Search for the Standard Model Higgs boson produced with an associated W boson and decaying to a pair of b quarks at $\sqrt{s} = 13$ TeV with the ATLAS detector

A Dissertation presented

by

Alyssa Rae Montalbano

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Physics

Stony Brook University

December 2017

CERN-THESIS-2017-459
29/09/2017

Copyright by
Alyssa Rae Montalbano
2017

Stony Brook University

The Graduate School

Alyssa Rae Montalbano

We, the dissertation committe for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation

John Hobbs - Dissertation Advisor
Professor of Physics and Astronomy

Rouven Essig - Chairperson of Defense
Associate Professor of Physics and Astronomy

Joanna Kiryluk - Internal Member
Assistant Professor of Physics and Astronomy

Mary Bishai - External Member
Physicist, Department of Physics, Brookhaven National Laboratory

This dissertation is accepted by the Graduate School

Charles Taber
Dean of the Graduate School

Abstract of the Dissertation

Search for the Standard Model Higgs boson produced with an associated W boson and decaying to a pair of $b\bar{b}$ quarks at $\sqrt{s} = 13$ TeV with the ATLAS detector

by

Alyssa Rae Montalbano

Doctor of Philosophy

in

Physics

Stony Brook University

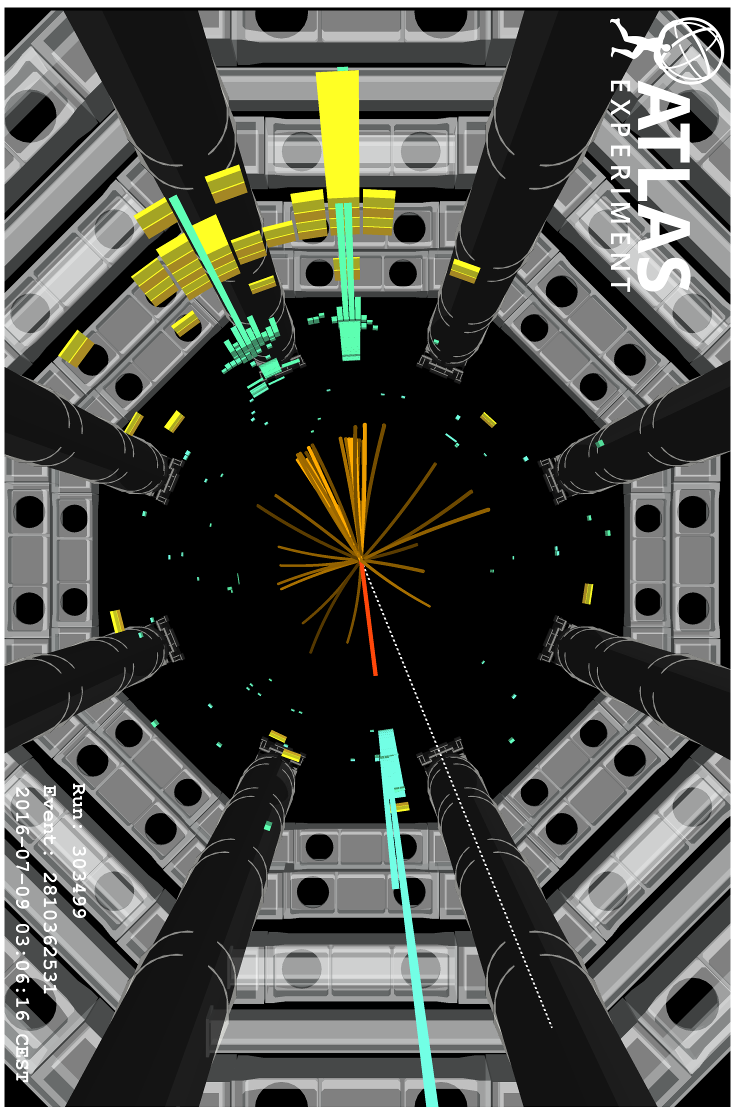
2017

A Higgs boson has been found in multiple decay modes, but not yet in the mode with the largest standard model (SM) branching fraction: $h \to b\bar{b}$. This thesis details a search for the SM Higgs when produced with an associated W boson. This search is performed with the ATLAS detector using data from Run 2 of the Large Hadron Collider operating at a center-of-mass energy of 13 TeV and corresponding to an integrated luminosity of 36.1 fb^{-1}. The final state of the W boson is its leptonic decay with an electron or muon. The measured signal strength relative to the SM Higgs prediction is $\mu = 1.35^{+0.40}_{-0.39}(\text{stat.})^{+0.56}_{-0.47}(\text{syst.})$ and the probability that observed data arises from a background fluctuation only gives a gaussian equivalent of 2.30 standard deviations compared to the expected value of 1.80 standard deviations. Combined with additional leptonic modes of associated vector boson production with a Z boson, the significance of $H \to b\bar{b}$ is observed at 3.5 standard deviations compared to an expected 3.0 standard deviations with a signal strength of $\mu = 1.20^{+0.24}_{-0.23}(\text{stat.})^{+0.34}_{-0.28}(\text{syst.})$ The combined search demonstrates evidence for both the Higgs boson's decay to b-quarks and for its production in association with a vector boson.

Dedication Page

I would like to dedicate this thesis to my friends and family who have supported me through my years in graduate school. Thank you, Dad, Mom, Mike and Grandma for always encouraging me to go for my dreams. And thank you to my friends who have supported me through both the ups and downs in undergrad and grad school, both in the states and abroad.

Frontispiece



Visualization of a WH candidate event with an electron, $E_T^{miss}$and two b-tagged jets taken in 2016 at the ATLAS experiment.

Table of Contents

Contents

List of Figures

List of Tables

List of Abbreviations

ALICE - Experiment at the Large Hadron Collider (A Large Ion Collider Experiment)
ATLAS - Experiment at the Large Hadron Collider (A Toroidal LHC ApparatuS)
BCM - Beam Condition Monitor
BDT - Boosted Decision Tree
BNL - Brookhaven National Laboratory
BR - Branching Ratio
BSM - Beyond the Standard Model
CB - Combined Muon
CKM - CabibboKobayashiMaskawa
CMLE - Conditional Maximum Likelihood Estimate
CMS - Experiment at the Large Hadron Collider (Compact Muon Solenoid)
CP - Charge Parity
CR - Control Region
CSC - Cathode Strip Chambers
CT - Calorimeter-tagged muons
DS - Diagram Subtraction
EM - Electromagnetic
EW - Electroweak
FF - Fake Factor
GUT - Grand Unified Theory
HL-LHC - High Luminosity LHC
HLT - High Level Trigger
HVT - Heavy Vector Triplet
IBL - Insertable B-Layer
ID - Inner Detector
JVT - Jet Vertex Tagger
KS - KolmogorovSmirnov
LAr - Liquid Argon
LEP - Large Electron-Positron Collider
LHC - Large Hadron Collider
LH - Likelihood

LHCb - Experiment at the LHCb (LHC beauty)
MC - Monte Carlo
MDT - Monitored Drift Tubes
ME - Extrapolated Muons
MET - Missing Transverse Energy
MS - Muon Spectrometer
MVA - Multivariate Analysis
NLO - Next to Leading Order
NNLO - Next to Next to Leading Order
NP - Nuisance Parameter
PD - Pixel Detector
PDF - Parton Distribution Functions
PDF - Probability Density Functions
PF - Particle Flow
PS - Proton Synchrotron
PS/UE - Parton Shower / Underlying Event
PV - Primary Vertex
QED - Quantum electrodynamics
QCD - Quantum chromodynamics
radUp - High Initial State Radiation
radDo - Low Initial State Radiation
RF - Radiofrequency
ROI - Region Of Interest
RPC - Resistive Plate Chambers
SCT - Semiconductor Tracker
SM - Standard Model
SPS - Super Proton Synchrotron
SR - Signal Region
ST - Segment-tagged Muons
TGC - Thin Gap Chambers
TRT - Transition Radiation Tracker
vdM - Van Der Meer
VEV - Vacuum Expectation Value

Acknowledgements

I would like to thank my advisor, Professor John Hobbs, for all of the support and advice that he has provided me during my doctoral studies. Professor Giacinto Piacquadio, thank you for our illuminating physics discussions. Thank you to Dr. Valerio Dao, Dr. Lily Morvaj, Dr. Garabed Halladjian and Dr. Chiara Debenedetti for their guidance in physics analysis and navigating the world of particle physics. Finally, a shout-out to my good friends who helped me survive a physics phd at Stony Brook and CERN. Thank you Sarah, Mike, Andy, Lydia, Evan, Lucie, Nic, Xiaoyue, Hannah and Zoya.

