			$m_H = 800$ GeV				
SR	99.2 ± 0.9	23.6 ± 0.5	8340 ± 39	2126 ± 9	912 ± 18	144 ± 54	11418 ± 107
WCR	10.9 ± 0.3	2.9 ± 0.2	6019 ± 35	656 ± 5	406 ± 13	162 ± 66	7218 ± 85
TCR	13.2 ± 0.4	2.8 ± 0.3	1206 ± 13	4553 ± 13	131 ± 9	53 ± 25	6014 ± 78
			$m_H = 900$ GeV				
SR	47.4 ± 0.5	15.7 ± 0.4	4942 ± 28	1220 ± 7	557 ± 14	68 ± 28	6635 ± 81
WCR	6.0 ± 0.2	2.1 ± 0.1	3262 ± 23	368 ± 4	223 ± 9	44 ± 17	3943 ± 63
TCR	6.5 ± 0.2	1.5 ± 0.1	736 ± 9	2566 ± 10	76 ± 7	25 ± 10	3420 ± 58
			$m_H = 1000$ GeV				
SR	22.3 ± 0.2	9.9 ± 0.3	1742 ± 14	376 ± 4	207 ± 8	25 ± 11	2243 ± 47
WCR	3.20 ± 0.09	1.4 ± 0.1	1108 ± 11	117 ± 2	83 ± 5	15 ± 7	1317 ± 36
TCR	2.84 ± 0.09	1.01 ± 0.08	287 ± 5	830 ± 5	26 ± 4	16 ± 7	1165 ± 34
			$m_H = 1100$ GeV				
SR	23.2 ± 0.21	10.0 ± 0.1	1742 ± 14	376 ± 4	207 ± 8	25 ± 11	2243 ± 47
WCR	3.53 ± 0.08	1.54 ± 0.05	1108 ± 11	117 ± 2	83 ± 5	15 ± 7	1317 ± 36
TCR	3.23 ± 0.08	0.82 ± 0.03	287 ± 5	830 ± 5	26 ± 4	16 ± 7	1165 ± 34
			$m_H = 1200$ GeV				
SR	23.8 ± 0.2	10.3 ± 0.1	1742 ± 14	376 ± 4	207 ± 8	25 ± 11	2243 ± 47
WCR	4.12 ± 0.09	1.72 ± 0.05	1108 ± 11	117 ± 2	83 ± 5	15 ± 7	1317 ± 36
TCR	3.31 ± 0.08	0.91 ± 0.04	287 ± 5	830 ± 5	26 ± 4	16 ± 7	1165 ± 34

Table 5.12: Event yields for VBF (NWA) category. The W+jets, Top and diboson/Z+jets samples are from Monte Carlo simulation with statistical error. The data driven multijet normalization is determined by a template fit, with systematic uncertainties given. Expected signal yields assume SM Higgs production cross-sections and branching ratios.

Region	Signal [ggF]	Signal [VBF]	W+jets	Top	Diboson/Z+jets	Multijet	Data
			$m_H = 300$ GeV				
SR	25 ± 3	41 ± 2	2011 ± 31	969 ± 6	155 ± 8	131 ± 48	3020 ± 55
WCR	6 ± 1	13 ± 1	3220 ± 39	771 ± 5	224 ± 10	226 ± 81	4215 ± 65
TCR	8 ± 2	8 ± 1	618 ± 17	2881 ± 10	61 ± 6	61 ± 23	3401 ± 58
			$m_H = 400$ GeV				
SR	53 ± 3	59 ± 2	1645 ± 26	765 ± 5	132 ± 7	88 ± 35	2480 ± 50
WCR	15 ± 2	11.6 ± 0.9	2546 ± 33	541 ± 4	177 ± 9	195 ± 70	3258 ± 57
TCR	18 ± 2	10.2 ± 0.8	491 ± 13	2297 ± 9	49 ± 5	51 ± 20	2742 ± 52
			$m_H = 500$ GeV				
SR	40 ± 2	48.2 ± 1	1103 ± 21	500 ± 4	88 ± 6	59 ± 25	1693 ± 41
WCR	8.0 ± 0.9	7.8 ± 0.5	1739 ± 27	346 ± 4	120 ± 7	127 ± 48	2241 ± 47
TCR	10 ± 1	7.8 ± 0.6	306 ± 10	1464 ± 7	31 ± 4	8 ± 4	1794 ± 42
			$m_H = 600$ GeV				
SR	24 ± 1	40 ± 1	775 ± 15	346 ± 4	67 ± 5	44 ± 20	1196 ± 35
WCR	5.5 ± 0.5	5.1 ± 0.4	1175 ± 20	235 ± 3	84 ± 6	72 ± 27	1548 ± 39
TCR	5.5 ± 0.5	5.4 ± 0.4	210 ± 7	981 ± 6	22 ± 3	8 ± 4	1200 ± 35
			$m_H = 700$ GeV				
SR	12.7 ± 0.5	26.6 ± 0.7	421 ± 11	179 ± 3	41 ± 4	24 ± 10	672 ± 26
WCR	2.4 ± 0.2	4.3 ± 0.3	610 ± 13	121 ± 2	50 ± 5	36 ± 13	806 ± 28
TCR	2.5 ± 0.2	3.4 ± 0.2	113 ± 5	466 ± 4	13 ± 3	7 ± 4	550 ± 23
			$m_H = 800$ GeV				
SR	7.5 ± 0.3	20.5 ± 0.5	260 ± 8	101 ± 2	27 ± 3	9 ± 4	394 ± 20
WCR	1.5 ± 0.1	2.7 ± 0.2	371 ± 10	73 ± 2	34 ± 4	11 ± 4	477 ± 22
TCR	1.2 ± 0.1	2.0 ± 0.2	58 ± 3	234 ± 3	5 ± 1	5 ± 3	292 ± 17
			$m_H = 900$ GeV				
SR	4.4 ± 0.1	15.6 ± 0.4	141 ± 5	55 ± 1	16 ± 2	7 ± 3	223 ± 15
WCR	0.7 ± 0.1	2.7 ± 0.2	208 ± 6	42 ± 1	25 ± 4	12 ± 5	279 ± 17
TCR	0.7 ± 0.1	1.3 ± 0.1	30 ± 2	119 ± 2	2.5 ± 0.9	0.5 ± 0.3	145 ± 12
			$m_H = 1000$ GeV				
SR	2.92 ± 0.09	13.6 ± 0.3	100 ± 4	38 ± 1	12 ± 2	4 ± 2	159 ± 13
WCR	0.61 ± 0.04	2.1 ± 0.1	154 ± 4	33 ± 1	20 ± 3	7 ± 3	215 ± 15
TCR	0.34 ± 0.03	1.0 ± 0.1	21 ± 2	76 ± 2	2.1 ± 0.8	0.3 ± 0.2	98 ± 10
			$m_H = 1100$ GeV				
SR	3.33 ± 0.08	14.2 ± 0.1	100 ± 4	38 ± 1	12 ± 2	4 ± 2	159 ± 13
WCR	0.73 ± 0.04	2.72 ± 0.06	154 ± 4	33 ± 1	20 ± 3	7 ± 3	215 ± 15
TCR	0.42 ± 0.03	0.92 ± 0.04	21 ± 2	76 ± 2	2.1 ± 0.8	0.3 ± 0.2	98 ± 10
			$m_H = 1200$ GeV				
SR	3.60 ± 0.08	15.5 ± 0.2	100 ± 4	38 ± 1	12 ± 2	4 ± 2	159 ± 13
WCR	0.94 ± 0.04	3.30 ± 0.07	154 ± 4	33 ± 1	20 ± 3	7 ± 3	215 ± 15
TCR	0.52 ± 0.03	1.03 ± 0.04	21 ± 2	76 ± 2	2.1 ± 0.8	0.3 ± 0.2	98 ± 10

5.12 Systematic Uncertainties

To account for the experimental systematic uncertainties[1], the initial event selection done in the **C**ommon **A**nalysis **F**ramework (CAF) can be run with one of 25 possible variations, which can adjust the energy scale or resolution of the reconstructed objects (such as the electron energy scale, and the electron energy resolution). The full list of systematic variations is summarized in Table 5.14, with the variations performed with the CAF indicated.

There are some uncertainties that are not evaluated using a dedicated CAF variation, such as those associated with the trigger and lepton scale factors, b-tagging scale factors, as well as those associated with the treatment of the interference, and the integrated luminosity. These are discussed further in this section.

5.12.1 Lepton Uncertainties

Both the muon and electron reconstructed energy scales and resolutions have associated uncertainties which must be accounted for. In the case of the muons, which use information from the independent muon spectrometer and inner detector subsystems, the resolution is affected by the resolution of both the ID and the MS, which are independent. Therefore, the muon resolution uncertainty has two uncorrelated components; one, which scales the resolution of the muon spectrometer, and the other that scales the resolution from the inner detector. This is in contrast to the electrons, which only have a single source of resolution uncertainty (coming from the calorimeters).

There are also systematics associated with the lepton and trigger scale factors introduced in Chapter 4. Each of the scale factors also has its own associated uncertainty variations.

5.12.2 Resolved Jet Uncertainties

The resolved jets have uncertainties associated with the jet energy scale (JES) and jet energy resolution (JER), that are evaluated by the ATLAS *JetEtMiss* group. These uncertainties are associated with the jet calibration procedure. The jet energy scale uncertainties are parameterized as a function of the jet's p_{T} and η and are further broken into thirteen components that enter into the limit setting as separate nuisance parameters (see Section 5.13). The thirteen parameters making up the total JES uncertainty are included in Table 5.14 along with a short description. These JER and JES uncertainties depend on the jet algorithm used, as well as the radius parameter. The merged C/A MDF jets have their own JES/JER uncertainties (as well as mass scale and mass resolution uncertainties), which are discussed separately in Section 5.12.3.

b-tagging Uncertainties

The ATLAS b-tagging algorithm has a related b-tag identification systematic that indicates the uncertainty in six uncorrelated components, using the eigenvector method. These are used together with systematic uncertainties on the mistag rates associated with the misidentification of light quarks and c-jets. These eight uncertainties change the b-tagging scale factor(s) applied to an event, and do not require a separate CAF dataset to evaluate. The eight separate b-tagging uncertainties act as separate nuisance parameters in the final fit.

[1] In ATLAS vernacular, systematic uncertainties are often referred to as *systematics*.

5.12.3 Large-R Jet Uncertainties

As for resolved jets, the large-R jets used in this analysis have associated scale and resolution systematics. However, since we are also interested in the *mass* of the merged jets, we must not only pay attention to the JES/JER uncertainties, but also to the associated jet mass scale (JMS) and jet mass resolution (JMR) uncertainties.

Large-R Jet Scale Uncertainties

The JES and JMS uncertainties for the merged jets are evaluated by the ATLAS *JetEtMiss* group. Like the uncertainties for resolved jets, the uncertainties are parameterized by the jet's p_{T} and η. However, the uncertainty is not broken up into subcategories. The systematics associated with the JMS and JES are on the order of a few percent. These uncertainties were only evaluated in the kinematic region of $p_{\mathrm{T}} \geq 100$ GeV and $|\eta| < 1.2$, which was the reason behind the object selection cuts applied for the merged jets, outlined in Section 5.5.1.

Large-R Jet Resolution

A study of the resolution of the large-R jet mass and energy resolutions (JER and JMR) in the 2010 dataset revealed that the measured resolution can vary up to 20% [126]. This conclusion was also validated with the 2011 and 2012 *pp* datasets. Therefore, the systematics associated with the large-R JER and JMR were evaluated by finding the mass and energy resolutions of our merged jets and then smearing them by a Gaussian distribution, such that the resolutions were increased by 20%. The JER and JMR were smeared in separate CAF variations, and like the rest of the experimental systematics, they are treated as independent nuisance parameters in the final fit.

To extract the resolutions, the jet energy (mass) for reconstructed MC jets in our signal samples were compared to a matched ($\Delta R \leq 0.8$) truth jet. The square root of the variance, σ, from a Gaussian fit of the central portion of the resulting distribution, is taken as the jet energy (mass) resolution. The resolutions vary as a function of the jet p_{T} and $|\eta|$; the jets were therefore binned into two bins in $|\eta|$ ($0 < |\eta| < 0.6$ and $0.6 < |\eta| < 1.2$) and three bins of p_{T} ($100 < p_{\mathrm{T}} < 200$, $200 < p_{\mathrm{T}} < 500$, and $p_{\mathrm{T}} > 500$ GeV) before fitting[1]. The resolution spectra and associated fits are found in Figure 5.31. The spectra/fits for the mass resolution are found in Figure 5.32. The extracted resolutions (before smearing) are summarized in Table 5.13.

In order to find the width of the Gaussian needed (σ_{smear}) to smear the jet energy/mass by 20%, recall that the convolution of two Gaussians with variances σ_1^2 and σ_2^2, result in a third Gaussian with a variance of $\sigma'^2 = \sigma_1^2 + \sigma_2^2$. Using this rule, and remembering that the final resolution increases the observed resolution by 20%, the variance of the smearing Gaussian can be solved for:

$$
\begin{aligned}
{\sigma'_{\mathrm{res}}}^2 &= \sigma_{\mathrm{res}}^2 + \sigma_{\mathrm{smear}}^2 \qquad [\textbf{N.B. } \sigma'_{\mathrm{res}} = 1.2 \times \sigma_{\mathrm{res}}] \\
1.44 \times \sigma_{\mathrm{res}}^2 &= \sigma_{\mathrm{res}}^2 + \sigma_{\mathrm{smear}}^2 \\
\sigma_{\mathrm{smear}}^2 &= 0.44 \times \sigma_{\mathrm{res}}^2 \\
\sigma_{\mathrm{smear}} &= (0.6633\ldots) \times \sigma_{\mathrm{res}}.
\end{aligned}
$$

[1] The binning was chosen empirically in order to maximize the granularity, while still maintaining significant statistics for the fit.

Therefore, the jet energy and mass are smeared by a unit Gaussian with $\sigma_{\text{smear}} = 0.6633 \times \sigma_{\text{res}}$ in order to retrieve the systematic variations on the large-R JMR and JER. This increase in the resolution is symmetrized to get the final uncertainty on the jet resolution.

Table 5.13: Extracted energy and mass resolutions for large-R groomed jets in bins of p_{T} and $|\eta|$. These resolutions are then used to smear the jet energy/mass in order to increase the resolutions by 20%. The transverse momentum is measured in units of GeV.

Energy Resolution	$100 < p_{\text{T}} < 200$	$200 < p_{\text{T}} < 500$	$p_{\text{T}} > 500$
$0 < \|\eta\| < 0.6$	7.58%	5.05%	3.76%
$0.6 < \|\eta\| < 1.2$	7.91%	5.51%	4.27%
Mass Resolution	$100 < p_{\text{T}} < 200$	$200 < p_{\text{T}} < 500$	$p_{\text{T}} > 500$
$0 < \|\eta\| < 0.6$	9.75%	7.26%	6.36%
$0.6 < \|\eta\| < 1.2$	9.81%	7.56%	6.92%

5.12.4 Top and W+jets Reweighting Uncertainty

An uncertainty is evaluated for the Top and W+jets contributions, originating from the uncertainty on the reweighting procedure. For both background reweightings, a 50% (symmetric) uncertainty on the scaling used is applied. While the magnitude of the uncertainty may appear to be large, the effect that this uncertainty has on the final limits was observed to be minor.

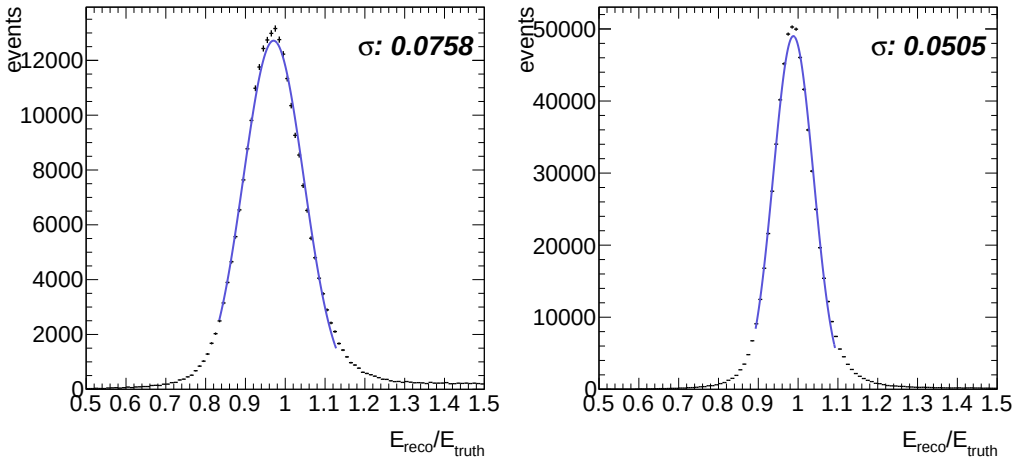
5.12.5 Multijet Background Systematic Uncertainty

The uncertainty on the MJ background normalization is evaluated by performing the same template fit as described in Section 5.8, but with an alternate definition for the loosened leptons used to make the templates. For electrons, the calorimeter and track isolation requirements are reversed, while for muons, the requirement on the impact parameter significance is dropped.

The change in the lepton selection requirements (providing an alternate definition of the MJCRs') results in an alternate MJ estimation, which is compared to the default one. The difference between the two normalization methods is taken as an estimate on the uncertainty.

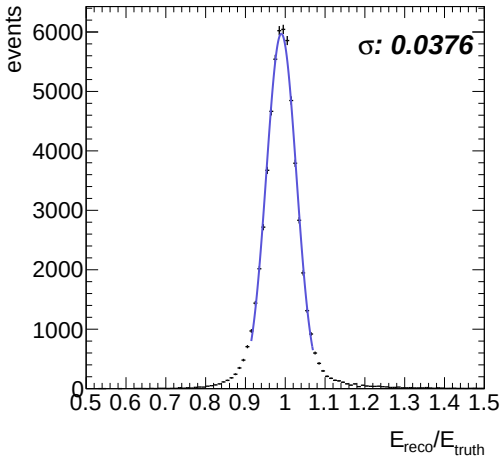
The uncertainty on the MJ normalization is then evaluated through examining the difference between the nominal MJ contribution, and the contribution determined by using the alternate template. A comparison of the uncertainties for all mass points resulted in the following systematics uncertainties being applied for the ggF category: 100% uncertainty on the electron channel, with an exception for the e^- channel when $m_H < 400$ GeV, where the uncertainty is 70%, due to better compatibility between the two MJ estimates. For the muon channel, the uncertainty in the μ^- channel is 30%, and ranges from 10–20% for μ^+, depending on m_H. Despite the large uncertainties, the multijet background forms a very small contribution to the total background, being heavily suppressed by the $E_{\text{T}}^{\text{miss}}$ cut applied in the preselection. For the higher masses, it contributes an even smaller amount due to the stricter selection criteria; therefore, the large uncertainties on this background have a very small effect on the final sensitivity of the analysis.

For the VBF category, it is the statistical contribution to the uncertainty that dominates relative to that from the alternate template fit. Therefore, the statistical error on the template fit is taken as the

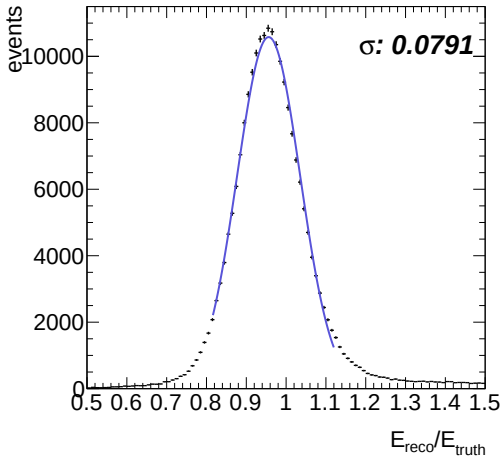


(a) $0 < |\eta| < 0.6$, $100 < p_{\mathrm{T}} < 200$ GeV

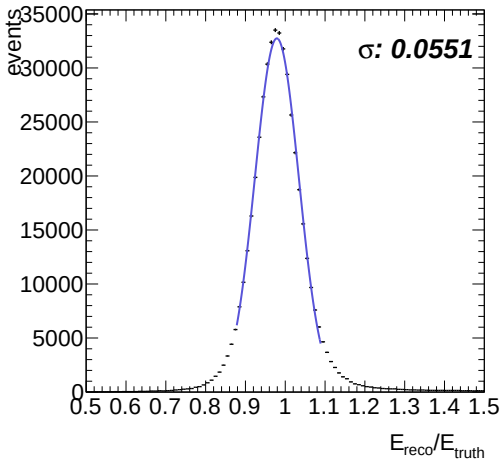
(b) $0 < |\eta| < 0.6$, $200 < p_{\mathrm{T}} < 500$ GeV



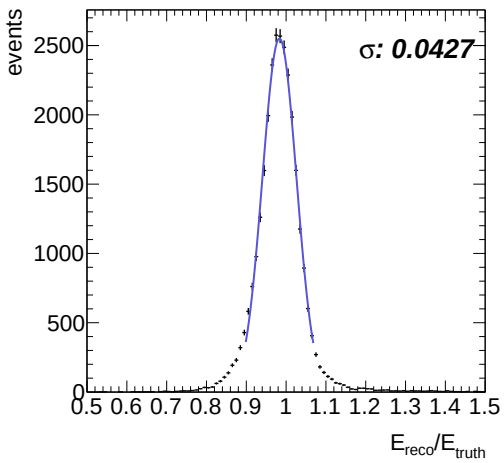
(c) $0 < |\eta| < 0.6$, $p_{\mathrm{T}} > 500$ GeV



(d) $0.6 < |\eta| < 1.2$, $100 < p_{\mathrm{T}} < 200$ GeV

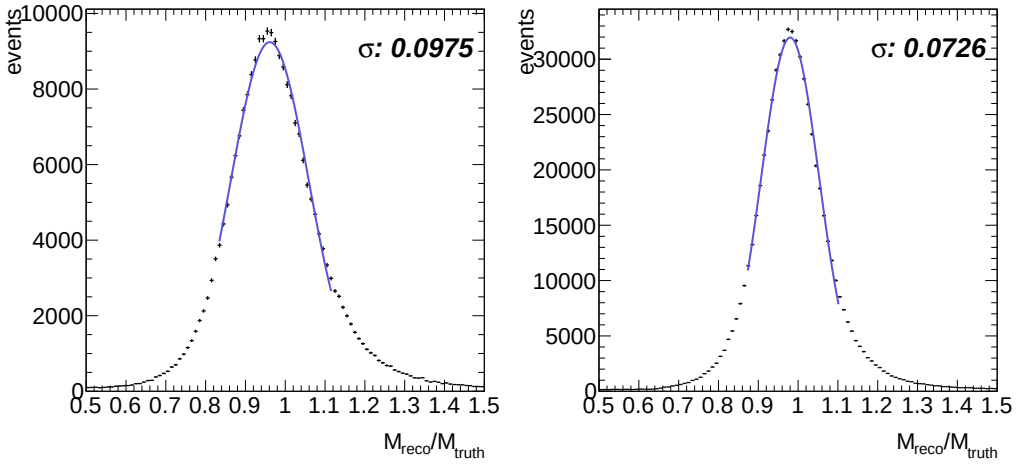


(e) $0.6 < |\eta| < 1.2$, $200 < p_{\mathrm{T}} < 500$ GeV



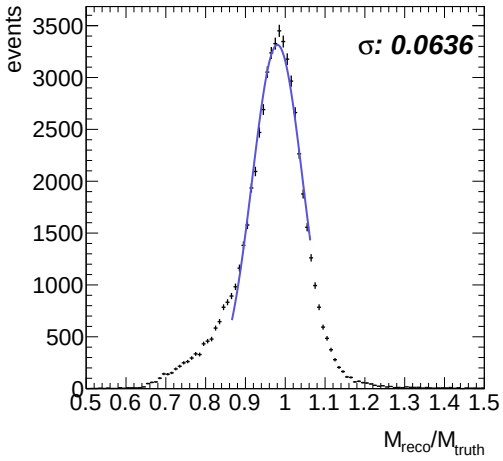
(f) $0.6 < |\eta| < 1.2$, $p_{\mathrm{T}} > 500$ GeV

Figure 5.31: Figures showing the distribution and associated fits used to extract the large-R jet energy resolution in bins of p_{T} and $|\eta|$.

