

Observation of the Higgs boson in the WW^* channel and search for Higgs boson pair production in the $b\bar{b}b\bar{b}$ channel with the ATLAS detector

A DISSERTATION PRESENTED

BY

TOMO LAZOVICH

TO

THE DEPARTMENT OF PHYSICS

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN THE SUBJECT OF

PHYSICS

HARVARD UNIVERSITY

CAMBRIDGE, MASSACHUSETTS

MAY 2016

©2016 – TOMO LAZOVICH
ALL RIGHTS RESERVED.

Thesis advisor: Professor Joao Guimaraes da Costa and Melissa Franklin

Tomo Lazovich

**Observation of the Higgs boson in the WW^* channel and search
for Higgs boson pair production in the $b\bar{b}b\bar{b}$ channel with the
ATLAS detector**

ABSTRACT

We measured things. And searched for other things. Here is what we found, please let me graduate.

Contents

o INTRODUCTION	I
I THEORETICAL AND EXPERIMENTAL BACKGROUND	2
I.1 THE PHYSICS OF THE HIGGS BOSON	3
I.1.1 The Standard Model of Particle Physics	3
I.1.2 Electroweak Symmetry Breaking and the Higgs	5
I.1.3 Higgs Boson Production and Decay	8
I.1.4 Higgs Pair Production in the Standard Model	12
I.1.5 Higgs Pair Production in Theories Beyond the Standard Model	14
I.1.6 Conclusion	18
I.2 THE ATLAS DETECTOR AND THE LARGE HADRON COLLIDER	19
I.2.1 The Large Hadron Collider	20
I.2.2 The ATLAS Detector	22
I.2.3 The ATLAS Muon New Small Wheel Upgrade	33
I.2.4 Object Reconstruction in ATLAS	37
II OBSERVATION AND MEASUREMENT OF HIGGS BOSON DECAYS TO WW^* IN LHC RUN I AT $\sqrt{s} = 7$ AND 8 TeV	45
II.1 $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ ANALYSIS STRATEGY	46
II.1.1 Introduction	46
II.1.2 The $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ signal in ATLAS	47
II.1.3 Background processes	48
II.1.4 Shared signal region selection requirements	52
II.1.5 Background reduction in same-flavor final states	54
II.1.6 Parameters of interest and statistical treatment	60
II.2 THE DISCOVERY OF THE HIGGS BOSON AND THE ROLE OF THE $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ CHANNEL	66
II.2.1 Introduction	66

4.2	Data and simulation samples	67
4.3	$H \rightarrow \gamma\gamma$ search	67
4.4	$H \rightarrow ZZ \rightarrow 4\ell$ search	69
4.5	$H \rightarrow WW \rightarrow e\nu\mu\nu$ search	72
4.6	Combined results	74
4.7	Conclusion	76
5	OBSERVATION OF VECTOR BOSON FUSION PRODUCTION OF $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$	79
5.1	Introduction	79
5.2	Data and simulation samples	80
5.3	Object selection	84
5.4	Analysis selection	87
5.5	Background estimation	96
5.6	Systematic uncertainties	107
5.7	Results	III
6	COMBINED RUN 1 $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ RESULTS	116
6.1	Introduction	116
6.2	Results of dedicated gluon fusion $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ search	117
6.3	Signal strength measurements in ggF and VBF production	119
6.4	Measurement of Higgs couplings to vector bosons and fermions	122
6.5	Higgs production cross section measurement	123
6.6	Conclusion	124
III	Search for Higgs pair production in the $HH \rightarrow b\bar{b}b\bar{b}$ channel in LHC Run 2 at $\sqrt{s} = 13$ TeV	126
7	SEARCH OVERVIEW	127
8	SEARCH FOR HIGGS PAIR PRODUCTION IN BOOSTED FINAL STATES	128
9	RESULTS WITH RUN 2 2015 DATASET	129
IV	Looking ahead	130
10	CONCLUSION	131
	REFERENCES	138

Listing of figures

1.1	The particles of the Standard Model and their properties[1].	4
1.2	The four most common Higgs boson production modes at the LHC: (a) gluon-gluon fusion, (b) vector boson fusion, (c) $W/Z + H$ production, (d) $t\bar{t}H$ production	8
1.3	Higgs production cross sections as a function of center of mass energy (\sqrt{s}) at a pp collider[2].	9
1.4	Higgs boson branching ratios as a function of m_H [2].	II
1.5	The two leading diagrams for Standard Model di-Higgs production at the LHC: (a) box diagram, (b) Higgs self coupling	13
1.6	Diagrams with new vertices for non-resonant Higgs pair production arising in composite Higgs models	14
1.7	Generic Feynman diagram for resonant Higgs pair production in BSM theories	15
1.8	Branching ratios for a spin-2 Randall-Sundrum graviton as a function of mass computed in MadGraph with the CP3-Origins implementation [3, 4]	16
1.9	$\sigma \times BR(HH)$ for RSG as a function of mass computed in MadGraph with the CP3-Origins implementation [3, 4]	16
1.10	RSG width as a function of mass computed in MadGraph with the CP3-Origins implementation [3, 4]	17
1.11	Branching ratios for heavy Higgs H in Type I (left) and Type II (right) 2HDM models with $\tan \beta = 1.5$ and $\cos(\beta - \alpha) = 0.1(0.01)$ for Type I (Type II). [5]	18
2.1	A schematic view of the LHC ring [6]	20
2.2	A full diagram of the ATLAS detector [7]	23
2.3	The ATLAS coordinate system	24
2.4	Layout of the ATLAS Inner Detector system [8]	25
2.5	Layout of the ATLAS calorimeter system [7]	27
2.6	Layout of the ATLAS muon system [7]	29
2.7	Predicted field integral as a function of $ \eta $ for the ATLAS magnet system [7]	31
2.8	ATLAS trigger rates for Level-1 triggers as a function of instantaneous luminosity in 2012 and 2015 operation. These are single object triggers for electromagnetic clusters (EM), muons (MU), jets (J), missing energy (XE), and τ leptons (TAU). The threshold of the trigger is given in the name in GeV. [9]	32
2.9	Instantaneous luminosity as a function of time for data recorded by ATLAS at different center of mass energies [10, 11]	33

2.10	MDT tube hit (solid) and segment (dashed) efficiency as a function of hit rate per tube [12]	35
2.11	Trigger rate as a function of p_T threshold with and without the NSW upgrade [12]	35
2.12	Illustrations of the geometry (left) and operating principle (right) of the micromegas detector [12]	36
2.13	Geometry of the sTGC detector [12]	37
2.14	Illustration of particle interactions in ATLAS [13]	38
2.15	Electron performance: (a) reconstruction efficiency as a function of electron E_T [14] (b) energy resolution in simulation as a function of $ \eta $ for different energy electrons [15]	39
2.16	Muon performance: (a) reconstruction efficiency as a function of muon p_T (b) dimuon mass resolution as a function of average p_T [16]	41
2.17	Jet energy response after calibration as a function of true p_T in simulation [17]	42
2.18	Light jet rejection (1/efficiency) vs. b -jet efficiency for MV1 and its input algorithms (a) [18] and MV2 (b) [19] in simulated $t\bar{t}$ events. The numbers in the algorithm names in (b) refer to the fraction of charm events used in the MV2 training.	43
3.1	Branching ratios for a WW system. q refers to quarks. ℓ can be either an electron or muon, and the leptonic branching ratios of the τ are included. For example, the $\ell\nu qq$ final state includes one W decaying to $e\nu$, $\mu\nu$, or $\tau\nu$. τ_h refer to hadronic decays of the τ	48
3.2	Feynman diagram for Standard Model WW production	49
3.3	Feynman diagrams for top pair production (left) and Wt production (right)	50
3.4	An example Feynman diagram of W +jets production	50
3.5	An example Feynman diagram of Z +jets production	51
3.6	An illustration of the unique analysis signal regions[20]	52
3.7	A graphical illustration of the $E_{T,\text{rel}}^{\text{miss}}$ calculation	54
3.8	Predicted backgrounds (compared with data) as a function of n_j (a and b) and n_b (c)	55
3.9	An event display of a Z/γ^* + jets event illustrating the effect of pileup interactions	56
3.10	The RMS of different missing transverse momentum definitions as a function of the average number of interactions per bunch crossing	57
3.11	The difference between the true and reconstructed values of the missing transverse momentum (a) and m_T (b) in a gluon fusion signal sample	58
3.12	Comparison of f_{recoil} distributions for Z/γ^* +jets, $H \rightarrow WW^*$, and other backgrounds with real neutrinos.	59
3.13	Signal significance as a function of required value for f_{recoil} and $p_{T,\text{rel}}^{\text{miss}(\text{trk})}$ in the ggF $H \rightarrow WW^*$ with $n_j = 0$	60

4.1	Diphoton mass spectrum in 7 and 8 TeV data. Panel a) shows the unweighted data distribution superimposed on the background fit, while panel c) shows the data where each event category is weighted by its signal to background ratio. Panels b) and d) show the respective distributions with background subtracted[21].	70
4.2	Four lepton invariant mass spectrum ($m_{4\ell}$) in 7 and 8 TeV data compared to background estimate. A 125 GeV SM Higgs signal is shown in blue[21].	72
4.3	m_T distribution in the $H \rightarrow WW \rightarrow e\nu\mu\nu n_j \leq 1$ channels for 8 TeV data[21].	74
4.4	Local p_0 distribution as a function of hypothesized Higgs mass for the $H \rightarrow ZZ^* \rightarrow 4\ell$ (a), $H \rightarrow \gamma\gamma$ (b), and $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ (c) channels. Dashed curves show expected results, while solid curves show observed. Red curves are from 7 TeV data, blue curves from 8 TeV, and black curved combined[21].	76
4.5	Combined 95% CL limits (a), local p_0 values (b), and signal strength measurement (c) as a function of Higgs mass[21].	77
4.6	Comparison of measured signal strength μ for a 126 GeV Higgs in the 7 and 8 TeV datasets[21].	78
4.7	Two dimensional likelihood as a function of signal strength μ and Higgs mass m_H [21].	78
5.1	A comparison of the subleading lepton p_T spectrum between VBF $H \rightarrow WW^*$ production and $t\bar{t}$ background	80
5.2	Leading jet η in VBF $H \rightarrow WW^*$ (red) and $t\bar{t}$ (black)	89
5.3	Distributions of (a) m_{jj} , (b) Δy_{jj} , (c) $C_{\ell 1}$, and (d) $\sum m_{\ell j}$, for the VBF analysis. The top panels compare simulation and data, while the bottom panels show normalized distributions for all background processes and signal[20].	91
5.4	A cartoon of the WW final state. Momenta are represented with thin arrows, spins with thick arrows. [20]	92
5.5	Event display of a VBF candidate event[20].	94
5.6	Distributions of $m_{\ell\ell}$ (top left), $\Delta\phi_{\ell\ell}$ (top right), and m_T (bottom) for the VBF BDT analysis. These are plotted after all of the BDT pre-training selection cuts[20].	95
5.7	Distributions of m_{jj} (top left), Δy_{jj} (top right), $\sum C_\ell$ (bottom left), and $\sum m_{\ell j}$ (bottom right) for the VBF BDT analysis. These are plotted after all of the BDT pre-training selection cuts[20].	96
5.8	Distributions of O_{BDT} for the VBF signal and associated backgrounds after the VBF pre-training selection[20].	97
5.9	Distributions of m_{jj} (a) and O_{BDT} (b) in the VBF $n_b = 1$ top CR[20].	100
5.10	Comparison of m_{jj} shape in a same flavor $Z \rightarrow \ell\ell$ control region and the VBF cut-based signal region.	101
5.11	General illustration of the ABCD region definitions for $Z/\gamma^* \rightarrow \ell\ell$ background estimation.	102
5.12	Distribution of m_{T2} in the WW validation region of the VBF analysis[20].	104

5.13	Extrapolation factors for the W +jets estimate derived for muons (a) and electrons (b) as a function of lepton p_T [20].	106
5.14	Background composition in final VBF signal region[20].	106
5.15	Variations in the top background extrapolation factor in the cut-based analysis due to PDF uncertainties, binned in m_T	109
5.16	Variations in the top background extrapolation factor in the cut-based analysis due to QCD scale uncertainties, binned in m_T	109
5.17	Postfit distributions in the cut-based VBF analysis. Panel (a) shows the one-dimensional m_T distribution, while (b) shows the data candidates split into the bins of m_T and m_{jj} used in the final fit[20].	113
5.18	Postfit distributions in the BDT VBF analysis[20].	114
5.19	Overlap between cut-based and BDT VBF signal region candidates in the m_{jj} - m_T plane. .	114
6.1	Post-fit m_T distribution in the $n_j \leq 1$ regions[20].	118
6.2	Best fit signal strength $\hat{\mu}$ as a function of hypothesized m_H [20].	120
6.3	Local p_0 as a function of m_H [20].	120
6.4	Likelihood as a function of μ_{VBF}/μ_{ggF} [20].	121
6.5	Likelihood scan as a function of μ_{VBF} and μ_{ggF} [20].	122
6.6	Likelihood scan as a function of κ_F and κ_V [20].	123
6.7	Comparison of signal strength measurements in different Higgs decay channels on ATLAS[22].	125

Listing of tables

1.1	Production cross sections for a 125 GeV Higgs boson at $\sqrt{s} = 8$ TeV with scale and PDF uncertainties [2].	10
1.2	Branching ratios for a 125 GeV Higgs boson[2].	12
1.3	Possible channels for Higgs searches. Checkmarks denote the most sensitive production modes [23].	12
1.4	Production cross sections for pair production of a 125 GeV Higgs boson at $\sqrt{s} = 14$ TeV with total uncertainty [24]. The uncertainties include QCD scale and PDF variations as well as uncertainties on α_S	13
2.1	Evolution of LHC machine conditions [25, 26]	22
2.2	Performance requirements for the ATLAS detector [7].	34
2.3	Signal efficiencies for WH production with $H \rightarrow b\bar{b}$ and $H \rightarrow WW^* \rightarrow \mu\nu qq$ under different trigger configurations [12].	37
3.1	A summary of backgrounds to the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ signal	51
4.1	Monte carlo generators used to model signal and background for the Higgs search[21].	67
4.2	A summary of selection efficiencies and expected resolutions for a 125 GeV Higgs in the $H \rightarrow ZZ \rightarrow 4\ell$ channel[21].	71
4.3	Summary of the expected and observed significance and measured signal strengths in the combined 7 and 8 TeV datasets for the Higgs discovery analysis[21].	75
5.1	Single lepton triggers used for electrons and muons. A logical “or” of the triggers listed for each lepton type is taken. Units are in GeV, and the i denotes an isolation requirement in the trigger.	81
5.2	Di-lepton triggers used for different flavor combinations. The two thresholds listed refer to leading and sub-leading leptons, respectively. The di-muon trigger only requires a single lepton at level-i.	81
5.3	Trigger efficiency for signal events and relative gain of adding a dilepton trigger on top of the single lepton trigger selection. The first lepton is the leading, while the second is the sub-leading. Efficiencies shown here are for the ggF signal in the $n_j = 0$ category but are comparable for the VBF signal.	82
5.4	Monte Carlo samples used to model the signal and background processes[20].	83

5.5	p_T dependent isolation requirements for muons. Muons are required to have the amount of calorimeter or track based cone sums be less than this fraction of their p_T	85
5.6	p_T dependent requirements for electrons. Electrons are required to have the amount of calorimeter or track based cone sums be less than this fraction of their E_T	86
5.7	Event selection for the $n_j \geq 2$ VBF analysis in the 8 TeV cut-based analysis[20].	93
5.8	Top normalization factors computed at each stage of the cut-based selection. Uncertainties are statistical only.	99
5.9	Top normalization factors computed for each bin of O_{BDT} . Uncertainties are statistical only.	99
5.10	$Z/\gamma^* \rightarrow \tau\tau$ correction factors for the VBF cut-based analysis. Uncertainties are statistical only.	102
5.11	$Z/\gamma^* \rightarrow \ell\ell$ normalization factors for cut-based and BDT analyses. Uncertainties are statistical only.	103
5.12	Systematic uncertainties for various processes in the BDT analysis, given in units of % change in yield. Values are given for the most sensitive BDT bin (bin 3), except where noted with a *, in which case the uncertainty affect the normalization in all BDT bins. Empty entries indicate that the uncertainty is negligible or not applicable to this background.	108
5.13	Systematic uncertainties for various processes in the cut-based VBF analysis, given in units of % change in yield. Values are given for the low m_{jj} signal region.	108
5.14	Composition of the post-fit uncertainties (in %) on the total signal (N_{sig}), total background (N_{bkg}), and individual background yields in the VBF analysis[20].	111
5.15	Event selection for the VBF BDT analysis. The event yields in (a) are shown after the pre-selection and the additional requirements applied before the BDT classification (see text). The event yields in (b) are given in bins in O_{BDT} after the classification[20].	112
6.1	All signal regions definitions input into final statistical fit[20].	117
6.2	Post-fit yields in the different ggF and VBF dedicated signal regions[20].	118

THIS IS THE DEDICATION.

Acknowledgments

LOREM IPSUM DOLOR SIT AMET, consectetur adipiscing elit. Morbi commodo, ipsum sed pharetra gravida, orci magna rhoncus neque, id pulvinar odio lorem non turpis. Nullam sit amet enim. Suspendisse id velit vitae ligula volutpat condimentum. Aliquam erat volutpat. Sed quis velit. Nulla facilisi. Nulla libero. Vivamus pharetra posuere sapien. Nam consectetur. Sed aliquam, nunc eget euismod ullamcorper, lectus nunc ullamcorper orci, fermentum bibendum enim nibh eget ipsum. Donec porttitor ligula eu dolor. Maecenas vitae nulla consequat libero cursus venenatis. Nam magna enim, accumsan eu, blandit sed, blandit a, eros.

0

Introduction

Part I

Theoretical and Experimental Background

1

The Physics of the Higgs Boson

This chapter presents an overview of the Standard Model of Particle Physics and in particular the physics of the Higgs boson. First, a brief overview of the Standard Model and its history are presented. Then, a description of the Higgs mechanism of electroweak symmetry breaking is given. Next, the physics of single Higgs boson production and decay is described. The Standard Model also allows for production of two Higgs bosons and this is detailed as well. Finally, di-Higgs production in two beyond the Standard Model (BSM) theories - Randall-Sundrum gravitons (RSG) and Two Higgs Doublet Models (2HDM) - is shown.

1.1 THE STANDARD MODEL OF PARTICLE PHYSICS

The Standard Model (SM) of Particle Physics is a quantum field theory describing the fundamental particles of nature and the forces that govern their interactions. Several comprehensive treatments of the SM already exist in the literature[1, 23, 27–30] and this section will not rehash those. Rather, this

section presents a brief overview of the SM particles and forces in order to define them for subsequent discussions.

The Standard Model consists of two primary categories of fundamental particles: fermions (spin 1/2 particles) and bosons (integer spin particles). The SM also describes three forces: electromagnetism, the weak nuclear force, and the strong nuclear force. Gravity is not included in the theory and is largely irrelevant at the scales currently probed by collider experiments. Within the fermions, there are both quarks (which interact via all three forces) and the leptons. The charged leptons interact via electromagnetic and weak interactions, while neutrinos (neutral leptons) interact only via the weak force. Within the bosons, there are the W^\pm and Z bosons (the mediators of the weak force), the gluon (g , the mediator of the strong force), and the photon (γ), the mediator of the electromagnetic force. Finally, there is the Higgs boson, a fundamental spin-0 particle resulting from the Higgs mechanism of electroweak symmetry breaking. Figure 1.1 summarizes the fermions and bosons of the SM.

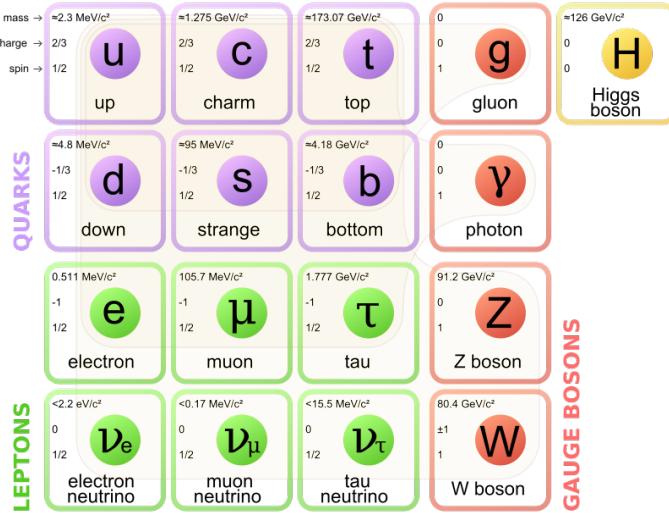


Figure 1.1: The particles of the Standard Model and their properties[1].

The Standard Model coalesced into a unified theoretical framework in the 1960s through the work of Glashow, Weinberg, Salam, and others on the theory of electroweak interactions[31–34]. This theory characterized both the electromagnetic and weak interactions as unified under a single gauge symmetry group, namely $SU(2) \times U(1)$. At low enough energy scales (on the order of the W and Z masses), the electroweak symmetry is broken, as evidenced by the fact that the weak bosons have mass while the

photon does not. The discovery of the Higgs boson in 2012 confirmed the Higgs mechanism as the most likely candidate for this electroweak symmetry breaking[21, 35]. The electroweak theory is then combined with the theory of quantum chromodynamics (which models the strong sector as a non-abelian $SU(3)$ gauge group) to form the complete SM[36].

1.2 ELECTROWEAK SYMMETRY BREAKING AND THE HIGGS

In the Standard Model Lagrangian, it is difficult to include mass terms for the W and Z bosons without breaking the fundamental gauge symmetry of the Lagrangian. A traditional mass term does not preserve the $SU(2) \times U(1)$ symmetry. Additionally, scattering of massive W and Z bosons violate unitarity and these diagrams diverge at high energy scales. In the 1960s, Higgs, Brout, Englert, Guralnik, Kibble, and Hagen developed a mechanism for spontaneous symmetry breaking via the addition of a complex scalar doublet to the SM. Three of the four real degrees of freedom of this complex field would go to the longitudinal modes of the W^\pm and Z , thus allowing them to have mass[37–40]. The remaining degree of freedom would manifest as an additional scalar, known now as the Higgs boson (Higgs was the first to predict the existence of the new particle).

The mechanism works by introducing a Lagrangian for the newly introduced field that still respects the symmetry of the Standard Model inherently, but with a minimum at a non-zero vacuum expectation value for the field. In this minimum of the potential, the electroweak symmetry is broken. Specifically, consider a complex scalar doublet Φ with four degrees of freedom, as shown in equation 1.1.

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1^+ + i\phi_2^+ \\ \phi_1^0 + i\phi_2^0 \end{pmatrix} \quad (1.1)$$

The minimal potential of a self-interacting Higgs that still respects the SM symmetry is given in equation 1.2.

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \quad (1.2)$$

If the μ^2 term of this potential is positive, then the potential has a minimum at $\Phi = 0$ and the SM

symmetry is preserved. However, if instead $\mu^2 < 0$, then the minimum is at a finite value of Φ , namely

$$\Phi_{\min} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad (1.3)$$

where $v = \sqrt{\mu^2/\lambda}$. Because this is the location of the minimum, it corresponds to the vacuum expectation value for the field ($\langle \Phi \rangle = \Phi_{\min}$). The excitations of the Higgs can then be parameterized as

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix} \quad (1.4)$$

The full scalar Lagrangian, including the kinetic term, is then given as

$$\mathcal{L}_s = (D^\mu \Phi)^\dagger (D_\mu \Phi) - V(\Phi) \quad (1.5)$$

where the covariant derivative is defined as

$$D_\mu = \partial_\mu + \frac{ig}{2} \tau^a W_\mu^a + ig' Y B_\mu \quad (1.6)$$

and W^1, W^2, W^3 and B are the $SU(2)$ and $U(1)$ gauge fields of the electroweak theory, respectively. g and g' are the corresponding coupling constants. With the scalar Lagrangian in place, the physical gauge fields can then be written as

$$W_\mu^\pm = \frac{1}{\sqrt{2}} (W_\mu^1 \mp i W_\mu^2) \quad (1.7)$$

$$Z_\mu = \frac{-g' B_\mu + g W_\mu^3}{\sqrt{g^2 + g'^2}} \quad (1.8)$$

$$A_\mu = \frac{g B_\mu + g' W_\mu^3}{\sqrt{g^2 + g'^2}} \quad (1.9)$$

Equation 1.7 corresponds to the charged W^+ and W^- bosons, equation 1.8 corresponds to the neu-

tral Z boson, and equation 1.9 corresponds to the neutral photon. The masses of the particles also arise from the Lagrangian. The photon has zero mass, while the masses of the W and Z bosons are given in equation 1.10.

$$\begin{aligned} M_W^2 &= \frac{1}{4}g^2v^2 \\ M_Z^2 &= \frac{1}{4}(g^2 + g'^2)v^2 \end{aligned} \tag{1.10}$$

The fermion masses also arise through a coupling with the Higgs via the Yukawa interaction (for a detailed description, see [30]). In this case the coupling between the Higgs and the fermions goes as

$$g_{Hf\bar{f}} = \frac{m_f}{v} \tag{1.11}$$

The full Lagrangian of Higgs interactions can be written as

$$\mathcal{L}_{\text{Higgs}} = -g_{Hf\bar{f}}\bar{f}fH + \frac{g_{HHH}}{6}H^3 + \frac{g_{HHHH}}{24}H^4 + \delta_V V_\mu V^\mu \left(g_{HVV}H + \frac{g_{HHV}}{2}H^2 \right) \tag{1.12}$$

with

$$\begin{aligned} g_{HVV} &= \frac{2m_V^2}{v} & g_{HHV} &= \frac{2m_V^2}{v^2} \\ g_{HHH} &= \frac{3m_H^2}{v} & g_{HHHH} &= \frac{3m_H^2}{v^2} \end{aligned} \tag{1.13}$$

Here, V refers to the W^\pm and Z , and $\delta_W = 1$ while $\delta_Z = 1/2$. Phenomenologically, there are a few features of this Lagrangian that are useful to note. First, note that the Higgs mass is a free parameter of the theory that must be determined experimentally. Second, note that the coupling of the Higgs to the vector bosons and fermions scales with the masses of these particles, a fact that is important when considering both the production and decays of the particle. Also note that the branching ratio of the Higgs to W bosons will be twice that of the branching ratio to Z if the Higgs mass is large enough to produce the particles on shell because of the extra symmetry factor associated with the W coupling. Finally, note the presence of the cubic and quartic Higgs self interaction terms, which can lead to final

states with multiple Higgs bosons produced.

1.3 HIGGS BOSON PRODUCTION AND DECAY

This section discusses the properties of Higgs production and decay mechanisms. The details presented here will focus on the properties of a 125 GeV Higgs boson, as this is the mass closest to that of the newly discovered Higgs.

1.3.1 HIGGS PRODUCTION

The Higgs is produced by four main production modes at the Large Hadron Collider - gluon-gluon fusion (ggF), vector boson fusion (VBF), associated production with a W or Z boson, or associated production with top quarks ($t\bar{t}H$). Figure 1.2 shows the Feynman diagrams for these four modes.

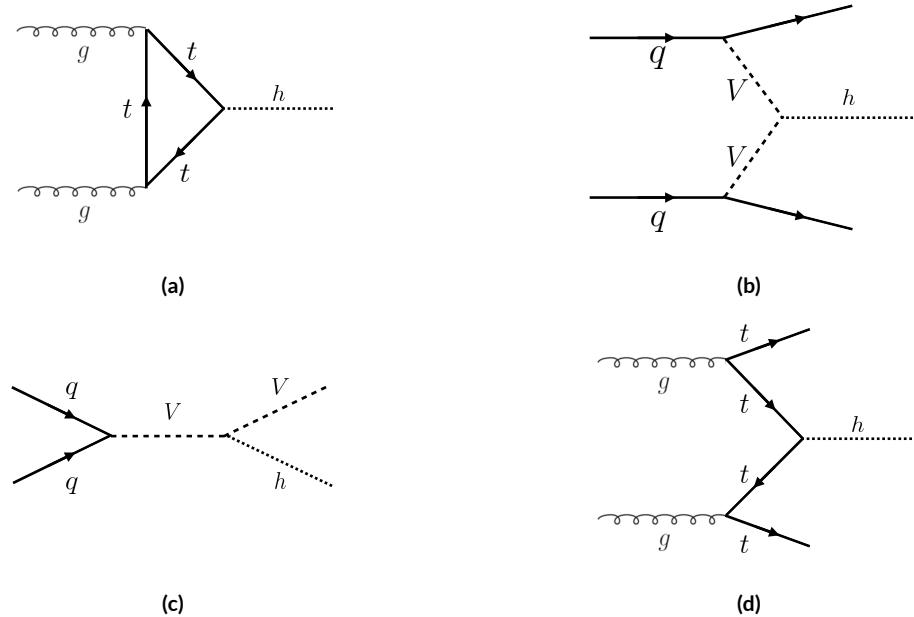


Figure 1.2: The four most common Higgs boson production modes at the LHC: (a) gluon-gluon fusion, (b) vector boson fusion, (c) $W/Z + H$ production, (d) $t\bar{t}H$ production

In gluon-gluon fusion, gluons from the incoming protons fuse via a top-quark loop to produce a Higgs. The top quark is the dominant contribution in the loop due to its heavy mass and the fact that the Higgs-fermion coupling constant scales with fermion mass. In vector boson fusion, the incoming

quarks each radiate a W or Z boson which fuse to produce the Higgs. This production mode results in a final state with a Higgs boson and two additional jets which tend to be forward because they carry the longitudinal momentum of the incoming partons. The Higgs can also be produced in association with a W or Z boson. The W/Z is produced normally and then radiates a Higgs (this mode is also sometimes known as “Higgs-strahlung”). Finally, the Higgs can be produced in association with two top quarks. Each incoming gluon splits into a $t\bar{t}$ pair, and one of the top pairs combines to create a Higgs.

Figure 1.3 shows the production cross section for a 125 GeV Higgs boson in each of these modes at a pp collider as a function of center of mass energy.

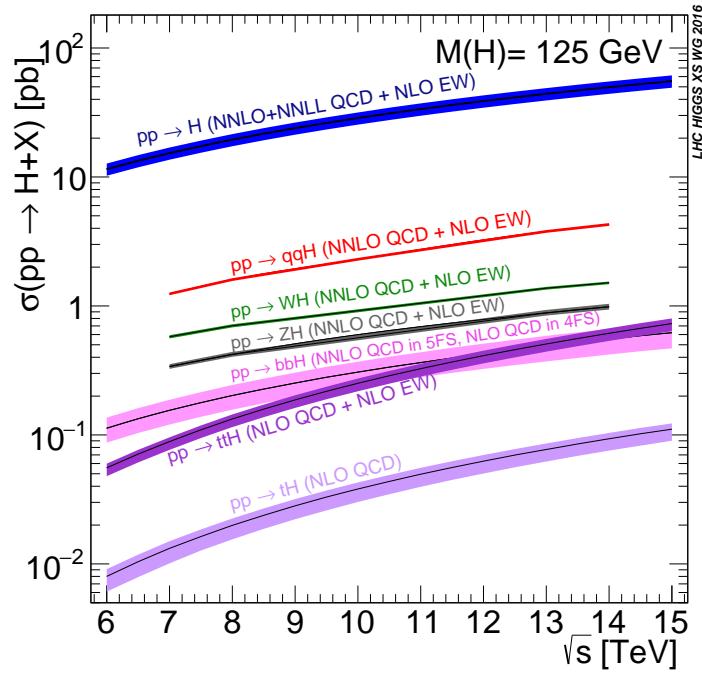


Figure 1.3: Higgs production cross sections as a function of center of mass energy (\sqrt{s}) at a pp collider[2].

In figure 1.3, note that gluon fusion has the largest cross section, while VBF is the second largest at approximately a factor of 10 smaller. The figure also includes the less commonly studied $b\bar{b}H$ and tH modes. The $b\bar{b}H$ and tH modes are not studied as commonly as $t\bar{t}H$ due to the larger background contributions and lower cross sections, respectively. At $\sqrt{s} = 8$ TeV, ggF production of a 125 GeV Higgs has a cross section of 19.47 pb, while VBF has a cross section of 1.601 pb[2]. The cross sections of all of the main Higgs production modes at this center of mass energy, as well as their uncertainties from

varying the renormalization and factorization scales and PDFs, are summarized in table 1.1 for a 125 GeV Higgs.

Production mode	σ (pb)	QCD scale uncert. (%)	PDF + α_s uncert. (%)
Gluon fusion	19.47	+7.3 / - 8.0	3.1
Vector boson fusion	1.601	+0.3 / - 0.2	2.2
WH	0.7026	+0.6 / - 0.9	2.0
ZH	0.4208	+2.9 / - 2.4	1.7
$t\bar{t}H$	0.1330	+4.1 / - 9.2	4.3
bbH	0.2021	+20.7 / - 22.3	
tH (t -channel)	0.01869	+7.3 / - 16.5	4.6
tH (s -channel)	1.214×10^{-3}	+2.8 / - 2.4	2.8

Table 1.1: Production cross sections for a 125 GeV Higgs boson at $\sqrt{s} = 8$ TeV with scale and PDF uncertainties [2].

1.3.2 HIGGS BRANCHING RATIOS

The fact that the Higgs couples more strongly to more massive particles is crucial for understanding its branching ratios. The width for Higgs decays to fermions is given in equation 1.14 [23].

$$\Gamma(H \rightarrow f\bar{f}) = \frac{N_c \sqrt{2} G_F m_f^2 m_H}{8\pi} \quad (1.14)$$

In this case, N_c is the number of colors, G_F is the Fermi constant, m_f is the mass of the fermion, and m_H is the mass of the Higgs. Note that the width scales with the square of the fermion mass. (This also assumes that the Higgs mass is large enough to decay with both the fermions on shell.)

The decay width to WW is given in equation 1.15 [23].

$$\Gamma(H \rightarrow W^+W^-) = \frac{\sqrt{2} G_F M_W^2 m_H}{16\pi} \frac{\sqrt{1-x_W}}{x_W} (3x_W^2 - 4x_W + 4) \quad (1.15)$$

where m_W is the mass of the W and $x_W = 4M_W^2/m_H^2$. To get the branching ratio to ZZ , the equation is divided by 2 to account for identical particles in the final state, and x_W is replaced with

$x_Z = 4M_Z^2/m_H^2$. This is shown in equation 1.16 [23].

$$\Gamma(H \rightarrow ZZ) = \frac{\sqrt{2}G_F M_Z^2 m_H}{32\pi} \frac{\sqrt{1-x_Z}}{x_Z} (3x_Z^2 - 4x_Z + 4) \quad (1.16)$$

These formulas can also be visualized as a function of Higgs mass. Figure 1.4 shows the branching ratios as a function of the Higgs mass.

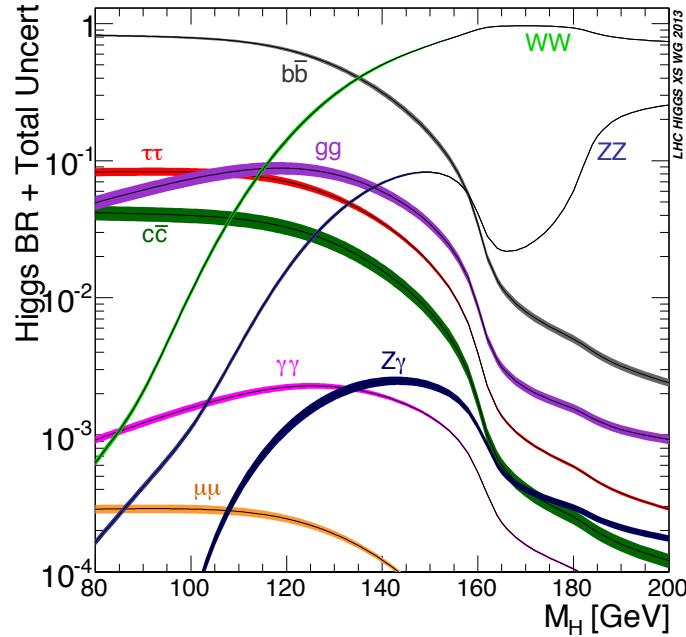


Figure 1.4: Higgs boson branching ratios as a function of m_H [2].

There are a few interesting features to note in this figure. First, note that at high Higgs masses, once on-shell production of both W and Z bosons is possible, these two decays are the dominant ones due to the large masses of the W/Z . Also note that the branching ratio to W s is twice that of Z s at these large masses due to the δ_V symmetry factor noted previously. At 125 GeV, the Higgs is accessible through many different decay modes. The largest branching ratio is the decay $H \rightarrow b\bar{b}$ at 58.24% [2]. This branching is larger than the WW/ZZ decays because one of the two bosons must be produced off-shell for $m_H = 125$ GeV. The second largest branching ratio is to WW^* at 21.37 % (before taking into account the branching ratios of the W). Table 1.2 summarizes the branching ratios for a 125 GeV Higgs. Note that there is in fact a Higgs branching ratio to $\gamma\gamma$ even though photons are massless. This

decay happens through a loop (the largest contributions to the loop are top and W) which suppresses the branching ratio.

Decay	Branching ratio (%)
$b\bar{b}$	58.24
WW^*	21.37
gg	8.187
$\tau\tau$	6.272
$c\bar{c}$	2.891
ZZ^*	2.619
$\gamma\gamma$	0.2270
$Z\gamma$	0.1533
$\mu\mu$	0.02176

Table 1.2: Branching ratios for a 125 GeV Higgs boson[2].

Note that the branching ratios alone do not tell the full story of which Higgs channels are the most sensitive. For example, a $H \rightarrow b\bar{b}$ search in gluon fusion production is incredibly difficult due to the large QCD dijet background at the LHC. However, in associated production of the Higgs, where a W or Z gives additional final state particles that can be used to reduce background, a search for $H \rightarrow b\bar{b}$ can be sensitive. The combinations of production and decay modes that are most commonly studied are summarized in table 1.3 [23].

Decay	Inclusive (incl. ggF)	VBF	WH/ZH	$t\bar{t}H$
$H \rightarrow \gamma\gamma$	✓	✓	✓	✓
$H \rightarrow b\bar{b}$			✓	✓
$H \rightarrow \tau^+\tau^-$		✓		
$H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$	✓	✓	✓	
$H \rightarrow ZZ \rightarrow 4\ell$	✓			
$H \rightarrow Z\gamma \rightarrow \ell\ell\gamma$	very low			

Table 1.3: Possible channels for Higgs searches. Checkmarks denote the most sensitive production modes [23].

1.4 HIGGS PAIR PRODUCTION IN THE STANDARD MODEL

The Standard Model also allows for processes that produce two Higgs bosons in the final state, known as Higgs pair production or di-Higgs production. The two main production mechanisms are shown in

figure 1.5.

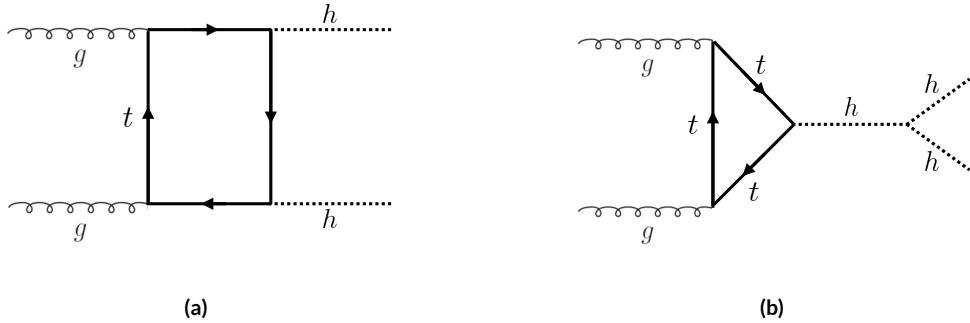


Figure 1.5: The two leading diagrams for Standard Model di-Higgs production at the LHC: (a) box diagram, (b) Higgs self coupling

The two diagrams in figure 1.5 interfere destructively with one another, resulting in a low overall cross section for di-Higgs production at the LHC. Nevertheless, Higgs pair production is quite interesting to study because it gives direct access to the λ parameter of the Higgs potential, also known as the Higgs self coupling. The diagram in figure 1.5(b) is sensitive to this coupling through the triple Higgs vertex.

One can substitute the gluon fusion production of diagram 1.5(b) with any of the other production modes previously discussed. These other modes do not suffer from interference with the box diagram in figure 1.5(a) due to the presence of additional particles in the final state. They still have a lower cross section than the gluon fusion mode, however. The cross sections for di-Higgs production in the different modes, as well as their uncertainties, are shown in table 1.4 [24]. These are shown for $\sqrt{s} = 14$ TeV as the higher center of mass energy is more sensitive to this process. Note that the scale of cross section quoted is now in fb rather than pb.