1 Theoretical Background

Elementary particle physics is the study of the most basic constituents of matter and how they interact with each other. Particle physicists seek to answer one of humanity's most fundamental questions: What are the building blocks of the world? Philosophers in ancient Greece provide the first recorded instance of an answer to this question. They developed a concept of "atomos" from which the word atom is derived. In Greek it represents that which is uncuttable - something which can no longer be broken into smaller pieces. Early 20th century scientists discovered protons, neutrons and electrons, and inferred the atom. They thought the atom was the most basic constituent of the world and that it composed elements, molecules and all matter. However, in the second half of the 20th century scientists discovered sub-atomic particles which are currently believed to be fundamental particles. With discoveries, scientists created a theory of elementary particle interactions via forces named the Standard Model (SM).

The constituents of matter and their properties are discussed in this section. Forces, the current understanding of how the particles interact with each other, are also discussed.

1.1 Standard Model

The SM is the theory that explains particle physicists' current understanding of the universe on a quantum scale. It illuminates the constituents and interactions of the physical world that can be experienced with human senses, as well as via experiments with the exception of gravity. The SM Lagrangian can most be succinctly stated as

$$\begin{aligned} \mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\ +i\bar{\psi}\not{D}\psi \\ +\psi_i y_{ij}\psi_j\phi + h.c. \\ +|D_\mu\phi|^2 - V(\phi). \end{aligned} \tag{1}$$

The above equation has $\mathrm{SU}(3)_c\times\mathrm{SU}(2)_L\times\mathrm{U}(1)_Y$ symmetry and is composed of multiple terms. The first is made of the electromagnetic field tensor $F_{\mu\nu}$, which gives a gauge-invariant kinetic term for the A_μ field in quantum electrodynamics. It also represents the field strength tensor for the gluon fields of quantum chromodynamics. Next are spinors (ψ) that represent fermions and require use of the gauge covariant derivative D. Feynman slashed notation ($\not{D}$) is used here to indicate a contraction with a gamma matrix. The hermitian conjugate ($h.c.$) must be included to account for antiparticles. The Yukawa coupling (y_{ij}) gives mass to the chiral fermions. The covariant derivative is also used to construct a kinetic term for complex scalar fields, ϕ, while the Higgs field potential is subtracted ($V(\phi)$).

First the elementary particles - leptons, quarks and bosons - and then the interactions the SM incorporates - quantum electrodynamics, quantum chromodynamics - as well as electroweak symmetry breaking will be introduced. Successes and limitations of the SM will be reviewed. The search for $H \to b\bar{b}$ with an associated vector boson will be motivated next. Finally the need for new physics beyond the SM (BSM) is discussed with specific attention paid to the search for heavy vector triplets which is very similar in method to the SM search described here.

1.1.1 Leptons

The first elementary particle discovered was the electron found by J. J. Thomson in 1897 [1]. It was the first charged lepton found, and it was followed by the discovery of its heavier cousin the muon (μ) observed in 1936 by C. Anderson and S. Neddermeyer. The last charged lepton, the tau (τ), was discovered by M. Perl et al in 1975 [2, 3]. The electron exists in most common-day matter, whether bound in atoms or free in electricity. The muon and tau are contained in the second and third generations of matter respectively. The first generation makes up everyday matter, while the second and third generation are distinguished in order of increasing mass. The basic properties of the charged leptons can be found in Table 1.

Every particle in the three generation has an associated antiparticle. The electron's antimatter partner is the positron. It is identical to the electron in mass, spin, the forces it couples to and their coupling strengths, but opposite in charge, lepton number and the corresponding flavor number (i.e. electron number.)

There is another partner of the electron: the electrically neutral electron neutrino (ν_e). Similarly the μ has a muon neutrino (ν_μ) partner and the τ also has a corresponding tau neutrino (ν_τ). The neutral leptons' properties are summarized in Table 1.

Leptons are represented in the SM Lagrangian as left-handed point-like

Property	e	ν_e	μ	ν_μ	τ	ν_τ
Charge (e)	-1	0	-1	0	-1	0
Mass (MeV)	0.51	<0.0000022	105.66	<0.17	1776.84	<15.5
Spin	1/2	1/2	1/2	1/2	1/2	1/2
Lifetime	stable	?	2.20×10^{-6}	?	2.91×10^{-13}	?
Colored	no	no	no	no	no	no
Hypercharge	-1(-2)	-1(-)	-1(-2)	-1(-)	-1(-2)	-1(-)
Weak Isospin T_3	-1/2 (0)	1/2(-)	-1/2(0)	1/2(-)	-1/2(0)	1/2(-)
Force	E,W,G	W,G	E,W,G	W,G	E,W,G	W,G
Lepton number	-1	+1	-1	+1	-1	+1
Electron number	+1	+1	0	0	0	0
Muon number	0	0	+1	+1	0	0
Tau number	0	0	0	0	+1	+1
Baryon number	0	0	0	0	0	0

Table 1: Properties of leptons including charge, mass, spin, whether it has color charge and what forces it interacts with. Additionally lepton number and flavor number used for conservation considerations are listed, with the understanding that the respective anti-particles exhibit the opposite of these conservation numbers as well as the opposite charge. Antiparticles are denoted with a bar on top of the particle's symbol. Certain properties are unknown. These are shown as a question mark, and others have limits placed upon them. Mass values are rounded for brevity's sake. The hypercharge and weak isospin component have both left(right)-handed components. The four forces are abbreviated as follows: electromagnetism (E), weak (W), strong (S), gravitational (G) [4].

half integer spinors. These weak-isospin doublets can be seen below.

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L$$

Example forms of the states are shown below for the first generation (L_e),

$$\nu_L \equiv \frac{1}{2}(1-\gamma_5)\nu$$

$$e_L \equiv \frac{1}{2}(1-\gamma_5)e.$$

The SM defines the right-handed singlets of electrically charged leptons, of the form

$$e_R \equiv \frac{1}{2}(1+\gamma_5)e$$

but not those of right-handed singlets of the electrically neutral neutrino. This means the SM only has massless left-handed neutrinos. Experiments [5–7] have shown, however, that neutrinos oscillate between different flavors and mass eigenstates, requiring that neutrinos have nonzero mass. A theoretical mechanism that gives mass to neutrinos still needs to be incorporated into the SM.

The electron neutrino was originally posited to explain the beta decay spectrum. The spontaneous decay of a neutron results in a proton and an electron (called a β particle at the time). A question arose when studying Bismuth-210's beta decay spectrum: the electron's kinematic energy spectrum was insufficient to conserve energy so where did it go? W. Pauli proposed an undetected particle that carried away the missing energy. It was the neutrino, a particle with mass negligible compared to that of the electron [8]. Confirmation of this novel particle came in 1956 from C. Cowan and F. Reines [9,10]. E. Fermi built upon Pauli's proposition of the neutrino to develop the first iteration of a theory of the weak interaction [11].

1.1.2 Quarks

The second category of elementary particles are quarks. Quarks make up mesons and baryons, composed of a quark-antiquark pair and three quarks or antiquarks respectively. Everyday matter is made of protons and neutrons which in turn are composed of the first generation of quarks, up (u) and

Property	up (u)	down (d)	charm(c)	strange(s)	top(t)	bottom(b)
Charge (e)	+2/3	-1/3	+2/3	-1/3	+2/3	-1/3
Mass (MeV)	2.3	4.8	1275	95	173210	4180
Spin	1/2	1/2	1/2	1/2	1/2	1/2
Colored	yes	yes	yes	yes	yes	yes
Hypercharge	1/3 (4/3)	1/3 (-2/3)	1/3 (4/3)	1/3 (-2/3)	1/3 (4/3)	1/3 (-2/3)
Weak Isospin (T_3)	1/2(0)	-1/2(0)	1/2(0)	-1/2(0)	1/2(0)	-1/2(0)
Force	E,S,W,G	E,S,W,G	E,S,W,G	E,S,W,G	E,S,W,G	E,S,W,G
Baryon number	1/3	1/3	1/3	1/3	1/3	1/3
Charmness	0	0	+1	0	0	0
Strangeness	0	0	0	-1	0	0
Topness	0	0	0	0	+1	0
Bottomness	0	0	0	0	0	-1
Lepton number	0	0	0	0	0	0

Table 2: Properties of quarks including charge, mass, spin, where it has color charge and what forces it interacts with. Additionally their baryon number is listed, with the understanding that the respective anti-particles exhibit the opposite of these as well as the opposite charge. Mass values are rounded for brevity's sake. Direct search bounds on masses are listed. The hypercharge and weak isospin have both left(right)-handed components. The four forces are abbreviated as follows: electromagnetism (E), weak (W), strong (S), gravitational (G) [4].

down (d). The additional generations of quarks have more mass and were discovered later. The second generation of quarks consists of strange (s) and charm (c). The third and final generation of quarks consists of bottom (b) and top (t). The properties of quarks can be found in Table 2.