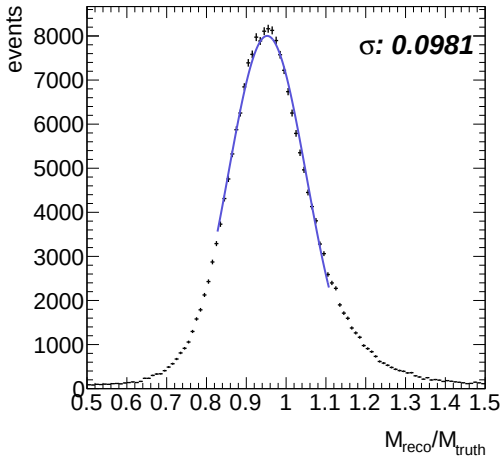


(a) $0 < |\eta| < 0.6$, $100 < p_{\mathrm{T}} < 200$ GeV

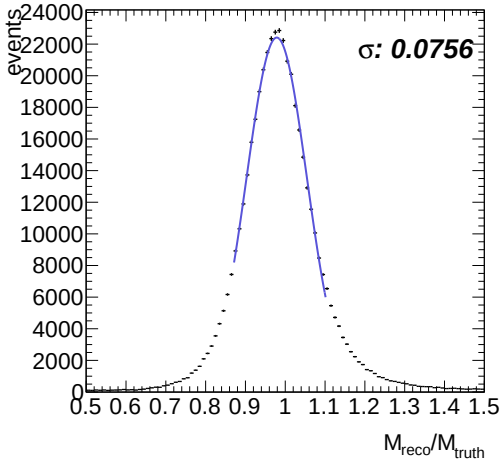
(b) $0 < |\eta| < 0.6$, $200 < p_{\mathrm{T}} < 500$ GeV



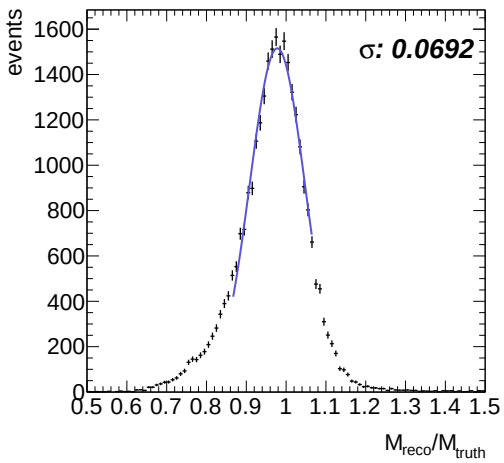
(c) $0 < |\eta| < 0.6$, $p_{\mathrm{T}} > 500$ GeV



(d) $0.6 < |\eta| < 1.2$, $100 < p_{\mathrm{T}} < 200$ GeV



(e) $0.6 < |\eta| < 1.2$, $200 < p_{\mathrm{T}} < 500$ GeV



(f) $0.6 < |\eta| < 1.2$, $p_{\mathrm{T}} > 500$ GeV

Figure 5.32: Figures showing the distribution and associated fits used to extract the large-R jet mass resolution in bins of p_{T} and $|\eta|$.

Table 5.14: List of experimental systematic variations, with short description. These systematics are used as input to the limit setting as nuisance parameters.

Category	Description
Leptons	Muon energy scale†
	Muon spectrometer track resolution†
	Inner detector muon track resolution†
	Electron energy scale†
	Electron energy resolution†
	Lepton isolation
	Lepton reconstruction efficiency
	Electron trigger efficiency
	Muon trigger efficiency
Missing transverse momentum	Soft term resolution†
	Soft term scale†
Resolved jets	Jet energy resolution†
	Jet energy scale, high-p_{T} jets†
	Jet energy scale, in-time pile-up†
	Jet energy scale, out-of-time pile-up†
	Jet energy scale, residual p_{T} dependence from pile-up†
	Jet energy scale, residual dependence on unrelated energy due to jet-area based pile-up correction†
	Jet energy scale, flavor composition†
	Jet energy scale, flavor response†
	Jet energy scale, b-jet response†
	Jet energy scale, in-situ modelling†
	Jet energy scale, in-situ detector related†
	Jet energy scale, statistical systematic†
	Jet energy scale, modelling systematic†
	Jet energy scale, closure systematic for AFII samples†
	b-tagging and mis-tagging uncertainty
Merged jets	Jet energy scale†
	Jet mass scale†
	Jet mass resolution†
	Jet energy resolution†
Interference	ggF interference cross-section uncertainty
	VBF interference modelling uncertainty
	VBF interference dependence on QCD scale
Background Modelling	Top reweighting uncertainty
	W+jets reweighting uncertainty
	Multijet normalization (electron channel)
	Multijet normalization (muon channel)
Misc.	Recorded integrated luminosity

†CAF variations

uncertainty on the VBF MJ estimate.

5.12.6 Missing Transverse Momentum

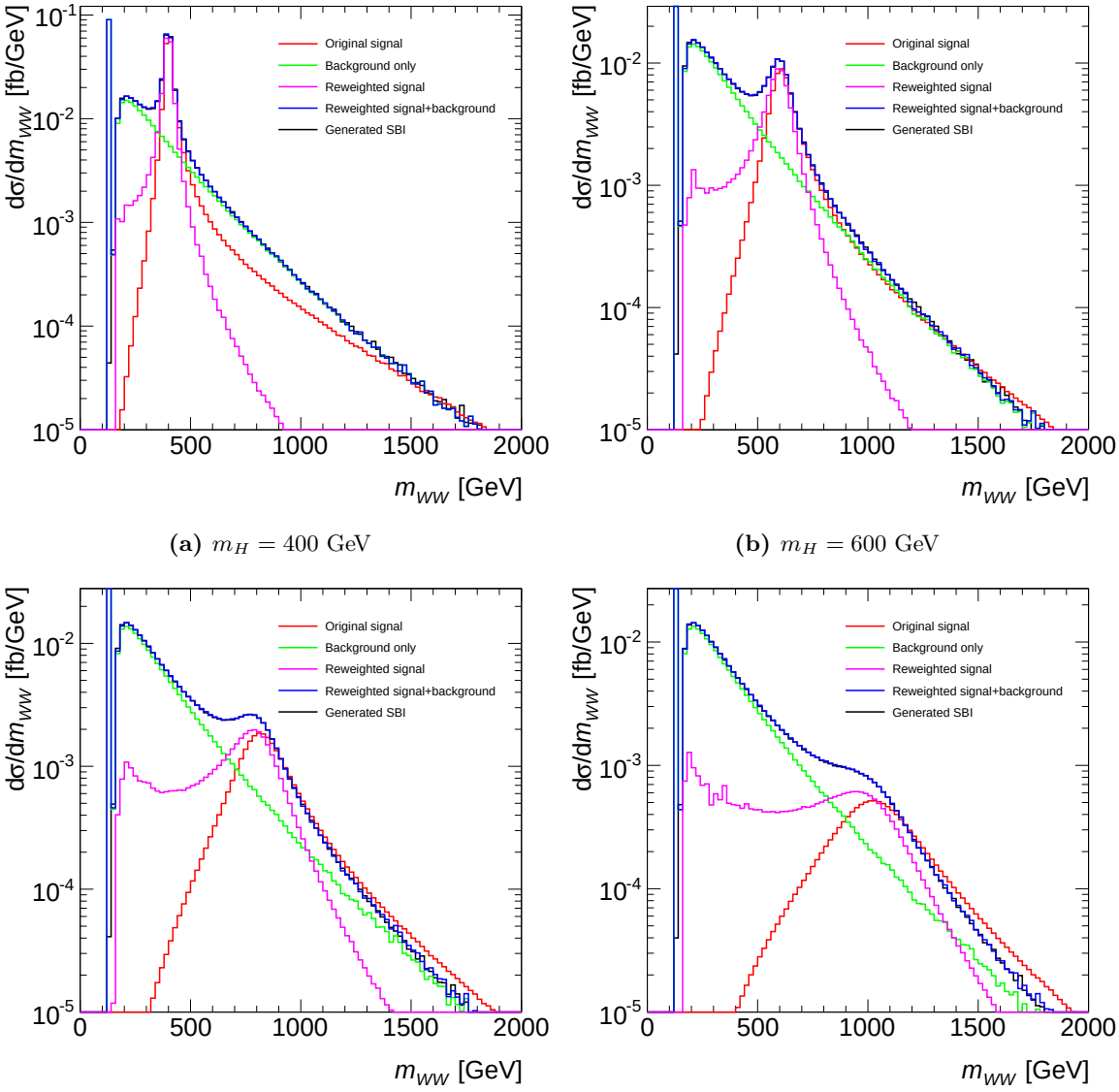
In Section 4.7 the expression for $E_{\mathrm{T}}^{\mathrm{miss}}$ in ATLAS is displayed in Equation 4.6. The equation is separated into components coming from *hard* energy deposits, which are reconstructed as objects, and the *soft terms* which are deposits of energy unassociated with any reconstructed objects. The systematic variations listed above for the reconstructed objects are therefore automatically pushed into the definition of the missing transverse momentum. The one component of the missing transverse momentum that is not varied by the reconstructed objects variations are the uncertainties on the soft terms. The uncertainties on the soft terms come in two components: the soft term energy resolution uncertainty, and the soft term energy scale. The CAF code is able to vary the energy scale and the resolution of the soft terms in order to evaluate these uncertainties.

5.12.7 Interference Uncertainty

The treatment of the uncertainty on the interference for the ggF process is discussed and motivated in Section 2.6, and further discussed in Section 5.4. The difference in magnitude of the interference weights between the *intermediate* and *additive/multiplicative* values form the up and down variations for the interference weight, for each bin of $m_{WW,truth}$.

In the case of the VBF interference, the systematic uncertainty comes from two separate sources. First, there is the modelling uncertainty, which is determined by comparing the modelling of the interference from Repolo with an alternate method of evaluating the interference, and performing a closure test. In the closure test, the $\mathcal{S}+\mathcal{I}$ spectrum obtained by reweighting the $\mathcal{S}$-only Powheg samples with the interference weights evaluated by Repolo is compared to the interference spectrum obtained by using VBFNLO alone. This is done by generating the $\mathcal{B}$, and $\mathcal{SBI}$ invariant m_{WW} spectra from VBFNLO alone and comparing the difference between the Repolo + Powheg $\mathcal{S}+\mathcal{I}$ spectrum and the VBFNLO generated $\mathcal{SBI}-\mathcal{B}$ shape. Figure 5.33 shows the different components of the m_{WW} spectrum that go into the evaluation of this modelling systematic. The symmetrized band (in Figure 5.34) that envelopes the ratio between these spectra (as a function of m_{WW}) is taken as the uncertainty. For $m_H \geq 500$ GeV, the uncertainty is approximately m_H invariant if the ratio is plotted as a function of m_{WW}/m_H. Therefore, for masses above 500 GeV, a common uncertainty is applied. The ratio of the interference weights, with assigned uncertainty band is shown in Figure 5.34, separately for the regions $m_H < 500$ GeV and $m_H \geq 500$ GeV. The symmetrized shaded bands indicate the applied systematic uncertainty in bins of m_{WW}/m_H. The parameterizations that define the uncertainty bands in the plots are given in Table 5.15.

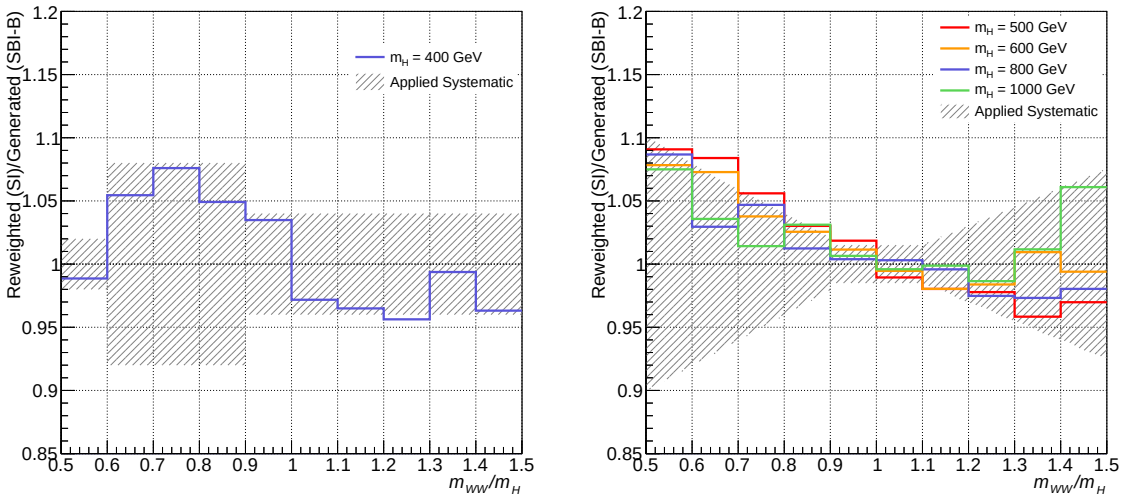
The effect of the QCD scale on the VBF interference weights is also evaluated using Repolo by varying the renormalization and factorization scales between $Q/2 < \mu_R, \mu_F < 2Q$, where Q is the nominal value of the normalization and factorization scales ($Q = m_H$). For these variations, the interference weights are recalculated, and the ratio between the interference spectra is evaluated. The effect was found to be minimal, and not mass dependant (when plotted as a function of m_{WW}/m_H). Therefore, a small overall uncertainty of 0.5% was applied to all interference weights, in addition to the modelling uncertainty.



(a) $m_H = 400$ GeV **(b)** $m_H = 600$ GeV

(c) $m_H = 800$ GeV **(d)** $m_H = 1000$ GeV

Figure 5.33: The various components of the WW invariant mass spectrum generated via different sources/methods. The signal alone, together with the background only and $\mathcal{SBI}$ spectrum are shown, as generated by VBFNLO. This is complemented by the reweighted signal (from Repolo), and the reweighted signal plus background only component. These form the basis for the systematic uncertainty assigned to the VBF interference reweighting method, with the systematic being from the difference of the reweighted signal compared to the VBFNLO generated $\mathcal{SBI} - \mathcal{B}$ spectrum.



(a) **(b)**

Figure 5.34: Parameterization of the modelling uncertainties for the VBF interference, split into the low mass region (a) and the high mass region (b). The ratio of the Repolo generated $\mathcal{S}+\mathcal{I}$ and VBFNLO $\mathcal{SBI}-\mathcal{B}$ – which informed the parameterization of the uncertainty – is also shown for multiple mass points.

Table 5.15: Parameterization of closure systematic as function of m_{WW}/m_H

m_{WW}/m_H	Assigned Systematic [%]
$400 \leq m_H < 500$ GeV	
$m_{WW}/m_H < 0.6$	2
$0.6 < m_{WW}/m_H < 0.9$	8
$m_{WW}/m_H > 0.9$	4
$m_H \geq 500$ GeV	
$m_{WW}/m_H < 0.9$	$(1 - m_{WW}/m_H) \times 20$
$0.9 < m_{WW}/m_H < 1.1$	1.5
$m_{WW}/m_H > 1.1$	$(m_{WW}/m_H - 1) \times 15$

Table 5.16: Binning range and number of bins used for final $m_{\ell\nu jj}$ fit in the ggF and VBF categories. Higgs masses falling inbetween the m_H values specified use the binning and fit range of the next lowest mass point.

Higgs Mass [GeV]	Fit Range [GeV]	Number of bins	
		ggF	VBF
300	200 – 900	35	17
350	210 – 910	31	15
400	220 – 920	28	14
450	240 – 940	25	12
500	260 – 960	23	11
550	280 – 980	21	10
600	300 – 1000	20	10
650	320 – 1020	18	9
700	340 – 1040	17	8
750	360 – 1060	16	8
800	380 – 1080	15	7
900	410 – 1110	14	7
900	440 – 1140	14	7
950	470 – 1170	13	6
≥ 1000	500 – 1200	12	6

5.13 Limit Setting and Statistics Framework

The limit setting and statistical framework utilizes the modified confidence interval method (CL_s) described in Refs. [127, 128] and implemented in `HistFactory` [129], as part of the ROOT software framework [130]. The final fit to determine the upper limit on the $\sigma \times$ BR for the $H \to WW \to \ell\nu jj$ process, is to the reconstructed Higgs invariant mass distribution, $m_{\ell\nu jj}$. Lepton charge and flavour categories are separated in the SR portion of the fit, but are summed in the WCR and TCR in order to increase statistics.

The binning and range of the $m_{\ell\nu jj}$ histogram were determined empirically, balancing individual bin statistics with bin granularity. The resulting binning is mass and category dependant and is documented in Table 5.16. For the case of the NWA limits, the binning and mass ranges above $m_H = 1000$ GeV are kept identical to the $m_H = 1000$ GeV case; this was done because of insufficient statistics at higher values of $m_{\ell\nu jj}$.

In order to scale the signal, and the backgrounds whose overall normalizations are extracted from the data (W+jets and Top), *strength parameters*, which float the normalizations of the signal and background, are used. For the signal, the strength parameter is given the symbol μ, and is defined as $\mu = {}^{\sigma_{meas.}}/_{\sigma_{theory}}$; the background strength parameter(s), μ_b, are analogously constructed. For the models used here, there are no concrete predictions for the value of σ_{theory}, and in our case, where the final limits are presented in terms of $\sigma \times$ BR (*not* on μ), the value for σ_{theory} is somewhat arbitrary. For our purposes, we assign $\sigma_{theory} = \sigma_{\mathrm{SM}}$, where σ_{SM} is m_H dependent. With this parameterization, an example binned likelihood, for one bin in the fit (and an associated control region for the background) can be written as the product of Poisson probabilities:

$$L(\mu, \mu_b) = P(N|\mu s + \mu_b b_{\mathrm{SR}}^{\mathrm{exp}}) \times P(M|\mu_b b_{\mathrm{CR}}^{\mathrm{exp}}). \tag{5.19}$$

Here $b_{\mathrm{SR}}^{\mathrm{exp}}$ is the expected background in the SR, $b_{\mathrm{CR}}^{\mathrm{exp}}$ is the expected background in the CRs, s is the expected signal in the SR, and N (M) is the observed number of events in the SR (CR). The expected background yields are evaluated for $\mu_{(b)} = 1$.

In order to help constrain the background contributions, the TCR and WCR events are also included in the likelihood, along with the SR contribution. However, the full information from the TCR/WCR is not used, and instead, each CR is merged into a single bin, thereby only contributing overall background normalization constraints. As mentioned in Section 5.3.4, the overall normalizations of the W+jets and Top backgrounds are obtained from the final fit, instead of from MC prediction. So, in addition to the signal strength parameter μ, the likelihood also incorporates further strength parameters that scale the W+jets and Top background normalization separately.

The systematic uncertainties outlined in Section 5.12 are incorporated into the statistical treatment in the form of nuisance parameters, with the full set forming the vector $\vec{\theta}$. Each of these nuisance parameters is constrained about its nominal value $N(\tilde{\theta}|\theta)$, where $\tilde{\theta}$ is the nominal expectation (determined from an auxiliary measurement), by a unit Gaussian. The signal and background rates are then parameterized in terms of these nuisance parameters (their values depend on $\vec{\boldsymbol{\theta}}$) so that the following holds, where s_0 is the nominal expectation of the number of signal events (or background events).

$$s = s_0 \times \prod_{i=\theta_i,\theta_2,\ldots} \nu(\theta_i) \tag{5.20}$$

The form of the function $\nu(\theta_i)$ can vary depending on the nuisance parameter, but in the case of this analysis, $\nu(\theta)$ is of the form $\nu(\theta) = \kappa^{\theta}$. Since θ is is normally distributed, $\nu(\theta)$ is lognormally distributed.

The complete likelihood can then then written as the product of the Poisson probabilities over all bins, along with the nuisance parameters:

$$L(\mu, \boldsymbol{\theta}) = \left\{ \prod_{i=\substack{e^+,e^- \\ \mu^+,\mu^-}} \prod_{j=\substack{\mathrm{ggF} \\ \mathrm{VBF}}} \prod_{k=1}^{k=N_{bins}} P\left(N_{ijk} \middle| \mu s_{ijk} + \sum_{Nbkg} b_{ijk} \right) \right\} \times \prod_{i=1}^{N_\theta} N(\tilde{\theta}_i|\theta_i). \tag{5.21}$$

With this definition of the likelihood, the profile likelihood, $\tilde{\lambda}(\mu)$, can be written:

$$\tilde{\lambda}(\mu) = \begin{cases} \frac{L(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(\hat{\mu},\hat{\boldsymbol{\theta}})} & 0 \leq \hat{\mu} \leq \mu \\ \frac{L(0,\hat{\hat{\boldsymbol{\theta}}}(0))}{L(\hat{\mu},\hat{\boldsymbol{\theta}})} & \hat{\mu} < 0 \end{cases} . \tag{5.22}$$

In Equation 5.22, $\hat{\boldsymbol{\theta}}$ and $\hat{\mu}$ are defined as the values of $\boldsymbol{\theta}$ and μ that unconditionally maximize the likelihood. $\hat{\hat{\boldsymbol{\theta}}}$ is the value of the nuisance parameters that maximize the likelihood for some given signal strength μ. The signal hypothesis is assumed to only contribute a positive value on top of the background only hypothesis. Therefore, in the case that the unconditional maximum likelihood estimation (MLE) for μ ($\hat{\mu}$) is less than zero, the profile likelihood sets $\mu = 0$.