Production mode	σ (fb)	Total uncert. (%)
Gluon fusion	33.89	+37.2 / - 27.8
Vector boson fusion	2.01	+7.6 / - 5.1
$W H H$	0.57	+3.7 / - 3.3
$Z H H$	0.42	+7.0 / - 5.5
$t \bar{t} H$	1.02	-

Table 1.4: Production cross sections for pair production of a 125 GeV Higgs boson at $\sqrt{s} = 14$ TeV with total uncertainty [24]. The uncertainties include QCD scale and PDF variations as well as uncertainties on α_S .

1.5 HIGGS PAIR PRODUCTION IN THEORIES BEYOND THE STANDARD MODEL

The Standard Model Higgs pair production cross section is rather small, and datasets on the scale of the full lifetime of the LHC will be required to obtain sensitive measurements of the Higgs self-coupling. However, the discovery of the Higgs also gives particle physicists a new tool that can be exploited in the search for new physics beyond the Standard Model. In particular, Higgs pair production is a promising channel in the search for new physics. The cross section for di-Higgs production can be altered through both resonant and non-resonant production of Higgs pairs. In non-resonant production, di-Higgs production vertices can arise from the presence of a new strong sector and additional colored particles[41–43]. Figure 1.6 shows examples of the types of vertices that can arise. In the resonant case, new heavy particle can decay to Higgs pairs. Such new particles can include heavy Higgs bosons arising in two Higgs doublet models (2HDM) or Higgs portal models as well as heavy gravitons in Randall-Sundrum theories[3, 5, 41, 44–48]. Figure 1.7 shows a generic diagram for a heavy resonance decaying to two Higgs bosons. In the 2HDM, X corresponds to the heavy CP-even scalar H . In the Randall-Sundrum model, X corresponds to a heavy spin-2 graviton G .

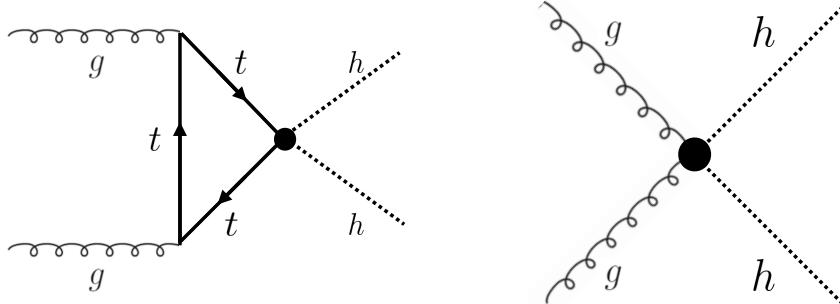


Figure 1.6: Diagrams with new vertices for non-resonant Higgs pair production arising in composite Higgs models

The next sections provide more detail on the phenomenology of resonant Higgs production in Randall-Sundrum and 2HDM models, as these models will later be tested in a dedicated search for resonant production of boosted Higgs pairs.

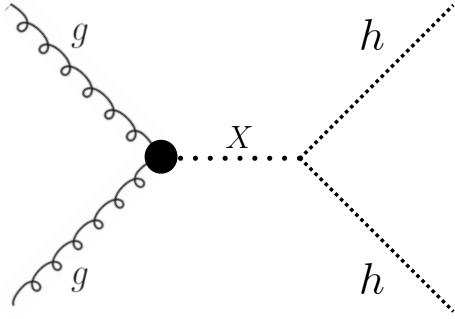


Figure 1.7: Generic Feynman diagram for resonant Higgs pair production in BSM theories

1.5.1 RANDALL-SUNDRUM GRAVITONS

The Randall-Sundrum model is a proposed solution to the hierarchy problem that posits a five-dimensional warped spacetime that contains two branes: one where the force of gravity is very strong and a second brane at the TeV scale corresponding to the known Standard Model sector [44]. In the theory, the branes are weakly coupled and the graviton probability function drops exponentially going from the gravity brane to the SM brane, rendering gravity weak on the SM brane. The experimental consequence of this theory is a tower of widely spaced (in mass) Kaluza-Klein graviton resonances. In theories where the fermions are localized to the SM brane, production of gravitons from fermion pairs is suppressed and the primary mode of production is gluon fusion[3]. These gravitons have a substantial branching fraction to Higgs pairs, ranging from 6.43% for gravitons with a mass of 500 GeV to 7.66% at 3 TeV. Figure 1.8 shows the branching ratios of the spin-2 Randall Sundrum graviton (RSG) as a function of its mass. The predominant decays are to $t\bar{t}$ above the mass threshold for that channel.

These models have two free parameters - the mass of the graviton and a curvature parameter k . Typically, rather than k , the theory is parameterized using $c \equiv k/\bar{M}_{\text{pl}}$, where \bar{M}_{pl} is the reduced Planck mass. The cross section for production of the RSG decreases as a function of mass and is strongly dependent on the gluon PDF. The increase in center of mass energy from 8 to 13 TeV in LHC Run 2 greatly increases the cross section at higher mass. Figure 1.9 shows the cross section as a function of graviton mass at $\sqrt{s} = 13$ TeV for RSG models with $c = 1.0$ and $c = 2.0$.

Another interesting feature of the theory is that the width of the graviton increases with both c and

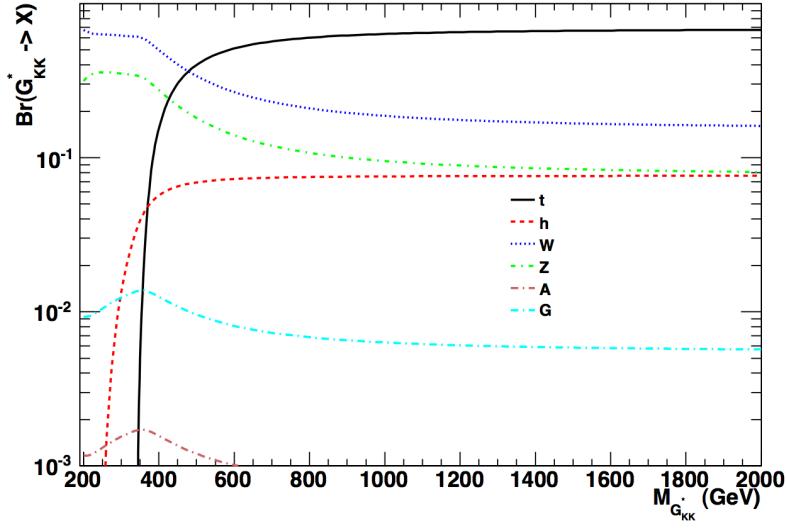


Figure 1.8: Branching ratios for a spin-2 Randall-Sundrum graviton as a function of mass computed in MadGraph with the CP3-Origins implementation [3, 4]

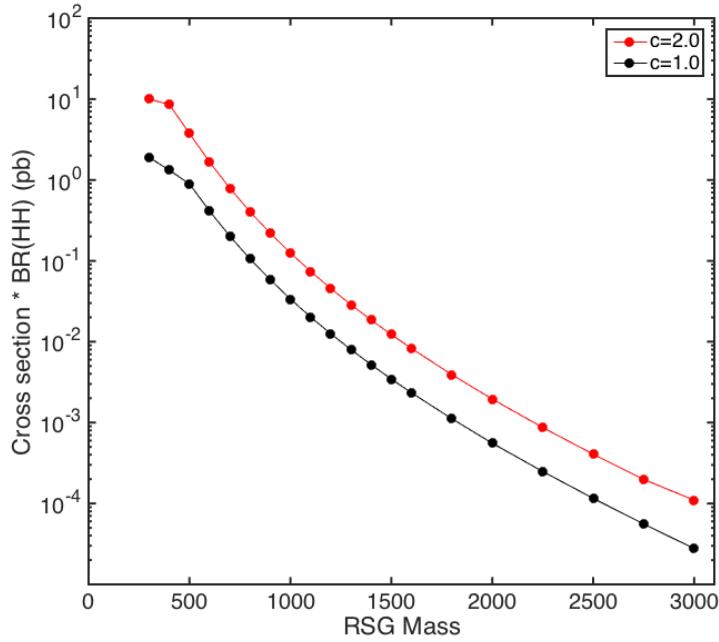


Figure 1.9: $\sigma \times \text{BR}(HH)$ for RSG as a function of mass computed in MadGraph with the CP3-Origins implementation [3, 4]

m_G . Figure 1.10 shows the graviton width for both $c = 1.0$ and $c = 2.0$ as a function of mass. In $c = 1.0$, the width starts at 8.365 GeV for a mass of 300 GeV and increases to 187.2 GeV at a mass of

3 TeV. Similarly, with $c = 2.0$, the width starts at 33.46 GeV for $m_G = 300$ GeV and increases to 748.8 GeV at a mass of 3 TeV.

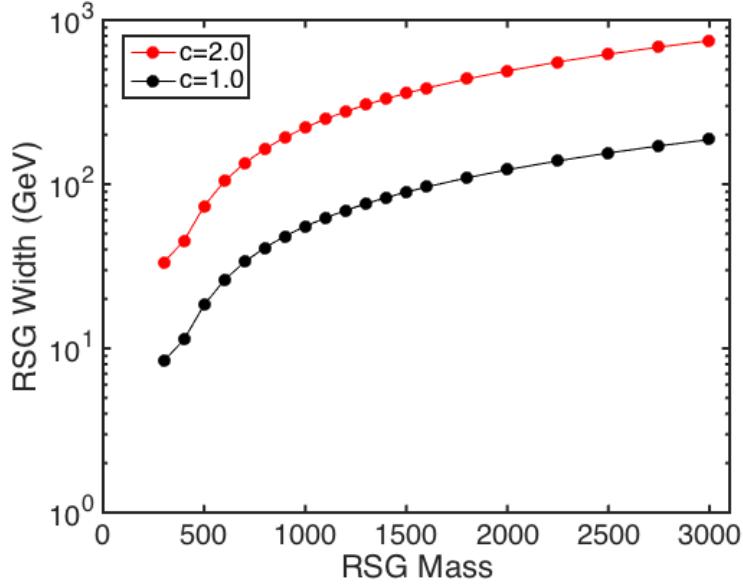


Figure 1.10: RSG width as a function of mass computed in MadGraph with the CP3-Origins implementation [3, 4]

1.5.2 TWO HIGGS DOUBLET MODELS

In Two Higgs Doublet Models (2HDM), a second complex scalar doublet is added to SM [5, 46, 47]. In this case, all four degrees of freedom in the second doublet correspond to new particles, meaning that there are five total scalars from the two Higgs doublets - h (light CP-even Higgs), H (heavy CP-even Higgs), A (heavy CP-odd Higgs), and H^\pm (charged Higgs). The model is parameterized by two main parameters. The first, $\tan \beta \equiv \frac{v_2}{v_1}$, is the ratio of the vacuum expectation values of the two Higgs doublets (where v_1 corresponds to the v in the SM Higgs model described above). The second parameter is α , a mixing angle between the heavy and light Higgs fields. Models are also often parameterized with $\cos(\beta - \alpha)$ rather than α directly. The limit where $\cos(\beta - \alpha) = 0$ is called the alignment limit, and it is in this limit that the light Higgs h has the same couplings as a Standard Model Higgs.

2HDM models are usually separated into two main types - Type I and Type II. In Type I models, the charged fermions only couple to the second Higgs doublet, leading to a fermiophobic light Higgs. In

Type II models, up-type quarks couple to the first doublet while down-type quarks couple to the second doublet. One specific realization of a Type II 2HDM is the Minimal Supersymmetric Standard Model (MSSM).

Resonant di-Higgs production in this model can proceed through decays of the heavy CP-even Higgs $H \rightarrow hh$. The branching ratio for $H \rightarrow hh$ depends on the model type as well as the values of $\tan \beta$ and $\cos \beta - \alpha$. Figure 1.11 shows the branching ratios as a function of the mass of the heavy scalar H for both Type I and Type II models. Depending on the type of model hh can be a substantial fraction of the decays of H .

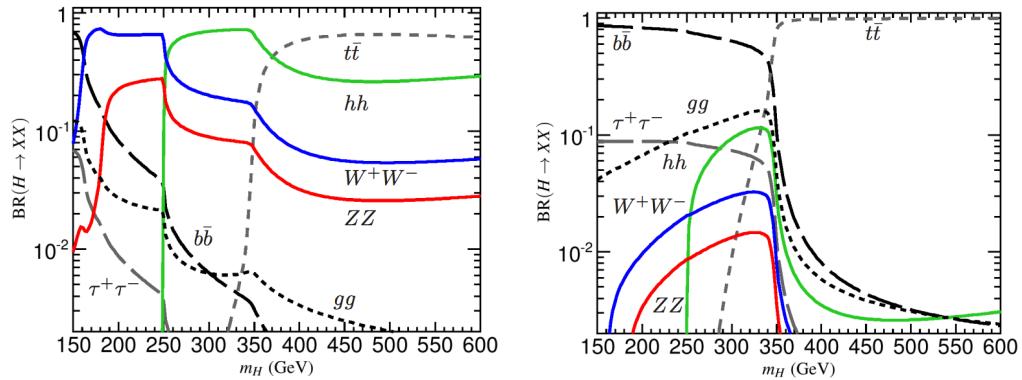


Figure 1.11: Branching ratios for heavy Higgs H in Type I (left) and Type II (right) 2HDM models with $\tan \beta = 1.5$ and $\cos(\beta - \alpha) = 0.1(0.01)$ for Type I (Type II). [5]

1.6 CONCLUSION

Studying the Higgs sector is essential for understanding the details of how mass arises in the Standard Model and how the electroweak symmetry is broken. The discovery of the Higgs boson also opens the door for its use as a tool to search for new physics, and Higgs pair production is an ideal candidate for this study. Even if no BSM physics is found in Higgs pair production, searches for Higgs pairs will put constraints on the Higgs self coupling and thus further knowledge of the Standard Model and the details of the Higgs potential.

This is some random quote to start off the chapter.

Firstname lastname

2

The ATLAS detector and the Large Hadron Collider

This chapter presents an overview of the experimental systems used to conduct the measurements presented in this thesis. First, a brief overview of the accelerator, the Large Hadron Collider, will be given. In this section, the accelerator conditions relevant to data-taking are presented as well. Next, an overview of the ATLAS experiment is given. The basics of each sub-detector's role are summarized, as well as the details of the datasets accumulated. Then, a brief interlude on the ATLAS Muon New Small Wheel upgrade is presented. While this new detector does not have a direct impact on any of the datasets taken so far, it will have an impact on future analyses and the work done on it is briefly summarized here. Finally, an overview of object reconstruction in ATLAS is given. While the details of all of the algorithms will not be presented in detail, aspects of the reconstruction performance such as object resolutions are shown as these are relevant to the two studies presented later in this thesis.

2.1 THE LARGE HADRON COLLIDER

The Large Hadron Collider (LHC) is a proton-proton collider at the CERN laboratory in Geneva, Switzerland[49]. It is designed for a maximum collision center of mass energy of $\sqrt{s} = 14 \text{ TeV}$ and has a circumference of 26.7 kilometers. Four main experiments are located at the interaction points (IP) of the accelerator: ATLAS (A Toroidal LHC ApparatuS), CMS (the Compact Muon Solenoid), ALICE (A Large Ion Collider Experiment), and LHCb [7, 50–52]. The studies performed in this thesis were all completed with the ATLAS detector.

Figure 2.1 shows a schematic of the LHC ring and the various experiments.

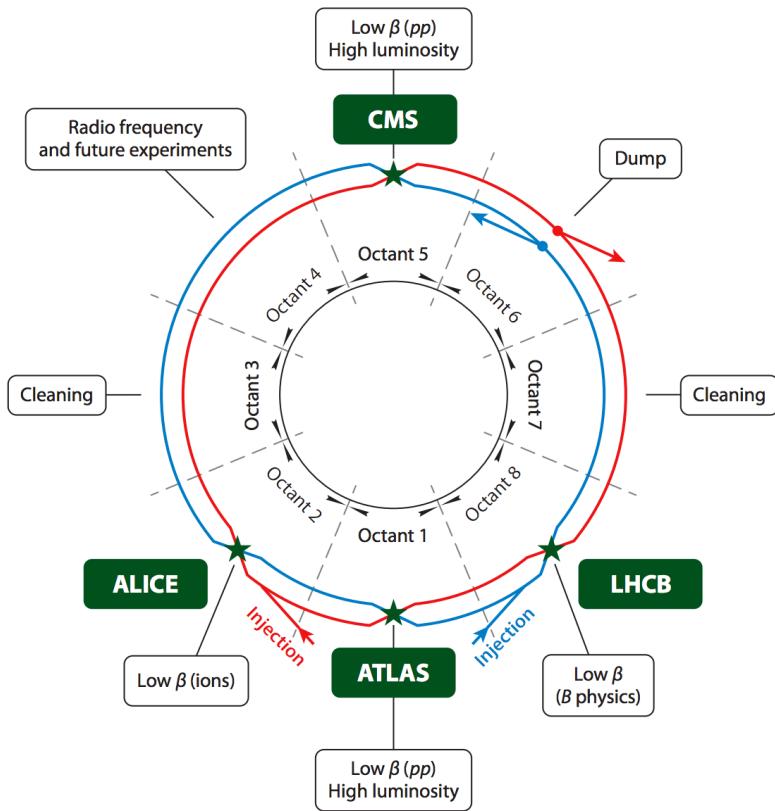


Figure 2.1: A schematic view of the LHC ring [6]

One of the most interesting features of the LHC is in its magnet design. Because the tunnel does not have room for separate superconducting magnets for each of the beam pipes, the LHC employs a twin-bore magnet design. Each magnet must hold an 8.3 Tesla magnetic field in order to bend the proton

beams at $\sqrt{s} = 14$ TeV. The superconducting magnets are cooled to a temperature of 1.9 Kelvin with superfluid helium.

2.1.1 INSTANTANEOUS LUMINOSITY

The rate of physics events expected from the accelerator is dependent on the instantaneous luminosity of the machine and the cross section of the physics process, $R_{\text{events}} = L\sigma$. Here, R_{events} is the number of events per second, L is the instantaneous luminosity of the machine, and σ is the cross section for the physics process being measured. The instantaneous luminosity of the LHC is determined by numerous factors related to machine conditions. Equation 2.1 gives the equation for instantaneous luminosity of Gaussian beam profile [6].

$$L = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi \epsilon_n \beta^*} F \quad (2.1)$$

The LHC collides protons in bunches, and in the above equation N_b is the number of protons per bunch while n_b is the number of bunches per beam. Nominally, the LHC can hold up to 2808 proton bunches. f_{rev} is the revolution frequency. ϵ_n is the normalized transverse beam emittance, a measurement of the average spread of the particles position-momentum space which has the dimension of length. β^* is the value of the *beta* function for the beam at the interaction point. It relates the emittance to the Gaussian width of the beam with $\sigma_{\text{beam}} = \sqrt{\epsilon \cdot \beta}$. F is a reduction factor that corrects for the fact that the beams are colliding at an angle at the IP.

Another way of writing the instantaneous luminosity is shown in equation 2.2. In this case, the instantaneous luminosity is written as the ratio of the rate of inelastic collisions with the inelastic cross section[53].

$$L = \frac{R_{\text{inel}}}{\sigma_{\text{inel}}} = \frac{\mu n_b f_{\text{rev}}}{\sigma_{\text{inel}}} \quad (2.2)$$

In this case, μ is the average number of interactions per bunch crossing in the accelerator. μ is a useful parameter for characterizing the amount of activity recorded in an experiment. As the instantaneous

luminosity and thus μ increase, there are more interactions per bunch crossing and more activity in the detector. This is often characterized with $\langle \mu \rangle$, the measured per bunch crossing μ value averaged over all bunch crossings. The interactions inside each bunch crossing that are not the main physics process of interest are often referred to as “pileup” interactions, and $\langle \mu \rangle$ is a measurement of the level of pileup in the detector.

2.1.2 EVOLUTION OF MACHINE CONDITIONS

This thesis uses datasets taken at three different center of mass energies: $\sqrt{s} = 7$ TeV data taken in the year 2011, $\sqrt{s} = 8$ TeV data taken in the year 2012, and $\sqrt{s} = 13$ TeV dataa taken in the year 2015. In addition to increasing center of mass energy, the instananeous luminosity and parameters that determine it were evolving. Table 2.1 summarizes that machine conditions in each of these datasets.

	2011	2012	2015	Design
\sqrt{s} [TeV]	7	8	13	14
Number of bunches	1380	1380	1825	2808
Max. protons per bunch	1.45×10^{11}	1.7×10^{11}		1.15×10^{11}
Bunch spacing [ns]	50	50	25	25
Max. instantaneous luminosity [$\text{cm}^{-2}\text{s}^{-1}$]	3.7×10^{33}	7.7×10^{33}	5×10^{33}	10^{34}
β^* [m]	1.0	0.6	0.8	0.55
$\langle \mu \rangle$	11.6	20.7	13.7	-

Table 2.1: Evolution of LHC machine conditions [25, 26]

2.2 THE ATLAS DETECTOR

The ATLAS detector is a multi-purpose particle detector experiment at the LHC’s Point 1 [7]. It has nearly 4π coverage in solid angle around the interaction point. It consists of an inner detector for measuring charged particles, electromagnetic and hadronic calorimeters, and a muon spectrometer. Figure 2.2 gives an overview of the detector.

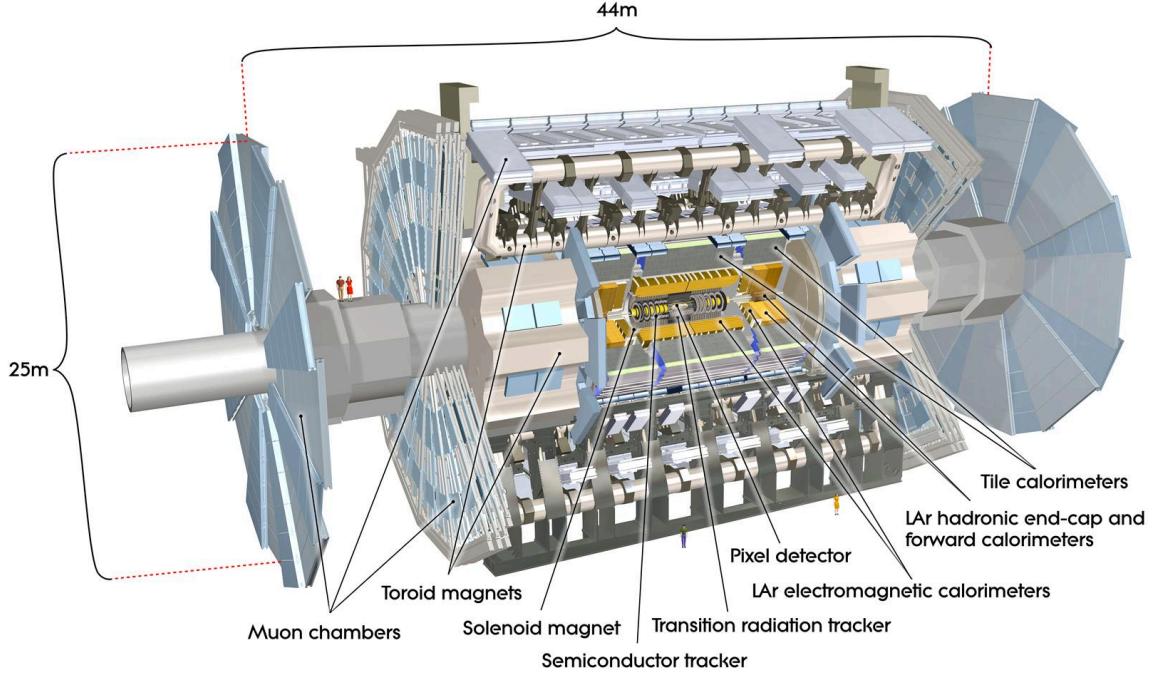


Figure 2.2: A full diagram of the ATLAS detector [7]

2.2.1 COORDINATE SYSTEM

Before defining the properties of the individual detectors, it is important to establish the coordinate system used. Figure 2.3 shows a schematic of the coordinate system. The azimuthal plane (perpendicular to the beam line) is defined as the x - y plane. The angle in this plane is referred to as ϕ . The angle relative to the beam axis is referred to as θ . Rather than using θ directly as a coordinate, the experiment often uses the pseudorapidity η . η is defined in equation 2.3.

$$\eta = \ln \left(\tan \left(\frac{\theta}{2} \right) \right) \quad (2.3)$$

Pseudorapidity is the massless approximation of rapidity, the angle used to parameterize boosts in special relativity. This is important for two reasons. First, it means that differences in η are Lorentz invariant. Second, particle production is roughly constant in pseudorapidity. Particles with η close to zero are referred to as “central”, while those at high $|\eta|$ are called “forward”. In general, two main detector

topologies can be seen in figure 2.2. There are “barrel” elements, which surround the beam line cylindrically and are in the central region of the detector. In the forward region, there are “endcap” regions which are arranged as disks perpendicular to the beam line.

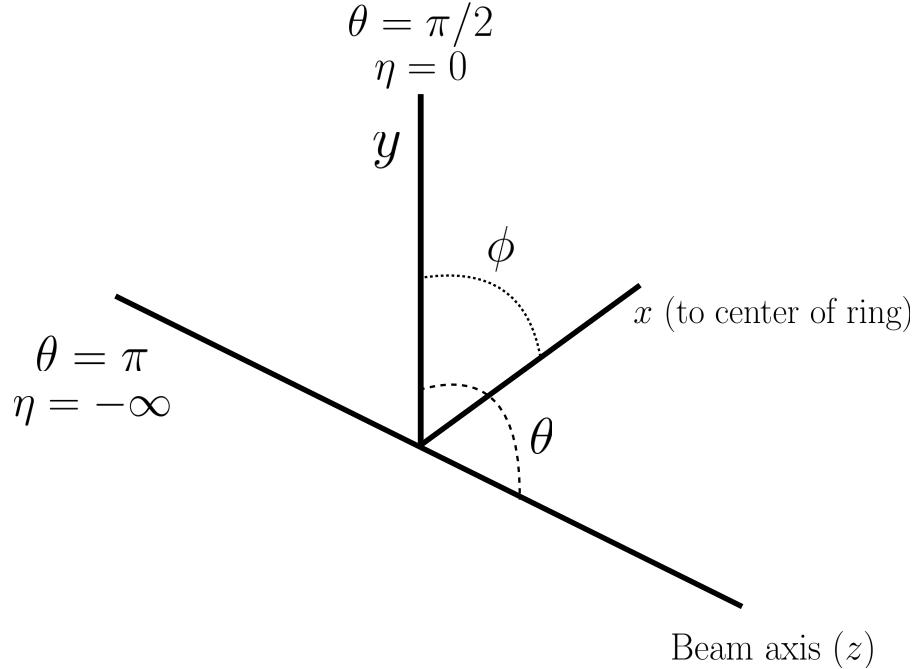


Figure 2.3: The ATLAS coordinate system

2.2.2 INNER DETECTOR

The ATLAS Inner Detector (ID) system is built for precision tracking of charged particles. It covers the range $|\eta| < 2.5$. In this range, approximately 1000 particles are generated every bunch crossing in the detector. This requires having fine granularity to achieve the resolutions required for good momentum measurement and vertex reconstruction.

The ID consists of three sub-components: the pixel detector, semiconductor tracker (SCT), and transition radiation tracker (TRT). It is surrounded by a solenoid providing a 2 T axial magnetic field which bends particles in the transverse plane to allow for momentum measurement. Figure 2.4 shows the layout of each of these components.

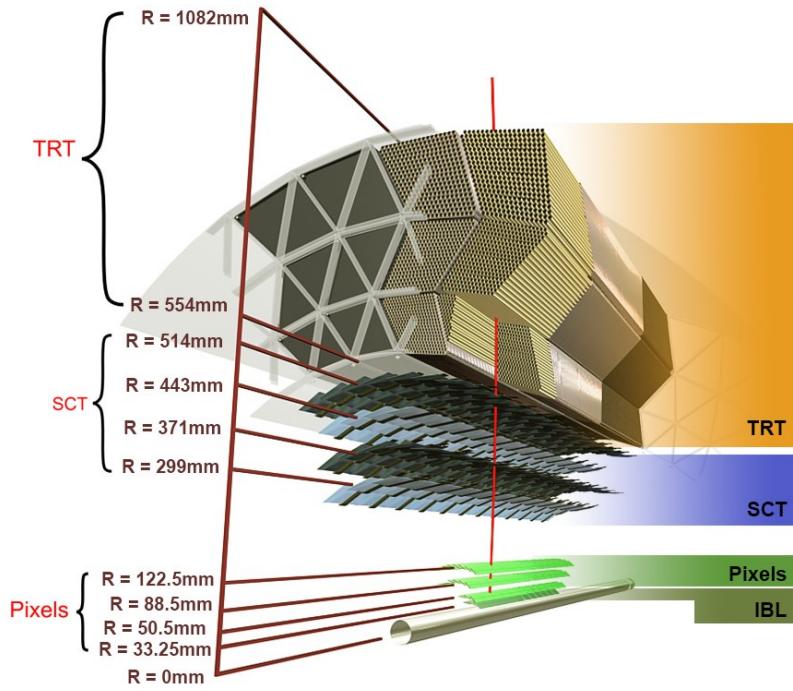


Figure 2.4: Layout of the ATLAS Inner Detector system [8]

PIXEL DETECTOR

The pixel detector is the first detector particles traverse after being generated in proton collisions and is the most granular detector. Its operation is crucial for precision tracking and vertex reconstruction as well as higher level object reconstruction like tagging of jets from b -quarks. The basic sensing element in this subdetector is a silicon pixel detector. The operating principle for the silicon pixels is that of a p - n junction. When a charged particle passes through, it creates electron-hole pairs that are then separated by the electric field. The sensors are $250\ \mu\text{m}$ thick and use oxygenated n -type wafers with readout pixels on the n^+ side of the detector [7]. Overall, the pixel detector has 1744 sensors and 80.4 million readout channels.

In the barrel region, the pixel detector has three concentric layers of sensors surrounding the beamline. In the endcap region, it consists of disks perpendicular to the beam axis. The detector is segmented in the R - ϕ plane and in z . Usually, three pixel layers are crossed by a charged particle track. The intrinsic accuracies of the sensors are $10\ \mu\text{m}$ in R - ϕ and $115\ \mu\text{m}$ in z (or R for the endcap).

INSERTABLE B-LAYER

In Run 2, a new innermost pixel layer, known as the insertable B-layer (IBL), was added to the Inner Detector [54]. This layer was added to cope with the higher luminosities planned in LHC Run 2 and at the high luminosity HL-LHC. Additionally it improves tracking position resolution which in turn improves the vertexing and b -tagging capabilities in ATLAS. The detector sits directly on a new beam pipe, only 33.25 mm away from the collision points in the azimuthal plane.

SEMICONDUCTOR TRACKER (SCT)

The semiconductor tracker (SCT) consists of silicon microstrips and comprises the next four layers of the ID. This sub-detector has 6.4cm long sensors that are daisy-chained into strips with a strip pitch of $80\ \mu\text{m}$ [7]. Some of the strips have a small stereo angle to allow for measurement of both angular coordinates. In total there are 6.3 million readout channels. The intrinsic accuracies are $17\ \mu\text{m}$ in $R\text{-}\phi$ and $580\ \mu\text{m}$ in z (or R in the endcap).

TRANSITION RADIATION TRACKER (TRT)

The transition radiation tracker (TRT) serves two purposes. First, it consists of 4mm diameter straw tubes filled with a 70/27/3% gas mixture of xenon, carbon dioxide, and oxygen to provide tracking of charged particles. Particles typically have 36 TRT straw tube hits per track. The material in between the straws is designed to induce transition radiation which can be useful for particle identification. As particles pass between media with different dielectric constants, they emit transition radiation that can cause additional showers in the TRT. In particular it is useful for discrimination between electrons and pions or other charged hadrons, as the amount of transition radiation is proportional to the Lorentz factor of the particle.

2.2.3 CALORIMETERS

The calorimeter system consists of two main sub-components: a fine granularity electromagnetic calorimeter tailored for the measurement of photons and electrons and multiple coarser hadronic calorimeters

dedicated to the measurement of hadronic showers [7]. The calorimeter system has broader coverage than the inner detector, covering the region out to $|\eta| < 4.9$. It is also designed to deliver good containment of showers so as to limit leakage into the muon system. Figure 2.5 shows the layout of the calorimeter system.

Both the electromagnetic and hadronic calorimeters are sampling calorimeters. They alternate active material for energy measurement with passive material for energy absorption. The materials used for each purpose vary based on the type of calorimeter and its location in the detector.

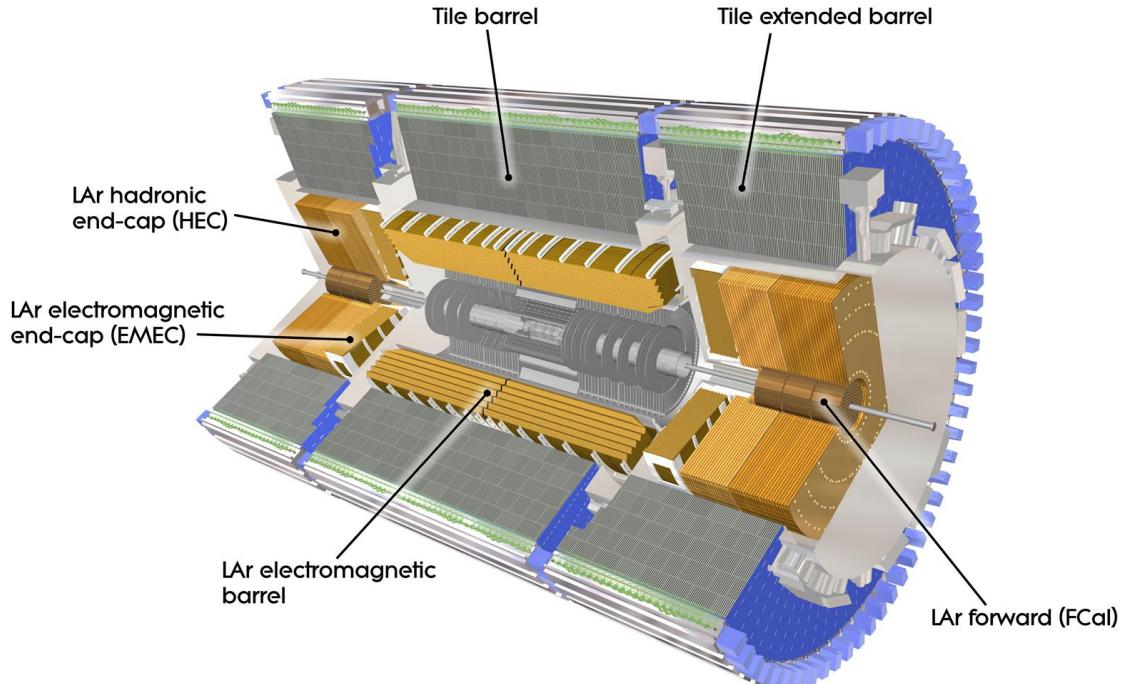


Figure 2.5: Layout of the ATLAS calorimeter system [7]

ELECTROMAGNETIC CALORIMETER

The electromagnetic calorimeter (EM calorimeter) use liquid Argon (LAr) as its active material and lead as its passive material. It is arrange in an accordion geometry to increase the absorption area while still allowing it to have no azimuthal cracks (complete symmetry in ϕ). The EM calorimeter is divided into a

barrel portion that extends to $|\eta| < 1.475$ and an endcap portion going from $1.375 < |\eta| < 3.2$. The region where these two units overlap is called the “transition region”.

In order to provide good containment the calorimeter depth must be optimized. Typically, for electromagnetic calorimeters the depth is measured in radiation lengths. In general, the intensity of a particle beam attenuates exponentially in distance with a constant equal to the radiation length. That is, $I(x) = I_0 e^{-x/X_0}$, where I is the intensity, x is the distance traveled, and X_0 is the radiation length. The ATLAS EM calorimeter is designed to have > 22 radiation lengths in the barrel and > 24 in the endcap [7].

HADRONIC CALORIMETERS

There are three types of hadronic calorimeters present in ATLAS: the tile calorimeter (TileCal), hadronic endcap (HEC), and forward calorimeter (FCal). Each one is optimized for stopping of hadronic showers and the materials chosen are specific to their placement in the detector.

The TileCal is a scintillating tile calorimeter placed directly outside the EM calorimeter. It uses steel as the absorber and plastic scintillator tiles as the active material. It has coverage in the barrel at $|\eta| < 1.0$ and in the “extended barrel” region of $0.8 < |\eta| < 1.7$.

The HEC had two wheels perpendicular to the beam line per endcap and is located directly behind the EM calorimeter endcap modules. The HEC covers the region from $1.5 < |\eta| < 3.2$, overlapping slightly with both the tile calorimeter and the forward calorimeter. Like the EM calorimeter, it uses liquid Argon as the active material, but it uses copper as the absorber.

The FCal covers the most forward regions of the calorimeter system, extending to the region of $3.1 < |\eta| < 4.9$. It again uses liquid argon as its active material. For absorber, it consists of an innermost module made of copper followed by a module made of tungsten.

The hadronic equivalent of radiation length is called the interaction length and is denoted as λ . In the barrel, the hadronic calorimeter depth is approximately 9.7λ , while in the endcap is is 10λ . The outer supports contribute an additional 1.3λ . This is been shown to be sufficient to limit punch-through of showers to the muon system [7].

2.2.4 MUON SPECTROMETER

The muon spectrometer is dedicated to measuring the momentum and position of muons. It consists of tracking and trigger chambers which are unique in the barrel and endcap regions. The magnetic field for bending of muons is provided by a system of three large air-core toroid magnets (from which ATLAS derives its name.) These magnets provide 1.5 to 5.5 Tm of bending power at $0 < |\eta| < 1.4$ and approximately 1 to 7.5 Tm in the endcap region of $1.6 < |\eta| < 2.7$. The entire muon system covers the range $0 < |\eta| < 2.7$. Monitored drift tubes (MDTs) are used for tracking in the barrel and the two outer layers of the endcap, while cathode strip chambers (CSCs) are used to provide tracking in the innermost endcap wheel. In the barrel, resistive plate chambers (RPCs) are used as trigger chambers while thin gap chambers (TGCs) are used in the endcap. Figure 2.6 shows the layout of the ATLAS muon system. The entire muon system is designed with the specification of providing a 10% momentum resolution for a 1 TeV muon.

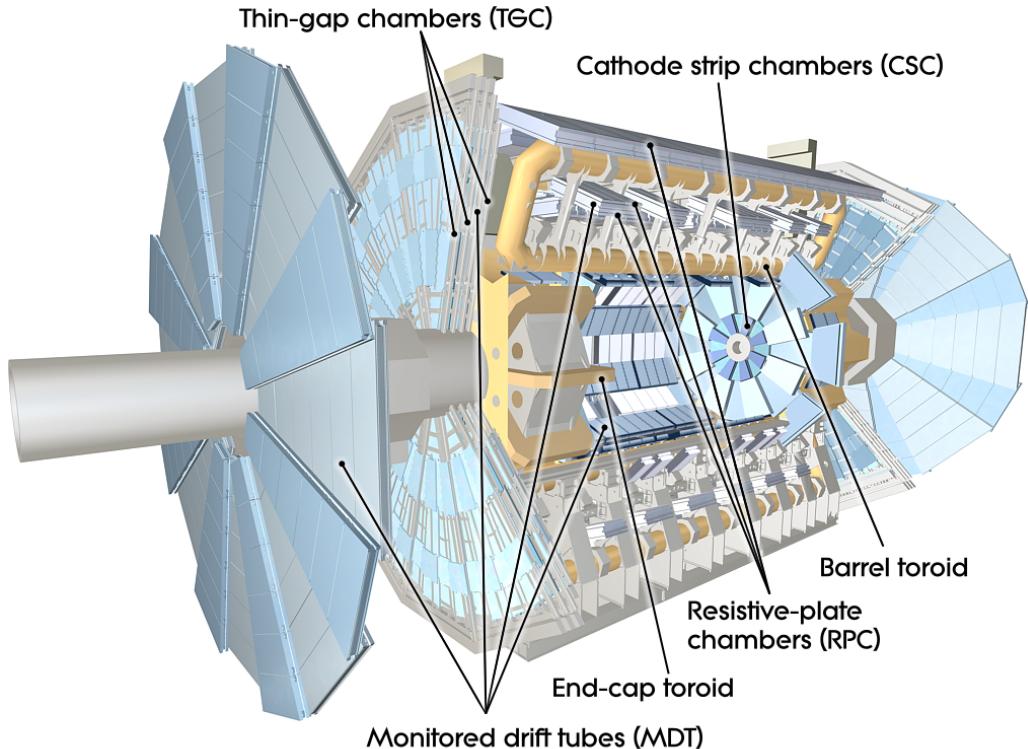


Figure 2.6: Layout of the ATLAS muon system [7]

MONITORED DRIFT TUBES (MDTs)

The monitored drift tubes (MDTs) are aluminum 3cm diameter tubes filled with a 93/7 % mixture of Argon and CO₂, with trace amounts of water. As a charged particle traverses the tube, it ionizes the gas and the ions drift to a wire at the center of the tube. The radial distance of traversal of the particle in the tube is determined by the drift time of the electrons, allowing for fine position resolution. The tubes have an average resolution of 80 μm per tube and a maximum drift time of approximately 700ns. The tubes are oriented so that they give precision measurement in η and run along ϕ . They cover $|\eta| < 2.7$, except in the innermost layer of the endcap where they only go to $|\eta| < 2.0$ [7].