Like leptons, quarks are also represented in the SM Lagrangian as left-handed point-like half-integer spinors:

$$\begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} c \\ s \end{pmatrix}_L, \begin{pmatrix} t \\ b \end{pmatrix}_L.$$

Quarks carry additional quantum information called color. Color can be expressed as three flavored doublets per generation. The first generation is shown below as an example.

$$\begin{pmatrix} u_{Red} \\ d_{Red} \end{pmatrix}_L \begin{pmatrix} u_{Green} \\ d_{Green} \end{pmatrix}_L \begin{pmatrix} u_{Blue} \\ d_{Blue} \end{pmatrix}_L$$

The right-handed equivalents are two weak isospin singlets,

$$R_u = u_R = \frac{1}{2}(1+\gamma_5)u$$
$$R_d = d_R = \frac{1}{2}(1+\gamma_5)d,$$

of the SU(3)$_c$ symmetry. They are governed by quantum chromodynamics, which does not have spontaneous symmetry breaking meaning the fermions do not require any additional mechanism to obtain mass.

To match experimental results, the Cabibbo current must be taken into account. The Cabibbo current adds a multiplicative factor for flavor-changing charged currents to describe quark mixing between strong (mass) and weak eigenstates. They are also used to interpret the probability that a quark will decay into another flavored quark via the weak interaction. These factors are formulated by the Cabibbo-Kobayashi-Maskawa (CKM) matrix using the Cabibbo angle θ_c, phrased in two formats in Equations 2 and 3 [12].

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \quad (2)$$

The CKM matrix in Equation 2 uses abbreviations for the three mixing angles: $s_{ij} = \sin\theta_{ij}, c_{ij} = \cos\theta_{ij}$. It also denotes the charge parity (CP) violating phase δ. One can see the couplings the CKM matrix describes explicitly in Equation 3 where the primed quark states are the weak eigenstates.

$$\begin{pmatrix} d' \\ s' \\ c' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ c \end{pmatrix} \equiv V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad (3)$$

The coupling constants between quarks in the CKM matrix experimentally obtained are [4]

$$|V| = \begin{bmatrix} 0.9743 & 0.2253 & 0.0035 \\ 0.2252 & 0.9734 & 0.0410 \\ 0.0086 & 0.0403 & 0.9992 \end{bmatrix}.$$

It is clear that the dominant quark couplings are on the diagonal: $u-d$, $c-s$ and $t-b$. Since the CKM matrix is unitary, the SM can be tested by examining the magnitude of the CKM matrix. While experiments determine the CKM matrix, the SM theory does not motivate any of the values nor does it exclude a fourth generation of quarks.

1.1.3 Bosons

Force can be understood as the interaction of a force carrier with an elementary particle. The most common graphical depiction of an interactions is a Feynman diagram. Introduced by R. Feynman, these diagrams sketch out particle interactions in time and space. They are direct representations of calculations when following accompanying Feynman rules (dependent on the forces in the diagram). The resulting calculation produces a transition matrix element can then be translated into a probability amplitude, decay rate and cross section [13].

Property	Photon (γ)	Gluon (g)	W^+/W^- (W)	Z^0 (Z)	Higgs (H)
Charge (e)	0	0	±1	0	0
Mass (GeV)	0	<1.3	80	91	125
Spin	1	1	1	1	0
Hypercharge	0	-	0	0	1
Weak Isospin	0	-	±1	0	-1/2
Scalar/Vector	Vector	Vector	Vector	Vector	Scalar
Field	Electroweak	Strong	Electroweak	Electroweak	Higgs
Baryon number	0	0	0	0	0
Lepton number	0	0	0	0	0

Table 3: Properties of force carriers and the Higgs boson including charge, mass, spin, what forces it carries. Additionally their lepton and baryon numbers are listed. Experimental mass values are rounded for brevity's sake [4].

The manifestations of four gauge fields: $W_\mu^1, W_\mu^2, W_\mu^3$ and A_μ are the gauge bosons. Properties of these four gauge bosons, also known as force carriers, are summarized in Table 3 with the addition of the Higgs boson. The mass eigenstates of the gauge bosons can be realized from the gauge fields. The mass eigenstates of the charged intermediate vector bosons, $W^\pm$ are

$$W_\mu^\pm = \frac{W_\mu^1 \mp iW_\mu^2}{\sqrt{2}}. \tag{4}$$

The neutral intermediate vector boson's, Z^0, mass eigenstate is given by

$$Z_\mu^0 = \frac{-g'\mathcal{A}_\mu + gW_\mu^3}{\sqrt{g^2 + g'^2}}, \tag{5}$$

while the physical photon's (γ) massless eigenstate is

$$A_\mu = \frac{g\mathcal{A}_\mu + g'W^3_\mu}{\sqrt{g^2+g'^2}} \tag{6}$$

with the requirement that the coupling constants' relation defined in Equation 7 holds true at tree level. The coupling constants satisfy the $\mathrm{SU}(2)_L$ and $\mathrm{U}(1)_Y$ groups, unifying electromagnetic and weak interactions on the electroweak symmetry breaking scale,

$$e = \frac{gg'}{\sqrt{g^2+g'^2}}. \tag{7}$$

The massless gauge fields of weak isospin ($W^{1,2,3}$) and hypercharge ($\mathcal{A}$) are rotated by the weak mixing angle, also known as the Weinberg angle, into the vector boson eigenstates:

$$\begin{pmatrix} Z^0_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} W^3_\mu \\ \mathcal{A}_\mu \end{pmatrix} \qquad g' \equiv g\tan\theta_W \tag{8}$$

As the mediator of color charge governed by the $\mathrm{SU}(3)_c$ group, the gluon field has eight color charges. The gluon field-strength tensor is $G_{\mu\nu} = \frac{1}{2}\boldsymbol{G}_{\mu\nu}\cdot\boldsymbol{\lambda}$, where λ is a color parallel of the Pauli isospin matrices based on the eight Gell-Mann matrices:

$$\begin{gathered}
\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}. \\
\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}
\end{gathered} \tag{9}$$

1.1.4 Quantum Electrodynamics

The quantum electrodynamics (QED) Lagrangian is

$$\mathcal{L}_{QED} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \psi(i\not{D})\psi \tag{10}$$

where $F_{\mu\nu}$ is the electromagnetic tensor defined in Equation 11, ψ is the spinor of fermions and D is the gauge covariant derivative. Feynman slashed notation is used to contract a gamma matrix with the slashed term. A_μ is the electromagnetic vector potential.

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{11}$$

$$D_\mu = \partial_\mu + ieA_\mu. \tag{12}$$

QED has U(1) symmetry, meaning it is invariant under the following gauge transformation:

$$\begin{aligned} \psi(x) &\to e^{i\alpha(x)}\psi(x) \\ A_\mu &\to A_\mu + \frac{1}{e}\partial_\mu\alpha(x). \end{aligned} \tag{13}$$

Quantum electrodynamics describes the interactions between relativistic charged particles interacting via electromagnetic force. The force carrier for QED is the photon, a manifestation of the A_μ field.

1.1.5 Quantum Chromodynamics

Quantum Chromodynamic's (QCD) decribes how partons (quarks and gluons) interact with each other via the strong force. Pulling QCD's contribution to the SM's Lagrangian from the first two lines of Equation 1, QCD's contribution can be more explicitly stated as in Equation 14,

$$\mathcal{L}_{QCD} = \bar{\psi}_i(i(\gamma^\mu D_\mu)_{ij} - m\delta_{ij})\psi_j - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a \tag{14}$$

where $G^a_{\mu\nu}$ is the gluon field strength tensor, which unlike the QED field tensor, has a quark interaction term

$$G^a_{\mu\nu} = \partial_\mu A^a_\nu(x) - \partial_\nu A^a_\mu(x) + gf_{abc}A^b_\mu(x)A^c_\nu(x)$$

where A_μ are the gauge field of gluons. The gauge-covariant derivative is $\mathcal{D}_\mu = \partial_\mu + igA_\mu$. The field strength tensor also depends on the coupling strength g, and f_{abc} are the fine structure constants of $SU(3)$. The color-triplet quarks make the composite spinor with q standing for a generic quark

$$\psi = \begin{pmatrix} q_{Red} \\ q_{Green} \\ q_{Blue} \end{pmatrix}.$$

QCD has SU(3) symmetry through the color charge and is invariant under the unitary transformation in Equation 15.

$$U(x)\psi(x) \rightarrow e^{-i\theta}\psi(x) \tag{15}$$

Atoms contain proton(s), neutron(s) and electron(s). Protons and neutrons are made of three quarks each - two u and one d quark are within each proton, and each neutron has one u and two d quarks. Quarks are held together as a baryon or a meson by the strong force. Interactions among partons (quarks and gluons) are all mediated by the vector force particles, gluons.