The test statistic, $\tilde{q}_\mu$, is used to determine the upper limits, with the condition that the case that $\hat{\mu} > \mu$ does not contribute to the rejection of the background-only hypothesis.

$$\tilde{q}_\mu = \begin{cases} -2\ln\tilde{\lambda}(\mu) & \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases} = \begin{cases} -2\ln\frac{L(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{L(\hat{\mu},\hat{\boldsymbol{\theta}})} & 0 \leq \hat{\mu} \leq \mu \\ -2\ln\frac{L(0,\hat{\hat{\boldsymbol{\theta}}}(0))}{L(\hat{\mu},\hat{\boldsymbol{\theta}})} & \hat{\mu} < 0 \\ 0 & \hat{\mu} > \mu \end{cases} \tag{5.23}$$

With this definition of the test statistic, lower values for $\tilde{q}_\mu$ indicate better compatibility between the data and the background plus signal (with strength μ), while larger values indicate incompatibility.

Finally the CL_s can be constructed as:

$$CL_s = \frac{p_{s+b}}{1 - p_b}, \tag{5.24}$$

$$p_{s+b} = \int_{\tilde{q}_\mu}^{\infty} f(\tilde{q}_\mu | \mu, \hat{\hat{\boldsymbol{\theta}}}(\mu))\ \mathrm{d}\tilde{q}_\mu, \tag{5.24a}$$

$$p_b = \int_{-\infty}^{\tilde{q}_{\mu=0}} f(\tilde{q}_{\mu=0} | 0, \hat{\hat{\boldsymbol{\theta}}}(\mu = 0))\ \mathrm{d}\tilde{q}_\mu. \tag{5.24b}$$

The values of p_{s+b} and p_b come from a fit of the data assuming either the presence of signal, with some $\mu > 0$ (in the case of p_{s+b}), or the p-value associated with the data fit to the background only hypothesis, with $\mu = 0$ (p_b). The upper limit of μ at the 95% CL is given by the μ value that yields 0.95 for CL_s. Once the upper limit on μ has been evaluated, the value is converted to $\sigma \times \mathrm{BR}$.

The same procedure is applied for each of the different mass hypotheses and signal model/lineshapes: SM-like, NWA, EWS. For the case of the EWS signal shapes, the fits are done independently for each mass and width, providing upper limits between $300 - 1000$ GeV for $0.2 \leq \kappa'^2 \leq 1.0$. The EWS signal histograms, which act as signal template inputs for the limit setting with the various widths considered, are found in Appendix C for reference.

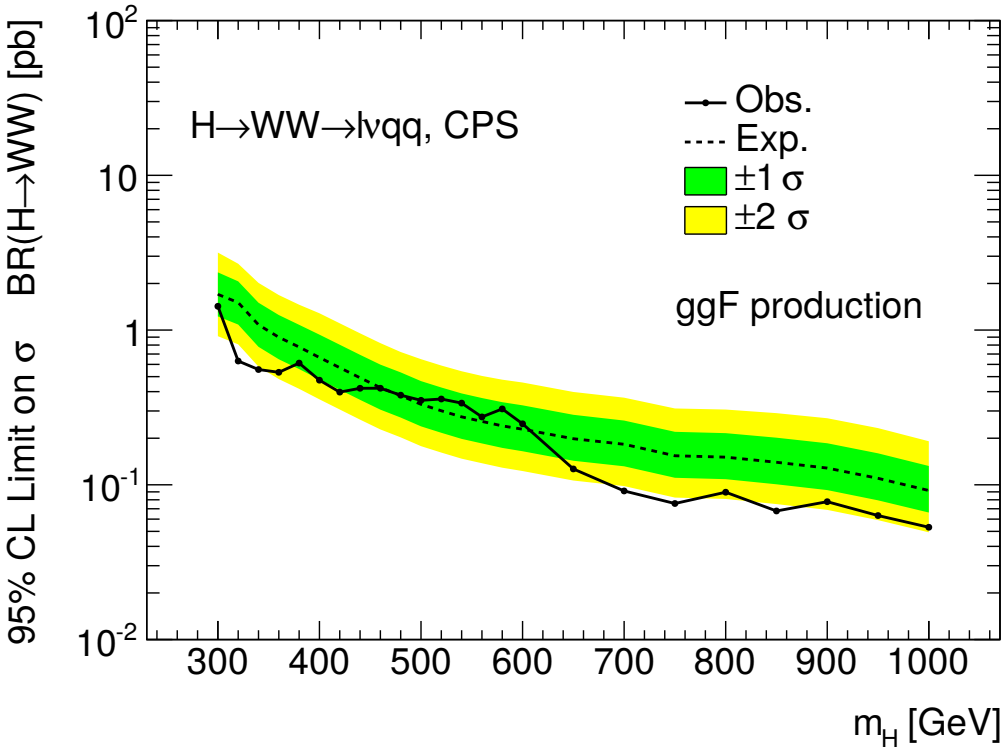
5.14 Results

The upper limits on $\sigma \times \mathrm{BR}$ for the three models being considered, SM-like, EWS, and NWA, are shown in Figures 5.35, 5.36, and 5.37 respectively. The EWS limits are shown as $\frac{\sigma}{\kappa'^2} \times \mathrm{BR}$ in order to have the limit curves for all κ'^2 values well separated.

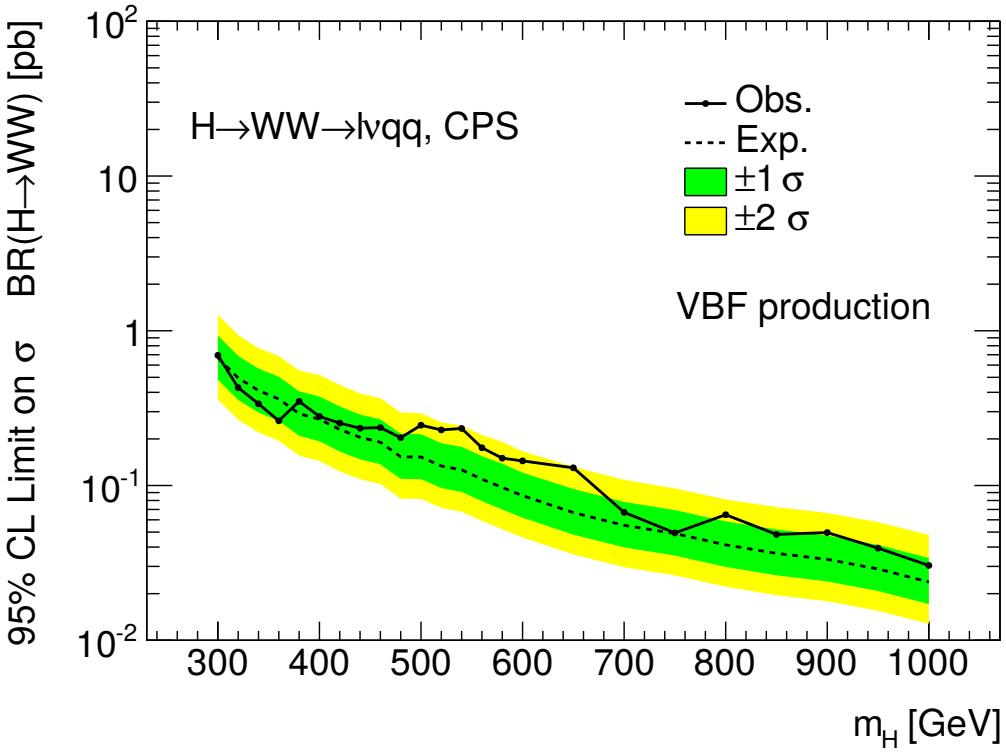
Although the NWA Higgs is considered in the range $300 \leq m_H \leq 2000$ GeV, the limits (in Figure 5.36) only are shown up to $m_H = 1200$ GeVas at higher masses the upper limit is constant. This is a result of the SR fit range used in the final limit setting (Table 5.16); above $m_H \approx 1200$ GeV all of the signal events are in the last bin of the $m_{\ell\nu jj}$ histogram, allowing no shape information to discriminate between the Higgs hypotheses above 1200 GeV.

In the Standard Model, the relative cross-sections of the ggF and VBF production mechanisms is known. For model-independent BSM Higgs theories, there is no defined relationship that is able to constrain the ratio between the two production mechanisms. Therefore, for the pseudo-model-independent limits set for the SM-like and NWA Higgs models, the ggF and VBF limits cannot be combined, and so are calculated and presented separately. In the case of the EWS model, the ratio between ggF/VBF production is assumed to obey the same relationship as in the SM. Therefore, the combined ggF+VBF limits can be calculated.

As expected, the limits in all channels improve as a function of Higgs mass hypothesis, due to lower expected backgrounds in the high $m_{\ell\nu jj}$ region. The impact of the much wider CPS/SM-like lineshape, compared to that of the NWA sample is also apparent, with the limit set on the latter signal model being more stringent over the full mass range considered.

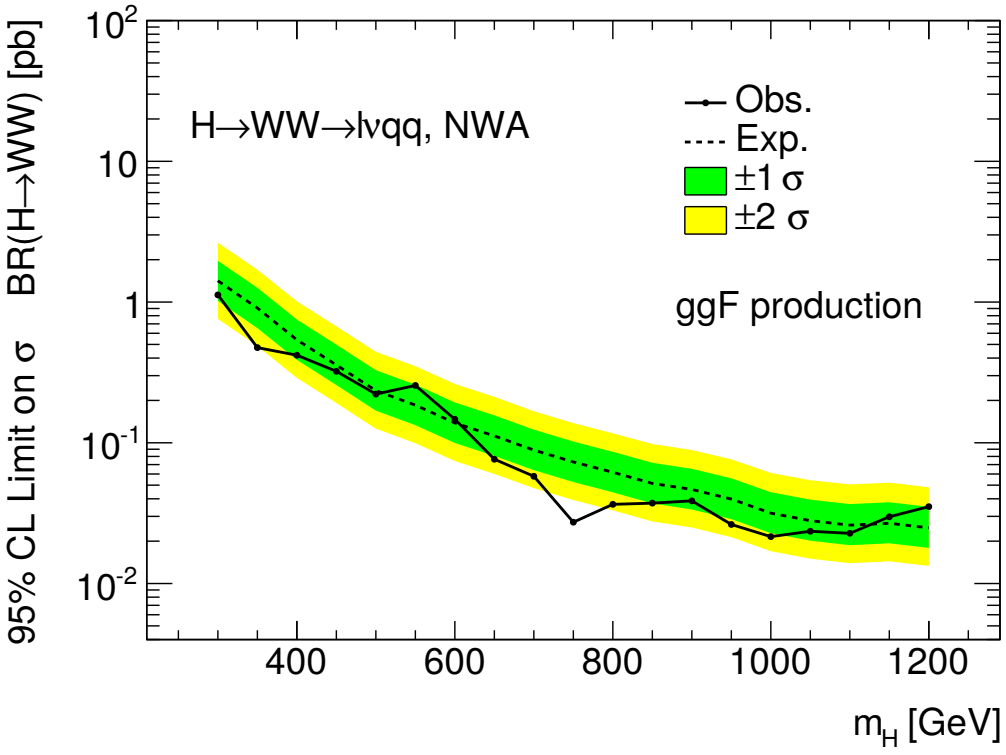


(a)

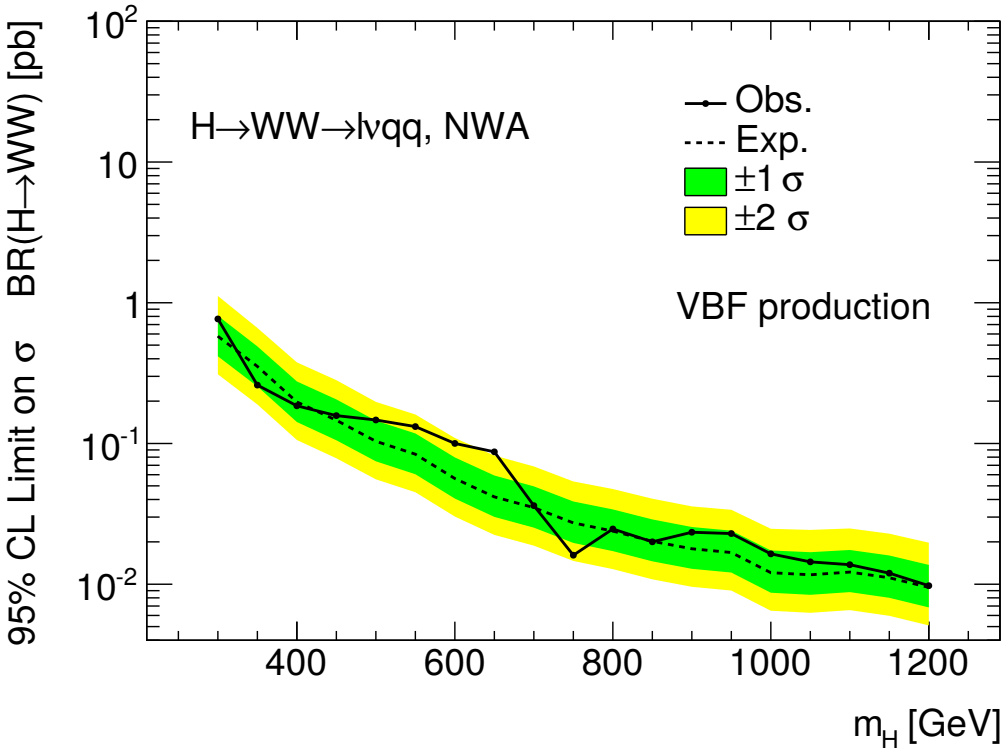


(b)

Figure 5.35: Limits on $\sigma \times$ BR (at 95% CL_S) for an SM-like Higgs for ggF production (a) and VBF production (b).



(a)



(b)

Figure 5.36: Limits on $\sigma \times \mathrm{BR}$ (at 95% $\mathrm{CL_s}$) for a Higgs produced with the NWA lineshape. The ggF category limits are shown in (a) and the VBF limits shown in (b).

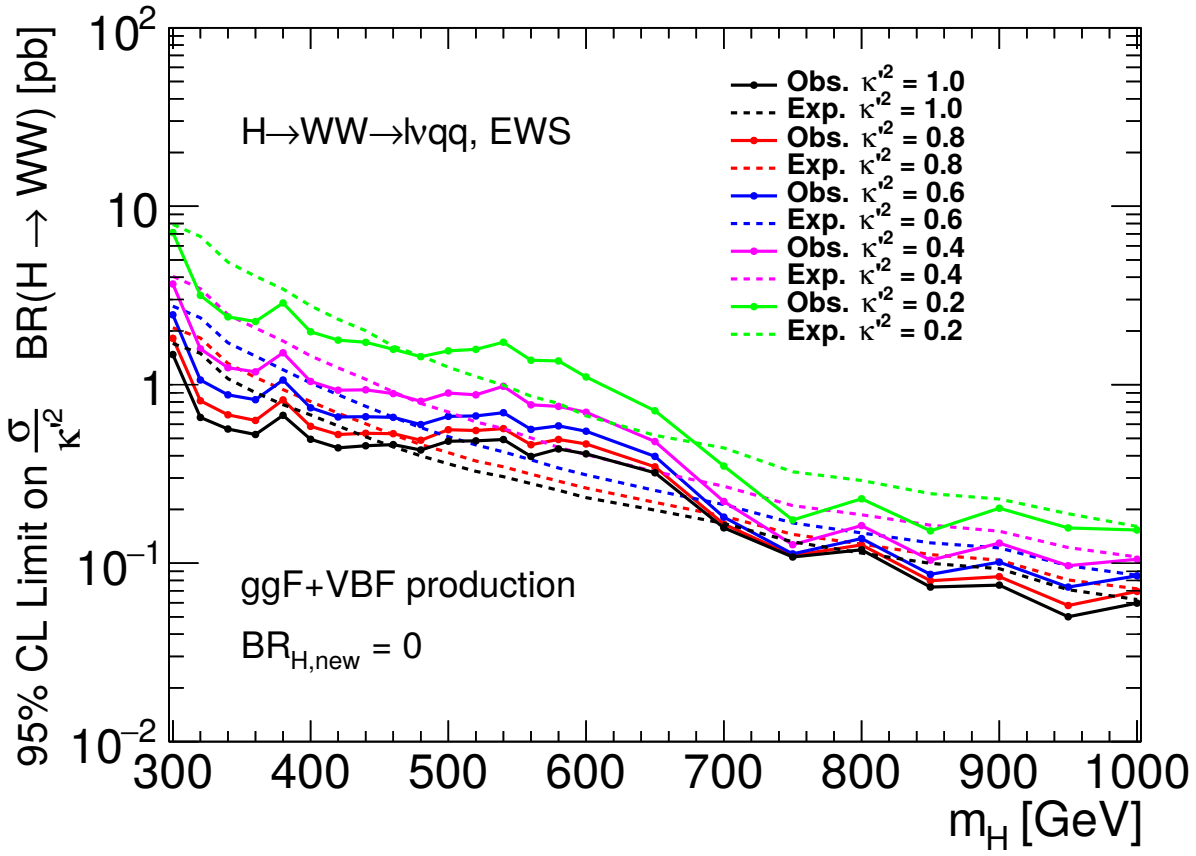


Figure 5.37: Expected and observed upper limits on $\frac{\sigma}{\kappa'^2} \times \mathrm{BR}$ for the EWS Higgs model, for $\kappa'^2 = 0.2$, 0.4, 0.6, 0.8 and 1.0 (ggF and VBF production modes combined).

Chapter 6

Conclusion

"Isn't it enough to see that a garden is beautiful without having to believe that there are fairies at the bottom of it too?"

Douglas Adams

The results of a search for a heavy Higgs boson in the $H \to WW \to \ell\nu jj$ channel have been reported (Section 5.14) for SM-like, EWS and NWA Higgs signal models/lineshapes. In all cases, and for all mass points (and widths in the case of the EWS model interpretation) no significant deviation from the background-only hypothesis was found, resulting in upper limits being set at the 95% confidence level (using the CL_s method). As mentioned in Chapter 1, the results presented here do not conclusively rule out an extended sector, and are not the last results expected from Run-I which probe an extended Higgs sector in the $H \to WW$ decay channel. The analysis presented here will be combined with the high-mass $H \to WW \to \ell\nu\ell\nu$ ($\ell = e, \mu$) analysis for publication. Further progress in this analysis channel can be made by analyzing the data from LHC's Run-II, which will deliver even more integrated luminosity, at a higher center of mass energy, enabling even more stringent limits (and perhaps an observation) in the future.

6.1 Combination with $H \to WW \to \ell\nu\ell\nu$

Currently, the analysis presented in this thesis is going through ATLAS approval in order to be published as an ATLAS result. The paper will present the results in combination with the $H \to WW \to \ell\nu\ell\nu$ high-mass analysis (SM-like, NWA and EWS Higgs signal models). The $H \to WW \to \ell\nu\ell\nu$ analysis, because of the distinctive di-lepton signature, has lower backgrounds throughout the range of m_H considered, and also has a lower contribution from the multijet background (which is large in the low mass region, and is only rejected in the $H \to WW \to \ell\nu jj$ analysis because of the strict $E_{\mathrm{T}}^{\mathrm{miss}}$ cut applied). However, the branching fraction of $W \to \ell\nu$ is lower than that of the hadronic W decays, lowering the signal yields. Finally, unlike the $H \to WW \to \ell\nu jj$ analysis, because of the presence of multiple neutrinos in the final state, the invariant mass of the Higgs cannot be calculated (the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector in this case will be the vector sum of the neutrino momenta), and therefore the final fit has to be done on the Higgs transverse mass, m_{T}, which is not a sharply peaked distribution, lowering the sensitivity.

Therefore, the $H \to WW \to \ell\nu jj$ analysis is the more sensitive analysis at high-mass.

The other dominant decay mode for heavy Higgs final states is $H \to ZZ$, which is also being studied in the context of the 2HDM model (based on limits set with the NWA). Because of the different signal models considered in the $H \to ZZ$ analysis, there is no plan to combine these results with the $H \to WW$ analyses.

6.2 LHC Run-II Prospects

For Run-II, the LHC plans to deliver $\approx$ 100 fb^{-1} of integrated luminosity, between 2015 and 2017. The additional luminosity will also be at a higher center of mass energy of $\sqrt{s} = 13$ TeV. The increased collision energy will impact the cross-sections for Higgs production, with the total SM Higgs cross-sections increasing by almost an order of magnitude for $m_H \leq 1000$ GeV. The pile-up conditions for Run-II are also more challenging; instead of 50 ns between each bunch-crossing, the LHC will fill colliding bunches with only 25 ns between collisions (the design bunch separation), and have a much larger average number of collisions per bunch crossing of $\sim$50 (compared to $\sim$21 in the 2012 run). The larger dataset, and production cross-sections, will allow even more stringent limits on additional Higgs bosons, whether it is through a direct search, or through precision measurements of the 125 GeV Higgs properties, which will apply indirect constraints on an extended Higgs sector.

The next iteration of measurements in this channel will not only benefit from the increased luminosity and cross-sections, but also the lessons learned from the 2012 analysis. For example, the mass dependant optimizations were seen to give an improvement to the overall limits, but severely impact the final event yields. This is due to the fact that many of the variables used are highly correlated to the reconstructed Higgs mass, $m_{\ell\nu jj}$ (which is the final discriminant). Further tuning of the mass dependant selection, which reduced the number of both the signal and background events greatly, could allow finer binning in the final discriminant histogram, resulting in better sensitivity for narrow resonances.

Additionally, the choice of large-R groomed jet (defining the merged jet category in the analysis) in our analysis could also use further optimizations. There has been extensive work done in ATLAS to validate boosted boson taggers, specifically designed in order to tag hadronic W/Z decays, which could further improve our background rejection for the higher Higgs mass hypotheses.

Run-I has been a spectacular success, with over 400 papers published (and more still on the way). Analysis of the Run-I dataset also resulted in the discovery of what appears to be the Standard Model Higgs boson; precision measurements of this Higgs, together with direct heavy Higgs searches (such as the one presented in this thesis) have already begun to place limits on a possible extended Higgs sector. The large increase in data that will be available at the end of Run-II, collected at a higher center of mass energy, will allow even stricter limits on BSM Higgs models (and BSM physics in general). Because Run-I has been incredibly successful, this has raised expectations, and the hopes for Run-II are even higher, and I genuinely look forward to what is around the corner.

Bibliography

[1] P. W. Higgs, Broken symmetries, massless particles and gauge fields, Phys. Lett. **12** (1964) pp. 132–133.