CATHODE STRIP CHAMBERS (CSCs)

The cathode strip chambers cover a narrow window of the innermost endcap region at $2.0 < |\eta| < 2.7$. In this region the background rates in the cavern are particularly high and the CSCs are designed to handle these higher rates. The CSCs are multiwire proportional chambers with wires pointing in the radial direction (away from the beam pipe). The wire serves as an anode and there are two types of segmented cathode strip, one perpendicular to the wires which gives the precision measurement and one parallel which provides the transverse coordinate. It has an 80/20 gas mixture of Argon and CO₂ [7].

RESISTIVE PLATE CHAMBERS (RPCs)

The resistive plate chambers (RPCs) are gaseous electrode-plate detectors covering the region $|\eta| < 1.05$. They consist of two resistive plates separated by a distance of 2 mm. The gas mixture used is a 94.7/5/0.3% mixture of C₂H₂F₄, Iso-C₄H₁₀, and SF₆. It has readout strips with a pitch of 23-35 mm for both η and ϕ measurement and thus provides measurement of the azimuthal coordinate in the barrel that the MDTs do not. The thin gas gap allows for a quick response time which makes it ideal for use in the trigger. There are three layers of RPCs which are referred to as the three trigger stations. They allow for both a low p_T and high p_T trigger. The coincidence of hits in the innermost chambers allows for triggering of muons between 6 and 9 GeV, while the outermost layer allows the trigger to select high momentum tracks in the range of 9 to 35 GeV [7].

THIN GAP CHAMBERS (TGCs)

The thin gap chambers (TGCs) are multiwire proportional chambers where the wire to cathode distance (1.4mm) is smaller than the wire-to-wire distance (1.8 mm). They contain a gas mixture of CO₂ and *n*-pentane and use a hih electric field to gain good time resolution. They serve two functions in the end-cap system. First, they serve as the trigger chambers. Second, they also provide azimuthal coordinate measurement which the MDTs do not. They sit on the inner and middle layers of the endcap. The outermost layer's azimuthal coordinate is determined by extrapolation [7].

2.2.5 MAGNET SYSTEM

As mentioned previously, there are two independent magnet systems in ATLAS. The first is a 2 T solenoid field in the inner detector which provides bending in the azimuthal plane. The second is an approximately 0.5 T toroidal field in the muon system which provides bending in η . Figure 2.7 shows the predicted field integral as a function of $|\eta|$ [7].

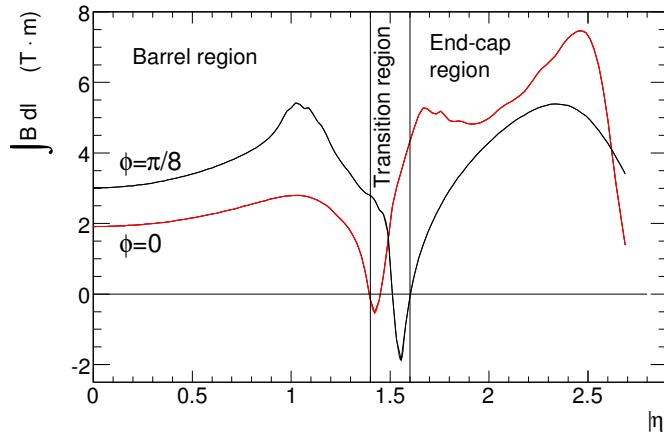


Figure 2.7: Predicted field integral as a function of $|\eta|$ for the ATLAS magnet system [7]

2.2.6 TRIGGER SYSTEM

The ATLAS trigger system searches for signatures of muons, electrons, photons, hadronically decaying τ leptons, and jets in order to save these events for further analysis. The trigger system in ATLAS

is designed to reduce the maximum LHC event rate of 40 MHz to a more reasonable rate that can be recorded. The trigger first consists of a fast, hardware based system called the Level-1 (L1) trigger. The L1 trigger consists of independent dedicated detector sub-components that can seed regions of interest (RoIs) for further analysis downstream. For muons, the RPCs and TGCs are used, while in the calorimeter coarsely grained sections of calorimeter cells called towers are used. Once regions of interest are seeded, a software based system called the High Level Trigger (HLT) is used to reconstruct objects and integrate information from different parts of the detector. In Run 1 of ATLAS, the HLT consisted of two separate stages: the level 2 (L2) trigger and the event filter (EF).

The maximum trigger rate that the L1 trigger can handle is 75 kHz. In the HLT, the rate of events written to disk is approximately 200 Hz. Figure 2.8 shows the trigger rates for different L1 triggers in 2012 and 2015 for ATLAS [9].

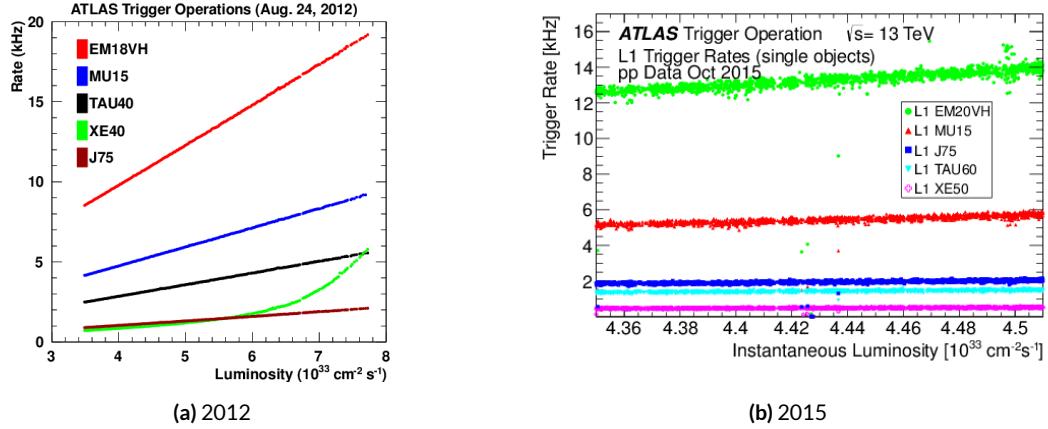


Figure 2.8: ATLAS trigger rates for Level-1 triggers as a function of instantaneous luminosity in 2012 and 2015 operation. These are single object triggers for electromagnetic clusters (EM), muons (MU), jets (J), missing energy (XE), and τ leptons (TAU). The threshold of the trigger is given in the name in GeV. [9]

2.2.7 ATLAS DATASETS

ATLAS has collected data at center of mass energies of 7, 8, and 13 TeV. Figure 2.9 shows the integrated luminosity as a function of time for each of the three collected datasets. At $\sqrt{s} = 7 \text{ TeV}$, ATLAS recorded 5.08 fb^{-1} . Increased instantaneous luminosity in 2012 led to a larger dataset of 21.3 fb^{-1} recorded at $\sqrt{s} = 8 \text{ TeV}$. After Long Shutdown 1 (LS1) of the LHC and a restart in 2015, ATLAS

recorded 3.9 fb^{-1} of data at $\sqrt{s} = 13 \text{ TeV}$. [10, 11]

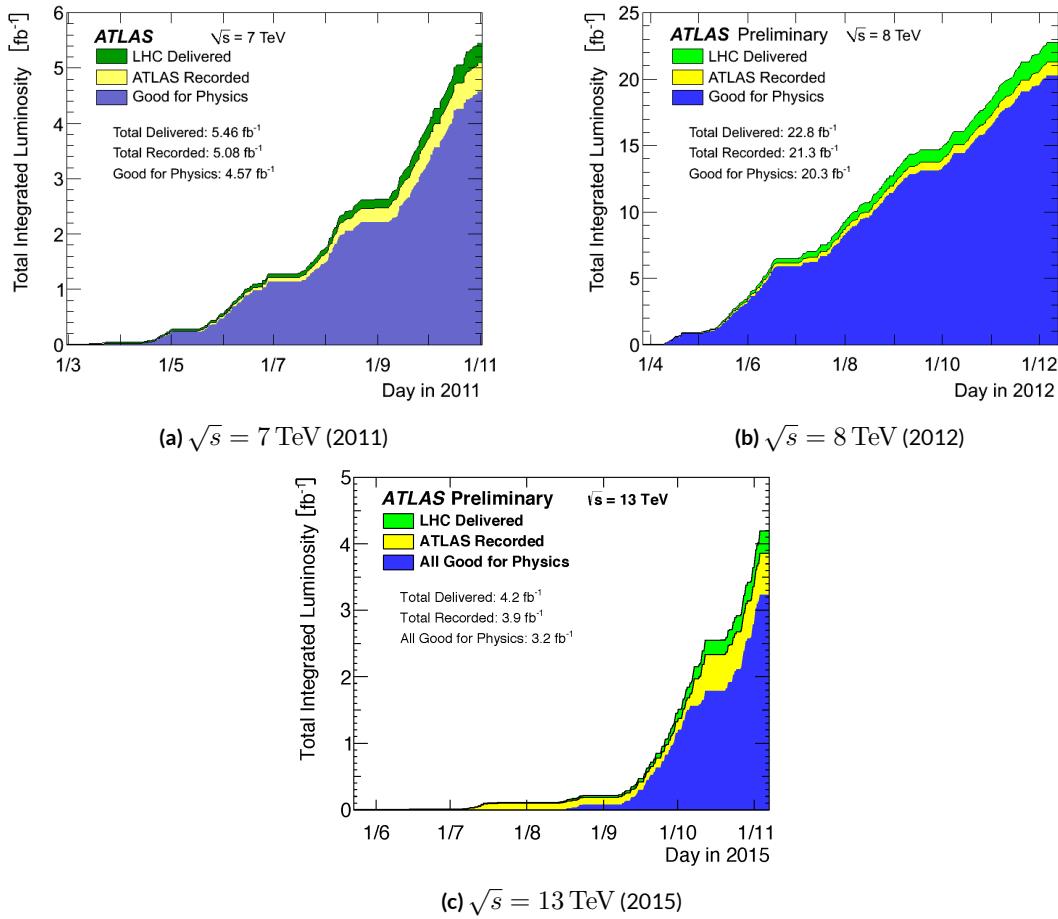


Figure 2.9: Instantaneous luminosity as a function of time for data recorded by ATLAS at different center of mass energies [10, 11]

2.2.8 DETECTOR PERFORMANCE

Table 2.2 summarizes the design requirements for each of the different sub-detectors. This table shows the energy and momentum resolution of each tracking, calorimetry, and muon measurements.

2.3 THE ATLAS MUON NEW SMALL WHEEL UPGRADE

As the LHC continues operation, it is scheduled to be upgraded in several phases to allow it to reach higher instantaneous luminosities and thus collect larger datasets. These conditions will open new doors

	Required resolution
Tracking	$\sigma_{p_T}/p_T = 0.05\% p_T \oplus 1\%$
EM calorimetry	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$
Hadronic calorimetry	
Barrel and end-cap	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$
Forward	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$
Muon spectrometer	σ_{p_T}/p_T at $p_T = 1$ TeV

Table 2.2: Performance requirements for the ATLAS detector [7].

for study of rare physics processes but will also present interesting challenges that must be faced. ATLAS will require new detector technologies to cope with the increased background rates in the cavern in these high luminosity conditions. One such upgrade, scheduled to be installed during Long Shutdown 2 (LS2) of the LHC in 2018, is the ATLAS Muon New Small Wheel (NSW) upgrade [12]. The NSW will replace the innermost end-cap wheel of the muon system with new technologies, as this is the part of the muon detector closest to the beam and thus suffers from the highest rates.

2.3.1 MOTIVATION

The motivation of the NSW is two-fold. First, the objective is to alleviate the decreased tracking efficiency that comes in a high rate environment. As figure 2.10, at the LHC design luminosity both the efficiency of recording hits and reconstructing track segments in the MDTs decreases at the LHC design luminosity.

Second, the NSW will work to alleviate the rate of fake triggers arising in the endcap. Figure 2.11 shows the extrapolated trigger rates as a function of the p_T threshold with and without the NSW upgrade. As the figure shows, the NSW upgrade will reduce the trigger rate by an order of magnitude compared to the current endcap trigger system.

2.3.2 NSW DETECTOR TECHNOLOGIES

The NSW will use two new detector technologies - micromesh gaseous structure detectors (micromegas) and small-strip thin gap chambers (sTGCs) [12, 55]. Unlike the previous detectors, both of these detector technologies can be used for tracking or trigger. However, the micromegas is more suited to tracking

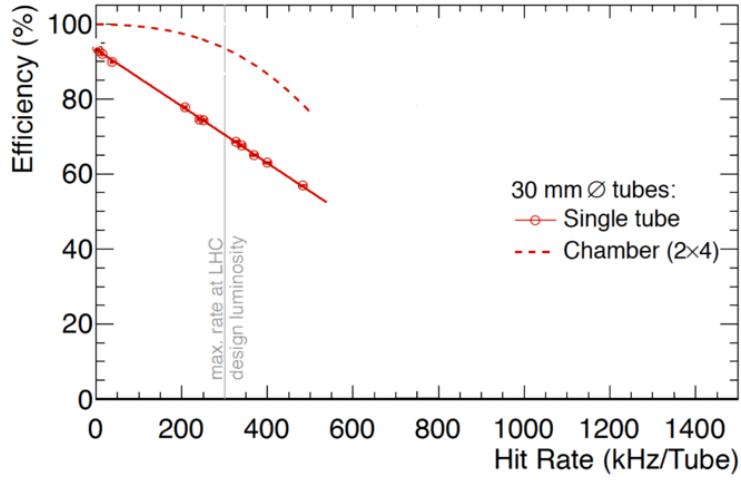


Figure 2.10: MDT tube hit (solid) and segment (dashed) efficiency as a function of hit rate per tube [12]

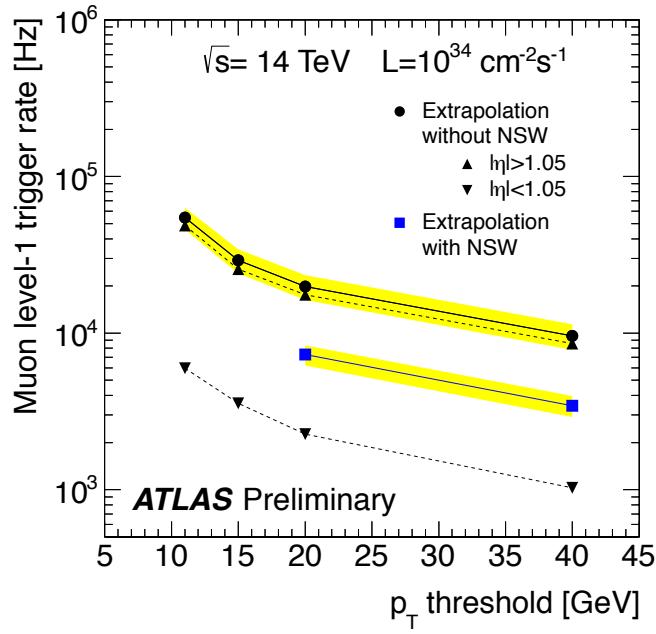


Figure 2.11: Trigger rate as a function of p_T threshold with and without the NSW upgrade [12]

because of its good spatial resolution, while the sTGCs have better time resolution and are more suited for the trigger. To maintain a fully redundant system, both technologies are used for both purposes.

MICROMEGAS

Micromegas detectors operate using a thin metallic mesh that sits approximately $100\ \mu\text{m}$ away from the readout electrodes to create the amplification region. Above this mesh, there is a drift region on the order of a few mm in length capped by a drift electrode. As a charged particle traverses the detector, it ionizes gas and the electrons drift down towards readout strips. The timing of the drift can be used to reconstruct the angle of traversal of the particle. This is illustrated in figure 2.12. The micromegas used in ATLAS will be resistive micromegas, where the readout electrodes are topped with resistive strips [56]. This alleviates the risk of sparking in the large area detectors that ATLAS will use.

In ATLAS, the micromegas drift gap will be 5 mm and the amplification gap will be $128\ \mu\text{m}$. They are filled with the same gas mixture as the MDTs. They will be stacked in an octuplet in an XXUV-UVXX geometry, where X refers to straight strips and U and V refer to stereo strips at an angle of $\pm 1.5^\circ$. This arrangement allows for measurement of the azimuthal coordinate and gives a large lever arm between the straight strips for triggering purposes. Figure 2.12 shows the geometry of a single micromegas detector as well as its operating principle [12].

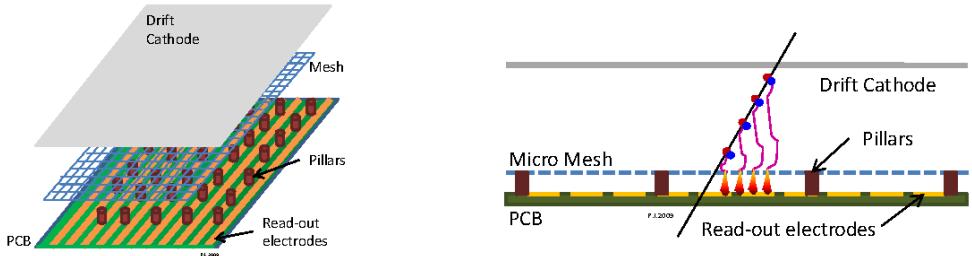


Figure 2.12: Illustrations of the geometry (left) and operating principle (right) of the micromegas detector [12]

sTGCs

The sTGCs are similar to the TGCs already described. They consist of gold-plated tungsten wires with a 1.8 mm pitch between two cathode planes 1.4 mm away from the wire plane. One cathode plane consists of strips with a 3.2 mm pitch (much smaller pitch than the TGCs), while the other consists of coarser pads that are used for defining regions of interest in the sTGC trigger algorithm. Figure 2.13 shows the basic detector geometry.

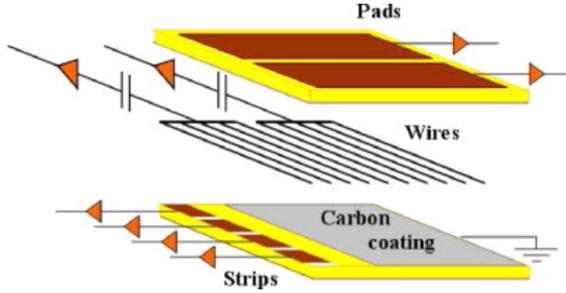


Figure 2.13: Geometry of the sTGC detector [12]

2.3.3 PHYSICS IMPACT

Maintaining low p_T thresholds for muons while still staying within the trigger rate budget at Level 1 (20 kHz) for the muon system is crucial for physics analyses to be successful in high luminosity conditions. One realm where the lepton trigger threshold is especially important is in Higgs physics. In the $H \rightarrow WW^*$ analysis, one of the W bosons is off shell and tends to decay to soft leptons. In associated production of a Higgs with a W , the lepton is also important because the lepton provides the main handle which allows the event to be triggered. Table 2.3 shows the impact of increasing the trigger thresholds on these analyses. It shows that either raising the threshold or using only the barrel both have significant impacts on the signal efficiency. With the NSW, the signal efficiency is largely maintained and the triggers can be unprescaled.

Threshold	$H \rightarrow b\bar{b}$ (%)	$H \rightarrow WW^*$ (%)
$p_T > 20$ GeV	93	94
$p_T > 40$ GeV	61	75
$p_T > 20$ GeV (barrel only)	43	72
$p_T > 20$ GeV (with NSW)	90	92

Table 2.3: Signal efficiencies for WH production with $H \rightarrow b\bar{b}$ and $H \rightarrow WW^* \rightarrow \mu\nu qq$ under different trigger configurations [12].

2.4 OBJECT RECONSTRUCTION IN ATLAS

ATLAS analyses first start by requiring the presence of certain reconstructed physics objects in the event. This section will present a brief overview of the algorithms used to reconstruct electrons, muons, jets

(including b -jets), and missing energy^{*}. The performance of object reconstruction and measurement will also be discussed as these are relevant to the analyses presented later. Figure 2.14 gives an overview of the different sub-detectors that each type of particle will interact with in ATLAS.

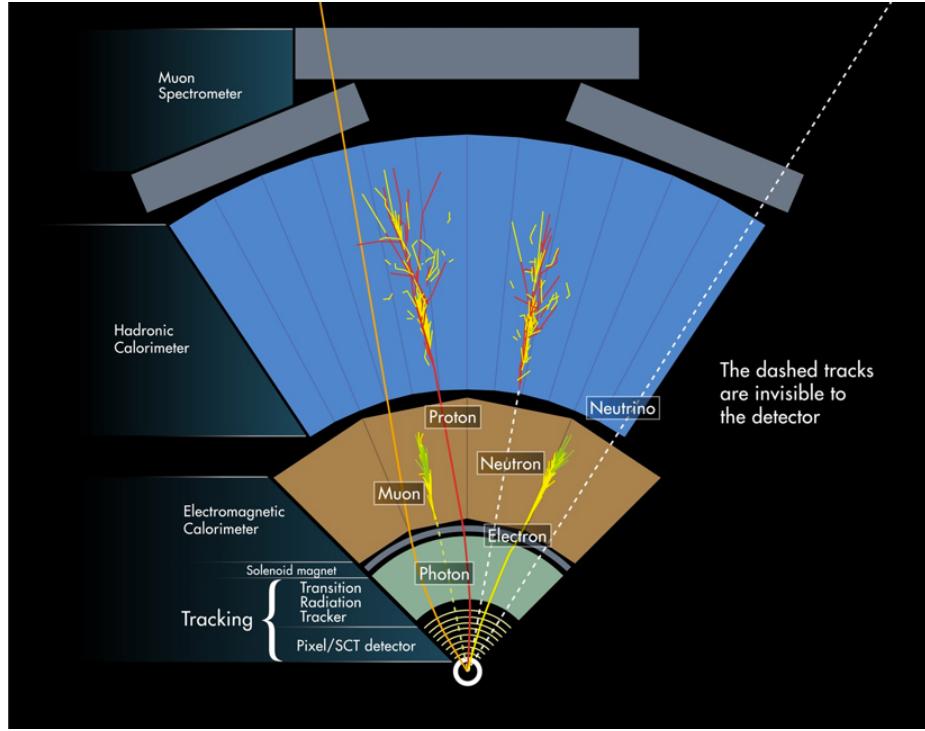


Figure 2.14: Illustration of particle interactions in ATLAS [13]

2.4.1 ELECTRONS

Electrons in ATLAS will leave tracks in the inner detector and energy deposits in the electromagnetic calorimeter. The algorithm for recognizing the signature of electrons proceeds in two steps: reconstruction and identification.

In reconstruction, an electron candidate is formed by matching EM calorimeter deposits with ID tracks. The algorithm first chooses seed clusters in the EM calorimeter by using a sliding window algorithm that searches for towers with transverse energy larger than 2.5 GeV. In addition to seed clusters, track candidates must be identified in the ID. The algorithm selects seed tracks with $p_T > 1$ GeV that

^{*}Reconstruction algorithms for other objects, such as photons and τ leptons, are not detailed here as these objects are not used in the presented studies.

do not fit well with a pion hypothesis. Once candidate tracks are selected, they are re-fit with a Gaussian Sum Filter (GSF) algorithm to estimate electron parameters [57]. Finally, an electron candidate is formed if at least one track matches to a seed cluster in the calorimeter. The full details of the reconstruction algorithm can be found in reference [14].

Once an electron candidate is present, identification criteria must be applied in order to reject fake electrons from background. Many different variables are used for this identification, most of them related to the shower shape in the EM calorimeter and the amount of leakage into the hadronic calorimeter, as well as information from the ID and in particular the TRT. There are both cut-based and likelihood-based criteria that range from “loose” to “very tight”. For details, see reference [14].

Figure 2.15 shows the algorithm’s reconstruction efficiency of true electrons for different identification criteria as well as the electron energy resolution in simulation [14, 15]. The reconstruction efficiency is measured using both Z and J/ψ tag and probe techniques.

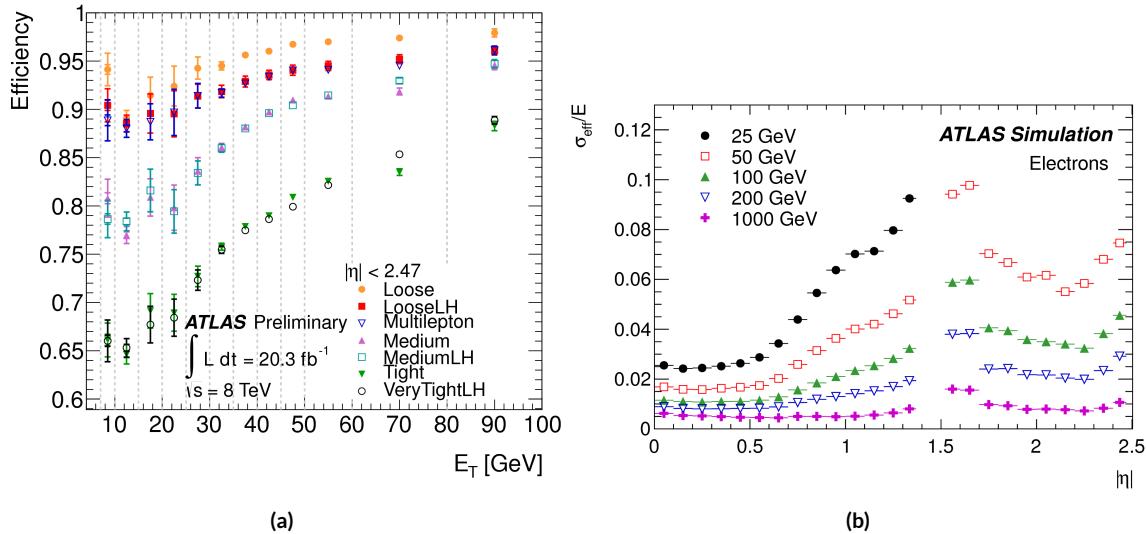


Figure 2.15: Electron performance: (a) reconstruction efficiency as a function of electron E_T [14] (b) energy resolution in simulation as a function of $|\eta|$ for different energy electrons [15]

2.4.2 MUONS

The ATLAS detector is designed to stop most particles before they reach the muon spectrometer. Muons, however, are minimum ionizing particles, meaning that they will not lose a significant amount of energy

through interactions with the detector and will thus pass through. Therefore, the muon reconstruction works to match tracks in the muon spectrometer with tracks in the inner detector.

The first step of reconstruction is to reconstruct local straight line tracks, called segments, in each muon chamber. Segments are then fit to larger tracks that traverse the entire muon spectrometer. Such muon tracks are referred to as “standalone” tracks (SA) as they only use information from the muon spectrometer. The standalone tracks are then matched to tracks in the inner detector to form “combined” (CB) muons, where the combined ID and MS fit are used to determine the momentum and direction of the muon. To improve acceptance, segment-tagged and calorimeter-tagged muons are also reconstructed. In these cases, ID tracks are matched to segments in the MS and calorimeter deposits consistent with a minimum ionizing particle, respectively. The details of the reconstruction can be found in reference [16].

As with electrons, once muon candidates are reconstructed they have identification criteria applied to reduce background. These criteria include the χ^2 match between the ID and MS tracks, the number of hits in the ID, overall ID and MS track fit quality, and additional variables [16]. The criteria range from “loose” to “tight” as with electrons.

Figure 2.16 shows the muon reconstruction efficiency (measured with Z and J/ψ tag and probe) and invariant mass resolution [16].

2.4.3 JETS

When a quark or gluon is produced in collisions, it is not measured directly in ATLAS. Rather, due to QCD effects, it produces a collimated spray of hadrons in the direction of the original parton, which is known as a jet. Jets are reconstructed in ATLAS using energy deposits in the hadronic calorimeter. The first step is build “topological clusters” out of energy deposits in calorimeter cells [58, 59]. This is done using strategy where seed cells are chosen by picking cells whose energy measurements are four times the amount of noise expected for that cell. Adjacent cells with at least 2σ energy measurements are added to the cluster, then a final layer of clusters with energy above 0σ are added. Once calorimeter clusters are formed, they are clustered further into jet candidates using the anti- k_T jet clustering algorithm [60].

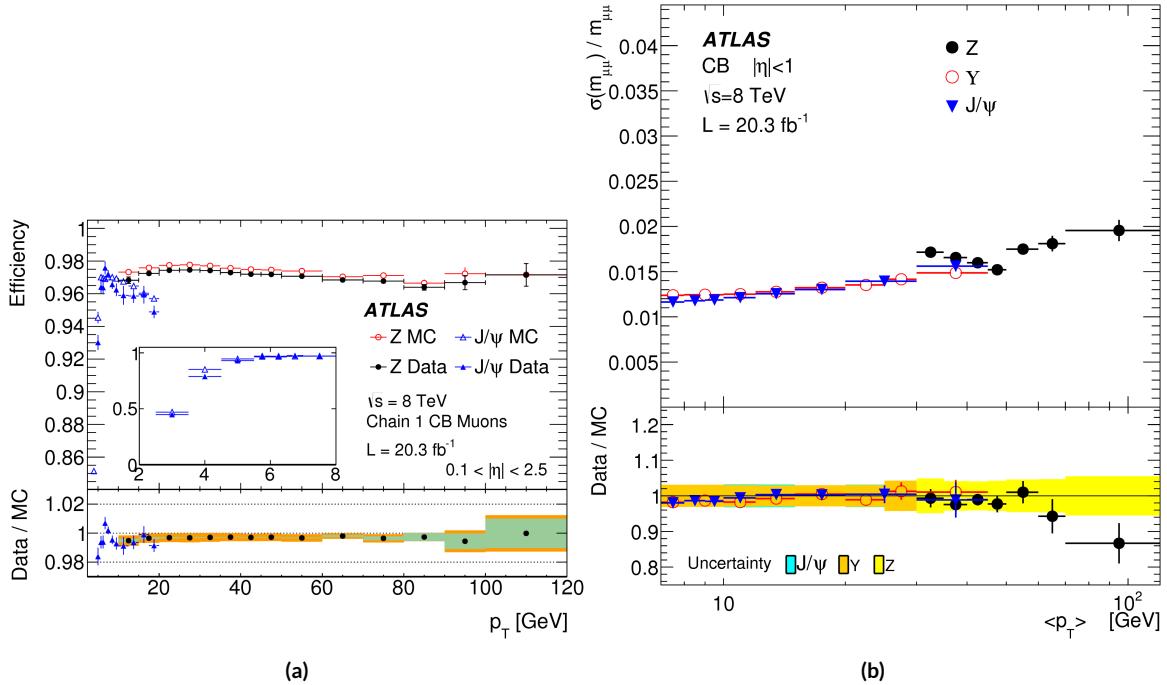


Figure 2.16: Muon performance: (a) reconstruction efficiency as a function of muon p_T (b) dimuon mass resolution as a function of average p_T [16]

This algorithm uses a parameter R that appears in the denominator of the clustering distance metric and defines the radial size of the jet in η - ϕ space.

The energy response of the calorimeter must be properly characterized in order to reconstruct jet energy. Calorimeter clusters can be calibrated either with the EM calibration, where each cluster is assumed to have come from the energy deposit of an electron or photon, or the LCW calibration, where local cluster weights are computed to allow for local calibration of clusters as hadronic or electromagnetic. The details of the jet energy calibration are not detailed here and are discussed in reference [17].

Figure 2.17 shows the jet energy response after calibration in Monte Carlo as a function of the true p_T of the jet [17].

2.4.4 b -TAGGING

One important aspect of jet physics is the task of identifying the flavor of parton that produced the measured jet. While in general this is very difficult, jets from b -quarks offer an interesting case where such identification is possible. B mesons have a lifetime on the order of 10^{-12} seconds, which makes a $c\tau$

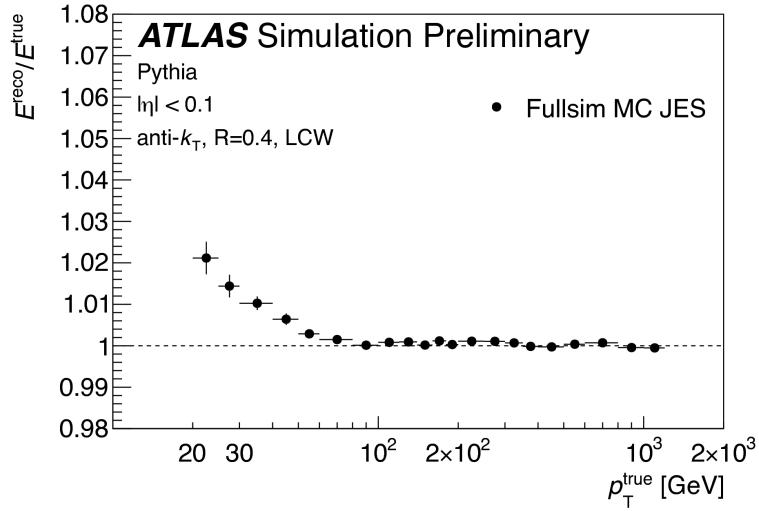


Figure 2.17: Jet energy response after calibration as a function of true p_T in simulation [17]

on the order of millimeters [1]. This type of displaced decay vertex can be identified in detectors like ATLAS and allows b -jets to be distinguished from other flavors of jets[†].

ATLAS uses a multivariate machine learning algorithm to identify jets from b -quarks. The inputs to this algorithm are determined from lower level reconstruction algorithms. There are three distinct algorithms that reconstruct variables which are used as input to the multivariate technique.

The first family is referred to as IPxD (where the x can either be 2 or 3). These algorithms use the transverse and longitudinal impact parameters d_0 and z_0 of the tracks inside a jet to determine their consistency with the primary vertex. They two or three dimensional (hence the x) templates for light flavor, charm, and bottom jets and then evaluate the likelihood of the jet coming from each of these types. The likelihood ratios are used as inputs to the multivariate algorithm.

The next two algorithms used as input are referred to as the secondary vertex (SV) and JetFitter (JF) algorithms. The SV algorithm uses tracks inside the jet to fit for vertices that are displaced from the primary vertex. The JF algorithm attempts to reconstruct the full flight path of the b by looking for multiple displaced vertices along the same line (as B decays often result in subsequent c decays).

In Run 1, the multivariate b -tagging algorithm used a neural network and was referred to as MV1.

[†]Jets from charm quarks can also be detected in this way but they do not live quite as long so the displacement of the vertex is harder to distinguish

The details of this algorithm and its inputs are given in reference [18]. In Run 2, the number of inputs was simplified and a boosted decision tree with 24 input variables was used, referred to as MV2. The details of this algorithm are in reference [19]. Figure 2.18 shows the performance of each of these algorithms.

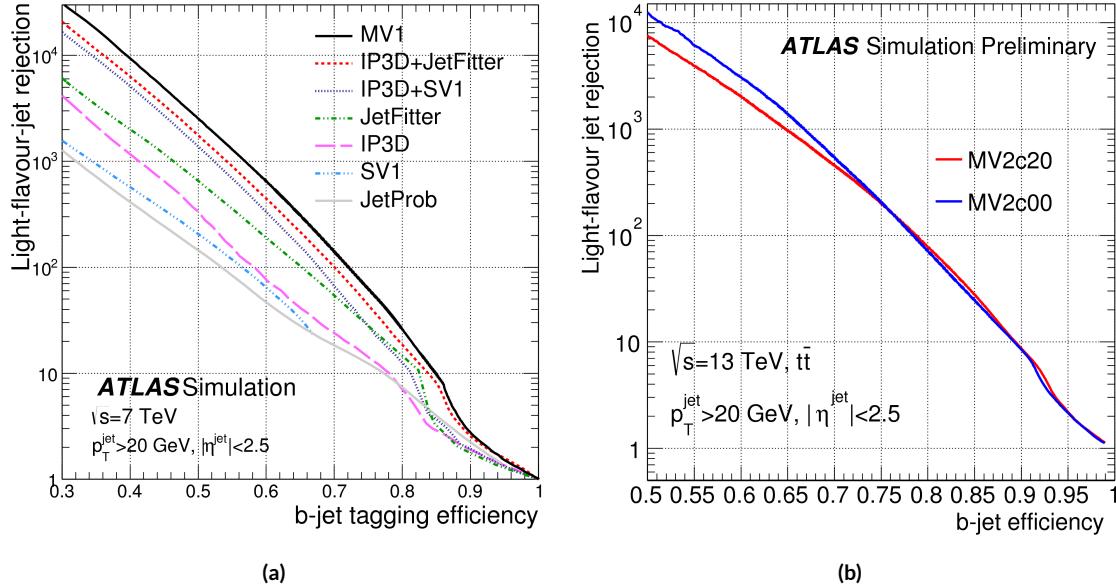


Figure 2.18: Light jet rejection ($1/\text{efficiency}$) vs. b -jet efficiency for MV1 and its input algorithms (a) [18] and MV2 (b) [19] in simulated $t\bar{t}$ events. The numbers in the algorithm names in (b) refer to the fraction of charm events used in the MV2 training.

2.4.5 MISSING TRANSVERSE ENERGY

As noted in figure 2.14, neutrinos produced in ATLAS will pass through the detector without interacting. The only way of detecting the presence of particles like neutrinos (or BSM particles that are long-lived) is to use missing transverse momentum. The basic principle of missing transverse energy is to use the momentum balance of the incoming protons to infer the presence of missing particles. The net longitudinal momentum of the incoming partons that collide is not known (since each carries an unknown fraction of the proton's momentum). However, the protons (and thus incoming partons) have no net momentum in the plane transverse to the beam line (the x - y) plane. Therefore, if there are no un-measured particles in the final state, the transverse momenta of all of the ifnal state particles should

balance. The magnitude of this imbalance is known as missing transverse momentum.

Part II

Observation and measurement of Higgs
boson decays to WW^* in LHC Run I at
 $\sqrt{s} = 7$ and 8 TeV

*Basic research is what I am doing when I don't know
what I am doing.*

Wernher von Braun

3

$H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ Analysis Strategy

3.1 INTRODUCTION

This chapter presents an overview of the strategy for searching for a Higgs boson in the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ decay topology. Its purpose is to define in broad terms how the search and measurement are undertaken, before going into details on the specific sub-categories within the larger analysis. First, the properties of the Higgs signal are discussed and the associated backgrounds are presented. Next, the observables used to enhance the signal to background ratio are defined. Finally, the parameters of interest in the search and measurement will be shown, along with a brief overview of the statistical treatment of the final Higgs candidates.

Following this chapter, the results of three different studies within the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ channel are shown. Chapter 4 presents a search for Higgs boson production in gluon fusion mode and the role of the $H \rightarrow WW^*$ channel in its discovery. Chapter 5 shows the search and first observation in ATLAS of the Vector Boson Fusion (VBF) production mode of the Higgs in the $H \rightarrow WW^*$ decay channel.

Finally, chapter 6 shows the combined Run 1 $H \rightarrow WW^*$ results for the measurement of the Higgs cross section and relative coupling strengths to other SM particles.

3.2 THE $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ SIGNAL IN ATLAS

The signal studied in this and subsequent chapters is the Higgs boson in the WW^* final state, where each W boson subsequently decays into a charged lepton and a neutrino. In its simplest decay path, the final state consists of two neutrinos and two charged leptons, each of which can be either an electron or a muon. If one or both of the W s decay to τ leptons, only leptonic decays of the τ are considered. This decay path produces additional neutrinos in the final state but still gives two charged leptons as before. Neutrinos are not detected in ATLAS, so the final state ultimately consists of two reconstructed leptons and missing transverse momentum (denoted as E_T^{miss}). Final states where both of the charged leptons are electrons or muons are referred to as the “same flavor” ($ee/\mu\mu$) final states, while those with one electron and one muon are referred to as “different flavor” ($e\mu$ or μe).

While the basic final state consists of two leptons and E_T^{miss} , there can be additional objects depending on the production mode of the Higgs. As described in detail in Chapter 1, if the Higgs is produced via vector boson fusion production, there will be two additional forward jets in the event. Even in gluon fusion, one or more jets can be produced through initial state radiation from the incoming gluons. Because of the varying background composition as a function of jet multiplicity, each bin in this variable has its own dedicated requirements applied in the search and measurement. The $n_j = 0$ and $n_j = 1$ bins are dedicated to gluon fusion production, while the $n_j \geq 2$ bin has separate dedicated searches for ggF and VBF production.