The strong force becomes asymptotically weaker due to a decrease of the QCD coupling constant at short distances or high momentum. This change in the coupling constant results in asymptotic freedom. It allows quarks and gluons inside a hadron to act like independent partons, and allows use of perturbative theory.

At low energies the strong force keeps partons from being seen in isolation. This phenomenon, color confinement, also describes how the strong force acts in the non-perturbative regime. Energy is required to pull partons apart. Before the energy reaches a magnitude where it could pull the partons apart far enough to be free of the other parton's influence, the strong force creates new parton-antiparton pairs out of the vacuum, which keeps the original two partons within their own composite particles.

1.1.6 Electroweak Symmetry Breaking

Symmetries play a very important role in particle physics. Noether's theorem associates every theoretical symmetry with a conserved current [14]. Table 4 shows examples of symmetries and conserved quantities appearing in particle physics. This is seen in quantum mechanics as gauge invariance in the Lagrangian ($\delta\mathcal{L} = 0$). Electroweak (EW) symmetry, the unification of electromagnetism and the weak force, must be broken to give masses to the

Symmetry		**Conservation Law**
Translation in time	$\leftrightarrow$	Energy
Translation in space	$\leftrightarrow$	Momentum
Rotation	$\leftrightarrow$	Angular momentum
Gauge transformation	$\leftrightarrow$	Charge

Table 4: A summary of E. Noether's theorem on the relation between symmetries and conservation laws [14].

W and Z bosons. This symmetry breaking is done by the Brout-Englert-Higgs mechanism.

The Brout-Englert-Higgs Mechanism is given by a scalar term in the Lagrangian [15, 16]

$$\mathcal{L}_{Higgs} = (\mathcal{D}^{\mu}\phi)^{\dagger}(\mathcal{D}_{\mu}\phi) - V(\phi^{\dagger}\phi), \tag{16}$$

where the gauge-covariant derivative is

$$\mathcal{D}_{\mu} \equiv \partial_{\mu} + \frac{ig'}{2}\mathcal{A}_{\mu}Y + \frac{ig}{2}\boldsymbol{\tau}\cdot\boldsymbol{W}_{\mu}$$

with $\boldsymbol{\tau}$ denoting the pauli matrices, Y the hypercharge, and the SU(2) and U(1) gauge bosons: $\boldsymbol{W} = W^1, W^2, W^3, \mathcal{A}$. The potential can be expanded as

$$V(\phi^{\dagger}\phi) = \mu^2(\phi^{\dagger}\phi) + |\lambda|(\phi^{\dagger}\phi)^2. \tag{17}$$

The Higgs scalar field is a SU(2) doublet of complex scalar fields (ϕ). By requiring the non-trivial solution $\mu^2 < 0$, the minimum of the Higgs field gives a vacuum expectation value (VEV) of

$$\langle\phi\rangle_0 = \begin{pmatrix} 0 \\ \sqrt{-\mu^2/2\lambda} \end{pmatrix},$$

which is non-zero. The VEV therefore breaks EW symmetry, turning a SU(2)$\bigotimes$U(1) symmetry into a U(1) symmetry. Incorporating this non-zero VEV back into the Lagrangian allows masses to be obtained for the W, Z, and h bosons and fermions, f, while the photon's mass (m_A) remains zero.

Equation 18 shows how the fermions and gauge bosons gain mass, with m_f, m_W, m_Z and m_h explicitly shown

$$\mathcal{L}_{EWSB} = \Sigma_f \bar{f}(i\not{\partial} - m_f)f - \frac{1}{4}A_{\mu\nu}A^{\mu\nu} - \frac{1}{2}W^+_{\mu\nu}W^{-\mu\nu} + m_W^2 W^+_\mu W^{-\mu} - \frac{1}{4}Z_{\mu\nu}Z^{\mu\nu} + \frac{1}{2}m_Z^2 Z_\mu Z^\mu + \frac{1}{2}(\partial^\mu H)(\partial_\mu H) - \frac{1}{2}m_H^2 H^2. \quad (18)$$

1.1.7 Successes of the SM

The SM has been extremely successful since it was first introduced. It has predicted unknown particles, interactions, and rates very precisely. The successful predictions of the intermediate vector bosons, Ω^- particle, charm, bottom and top quarks as well as the Higgs boson will be discussed.

Glashow, Weinberg and Salam unified the EM and weak theories [17]. Their EW theory posited the existence of the Z boson and explained the W boson present in beta decay. Combined with the Brout-Englert-Higgs mechanism, their model predicts the leading-order masses of the intermediate vector bosons as

$$m_W = \frac{1}{2}vg$$
$$m_Z = \frac{1}{2}v\sqrt{g^2 + g'^2},$$

where v is the Higgs boson's VEV. The W bosons were discovered at the Super Proton Synchrotron (SPS) in 1983 by the UA1 and UA2 experiments [18, 19]. The Z boson was discovered a few months later at the same experiments [20, 21]. These discoveries led to a Nobel Prize being awarded to Rubbia and S. van de Meer.

Once the strange quark was discovered [22], there were four known leptons (e, ν_e, μ, ν_μ) and three known quarks (d, u, s). Unhappy with four leptons but only three quarks, Glashow, Iliopoulous and Maiani theorized a fourth quark to maintain symmetry [23]. The fourth quark was necessary to allow weak interactions that change strangeness by one in charged current interaction but suppress interactions that change strangeness by two. They were proven right by the discovery of the ψ particle composed of a c quark and $\bar{c}$ [24]. The c quark completed the second generation of matter - s, c, μ, ν_μ.

The third generation of matter was first observed with the τ lepton in 1975 [3]. To complete the third generation, a corresponding τ neutrino and pair of up-like and down-like quarks were posited (t, b). Symmetry was conserved between the leptons and quarks in the third generation by the discoveries of the tau neutrino and top and bottom quarks [25–28].

In 1961, before the quark model was developed, Gell-Mann was cataloging mesons and baryons in his Eightfold Way. The Eightfold way is a geometrical pattern based on the charmness, strangeness and charge of particles. While sorting the known particles this way, Gell-Mann predicted a missing particle in his pattern - the Ω^- particle. Not only did he suggest this particle's existence, he also foretold its characteristics - its mass, lifetime, charge, charmness and strangeness. This enabled him to tell experimentalists exactly where to find the particle [29]. This prediction lent credence to the quark model, as the interpretation of the Eightfold Way was determined to be a way to differentiate particles based on their quark composition.

Perhaps the SM's most well known success is of the Higgs boson. Initially posited independently by Brout, Englert and Higgs in the 1960s, the higgs boson gives an explanation for the source of mass of elementary particles as in the Glashow-Weinberg-Salam model [15–17]. Without this boson and their mechanism, there was a fundamental theoretical problem in the SM. The intermediate vector bosons must have zero mass without the Brout-Englert-Higgs mechanism yet experimentally the vector bosons were observed with non-zero mass. Despite the need for the boson to complete the model, the boson remained undiscovered for many decades despite searches at multiple colliders. In 2012 a particle consistent with the SM Higgs boson was found by the ATLAS and CMS collaborations at the Large Hadron Collider (LHC) through multiple decay modes, $H \to WW, H \to ZZ$, and $H \to \gamma\gamma$ [30, 31]. This experimental verification directly led to the 2013 Nobel Prize in Physics being awarded to F. Englert and P. W. Higgs "for the theoretical discovery of a mechanism that contributes to our understanding of the origin of mass of subatomic particles, and which recently was confirmed through the discovery of the predicted fundamental particle, by the ATLAS and CMS experiments at CERN's Large Hadron Collider" [32].

1.1.8 Limitations of the SM

Gravity is the only known force absent from the SM. Calculations of gravity's effect on the known universe show that the SM only explains 4% of the mass-energy in the universe. The remaining 96% is split between dark energy and dark matter [33]. No particle has been experimentally identified as a dark matter candidate has been found, nor has an experimentally verified theory incorporated dark energy. Gravity needs to be added to the SM somehow, and theoretically the graviton has been proposed as gravity's force carrier. The hierarchy problem remains: the vast difference in scale, a factor of 10^{24}, between gravity and the weak force put in by hand with no fundamental explanation.