[2] P. W. Higgs, Broken Symmetries and the Masses of Gauge Bosons, Phys. Rev. Lett. **13** (1964) pp. 508–509.

[3] P. W. Higgs, Spontaneous Symmetry Breakdown without Massless Bosons, Phys. Rev. **145** (1966) pp. 1156–1163.

[4] G. Guralnik et al., Global conservation laws and massless particles, Phys. Rev. Lett. **13** (1964) pp. 585–587.

[5] T. W. B. Kibble, Symmetry breaking in non-Abelian gauge theories, Phys. Rev. **155** (1967) pp. 1554–1561.

[6] F. Englert et al., Broken Symmetry and the Mass of Gauge Vector Mesons, Phys. Rev. Lett. **13** (1964) pp. 321–323.

[7] LEP Collaborations (ALEPH, DELPHI, L3 and OPAL), LEP Electroweak Working Group and SLD Heavy Flavour Group, A Combination of preliminary electroweak measurements and constraints on the standard model (2001), arXiv: `hep-ex/0103048 [hep-ex]`.

[8] LEP Collaborations (ALEPH, DELPHI, L3 and OPAL), Search for the standard model Higgs boson at LEP, Phys. Lett. **B565** (2003) pp. 61–75, arXiv: `hep-ex/0306033 [hep-ex]`.

[9] CDF and DØ Collaborations, Higgs Boson Studies at the Tevatron, Phys. Rev. **D88**.5 (2013) p. 052014, arXiv: `1303.6346 [hep-ex]`.

[10] A. Djouadi, The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model, Phys. Rept. **457** (2008) pp. 1–216, arXiv: `hep-ph/0503172 [hep-ph]`.

[11] ATLAS Collaboration, Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC, Phys. Lett. B **716** (2012) 1–29. 39 p, arXiv: `1207.7214 [hep-ph]`.

[12] CMS Collaboration, Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC, Phys. Lett. B **716** (2012) 30–61. 59 p, arXiv: `1207.7235 [hep-ph]`.

[13] ATLAS Collaboration, Search for the Higgs boson in the $H \to WW \to \ell\nu jj$ decay channel at $\sqrt{s} = 7$ TeV with the ATLAS detector, Phys. Lett. **B718** (2012) pp. 391–410, arXiv: `1206.6074 [hep-ex]`.

[14] LHC Higgs Cross Section Working Group, Handbook of LHC Higgs Cross Sections: 3. Higgs Properties (2013), arXiv: `1307.1347 [hep-ph]`.

[15] S. L. Glashow, Partial symmetries of weak interactions, Nucl. Phys. **22** (1961) pp. 579–588.

[16] S. Weinberg, A model of leptons, Phys. Rev. Lett. **19** (1967) pp. 1264–1266.

[17] A. Salam, *Elementary particle theory*, Stockholm: Almqvist and Wiksells, 1968.

[18] K. Olive et al., Review of Particle Physics, Chin. Phys. **C38** (2014) p. 090001.

[19] C. Quigg, *Gauge theories of the strong, weak, and electromagnetic interactions; 2nd ed.* Princeton, NJ: Princeton University Press, 2013.

[20] Y. Nambu, Quasiparticles and Gauge Invariance in the Theory of Superconductivity, Phys. Rev. **117** (1960) pp. 648–663.

[21] J. Goldstone, Field Theories with Superconductor Solutions, Nuovo Cim. **19** (1961) pp. 154–164.

[22] J. Goldstone et al., Broken Symmetries, Phys. Rev. **127** (1962) pp. 965–970.

[23] C. P. Burgess et al., *The Standard Model: A Primer*, Cambridge: Cambridge Univ. Press, 2007.

[24] S. Dittmaier, C. Mariotti, G. Passarino and R. Tanaka, *The Higgs Cross Section Working Group web page*, http://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWGCrossSectionsFigures, 2015.

[25] LHC Higgs Cross Section Working Group, Handbook of LHC Higgs Cross Sections: 2. Differential Distributions, CERN-2012-002 (CERN, Geneva, 2012), arXiv: `1201.3084 [hep-ph]`.

[26] S. Goria et al., The Higgs Boson Lineshape (2012), arXiv: `1112.5517 [hep-ph]`.

[27] ATLAS Collaboration, Evidence for the spin-0 nature of the Higgs boson using ATLAS data, Phys. Lett. B **726** (2013) p. 120, arXiv: `1307.1432 [hep-ex]`.

[28] CMS Collaboration, Constraints on the spin-parity and anomalous HVV couplings of the Higgs boson in proton collisions at 7 and 8 TeV (2014), arXiv: `1411.3441 [hep-ex]`.

[29] CMS Collaboration, Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV (2014), arXiv: `1412.8662 [hep-ex]`.

[30] ATLAS Collaboration, Observation and measurement of Higgs boson decays to WW^* with the ATLAS detector (2014), arXiv: `1412.2641 [hep-ex]`.

[31] ATLAS Collaboration, Measurements of Higgs boson production and couplings in the four-lepton channel in *pp* collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector, Phys. Rev. D **91** (2015) p. 012006, arXiv: `1408.5191 [hep-ex]`.

[32] ATLAS Collaboration, Measurement of Higgs boson production in the diphoton decay channel in *pp* collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector, Phys. Rev. D **90** (2014) p. 112015, arXiv: `1408.7084 [hep-ex]`.

[33] G. Branco et al., Theory and phenomenology of two-Higgs-doublet models, Phys. Rept. **516** (2012) pp. 1–102, arXiv: `1106.0034 [hep-ph]`.

[34] H. E. Haber et al., The Search for Supersymmetry: Probing Physics Beyond the Standard Model, Phys. Rept. **117** (1985) pp. 75–263.

[35] G. Passarino, Higgs Interference Effects in $gg \to ZZ$ and their Uncertainty, JHEP **1208** (2012) p. 146, arXiv: `1206.3824 [hep-ph]`.

[36] J. M. Campbell et al., Gluon-gluon contributions to W+ W- production and Higgs interference effects, JHEP **1110** (2011) p. 005, arXiv: 1107.5569 [hep-ph].

[37] N. Kauer, Signal-background interference in gg →H →VV, PoS **RADCOR2011** (2011) p. 027, arXiv: 1201.1667 [hep-ph].

[38] A. Ballestrero et al., PHANTOM: A Monte Carlo event generator for six parton final states at high energy colliders, Comput.Phys.Commun. **180** (2009) pp. 401–417, arXiv: 0801.3359 [hep-ph].

[39] 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, JHEP **1407** (2014) p. 079, arXiv: 1405.0301 [hep-ph].

[40] K. Arnold et al., Release Note – Vbfnlo-2.6.0 (2012), arXiv: 1207.4975 [hep-ph].

[41] K. Arnold et al., VBFNLO: A parton level Monte Carlo for processes with electroweak bosons – Manual for Version 2.5.0 (2011), arXiv: 1107.4038 [hep-ph].

[42] K. Arnold et al., VBFNLO: A parton level Monte Carlo for processes with electroweak bosons, Comput. Phys. Commun. **180** (2009) p. 1661, arXiv: 0811.4559 [hep-ph].

[43] J. M. Campbell et al., An Update on vector boson pair production at hadron colliders, Phys. Rev. **D60** (1999) p. 113006, arXiv: hep-ph/9905386 [hep-ph].

[44] T. Figy et al., Next-to-leading order jet distributions for Higgs boson production via weak boson fusion, Phys. Rev. **D68** (2003) p. 073005, arXiv: hep-ph/0306109 [hep-ph].

[45] M. Ciccolini et al., Strong and electroweak corrections to the production of Higgs+2 jets via weak interactions at the LHC, Phys. Rev. Lett. **99** (2007) p. 161803, arXiv: 0707.0381 [hep-ph].

[46] M. Ciccolini et al., Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC, Phys. Rev. **D 77** (2008) p. 013002, arXiv: 0710.4749 [hep-ph].

[47] P. Bolzoni et al., Higgs production via vector-boson fusion at NNLO in QCD, Phys. Rev. Lett. **105** (2010) p. 011801, arXiv: 1003.4451 [hep-ph].

[48] P. Bolzoni et al., Vector boson fusion at NNLO in QCD: SM Higgs and beyond, Phys. Rev. **D85** (2012) p. 035002, arXiv: 1109.3717 [hep-ph].

[49] L. Evans et al., LHC Machine, JINST **3** (2008), ed. by L. Evans S08001.

[50] ALICE Collaboration, The ALICE experiment at the CERN LHC, JINST **3** (2008) S08002.

[51] ATLAS Collaboration, The ATLAS experiment at the CERN Large Hadron Collider, JINST **3** (2008) S08003.

[52] LHCb Collaboration, The LHCb Detector at the LHC, JINST **3** (2008) S08005.

[53] CMS Collaboration, The CMS experiment at the CERN LHC, JINST **3** (2008) S08004.

[54] P. Grafstrm et al., Luminosity determination at proton colliders, Prog.Part.Nucl.Phys. **81** (2015) pp. 97–148.

[55] O. S. Bruning et al., LHC Design Report Vol.1: The LHC Main Ring (2004), ed. by O. S. Bruning.

[56] J. Haffner, The CERN accelerator complex. Complexe des acclrateurs du CERN (2013), General Photo, URL: https://cdsweb.cern.ch/record/1621894.

[57] ATLAS Collaboration, *ATLAS detector and physics performance: Technical Design Report, 2*, Technical Design Report ATLAS, Geneva: CERN, 1999, URL: http://cds.cern.ch/record/391177.

[58] V Cindro et al., The ATLAS Beam Conditions Monitor, Journal of Instrumentation **3**.02 (2008) P02004, URL: http://stacks.iop.org/1748-0221/3/i=02/a=P02004.

[59] ATLAS Collaboration, ATLAS pixel detector electronics and sensors, Journal of Instrumentation **3**.07 (2008) P07007, URL: http://stacks.iop.org/1748-0221/3/i=07/a=P07007.

[60] A. Ahmad et al., The silicon microstrip sensors of the ATLAS semiconductor tracker, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **578**.1 (2007) pp. 98–118, ISSN: 0168-9002, URL: http://www.sciencedirect.com/science/article/pii/S0168900207007644.

[61] The ATLAS TRT Collaboration, The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: design and performance, Journal of Instrumentation **3**.02 (2008) P02013, URL: http://stacks.iop.org/1748-0221/3/i=02/a=P02013.

[62] M. Alam et al., The ATLAS silicon pixel sensors, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **456**.3 (2001) pp. 217 –232, ISSN: 0168-9002, URL: http://www.sciencedirect.com/science/article/pii/S016890020000574X.

[63] M. L. Cherry et al., Transition radiation from relativistic electrons in periodic radiators, Phys. Rev. **D10** (1974) pp. 3594–3607.

[64] ATLAS Collaboration, Luminosity Determination in *pp* Collisions at $\sqrt{s} = 7$ TeV Using the ATLAS Detector at the LHC, Eur. Phys. J. C **71** (2011) p. 1630, arXiv: 1101.2185 [hep-ex].

[65] R. Wigmans, *Calorimetry: Energy Measurement in Particle Physics*, International Series of Monographs on Physics, Clarendon Press, 2000, ISBN: 9780198502968.

[66] C. Fabjan et al., Calorimetry for particle physics, Rev. Mod. Phys. **75** (2003) pp. 1243–1286.

[67] ATLAS Collaboration, *ATLAS liquid-argon calorimeter: Technical Design Report*, Technical Design Report ATLAS, Geneva: CERN, 1996, URL: http://cds.cern.ch/record/331061.

[68] ATLAS Collaboration, *ATLAS detector and physics performance: Technical Design Report, 1*, Technical Design Report ATLAS, Geneva: CERN, 1999, URL: https://cds.cern.ch/record/391176.

[69] J Abdallah et al., "Design, Construction and Installation of the ATLAS Hadronic Barrel Scintillator-Tile Calorimeter", tech. rep. ATL-TILECAL-PUB-2008-001. ATL-COM-TILECAL-2007-019, CERN, 2007.

[70] ATLAS Collaboration, Drift Time Measurement in the ATLAS Liquid Argon Electromagnetic Calorimeter using Cosmic Muons, Eur.Phys.J. **C70** (2010) pp. 755–785, arXiv: 1002.4189 [physics.ins-det].

[71] A Artamonov et al., The ATLAS Forward Calorimeter, Journal of Instrumentation **3**.02 (2008) P02010, URL: http://stacks.iop.org/1748-0221/3/i=02/a=P02010.

[72] N. Buchanan et al., ATLAS liquid argon calorimeter front end electronics, JINST **3** (2008) P09003.

[73] A. Bazan et al., ATLAS liquid argon calorimeter back end electronics, JINST **2** (2007) P06002.

[74] N. Buchanan et al., Design and implementation of the Front End Board for the readout of the ATLAS liquid argon calorimeters, JINST **3** (2008) P03004.

[75] W. Cleland et al., Signal processing considerations for liquid ionization calorimeters in a high rate environment, Nucl.Instrum.Meth. **A338**.2-3 (1994) pp. 467–497.

[76] S. Palestini, The Muon Spectrometer of the ATLAS experiment, Nucl. Phys. Proc. Suppl. **125** (2003) pp. 337–345.

[77] A. Yamamoto et al., The ATLAS central solenoid, Nucl. Instrum. Meth. **A584** (2008) pp. 53–74.

[78] A. Yamamoto et al., Progress in ATLAS central solenoid magnet, IEEE Trans. Appl. Supercond. **10** (2000) pp. 353–356.

[79] *ATLAS Luminosity Public Results*, 2015, URL: `http://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults`.

[80] *ATLAS Data Quality Information*, 2015, URL: `http://twiki.cern.ch/twiki/bin/view/AtlasPublic/RunStatsPublicResults2010`.

[81] ATLAS Collaboration, Performance of the ATLAS muon trigger in pp collisions at $\sqrt{s} = 8$ TeV (2014), arXiv: `1408.3179 [hep-ex]`.

[82] W Lampl et al., Calorimeter Clustering Algorithms: Description and Performance, ATL-LARG-PUB-2008-002 (2008), URL: `http://cdsweb.cern.ch/record/1099735`.

[83] ATLAS Collaboration, Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data, Eur. Phys. J. **C74**.10 (2014) p. 3071, arXiv: `1407.5063 [hep-ex]`.

[84] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton–proton collision data*, ATLAS-CONF-2014-032, 2014, URL: `https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2014-032`.

[85] ATLAS Collaboration, Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data, Eur. Phys. J. **C74**.7 (2014) p. 2941, arXiv: `1404.2240 [hep-ex]`.

[86] ATLAS Collaboration, Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data, Eur. Phys. J. **C74**.11 (2014) p. 3130, arXiv: `1407.3935 [hep-ex]`.

[87] ATLAS Collaboration, Expected performance of the ATLAS experiment - detector, trigger and physics (2009), arXiv: `0901.0512 [hep-ex]`.

[88] S. Catani et al., Longitudinally invariant K_t clustering algorithms for hadron hadron collisions, Nucl. Phys. **B406** (1993) pp. 187–224.

[89] S. D. Ellis et al., Successive combination jet algorithm for hadron collisions, Phys. Rev. **D48** (1993) pp. 3160–3166, arXiv: `hep-ph/9305266 [hep-ph]`.

[90] M. Wobisch et al., Hadronization corrections to jet cross-sections in deep inelastic scattering (1998), arXiv: `hep-ph/9907280 [hep-ph]`.

[91] M. Cacciari et al., The Anti-k(t) jet clustering algorithm, JHEP **0804** (2008) p. 063, arXiv: `0802.1189 [hep-ph]`.

[92] ATLAS Collaboration, Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector, Eur. Phys. J. **C75**.1 (2015) p. 17, arXiv: 1406.0076 [hep-ex].

[93] T Barillari et al., "Local Hadronic Calibration", tech. rep. ATL-LARG-PUB-2009-001-2. ATL-COM-LARG-2008-006. ATL-LARG-PUB-2009-001, CERN, 2008.

[94] ATLAS Collaboration, Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV (2011), Submitted to Eur. Phys. J. C, arXiv: 1112.6426 [hep-ex].

[95] S. Agostinelli et al., GEANT4: A Simulation toolkit, Nucl. Instrum. Meth. **A506** (2003) pp. 250–303.

[96] J. Allison et al., Geant4 developments and applications, IEEE Trans. Nucl. Sci. **53** (2006) p. 270.

[97] M. Cacciari et al., Pileup subtraction using jet areas, Phys. Lett. **B659** (2008) pp. 119–126, arXiv: 0707.1378 [hep-ph].

[98] ATLAS Collaboration, *Pile-up subtraction and suppression for jets in ATLAS*, ATLAS-CONF-2013-083, 2013, URL: http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2013-083.

[99] ATLAS Collaboration, Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data, ATLAS-CONF-2011-102 (2011), http://cdsweb.cern.ch/record/1369219.

[100] ATLAS Collaboration, *Calibration of b-tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment*, ATLAS-CONF-2014-004, 2014, URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2014-004/.

[101] ATLAS Collaboration, *Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 ATLAS data*, ATLAS-CONF-2014-046, 2014, URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2014-046.

[102] ATLAS Collaboration, *Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton–Proton Collisions recorded in 2012 at* $\sqrt{s} = 8$ *TeV*, ATLAS-CONF-2013-082, 2013, URL: http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2013-082.

[103] ATLAS Collaboration, Performance of missing transverse momentum reconstruction in proton–proton collisions at $\sqrt{s} = 7$ TeV with ATLAS, Eur. Phys. J. C **72** (2012) p. 1844, arXiv: 1108.5602 [hep-ex].

[104] ATLAS Collaboration, The ATLAS Simulation Infrastructure, Eur. Phys. J. C **70** (2010) p. 823, arXiv: 1005.4568 [hep-ex].

[105] G. Corcella et al., HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes), JHEP **0101** (2001) p. 010, arXiv: hep-ph/0011363 [hep-ph].

[106] J. Butterworth et al., Multiparton interactions in photoproduction at HERA, Z. Phys. **C72** (1996) pp. 637–646, arXiv: hep-ph/9601371 [hep-ph].

[107] T. Gleisberg et al., Event generation with SHERPA 1.1, JHEP **0902** (2009) p. 007, arXiv: 0811.4622 [hep-ph].

[108] S. Höche et al., QCD matrix elements and truncated showers, JHEP **0905** (2009) p. 053, arXiv: 0903.1219 [hep-ph].

[109] T. Gleisberg et al., Comix, a new matrix element generator, JHEP **0812** (2008) p. 039, arXiv: 0808.3674 [hep-ph].

[110] S. Schumann et al., A Parton shower algorithm based on Catani-Seymour dipole factorisation, JHEP **0803** (2008) p. 038, arXiv: 0709.1027 [hep-ph].

[111] T. Sjöstrand et al., PYTHIA 6.4 Physics and Manual, JHEP **0605** (2006) p. 026, arXiv: hep-ph/0603175.

[112] T. Sjostrand et al., A Brief Introduction to PYTHIA 8.1, Comput. Phys. Commun. **178** (2008) pp. 852–867, arXiv: 0710.3820 [hep-ph].

[113] M. L. Mangano et al., ALPGEN, a generator for hard multiparton processes in hadronic collisions, JHEP **0307** (2003) p. 001, arXiv: hep-ph/0206293 [hep-ph].

[114] B. P. Kersevan et al., The Monte Carlo event generator AcerMC versions 2.0 to 3.8 with interfaces to PYTHIA 6.4, HERWIG 6.5 and ARIADNE 4.1, Comput. Phys. Commun. **184** (2013) pp. 919–985, arXiv: hep-ph/0405247 [hep-ph].

[115] P. Nason, A New method for combining NLO QCD with shower Monte Carlo algorithms, JHEP **0411** (2004) p. 040, arXiv: hep-ph/0409146 [hep-ph].

[116] S. Frixione et al., Matching NLO QCD computations with Parton Shower simulations: the POWHEG method, JHEP **0711** (2007) p. 070, arXiv: 0709.2092 [hep-ph].

[117] S. Alioli et al., A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX, JHEP **1006** (2010) p. 043, arXiv: 1002.2581 [hep-ph].

[118] J. Pumplin et al., New generation of parton distributions with uncertainties from global QCD analysis, JHEP **07** (2002) p. 012, arXiv: hep-ph/0201195.

[119] H.-L. Lai et al., New parton distributions for collider physics, Phys.Rev. **D82** (2010) p. 074024, arXiv: 1007.2241 [hep-ph].

[120] P. Z. Skands, Tuning Monte Carlo Generators: The Perugia Tunes, Phys. Rev. **D82** (2010) p. 074018, arXiv: 1005.3457 [hep-ph].

[121] J. Alwall et al., A Standard format for Les Houches event files, Comput. Phys. Commun. **176** (2007) pp. 300–304, arXiv: hep-ph/0609017 [hep-ph].

[122] ATLAS Collaboration, Performance of jet substructure techniques for large-R jets in proton–proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector, JHEP **1309** (2013) p. 076, arXiv: 1306.4945 [hep-ex].

[123] J. M. Butterworth et al., Jet substructure as a new Higgs search channel at the LHC, Phys. Rev. Lett. **100** (2008) p. 242001, arXiv: 0802.2470 [hep-ph].

[124] ATLAS Collaboration, Measurements of normalized differential cross-sections for $t\bar{t}$ production in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector, Phys. Rev. D **90** (2014) p. 072004, arXiv: 1407.0371 [hep-ex].

[125] ATLAS Collaboration, Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector (2015), arXiv: 1503.05066 [hep-ex].

[126] ATLAS Collaboration, Jet mass and substructure of inclusive jets in $\sqrt{s} = 7$ TeV pp collisions with the ATLAS experiment, JHEP **1205** (2012) p. 128, arXiv: 1203.4606 [hep-ex].

[127] G. Cowan et al., Asymptotic formulae for likelihood-based tests of new physics, Eur. Phys. J. C **71**, 1554 (2011) p. 1554, arXiv: 1007.1727 [physics.data-an].

[128] A. L. Read, Presentation of search results: The CL(s) technique, J. Phys. **G28** (2002) pp. 2693–2704.

[129] K. Cranmer et al., "HistFactory: A tool for creating statistical models for use with RooFit and RooStats", tech. rep. CERN-OPEN-2012-016, New York U., 2012.

[130] R. Brun et al., ROOT: An object oriented data analysis framework, Nucl. Instrum. Meth. **A389** (1997) pp. 81–86.

Appendices

Appendix A

Multijet Fits

The full set of multijet template fits is displayed here (see Section 5.8) for the event preselection, split my lepton flavour and charge (as well as the summed lepton flavour charge fits, which were shown in the main text). The fits for the ggF WCR (TCR) are shown in Figure A.1 (Figure A.2). For VBF, the WCR (TCR) fits are in Figure A.3 (Figure A.4).