Figure 3.1 shows the relative branching fractions for the $H \rightarrow WW^*$ process, calculated from the Particle Data Group values for the W and τ branching ratios[?]. The largest branching ratio is both W bosons decaying to quark pairs at 45.44%. The next largest is one W decaying leptonically and the other decaying to quarks, a branching ratio of 34.18%. In all cases, ℓ denotes either an electron or muon, and the leptonic branching ratios of the τ are included. For example, the $\ell\nu qq$ final state includes one W decaying to $e\nu$, $\mu\nu$, or $\tau\nu$. In the case of the $W \rightarrow \tau\nu$ decay, the τ lepton then decays to an electron or

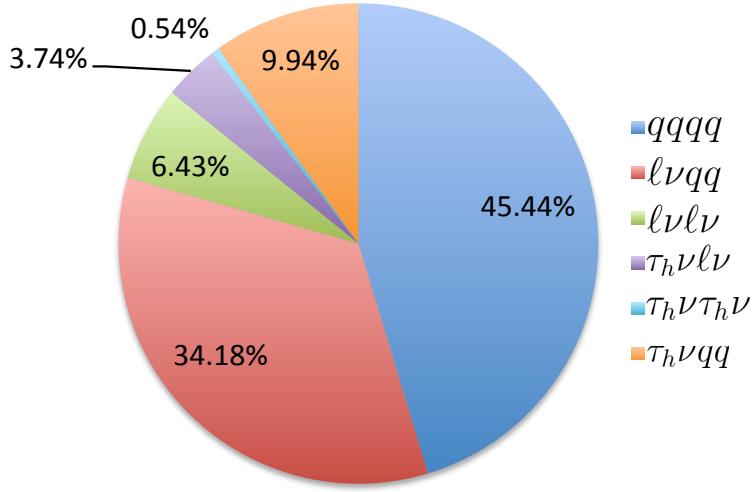


Figure 3.1: Branching ratios for a WW system. q refers to quarks. ℓ can be either an electron or muon, and the leptonic branching ratios of the τ are included. For example, the $\ell\nu qq$ final state includes one W decaying to $e\nu$, $\mu\nu$, or $\tau\nu$. τ_h refer to hadronic decays of the τ .

muon via $\tau \rightarrow \nu_\tau \ell \nu_\ell$. Final states with a τ_h refer to hadronic decays of the τ . The branching ratio to the $\ell\nu\ell\nu$ final state is 6.43%.

While the $\ell\nu\ell\nu$ final state is not a large fraction of the branching ratio, there are significant advantages in this channel. First, both the $qqqq$ and $\ell\nu qq$ channels suffer from a large QCD multijet background, which is often difficult to model. Second, events in the the $\ell\nu\ell\nu$ channel in data can be triggered more efficiently due to the presence of two leptons.

3.3 BACKGROUND PROCESSES

Many processes from the Standard Model can also produce a final state with two leptons and missing transverse momentum . This section lists the dominant backgrounds to Higgs production. It gives general descriptions of how the backgrounds mimic Higgs production and how they can be reduced. Table 3.1 summarizes the different processes.

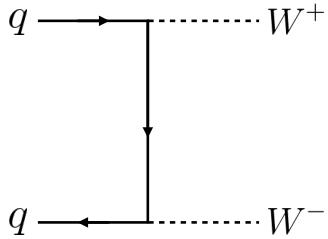


Figure 3.2: Feynman diagram for Standard Model WW production

3.3.1 STANDARD MODEL WW PRODUCTION

Non-resonant Standard Model diboson production, as shown in figure 3.2, is an irreducible background to Higgs boson production in the WW final state. It produces the same exact final state objects, namely leptonically decaying W bosons. There are no additional objects in the final state that allow for background reduction. Therefore the analysis solely relies on the correlations between the leptons to reduce this background.

3.3.2 TOP QUARK PRODUCTION

Production of top quarks, either in pairs ($t\bar{t}$ production) or singly (e.g. Wt production), can also mimic Higgs production. Because top quarks decay via $t \rightarrow Wb$, top pair production can produce a final state with two W bosons that then decay leptonically. In this case, however, there are two additional jets from the bottom quarks in the final state. This allows the analysis to veto on the presence of jets identified as originating from a b in order to reduce the size of the background.

Single top production can occur via s -channel, t -channel, or associated production (Wt). The mode which most closely resembles the Higgs final state is Wt . In this case, there are two real W bosons produced, as with $t\bar{t}$. However, the decay of the single top quark will still also produce one b -jet, meaning a b veto will reduce this background as well.

Figure 3.3 shows the Feynman diagrams for $t\bar{t}$ and Wt production.

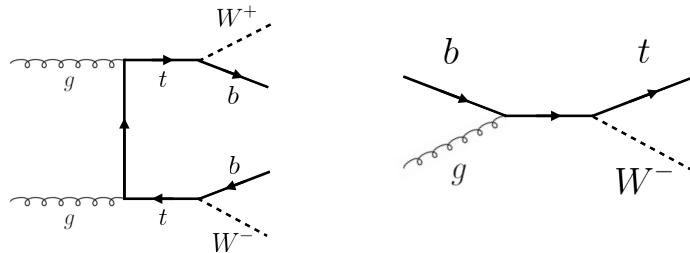


Figure 3.3: Feynman diagrams for top pair production (left) and Wt production (right)

3.3.3 W +JETS BACKGROUND

Single W boson production, in association with jets, is a unique background. The other background considered so far have all included real leptons in the final state. In this case, however, only one real lepton from the decay of a W exists in the final state. The second reconstructed lepton can arise from two different cases. First, the lepton may truly be an algorithm “fake”, or a jet misidentified as a lepton by either the electron or muon reconstruction algorithms. Second, the lepton may be a real lepton but coming from semi-leptonic decays of particles inside the shower of the jet. This background can be reduced by requiring that the reconstructed lepton have little activity surrounding it in the calorimeter (also known as an “isolated” lepton). Figure 3.4 shows the Feynman diagram for W +jets production.

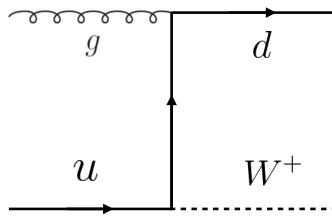


Figure 3.4: An example Feynman diagram of W +jets production

3.3.4 Z/γ^* +JETS BACKGROUND

Production of a Z/γ^* in association with jets (also known as Drell-Yan) is also a background to Higgs production. In particular, the same flavor final states have a large Z +jets background, as the Z decays into two leptons of the same flavor. (This background also enters the different flavor final state through

the leptonic decays of $Z \rightarrow \tau\tau$). Figure 3.5 shows the production of a Z in association with one jet. Because there are no neutrinos in this final state, variables like E_T^{miss} can be used to reduce the background.

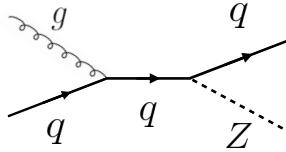


Figure 3.5: An example Feynman diagram of $Z + \text{jets}$ production

3.3.5 OTHER (SUBDOMINANT) BACKGROUNDS

There are additional processes which contribute to the background composition but are not produced as frequently as those listed already. The first of these are referred to as VV or “Other diboson” processes and include multiple Standard Model diboson processes, including WZ , ZZ , $W\gamma$, $W\gamma^*$, and $Z\gamma$ production. Additionally, there is background from QCD multijet production, where two jets are misidentified as leptons.

Category	Process	Description
SM WW	$WW \rightarrow \ell\nu\ell\nu$	Real leptons and neutrinos
Top quark production	$t\bar{t} \rightarrow WbWb \rightarrow \ell\nu b\bar{b}\nu\bar{b}$ $tW \rightarrow WbW \rightarrow \ell\nu\ell\nu b$ $t\bar{b}, t\bar{q}\bar{b}$	Real leptons, untagged b s Real leptons, untagged b Untagged b , jet misidentified as lepton
Drell-Yan	$Z/\gamma^* \rightarrow ee, \mu\mu$ $Z/\gamma^* \rightarrow \tau\tau \rightarrow \ell\nu\ell\nu\bar{\nu}$	“Fake” E_T^{miss} Real leptons and neutrinos
Other dibosons	$ZZ \rightarrow \ell\ell\nu\nu$ $W\gamma^*, WZ \rightarrow \ell\nu\ell\ell, ZZ \rightarrow \ell\ell\ell\ell$ $W\gamma, Z\gamma$	Real leptons and neutrinos Unreconstructed leptons γ reconstructed as e , unreconstructed lepton
$W + \text{jets}$	$Wj \rightarrow \ell\nu j$	Jet reconstructed as lepton
QCD multijet	jj	Jets reconstructed as leptons

Table 3.1: A summary of backgrounds to the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ signal

3.4 SHARED SIGNAL REGION SELECTION REQUIREMENTS

As presented in section 3.2, there are many different combinations of objects that can define a $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ final state. The multiplicity of jets and the flavor combinations of the leptons both lead to many potential signal regions. Additionally, signal regions can be optimized separately to be sensitive to the distinct production modes of the Higgs. Gluon fusion, vector boson fusion, and associated production of a Higgs all lead to unique final state topologies. Figure 3.6 delineates the different signal regions used in the gluon fusion and vector boson fusion $H \rightarrow WW^*$ analyses. While there are different optimizations possible in each signal region, there are also some commonly shared selections that will be described here.

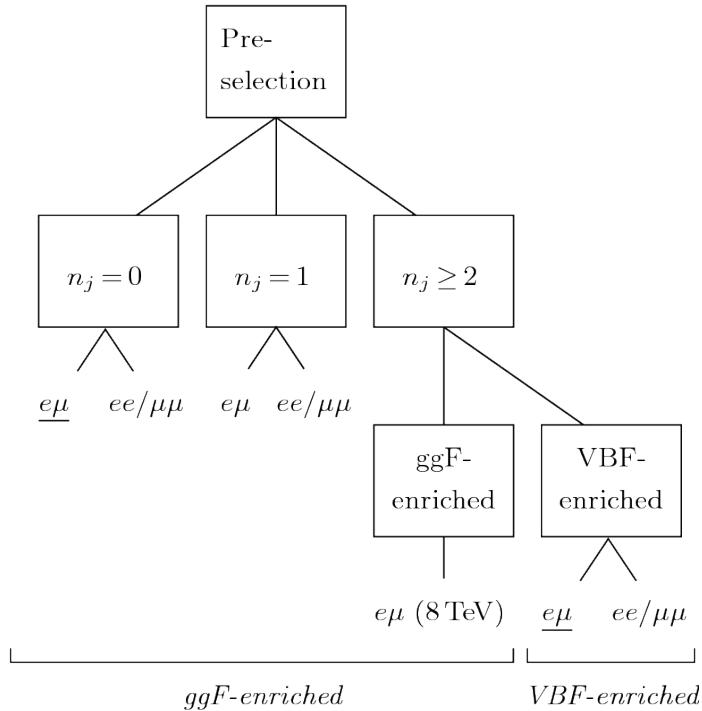


Figure 3.6: An illustration of the unique analysis signal regions[20]

3.4.1 EVENT PRE-SELECTION

Before being sorted into the distinct signal regions, basic requirements are applied on the reconstructed objects in the event to select Higgs-like event candidates. First, two oppositely charged leptons are re-

quired.

Once the leptons are selected, the last requirement for event pre-selection is the presence of neutrinos. As neutrinos cannot be detected directly in ATLAS, E_T^{miss} can be used as a proxy for the combined neutrino momentum in the transverse plane. In general, it is expected that the signal should have a harder E_T^{miss} spectrum than backgrounds, especially if those backgrounds did not contain neutrinos. One additional consideration when using E_T^{miss} is the fact that mis-measurements of objects in the detector can lead to imbalances in the transverse plane that are not due to real particles escaping the detector. One indicator that this is the case is that the E_T^{miss} vector in the transverse plane will be pointing in the same direction as the mis-measured object. Therefore, a new variable, $E_{T,\text{rel}}^{\text{miss}}$, is used in the pre-selection. $E_{T,\text{rel}}^{\text{miss}}$ is defined in equation 3.1.

$$E_{T,\text{rel}}^{\text{miss}} = \begin{cases} E_T^{\text{miss}} \sin \Delta\phi_{\text{near}} & \text{if } \Delta\phi_{\text{near}} < \pi/2 \\ E_T^{\text{miss}} & \text{otherwise,} \end{cases} \quad (3.1)$$

If the closest object to the E_T^{miss} vector is within $\pi/2$ radians in the transverse plane, the E_T^{miss} is projected away from this object. Otherwise, the normal E_T^{miss} vector is used. Figure 3.7 shows a graphical illustration of this concept.

Once both the lepton and E_T^{miss} pre-selections are made, the analysis can be divided into different regions according to jet multiplicity.

3.4.2 JET MULTIPLICITY

Jet multiplicity, denoted as n_j , is used to sub-divide the analysis into its distinct signal regions. The reason for this is twofold. First, different jet multiplicity bins will be more or less sensitive to different Higgs production modes. For example, the $n_j \geq 2$ region is more sensitive to VBF production because of the two high momentum jets produced at matrix element level. For gluon fusion production to enter this bin, two initial state radiation jets must be emitted. Second, background composition varies greatly in different bins of n_j . Figure 3.8 shows the jet multiplicity in both the different flavor and same flavor regions. It also shows the background composition in the bins of n_b . There are a few clear trends from

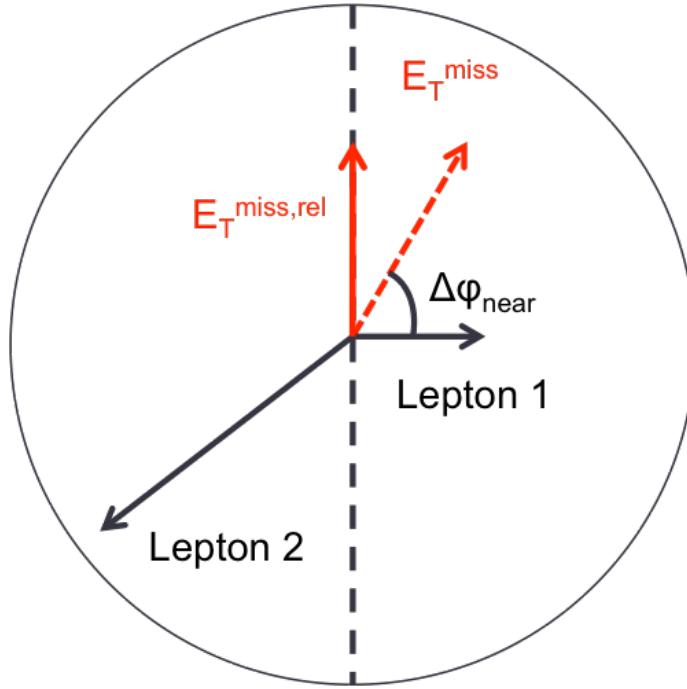


Figure 3.7: A graphical illustration of the $E_{T,\text{rel}}^{\text{miss}}$ calculation

this distribution. The first is that the Drell-Yan background dominates in the same flavor channels for $n_j \leq 1$. Second, the top background becomes a clear contributor to the total background for $n_j \geq 1$. Lastly, the SM WW production dominates in the $n_j = 0$ bin, as it is an irreducible background to $H \rightarrow WW^*$ production. Because of these distinct features, each jet multiplicity bin is treated separately.

3.5 BACKGROUND REDUCTION IN SAME-FLAVOR FINAL STATES

As described in section 3.4.2, the background composition of the same flavor final states is unique to that of the different flavor states. In particular, Drell Yan processes play a much larger role because the Z/γ^* decays to same flavor leptons. Because real neutrinos are absent in the Z/γ^* decays to ee and $\mu\mu$, a requirement on E_T^{miss} should largely reduce the background. However, as this section will demonstrate, with increasing pileup conditions the resolution of the calorimeter-based E_T^{miss} degrades greatly. Therefore, two new variables for Z/γ^* background reduction are constructed and described in this section.

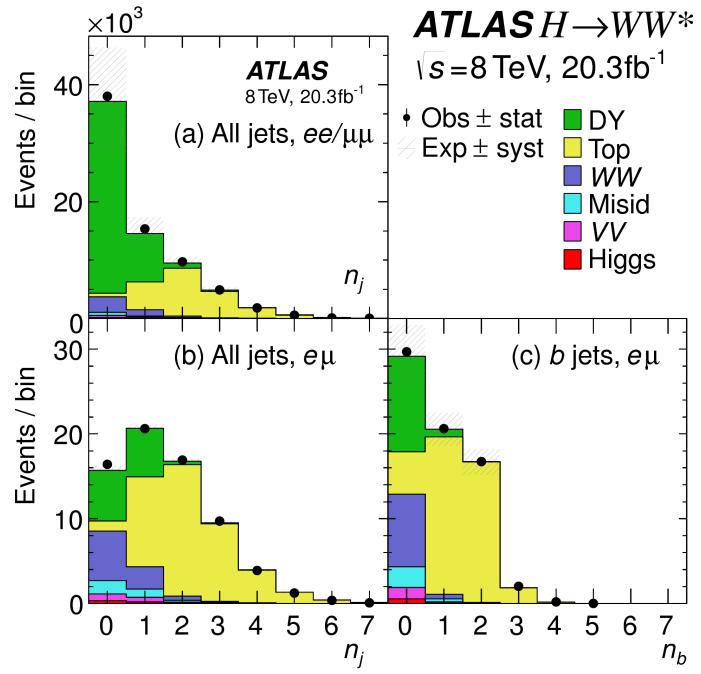


Figure 3.8: Predicted backgrounds (compared with data) as a function of n_j (a and b) and n_b (c)

3.5.1 PILEUP AND E_T^{miss} RESOLUTION

Secondary interactions of protons in the colliding bunches of the LHC (known as pileup interactions, described in detail in Chapter 2) deposit energy into the ATLAS calorimeter on top of the energy that comes from the hard scatter process that is being searched for or analyzed. The calculation of E_T^{miss} is fundamentally Poissonian, as summing up all of the energy deposits in individual calorimeter cells or clusters is similar to a counting experiment. Thus, the energy resolution scales as \sqrt{E} , just as the error on a mean of N in a Poisson distribution is \sqrt{N} . As more energy is deposited in the calorimeter, the E_T^{miss} resolution degrades, meaning that the E_T^{miss} resolution is particularly sensitive to LHC instantaneous luminosity conditions.

Figure 3.9 shows an event display of a $Z/\gamma^* + \text{jets}$ event candidate with the twenty-five reconstructed primary vertices. This display illustrates that while the interaction of interest only has tracks coming from the hardest primary vertex, all of the secondary interactions will deposit energy in the calorimeter as well.

Figure 3.10 shows the RMS of the E_T^{miss} distribution in $Z \rightarrow \mu\mu$ events (where there are no real neu-

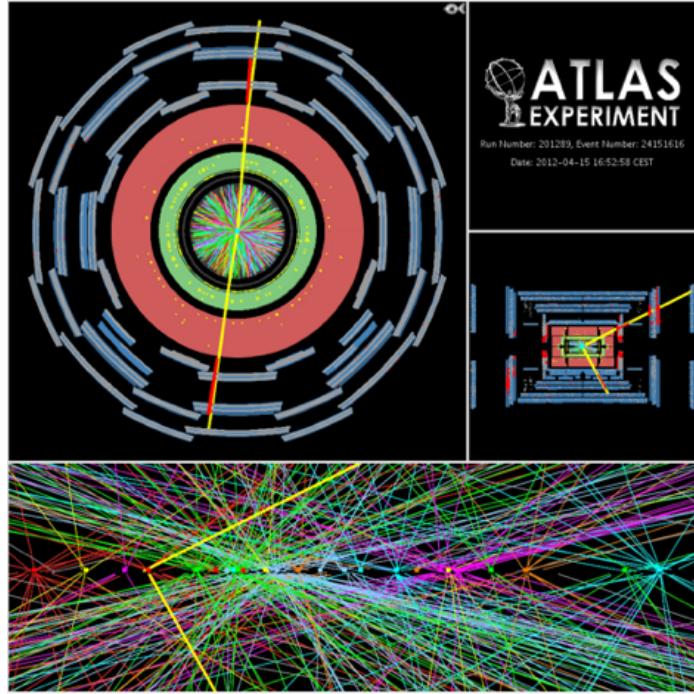


Figure 3.9: An event display of a Z/γ^* + jets event illustrating the effect of pileup interactions

trinos) as a function of the number of the average number of interactions. Under 2011 LHC conditions, this RMS was approximately 9 GeV, while under 2012 running conditions the resolution worsened to 12 GeV. This worsening dilutes the E_T^{miss} variable's ability to reduce the Z/γ^* background.

3.5.2 TRACK-BASED DEFINITIONS OF MISSING TRANSVERSE MOMENTUM

Because the increasing number of secondary proton-proton interactions degrades calorimeter-based E_T^{miss} resolution, a new variable using only contributions from the primary interaction vertex is necessary to further reduce the Z/γ^* background. While it is not possible to associate calorimeter energy deposits with a particular vertex, individual charged particle tracks in the Inner Detector are associated to unique vertices. Thus, two track-based definitions of missing transverse momentum , using only tracks coming from the primary vertex in the event, are used in the analysis. The simplest variable, $p_T^{\text{miss}(\text{trk})}$, is the vectorial sum of the p_T of all of the tracks from the primary vertex and the selected leptons (excluding the tracks associated with the selected leptons to avoid double counting). This is defined in equa-

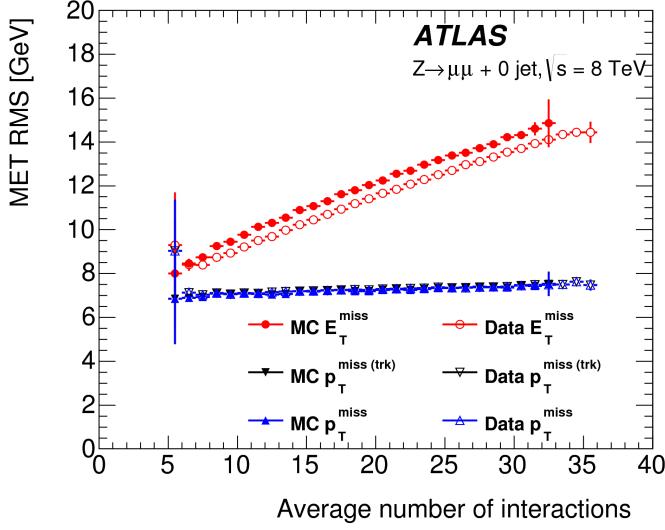


Figure 3.10: The RMS of different missing transverse momentum definitions as a function of the average number of interactions per bunch crossing

tion 3.2.

$$\mathbf{p}_T^{\text{miss}(\text{trk})} = - \left(\sum_{\text{selected leptons}} \mathbf{p}_T + \sum_{\text{other tracks}} \mathbf{p}_T \right), \quad (3.2)$$

In events with hard jets, a better resolution on the missing transverse momentum is obtained by including the calorimeter based measurement of the hard jets rather than the track based measurements. Thus, another variable, p_T^{miss} , is defined, using the nominal measurements of p_T for the selected leptons and jets and using tracks rather than calorimeter clusters for the soft component of the missing transverse momentum. This is defined in equation 3.3.

$$\mathbf{p}_T^{\text{miss}} = - \left(\sum_{\text{selected leptons}} \mathbf{p}_T + \sum_{\text{selected jets}} \mathbf{p}_T + \sum_{\text{other tracks}} \mathbf{p}_T \right), \quad (3.3)$$

Figure 3.10 illustrates that these two new variables accomplish their intended purpose. The resolution as a function of mean number of interactions for both $p_T^{\text{miss}(\text{trk})}$ and p_T^{miss} is much flatter compared to the dependence for E_T^{miss} .

Figure 3.11a shows the difference between the true and reconstructed values of missing transverse mo-

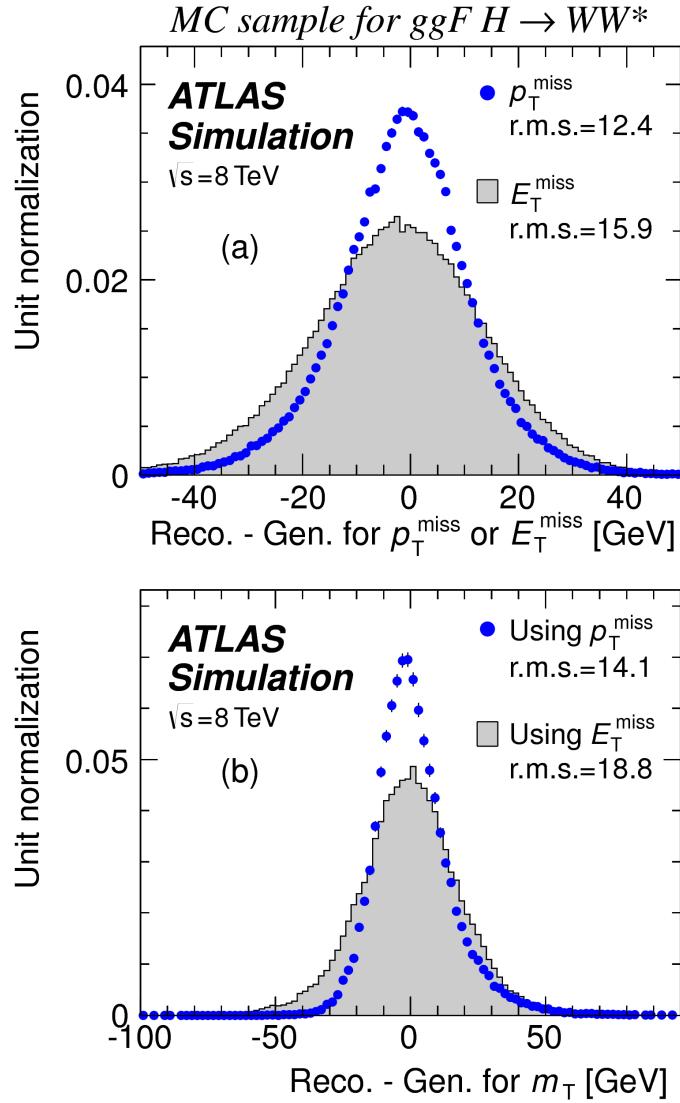


Figure 3.11: The difference between the true and reconstructed values of the missing transverse momentum (a) and m_T (b) in a gluon fusion signal sample

momentum using both the track-based p_T^{miss} and calorimeter based E_T^{miss} . The RMS of the distribution improves by 3.5 GeV when using p_T^{miss} .

3.5.3 DISTINGUISHING Z/γ^* +JETS AND $H \rightarrow WW^*$ TOPOLOGIES

The track-based definitions of missing transverse momentum were constructed to mitigate degrading performance as a function of pileup. However, an additional variable can be constructed to exploit kine-

matic and topological differences between the Z/γ^* background and $H \rightarrow WW^*$ signal. Because there are no real neutrinos in the final state (in the case of $Z/\gamma^* \rightarrow ee, \mu\mu$ decays), the dilepton system of a Z/γ^* will be balanced with the jets produced in the hard scatter. A new variable, f_{recoil} , is constructed to estimate the balance between the dilepton system and the jets in the quadrant opposite the dilepton vector in the transverse plane. It is defined in equation 3.4. The numerator of f_{recoil} is the magnitude of the vectorial sum of the p_T of jets in the quadrant opposite the dilepton system, weighted by each jet's Jet Vertex Fraction (JVF, described in chapter 2). The denominator is the magnitude of the dilepton p_T .

$$f_{\text{recoil}} = \left| \sum_{\text{jets } j \text{ in } \wedge} \text{JVF}_j \cdot \mathbf{p}_T^j \right| / p_T^{\ell\ell}. \quad (3.4)$$

Figure 3.12 shows a shape comparison of the distribution of f_{recoil} in a simulated $Z/\gamma^* + \text{jets}$ sample, a $H \rightarrow WW^*$ signal sample, and other backgrounds that contain real neutrinos. The $Z/\gamma^* + \text{jets}$ events tend to be more balanced between the dilepton system and recoiling jets, while the processes containing real neutrinos are less balanced in the transverse plane. Thus, a requirement on f_{recoil} will also reduce the $Z/\gamma^* + \text{jets}$ background while maintaining a good signal efficiency.

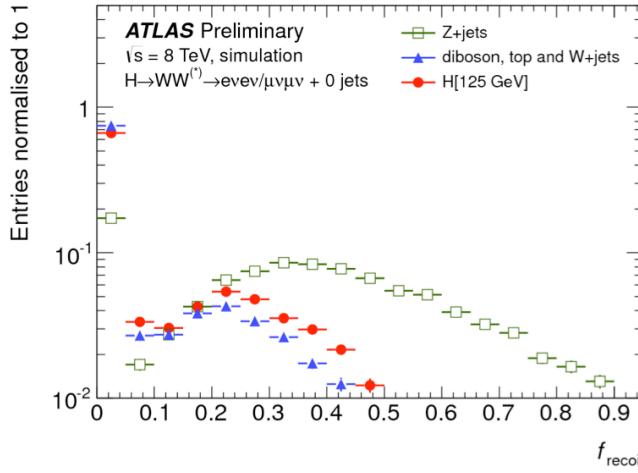


Figure 3.12: Comparison of f_{recoil} distributions for $Z/\gamma^* + \text{jets}$, $H \rightarrow WW^*$, and other backgrounds with real neutrinos.

3.5.4 OPTIMIZING BACKGROUND REDUCTION SELECTION REQUIREMENTS

The requirements on $p_T^{\text{miss(trk)}}$ and f_{recoil} used to reduce the Z+jets background must be optimized to maximize their efficacy. Figure 3.13 shows an early attempt to optimize the combination of the two requirements in the gluon fusion zero jet bin. Each bin shows the expected signal significance if the $p_{T,\text{rel}}^{\text{miss(trk)}}$ is required to be greater than the left edge of the bin and the f_{recoil} is required to be less than the top edge of the bin. The figure shows that the best signal significance comes from requiring low values of $f_{\text{recoil}} (< 0.05)$ and $p_{T,\text{rel}}^{\text{miss(trk)}}$ values greater than 45 GeV.

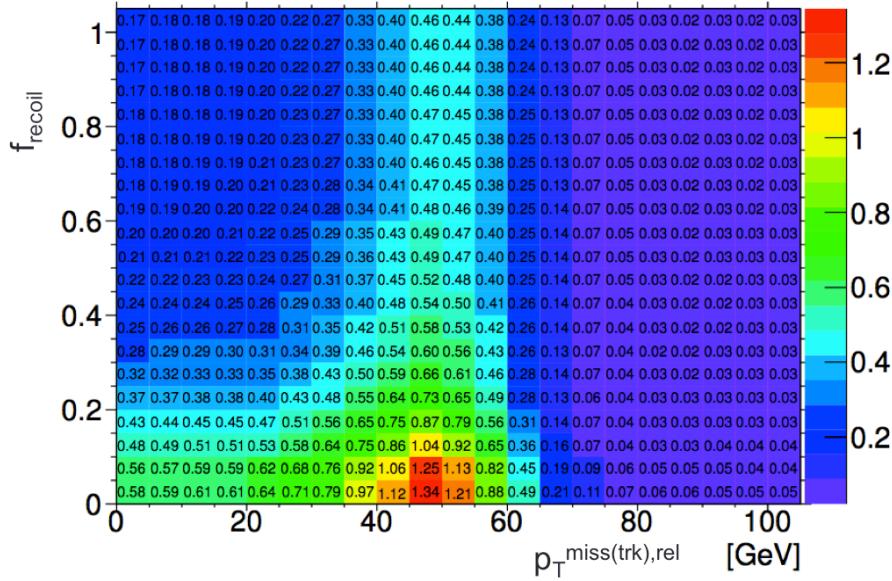


Figure 3.13: Signal significance as a function of required value for f_{recoil} and $p_{T,\text{rel}}^{\text{miss(trk)}}$ in the ggF $H \rightarrow WW^*$ with $n_j = 0$

3.6 PARAMETERS OF INTEREST AND STATISTICAL TREATMENT

As with any search or measurement, there are particular parameters of the Higgs that the $H \rightarrow WW^*$ analysis is interested in measuring. In this case, the parameters of interest are the mass of the Higgs boson and its production cross section. Because the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ process does not have a closed final state, it is not possible to measure the full invariant mass of the particle that may have produced the final state. However, a proxy for the invariant mass using transverse plane information can be de-

fined. This is described in more detail in section 3.6.1. The second parameter of interest is the ratio of the measured cross section to that expected from the Standard Model Higgs, which is denoted a μ . This is defined in equation 3.5.

$$\mu = \frac{\sigma}{\sigma_{\text{SM}}} \quad (3.5)$$

All of the likelihoods used in the statistical analysis of the final signal region events are paramaterized as a function of μ . μ is a natural variable for hypothesis testing, as $\mu = 0$ corresponds to a background only hypothesis and $\mu = 1$ corresponds exactly to a Standard Model Higgs.

3.6.1 TRANSVERSE MASS

Because the longitudinal information about the neutrinos is not attainable, the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ analysis uses a mass variable, the transverse mass, that exploits information in the transverse plane as a proxy for the full invariant mass. The transverse mass is defined in equation 3.6.

$$m_T = \sqrt{(E_T^{\ell\ell} + p_T^{\text{miss}})^2 - |\vec{p}_T^{\ell\ell} + \vec{p}_T^{\text{miss}}|^2}, \quad (3.6)$$

Here the $E_T^{\ell\ell}$ and $p_T^{\ell\ell}$ are the transverse energy and momentum of the dilepton system, while p_T^{miss} is a proxy for the transverse momentum of the di-neutrino system. The track-based p_T^{miss} is used in the m_T rather than the calorimeter based E_T^{miss} because it has a better resolution on the true transverse mass. Figure 3.11b shows the improvement in the RMS of the difference between the true and reconstructed transverse mass in a ggF signal sample. The RMS improves by 4.7 GeV using p_T^{miss} in the m_T calculation.

3.6.2 STATISTICAL TREATMENT^{*}

LIKELIHOOD FUNCTION

The statistical analysis of final event candidates is framed as a hypothesis test, where the null hypothesis is background-only (no Standard Model Higgs). The first step in the analysis is to form a likelihood function for the data. In its simplest form, this likelihood is the probability of observing the number of events seen in the final signal region given knowledge of the signal strength. Because observation of events is fundamentally a Poisson counting experiment, this simple likelihood can be expressed as a Poisson probability of observing N events given a total number of predicted signal and background events. This basic likelihood is shown in equation 3.7.

$$\mathcal{L}(\mu) = P(N|\mu S + B) \quad (3.7)$$

Here, P is the Poisson probability density function, N is the total number of observed events, μ is the signal strength, S is the predicted number of signal events, and B is the predicted number of background events.

In particle physics, certain background estimates are commonly normalized in so-called “control” regions and those predictions are scaled by the same normalization factor in the signal region. This leads to a slightly more complicated likelihood, which is a function of both the signal strength and the background normalization. This is shown in equation 3.8.

$$\mathcal{L}(\mu, \theta) = P(N|\mu S + \theta B) P(N_{\text{CR}}|\theta B_{\text{CR}}) \quad (3.8)$$

Here, θ is a so-called “nuisance parameter”, a parameter that is not a primary parameter of interest but still enters the likelihood. The second Poisson term adds an extra term to the likelihood, enforcing the fact that the background normalization must be consistent with the number of observed events in data in the control region, N_{CR} .

So far, these two formulations of likelihoods have assumed a single signal region and do not take into

^{*}Many thanks to Aaron Armbruster, whose thesis[61] inspired parts of this section.

account any shape information of potential discriminating variables. The $H \rightarrow WW^*$ analysis is divided into many different categories, and we can perform the same counting experiment described above in each individual category. As mentioned in section 3.6.1, the transverse mass is used as the primary discriminating variable in many of the $H \rightarrow WW^*$ sub-analyses, so additionally we can perform the same counting experiment in each bin of the m_T distribution to incorporate some shape information. Thus, the total likelihood becomes a product over signal regions and bins of the m_T distribution. Finally, there are usually many backgrounds that are normalized in control regions, so the new formulation of the likelihood takes this into account as well by including a product over control regions in the second Poisson term. All of these modifications are shown in equation 3.9.

$$\mathcal{L}(\mu, \theta) = \prod_{\substack{\text{SRs } i \\ \text{bins } b}} P \left(N_{ib} \middle| \mu S_{ib} + \sum_{\text{bkg } k} \theta_k B_{kib} \right) \prod_{\text{CRs } l} P \left(N_l \middle| \sum_{\text{bkg } k} \theta_k B_{kl} \right) \quad (3.9)$$

The final step to get the full likelihood used in the analysis is to add nuisance parameters for the systematic uncertainties. In cases where the uncertainty does not affect the shape of m_T bin-by-bin, each systematic uncertainty ϵ is allowed to affect the expected event yields through a linear response function of the nuisance parameter, namely $\nu(\theta) = (1 + \epsilon)^\theta$. If instead the uncertainty does affect the shape, the effect is instead parameterized by $\nu_b(\theta) = 1 + \epsilon_b \theta$. The value of the nuisance parameters for the systematic uncertainty are constrained with a Gaussian term that is added to the likelihood as well. This is of the form $g(\delta|\theta) = e^{-(\delta-\theta)^2/2}/\sqrt{2\pi}$, where δ is the central value and θ is a nuisance parameter. Finally, a last term is added to account for the statistical uncertainty in the Monte Carlo samples used, which adds an additional poisson term. The full likelihood used in the final statistical analysis is defined

in equation 3.10.

$$\begin{aligned} \mathcal{L}(\mu, \boldsymbol{\theta}) = & \prod_{\substack{\text{SRs i} \\ \text{bins b}}} P \left(N_{ib} \middle| \mu S_{ib} \cdot \prod_{\substack{\text{sig.} \\ r}} \nu_{br}(\theta_r) + \sum_{\text{bkg k}} \theta_k B_{kib} \cdot \prod_{\substack{\text{bkg.} \\ s}} \nu_{bs}(\theta_s) \right) \\ & \cdot \prod_{\text{CRs l}} P \left(N_l \middle| \sum_{\text{bkg k}} \theta_k B_{kl} \right) \\ & \cdot \prod_{\substack{\text{syst} \\ t}} g(\delta_t | \theta_t) \cdot \prod_{\text{bkg k}} P(\xi_k | \zeta_k \theta_k) \end{aligned} \quad (3.10)$$

In the fourth term of the equation, quantifying uncertainty due to finite Monte Carlo sample size, ξ represents the central value of the background prediction, θ is the associated nuisance parameter, $\zeta = (B/\delta B)^2$, where δB is the statistical uncertainty of B .

The best fit value of the signal strength μ is determined by finding the values of μ and $\boldsymbol{\theta}$ that maximize the likelihood, while setting $\delta = 0$ and $\xi = \zeta$.

Once the likelihood is defined, a test statistic must be built for use in hypothesis testing.

TEST STATISTIC

To distinguish whether the data match a background only or background and signal hypothesis, a test statistic must be used. The $H \rightarrow WW^*$ analysis used the profile likelihood technique[62]. The first step in formulating this test statistic is to define the profile likelihood ratio, shown in equation 3.11.

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\theta}_\mu)}{\mathcal{L}(\hat{\mu}, \hat{\theta})} \quad (3.11)$$

Here $\hat{\theta}_\mu$ is the value of θ that maximizes the likelihood for the choice of μ being tested. Additionally, $\hat{\theta}$ and $\hat{\mu}$ represent the values of θ and μ that gives the overall maximum value of the likelihood.

Once this is defined, a test statistic q_μ is constructed. This is shown in equation 3.12.

$$q_\mu = -2 \ln \lambda(\mu) \quad (3.12)$$

A higher value of q_μ means that the data are more incompatible with the hypothesized value of μ , and q_0 then corresponds to the value of the test statistic for the background only hypothesis. A p_0 value is then defined to quantify the compatibility between the data and the null hypothesis. The p_0 value is the probability of obtaining a value of q_0 larger than the observed value, and this is shown in equation 3.13.

$$p_0 = \int_{q_0^{\text{obs}}}^{\infty} f(q_\mu | \mu = 0) dq_\mu \quad (3.13)$$

Here $f(q_\mu)$ is the probability distribution function of the test statistic. Finally, the p_0 value can be converted into a signal significance, using the formula in equation 3.14, or the one-sided tail of the Gaussian distribution.

$$Z_0 = \sqrt{2} \operatorname{erf}^{-1}(1 - 2p_0) \quad (3.14)$$

The threshold for discovery used in particle physics is $Z_0 \geq 5$, more commonly known as a value of 5σ .

The real voyage of discovery consists not in seeking new landscapes, but in having new eyes.