The goal of the Grand Unified Theory (GUT) is to combine the three most important forces at high energy: electromagnetic, strong and weak. GUT is the intermediate step on the way to the "Theory of Everything" which aims to, quite aptly, describe everything in the universe. The Theory of Everything requires the description of the quantum world to be reconciled with the world at astronomical scales described by general relativity. Physics is able to describe the forces of gravity and electromagnetism well on a macroscopic scale. On the quantum scale the weak and strong forces are also well described, and the EM force is able to be combined with the weak force into the "Electroweak" force described by the Glashow-Weinberg-Salam model [17].

The SM Higgs sector must take into account large radiative corrections and therefore requires fine-tuning. Fine-tuning is the inclusion of dimensionless ratios between free parameters and physical constants that are far from $\mathcal{O}(1)$. It is possible that fine tuning will cancel quadratic radiative corrections and therefore solve the problem, however, introducing a specific term to cancel divergences is not well motivated in a theory so dependent on symmetries. In other words, fine-tuning cancels the naturalness of the SM theory [34].

The neutrino has no mass like the photon in the SM. However, a deficit of neutrinos were found in the Homestake experiment, leading to the idea of neutrino oscillation [5]. Neutrino oscillation shows neutrino flavor changes as a function of distance and energy. The neutrinos interact as three flavor eigenstates, but propagate as three mass eigenstates. This propagation necessitates that neutrinos have a nonzero, albeit experimentally determined very small, mass.

Another unexplained question regards matter-antimatter asymmetry. Immediately after the big bang when all of the universe was theorized to be created, the SM gives no reason to explain the vastly different amounts of matter and antimatter currently seen in the universe. There is also no evidence for any theory that contradicts the symmetry of matter and antimatter being produced in equal and opposite concentrations as in pair production. If matter and antimatter were created in equal amounts, it is not understood how the universe came to be dominated by matter. The SM offers no explanation for the imbalance between the two, or why the original matter created did not simply annihilate with the corresponding antimatter.

Neutrinos offer a possible solution to this asymmetry because they oscillate into different flavors with no cross-flavor interactions. This means that if an electron neutrino - electron anti-neutrino pair was produced, then either could oscillate into a muonic neutrino/anti-neutrino at their next interaction. The conservation of flavored leptonic number demands that an electron neutrino and muon anti-neutrino are unable to interact directly with each other due to conservation laws. It is possible that this neutrino oscillation contributed to the excess matter present in the universe. However, current estimates show that neutrinos by themselves are insufficient to explain the matter-antimatter asymmetry. Additionally the CKM framework contributes a small amount of CP violation in the quark sector, yet combined with neutrino oscillations, the two are still not sufficient to answer this question. This leaves open an unknown process that must take majority responsibility for the asymmetry.

1.2 Search for $H \to b\bar{b}$ with an associated vector boson

The Higgs boson was the final piece of the SM that was missing experimental confirmation to complete the a priori predictions of the SM. Once a boson consistent with the SM was discovered at a mass of 125 GeV, the particle physics world was ablaze with new questions asking if it was *the* SM Higgs boson, or if it was an exotic Higgs boson from another theory [30, 31]. Experimentally, the higgs boson was discovered in just a few decay modes: WW, $\gamma\gamma$, ZZ. Combined these modes represent the SM Higgs boson's probable decay scenarios less than half of the time as seen in Figure 1.

The SM Higgs decay mode, $b\bar{b}$ was not used for discovery. With the largest branching ratio

$$BR_{H\to b\bar{b}} = \frac{\sigma(H \to b\bar{b})}{\sigma(H \to All)} \tag{19}$$

of $57.7^{+3.2\%}_{-3.3\%}$, the $b\bar{b}$ decay mode naïvely would be the expected discovery

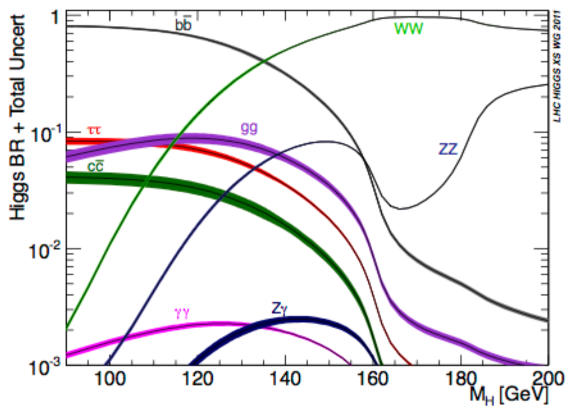


Figure 1: The Higgs boson's predicted decay modes as a function of probability over a span of possible Higgs masses [35].

mode. This is not the case due to large $b\bar{b}$ backgrounds at the LHC. The cross sections of backgrounds with experimentally identical final states far surpasses that of the direct $H \to b\bar{b}$.

In this analysis, the very large backgrounds are reduced by requiring the higgs boson to be produced with a vector boson. The higgs boson and vector boson are required to have large transverse momenta to further suppress backgrounds enough to make the search viable. The cross section of the two processes, WH and ZH, can be compared to direct production of the Higgs boson in Figure 2. The WH process has a larger predicted cross section than the ZH process but both are over an order of magnitude suppressed compared to direct production.

While the associated vector boson requirement reduces the number of signal events, this is more than compensated for by corresponding background reduction. The selected decay modes of the vector bosons are fully leptonic:

- $Z \to \nu\nu$
- $W \to \ell\nu$
- $Z \to \ell\ell$

Here ℓ stands for a charged lepton of either electron or muon flavor, and ν represents a neutrino. Both particle and antiparticle are included in this notation. The leptonic vector boson requirement further reduces the final state

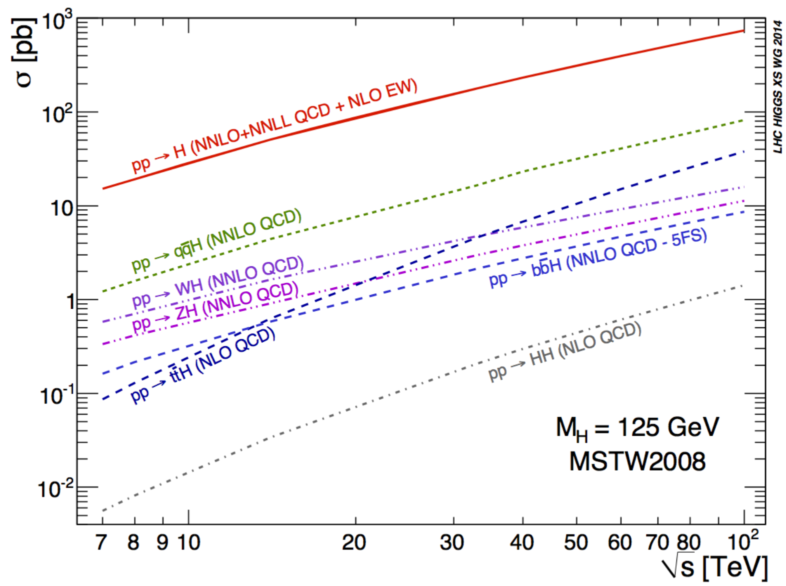


Figure 2: Predicted cross section of various higgs production modes as a function of centre-of-mass energy. Direct production of the Higgs boson is the red solid line, the associated production of a W boson is the purple dashed line and the associated production of a Z boson is the magenta dashed line [35].

topologies available by at least one-third but because the leptonic decays are significantly isolated and identifiable, they decrease the large hadronic backgrounds. While the reconstructed final state only explicitly targets electrons and muons, it is possible to recover some tau lepton decays through its subsequent decay to an electron or muon. The three different VH modes are differentiated by the number of charged leptons in the final state. The leading order Feynman diagrams of these processes can be seen Figure 3.

The following analysis focuses on the WH final state with a single charged lepton. This dissertation will also state results of the combination of WH and ZH, referred to as VH where V stands for both intermediate vector bosons. When combined, the total VH process has enough expected sensitivity to see evidence of a SM Higgs boson decaying to a pair of b quarks. The first questions are whether the $H \to b\bar{b}$ and VH processes can be observed. The next question is whether the $H \to b\bar{b}$ branching ratio is consistent with the SM prediction. If there is a significant deviation from prediction, that would indicate that the Higgs boson found at $m_H = 125$ GeV is not the SM Higgs boson. It would be particularly interesting if there is a dearth of events found because that would indicate that there is available phase space in Higgs boson decay schema for non-SM decays. This would indicate an exotic Higgs boson, opening up an entirely new avenue of physics to explore.