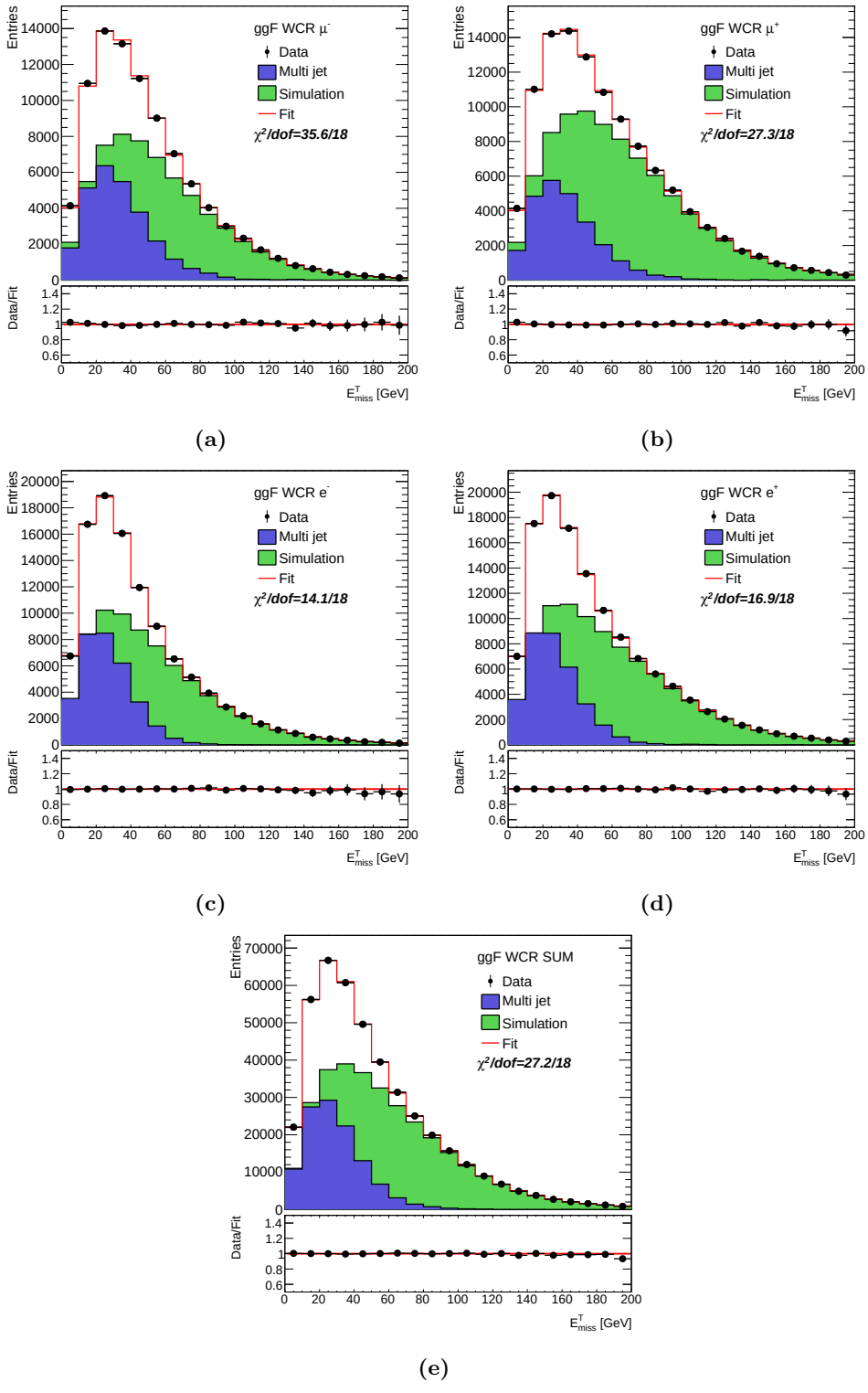


Figure A.1: Multijet background template fits for the ggF WCR after preselection. Fits are separated into separated charge/flavour channels, and also shown with all charge/flavour channels merged.

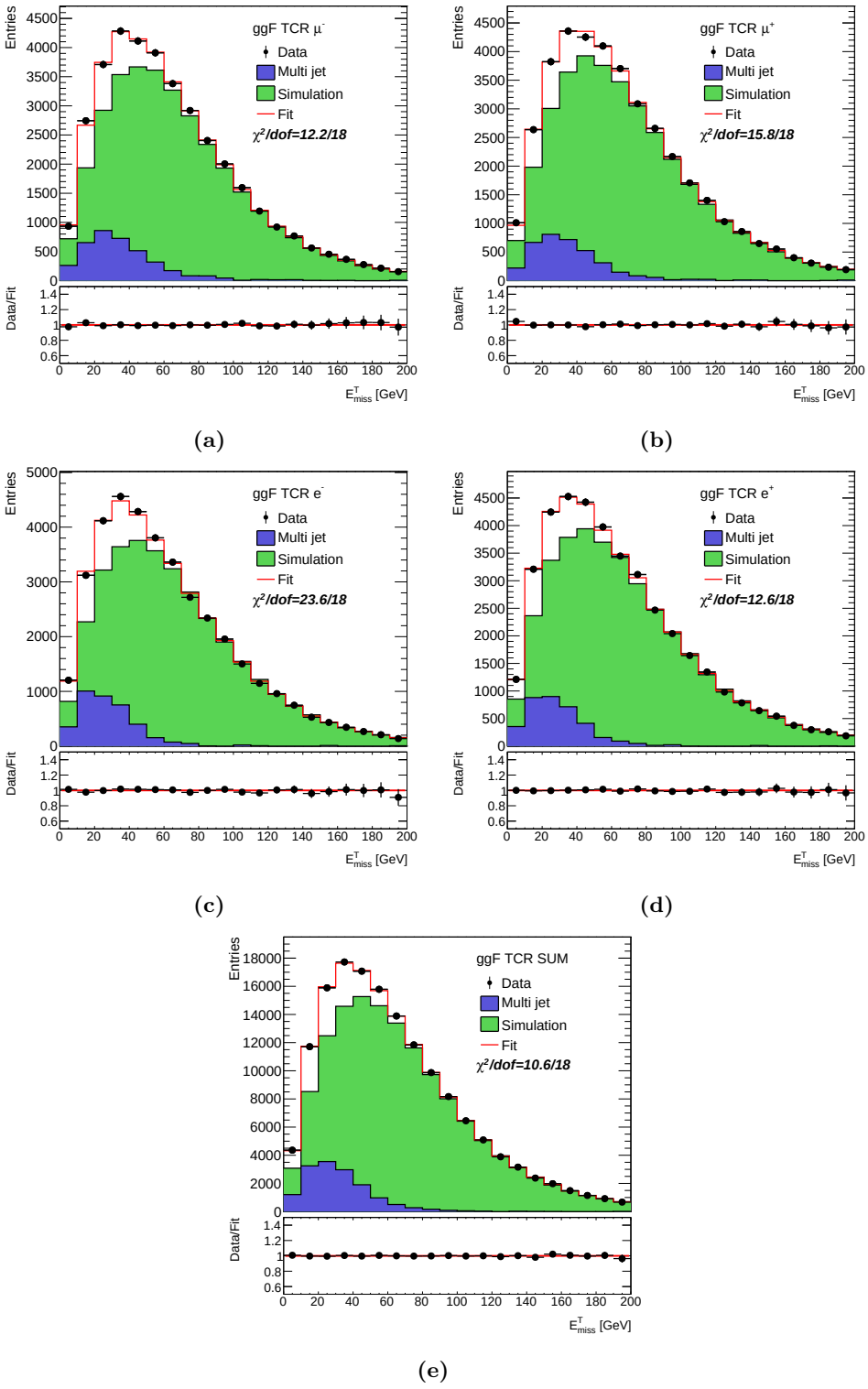


Figure A.2: Multijet background template fits for the ggF TCR after preselection. Fits are separated into separated charge/flavour channels, and also shown with all charge/flavour channels merged.

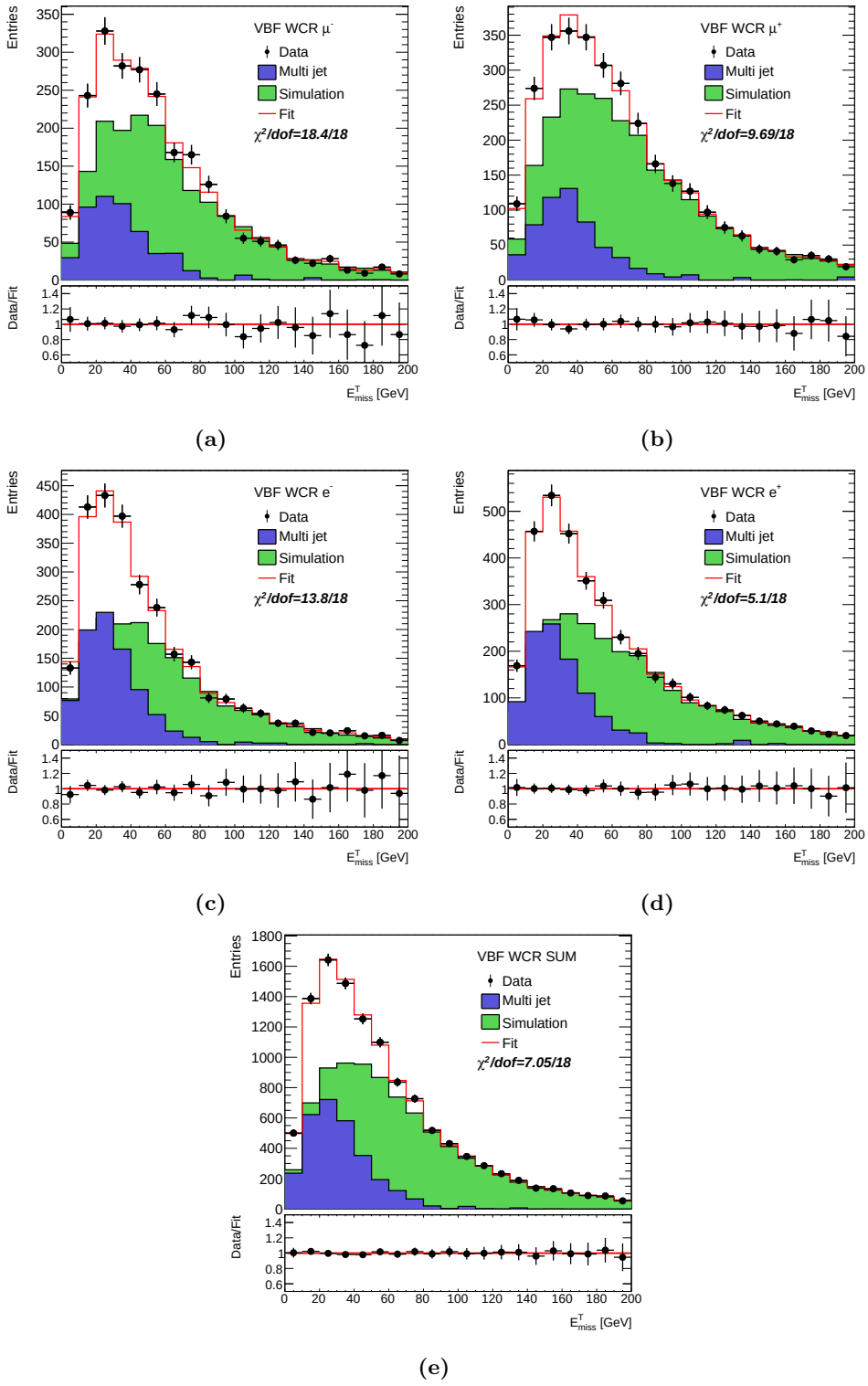


Figure A.3: Multijet background template fits for the VBF WCR after preselection. Fits are separated into separated charge/flavour channels, and also shown with all charge/flavour channels merged.

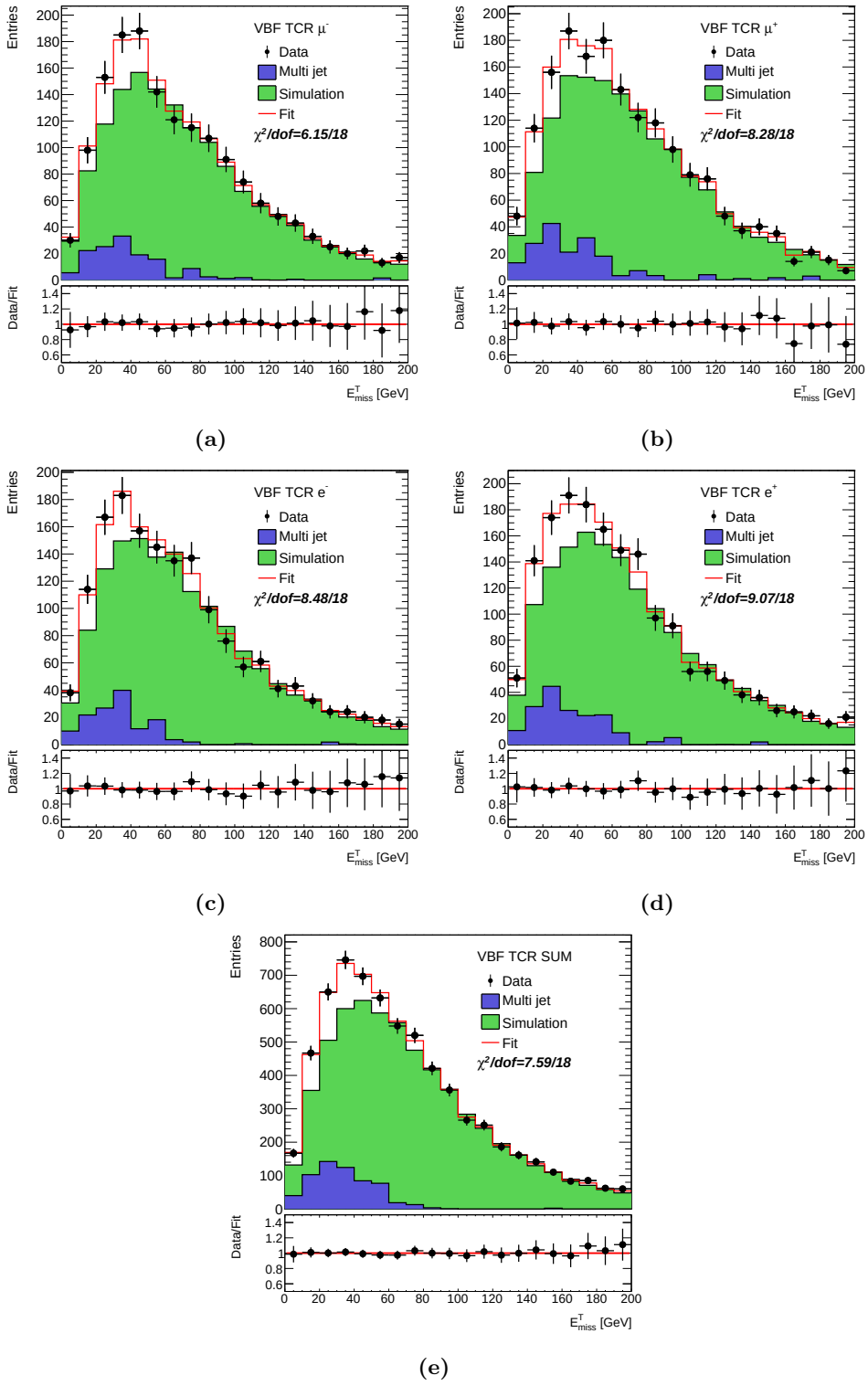
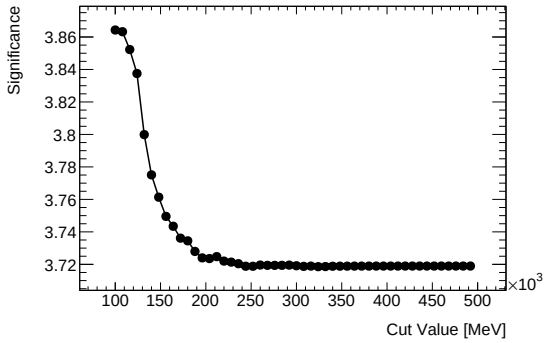


Figure A.4: Multijet background template fits for the VBF TCR after preselection. Fits are separated into separated charge/flavour channels, and also shown with all charge/flavour channels merged.

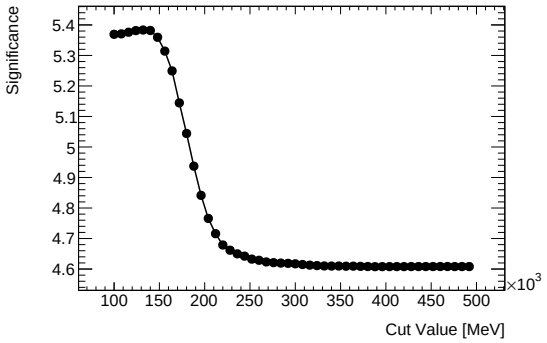
Appendix B

Mass Dependant Plots

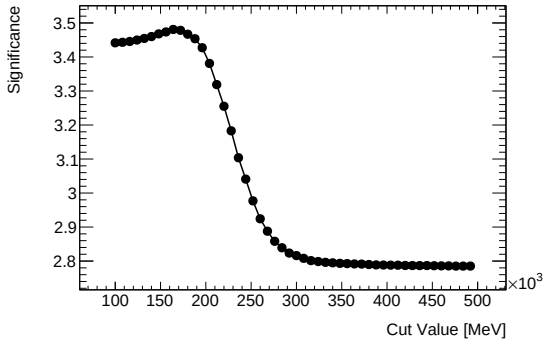
The full set of mass optimization plots used to determine the mass dependant selection (introduced in Section 5.9), including full set of cutscans for all variables and mass points considered. The ggF cutscan plots are in Figures B.1–B.7, and the VBF cutscan plots are in Figures B.8–B.10.



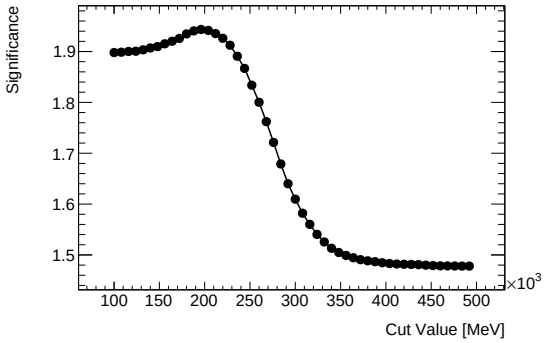
(a) $m_H = 300$ GeV



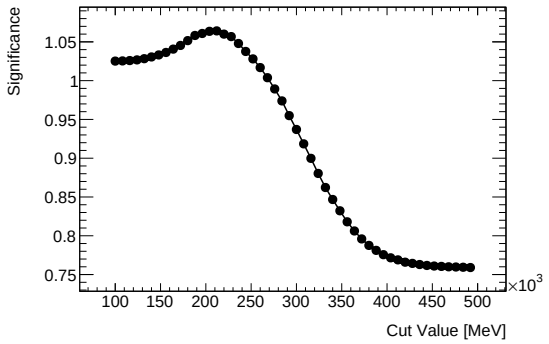
(b) $m_H = 400$ GeV



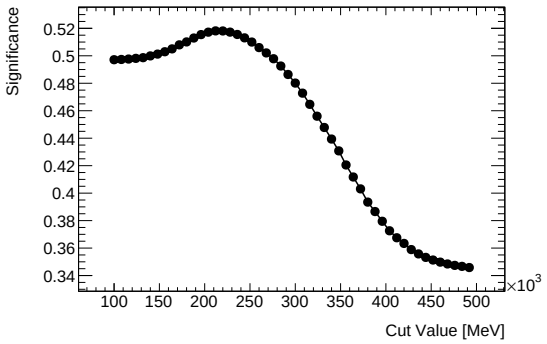
(c) $m_H = 500$ GeV



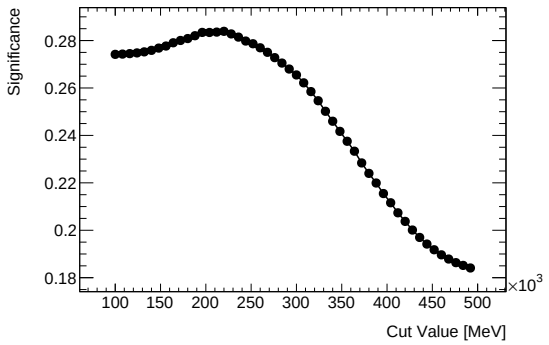
(d) $m_H = 600$ GeV



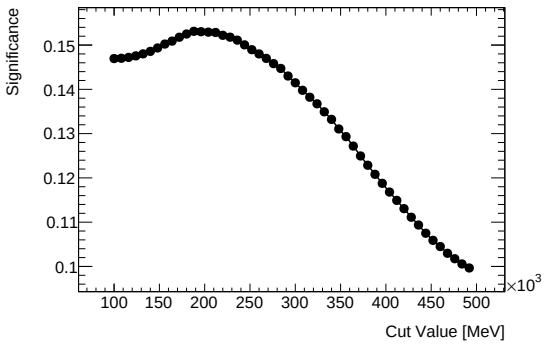
(e) $m_H = 700$ GeV



(f) $m_H = 800$ GeV

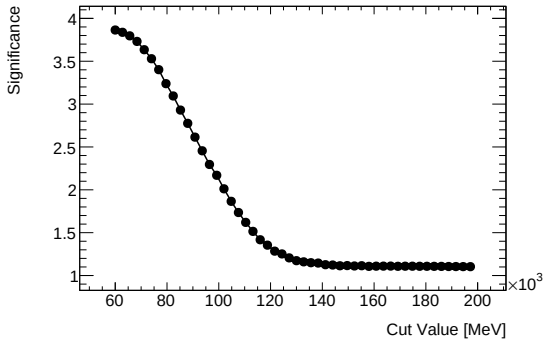


(g) $m_H = 900$ GeV

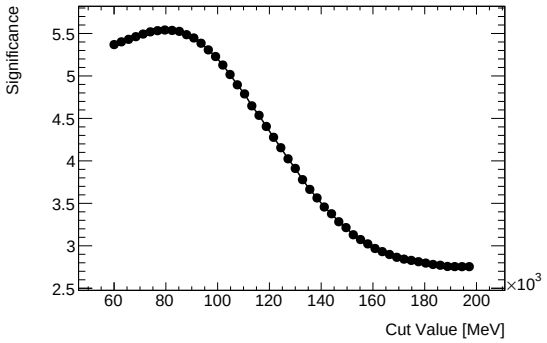


(h) $m_H = 1000$ GeV

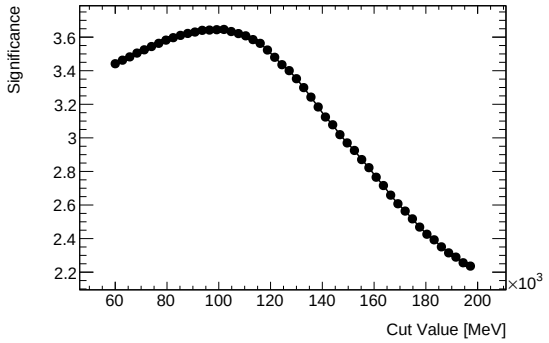
Figure B.1: Mass dependant ggF cut scan for $p_{\mathrm{T},J}$