Marcel Proust

4

The discovery of the Higgs boson and the role of the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ channel

4.1 INTRODUCTION

This chapter presents the results of the search for the Higgs boson in 4.8 fb^{-1} collected at $\sqrt{s} = 7 \text{ TeV}$ and 5.8 fb^{-1} at $\sqrt{s} = 8 \text{ TeV}$. The results of three searches at $\sqrt{s} = 8 \text{ TeV}$ in the $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ \rightarrow 4\ell$, and $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ channels are combined with results of searches at $\sqrt{s} = 7 \text{ TeV}$ in the same search channels (as well as the $H \rightarrow \tau\tau$ production and associated production searches for $H \rightarrow b\bar{b}$). The results of this combination are a 5.9σ detection of a new particle consistent with a Higgs boson. Rather than going into detail for all of the different Higgs decay searches, this chapter will discuss the three most sensitive channels and in particular give a comparison of the $\gamma\gamma$ and 4ℓ channels to the $\ell\nu\ell\nu$ channel. This chapter is a summary of results presented in the ATLAS Higgs discovery

publication[21].

4.2 DATA AND SIMULATION SAMPLES

The data sample used for the following results was taken in 2011 and 2012 at center of mass energies of 7 and 8 TeV, respectively, with 4.8 fb^{-1} collected at 7 TeV and 5.8 fb^{-1} collected at 8 TeV. Higgs production in the gluon fusion and vector boson fusion modes is modeled with **POWHEG** for the hard scattering event and **PYTHIA** for the showing and hadronization. Associated production of a Higgs with a vector boson or top quarks is modeled via **PYTHIA**.

Table 4.1 shows the Monte Carlo generators used for modeling the signal and background processes relevant for the three analyses to be discussed.

Process	Generator
$ggF, VBF H$	POWHEG + PYTHIA
$WH, ZH, t\bar{t}H$	PYTHIA
$W + \text{jets}, Z/\gamma^* + \text{jets}$	ALPGEN + HERWIG
$t\bar{t}, tW, tb$	MC@NLO + HERWIG
tqb	ACERMC + PYTHIA
$q\bar{q} \rightarrow WW$	MC@NLO + HERWIG
$gg \rightarrow WW$	GG2WW+ HERWIG
$q\bar{q} \rightarrow ZZ$	POWHEG + PYTHIA
$gg \rightarrow ZZ$	GG2ZZ+ HERWIG
WZ	MADGRAPH+ PYTHIA , HERWIG
$W\gamma + \text{jets}$	ALPGEN + HERWIG
$W\gamma^*$	MADGRAPH+ PYTHIA
$q\bar{q}/gg \rightarrow \gamma\gamma$	SHERPA

Table 4.1: Monte carlo generators used to model signal and background for the Higgs search[21].

4.3 $H \rightarrow \gamma\gamma$ SEARCH

The $H \rightarrow \gamma\gamma$ search is in essence a search for a peaked excess above the falling SM diphoton mass spectrum, with $m_{\gamma\gamma}$ as the ultimate discriminating variable. Events are selected by requiring two isolated photons, with the leading (sub-leading) photon required to have $E_T > 40(30)$ GeV. In the 8 TeV data, the photons are required to pass cut-based identification criteria consistent with a photon in the electro-

magnetic calorimeter and little leakage in the hadronic calorimeter. The main challenges for this analysis are accurate mass reconstruction and background estimation. These are the issues that will be discussed in this section.

4.3.1 DIPHOTON INVARIANT MASS RECONSTRUCTION

In order to accurately reconstruct the invariant mass of the di-photon system, both the energy and direction of the photons must be measured well. The η coordinate of the photons is calculated using the position of the primary vertex identified in the event as well as the entry points of the photons in the calorimeter. Therefore, the identification of the primary vertex of the hard interaction is particularly important, and is done using a likelihood which combines the flight direction of the photons, the beam position measurement, and the scalar sum of p_T^2 for tracks at the primary vertex. This gives a resolution of 15 mm on the z position of the primary vertex (which improves to 6 mm for photons which convert to electron-positron pairs in the ID material).

In the inclusive sample, there are 35251 events observed in the 8 TeV sample and 23788 observed in the 7 TeV sample. The total expected number of events for a Higgs at 126.5 GeV is 110.5 at 8 TeV and 79.6 at 7 TeV. The full width at half maximum (FWHM) of the mass distribution at this mass is 3.9 GeV in the inclusive sample. The mass distribution is modeled with a crystal ball function for the core and a Gaussian for the tails, and the width of the core is 1.6 GeV in the inclusive sample. The events are additionally categorized based on the types of photons in the event, leading to ten distinct categories which have varying S/B and mass resolution. The best expected FWHM of 3.2 GeV comes from the category with unconverted photons at high p_{Tt} (the component of the diphoton p_T orthogonal to the axis defined by the vectorial difference of their momenta). The worst expected FWHM of 6.1 GeV comes from converted photons in the transition region of the calorimeter.

4.3.2 BACKGROUND ESTIMATION

The background is modeled with a falling spectrum in $m_{\gamma\gamma}$ that is parameterized by different functions depending on the category of the event. An exponential function of a second order polynomial is used

for the categories involving low p_{Tt} photons in the central region, as well as photon conversions in the transition region. A fourth order Bernstein polynomial is used the low p_{Tt} categories outside of the central region as well as the inclusive category. An exponential function is used for all other categories.

One important consideration when using such a background model is how much the model could be biased by a signal arising in the data. This level of potential bias from a signal is estimated using the mass spectrum of background plus signal from simulation (with the background shape in simulation validated using data-driven control regions). For this test, the signal shape is taken from simulation but the yield is left as a free parameter. The bias is then taken as the largest absolute yield coming from the likelihood fit. This potential bias is taken as a systematic uncertainty and amounts to $\pm(0.3\text{-}6.8)$ events in the 8 TeV data sample.

4.3.3 RESULTS

The resulting diphoton mass spectrum is shown in figure 4.1. The best fit mass value in the $\gamma\gamma$ channel alone in the combined 7 and 8 TeV data is 126.5 GeV. The local significance at this point is 4.5σ , with an expected significance of 2.5σ . Therefore, the measured signal strength μ is 1.8 ± 0.5 in this channel.

4.4 $H \rightarrow ZZ \rightarrow 4\ell$ SEARCH

The $H \rightarrow ZZ \rightarrow 4\ell$ analysis searches for a Standard Model Higgs boson decaying to two Z bosons, each of which decays to a pair of same flavor, opposite charge isolated leptons. The ultimate discriminating variable is $m_{4\ell}$, or the invariant mass of the four selected leptons. The ℓ denotes an e or μ as with the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ analysis. Electrons must have $p_T > 7$ GeV and $|\eta| < 2.47$, while muons must have $p_T > 6$ GeV and $|\eta| < 2.7$. There are then additional p_T cuts depending on the relative ranking of the lepton p_T , with the leading lepton required to have $p_T > 20$ GeV, subleading lepton $p_T > 15$ GeV, and third leading lepton $p_T > 10$ GeV.

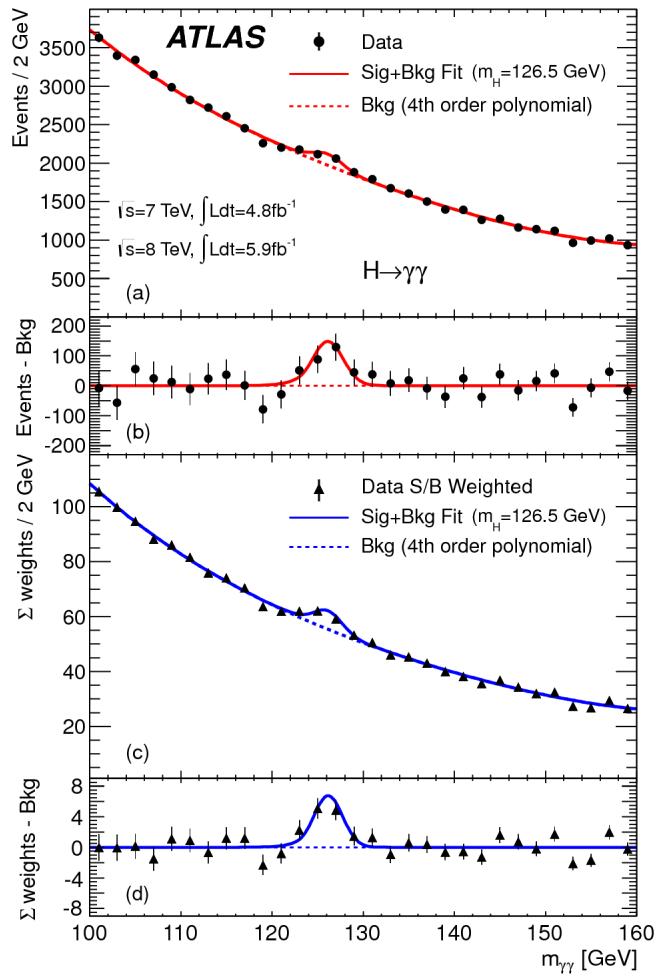


Figure 4.1: Diphoton mass spectrum in 7 and 8 TeV data. Panel a) shows the unweighted data distribution superimposed on the background fit, while panel c) shows the data where each event category is weighted by its signal to background ratio. Panels b) and d) show the respective distributions with background subtracted[21].

4.4.1 LEPTON PAIR CONSTRUCTION

One of the difficult aspects of this channel is that there are multiple possible same flavor, opposite charge lepton pairs that can be produced out of the four selected leptons. The lepton pair whose mass is closest to m_Z is referred to as the leading pair and its mass is denoted as m_{12} , while the other lepton pair's mass is m_{34} . Events are required to have m_{12} consistent with a Z mass, or $50 < m_{12} < 106$ GeV. m_{34} is required to be less than 115 GeV. The minimum requirement for m_{min} varies from 17.5 GeV at $m_{4\ell} = 120$ GeV to 50 GeV at $m_{4\ell} = 190$ GeV.

Four distinct signal regions are constructed depending on the flavors of the final state, additionally separated by the flavor of the leading lepton pair. These are referred to as $4e$, $2e2\mu$, $2\mu2e$, 4μ .

Table 4.2 summarizes the mass resolutions and selection efficiencies for the four signal regions. The mass resolution at 125 GeV is dominated by experimental resolution, ranging from 1.7 GeV to 2.3 GeV.

Channel	Efficiency	Mass resolution
4μ	36%	1.7 GeV
$2e2\mu$	22%	1.7 GeV
$2\mu2e$	22%	2.2 GeV
$4e$	20%	2.3 GeV

Table 4.2: A summary of selection efficiencies and expected resolutions for a 125 GeV Higgs in the $H \rightarrow ZZ \rightarrow 4\ell$ channel[21].

4.4.2 BACKGROUND ESTIMATION

The main backgrounds in the $H \rightarrow ZZ \rightarrow 4\ell$ search are continuum ZZ^* production, $Z + \text{jets}$ production, and $t\bar{t}$. The $m_{4\ell}$ distribution for background is estimated from simulation. The normalization of the SM ZZ^* background is also taken from MC simulation, while the $Z + \text{jets}$ and $t\bar{t}$ normalizations are taken from data-driven methods. The data-driven methods are separated into $\ell\ell + \mu\mu$ and $\ell\ell + ee$ methods.

In the $\ell\ell + \mu\mu$ case, the control region is defined by relaxing the requirements on the $\mu\mu$ pair to create a region enriched in background. The $t\bar{t}$ and $Z + \text{jets}$ backgrounds are then fit together from the leading mass distribution m_{12} . A similar method is used in the $\ell\ell + ee$ case, but there is an additional sorting into categories based on identification criteria for electrons that come from heavy flavor decays or photon conversions.

4.4.3 RESULTS

Figure 4.2 shows the $m_{4\ell}$ spectrum measured in the 7 and 8 TeV datasets. The total number of events observed in the window between 120 and 130 GeV is 13, with 6 events in the 4μ channel, 2 events in the $4e$ channel, and 5 events in the $2e2\mu/2\mu2e$. The best fit μ value in the combined 7 and 8 TeV data

occurs at 125 GeV and is measured to be 1.2 ± 0.6 . The observed significance at this mass is 3.6σ , with an expected significance of 2.7σ .

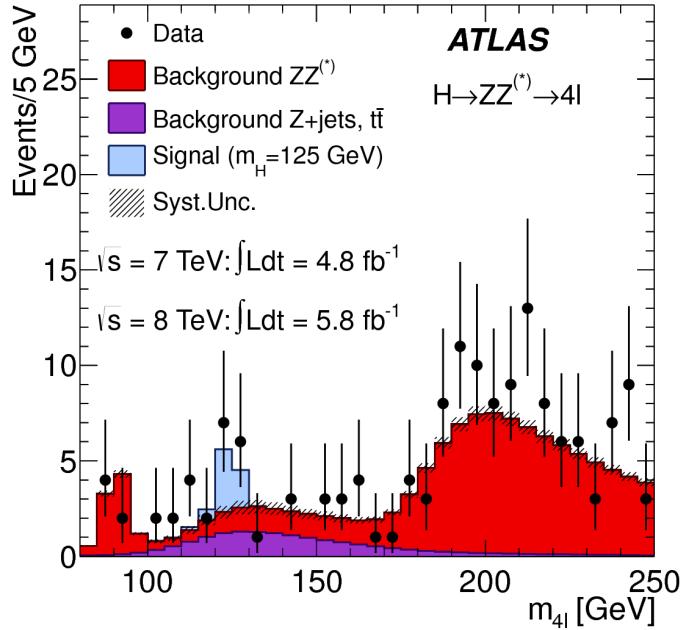


Figure 4.2: Four lepton invariant mass spectrum ($m_{4\ell}$) in 7 and 8 TeV data compared to background estimate. A 125 GeV SM Higgs signal is shown in blue[21].

4.5 $H \rightarrow WW \rightarrow e\nu\mu\nu$ SEARCH

The $H \rightarrow WW \rightarrow e\nu\mu\nu$ search is unique compared to the ZZ and $\gamma\gamma$ channels. The Higgs mass cannot be fully reconstructed due to the presence of neutrinos in the final state, so the transverse mass m_T is used as the final discriminating variable. Compared to the other channels, there are more backgrounds here as well, as discussed in chapter 3. The same flavor final states are excluded from this search due to high pileup in the 8 TeV dataset.

4.5.1 EVENT SELECTION

The analysis requires two opposite charge isolated leptons, with the leading (sub-leading) lepton required to have $p_T > 25(15)$ GeV. The events are separated into different signal regions depending on which

flavor of lepton is leading ($e\mu$ for leading electron, μe for leading muon). Strict lepton quality cuts are applied to the sample to reduce backgrounds from fake leptons.

Jets are reconstructed with the anti- k_T algorithm with a radius parameter $R = 0.4$. The jets are required to have $p_T > 25 \text{ GeV}$ and $|\text{eta}| < 4.5$, with jets in the tracking volume required to have a jet vertex fraction of 0.5 and jets in the forward region required to have $p_T > 30 \text{ GeV}$. The analysis is separated into three different signal regions based on jet multiplicity: $n_j = 0, 1, \geq 2$.

To indicate the presence of neutrinos in the event, a requirement of $E_{\text{T},\text{rel}}^{\text{miss}} > 25 \text{ GeV}$ is made. This requirement significantly reduces the QCD multijet and $Z/\gamma^* + \text{jets}$ backgrounds.

Additional selections are applied to require the dilepton topology to correspond to that of a SM Higgs. The requirements are presented here - more detailed discussion on the motivation for each requirement is saved for Chapter 5. In all of the jet multiplicity channels, the dilepton system is required to have a small gap in azimuthal angle, $\Delta\phi_{\ell\ell} < 1.8$. Similarly, the $m_{\ell\ell}$ is required to be less than 50 GeV in the lower jet multiplicity channels and less than 80 GeV in the $n_j \geq 2$ channel. In the $n_j = 0$ channel, the magnitude of the dilepton p_T , $p_T^{\ell\ell}$, is required to be greater than 30 GeV.

In the higher jet multiplicity channels ($n_j \geq 1$), the top background is a more important component and must be reduced. The total transverse momentum p_T^{sum} is thus required to be less than 30 GeV. Additionally, the di- τ invariant mass $m_{\tau\tau}$ (dilepton mass computed under the assumption that the neutrinos from the τ decay are emitted collinear to the charged leptons) is used to reject $Z \rightarrow \tau\tau$ events by requiring $|m_{\tau\tau} - m_Z| > 25 \text{ GeV}$. These variables are also discussed in more detail in Chapter 5.

In the $n_j \geq 2$ channel, requirements are made to isolate the VBF contribution to Higgs production. The kinematics of the two leading jets are used to make these requirements. In particular, the event must have $\Delta y_{jj} > 3.8$ and $m_{jj} > 500 \text{ GeV}$, along with a veto on having any additional jets with rapidity between the two leading jets.

4.5.2 BACKGROUND ESTIMATION

The details of the background estimation techniques used in the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ analysis are discussed in section 5.5. As that section refers to a later iteration of the analysis, a general discussion is given

here for completeness. The dominant backgrounds are SM WW production and top (both pair and single) production, and these backgrounds have their normalizations estimated from dedicated control regions while their shapes are taken from simulation. The $W+jets$ background estimate is taken entirely from data using a control sample with one well reconstructed lepton and one anti-identified lepton. All other backgrounds are taken purely from simulation.

4.5.3 RESULTS

Figure 4.3 shows the m_T distribution in the $n_j \leq 1$ channels for 8 TeV data. (No events are observed in data in the $n_j \geq 2$ channels in this dataset). The excess shown here relatively flat as a function of hypothesized Higgs mass. The combined 7 and 8 TeV data gives an excess with local significance of 2.8σ with an expected significance of 2.3σ , corresponding to a μ measurement of 1.3 ± 0.5 .

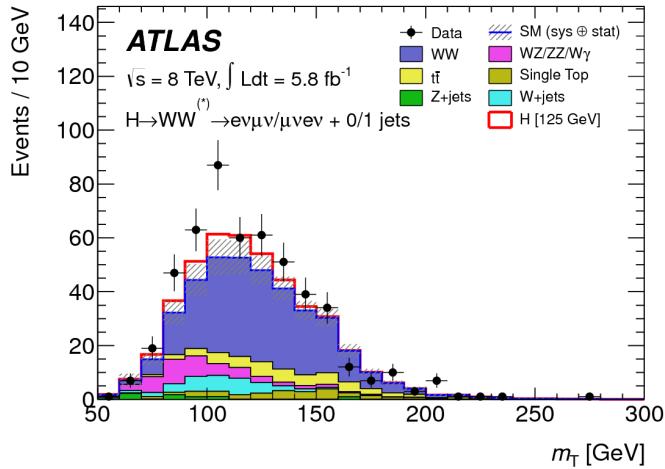


Figure 4.3: m_T distribution in the $H \rightarrow WW \rightarrow e\nu\mu\nu$ $n_j \leq 1$ channels for 8 TeV data[21].

4.6 COMBINED RESULTS

The statistical interpretation of the combined results is undertaken as described in section 3.6.2, with a hypothesis test based on a likelihood ratio parameterized by the Higgs signal strength μ . The null hypothesis corresponds to $\mu = 0$, while the SM Higgs corresponds to $\mu = 1$.

Table 4.3 summarizes the properties of the individual channels as well as the significances of the excesses seen. The most significant observed local excess comes from the $\gamma\gamma$ channel. Figure 4.4 shows a comparison of the observed local p_0 values as a function of hypothesized mass for the three different search channels. Both the ZZ^* and $\gamma\gamma$ channels have very peaked excesses, while the WW^* excess can be seen as very broad because the m_T distribution does not provide detailed information about the true Higgs mass.

Channel	Fit var.	Observed Z_l	Expected Z_l	$\hat{\mu}$
$H \rightarrow ZZ^* \rightarrow 4\ell$	$m_{4\ell}$	3.6	2.7	1.2 ± 0.6
$H \rightarrow \gamma\gamma$	$m_{\gamma\gamma}$	4.5	2.5	1.8 ± 0.5
$H \rightarrow WW^* \rightarrow e\nu\mu\nu$	m_T	2.8	2.3	1.3 ± 0.5
Combined	-	6.0	4.9	1.4 ± 0.3

Table 4.3: Summary of the expected and observed significance and measured signal strengths in the combined 7 and 8 TeV datasets for the Higgs discovery analysis[21].

Figure 4.5 shows the combined exclusion limit, p_0 , and signal strength. The highest local excess comes at a value of 126.5 GeV and corresponds to a 6.0σ observed excess.

Figure 4.6 shows a comparison of the measured signal strengths between the different Higgs search channels. All measured μ are consistent with unity within their uncertainty, and the combined μ measurement is 1.4 ± 0.3 .

The likelihood can also be computed in a two-dimensional plane of m_H and μ , and this is shown in figure 4.7. The figure shows that while the $\gamma\gamma$ and ZZ^* channels have very good mass resolution, the excess in WW^* covers a broad mass range. The banana shape of the WW^* result is due to the fact that the excess in this channel can either be explained by increasing the signal strength or by changing the mass (and thus the cross section). The two parameters are correlated due to the lack of mass sensitivity in this channel.

Because multiple Higgs mass points are searched for, the local significance must be corrected for a look-elsewhere effect to compute a true global significance. The global significance for finding a Higgs anywhere in the mass range of 110 GeV to 600 GeV is 5.1σ . This increases slightly to 5.3σ if only mass range from 110 to 150 GeV.

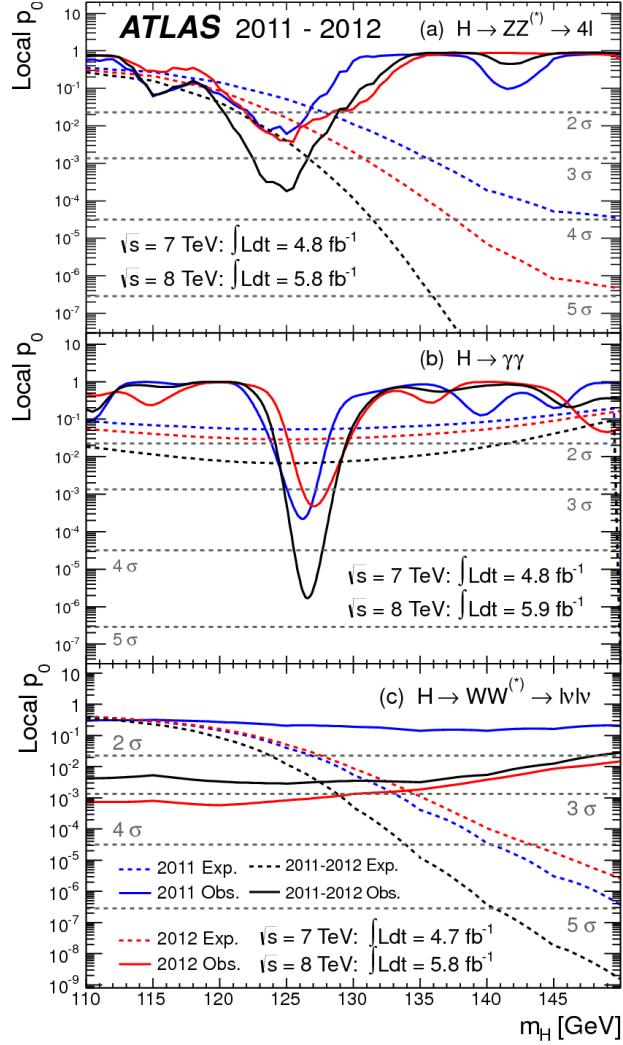


Figure 4.4: Local p_0 distribution as a function of hypothesized Higgs mass for the $H \rightarrow ZZ^* \rightarrow 4\ell$ (a), $H \rightarrow \gamma\gamma$ (b), and $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ (c) channels. Dashed curves show expected results, while solid curves show observed. Red curves are from 7 TeV data, blue curves from 8 TeV, and black curved combined[21].

4.7 CONCLUSION

A search for the production of a Standard Model Higgs boson was conducted in 4.8 fb^{-1} collected at $\sqrt{s} = 7 \text{ TeV}$ and 5.8 fb^{-1} at $\sqrt{s} = 8 \text{ TeV}$. A new particle consistent with the Higgs boson was observed, with a mass of 126.5 GeV and a global (local) significance of $5.1(6.0)\sigma$. This is the first discovery level observation of a particle consistent with the Higgs.

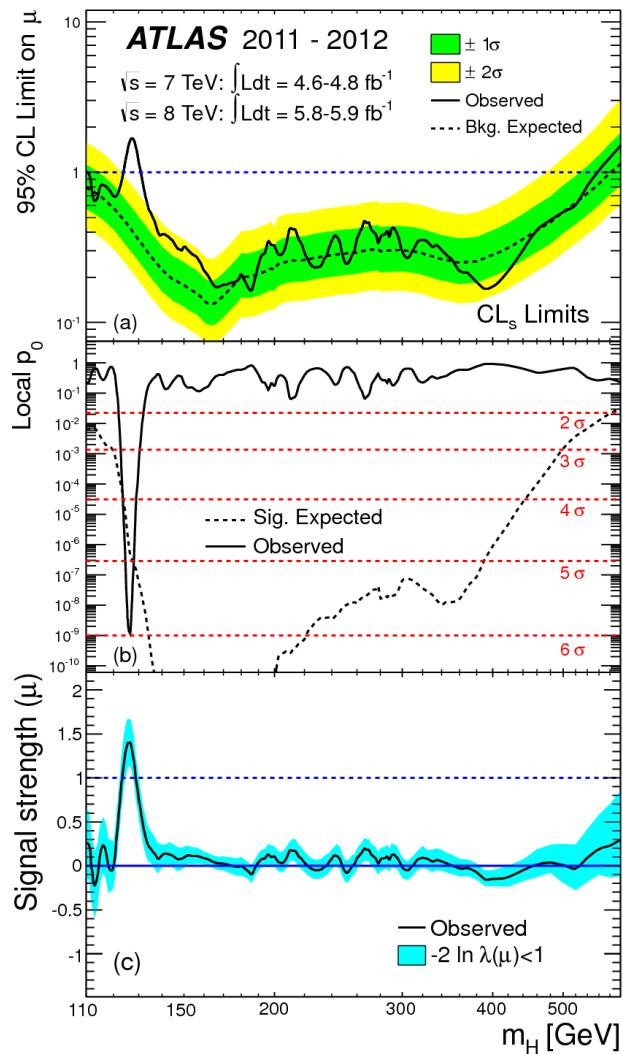


Figure 4.5: Combined 95% CL limits (a), local p_0 values (b), and signal strength measurement (c) as a function of Higgs mass[21].

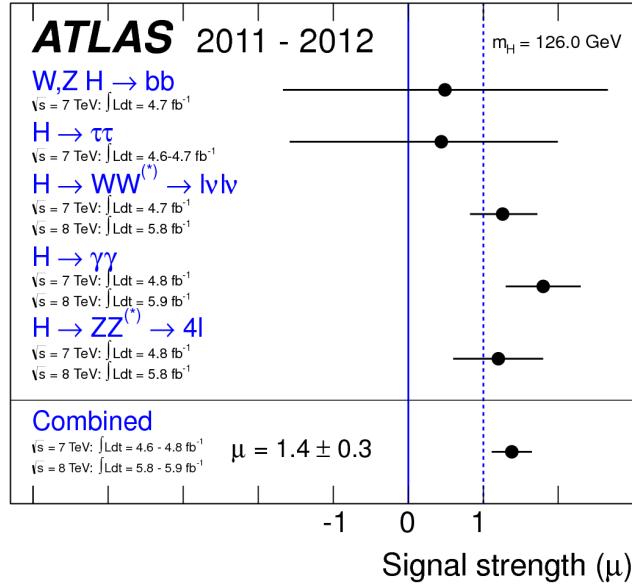


Figure 4.6: Comparison of measured signal strength μ for a 126 GeV Higgs in the 7 and 8 TeV datasets[21].

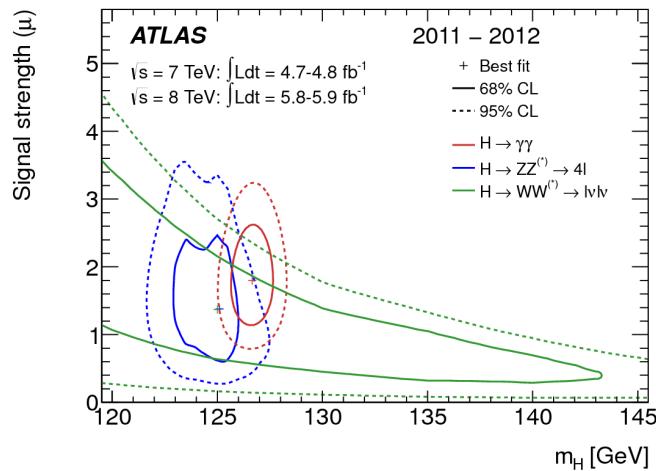


Figure 4.7: Two dimensional likelihood as a function of signal strength μ and Higgs mass m_H [21].

*The imagination of nature is far, far greater than the
imagination of man.*

Richard Feynman

5

Observation of Vector Boson Fusion

production of $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$

5.1 INTRODUCTION

After the discovery of a particle consistent with the Higgs boson, the $H \rightarrow WW^*$ analysis had two main goals. The first goal was to increase the sensitivity of the analysis to fully confirm that the $H \rightarrow WW^*$ process did indeed exist. The second goal was to characterize the particle as much as possible, including searching for the lower cross-section production modes, in order to confirm that it was indeed a Higgs boson. This chapter presents a dedicated search for Vector Boson Fusion (VBF) production of a Higgs boson decaying via the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ mode. First, basics of the topology of VBF production are presented. Then, the details of the analysis are shown, including signal region definition, background estimation techniques, and systematic uncertainties. Finally, the results of the analysis are shown. As will

be shown, this analysis is the first and most sensitive observation of the VBF production mode of the Higgs on ATLAS.

5.2 DATA AND SIMULATION SAMPLES

The results presented here are with 20.3 fb^{-1} taken at $\sqrt{s} = 8 \text{ TeV}$ and 4.5 fb^{-1} taken at $\sqrt{s} = 7 \text{ TeV}$. The details of the LHC and detector conditions during this period are given in Chapter 2. The trigger selection defining the dataset is discussed in section 5.2.1. The simulation samples used for signal and background modeling are given in section 5.2.2.

5.2.1 TRIGGERS

The analysis uses a combination of single lepton and dilepton triggers to allow lowering of the p_T thresholds and increased signal acceptance. The p_T threshold on the leptons is a particularly important consideration for this signal. Because the second W produced in the decay can be off-shell, it tends to produce lower momentum leptons. Thus, being able to lower the p_T threshold while still maintaining a low background rate is critical. Figure 5.1 shows an example of the subleading lepton p_T for a VBF $H \rightarrow WW^*$ signal compared to the corresponding $t\bar{t}$ background. Note that the lepton p_T spectrum is considerably softer in the signal sample.

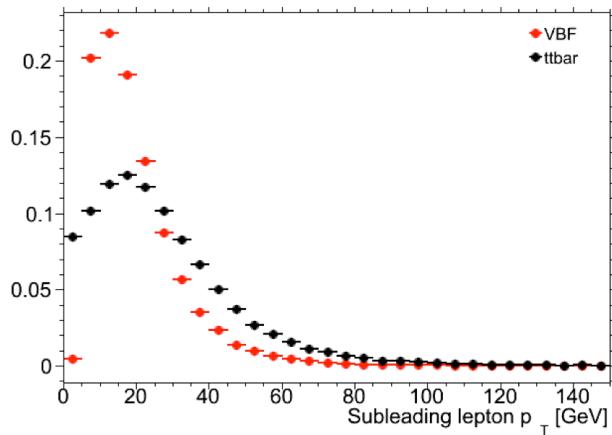


Figure 5.1: A comparison of the subleading lepton p_T spectrum between VBF $H \rightarrow WW^*$ production and $t\bar{t}$ background

As discussed in Chapter 2, there are multiple levels in the ATLAS trigger system, and there are different p_T thresholds imposed for the leptons at each level. Additionally, some triggers have a loose selection on the isolation of the lepton (looser than that applied offline in the analysis object selection). Table 5.1 shows the thresholds used for single lepton triggers, while table 5.2 shows the thresholds coming from di-lepton triggers. The single lepton trigger efficiency for muons that pass the analysis object selection is 70% for muons in the barrel region ($|\eta| < 1.05$) and 90% in the endcap region. The electron trigger efficiency increases with electron p_T but the average is approximately 90%. These efficiencies are measured by combined performance and trigger signature groups[63, 64].

	Level-1 threshold	High-level threshold
Electron	18	$24i$
	30	60
Muon	15	$24i$
		36

Table 5.1: Single lepton triggers used for electrons and muons. A logical “or” of the triggers listed for each lepton type is taken. Units are in GeV, and the i denotes an isolation requirement in the trigger.

	Level-1 threshold	High-level threshold
ee	10 and 10	12 and 12
$\mu\mu$	15	18 and 8
$e\mu$	10 and 6	12 and 8

Table 5.2: Di-lepton triggers used for different flavor combinations. The two thresholds listed refer to leading and sub-leading leptons, respectively. The di-muon trigger only requires a single lepton at level-1.

The combination of all triggers shown gives good efficiency for signal events. This efficiency is summarized in table 5.3. The relative improvement in efficiency by adding the dilepton triggers is also shown in the same table. The largest gain comes in the $\mu\mu$ channel. Overall the trigger selection shows a good efficiency for $H \rightarrow WW^*$ signal events.

5.2.2 MONTE CARLO SAMPLES

Modeling of signal and background processes in the signal region, in particular for the m_T distribution, is an important consideration for the final interpretation of the analysis. Therefore, careful consideration

Channel	Trigger efficiency	Gain from 2ℓ trigger
ee	97%	9.1%
$\mu\mu$	89%	18.5%
$e\mu$	95%	8.3%
μe	81%	8.2%

Table 5.3: Trigger efficiency for signal events and relative gain of adding a dilepton trigger on top of the single lepton trigger selection. The first lepton is the leading, while the second is the sub-leading. Efficiencies shown here are for the ggF signal in the $n_j = 0$ category but are comparable for the VBF signal.

must be paid to which Monte Carlo (MC) generators are used for specific processes. With the exception of the $W + \text{jet}$ and multijet backgrounds, the m_T shape used as the final discriminant is taken from simulation. (Many backgrounds are normalized from data, as described in section 5.5).

Table 5.4 shows the MC generators used for the signal and background processes, as well as their cross sections. In order to include corrections up to next-to-leading order (NLO) in the QCD coupling constant α_s , the **POWHEG** [65] generator is often used. In some cases, only leading order generators like **ACERMC** [66] and **GG2VV** [67] are available for the process in question. If the process requires good modeling for very high parton multiplicities, the **SHERPA** [68] and **ALPGEN** [69] generators are used to provide merged calculations for five or fewer additional partons. These matrix element level calculations must then be additionally matched to models of the underlying event, hadronization, and parton shower. There are four possible generators for this: **SHERPA**, **PYTHIA 6**[70], **PYTHIA 8**[71], or **HERWIG** [72] + **JIMMY** [73]. The simulation additionally requires an input parton distribution function (PDF). The **CT10**[74] PDFs are used for **SHERPA** and **POWHEG** simulated samples, while **CTEQ6Li**[75] is used for **ALPGEN + HERWIG** and **ACERMC** simulations. The Drell-Yan samples are reweighted to the **MRST** [76] PDFs, as these are found to give the best agreement between data and simulation.

Once the basic hard scattering process is simulated, it must be passed through a detector simulation and additional pile-up events must be overlaid. The pile-up events are modeled with **PYTHIA 8**, and the ATLAS detector is simulated with **GEANT4**[77]. Because of the unique phase space of the $H \rightarrow WW^*$ analysis, events are sometimes filtered at generator level to allow for more efficient generation of relevant events. The efficiency of the trigger in MC simulation does not always match the measured efficiency in data, so trigger scale factors are applied to correct the MC efficiency to the data. These are derived by the

Process	MC generator	$\sigma \cdot \mathcal{B}$ (pb)
Signal		
ggF	$H \rightarrow WW^*$ POWHEG +PYTHIA 8	0.435
VBF	$H \rightarrow WW^*$ POWHEG +PYTHIA 8	0.0356
VH	$H \rightarrow WW^*$ PYTHIA 8	0.0253
WW		
$q\bar{q} \rightarrow WW$ and $qg \rightarrow WW$	POWHEG +PYTHIA 6	5.68
$gg \rightarrow WW$	GG2VV +HERWIG	0.196
$(q\bar{q} \rightarrow W) + (q\bar{q} \rightarrow W)$	PYTHIA 8	0.480
$q\bar{q} \rightarrow WW$	SHERPA	5.68
VBS $WW + 2$ jets	SHERPA	0.0397
Top quarks		
$t\bar{t}$	POWHEG +PYTHIA 6	26.6
Wt	POWHEG +PYTHIA 6	2.35
$tq\bar{b}$	ACERMC +PYTHIA 6	28.4
$t\bar{b}$	POWHEG +PYTHIA 6	1.82
Other dibosons (VV)		
$W\gamma$ ($p_T^\gamma > 8$ GeV)	ALPGEN +HERWIG	369
$W\gamma^*$ ($m_{\ell\ell} \leq 7$ GeV)	SHERPA	12.2
WZ ($m_{\ell\ell} > 7$ GeV)	POWHEG +PYTHIA 8	12.7
VBS $WZ + 2$ jets	SHERPA	0.0126
($m_{\ell\ell} > 7$ GeV)		
$Z\gamma$ ($p_T^\gamma > 8$ GeV)	SHERPA	163
$Z\gamma^*$ (min. $m_{\ell\ell} \leq 4$ GeV)	SHERPA	7.31
ZZ ($m_{\ell\ell} > 4$ GeV)	POWHEG +PYTHIA 8	0.733
$ZZ \rightarrow \ell\ell\nu\nu$ ($m_{\ell\ell} > 4$ GeV)	POWHEG +PYTHIA 8	0.504
Drell-Yan		
Z ($m_{\ell\ell} > 10$ GeV)	ALPGEN +HERWIG	16500
VBF $Z + 2$ jets	SHERPA	5.36
($m_{\ell\ell} > 7$ GeV)		

Table 5.4: Monte Carlo samples used to model the signal and background processes[20].

combined performance groups[63, 64].

5.3 OBJECT SELECTION

In order to define the signal region, the analysis must first select the objects to be considered. The details of the object reconstruction algorithms are discussed in Chapter 2, while this section gives specific selection cuts used in the $H \rightarrow WW^*$ analysis.

The first step in this process is to select a primary vertex candidates. The event's primary vertex is the vertex with the largest sum of p_T^2 for associated tracks and is required to have at least three tracks with $p_T > 450$ MeV. Many of the object selection cuts are then made relative to this chosen primary vertex.

5.3.1 MUONS

The analysis uses combined muon candidates, where a track in the Inner Detector has been matched to a standalone track in the Muon Spectrometer. The track parameters are combined statistically in the muon reconstruction algorithm[16]. The muons are required to be within $|\eta| < 2.5$ and have a $p_T > 10$ GeV. To reduce backgrounds coming from mis-reconstructed leptons, there are requirements on the impact parameter of the muon relative to the primary vertex. The transverse impact parameter d_0 is required to be small relative to its estimated uncertainty, the exact cut value being $d_0/\sigma_{d_0} < 3$. The longitudinal impact parameter z_0 must satisfy $|z_0 \sin \theta| < 1$ mm.

As discussed previously, the muons must also be isolated. There are two types of lepton isolations that are calculated: track-based and calorimeter-based. For muons, the track-based isolation is defined using the scalar sum $\sum p_T$ for tracks with $p_T > 1$ GeV (excluding the muon's track) within a cone of $\Delta R = 0.3$ (0.4) for muon with $p_T > 15$ GeV ($10 < p_T < 15$ GeV). The final isolation requirement is made by requiring that this scalar sum be no more than a certain fraction of the muon's p_T . This requirement varies with muon p_T and the exact cuts are defined in table 5.5.

The calorimeter-based muon isolation is defined using as a $\sum E_T$ calculated from calorimeter cells using the same cone size as the track-based isolation but excluding cells with $\Delta R < 0.05$ around the muon. This requirement is also defined as a cut on the ratio of the sum to the muon p_T and varies with muon p_T . The cut values are also given in table 5.5.

The isolation requirements loosen as a function of p_T to allow for larger signal acceptance. At low p_T ,

the isolation is tightened to reduce the $W + \text{jets}$ background which arises from a misidentified lepton.

p_T range (GeV)	Calorimeter isolation	Track isolation
10 – 15	0.06	0.06
15 – 20	0.12	0.08
20 – 25	0.18	0.12
> 25	0.30	0.12

Table 5.5: p_T dependent isolation requirements for muons. Muons are required to have the amount of calorimeter or track based cone sums be less than this fraction of their p_T .