1.3 Beyond the Standard Model

If the $H \to b\bar{b}$ branching ratio measured is less than the 58% predicted by the SM Higgs boson, the Higgs boson at $m_H = 125$ GeV is likely not from the SM. In such a case, the boson discovered in 2012 would be classified as an exotic or BSM Higgs boson. This is an exciting prospect because collider based particle physics currently reinforces the SM's predictions and rarely shows anything inconsistent.

One inconsistent result is neutrinos with non-zero mass. The Neutrino oscillation mixing matrices shows the relationship between their masses and flavors (see Equation 20). The data collected at the Large Electron-Positron collider (LEP) fits the Z^0's decay width to 2.984 ± 0.008 light flavors of neutrinos, consistent with the three known neutrinos [36]. If there are any additional neutrinos they would be heavier and/or sterile meaning they would not interact with the $SU(2)_L \bigotimes U(1)_L$ EW force. Sterile neutrinos are possible candidates for dark matter, and therefore BSM physics [37].

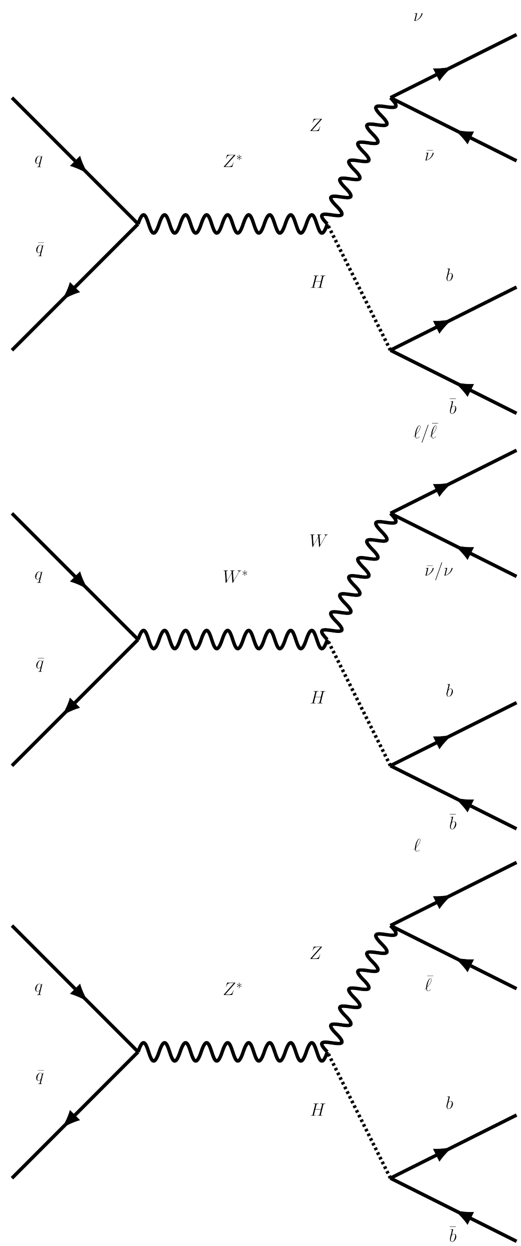


Figure 3: Leading order Feynman diagrams from top to bottom: $ZH \to \nu\bar{\nu}b\bar{b}$, $WH \to \ell\nu b\bar{b}$ and $ZH \to \ell\ell b\bar{b}$. In each scenario, two incoming quarks result in a vector boson and a Higgs boson through a virtual/excited intermediate vector boson. The final state is composed of the Higgs boson's decay to $b\bar{b}$ and the vector bosons decay leptonically in 0, 1 and 2 charged lepton modes.

$$V = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -c_{23} & c_{23} \end{bmatrix} \begin{bmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{bmatrix} \begin{bmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{20}$$

The above equation uses the abbreviations where $s_{ij} = \sin\theta_{ij}$, $c_{ij} = \cos\theta^{\nu}_{ij}$ and δ is a charge-parity violating phase. If the CP violating phase is nonzero it would indicate a new symmetry being broken and therefore new physics.

Another potentially broken assumption is lepton universality. Lepton universality holds that, accounting for the mass of making heavier leptons, particle reactions will have no preference for which flavor is created, e, μ or τ. Currently some B meson decay rate measurements are starting to question lepton universality. Were lepton universality violated, this would require new physics [38]. Other possible BSM physics could be unearthed by evidence of any theory that introduces phenomena outside of the SM such as supersymmetry.

1.3.1 New Physics through Heavy Vector Triplets

The limitations of the SM inspire theoretical physicists to either improve it or develop a new theory to encompass the current knowledge and solve some or all of the limitations. Unfortunately, despite many searches, no experimental evidence has been found in support of these theories. Possible new avenues of physics exist through non-SM Higgs bosons with extra decays compared to the SM. Since $H \to b\bar{b}$ is the Higgs boson's most common decay it holds the largest avenue for a possible decay to take phase space from.

A related avenue of possible new physics is a search for heavy vector triplets (HVT). HVT's production and decay result in a very similar final state as this analysis. The main difference is that the higgs is not restricted to the supposed 'SM' mass of 125 GeV, and so a mass search is performed for the inferred exotic higgs in a simplified model [39–41]. A simplified model is not a complete theory but instead makes assumptions with a phenomenological Lagrangian that has a restricted range of validity in parameter space. This allows experimental input and phenomenological parameterization to be used to describe the signal without a complete theory. It allows physicists to test for a composite Higgs boson.

The HVT phenomenological Lagrangian [41] is given by

$$
\begin{aligned}
\mathcal{L}_V = -\frac{1}{4} D_{[\mu} V_{\nu]}^a D^\mu V^{\nu]a} + \frac{m_V^2}{2} V_\mu^a V^{\mu a} + i g_V c_H V_\mu^a H^\dagger \tau^a \overleftrightarrow{D}^\mu H + \\
\frac{g^2}{g_V} c_F V_\mu^a J_F^{\mu a} + \frac{g_V}{2} c_{VVV} \epsilon_{abc} V_\mu^a V_\nu^b D^{[\mu} V^{\nu]c} \\
+ g_V^2 c_{VVHH} V_\mu^a V^{\mu a} H^\dagger H \\
- \frac{g}{2} C_{VVW} \epsilon_{abc} W^{\mu\nu a} V_\mu^b V_\nu^c .
\end{aligned} \tag{21}
$$

The eponymous heavy vector triplets are $V = (V^+, V^-, V^0)$. The coupling constants g and g_V determine the specific model, weakly or strongly interacting. The typical strength of vector boson interactions is $g_V \approx g \approx 1$ in the weakly coupled model, while in the strongly coupled model the value is $1 < g_V \leq 4\pi$.

The third term of the Lagrangian couples to a pair SM vectors (W_L, Z_L, H) with a coupling strength of $\approx g_V c_H$. The fourth term couples to a pair SM fermions with $\approx \frac{g^2}{g_V} c_F$.

The latter terms describe the couplings among vectors and not to the V decays meaning it is irrelevant for phenomenology in a search incorporating SM vector bosons. Figuring out the dimensionless coefficients c_H and c_F can also help distinguish between the weakly coupled model and the strongly coupled model. In the weakly coupled model, c_H is $\approx -\frac{g^2}{g_V^2}$ and c_F is ≈ 1. Both coefficients are approximately 1 in the strongly coupled model.

With these coefficients known in the different models, measurable phenomenological properties such as cross sections and decay widths can be computed analytically for various possible HVT's masses. This allows searches to be done. Current results do not show any significant excesses and limits have been set for different models and mass of the HVT resonance [39]. Future searches will be done at higher energies and larger luminosities.

2 Experiment

2.1 LHC - The Large Hadron Collider

The Large Hadron Collider is the highest energy collider ring in high energy physics. The LHC provides the best opportunity for particle physicists to test the standard model and search for new physics. The LHC is operated by CERN and is located outside Geneva, Switzerland and crosses the border to France. The LHC and its experiments are home for the largest international scientific collaborations in the world.

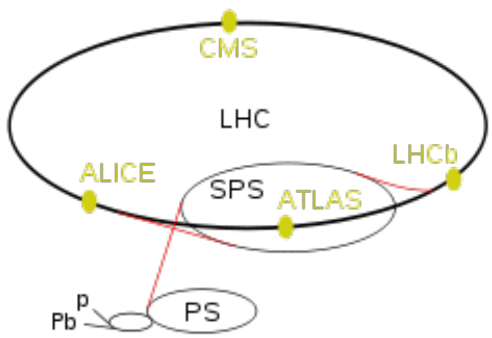


Figure 4: Diagram of the injector accelerators (PS, SPS) leading into the LHC with its experiments in yellow [42].

2.1.1 Delivery System

The previous incarnations of particle accelerators have been repurposed to provide the input to the LHC. These reincarnations are known as injectors.