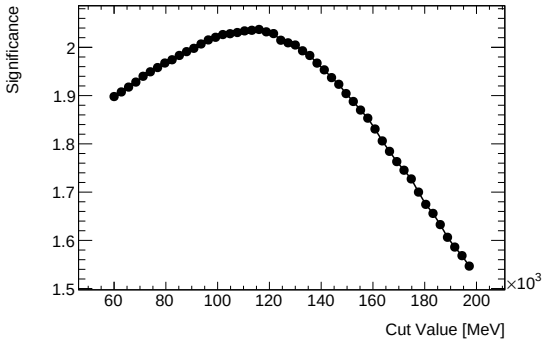
(a) $m_H = 300$ GeV



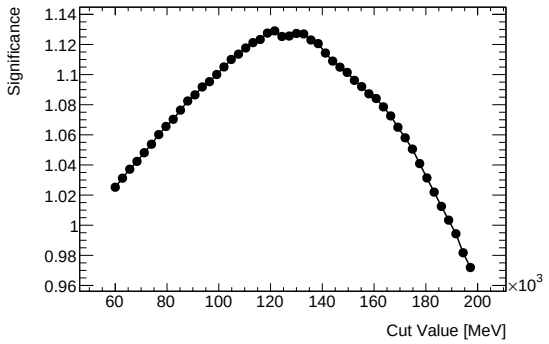
(b) $m_H = 400$ GeV



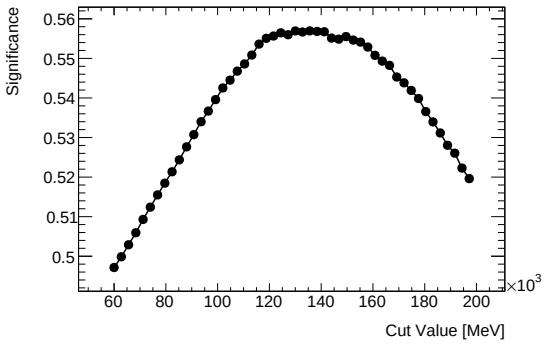
(c) $m_H = 500$ GeV



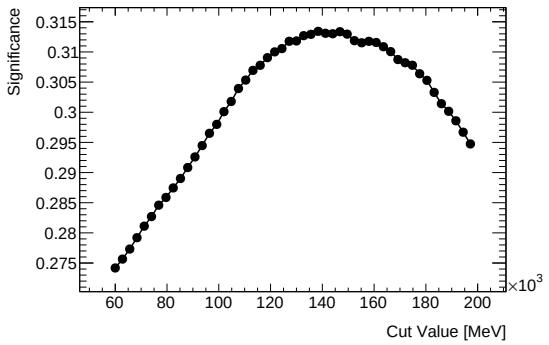
(d) $m_H = 600$ GeV



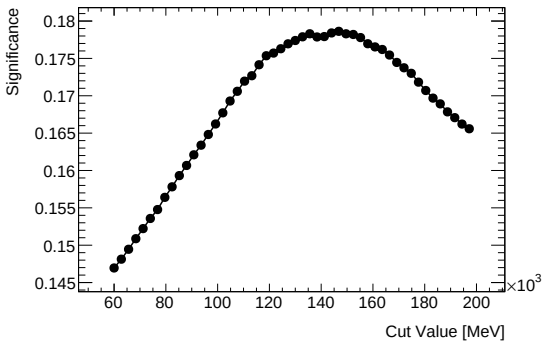
(e) $m_H = 700$ GeV



(f) $m_H = 800$ GeV

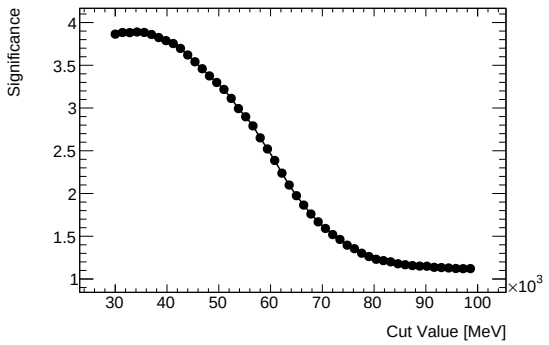


(g) $m_H = 900$ GeV

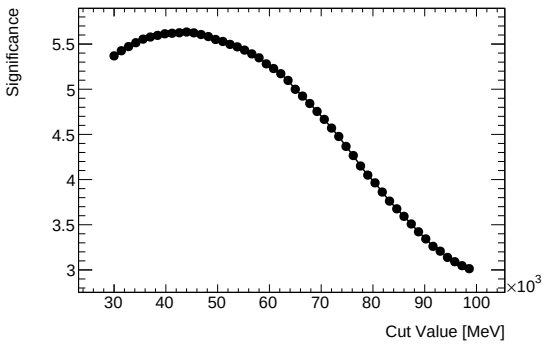


(h) $m_H = 1000$ GeV

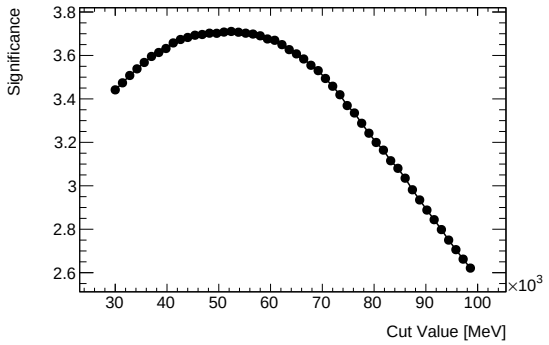
Figure B.2: Mass dependant ggF cut scan for $p_{\mathrm{T},j1}$



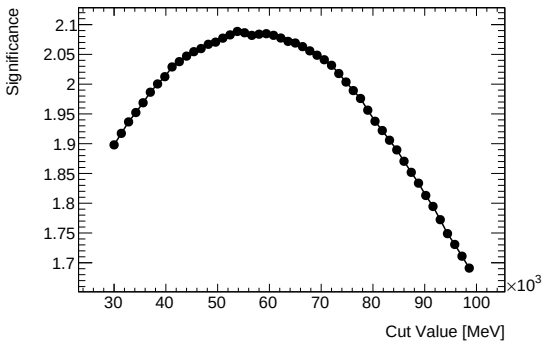
(a) $m_H = 300$ GeV



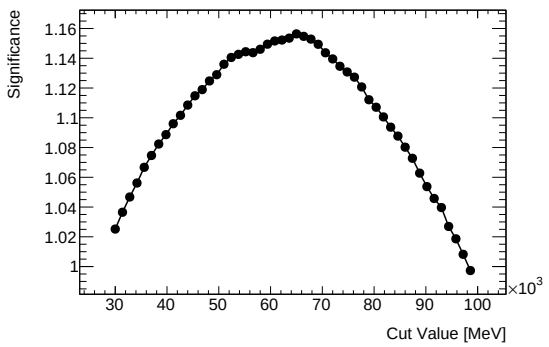
(b) $m_H = 400$ GeV



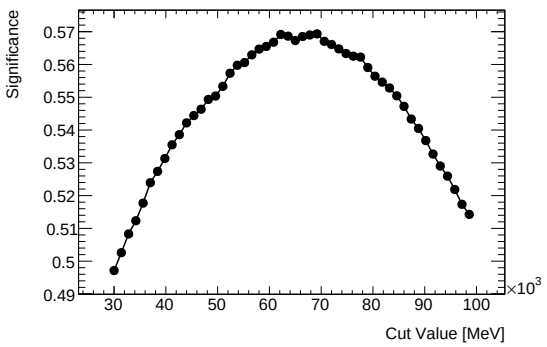
(c) $m_H = 500$ GeV



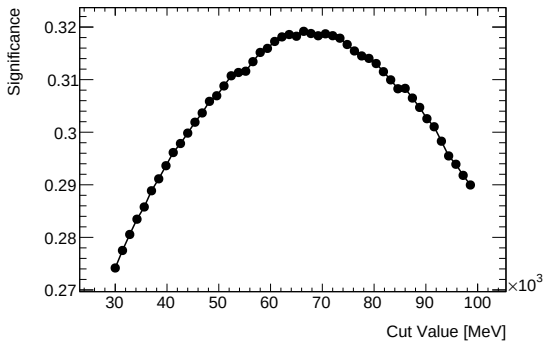
(d) $m_H = 600$ GeV



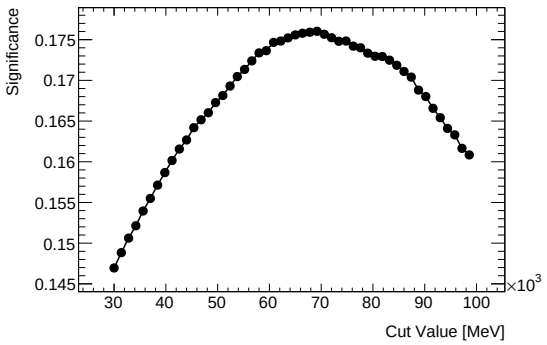
(e) $m_H = 700$ GeV



(f) $m_H = 800$ GeV

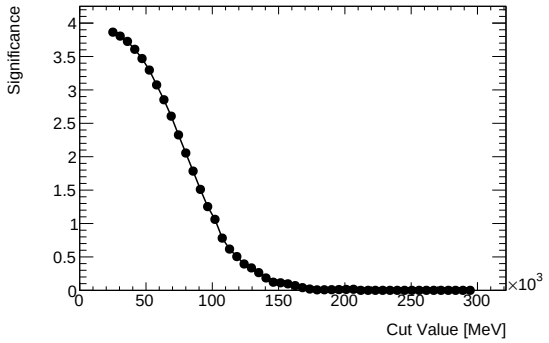


(g) $m_H = 900$ GeV

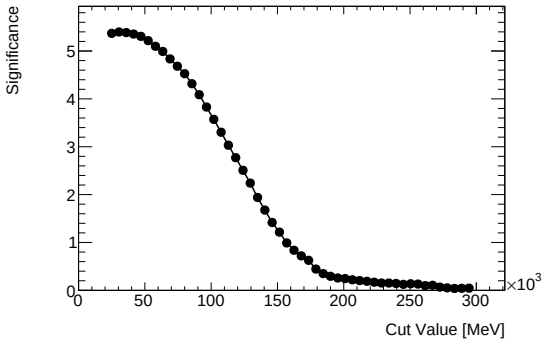


(h) $m_H = 1000$ GeV

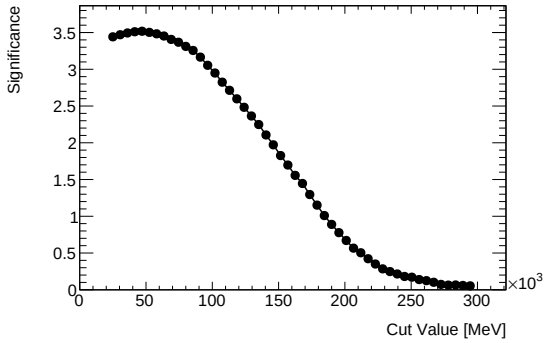
Figure B.3: Mass dependant ggF cut scan for $p_{\mathrm{T},j2}$



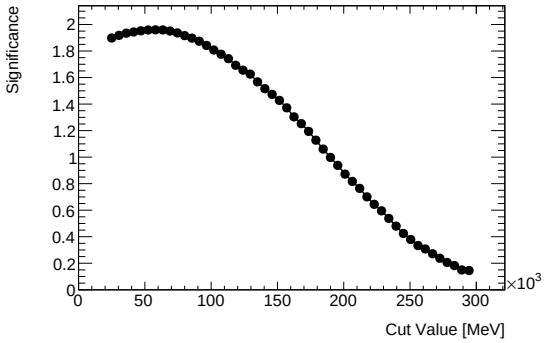
(a) $m_H = 300$ GeV



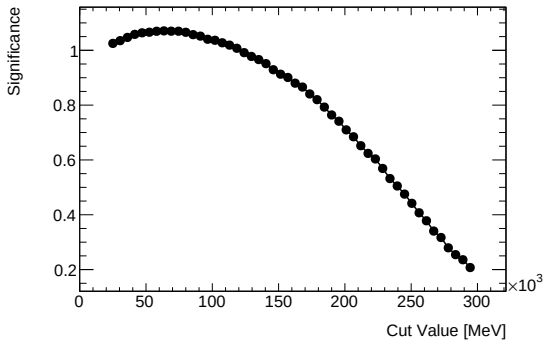
(b) $m_H = 400$ GeV



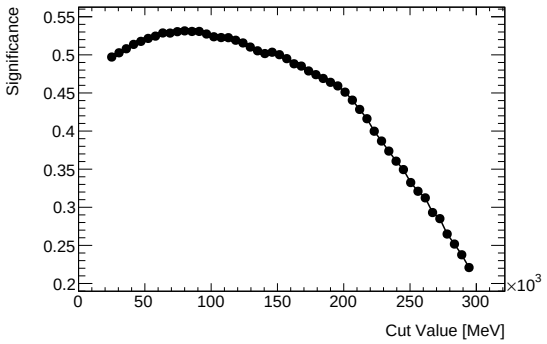
(c) $m_H = 500$ GeV



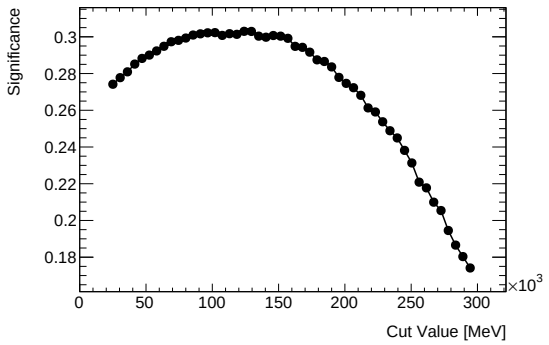
(d) $m_H = 600$ GeV



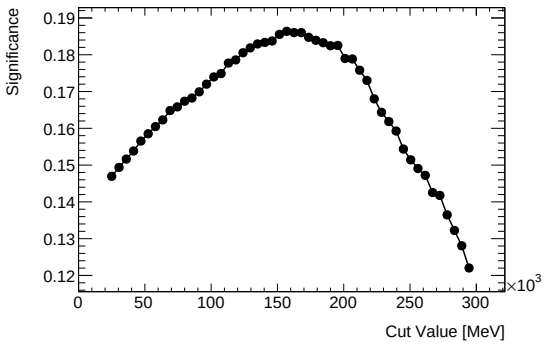
(e) $m_H = 700$ GeV



(f) $m_H = 800$ GeV

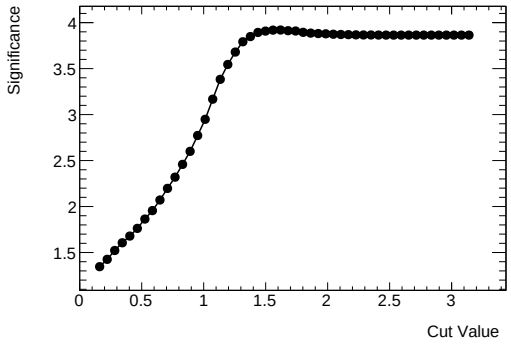


(g) $m_H = 900$ GeV

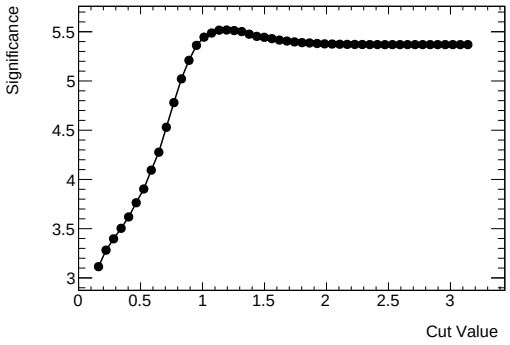


(h) $m_H = 1000$ GeV

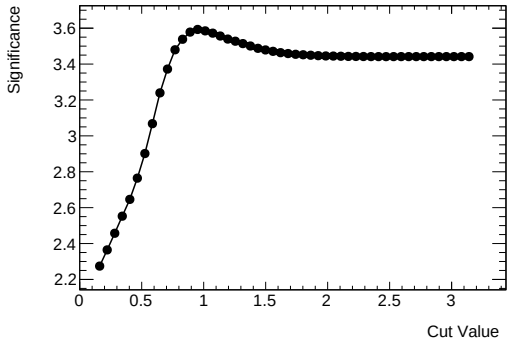
Figure B.4: Mass dependant ggF cut scan for $p_{\mathrm{T},\ell}$



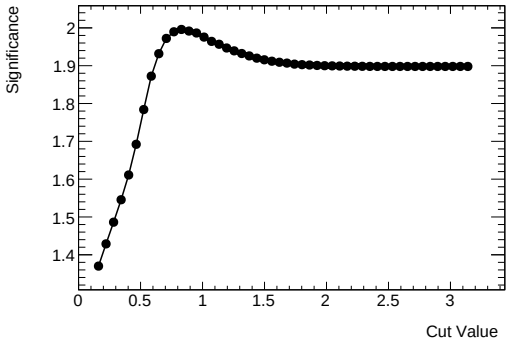
(a) $m_H = 300$ GeV



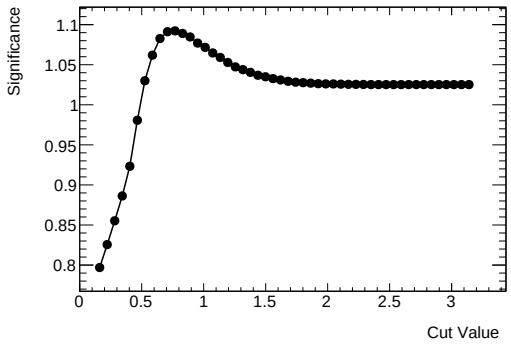
(b) $m_H = 400$ GeV



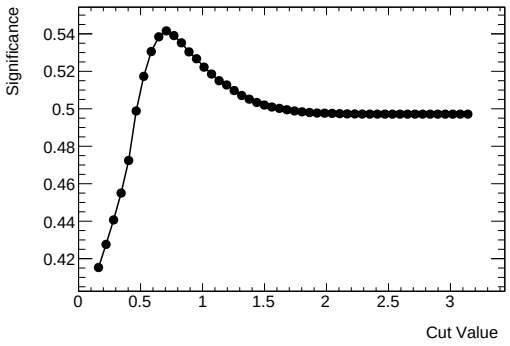
(c) $m_H = 500$ GeV



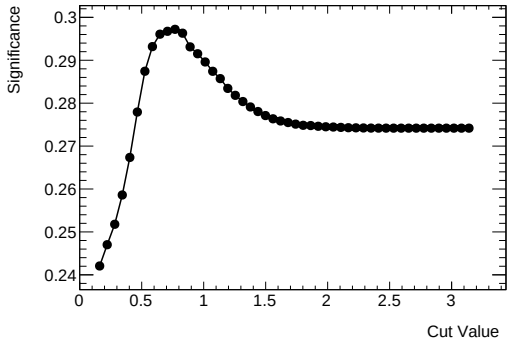
(d) $m_H = 600$ GeV



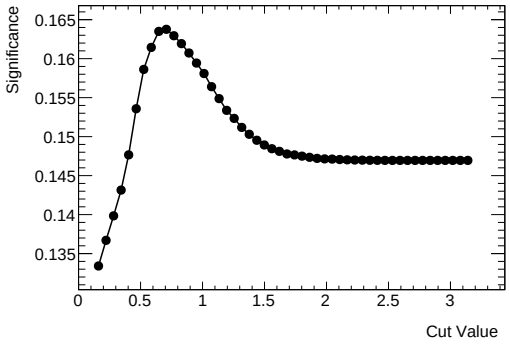
(e) $m_H = 700$ GeV



(f) $m_H = 800$ GeV

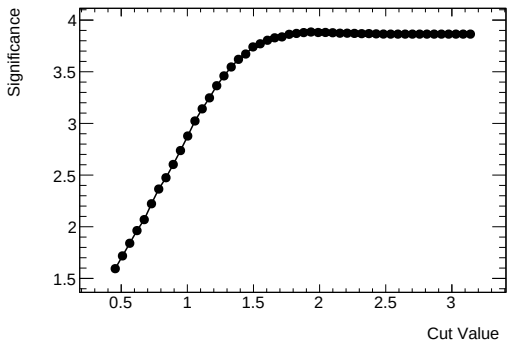


(g) $m_H = 900$ GeV

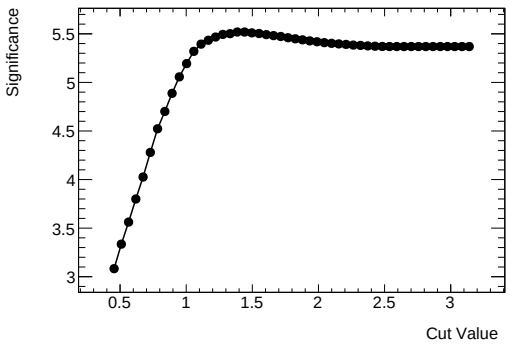


(h) $m_H = 1000$ GeV

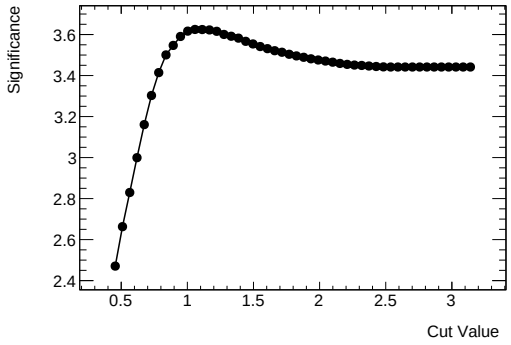
Figure B.5: Mass dependant ggF cut scan for $|\Delta\phi_{jj}|$