5.3.2 ELECTRONS

Electrons are identified by matching reconstructed clusters in the electromagnetic calorimeter with tracks in the inner detector. The electrons are identified using a likelihood based method[14, 57] which takes into account the shower shapes in the calorimeter, the matching of tracks to clusters, and the amount of transition radiation in the TRT. The electrons are required to have $|\eta| < 2.47$, and candidates in the transition region between the barrel and endcap ($1.37 < |\eta| < 1.52$) are excluded. As the muons, the electrons are required to have transverse impact parameter significance < 3 , while in the longitudinal direction they must have $|z_0 \sin \theta| < 0.4$ mm. Some electron requirements also vary with electron E_T , and these requirements are summarized in table 5.6.

The isolation for electrons are defined similarly to the muons but with unique cuts on the objects included. The track-based isolation is defined using tracks with $p_T > 400$ MeV with cone sizes as defined previously. The calorimeter-based isolation also uses the same cone size as the muon, but here the cells within a 0.125×0.175 area in $\eta \times \phi$ around the electron cluster's barycenter are excluded. The other difference with respect to muons is that the denominator of the isolation ratio is the electron's E_T rather than p_T . The isolation cuts very with electron E_T and are defined in table 5.6.

The electron is also required to not be consistent with a vertex coming from a photon conversion.

5.3.3 JETS

Jets are clustered with the anti- k_T reconstruction algorithm using a radius parameter of $R = 0.4$. They are required to have a jet vertex fraction (JVF) of at least 50%, meaning that half of the tracks associated

p_T range (GeV)	Quality cut	Calorimeter isolation	Track isolation
10 – 15	Very tight LH	0.20	0.06
15 – 20	Very tight LH	0.24	0.08
20 – 25	Very tight LH	0.28	0.10
> 25	Medium	0.28	0.10

Table 5.6: p_T dependent requirements for electrons. Electrons are required to have the amount of calorimeter or track based cone sums be less than this fraction of their E_T .

with the jet originated from the primary vertex. Jets with no tracks associated (i.e. those outside the acceptance of the ID) do not have this requirement applied. Jets are required to have $p_T > 25$ GeV if they are within the tracking acceptance ($|\eta| < 2.4$). Jets with $2.4 < |\eta| < 4.5$ are required to have $p_T > 30$ GeV. This tighter requirement reduces jets from pileup in the region where JV requirements cannot be applied. The two highest p_T jets in the event are referred to as the “VBF” jets and used to compute various analysis selections later.

Identification of b -jets is done using the MV1 algorithm and is limited to the acceptance of the ID ($|\eta| < 2.5$). The operating point of MV1 that is used is the one that is 85% efficient for identifying true b -jets. This operating point has a 10.3% of mis-tagging a light quark jet as a b -jet. In order to improve the rejection of b -jets, a lower threshold than the nominal p_T threshold described above is used. For the purposes of counting the number of b -jets, jets with p_T down to 20 GeV are used.

5.3.4 OVERLAP REMOVAL

There are some cases where certain reconstructed objects will overlap and one will have to be chosen (for example, an electron and a jet in the calorimeter). First, the case of lepton overlap is dealt with. If an electron candidate extends into the muon spectrometer, it is removed. If a muon or electron have a $\Delta R < 0.1$, the electron is removed and the muon is kept. If two electron candidates overlap within the same radius, then the higher E_T electron is kept. Next, the overlap between leptons and jets is considered. If an electron and jet are within $\Delta R < 0.3$ of one another, the electron is kept and the jet is removed. However, if a muon and jet overlap within $\Delta R < 0.3$, the jet is kept (as it is likely that the muon is the result of a semileptonic decay inside the jet).

Once the overlap removal is complete, the final set of objects used in the analysis is defined.

5.4 ANALYSIS SELECTION

The VBF analysis uses two distinct selections. The first is a looser selection that uses a Boosted Decision Tree (BDT) score as the final discriminator in order to take advantage of the detailed correlations between the VBF variables. The second is a more standard selection, referred to as “cut-based”, that applies cuts on the VBF variables and uses m_T as the final discriminating variable. While the BDT analysis is ultimately more sensitive, the cut-based serves as an important component of the analysis. First, the cut-based allows for confirming the modeling and validity of many variables used as input to the BDT. Second, because this is the first use of such an MVA technique in the $H \rightarrow WW^*$ analysis, the cut-based selection allows confirmation of the final BDT result with a more traditional analysis. Both analyses will be discussed here.

One important note is that because this analysis is dedicated to the measurement of the VBF production mode of the Higgs, events coming from gluon fusion production with the Higgs decaying via $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ are treated as background events. This will be seen throughout the various cutflow tables and yields shown.

5.4.1 COMMON PRE-SELECTION

Both the BDT and cut-based analyses have a common pre-selection that is applied before their main cuts. The cuts on leptons are common to all n_j bins. The analysis requires two oppositely charged leptons, with the leading lepton required to have $p_T > 22$ GeV while the subleading lepton must have $p_T > 10$ GeV. Next, to cut out low mass Z/γ^* events, a cut on the dilepton mass $m_{\ell\ell} > 10$ (12) GeV is applied in the different (same) flavor channel. In the same flavor channels, there is an additional veto placed on the region around the Z peak, requiring that $|m_{\ell\ell} - m_Z| > 15$ GeV.

There are also requirements on the amount of missing transverse momentum in the event. These are only applied in the same flavor channels, as in the different flavor channels $t\bar{t}$ is the dominant background in $n_j \geq 2$. The BDT analysis requires $p_T^{\text{miss}} > 40$ GeV and $E_T^{\text{miss}} > 45$ GeV. The cut-based analysis must cut tighter on these variables to have maximal sensitivity and thus requires $p_T^{\text{miss}} > 50$ GeV and $E_T^{\text{miss}} > 55$ GeV.

Finally, because this analysis is focused on VBF, a requirement on the jet multiplicity is placed, with $n_j \geq 2$. Additionally, the analysis requires that there are no jets identified as b-quarks in the event, or $n_b = 0$.

5.4.2 CUT-BASED SELECTION

The cut-based selection places sequential requirements on variables reconstructed from the VBF jets in order to increase the signal to background ratio.

GENERAL BACKGROUND REDUCTION

Top pair production is the primary background in the $n_j \geq 2$ bin. Even though $n_b = 0$ is required, an additional variable is constructed to further suppress the top background. There is often additional QCD radiation that accompanies the $t\bar{t}$ system when it is produced. Therefore, a variable which tests for the presence of this additional radiation, p_T^{sum} , is constructed. It is defined in equation 5.1.

$$p_T^{\text{sum}} = p_T^{\ell\ell} + p_T^{\text{miss}} + \sum p_T^j \quad (5.1)$$

The first cut after pre-selection in the cut-based analysis requires $p_T^{\text{sum}} < 15$ GeV to further suppress $t\bar{t}$ production.

In the different flavor channels, a cut is made to reduce the contamination from $Z \rightarrow \tau\tau$ decays. The di- τ invariant mass, $m_{\tau\tau}$, is constructed by assuming that the neutrinos from the τ decays were collinear with the leptons[?]. The analysis requires that this mass not be consistent with a Z by requiring $m_{\tau\tau} < m_Z - 25$ GeV.

VBF TOPOLOGICAL CUTS

The characteristic feature of VBF production of the Higgs is the presence of two additional forward jets coming from the incoming partons which radiate the vector bosons that make the Higgs. These jets are forward because the outgoing partons still carry the longitudinal momentum of the incoming partons. Figure 5.2 shows the distribution of the η for the leading jet in a VBF event compared to a background

top pair production event. As can be seen, the VBF jets tend to be more forward in η , while the $t\bar{t}$ jets are more central.

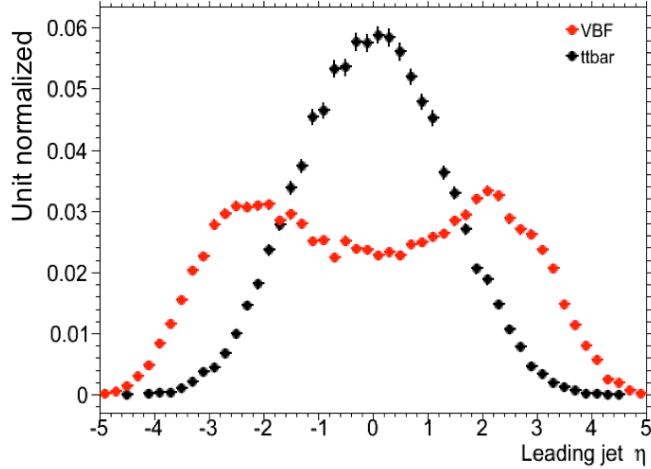


Figure 5.2: Leading jet η in VBF $H \rightarrow WW^*$ (red) and $t\bar{t}$ (black)

Because the cross section for VBF production is an order of magnitude smaller than gluon fusion production, these forward jets must be used in order to better reduce background and achieve a good signal to background ratio. The dedicated VBF search selection requirements are constructed to maximally exploit the features of the unique VBF topology.

Requirements on the VBF jets are collectively referred to as the “VBF topological cuts”. First, a requirement on the dijet invariant mass of the VBF jets, m_{jj} , is placed, requiring $m_{jj} > 600$ GeV. Next, the event is required to have a large gap in rapidity between the two VBF jets, or $\Delta y_{jj} > 3.6$. Both of these cuts put tight requirements on the presence of two forward, high p_T jets moving in opposite directions in the longitudinal plane.

Beyond requiring the presence of the two forward VBF jets, the analysis also vetoes on the presence of any additional jets that fall between the two VBF jets. This cut is referred to as the central jet veto, or CJV. Any events with a third jet with $p_T > 20$ GeV whose rapidity is between the region defined by the two VBF jets are vetoed. This can be expressed in terms of a variable called the jet centrality, defined in

equation 5.2.

$$C_{j3} = \left| \eta_{j3} - \frac{\eta_{j1} + \eta_{j2}}{2} \right| / \frac{|\eta_{j1} - \eta_{j2}|}{2}, \quad (5.2)$$

Here, η_{j1} and η_{j2} are the pseudorapidities of the leading and subleading jets, respectively, while η_{j3} is the pseudorapidity of the extra jet in the event (if one exists). Intuitively, C_{j3} is zero when η_{j3} is directly centered between the two jets and unity when η_{j3} is aligned with either of the VBF jets. Thus, the CJV can be expressed as a requirement that $C_{j3} > 1$.

The decay products of the Higgs tend to be central as well. Thus, the analysis also requires that both leptons in the analysis fall within the rapidity gap defined by the jets. This cut is referred to as the outside lepton veto, or OLV. A quantitative way to define the cut is to require that the centrality of each lepton (defined analogously to that of the third jet in equation 5.2) correspond to the lepton being within the jet rapidity gap, or $C_\ell < 1$ for both leptons.

Figure 5.3a-c shows the m_{jj} , Δy_{jj} , and $C_{\ell 1}$ variables at the stage where all previous cuts in the sequence have been made. The agreement between data and Monte Carlo is good, and the bottom panels show their power in discriminating the VBF signal from the background processes.

The final signal region is also split into two bins of m_{jj} , with the first bin corresponding to $600 \text{ GeV} < m_{jj} < 1 \text{ TeV}$ and the second bin corresponding to $m_{jj} > 1 \text{ TeV}$. The first bin has more statistics but also a larger contribution from background, while the second bin has lower statistics but a 1:1 signal to background ratio.

HIGGS TOPOLOGICAL CUTS

The final state leptons will exhibit unique correlations due to the fact that they are arising from the decay of a spin zero resonance. In particular, the spins of the final state leptons and neutrinos must all cancel, as shown in figure 5.4. Because the neutrino has a left handed chirality and the anti-neutrino has a right handed chirality (in the massless neutrino approximation), the spin and momentum of the particles will be anti-aligned and aligned, respectively. In the transverse plane, the momenta of all four final state objects must cancel as well. With the constraint of having both the momenta and the spin alignments

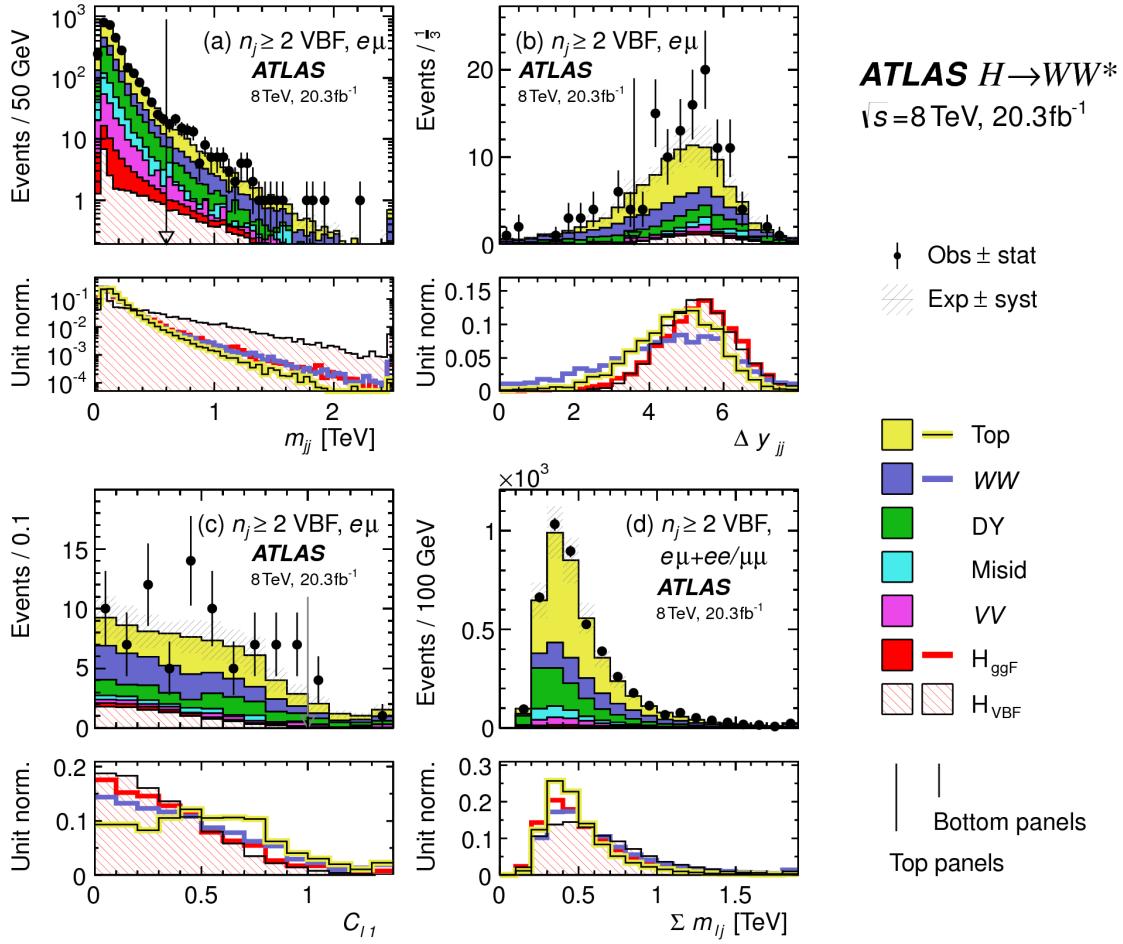


Figure 5.3: Distributions of (a) m_{jj} , (b) Δy_{jj} , (c) $C_{\ell 1}$, and (d) $\sum m_{ij}$, for the VBF analysis. The top panels compare simulation and data, while the bottom panels show normalized distributions for all background processes and signal[20].

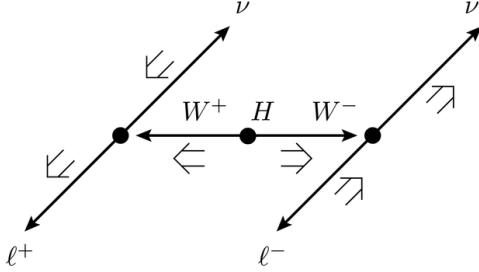


Figure 5.4: A cartoon of the WW final state. Momenta are represented with thin arrows, spins with thick arrows. [20]

cancel, the final state kinematics strongly prefer having a small angle between the leptons in the transverse plane (low $\Delta\phi_{\ell\ell}$). This angular correlation will also lead to low values of the di-lepton invariant mass $m_{\ell\ell}$. These unique signal final state kinematic correlations will be exploited to define the ultimate signal region.

The analysis places additional requirements on the final state leptons. Two requirements on dilepton kinematics are made that are common with lower multiplicity jet bins as well. The angle between leptons in the transverse plane, $\Delta\phi_{\ell\ell}$, is required to be less than 1.8 radians. Additionally, the dilepton mass $m_{\ell\ell}$ is required to be less than 50 GeV.

The cut-based analysis uses m_T as the final discriminating variable as in the ggF focused analysis. The optimal number of bins in m_T was found to be three bins, with the bin boundaries at 80 and 130 GeV.

Table 5.7 shows the data and estimated signal and background yields from simulation as each cut described above is made. The table shows how each cut reduces specific backgrounds and how the overall signal to background ratio grows through the cutflow.

Figure 5.5 shows an ATLAS event display of a candidate event in the final signal region.

5.4.3 BDT BASED SELECTION

The boosted decision tree based analysis takes a different philosophy compared to the cut-based. Rather than cutting sequentially on many variables, the BDT analysis uses many of these variables as inputs to the BDT and the output BDT score as the final discriminant. The BDT is trained with the VBF $H \rightarrow WW^*$ simulation as the signal samples and all other processes as background, including ggF

Table 5.7: Event selection for the $n_j \geq 2$ VBF analysis in the 8 TeV cut-based analysis[20].

Selection	Summary										Composition of N_{bkg}							
	$N_{\text{obs}}/N_{\text{bkg}}$	N_{obs}	N_{bkg}	N_{signal}	N_{ggF}	N_{VBF}	N_{VH}	N_{WW}^{QCD}	N_{WW}^{EW}	$N_{t\bar{t}}$	N_t	N_{top}	N_{Wj}	N_{jj}	N_{VV}	N_{misid}	$N_{\text{Drell-Yan}}$	$N_{ee/\mu\mu} N_{\tau\tau}^{\text{QCD}}$
$e\mu$ sample	1.00 ± 0.00	61434	61180	85	32	26	1350	68	51810	2970	847	308	380	51	3260	46		
$n_b = 0$	1.02 ± 0.01	7818	7700	63	26	16	993	43	3000	367	313	193	273	35	2400	29		
$p_T^{\text{sum}} < 15$	1.03 ± 0.01	5787	5630	46	23	13	781	38	1910	270	216	107	201	27	2010	23		
$m_{\tau\tau} < m_Z - 25$	1.05 ± 0.02	3129	2970	40	20	9.9	484	22	1270	177	141	66	132	7.6	627	5.8		
$m_{jj} > 600$	1.31 ± 0.12	131	100	2.3	8.2	—	18	8.9	40	5.3	1.8	2.4	5.1	0.1	15	1.0		
$\Delta y_{jj} > 3.6$	1.33 ± 0.13	107	80	2.1	7.9	—	11.7	6.9	35	5.0	1.6	2.3	3.3	—	11.6	0.8		
$C_{j3} > 1$	1.36 ± 0.18	58	43	1.3	6.6	—	6.9	5.6	14	3.0	1.3	1.3	2.0	—	6.8	0.6		
$C_{\ell 1} < 1, C_{\ell 2} < 1$	1.42 ± 0.20	51	36	1.2	6.4	—	5.9	5.2	10.8	2.5	1.3	1.3	1.6	—	5.7	0.6		
$m_{\ell\ell}, \Delta\phi_{\ell\ell}, m_\tau$	2.53 ± 0.71	14	5.5	0.8	4.7	—	1.0	0.5	1.1	0.3	0.3	0.3	0.6	—	0.5	0.2		
<hr/>																		
$ee/\mu\mu$ sample	0.99 ± 0.01	26949	27190	31	14	10.1	594	37	23440	1320	230	8.6	137	690	679	16		
$n_b, p_T^{\text{sum}}, m_{\tau\tau}$	1.03 ± 0.03	1344	1310	13	8.0	4.0	229	12.0	633	86	26	0.9	45	187	76	1.5		
$m_{jj}, \Delta y_{jj}, C_{j3}, C_\ell$	1.39 ± 0.28	26	19	0.4	2.9	0.0	3.1	3.1	5.5	1.0	0.2	0.0	0.7	3.8	0.7	0.1		
$m_{\ell\ell}, \Delta\phi_{\ell\ell}, m_\tau$	1.63 ± 0.69	6	3.7	0.3	2.2	0.0	0.4	0.2	0.6	0.2	0.2	0.0	0.1	1.5	0.3	0.1		

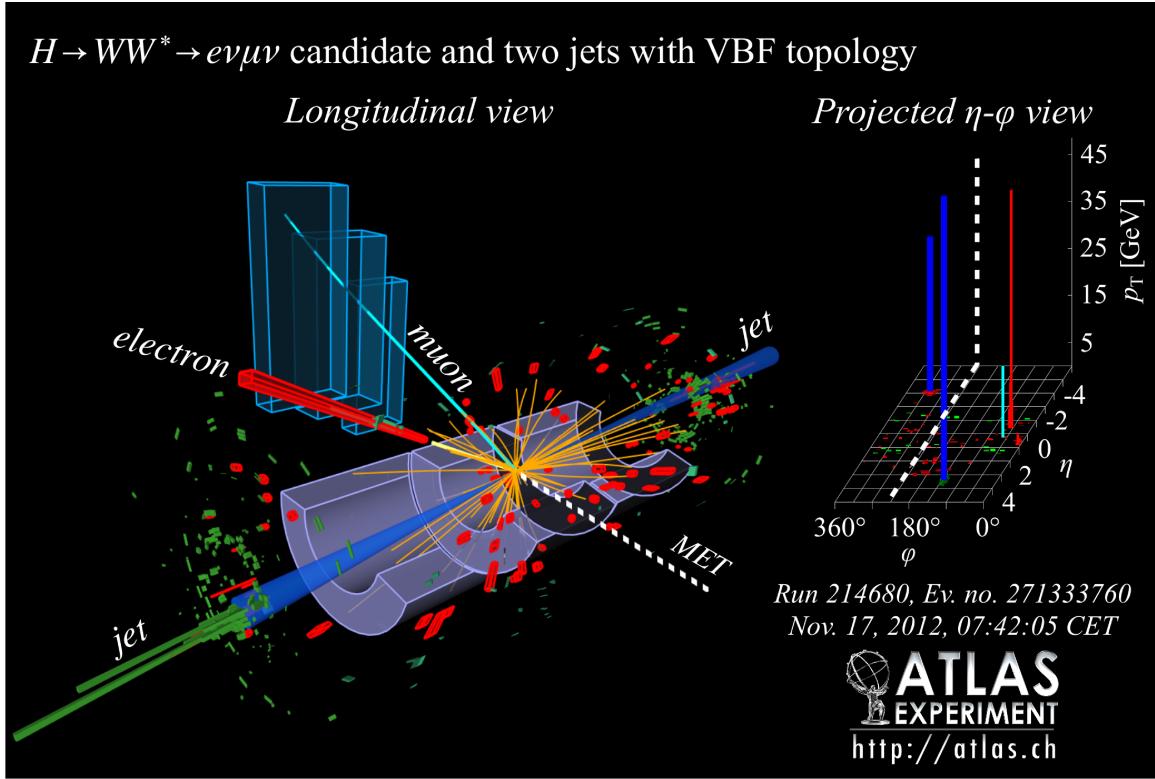


Figure 5.5: Event display of a VBF candidate event[20].

$H \rightarrow WW^*$ production.

PRE-TRAINING SELECTION AND BDT INPUTS

Before training, the common preselection cuts described in section 5.4.1 are applied. Additionally, the central jet veto and outside lepton veto described in section 5.4.2 are applied. The BDT has eight input variables, six of which are also variables that are used in the cut-based analysis. The six shared variables are p_T^{sum} , m_{jj} , Δy_{jj} , $m_{\ell\ell}$, $\Delta\phi_{\ell\ell}$, and m_T . The seventh variable input in the BDT is a combination of the variables used to do the OLV in the cut-based analysis. The BDT uses as input the sum of lepton centralities, or $\sum C_\ell = C_{\ell 1} + C_{\ell 2}$. The final BDT input variable, $\sum m_{\ell j}$, is constructed to account for the correlations between the jets and leptons in the event. It is the sum of the invariant masses of all four

possible lepton-jet combinations, shown in equation 5.3.

$$\Sigma m_{\ell j} = m_{\ell 1,j1} + m_{\ell 1,j2} + m_{\ell 2,j1} + m_{\ell 2,j2} \quad (5.3)$$

Figure 5.3d shows the agreement between data and simulation for the $\Sigma m_{\ell j}$ variable, as well as showing its discriminating power. Figure 5.6 shows the distributions of the Higgs topological variables used as BDT inputs for VBF signal and corresponding backgrounds. Figure 5.7 shows the distributions of the VBF topological variables used as BDT inputs. In both cases, the VBF yield has been scaled by a factor of 50 to better show the shape difference compared to the backgrounds.

Table ?? summarizes the cuts applied for the BDT and cut-based analyses, as well as which variables are used as input to the BDT.

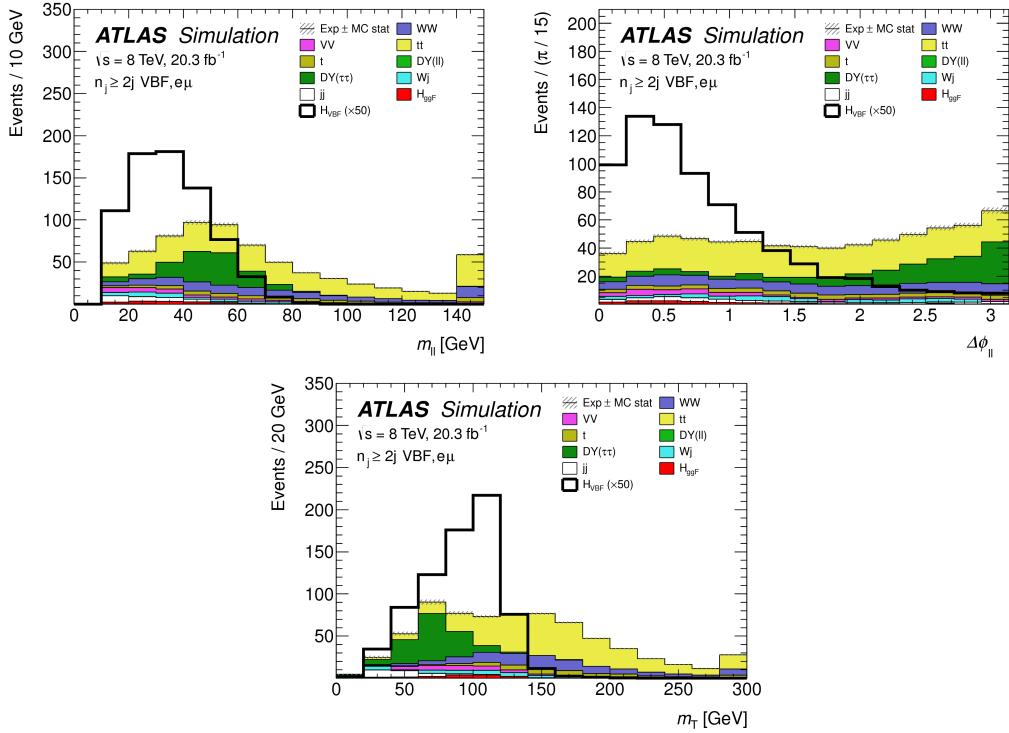


Figure 5.6: Distributions of m_{ll} (top left), $\Delta\phi_{ll}$ (top right), and m_T (bottom) for the VBF BDT analysis. These are plotted after all of the BDT pre-training selection cuts[20].

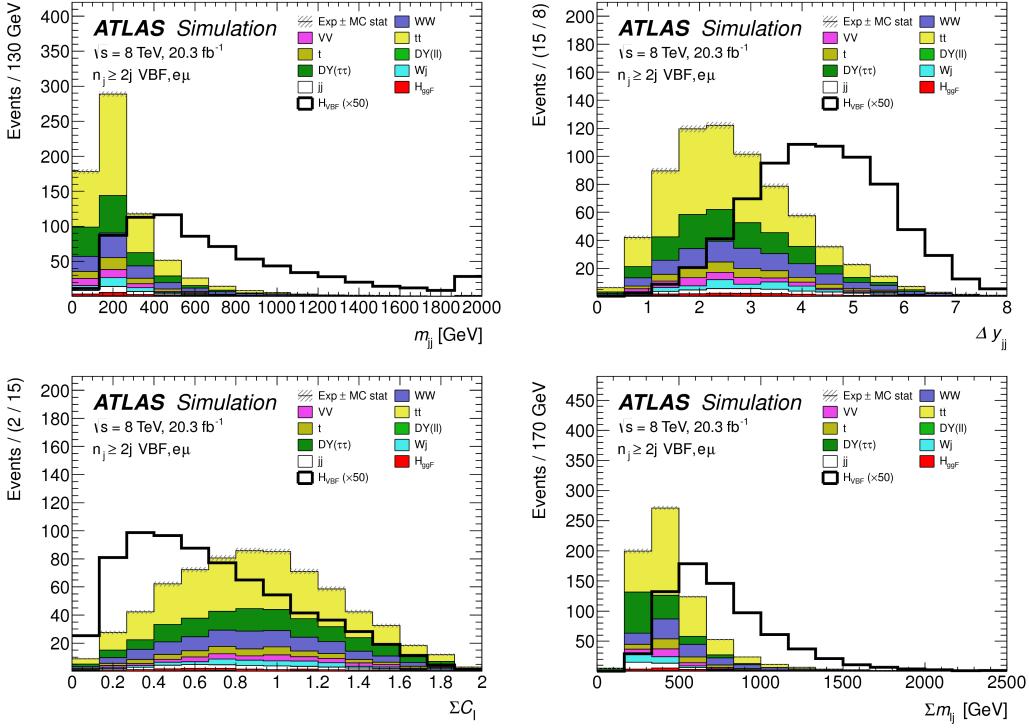


Figure 5.7: Distributions of m_{jj} (top left), Δy_{jj} (top right), $\sum C_\ell$ (bottom left), and $\sum m_{\ell j}$ (bottom right) for the VBF analysis. These are plotted after all of the BDT pre-training selection cuts[20].

BDT OUTPUT

After training, the BDT outputs a score (O_{BDT}) which is in the range $[-1, 1]$, where -1 corresponds to background-like events and $+1$ corresponds to signal-like events. Figure 5.8 shows the output BDT distribution in both the different flavor and same flavor channels. For the final discriminant analysis, the O_{BDT} distribution is divided into four bins, with boundaries at $[-1, -0.48, -0.3, 0.78, 1]$. The bins are numbered from 0, 1, 2, 3 respectively. Because bin 0 is predominantly background, it is excluded from the likelihood analysis.

5.5 BACKGROUND ESTIMATION

This section describes the procedures used to estimate backgrounds for the VBF analysis in both the cut-based and BDT analyses.

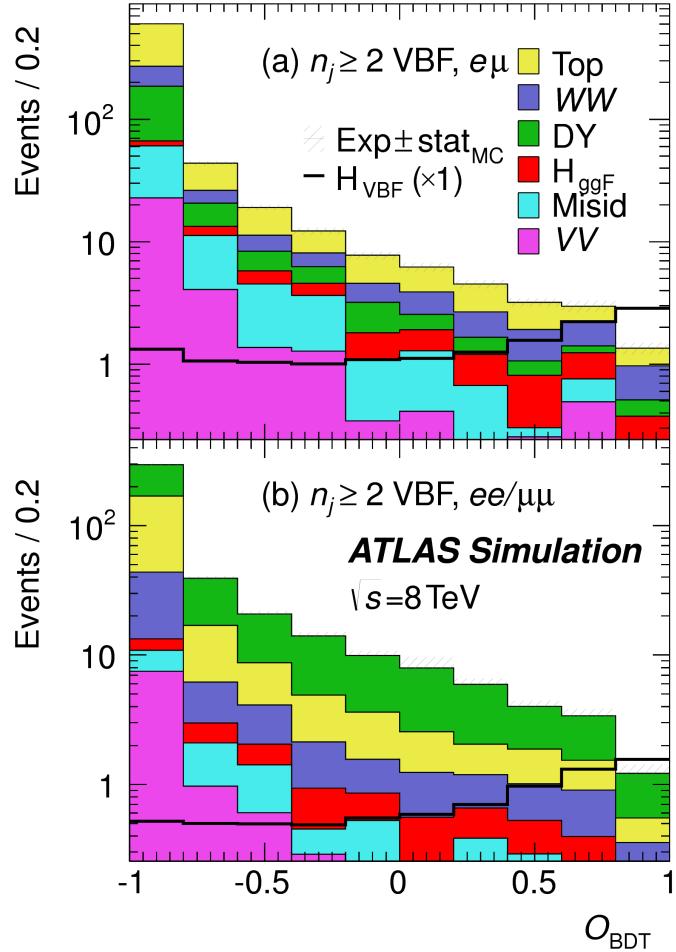


Figure 5.8: Distributions of O_{BDT} for the VBF signal and associated backgrounds after the VBF pre-training selection[20].

5.5.1 GENERAL STRATEGY

Most of the backgrounds in the VBF analysis have shapes estimated from Monte Carlo simulation but normalizations derived from control regions in data. In essence, a normalization factor (denoted with β or abbreviated as NF) is derived by scaling the MC yield in the control region to the corresponding yield in data. Once this factor is derived, it can be used to scale the MC estimate of the background in the signal region. This is illustrated in equation 5.4.

$$B_{\text{SR}}^{\text{est}} = B_{\text{SR}} \times \frac{N_{\text{CR}}}{B_{\text{CR}}} \equiv B_{\text{SR}} \times \beta \quad (5.4)$$

Here, B denotes the MC yield prediction in the denoted region, while N denotes the observed number of events in data in the denoted region.

Another way of writing the same equation, in terms of an extrapolation factor α rather than a normalization factor β . The overall calculation is exactly the same. However, when phrased in this way, it shows how the uncertainty on the background estimation can be reduced. This is shown in equation 5.5.

$$B_{\text{SR}}^{\text{est}} = N_{\text{CR}} \times \frac{B_{\text{SR}}}{B_{\text{CR}}} \equiv N_{\text{CR}} \times \alpha \quad (5.5)$$

Phrased this way, the equation shows that with enough statistics in the control region, a large theoretical uncertainty on the overall background yield in the signal region can be replaced by a small statistical uncertainty coming from the number of data events in the CR and a smaller theoretical uncertainty on the extrapolation from the control region to the signal region.

5.5.2 TOP BACKGROUND

The normalization factor β_t for the top background in the VBF analysis is derived in a region required to have one b-tagged jet, or $n_b = 1$. In the cut-based analysis, normalization factors are computed at every stage of the cutflow by applying the appropriate cuts in the CR. These NF are then applied to the $t\bar{t}$ and single top event yields in the SR. In the BDT analysis, a single normalization factor is computed for each bin of O_{BDT} after applying the BDT pre-training cuts described previously. The computed normalization factors are derived with all flavor combinations combined in order to decrease statistical uncertainty. Additionally, in the BDT analysis, BDT bins 2 and 3 are merged for the same reason.

Table 5.8 shows the evolution of the β_t through the cut-based selection. Table 5.9 shows the value of the β_t in each bin of O_{BDT} . In all cases, the computed factors are relatively consistent with unity, with the largest discrepancy coming in bin 1 of O_{BDT} .

Figure 5.9 shows the m_{jj} and O_{BDT} distributions in the top control region. Overall the modeling looks consistent with the data.

While these normalization factors can be computed and applied to the expected background yields listed in tables like table 5.7, in the end the normalization of the top background is profiled (meaning

Cut	β_t
$p_T^{\text{sum}} < 15 \text{ GeV}$	1.03 ± 0.01
$m_{\tau\tau} < m_Z - 25$	1.05 ± 0.01
$m_{jj} > 600 \text{ GeV}$	0.96 ± 0.06
$\Delta y_{jj} > 3.6$	1.02 ± 0.08
CJV	1.13 ± 0.16
OLV	1.01 ± 0.19
$m_{jj} < 1 \text{ TeV}$	0.94 ± 0.19
$m_{jj} > 1 \text{ TeV}$	1.48 ± 0.66

Table 5.8: Top normalization factors computed at each stage of the cut-based selection. Uncertainties are statistical only.

O_{BDT}	β_t
Bin0	1.09 ± 0.02
Bin1	1.58 ± 0.15
Bin2	0.95 ± 0.31
Bin3	0.95 ± 0.31

Table 5.9: Top normalization factors computed for each bin of O_{BDT} . Uncertainties are statistical only.

there is a dedicated Poisson constraint) and allowed to float in the final statistical fit.

5.5.3 $Z/\gamma^* \rightarrow \tau\tau$ BACKGROUND

In the different flavor channels, the $Z/\gamma^* \rightarrow \tau\tau$ background is an important one. Di-tau production can produce an $e\mu$ final state if each τ lepton decays to a different flavor lepton.

In the BDT analysis, a single normalization factor for the background is derived. A control region is defined using the pre-training selection cuts, except requiring that $|m_{\tau\tau} - m_Z| < 25 \text{ GeV}$ so that the region is enriched in $Z/\gamma^* \rightarrow \tau\tau$ background. Additional requirements of $m_{\ell\ell} < 80(75) \text{ GeV}$ in the different (same) flavor channel, as well as $O_{\text{BDT}} > -0.48$ are applied to increase the purity of the region. The final $\beta_{Z/\gamma^*\rightarrow\tau\tau}$ is calculated to be 0.9 ± 0.3 (statistical uncertainty only). Because of the small contribution of this background in the BDT analysis and the large statistical uncertainty, no additional systematics are calculated. The final SR estimate is scaled by this β and not allowed to float in the fit.

The cut-based corrections are a bit more involved because they need to be applied cut by cut through the cutflow, as well as in the final signal region for the fit. The region is defined including all SR cuts

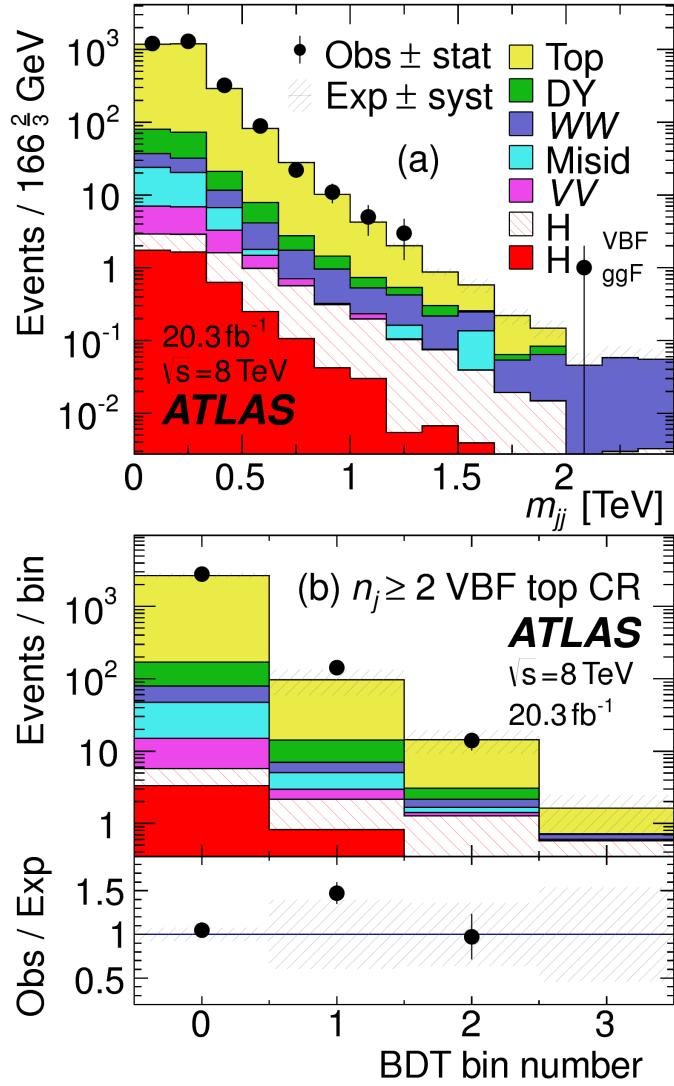


Figure 5.9: Distributions of m_{jj} (a) and O_{BDT} (b) in the VBF $n_b = 1$ top CR[20].

up to the $Z/\gamma^* \rightarrow \tau\tau$ veto, which is instead made into a Z mass peak requirement as for the BDT region. The $m_{\ell\ell}$ cut from the BDT region is included as well. The cut-based approach aims to correct the normalization of the $Z/\gamma^* \rightarrow \tau\tau$ background in two ways. First, an overall normalization factor is computed from the control region. However, the VBF topological cuts are not included in this region, and applying them as is done in the top CR is not feasible due to limited statistics. So, instead, correction factors (CF) to the cut efficiencies of the VBF cuts are derived in a same flavor $Z \rightarrow \ell\ell$ control region, which has significantly more statistics. The CF is simply the ratio of the cut efficiencies in data and MC

derived in this region. In the end, the overall background estimate is given by equation 5.6.

$$N_{Z/\gamma^* \rightarrow \tau\tau}^{\text{est}} = B_{Z/\gamma^* \rightarrow \tau\tau}^{\text{SR}} \times \beta_{\tau\tau} \times \frac{\epsilon_{\text{VBF cuts}}^{\text{data}}}{\epsilon_{\text{VBF cuts}}^{\text{MC}}} \quad (5.6)$$

The hypothesis is that while the normalization correction must be derived in a dedicated region, the efficiency of the VBF cuts should not be sensitive to the type of Z/γ^* process and thus the larger control region can be exploited to derive the CF. Figure 5.10 shows a shape comparison for the m_{jj} variable in $Z \rightarrow \tau\tau$ events in the signal region and $Z \rightarrow \ell\ell$ events in the control region. The figure shows that the shapes are indeed comparable and thus any CF derived in the same flavor control region can reliably be applied in the signal region.