The beginning of the injection chain (shown in Figure 4) is Linear Accelerator 2, otherwise known as Linac2. Protons that will eventually collide within the LHC's rings start here. These protons originate from hydrogen atoms which are passed through an electric field at the beginning of their journey. The electric field strips the hydrogen atoms of their electrons leaving behind protons at the beginning of the Linac2.

Attempting to collide a single specific proton with another single specific proton is a fool's errand due to the extremely small size of the proton. Instead, a large amount of protons are injected at the same time to form a proton beam bunch. Within this bunch are $1.15x10^{11}$ protons. At fully intensity, a single proton beam consists of 2,808 beam bunches. This large number of beam bunches leads to a much higher chance of collision per beam bunch crossing. The LHC ran at 50ns and 25ns beam bunch spacing during its second run, which is when the data for this analysis was collected.

Once a proton enters a linear accelerator it is accelerated by alternating radiofrequency (RF) cavities. The EM fields produced by the RF cavities are tuned to oscillate at a precise frequency that allows for acceleration of each beam bunch. At the end of the Linac2, each proton has reached an energy of 50 MeV.

After the Linac2, the proton beam bunches enter the Proton Synchrotron (PS) Booster. Made up of four superimposed synchrotron rings, the PS Booster takes 50 MeV protons and accelerates them to 1.4 GeV using RF cavities to accelerate and variable field magnets to provide the circular bending. It is the first circular accelerator in the accelerator chain and has a circumference of 157 meters.

Linear accelerators only have one opportunity to accelerate a particle as it passes through. Circular accelerators, on the other hand, are able to allow particles to make multiple turns, accelerating the particles until a maximum velocity, or energy in these relativistic cases, is achieved. This maximum velocity is due to synchrotron radiation, or braking radiation. Any object moving in a circle is constantly accelerating due to its changes in direction. When a particle does this at high energy, it loses energy via radiative emission. When this energy loss is equivalent to the energy gained by acceleration, the circular collider is at its maximum energy. This maximum energy is proportional to the radius of the accelerator, meaning larger accelerators are needed for higher energies.

The protons leave the PS Booster and are injected into the PS. The PS takes 1.4 GeV protons and accelerates them to 26 GeV within its 628.3 meter circumference. The last step in the injection complex is the SPS, which uses the same mechanics as the PS to accelerate the 26 GeV to 450 GeV in a 6.9 kilometer circumference.

2.1.2 Large Hadron Collider

The final accelerator is the LHC itself. Once the 450 GeV protons are injected into the LHC's 27 kilometer circumference rings, they are accelerated to the collision's center of mass energy, $\sqrt{s}$. In Run 2, the LHC was run at

$\sqrt{s} = 13$ TeV for periods up to an entire 24 hour day with energy symmetric beams of 6.5 TeV each.

The LHC is composed of two beampipes in which the proton bunches travel in opposite directions. Each proton beam travels inside the ring in a vacuum at 10^{-10} Torr. They are steered by dipole magnets. As the proton beams are steered, the beam spot will broaden, and is continually refocused via quadrupole magnets. A cooling system using superfluid helium is used to keep the superconducting magnets at the operating temperature of 1.9 K.

The bunch spacing of 25ns gives a bunch-bunch collision rate of 40MHz. The proton beams collide at four points within the ring, inside the four major experiments at the LHC: ALICE, ATLAS, CMS and LHCb.

ALICE (A Large Ion Collider Experiment) is designed to study heavy ion collisions and the quark-gluon plasma created at the collision point. A specialized run period of heavy ions is run at the end of each year, during which heavy ions are collided instead of protons. The lead ions actually originate in Linac3 instead of Linac2, but use the same accelerator chain as protons for the rest of their acceleration process.

Another interaction point hosts the LHCb experiment (LHC beauty). LHCb is designed to study B-physics and other tests of CP symmetry, resulting in an optimized detector that has an asymmetric design with respect to the beam pipe.

The CMS (Compact Muon Solenoid) experiment is the companion to the ATLAS detector. It is also a general purpose detector designed to study EW physics and discover the Higgs boson, as well as to cross-check the ATLAS experiment.

The ATLAS (A Toroidal LHC ApparatuS) experiment will be discussed in detail in section 2.2.

2.1.3 Luminosity

Luminosity is the metric by which the amount of data collected is measured. The unit of luminosity used by ATLAS is the inverse femtobarn (fb^{-1}). There are two types of luminosity: instantaneous and integrated. The instantaneous luminosity ($\mathcal{L}$) summed over time is the integrated luminosity (L) defined in Equation 22. Additionally, the latter can be separated into delivered and recorded luminosity. Delivered integrated luminosity is the amount of luminosity that the LHC has created in a certain time period, while the recorded integrated luminosity is the amount of luminosity that an experiment was exposed to when in data-taking mode during that time period. The amount of delivered and recorded integrated luminosity during the pp collisions in 2016 can be found in Figure 5.

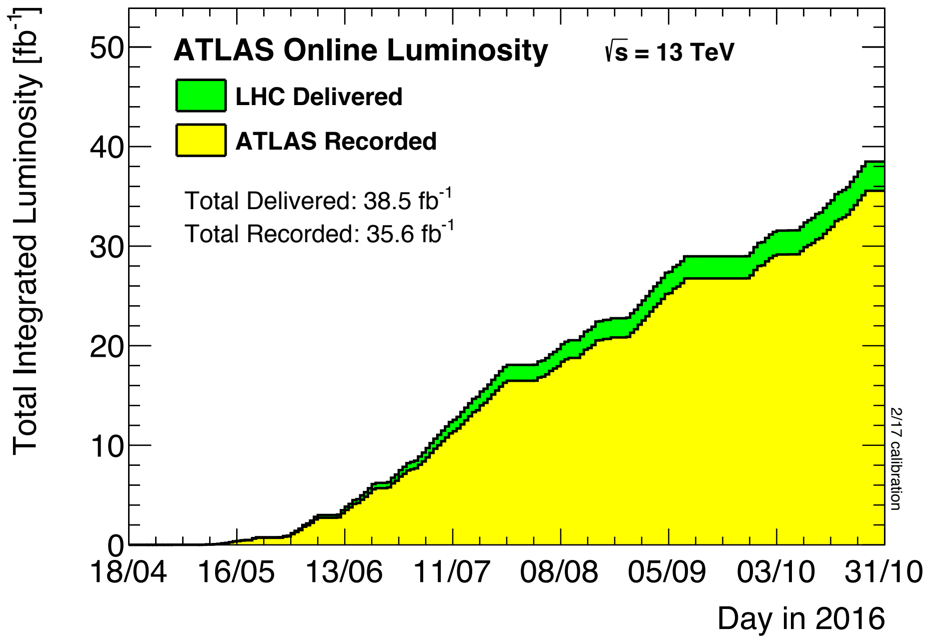


Figure 5: The total integrated luminosity delivered to ATLAS by the LHC (green) and recorded by ATLAS (yellow) in the 2016 pp run [43].

$$L = \int \mathcal{L} dt \tag{22}$$

Instantaneous luminosity is defined below in Equation 23. The two bunches, containing a certain number of particles (n_1 and n_2), colliding at frequency f, give luminosity $\mathcal{L}$ dependent upon the root-mean-square transverse sizes of the beam in the horizontal and vertical directions (σ_x and σ_y).

$$\mathcal{L} = f \frac{n_1 n_2}{4\pi \sigma_x \sigma_y} \tag{23}$$

The number of times a given reaction ($pp \rightarrow X$) happens, N_{events}, is determined by the cross section, $\sigma(pp \rightarrow X)$, and the instantaneous luminosity integrated over time, as defined in Equation 24.

$$N_{events} = \sigma_{pp\to X} \int \mathcal{L}dt \quad (24)$$

Within the ATLAS experiment, there are dedicated detectors called Beam Condition Monitors (BCM) to measure the delivered instantaneous luminosity to the ATLAS experiment as a function of bunch-bunch crossings. The calibration of the BCM as a luminosity detector is done using a technique called the van der Meer (vdM) scan. The vdM scan measures the convolved beam sizes in vertical and horizontal directions (σ_x and σ_y used in Equation 23) and is combined with the beam intensity to give an integrated luminosity calibration constant. This uncertainty in integrated luminosity ($\delta L/L$) is propagated into the analysis' results.