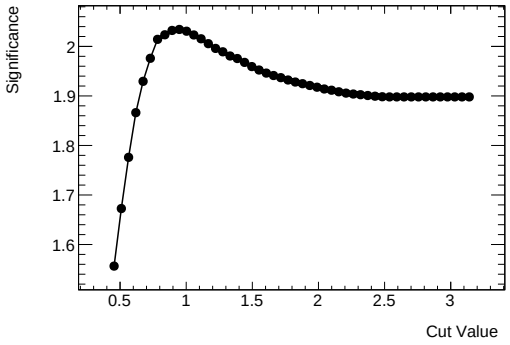
(a) $m_H = 300$ GeV



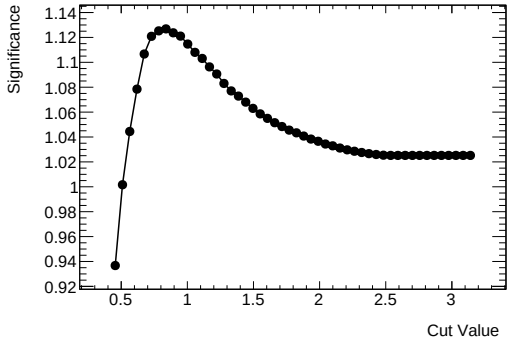
(b) $m_H = 400$ GeV



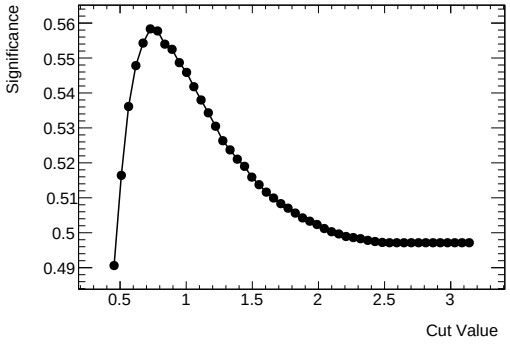
(c) $m_H = 500$ GeV



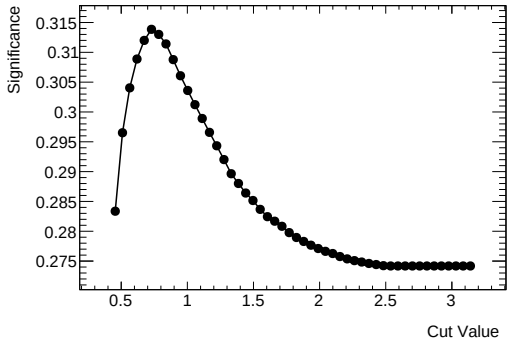
(d) $m_H = 600$ GeV



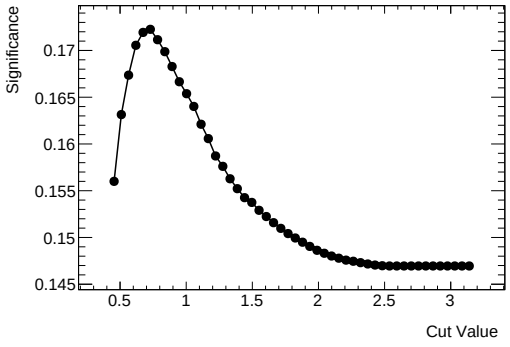
(e) $m_H = 700$ GeV



(f) $m_H = 800$ GeV

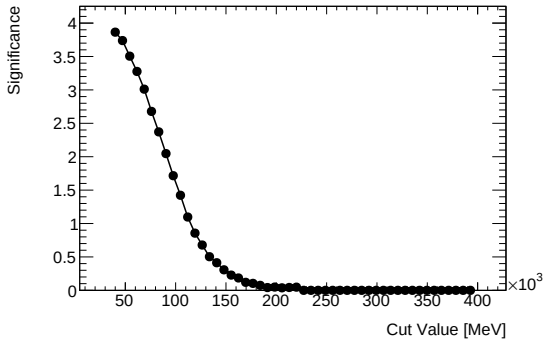


(g) $m_H = 900$ GeV

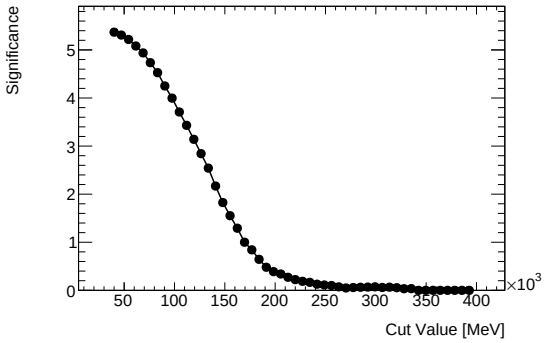


(h) $m_H = 1000$ GeV

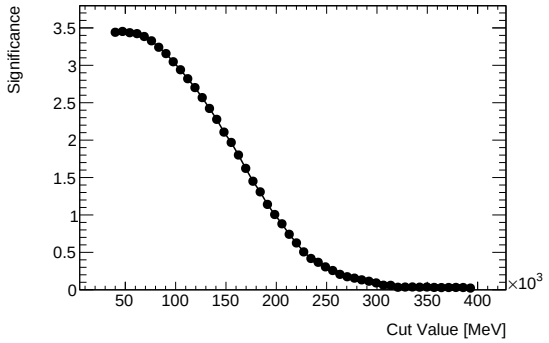
Figure B.6: Mass dependant ggF cut scan for $|\Delta\phi_{\ell\nu}|$



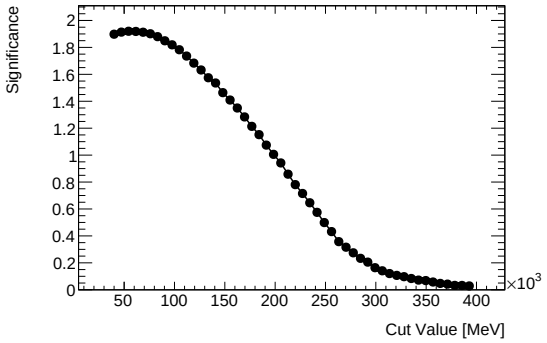
(a) $m_H = 300$ GeV



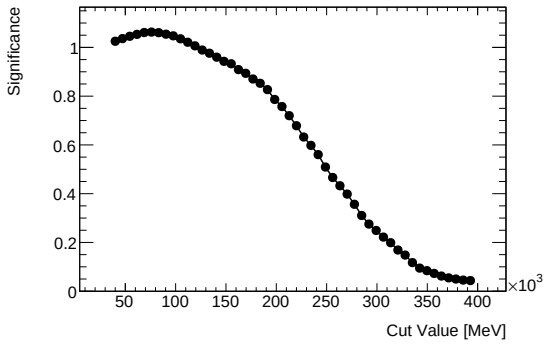
(b) $m_H = 400$ GeV



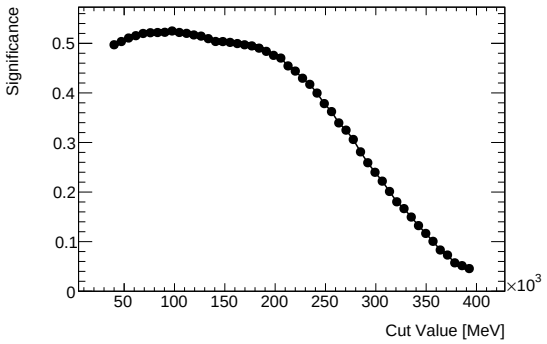
(c) $m_H = 500$ GeV



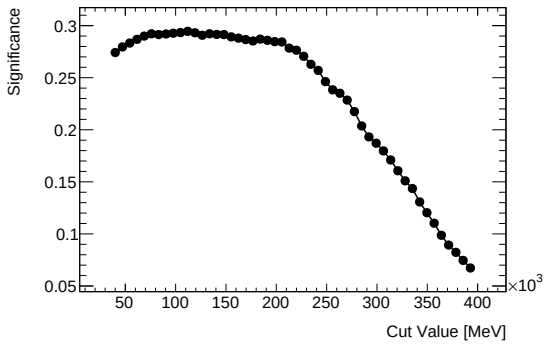
(d) $m_H = 600$ GeV



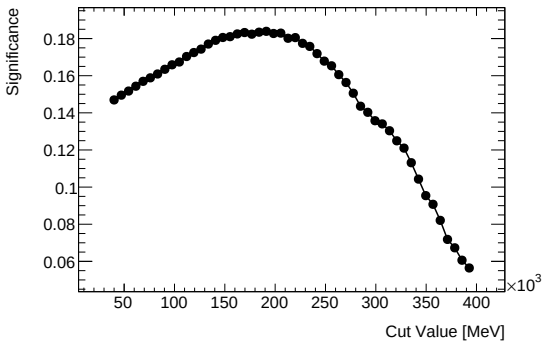
(e) $m_H = 700$ GeV



(f) $m_H = 800$ GeV

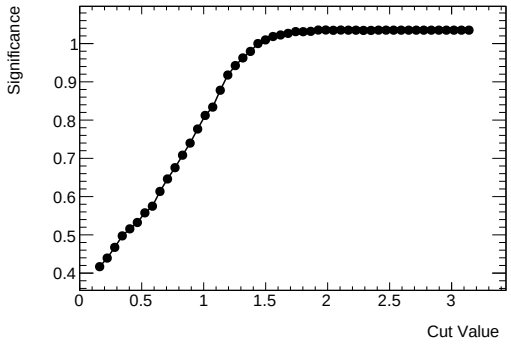


(g) $m_H = 900$ GeV

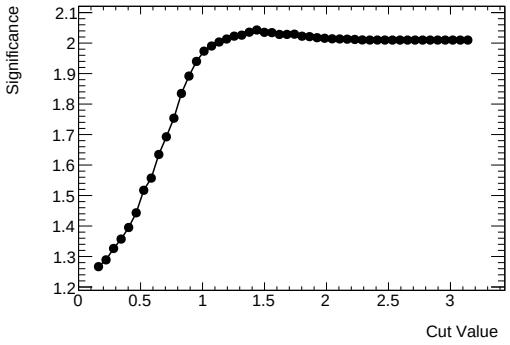


(h) $m_H = 1000$ GeV

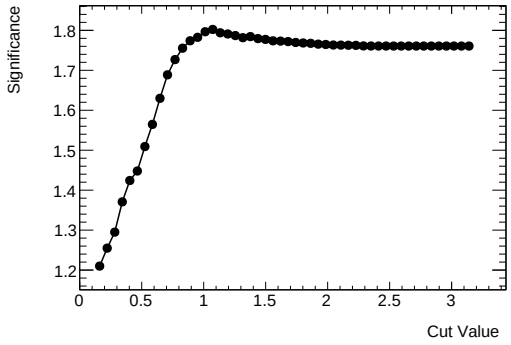
Figure B.7: Mass dependant ggF cut scan for $E_{\mathrm{T}}^{\mathrm{miss}}$



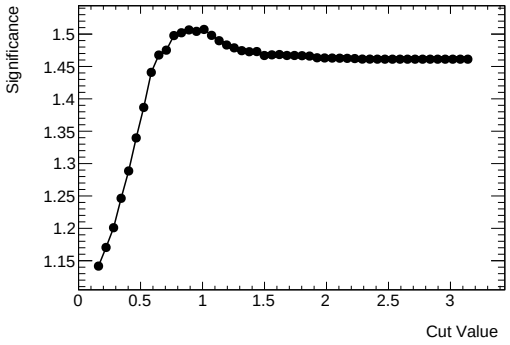
(a) $m_H = 300$ GeV



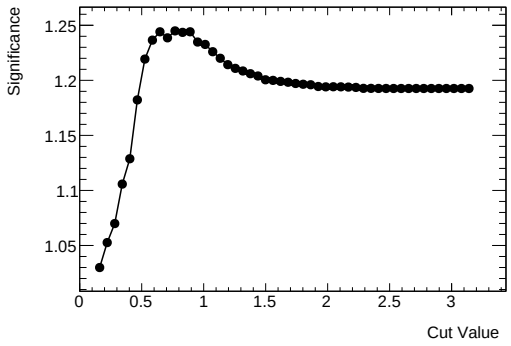
(b) $m_H = 400$ GeV



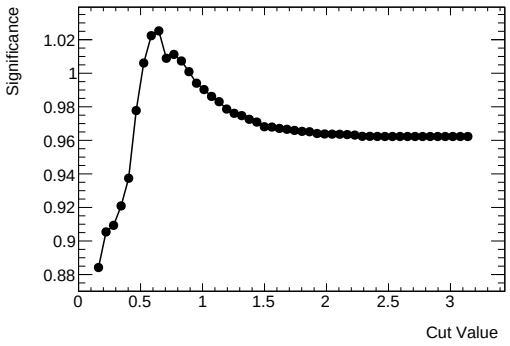
(c) $m_H = 500$ GeV



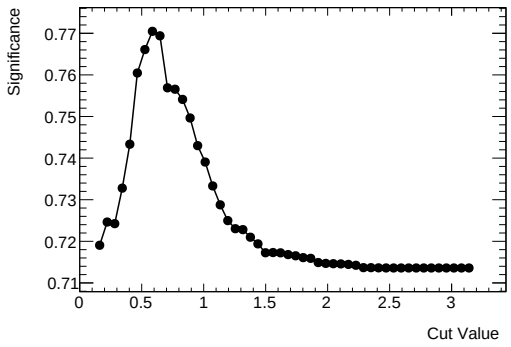
(d) $m_H = 600$ GeV



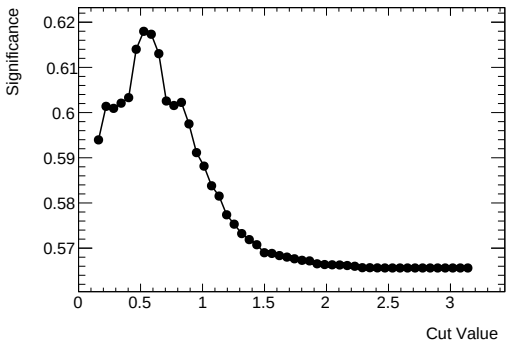
(e) $m_H = 700$ GeV



(f) $m_H = 800$ GeV

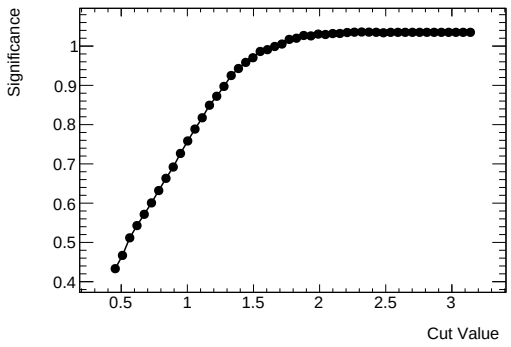


(g) $m_H = 900$ GeV

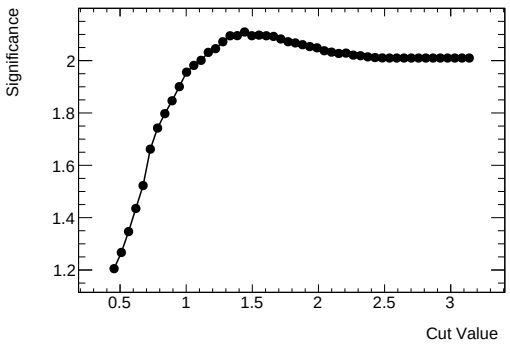


(h) $m_H = 1000$ GeV

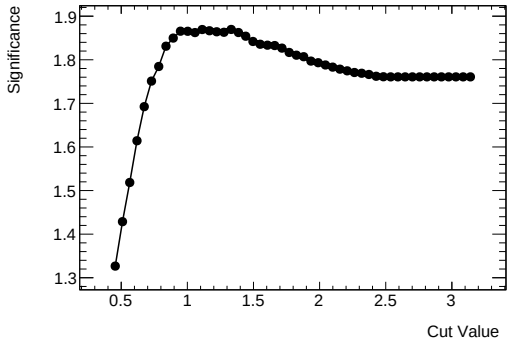
Figure B.8: Mass dependant VBF cut scan for $|\Delta\phi_{jj}|$



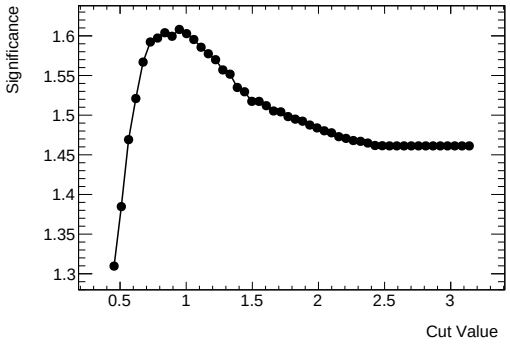
(a) $m_H = 300$ GeV



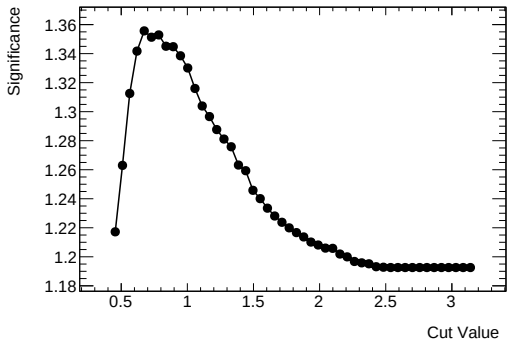
(b) $m_H = 400$ GeV



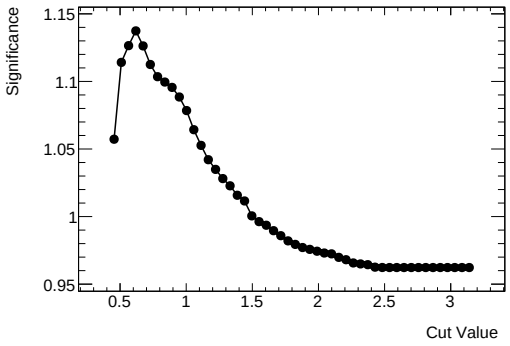
(c) $m_H = 500$ GeV



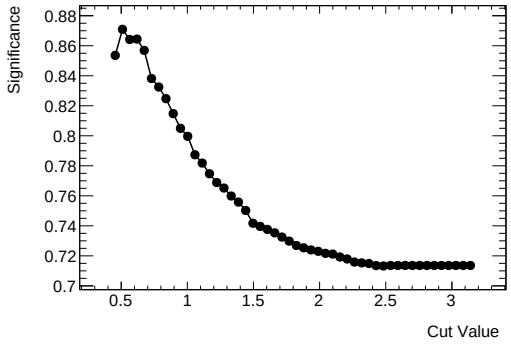
(d) $m_H = 600$ GeV



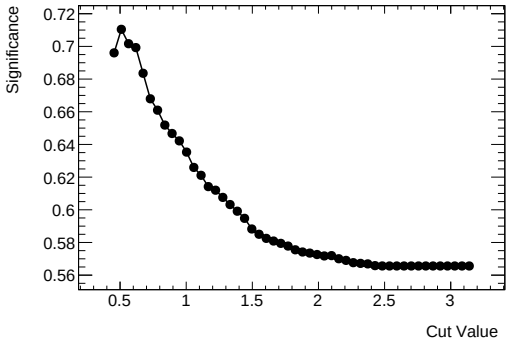
(e) $m_H = 700$ GeV



(f) $m_H = 800$ GeV

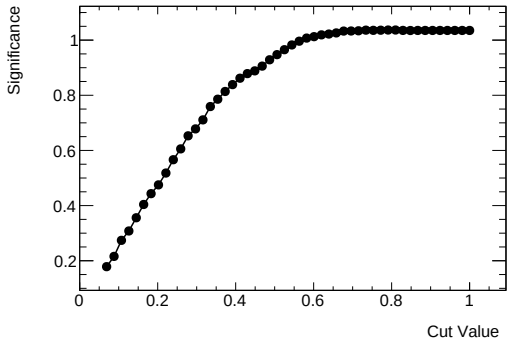


(g) $m_H = 900$ GeV

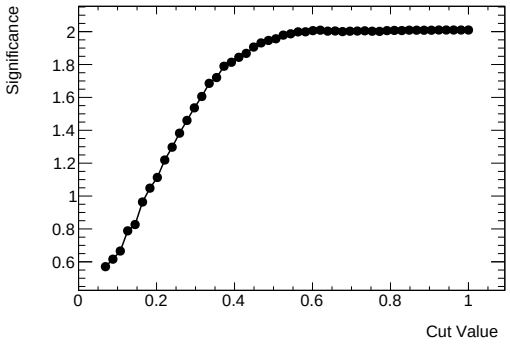


(h) $m_H = 1000$ GeV

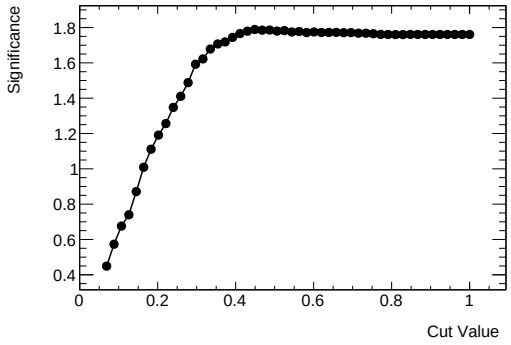
Figure B.9: Mass dependant VBF cut scan for $|\Delta\phi_{\ell\nu}|$



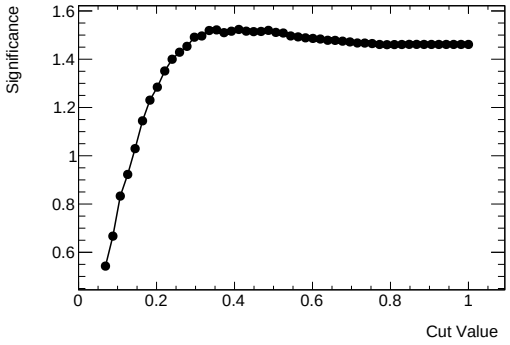
(a) $m_H = 300$ GeV



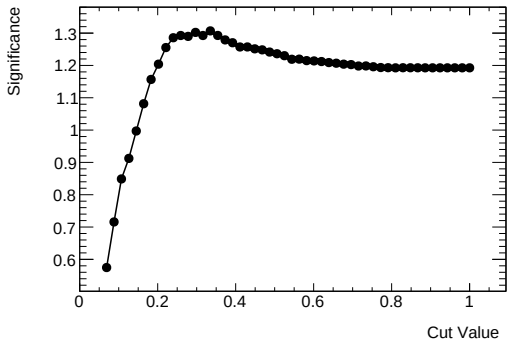
(b) $m_H = 400$ GeV



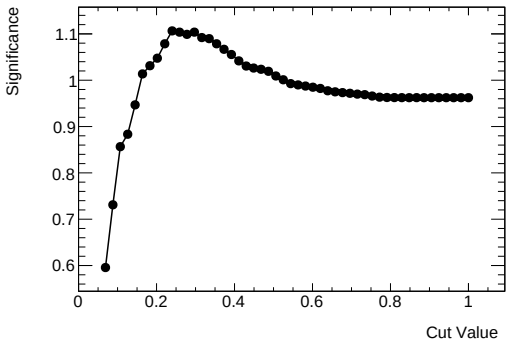
(c) $m_H = 500$ GeV



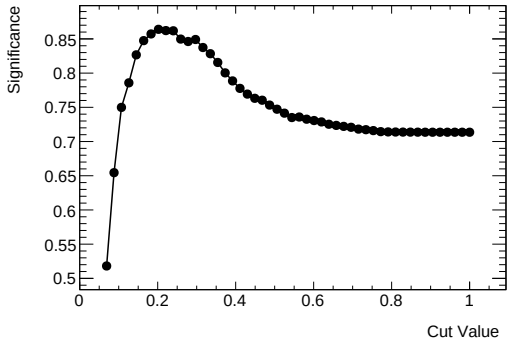
(d) $m_H = 600$ GeV



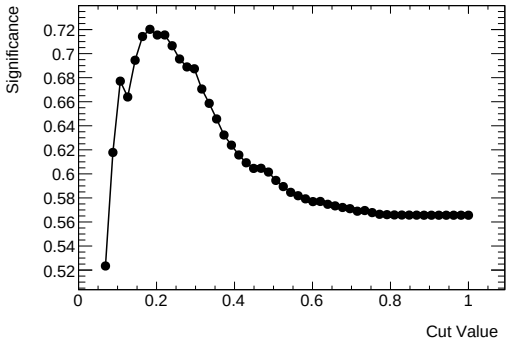
(e) $m_H = 700$ GeV



(f) $m_H = 800$ GeV



(g) $m_H = 900$ GeV



(h) $m_H = 1000$ GeV

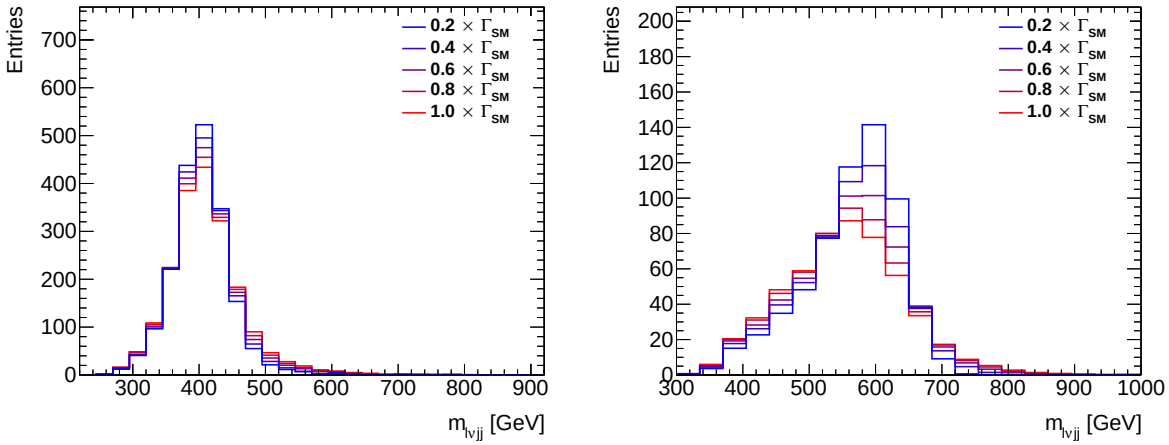
Figure B.10: Mass dependant VBF cut scan for p_{T}-balance.