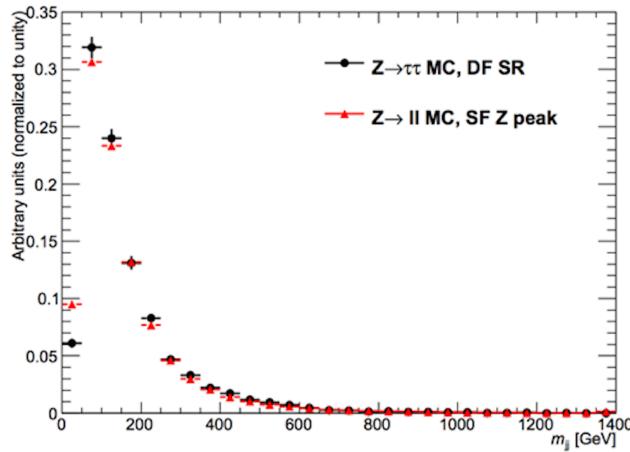


Figure 5.10: Comparison of m_{jj} shape in a same flavor $Z \rightarrow \ell\ell$ control region and the VBF cut-based signal region.

Table 5.10 shows the overall normalization factor $\beta_{\tau\tau}$ and the efficiency correction factors for the various VBF topological cuts. In general, the statistical uncertainties on the cut efficiency corrections are quite good, and the MC tends to underestimate the efficiency of the VBF cuts for the $Z/\gamma^* \rightarrow \tau\tau$ background.

5.5.4 $Z/\gamma^* \rightarrow \ell\ell$ BACKGROUND

In the same flavor channels, the $Z/\gamma^* \rightarrow \ell\ell$ background is dominant and thus must be estimated correctly. In both the BDT and cut-based analyses, the background is estimated using the so-called

$\beta_{\tau\tau}$	0.97 ± 0.04
Cut	Correction factors
$m_{jj} > 600 \text{ GeV}$	1.09 ± 0.01
$\Delta y_{jj} > 3.6$	1.14 ± 0.02
CJV	1.20 ± 0.02
OLV	1.17 ± 0.03
$m_{jj} < 1 \text{ TeV}$	1.17 ± 0.06
$m_{jj} > 1 \text{ TeV}$	1.18 ± 0.13

Table 5.10: $Z/\gamma^* \rightarrow \tau\tau$ correction factors for the VBF cut-based analysis. Uncertainties are statistical only.

“ABCD” method. The ABCD method creates four different regions by defining cuts on two variables. One of the regions (A) is the signal region, while the other regions are defined by inverting one of both of the cuts. in this case, the two variables used are $m_{\ell\ell}$ and E_T^{miss} , because inverting either of the SR cuts on these variables will give regions rich in the $Z/\gamma^* \rightarrow \ell\ell$ background. Figure 5.11 illustrates the general strategy for each region.

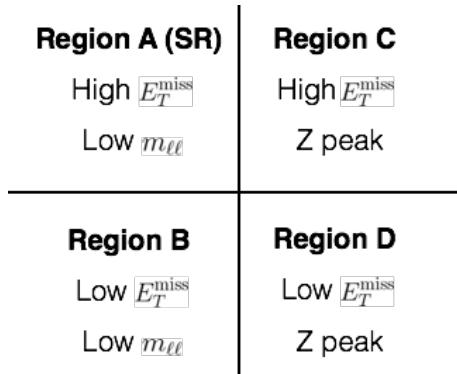


Figure 5.11: General illustration of the ABCD region definitions for $Z/\gamma^* \rightarrow \ell\ell$ background estimation.

In both of the BDT and cut-based analyses, the Z peak region is defined with $|m_{\ell\ell} - m_Z| < 15 \text{ GeV}$. In the BDT analysis, low $m_{\ell\ell}$ corresponds to $m_{\ell\ell} < 75 \text{ GeV}$ while in the cut-based it is $m_{\ell\ell} < 50 \text{ GeV}$ (as this defines the cut-based SR). The BDT low E_T^{miss} region is between 25 and 45 GeV, while the high E_T^{miss} region is $E_T^{\text{miss}} > 45 \text{ GeV}$. In the cut-based, high and low E_T^{miss} are defined as opposite ends of the 55 GeV cut applied for the signal region definition.

Once the regions are defined, the final signal region background estimate is done by taking the estimate in region B and extrapolating it to the signal region (A) by multiplying it by the ratio of regions

C and D. Effectively, the Z peak region is used to estimate the efficiency of the E_T^{miss} cut in data, and then this efficiency is applied in the low $m_{\ell\ell}$ region. An additional correction is also applied for the non-closure of the method in MC. This is summarized in equations 5.7 and 5.8.

$$N_{Z/\gamma^* \rightarrow \ell\ell}^{\text{SR}} = N_{Z/\gamma^* \rightarrow \ell\ell}^{\text{B}} \times \frac{N_{Z/\gamma^* \rightarrow \ell\ell}^{\text{C}}}{N_{Z/\gamma^* \rightarrow \ell\ell}^{\text{D}}} \times f_{\text{corr}} \quad (5.7)$$

$$f_{\text{corr}} = \frac{B_{\text{MC}}^{\text{A}}/B_{\text{MC}}^{\text{B}}}{B_{\text{MC}}^{\text{C}}/B_{\text{MC}}^{\text{D}}} \quad (5.8)$$

Here, the N refer to data yields in each region with the non Z/γ^* backgrounds subtracted, while B refer to the Z/γ^* yields in MC in each region.

A normalization factor $\beta_{\ell\ell}$ is computed for each analysis as the ratio of the predicted data yield to the MC yield in the SR. The shape of the BDT distribution is taken from data region B, while the shape of the m_T distribution in the cut-based analysis is taken from Z/γ^* MC in the SR. The values of the $\beta_{\ell\ell}$ in the BDT and cut-based analyses from this method are summarized in table 5.11. In the cut-based analysis, the same cut efficiency correction factors shown in table 5.10 are also applied (in product with the $\beta_{\ell\ell}$) in the same flavor channels to this background, as they were derived in the Z peak region.

	β_t
BDT Bin 1	1.01 ± 0.15
BDT Bin 2	0.89 ± 0.28
Cut-based	0.81 ± 0.21

Table 5.11: $Z/\gamma^* \rightarrow \ell\ell$ normalization factors for cut-based and BDT analyses. Uncertainties are statistical only.

5.5.5 WW AND OTHER DIBOSON BACKGROUNDS

The WW and other diboson backgrounds have both their shape and normalization taken from MC simulation. They are validated in dedicated control regions and found to agree with data well.

As WW is the largest of these backgrounds and is irreducible, validating the estimate is of particular importance. The validation region is constructed by requiring the pre-selection cuts on leptons and $m_{\ell\ell}$, $n_b = 0$, and $m_T > 100$ GeV. The m_{T2} variable[?] is an additional discriminant that will isolate

the WW background, and a requirement of $m_{T2} > 160$ GeV is placed to define the WW validation region. This cut gives a 60% purity for the validation region. The derived normalization factor in the region is 1.15 ± 0.19 and is thus consistent with unity. Figure 5.12 shows the m_{T2} distribution and how it distinguishes the WW background.

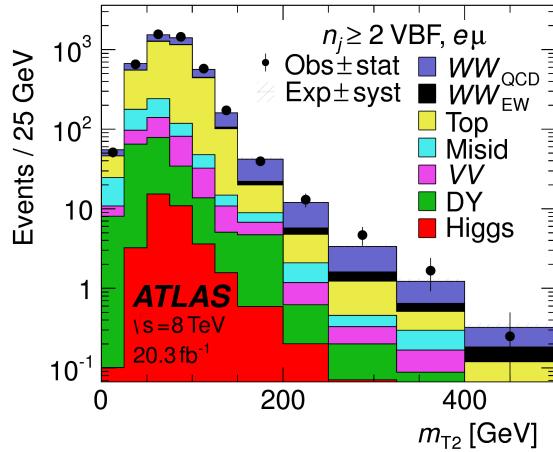


Figure 5.12: Distribution of m_{T2} in the WW validation region of the VBF analysis[20].

5.5.6 HIGGS PRODUCTION VIA GLUON-GLUON FUSION

Because this analysis is dedicated to measuring the VBF contribution to Higgs production, the component of Higgs production from gluon-gluon fusion is treated as a background. The shape is taken directly from simulation, using the generators described in table 5.4. In the final combined fit of all different signal regions, the normalization is controlled by either a combined signal strength parameter μ , which controls the normalization of both ggF and VBF production, or a separate parameter μ_{ggF} depending on the interpretation being presented in the final results.

5.5.7 BACKGROUNDS WITH MISIDENTIFIED LEPTONS

As discussed previously, the $W+jets$ and QCD multijet backgrounds are derived with fully data-driven methods. These backgrounds do not make a large contribution to the final VBF signal region but their estimation methods are discussed briefly here.

W +JETS BACKGROUND

The W +jets background enters the signal region by having one of the jets mis-reconstructed as a lepton. The background is estimated by constructing a control sample with two leptons, where one lepton passes the usual lepton quality cuts but the second lepton fails one of those cuts (also known as the “anti-identified” lepton). This control region is rich in the W +jets contribution because if a second lepton is reconstructed in a W +jets event it is likely to be poor quality. The purity of this W +jets control sample is 85% to 90% depending on the exact configuration of leptons in the final state.

The signal region estimate of W +jets is estimated by extrapolation from the control sample to the signal region using extrapolation factors derived in a Z +jets control sample in data. The extrapolation factor is the ratio of the number of lepton candidates satisfying all quality criteria to the number of lepton candidates anti-identified. This ratio is measured in bins of p_T and η . Thus, the final signal region estimate (binned as the extrapolation factor is binned) is simply the number of events in the anti-identified lepton control sample multiplied by the extrapolation factor derived from the Z +jets control sample. Figure 5.13 shows the extrapolation factors derived for electrons and muons.

QCD MULTIJET BACKGROUND

The method for estimating the multijet background is very similar to the W +jets estimation method. The control sample in this case has two anti-identified leptons but otherwise satisfies all signal region requirements. The extrapolation factor is estimated from a multijet sample and applied twice to the control sample.

5.5.8 BACKGROUND COMPOSITION IN FINAL SIGNAL REGION

After all of these estimation procedures, the final signal region composition can be calculated. The estimated yields are all shown in table 5.7. Figure 5.14 shows the relative percentages of the different background for the different flavor and same flavor final states. In $e\mu$, the leading backgrounds are top backgrounds, ggF Higgs, and SM WW production. In $ee/\mu\mu$, the leading background is Drell-Yan, followed by top and ggF Higgs.

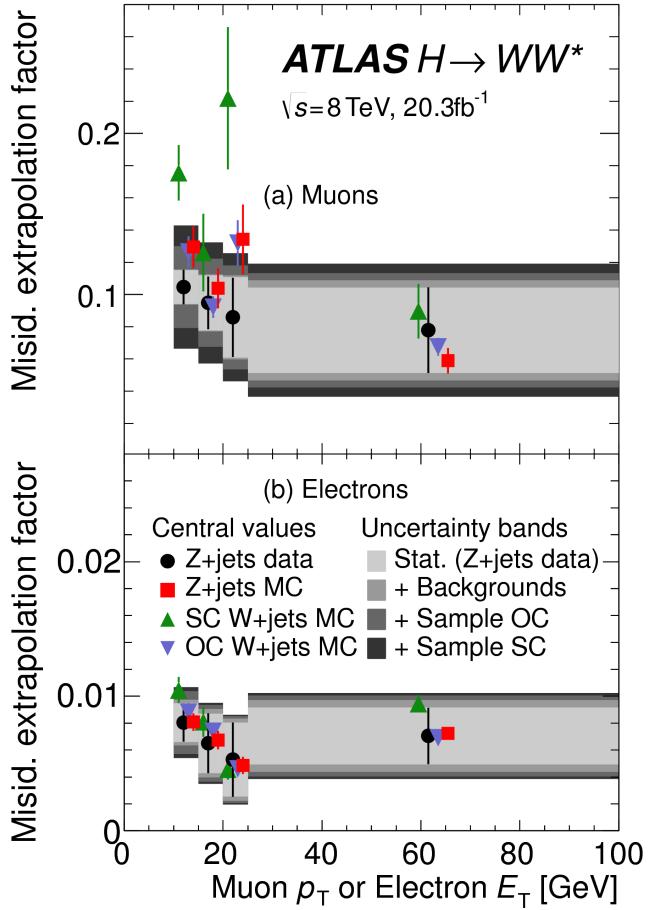


Figure 5.13: Extrapolation factors for the W +jets estimate derived for muons (a) and electrons (b) as a function of lepton p_T [20].

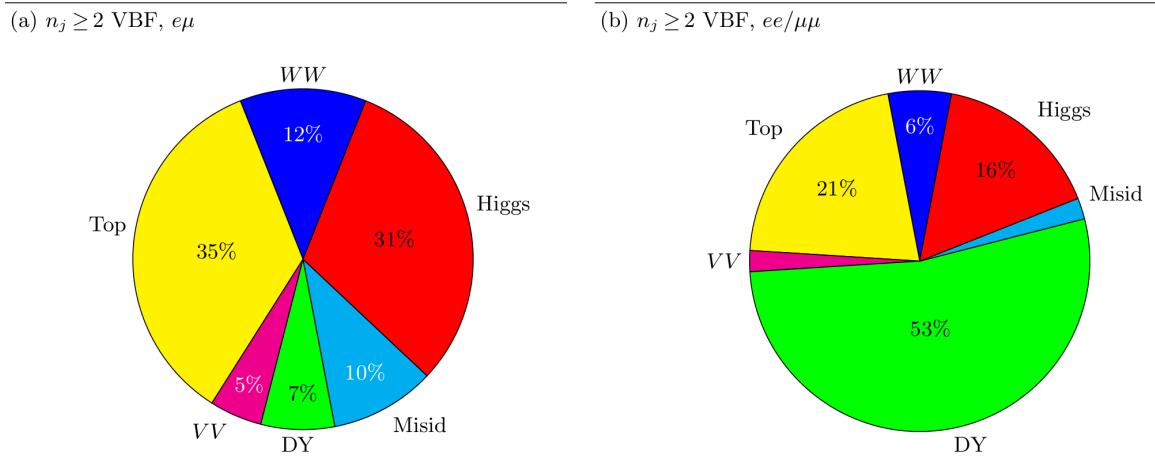


Figure 5.14: Background composition in final VBF signal region[20].

5.6 SYSTEMATIC UNCERTAINTIES

There are two main types of systematic uncertainties that are assessed for the analysis. First, theoretical uncertainties associated with the various signal and background yield estimates are discussed. Then, experimental uncertainties due to detector effects are shown. Normalization uncertainties refer to uncertainties that affect the cross section of the process in question in the signal region being probed. Shape uncertainties refer to systematic uncertainties that affect the shape of the final discriminating variable (either m_T or O_{BDT}).

5.6.1 THEORETICAL UNCERTAINTIES

There are four main components to theoretical uncertainties assigned to signal and background processes taken from Monte Carlo. Each one is a different source of variation in the overall acceptance for that process. The first involves variation of the QCD renormalization and factorization scales used in the calculation. In this case, the two scales are varied independently and simultaneously by factors of two high or low and quantifying the resulting variation in normalization and shape for the process. This approximates the correction to the cross section that would come from including the next order of the QCD calculation (referred to as scale uncertainty). Next, there is an uncertainty associated with the PDF set used in generating the events. The uncertainty eigenvectors for the given PDF set are studied, and the envelope of maximal variation is taken as an uncertainty. Finally, there are two uncertainties associated with the choice of MC software (referred to as PDF uncertainty). An uncertainty associated with the generator chosen for the hard scattering process is evaluated by keeping the parton showering software constant but varying the matrix element generator and taking the maximal variation as an uncertainty (referred to as the generator uncertainty). The converse variation can also be done, where the matrix element generator remains constant and the generator used for the underlying event/parton shower modeling is varied (referred to as the UE/PS uncertainty). In cases where the background is normalized in a control region, the systematic uncertainty arises from variations of the extrapolation factor α between the CR and the SR, which can affect the normalization of the background in the SR.

There are two additional uncertainties that are applied to the Higgs processes as well. First, there are

	VBF H	ggF H	Top	<u>QCD</u> WW
Total σ	2.7	7.2	-	-
Jet binning	-	29	-	-
Scale	3.0	48	5.0	27*
Generator	4.2	-	21	12
PDF	-	3.0	-	4*
UE/PS	14	15	-	2

Table 5.12: Systematic uncertainties for various processes in the BDT analysis, given in units of % change in yield. Values are given for the most sensitive BDT bin (bin 3), except where noted with a *, in which case the uncertainty affect the normalization in all BDT bins. Empty entries indicate that the uncertainty is negligible or not applicable to this background.

uncertainties assigned to the Higgs total production cross section. Then, there are uncertainties assigned based on the fact that the analysis is done in exclusive jet bins and it is possible for signal events to migrate from one bin to the next depending on the presence or absence of jets. These are assigned using the Jet Veto Efficiency (JVE) procedure[?] for ggF events and the Stewart-Tackmann (ST) method[78] for VBF production.

Table 5.12 shows how the different theory uncertainty components contribute to variation in the yields of dominant background processes and signal in the BDT analysis. Table 5.13 shows the total theory uncertainties on the backgrounds in the cut-based analysis. The $Z/\gamma^* \rightarrow \tau\tau$ background is included here because it plays a larger role while in the BDT analysis it is negligible.

Process	Theory syst. (%)
ggF H	48
Top	26
<u>QCD</u> WW	37
$Z/\gamma^* \rightarrow \tau\tau$	6.1

Table 5.13: Systematic uncertainties for various processes in the cut-based VBF analysis, given in units of % change in yield. Values are given for the low m_{jj} signal region.

While the estimate for the same-flavor $Z/\gamma^* \rightarrow \ell\ell$ background is data-driven, there is still a systematic uncertainty taken for the non-closure of the method in Monte Carlo. This is taken as the maximum of the deviation of the non-closure factor f_{corr} from unity and its uncertainty, or $\max(|1 - f_{\text{corr}}|, \delta f_{\text{corr}})$. For the BDT analysis, this uncertainty is 17%, while for the cut-based analysis it is 23%.

For illustration, figures 5.15 and 5.16 show the variations in the extrapolation factor from the PDF and

QCD uncertainties on the top background estimate, binned in m_T , for the cut-based analysis. In both cases, there was no significant shape uncertainty but normalization uncertainties were assigned according to the maximal variation.

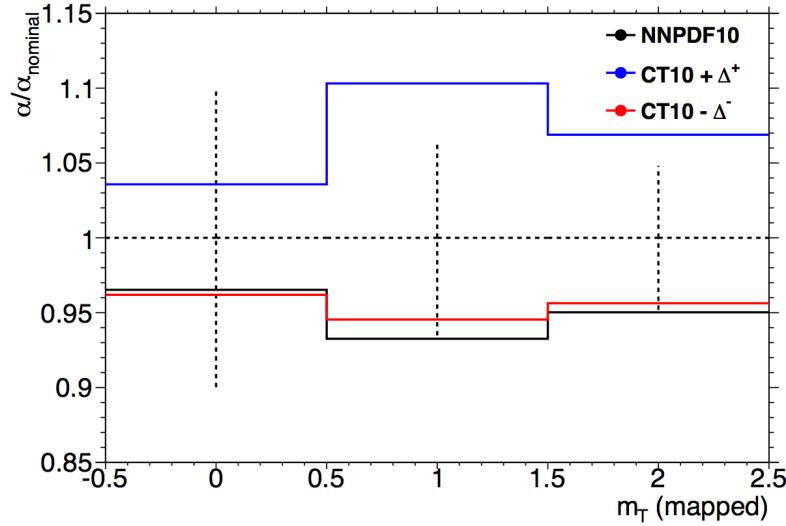


Figure 5.15: Variations in the top background extrapolation factor in the cut-based analysis due to PDF uncertainties, binned in m_T .

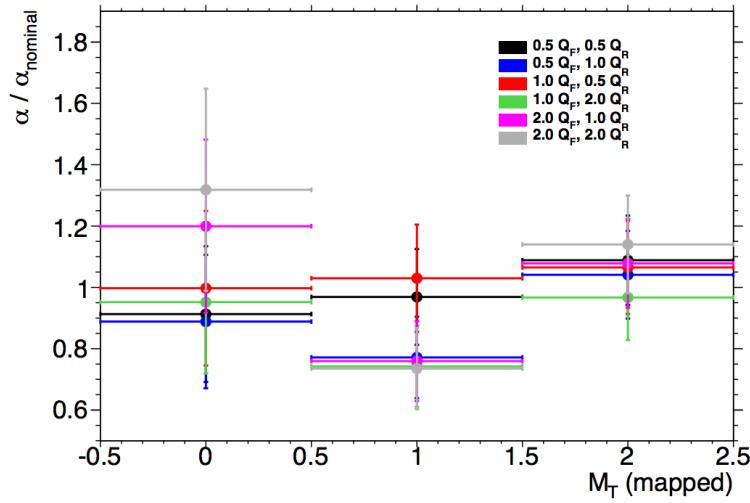


Figure 5.16: Variations in the top background extrapolation factor in the cut-based analysis due to QCD scale uncertainties, binned in m_T .

5.6.2 EXPERIMENTAL UNCERTAINTIES

In this analysis, the theoretical uncertainties end up being the most dominant, but there are some experimental uncertainties that make a contribution as well. The first is the uncertainty on the measured integrated luminosity, which affects backgrounds whose normalization is taken from MC and is measured to be 2.8% in the 8 TeV dataset [79]. The dominant sources of uncertainty overall are uncertainties on the jet energy scale and resolution and the b -tagging efficiency. Additional sources include lepton uncertainties on identification, resolution, and trigger efficiency, as well as uncertainties on the missing transverse momentum .

The jet energy scale uncertainty is split into several independent components, including jet-flavor dependent calorimeter response uncertainties, uncertainties on modeling of pile-up interactions, uncertainties on extrapolation from the central to forward detector regions, and MC non-closure [80]. The uncertainty on energy scale for jets used in this analysis ranges from 1% to 7% depending on the jet p_T and η . The jet energy resolution varies from 5% to 20%, with uncertainties ranging from 2% to 40% (the largest uncertainties occurring at the selection threshold).

The b -tagging efficiency is independently measured in data samples enriched in dileptonic decays of $t\bar{t}$ events or in events where a muon is reconstructed in the vicinity of a jet[81, 82]. The efficiencies and their uncertainties are binned in p_T and decomposed into uncorrelated components using an eigenvector method[?]. Uncertainties on the efficiency range from 1% to 7.8%. The uncertainty on the rate of misidentification of c -jets as b -jets ranges from 6-14%, while the uncertainty on the rate of light jet mis-tagging ranges from 9-19% depending on p_T and η .

The total experimental uncertainties on different signal and background components are summarized in table 5.14. They are compared to the level of other statistical and systematic uncertainties as well. Overall, the experimental uncertainties are sub-dominant compared to the statistical and theoretical uncertainties.

Sample	Total error	Stat. error	Expt. syst. err.	Theo. syst. err.
$n_j \geq 2$ VBF-enriched				
N_{sig}	13	-	6.8	12
N_{bkg}	9.2	4.7	6.4	4.5
N_{WW}	32	-	14	28
N_{top}	15	9.6	7.6	8.5
N_{misid}	22	-	12	19
N_{VV}	20	-	12	15
$N_{\tau\tau} (\text{DY})$	40	25	31	2.9
$N_{ee/\mu\mu} (\text{DY})$	19	11	15	-

Table 5.14: Composition of the post-fit uncertainties (in %) on the total signal (N_{sig}), total background (N_{bkg}), and individual background yields in the VBF analysis[20].

5.7 RESULTS

While the combined results of all the $H \rightarrow WW^*$ sub-analyses will be discussed in the next chapter, this section presents the results of the VBF specific analysis and interpretations.

As table 5.7 shows, the final cut-based signal region contains 20 events in data with $m_T < 150$ GeV, 14 coming from the $e\mu$ channel and 6 coming from the $ee + \mu\mu$ channel. The BDT analysis has many more candidates due to its looser selection, and the yields in each bin of O_{BDT} are shown in table 5.15.

Figure 5.17(a) shows the final distribution of data candidates compared to the expected m_T distribution for signal and background. The data are very consistent with a VBF Higgs hypothesis. Figure 5.17(b) shows where the data candidates fall in the two-dimensional binning of m_T and m_{jj} used in the fit for the cut-based analysis.

Figure 5.18 shows the distributions of O_{BDT} and m_T in the VBF BDT analysis. Again the data are quite consistent with a VBF Higgs hypothesis.

Because the cut-based result is used as a validation for the BDT analysis and the two signal regions are not fully orthogonal, it is interesting to explore which events overlap between the two analyses. Of the twenty events in the cut-based signal region, only seven were not selected by the BDT analysis, while the other thirteen also enter the BDT signal region. Figure ?? shows where the different analysis candidates lie in the m_{jj} - m_T plane. This shows clearly that the advantage of the BDT analysis is that it can extract

(a) Before the BDT classification

Selection	Summary						Composition of N_{bkg}											
	$N_{\text{obs}}/N_{\text{bkg}}$	N_{obs}	N_{bkg}	N_{signal}	N_{ggF}	N_{VBF}	N_{VH}	N_{WW}	$N_{\text{WW}}^{\text{NEW}}$	$N_{\text{WW}}^{\text{QCD}}$	N_t	$N_{W\bar{W}}$	N_{missid}	N_{ij}	N_{VV}	$N_{e\bar{e}/\mu\mu}$	$N_{\tau\tau}^{\text{QCD}}$	$N_{\tau\tau}^{\text{NEW}}$
$e\mu$ sample	1.04 ± 0.04	718	689	13	15	2.0	90	II	327	42	29	23	31	—	2.2	130	2	
$ee/\mu\mu$ sample	1.18 ± 0.08	469	397	6.0	7.7	0.9	37	3	132	17	5.2	1.2	10.1	168	23	1	—	—

(b) Bins in O_{BDT}

$e\mu$ sample	Composition of N_{bkg}																	
	Bin 0 (not used)	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6	Bin 7	Bin 8	Bin 9	Bin 10	Bin 11	Bin 12	Bin 13	Bin 14	Bin 15	Bin 16	Bin 17
Bin 0 (not used)	1.02 \pm 0.04	661	650	8.8	3.0	1.9	83	9	313	40	26	21	28	—	2.2	126	1	
Bin 1	0.99 \pm 0.16	37	37	3.0	4.2	0.1	5.0	1.0	17	3.1	3.3	1.8	2.6	—	—	4.0	0.2	
Bin 2	2.26 \pm 0.63	14	6.2	1.2	4.2	—	1.5	0.5	1.8	0.3	0.4	0.3	0.8	—	—	0.3	0.3	
Bin 3	5.41 \pm 2.32	6	1.1	0.4	3.1	—	0.3	0.2	0.3	0.1	—	—	0.1	—	—	0.1	0.1	
$ee/\mu\mu$ sample	Bin 0 (not used)	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6	Bin 7	Bin 8	Bin 9	Bin 10	Bin 11	Bin 12	Bin 13	Bin 14	Bin 15	Bin 16	Bin 17
Bin 0 (not used)	1.91 \pm 0.08	396	345	3.8	1.3	0.8	33	2	123	16	4.1	1.1	8.8	137	20.5	0.5		
Bin 1	0.82 \pm 0.14	53	45	1.5	2.2	0.1	3.0	0.5	10.4	1.8	0.8	0.2	0.9	26	1.7	0.1		
Bin 2	1.77 \pm 0.49	14	7.9	0.6	2.5	—	0.8	0.3	1.1	0.2	0.2	—	0.3	4.4	0.3	0.1		
Bin 3	6.52 \pm 2.87	6	0.9	0.2	1.7	—	0.1	0.2	0.2	—	—	—	—	0.7	—	—		

Table 5.15: Event selection for the VBF BDT analysis. The event yields in (a) are shown after the preselection and the additional requirements applied before the BDT classification (see text). The event yields in (b) are given in bins in O_{BDT} after the classification[20].

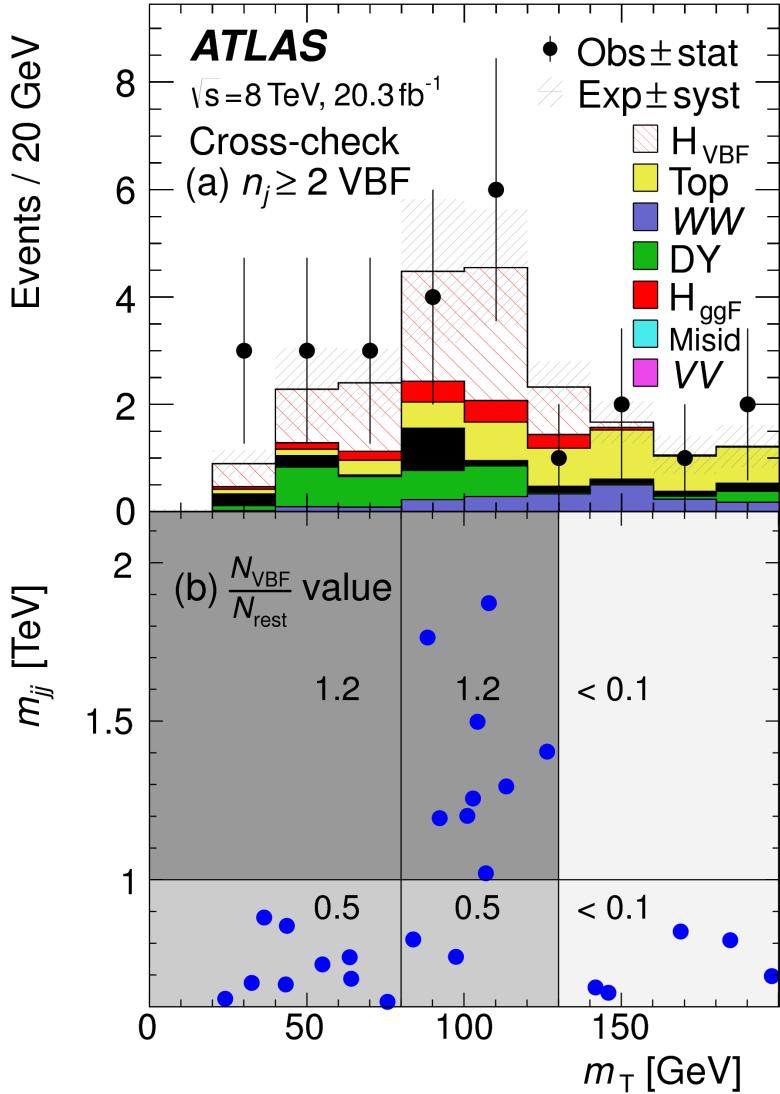


Figure 5.17: Postfit distributions in the cut-based VBF analysis. Panel (a) shows the one-dimensional m_T distribution, while (b) shows the data candidates split into the bins of m_T and m_{jj} used in the final fit[20].

signal candidates lower m_{jj} region due to its ability to recognize correlations with other variables.

While the context of these results in the broader $H \rightarrow WW^*$ statistical analysis will be presented in the next chapter, the significance of the VBF observation can be shown here. In the BDT analysis, the expected signal significance was 2.7σ , while the observed significance was 3.1σ . In the cut-based analysis, the expected significance was 2.1σ and the observed significance was 3.0σ . The compatibility between these two results can be evaluated by computing the probability of observing a larger difference

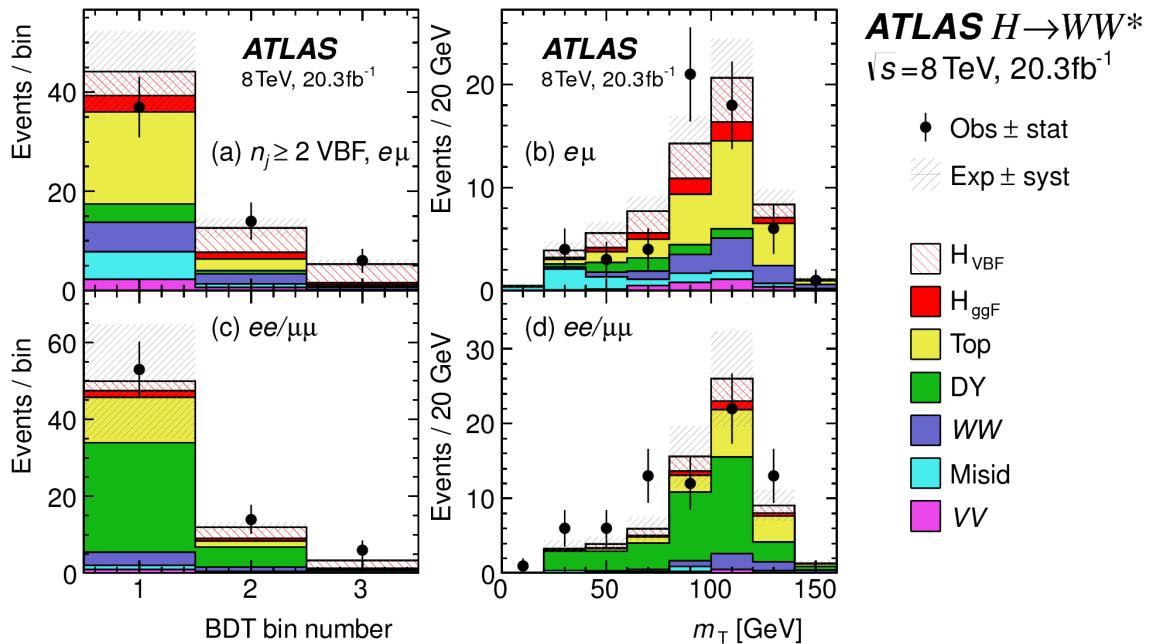


Figure 5.18: Postfit distributions in the BDT VBF analysis[20].

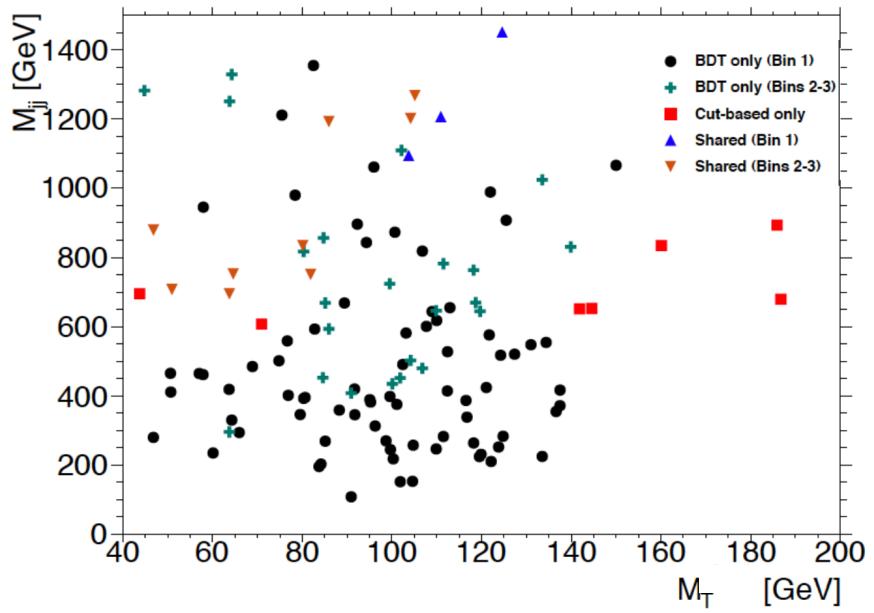


Figure 5.19: Overlap between cut-based and BDT VBF signal region candidates in the m_{jj} - m_T plane.

in Z_0 values than the one measured. Using toy Monte Carlo with the ggF signal strength fixed to unity and considering only statistical uncertainties, this probability is computed to be 79%, indicating good agreement between the analyses.

This result represents the first observation of the vector boson fusion production of a Higgs boson.

*The feeling is less like an ending than just another
starting point.*

Chuck Palahniuk

6

Combined Run I $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ results

6.1 INTRODUCTION

In the final statistical analysis of $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$, the dedicated gluon-gluon fusion and vector boson fusion sensitive signal regions are all combined into a single fit to determine the main parameters of interest, the Higgs signal strength μ and mass m_H . Therefore, while the specific requirements applied for the VBF sensitive analysis are discussed in chapter 5, the final measurement of these parameters can only be discussed in combination with the results of the ggF dedicated analysis. For example, because ggF Higgs production is considered a background in the VBF analysis, the ggF dedicated signal regions can actually constrain the normalization of this background in the VBF dedicated region.

This chapter presents the combined interpretation of results in the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ analysis

SR category i				Fit var.	
n_j , flavor	$\otimes m_{\ell\ell}$	$\otimes p_T^{\ell 2}$	$\otimes \ell_2$		
$n_j = 0$	$e\mu$	$\otimes [10, 30, 55]$	$\otimes [10, 15, 20, \infty]$	$\otimes [e, \mu]$	m_T
	$ee/\mu\mu$	$\otimes [12, 55]$	$\otimes [10, \infty]$		m_T
$n_j = 1$	$e\mu$	$\otimes [10, 30, 55]$	$\otimes [10, 15, 20, \infty]$	$\otimes [e, \mu]$	m_T
	$ee/\mu\mu$	$\otimes [12, 55]$	$\otimes [10, \infty]$		m_T
$n_j \geq 2$ ggF	$e\mu$	$\otimes [10, 55]$	$\otimes [10, \infty]$		m_T
$n_j \geq 2$ VBF	$e\mu$	$\otimes [10, 50]$	$\otimes [10, \infty]$		O_{BDT}
	$ee/\mu\mu$	$\otimes [12, 50]$	$\otimes [10, \infty]$		O_{BDT}

Table 6.1: All signal regions definitions input into final statistical fit[20].

for gluon fusion and vector boson fusion Higgs production. First, the results of the dedicated gluon fusion search are presented. Then, a comparison of the individual production mode signal strengths (μ_{ggF} and μ_{VBF} and a measurement of the combined signal strength (μ) are shown. Subsequently, the measured values of the Higgs couplings to fermions and vector bosons is presented. Finally, the cross section measurement for ggF and VBF production are shown.

6.2 RESULTS OF DEDICATION GLUON FUSION $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ SEARCH

The details of the dedicated gluon fusion $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ search are not discussed in this thesis and instead left to more comprehensive sources[20]. However, a brief summary of the results are essential for describing the results of the full analysis and interpreting the results of the dedicated VBF search in this broader context.

Table 6.1 shows the individual signal regions that were input into the final statistical fit. The ggF dedicated bins use m_T as their discriminating variable and are separated into bins of p_T of the subleading lepton as well. The VBF dedicated bin uses the O_{BDT} distribution as its final discriminant.

Table 6.2 shows the yields in the various signal regions in both data and expected signal and back-

grounds. The yields for signal and background are all scaled according to the final normalizations calculated in the fit.

	N_{obs}	N_{bkg}	N_{ggF}	N_{VBF}
$n_j = 0$	3750	3430 ± 90	300 ± 50	8 ± 4
$n_j = 1$	1596	1470 ± 40	102 ± 26	17 ± 5
$n_j \geq 2, \text{ggF } e\mu$	1017	960 ± 40	37 ± 11	13 ± 1.4
$n_j \geq 2, \text{VBF}$	130	99 ± 9	7.7 ± 2.6	21 ± 3

Table 6.2: Post-fit yields in the different ggF and VBF dedicated signal regions[20].

Figure 6.1 shows the final post-fit m_T distribution in the $n_j \leq 1$ regions. The data are very consistent with the hypothesis of ggF Higgs production.