Appendix C

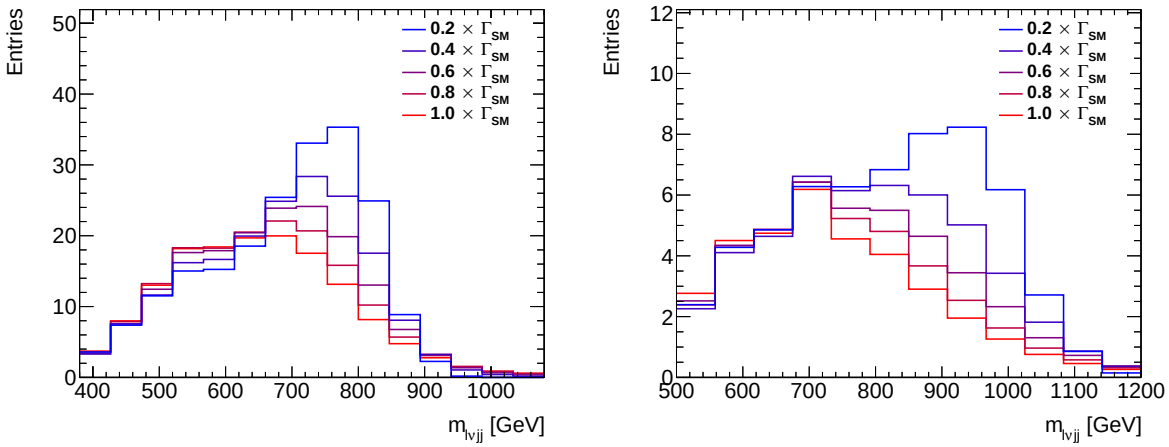
EWS Signal Shape Inputs

Presented here are the limit setting signal inputs for the EWS model. All figures are shown with relevant mass dependant selection already applied. Figures C.1 and C.2 show the inputs for ggF. Figure C.1 has the distributions for all the different widths normalized with respect to each other (set to σ_{SM} expectation), while Figure C.2 shows the distributions scaled with respect to each other, obeying the relations in Equations 2.24–2.26 (with $\text{BR}_{\text{new}} = 0$). The analogous figures for the VBF process are found in Figures C.3 and C.4.

Note that the lineshapes between the widths are not identical, which is a result of the interference weights. While this may seem counterintuitive for the case of the VBF samples (which have the same interference spectrum, independant of Γ_H, see Section 5.4.3), remember that while the inteference spectrum is identical, the impact of the interference spectrum (which is a function of m_{WW}) will more severly impact the shape of a wider Higgs.

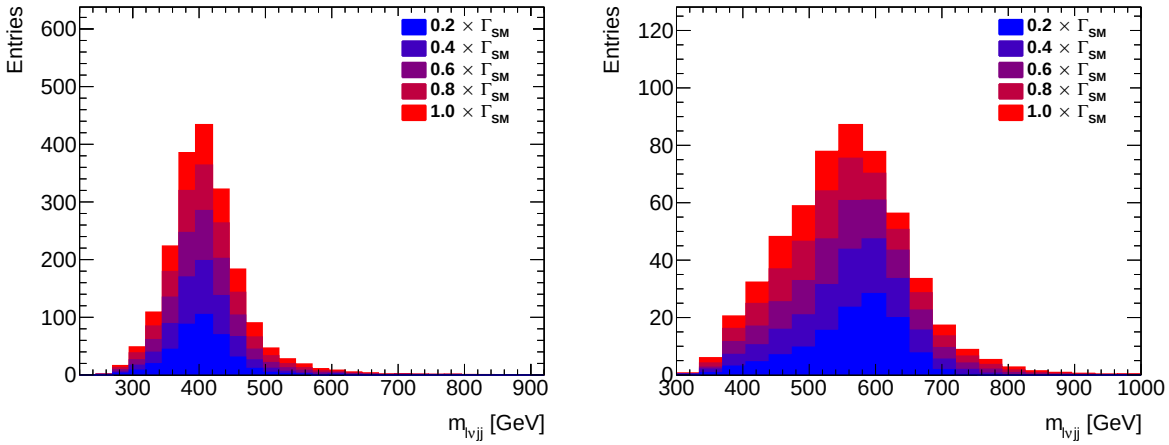


(a) $m_H = 400$ GeV **(b)** $m_H = 600$ GeV

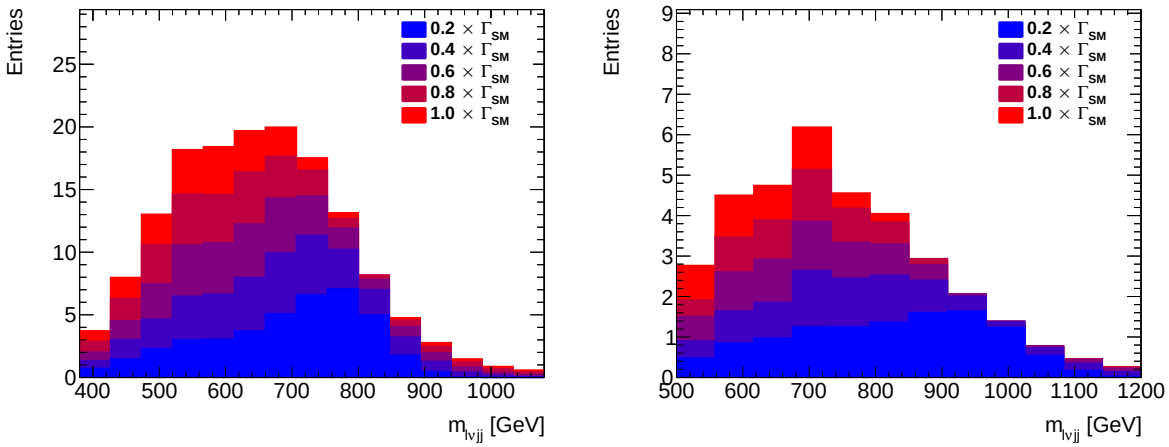


(c) $m_H = 800$ GeV **(d)** $m_H = 1000$ GeV

Figure C.1: Comparison of signal shapes (charge/flavour summed) for the different Higgs widths considered as part of the EWS interpretation, in the ggF channel. These histograms are input into the final limit setting, with the appropriate interference weights and mass dependant selection applied. For all curves, σ_{SM} is assumed.

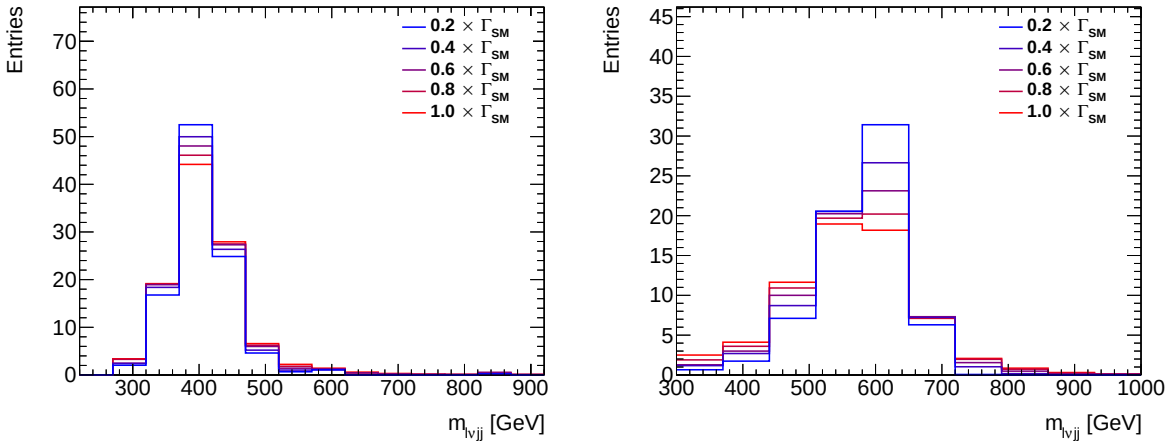


(a) $m_H = 400$ GeV **(b)** $m_H = 600$ GeV

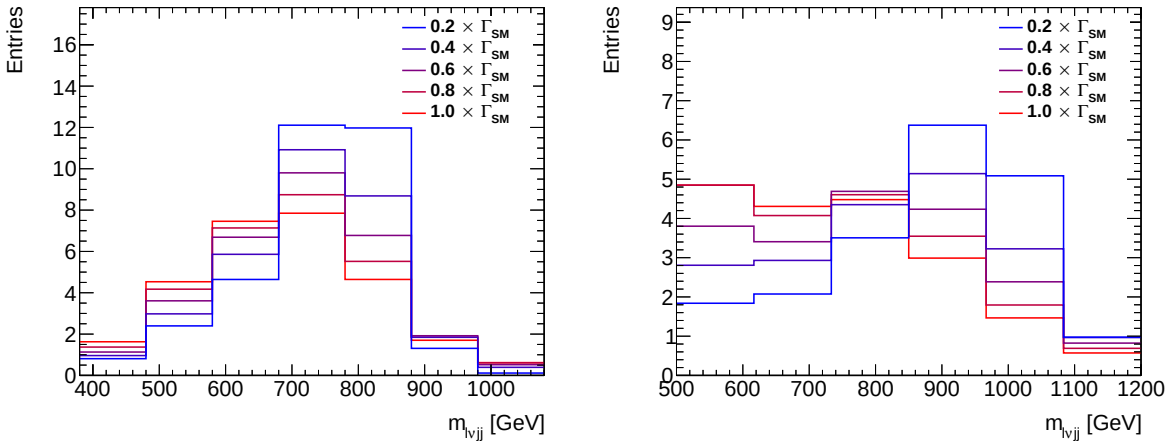


(c) $m_H = 800$ GeV **(d)** $m_H = 1000$ GeV

Figure C.2: Comparison of signal shapes and normalizations (charge/flavour summed) for the different Higgs widths considered as part of the EWS interpretation, in the ggF channel. These histograms are input into the final limit setting, with the appropriate interference weights and mass dependant selection applied. Relative normalizations assume $\mathrm{BR}_{\mathrm{new}} = 0$.

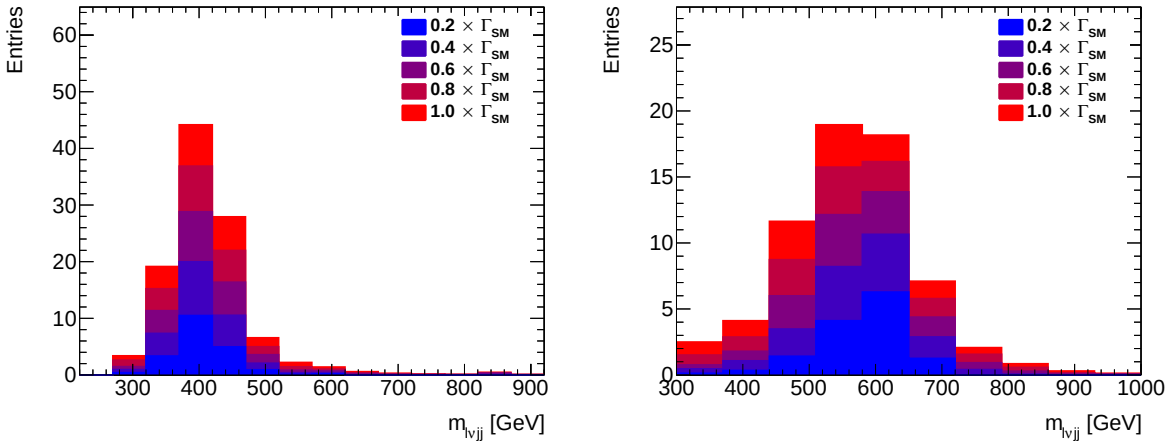


(a) $m_H = 400$ GeV **(b)** $m_H = 600$ GeV

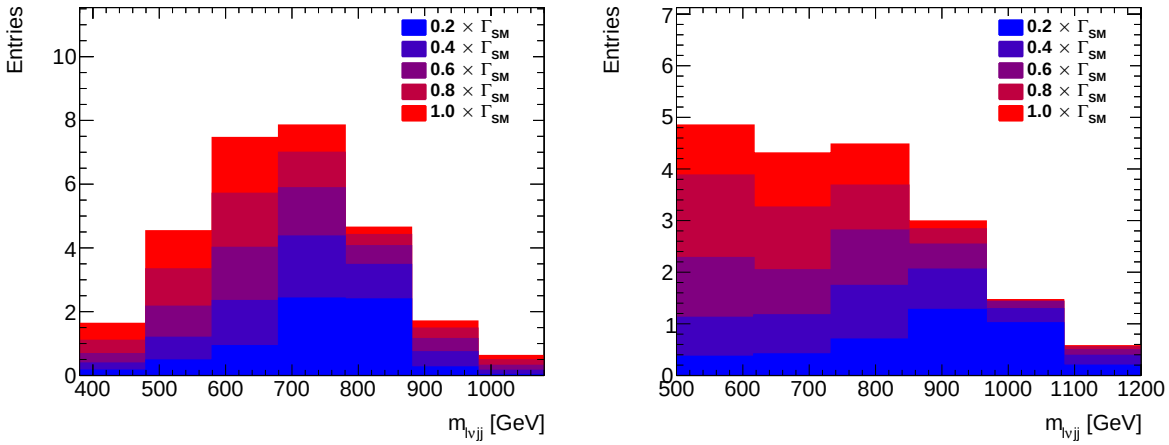


(c) $m_H = 800$ GeV **(d)** $m_H = 1000$ GeV

Figure C.3: Comparison of signal shapes (charge/flavour summed) for the different Higgs widths considered as part of the EWS interpretation, in the VBF channel. These histograms are input into the final limit setting, with the appropriate interference weights and mass dependant selection applied. For all curves, σ_{SM} is assumed.



(a) $m_H = 400$ GeV **(b)** $m_H = 600$ GeV



(c) $m_H = 800$ GeV **(d)** $m_H = 1000$ GeV

Figure C.4: Comparison of signal shapes and normalizations (charge/flavour summed) for the different Higgs widths considered as part of the EWS interpretation, in the VBF channel. These histograms are input into the final limit setting, with the appropriate interference weights and mass dependant selection applied. Relative normalizations assume $\mathrm{BR}_{\mathrm{new}} = 0$.

Appendix D

Contributions to the ATLAS Collaboration

During my time on ATLAS I have spent time performing service and technical work, in addition to my role in the $H \to WW \to \ell\nu jj$ analysis documented in this thesis.

My technical work started in the ATLAS Liquid Argon Calorimeter group with an analysis of the "bad" FCal channels. I studied the FCal calibration waveforms, attempting to diagnose faulty electronics in the readout chain based on deformations in the observed pulse shape. For some of these faults, it was unknown what impact there would be on the reconstructed energy (for others it was known, but needed to be verified) and so I then studied the relative energy response of these "bad" channels when compared to the response of neighbouring channels, and channels in a similar radial position. These studies were both meant to evaluate an energy correction that could be applied to the channel, and also discover channels which did not have the expected energy response, even though there was no fault present in the calibration waveform. This work then moved on to using a similar technique to study the positioning of the FCal (and therefore the entire LAr end-cap) with respect to the beampipe after long shutdowns.

My work with the FCal allowed me to contribute to the ATLAS Luminosity Taskforce where I was responsible for evaluating the long-term stability of the ATLAS luminosity algorithms, and evaluating the luminosity as determined by the FCal HV currents. Initially, my work was meant to focus on developing a new calibration of the FCal to account for observed non-linear behaviour of the original calibration. However, this work was quickly replaced with understanding the source of the "drift" of the FCal's response with respect to the main luminosity algorithm in ATLAS: `BCMV_EventOR`. This involved a detailed study of a variety of topics, such as the beam-spot drift, beam backgrounds, material activation, drifting calibrations, and faulty HV modules. It also lead to the derivation of a new luminosity algorithm, which I helped develop, using the pile-up noise in the FCal as a measure of the instantaneous luminosity.

I was also heavily involved in the operations of the Liquid Argon Calorimeter. In 2011, I started as one of the hardware on-call experts, responsible for reacting to any hardware related faults, which was most often a trip or other fault in the high-voltage system. After the 2011 *pp* run, I was asked to serve as one of the ATLAS Run Coordinators for the 2011 and 2013 heavy ion runs, and the 2012 *pp* run. My last contribution to the operations of the LAr Calorimeters was as the Calo Shift Coordinator. I was responsible for coordinating the merge of the LAr, Forward and Tile Calorimeter desks in the ATLAS Control Room, managing shift booking, shifter training, and managing/recruiting the team of

Super Shifters. The Super Shifters were responsible for helping to train new shifters, managing the shift environment, and acted as a liaison between the shift crew and operations crew at the Calo desk.

The analysis presented in this thesis is also in some ways a contribution to the ATLAS Collaboration. However, much of the technical work was not relevant to the thesis, and was left out of this document. One of my largest contributions to the $H \to WW \to \ell\nu jj$ analysis was as the software developer of the group's common analysis framework (CAF), and the production coordinator responsible for the creation of the group's datasets and managing the dataset distribution on the remote storage sites. This code formed the skeleton around which the analysis presented in this thesis was built upon. The other large contribution was evaluating the interference between the resonant $H \to WW$ and continuum SM WW processes in the VBF mode, as well as determining the uncertainty. This work was used by both HSG3 high-mass WW analyses. This work also led to me working with ATLAS' HSG2, where I contributed to the off-shell coupling analysis by determining the effect of a variable Higgs width on the interference spectrum using the MC program Phantom.

Selected Publications

- ATLAS Collaboration, Search for a high-mass Higgs boson decaying to a W boson pair in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector (2015), arXiv: `1509.00389 [hep-ex]`
- ATLAS Collaboration, Constraints on the off-shell Higgs boson signal strength in the high-mass ZZ and WW final states with the ATLAS detector, Eur. Phys. J. **C75**.7 (2015) p. 335, arXiv: `1503.01060 [hep-ex]`
- H. AbouZeid, Measurement of Properties of the Higgs Boson in the WW(lvlv) Channel using the ATLAS Detector, European Physical Journal Web of Conferences **71**, 00001 (2014) p. 1
- ATLAS Collaboration, Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC, Eur. Phys. J. C **73** (2013) p. 2518, arXiv: `1302.4393 [hep-ex]`
- H. AbouZeid, Status of the ATLAS Liquid Argon Calorimeter; Performance after 2 years of LHC operation, Journal of Physics: Conference Series **404**.1 (2012) p. 012004

One further publication in process of receiving ATLAS review/approval on the ATLAS luminosity determination for the 2012/2013 runs.

Selected Public Talks

August 2015 *"The Search for a Heavy Higgs Boson in the $WW \to \ell\nu jj$ Decay Channel with the ATLAS Detector"*
Santa Cruz Institute for Particle Physics, Santa Cruz, USA.

June 2014 *"The Search for a Heavy Higgs-like Boson in the $WW(\ell\nu jj)$ Decay Channel with the ATLAS Detector"*
Canadian Association of Physicists Congress, Sudbury, Canada.

Sept 2013 *"Measurement of Properties of the Higgs Boson in the $WW(\ell\nu\ell\nu)$ Channel using the ATLAS Detector"*
International Conference on New Frontiers in Physics, Kolymbari, Greece.

Sept 2013 *"Measurement of Properties of the Higgs Boson in the $WW(\ell\nu\ell\nu)$ Channel using the ATLAS Detector"*
Niels Bohr Institute, Copenhagen, Denmark.

June 2012 *"Status of the ATLAS Liquid Argon Calorimeter; Performance after 2 years of LHC Operation"*
International Conference on Calorimetry in High Energy Physics, Sante Fe, USA.