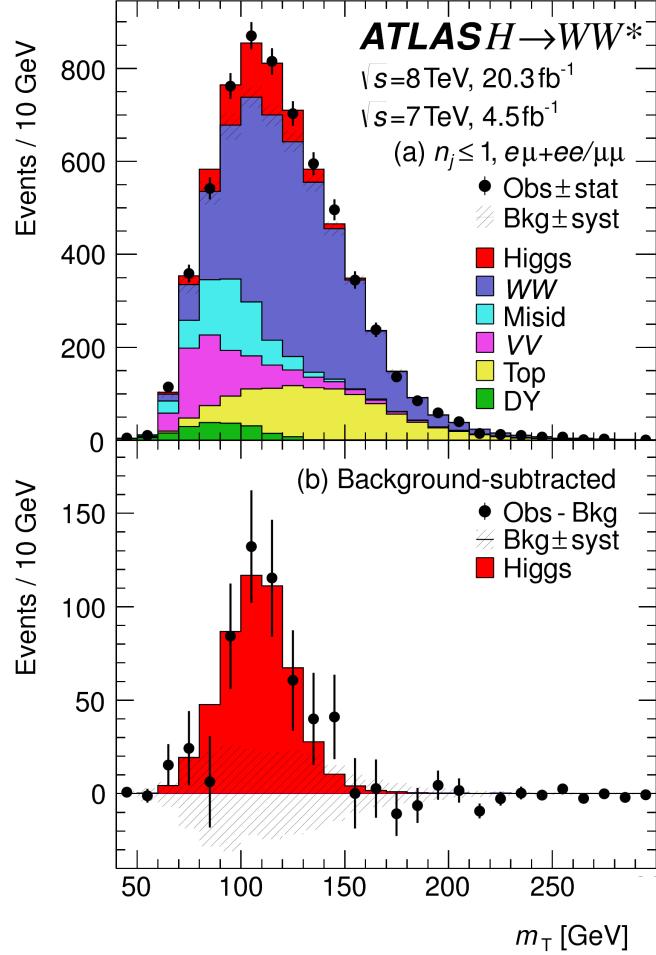


Figure 6.1: Post-fit m_T distribution in the $n_j \leq 1$ regions[20].

These yields are used as input, along with the VBF results in chapter 5, for the physical interpretation of results presented in subsequent sections.

6.3 SIGNAL STRENGTH MEASUREMENTS IN ggF AND VBF PRODUCTION

When all of the signal regions are combined in the fit, there can be a combined measurement of the signal strength as well as the individual ggF and VBF signal strengths. The combined signal strength is the ratio of the sum of the gluon fusion and VBF cross sections to the theory prediction, or a singal strength for the total Higgs production cross section that this analysis is sensitive to. The final measured combined signal strength μ is measured shown in equation 6.1.

$$\begin{aligned}\mu &= 1.09 \quad {}^{+0.16}_{-0.15} (\text{stat.}) \quad {}^{+0.08}_{-0.07} \left(\begin{array}{l} \text{expt} \\ \text{syst} \end{array} \right) \quad {}^{+0.15}_{-0.12} \left(\begin{array}{l} \text{theo} \\ \text{syst} \end{array} \right) \quad \pm 0.03 \left(\begin{array}{l} \text{lumi} \\ \text{syst} \end{array} \right) \\ &= 1.09 \quad {}^{+0.16}_{-0.15} (\text{stat}) \quad {}^{+0.17}_{-0.14} (\text{syst}) \\ &= 1.09 \quad {}^{+0.23}_{-0.21}.\end{aligned}\tag{6.1}$$

Figure 6.2 gives the best fit signal strength $\hat{\mu}$ as a function of hte hypothesized Higgs mass. The value at 125.36 GeV corresponds to the μ quoted in equation 6.1. This value of the Higgs mass is used because it is the most precise mass measurement from ATLAS, a result of the combined $\gamma\gamma$ and ZZ mass measurements[?].

As explained in chapter 3, a probability p_0 can be computed using the test statistic q_0 to quantify the probability that the background could fluctuate to produce an excess at least as large as the one observed in the data. The local p_0 value is shown in figure 6.3 as a function of m_H . The minimum p_0 value is at $m_H = 130$ GeV and coresponds to a significance of 6.1σ . The curve is relatively flat and the significance is the same at 125.36 GeVwithin the quoted precision. The expected significance for a signal with strength $\mu = 1.0$ is 5.8σ . This represents the first discovery level significance measurement in the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ analysis.

All the results presented so far in this section have been for the combined gluon fusion and VBF production modes. However, each signal strength can be calculated separately in the likelihood as well. There are two ways to do this. First, the likelihood can be parameterized in terms of a single parameter,

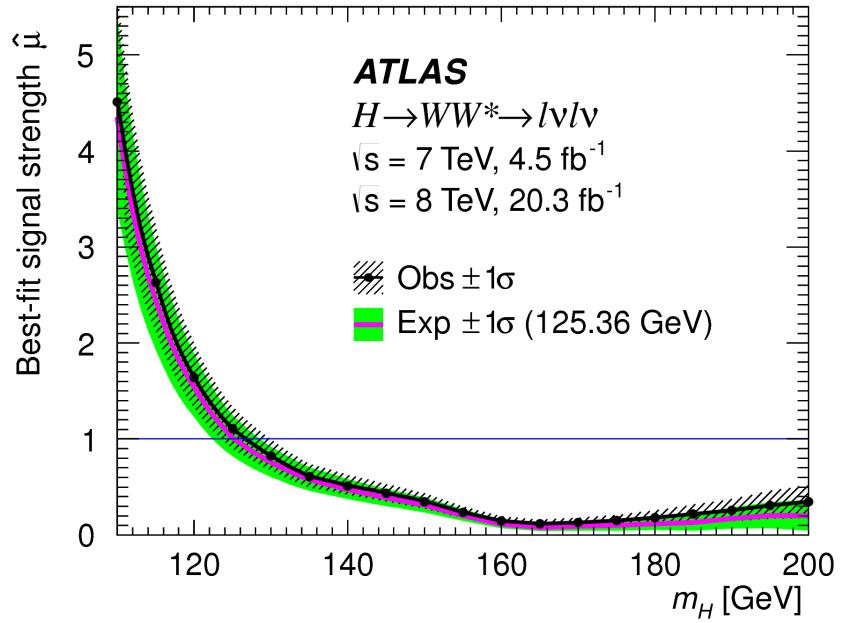


Figure 6.2: Best fit signal strength $\hat{\mu}$ as a function of hypothesized m_H [20].

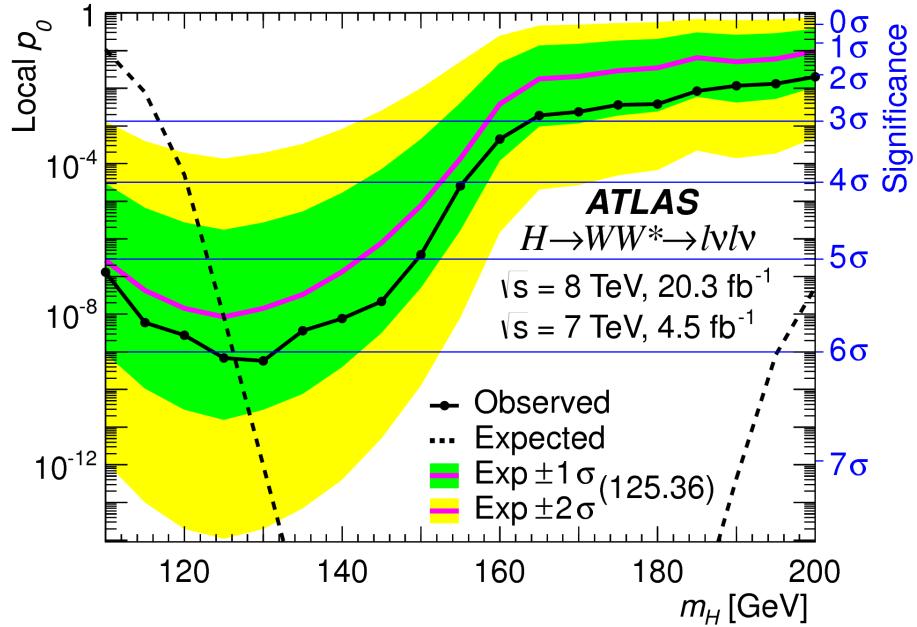


Figure 6.3: Local p_0 as a function of m_H [20].

the ratio of the VBF and gluon fusion signal strengths. With this method, the significance of the VBF observation can be evaluated. Figure 6.4 shows the likelihood as a function of the ratio $\mu_{\text{VBF}}/\mu_{\text{ggF}}$.

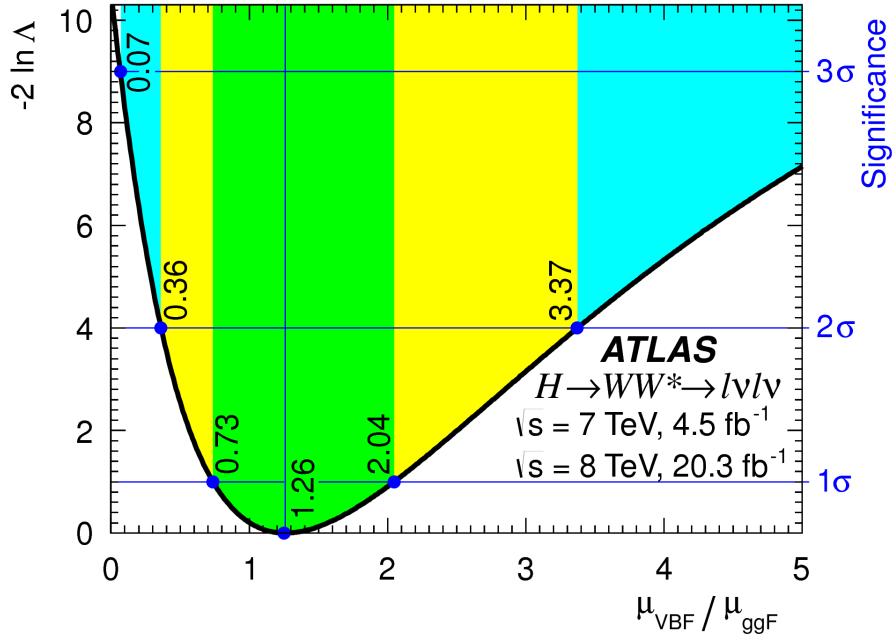


Figure 6.4: Likelihood as a function of $\mu_{\text{VBF}} / \mu_{\text{ggF}}$ [20].

The best fit value of the ratio of signal strengths is shown in equation 6.2. Within the quoted uncertainties, it is consistent with a ratio of unity.

$$\frac{\mu_{\text{VBF}}}{\mu_{\text{ggF}}} = 1.26^{+0.61} (\text{stat.})^{+0.50} (\text{syst.}) = 1.26^{+0.79}_{-0.53} \quad (6.2)$$

The null hypothesis for VBF production corresponds to a ratio of $\mu_{\text{VBF}} / \mu_{\text{ggF}} = 0$. The likelihood in figure 6.4 gives a significance of 3.2σ at $\mu_{\text{VBF}} / \mu_{\text{ggF}} = 0$, as quoted in chapter 5.

In addition to the ratio of signal strengths, each signal strength can be varied independently in the likelihood as well. Figure 6.5 shows the two dimensional likelihood scan in the $\mu_{\text{ggF}}-\mu_{\text{VBF}}$ plane. The best fit values of the two signal strengths are shown in equation 6.3. Both are consistent with unity within their uncertainties.

$$\begin{aligned} \mu_{\text{ggF}} &= 1.02 \pm 0.19^{+0.22}_{-0.18} = 1.02^{+0.29}_{-0.26} \\ \mu_{\text{VBF}} &= 1.27 \pm 0.44^{+0.29}_{-0.40} = 1.27^{+0.53}_{-0.45}. \end{aligned} \quad (6.3)$$

(stat.) (syst.)

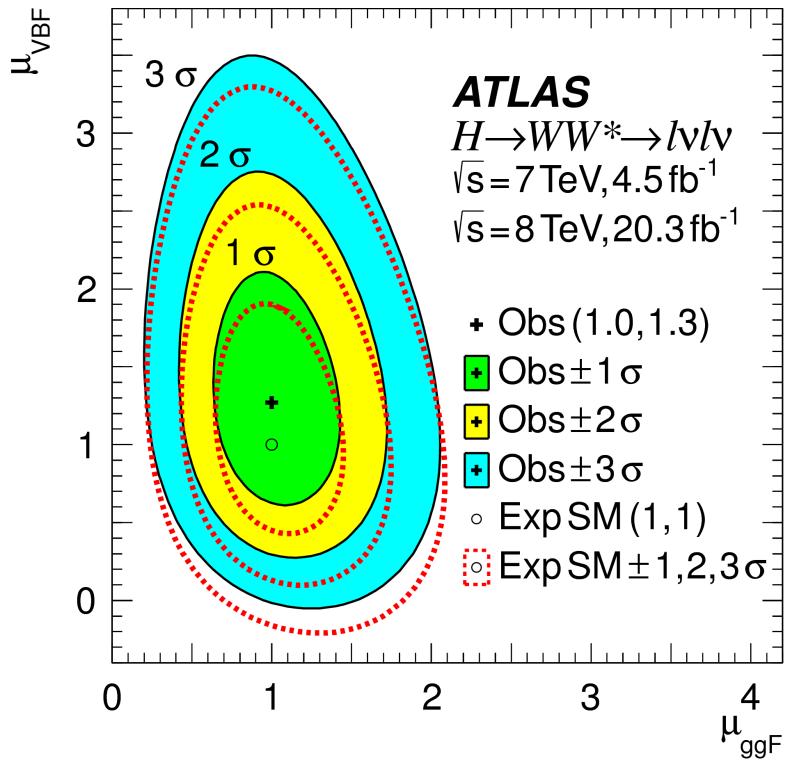


Figure 6.5: Likelihood scan as a function of μ_{VBF} and μ_{ggF} [20].

6.4 MEASUREMENT OF HIGGS COUPLINGS TO VECTOR BOSONS AND FERMIONS

Similar to the parameterization of signal strength, the couplings of the Higgs to fermions and bosons can also be parameterized. The parameter of interest in this case is κ , or the ratio of the measured coupling to the standard model expectation. Both the fermion and boson couplings have these so-called scale factors, κ_F for fermions and κ_V for bosons. Gluon fusion production is sensitive to the fermion couplings through the top quark loops in its production, while VBF production is sensitive to the vector boson couplings in its production. Both modes are sensitive to the vector boson couplings in their decays. The signal strengths will have dependence on the coupling scale factors as described in equation 6.4[2].

$$\begin{aligned} \mu_{\text{ggF}} &\propto \frac{\kappa_F^2 \cdot \kappa_V^2}{(\mathcal{B}_{H \rightarrow f\bar{f}} + \mathcal{B}_{H \rightarrow gg}) \kappa_F^2 + (\mathcal{B}_{H \rightarrow VV}) \kappa_V^2} \\ \mu_{\text{VBF}} &\propto \frac{\kappa_V^4}{(\mathcal{B}_{H \rightarrow f\bar{f}} + \mathcal{B}_{H \rightarrow gg}) \kappa_F^2 + (\mathcal{B}_{H \rightarrow VV}) \kappa_V^2}. \end{aligned} \quad (6.4)$$

Figure 6.6 shows the two-dimensional likelihood scan of κ_F and κ_V . The best-fit values are given in equation 6.5. The best-fit values are consistent with unity within their uncertainties.

$$\begin{aligned} \kappa_F &= 0.93 & +0.24 & +0.21 & = 0.93 & +0.32 \\ && -0.18 & -0.14 && -0.23 \\ \kappa_V &= 1.04 & +0.07 & +0.07 & = 1.04 & \pm 0.11. \end{aligned} \quad (6.5)$$

(stat.) (syst.)

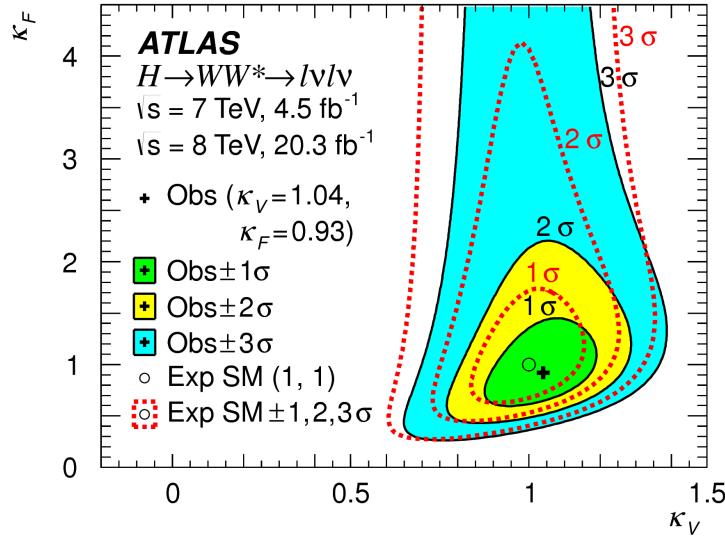


Figure 6.6: Likelihood scan as a function of κ_F and κ_V [20].

6.5 HIGGS PRODUCTION CROSS SECTION MEASUREMENT

Another measurement that comes naturally from the signal strength numbers quoted earlier is the production cross section at 7 and 8 TeV for both gluon fusion and VBF production. The general equation for calculating the cross section is given in equation 6.6.

$$\begin{aligned} (\sigma \cdot \mathcal{B}_{H \rightarrow WW^*})_{\text{obs}} &= \frac{(N_{\text{sig}})_{\text{obs}}}{\mathcal{A} \cdot \mathcal{C} \cdot \mathcal{B}_{WW \rightarrow \ell\nu\ell\nu}} \cdot \frac{1}{\int L dt} \\ &= \hat{\mu} \cdot (\sigma \cdot \mathcal{B}_{H \rightarrow WW^*})_{\text{exp}} \end{aligned} \quad (6.6)$$

$(N_{\text{sig}})_{\text{obs}}$ is the number of events observed in data. \mathcal{A} is the geometric and kinematic acceptance of the detector, while \mathcal{C} is the efficiency of the signal region selection for events that are reconstructed in the detector. The branching ratio of a WW system to leptons must also be divided out. The production cross section depends on the center of mass energy and the production mode desired (gluon fusion or VBF), and so three separate cross section measurements are quoted in equation 6.7.

$$\begin{aligned} \sigma_{\text{ggf}}^{\text{7TeV}} \cdot \mathcal{B}_{H \rightarrow WW^*} &= 2.0 \pm 1.7 {}^{+1.2}_{-1.1} = 2.0 {}^{+2.1}_{-2.0} \text{ pb} \\ \sigma_{\text{ggf}}^{\text{8TeV}} \cdot \mathcal{B}_{H \rightarrow WW^*} &= 4.6 \pm 0.9 {}^{+0.8}_{-0.7} = 4.6 {}^{+1.2}_{-1.1} \text{ pb} \\ \sigma_{\text{VBF}}^{\text{8TeV}} \cdot \mathcal{B}_{H \rightarrow WW^*} &= 0.51 {}^{+0.17}_{-0.15} {}^{+0.13}_{-0.08} = 0.51 {}^{+0.22}_{-0.17} \text{ pb.} \end{aligned} \quad (6.7)$$

(stat.) (syst.)

The predicted cross section values for gluon fusion are 3.3 ± 0.4 pb at 7 TeV and 4.2 ± 0.5 pb at 8 TeV, consistent with the measured values within their uncertainties. For vector boson fusion, the predicted cross section is 0.35 ± 0.02 pb, again consistent with the measured value.

6.6 CONCLUSION

The combined analysis of the gluon fusion and vector boson fusion processes in $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ in the 7 and 8 TeV datasets has yielded the first discovery level significance for Higgs production in this decay channel. Additionally, precise measurements of the couplings to vector bosons and fermions are given. Finally, signal strengths and cross sections for each production mode are measured. Figure 6.7 shows the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ measurements in comparison with other Higgs decay channels in ATLAS. The measurement of signal strength from this channel remains the most sensitive in both the gluon fusion and VBF production modes for the Run 1 dataset.

ATLAS

Individual analysis

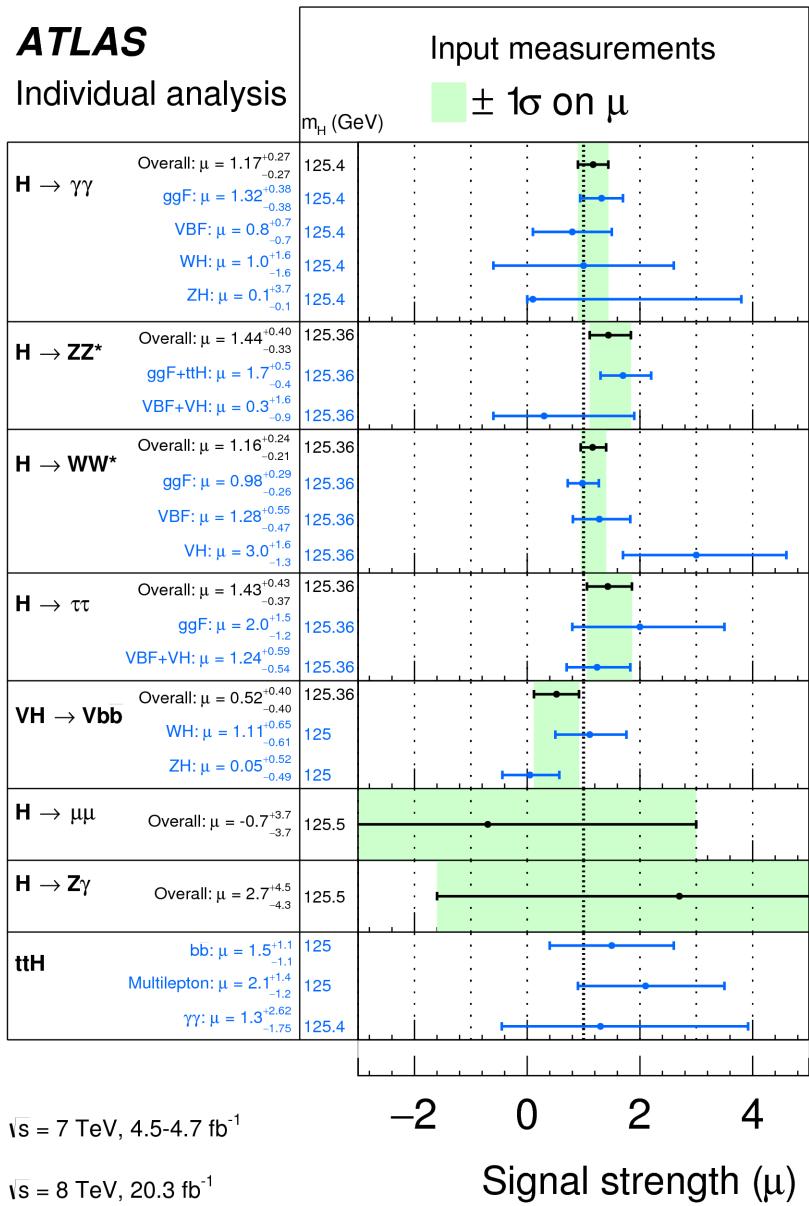


Figure 6.7: Comparison of signal strength measurements in different Higgs decay channels on ATLAS[22].

Part III

Search for Higgs pair production in the
 $HH \rightarrow b\bar{b}b\bar{b}$ channel in LHC Run 2 at $\sqrt{s} =$
13 TeV

7

Search overview

8

Search for Higgs pair production in boosted
final states

9

Results with Run 2 2015 dataset

Part IV

Looking ahead

10

Conclusion

We found the Higgs. Then measured it. Then used it to look for new physics. What a time to be alive!

References

- [1] K. A. Olive et al. Review of Particle Physics. *Chin. Phys.*, C38:090001, 2014. doi: 10.1088/1674-1137/38/9/090001.
- [2] LHC Higgs Cross Section Working Group, S. Heinemeyer, C. Mariotti, G. Passarino, and R. Tanaka (Eds.). Handbook of LHC Higgs Cross Sections: 3, Higgs Properties. 2013.
- [3] Kaustubh Agashe, Hooman Davoudiasl, Gilad Perez, and Amarjit Soni. Warped Gravitons at the LHC and Beyond. *Phys. Rev.*, D76:036006, 2007. doi: 10.1103/PhysRevD.76.036006.
- [4] Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer. MadGraph 5:Going Beyond. *JHEP*, 1106:128, 2011. doi: 10.1007/JHEP06(2011)128.
- [5] Howard E. Haber and Oscar Stål. New LHC benchmarks for the \mathcal{CP} -conserving two-Higgs-doublet model. *Eur. Phys. J.*, C75(10):491, 2015. doi: 10.1140/epjc/s10052-015-3697-x.
- [6] Lyndon Evans. The Large Hadron Collider. In Holstein, BR and Haxton, WC and Jawahery, A, editor, *ANNUAL REVIEW OF NUCLEAR AND PARTICLE SCIENCE, VOL 61*, volume 61 of *Annual Review of Nuclear and Particle Science*, pages 435–466. 2011. doi: {10.1146/annurev-nucl-102010-130438}.
- [7] ATLAS Collaboration. The ATLAS experiment at the CERN Large Hadron Collider. *JINST*, 3: S08003, 2008. doi: 10.1088/1748-0221/3/08/S08003.
- [8] Track Reconstruction Performance of the ATLAS Inner Detector at $\sqrt{s} = 13$ TeV. Technical Report ATL-PHYS-PUB-2015-018, CERN, Geneva, Jul 2015. URL <http://cds.cern.ch/record/2037683>.
- [9] ATLAS Collaboration. ATLAS Trigger Operations Public Results. 2015. URL <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/TriggerOperationPublicResults>.
- [10] ATLAS Collaboration. ATLAS Luminosity Public Results, Run 1. 2012. URL <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults>.
- [11] ATLAS Collaboration. ATLAS Luminosity Public Results, Run 2. 2015. URL <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2>.

- [12] T Kawamoto, S Vlachos, L Pontecorvo, J Dubbert, G Mikenberg, P Iengo, C Dallapiccola, C Amelung, L Levinson, R Richter, and D Lellouch. New Small Wheel Technical Design Report. Technical Report CERN-LHCC-2013-006. ATLAS-TDR-020, CERN, Geneva, Jun 2013. URL <https://cds.cern.ch/record/1552862>. ATLAS New Small Wheel Technical Design Report.
- [13] Joao Pequenao and Paul Schaffner. An computer generated image representing how ATLAS detects particles. Jan 2013. URL <https://cds.cern.ch/record/1505342>.
- [14] Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data. Technical Report ATLAS-CONF-2014-032, CERN, Geneva, Jun 2014. URL <https://cds.cern.ch/record/1706245>.
- [15] Georges Aad et al. Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data. *Eur. Phys. J.*, C74(10):3071, 2014. doi: 10.1140/epjc/s10052-014-3071-4.
- [16] Georges Aad et al. Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton–proton collision data. *Eur. Phys. J.*, C74(11):3130, 2014. doi: 10.1140/epjc/s10052-014-3130-x.
- [17] Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS. Technical Report ATLAS-CONF-2015-037, CERN, Geneva, Aug 2015. URL <http://cds.cern.ch/record/2044941>.
- [18] Georges Aad et al. Performance of b -Jet Identification in the ATLAS Experiment. 2015.
- [19] Expected performance of the ATLAS b -tagging algorithms in Run-2. Technical Report ATL-PHYS-PUB-2015-022, CERN, Geneva, Jul 2015. URL <http://cds.cern.ch/record/2037697>.
- [20] ATLAS Collaboration. Observation and measurement of Higgs boson decays to WW^* with the ATLAS detector. *Phys. Rev. D*, 92(012006), 2015.
- [21] Georges Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett.*, B716:1–29, 2012. doi: 10.1016/j.physletb.2012.08.020.
- [22] Georges Aad et al. Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at $\sqrt{s} = 7$ and 8 TeV in the ATLAS experiment. *Eur. Phys. J.*, C76(1):6, 2016. doi: 10.1140/epjc/s10052-015-3769-y.
- [23] Christopher G. Tully. *Elementary particle physics in a nutshell*. 2011.

[24]

- [25] Mike Lamont for the LHC team. The First Years of LHC Operation for Luminosity Production. International Particle Accelerator Conference, 2013. URL https://accelconf.web.cern.ch/accelconf/IPAC2013/talks/moyab101_talk.pdf.
- [26] Paul Collier for the LHC team. LHC Machine Status. CERN Resource Review Board, 2015. URL <https://cds.cern.ch/record/2063924/files/CERN-RRB-2015-119.PDF>.
- [27] David Griffiths. *Introduction to elementary particles*. 2008.
- [28] F. Halzen and Alan D. Martin. *QUARKS AND LEPTONS: AN INTRODUCTORY COURSE IN MODERN PARTICLE PHYSICS*. 1984. ISBN 0471887412, 9780471887416.
- [29] Matthew D. Schwartz. *Quantum Field Theory and the Standard Model*. Cambridge University Press, 2014. ISBN 1107034736, 9781107034730. URL <http://www.cambridge.org/us/academic/subjects/physics/theoretical-physics-and-mathematical-physics/quantum-field-theory-and-standard-model>.
- [30] S. Dawson. Introduction to electroweak symmetry breaking. In *High energy physics and cosmology. Proceedings, Summer School, Trieste, Italy, June 29-July 17, 1998*, pages 1–83, 1998. URL <http://alice.cern.ch/format/showfull?sysnb=0301862>.
- [31] S. L. Glashow. Partial Symmetries of Weak Interactions. *Nucl. Phys.*, 22:579–588, 1961. doi: [10.1016/0029-5582\(61\)90469-2](https://doi.org/10.1016/0029-5582(61)90469-2).
- [32] Steven Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, 1967. doi: [10.1103/PhysRevLett.19.1264](https://doi.org/10.1103/PhysRevLett.19.1264).
- [33] A. Salam. *Elementary Particle Theory*. Almqvist and Wiksell, Stockholm, 1968.
- [34] J. Iliopoulos S.L. Glashow and L. Maiani. D2:1285, 1970.
- [35] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett.*, B716:30–61, 2012. doi: [10.1016/j.physletb.2012.08.021](https://doi.org/10.1016/j.physletb.2012.08.021).
- [36] R. Keith Ellis, W. James Stirling, and B. R. Webber. QCD and collider physics. *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.*, 8:1–435, 1996.
- [37] P. W. Higgs. Broken symmetries and the masses of gauge bosons. 13:508, 1964.
- [38] P. W. Higgs. Spontaneous symmetry breakdown without massless bosons. 145:1156, 1966.
- [39] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. 13:321, 1964.

- [40] G. S. Guralnik, C. R. Hagen, and T. W. .B. Kibble. Global conservation laws and massless particles. *Phys. Rev. Lett.*, 13:585, 1964. doi: 10.1103/PhysRevLett.13.585.
- [41] Matthew J. Dolan, Christoph Englert, and Michael Spannowsky. New Physics in LHC Higgs boson pair production. *Phys. Rev.*, D87(5):055002, 2013. doi: 10.1103/PhysRevD.87.055002.
- [42] Roberto Contino, Margherita Ghezzi, Mauro Moretti, Giuliano Panico, Fulvio Piccinini, and Andrea Wulzer. Anomalous Couplings in Double Higgs Production. *JHEP*, 08:154, 2012. doi: 10.1007/JHEP08(2012)154.
- [43] R. Grober and M. Muhlleitner. Composite Higgs Boson Pair Production at the LHC. *JHEP*, 06:020, 2011. doi: 10.1007/JHEP06(2011)020.
- [44] Lisa Randall and Raman Sundrum. A Large mass hierarchy from a small extra dimension. *Phys. Rev. Lett.*, 83:3370–3373, 1999. doi: 10.1103/PhysRevLett.83.3370.
- [45] A. Liam Fitzpatrick, Jared Kaplan, Lisa Randall, and Lian-Tao Wang. Searching for the Kaluza-Klein Graviton in Bulk RS Models. *JHEP*, 09:013, 2007. doi: 10.1088/1126-6708/2007/09/013.
- [46] Julien Baglio, Otto Eberhardt, Ulrich Nierste, and Martin Wiebusch. Benchmarks for Higgs Pair Production and Heavy Higgs boson Searches in the Two-Higgs-Doublet Model of Type II. *Phys. Rev.*, D90(1):015008, 2014. doi: 10.1103/PhysRevD.90.015008.
- [47] G. C. Branco, P. M. Ferreira, L. Lavoura, M. N. Rebelo, Marc Sher, and Joao P. Silva. Theory and phenomenology of two-Higgs-doublet models. *Phys. Rept.*, 516:1–102, 2012. doi: 10.1016/j.physrep.2012.02.002.
- [48] Jose M. No and Michael Ramsey-Musolf. Probing the Higgs Portal at the LHC Through Resonant di-Higgs Production. *Phys. Rev.*, D89(9):095031, 2014. doi: 10.1103/PhysRevD.89.095031.
- [49] Lyndon R Evans and Philip Bryant. LHC Machine. *J. Instrum.*, 3:S08001. 164 p, 2008. URL <https://cds.cern.ch/record/1129806>. This report is an abridged version of the LHC Design Report (CERN-2004-003).
- [50] CMS Collaboration. The cms experiment at the cern lhc. *Journal of Instrumentation*, 3(08):S08004, 2008. URL <http://stacks.iop.org/1748-0221/3/i=08/a=S08004>.
- [51] LHCb Collaoration. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008. doi: 10.1088/1748-0221/3/08/S08005.
- [52] ALICE Collaboration. The alice experiment at the cern lhc. *Journal of Instrumentation*, 3(08):S08002, 2008. URL <http://stacks.iop.org/1748-0221/3/i=08/a=S08002>.

- [53] ATLAS Collaboration. Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV Using the ATLAS Detector at the LHC. *Eur. Phys. J.*, C 71:1630, 2011. doi: 10.1140/epjc/s10052-011-1630-5.
- [54] M Capeans, G Darbo, K Einsweiller, M Elsing, T Flick, M Garcia-Sciveres, C Gemme, H Pernegger, O Rohne, and R Vuillermet. ATLAS Insertable B-Layer Technical Design Report. Technical Report CERN-LHCC-2010-013, ATLAS-TDR-19, CERN, Geneva, Sep 2010. URL <https://cds.cern.ch/record/1291633>.
- [55] Y Giomataris, Ph Rebougeard, J.P. Robert, and G Charpak. Micromegas: a high-granularity position-sensitive gaseous detector for high particle-flux environments. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 376(1):29 – 35, 1996. ISSN 0168-9002. doi: [http://dx.doi.org/10.1016/0168-9002\(96\)00175-1](http://dx.doi.org/10.1016/0168-9002(96)00175-1). URL <http://www.sciencedirect.com/science/article/pii/0168900296001751>.
- [56] T Alexopoulos, J Burnens, R de Oliveira, G Glonti, O Pizzirusso, V Polychronakos, G Sekhniaidze, G Tsipolitis, and J Wotschack. A spark-resistant bulk-micromegas chamber for high-rate applications. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 640(1):110 – 118, 2011. ISSN 0168-9002. doi: <http://dx.doi.org/10.1016/j.nima.2011.03.025>. URL <http://www.sciencedirect.com/science/article/pii/S0168900211005869>.
- [57] Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung. Technical Report ATLAS-CONF-2012-047, CERN, Geneva, May 2012. URL <https://cds.cern.ch/record/1449796>.
- [58] W Lampl, S Laplace, D Lelas, P Loch, H Ma, S Menke, S Rajagopalan, D Rousseau, S Snyder, and G Unal. Calorimeter Clustering Algorithms: Description and Performance. Technical Report ATL-LARG-PUB-2008-002, ATL-COM-LARG-2008-003, CERN, Geneva, Apr 2008. URL <https://cds.cern.ch/record/1099735>.
- [59] Georges Aad et al. Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. 2016.
- [60] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Anti-k(t) jet clustering algorithm. *JHEP*, 04:063, 2008. doi: 10.1088/1126-6708/2008/04/063.
- [61] Aaron James Armbruster. Discovery of a Higgs Boson with the ATLAS detector. 2013. CERN-THESIS-2013-047.
- [62] G Cowan, K Cranmer, E Gross, and O Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J.*, C 71:1554, 2011. doi: 10.1140/epjc/s10052-011-1554-0.

- [63] ATLAS Collaboration. Performance of the ATLAS muon trigger in pp collisions at $\sqrt{s} = 8$ TeV. *Eur. Phys. J. C*, (arXiv:1408.3179. CERN-PH-EP-2014-154):75, 19 p, Aug 2014. URL <https://cds.cern.ch/record/1749694>.
- [64] ATLAS collaboration. Electron trigger performance in 2012 ATLAS data, 2015. ATLAS-COM-DAQ-2015-091.
- [65] Paolo Nason. A new method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004.
- [66] B. P. Kersevan and E. Richter-Was. The Monte Carlo event generator AcerMC version 2.0 with interfaces to PYTHIA 6.2 and HERWIG 6.5. 2004.
- [67] Nikolas Kauer and Giampiero Passarino. Inadequacy of zero-width approximation for a light Higgs boson signal. 2012.
- [68] T. Gleisberg, Stefan Hoeche, F. Krauss, M. Schonherr, S. Schumann, et al. Event generation with SHERPA 1.1. *JHEP*, 0902:007, 2009. doi: [10.1088/1126-6708/2009/02/007](https://doi.org/10.1088/1126-6708/2009/02/007).
- [69] Michelangelo L. Mangano et al. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 0307:001, 2003. doi: [10.1088/1126-6708/2003/07/001](https://doi.org/10.1088/1126-6708/2003/07/001).
- [70] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 0605:026, 2006. doi: [10.1088/1126-6708/2006/05/026](https://doi.org/10.1088/1126-6708/2006/05/026).
- [71] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. A Brief Introduction to PYTHIA 8.1. *Comput.Phys.Commun.*, 178:852–867, 2008. doi: [10.1016/j.cpc.2008.01.036](https://doi.org/10.1016/j.cpc.2008.01.036).
- [72] G. Corcella et al. HERWIG 6: An event generator for hadron emission reactions with interfering gluons (including super-symmetric processes). *JHEP*, 01:010, 2001. doi: [10.1088/1126-6708/2001/01/010](https://doi.org/10.1088/1126-6708/2001/01/010).
- [73] J. M. Butterworth, Jeffrey R. Forshaw, and M. H. Seymour. Multiparton interactions in photo-production at HERA. *Z. Phys., C* 72:637, 1996. doi: [10.1007/s002880050286](https://doi.org/10.1007/s002880050286).
- [74] Jun Gao, Marco Guzzi, Joey Huston, Hung-Liang Lai, Zhao Li, et al. The CT10 NNLO Global Analysis of QCD. *Phys.Rev., D* 89:033009, 2014. doi: [10.1103/PhysRevD.89.033009](https://doi.org/10.1103/PhysRevD.89.033009).
- [75] P. M. Nadolsky. Implications of CTEQ global analysis for collider observables. *Phys. Rev., D* 78: 013004, 2008. doi: [10.1103/PhysRevD.78.013004](https://doi.org/10.1103/PhysRevD.78.013004).
- [76] A. Sherstnev and R. S. Thorne. Parton distributions for the LHC. *Eur. Phys. J., C* 55:553, 2009. doi: [10.1140/epjc/s10052-008-0610-x](https://doi.org/10.1140/epjc/s10052-008-0610-x).

- [77] S. Agostinelli et al. GEANT4, a simulation toolkit. *Nucl. Instrum. Meth.*, A 506:250, 2003. doi: 10.1016/S0168-9002(03)01368-8.
- [78] I. Stewart and F. Tackmann. Theory uncertainties for Higgs mass and other searches using jet bins. *Phys. Rev.*, D 85:034011, 2012. doi: 10.1103/PhysRevD.85.034011.
- [79] ATLAS Collaboration. Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV Using the ATLAS Detector at the LHC. *Eur. Phys. J.*, C 71:1630, 2011. doi: 10.1140/epjc/s10052-011-1630-5.
- [80] Jet energy scale and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ tev with atlas 2011 data. *ATLAS-CONF-2013-004*, 2013.
- [81] Calibrating the b -tag efficiency and mistag rate in 35 pb^{-1} of data with the atlas detector. *ATLAS-CONF-2011-089*, 2011.
- [82] ATLAS Collaboration. Measurement of the b -tag Efficiency in a Sample of Jets Containing Muons with 5 fb^{-1} of Data from the ATLAS Detector. *ATLAS-CONF-2012-043*, 2012. URL <http://cdsweb.cern.ch/record/1435197>.



THIS THESIS WAS TYPESET using L^AT_EX, originally developed by Leslie Lamport and based on Donald Knuth's T_EX.

The body text is set in 11 point Egenolff-Berner Garamond, a revival of Claude Garamont's humanist typeface. The above illustration, *Science Experiment 02*, was created by Ben Schlitter and released under CC BY-NC-ND 3.0. A template that can be used to format a PhD dissertation with this look & feel has been released under the permissive AGPL license, and can be found online at github.com/asm-products/Dissertate or from its lead author, Jordan Suchow, at suchow@post.harvard.edu.