2.2 ATLAS - A Toroidal LHC ApparatuS

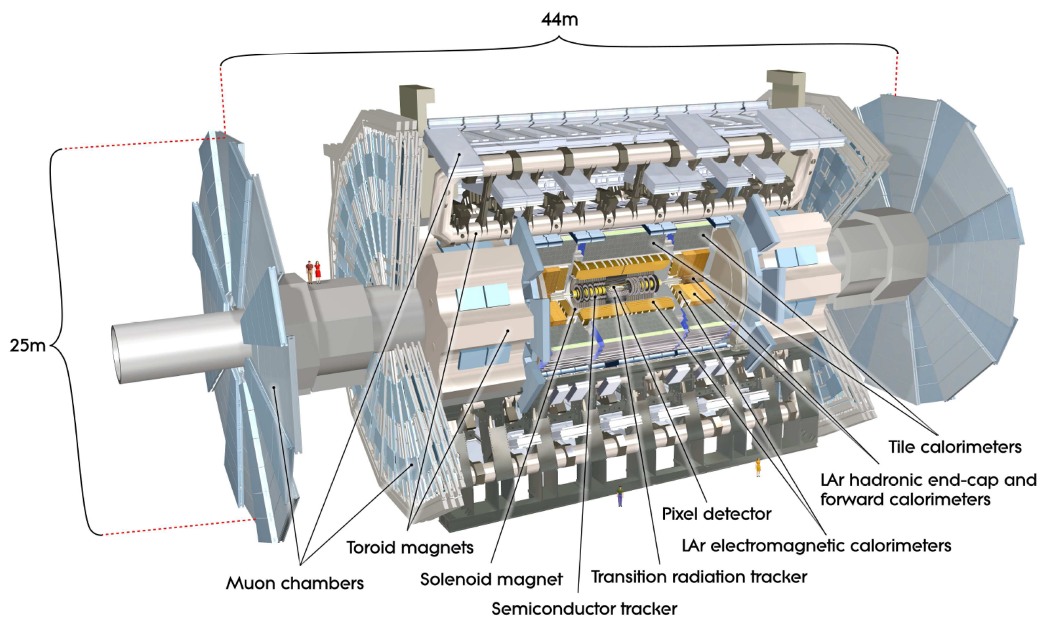


Figure 6: Computer generated view of the ATLAS experiment with the subdetector systems labeled and the dimensions listed [44].

The ATLAS experiment (shown in Figure 6) is a general purpose high energy physics experiment designed to discover the Higgs boson and test the SM.

ATLAS is an extremely large experiment measuring 44m lengthwise, 25m in diameter and weighing 7000 tons [45]. It is situated about 100 meters underground and is composed of the sub-detector elements as detailed below. ATLAS uses a solenoid magnet to produce a 2T field and bend charged particles in curved paths in order to get momentum measurements in its innermost detectors. The outermost elements of the experiment are toroidal magnets which generate a non-uniform magnetic field that varies between 2 and 8T depending on the position.

The center of the detector defines the origin of ATLAS' coordinate system and the nominal interaction point. The z-axis extends along the beam pipe in the direction protons travel. The detector is split into two halves, Sides A and C, with Side A being the positive side of z and Side C the negative. Perpendicular to the z-axis is the x-y, or transverse plane. Positive x points to the center of the LHC ring and positive y points towards the surface of the earth and then via a right-hand requirement, the positive z direction is defined. In the transverse plane cylindrical coordinates (r, ϕ) are used. Pseudorapidty, η, is defined as $\eta = -\ln\tan(\theta/2)$ with θ being the polar angle. With these quantities, the angular distance can be defined between two objects is $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

2.2.1 Inner Detector

When the protons collide inside of the beam pipe, the first detectors the resultant particles encounter are part of the Inner Detector (ID.) Concentrically from the center of the experiment outwards, a particle will pass through the ID in the following order: the Insertable B-layer (IBL), Pixel Detector (PD), Semiconductor Tracker (SCT) and the Transition Radiation Tracker (TRT). Each detector contains a central portion of roughly cylindrical geometry co-axial with the beam directions called the barrel of the detectors, and at higher η there are rings of detectors perpendicular to the beam directions called the end-cap detectors.

The Insertable B-layer is the most recent addition to the ID [46]. Added during the shutdown in 2014, this layer is the closest to the beam pipe at a radius of 33.25 mm. It contains planar silicon detectors in the center and 3D silicon detectors on the high pseudorapidity edges. The IBL upgrade allows for improved resolution of impact parameters and distance of closest approach, which are two major variables used in B hadron identification. This will be discussed further in section 2.3.4.

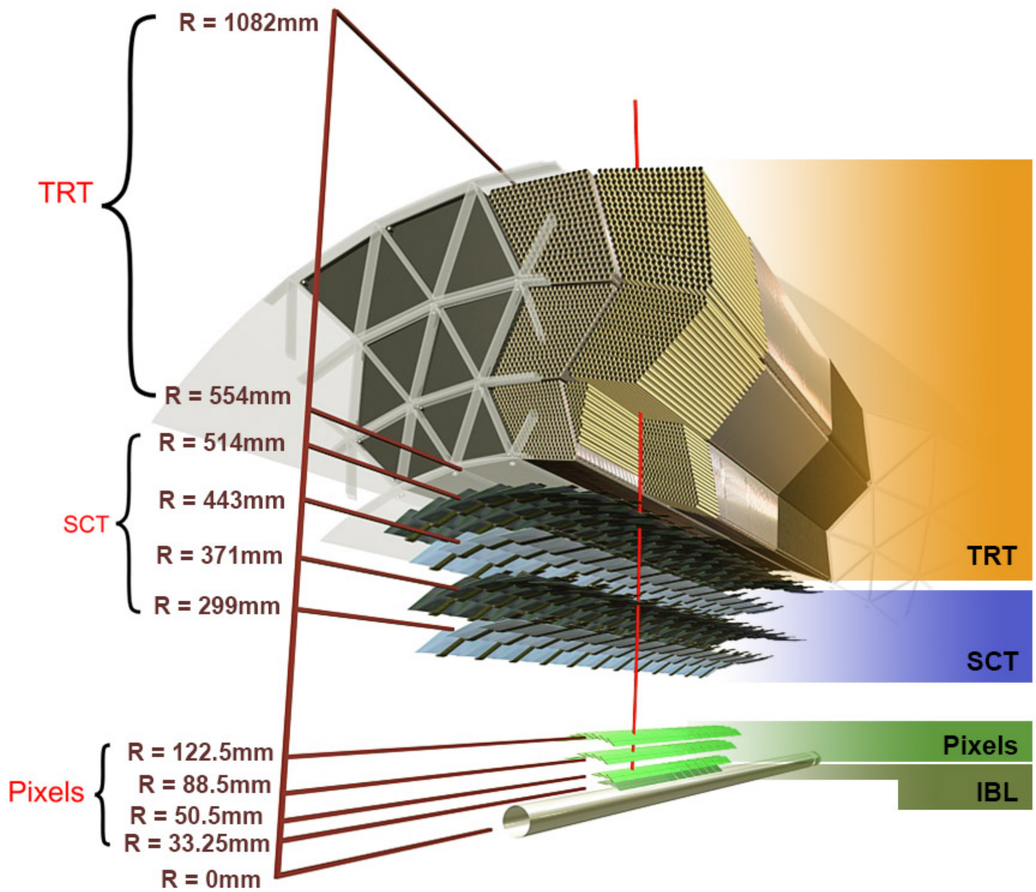


Figure 7: A computer generated slice of the Inner Detector with the subsystems labelled [44].

The PD has three pixel layers ranging from a radius of 50.5 mm to 122.5 mm. It has pseudorapidity coverage of $|\eta| < 2.5$ and is vital for charged particle reconstruction due to its fine granularity [45]. As a charged particle passes through the silicon pixel detectors it ionizes the silicon, releasing electron-hole pairs proportional to the particle's energy that are read out as signal. Each pixel has a resolution of $14\text{x}115\mu m^2$, and the barrel detector has 80 million channels to read out all pixels. The endcaps each have three disks with 6.6 million channels to capture particles further forward in η by positioning the pixels in a radial geometry perpendicular to the z axis [45].

The particle enters the SCT after the pixel detector. The SCT is a silicon microstrip tracker with 4 cylindrical barrel layers and 18 planar endcap discs. The SCT is only able to resolve one directional position, transverse to the strips which are separated by an 70-90 μm pitch with a 40 mrad stero angle between sensors so that upper and lower detector pairs overlap. The accuracy of the SCT's resolution is $17\mu m$ per layer up to a pseudorapidty of $|\eta| < 2.5$ [45].

A charged particle will next interact with the TRT if it is within $|\eta| < 2.0$. The TRT consists of Kapton drift tubes with a 4mm diameter and a gold-plated tungsten wire, 0.03mm in diameter, at the center of each tube [45]. The barrel is composed of 50,000 straws with a length of 144 cm while the end caps have 250,000 straws but are only 39 cm long. Inside the straw tubes is a gaseous mixture of 70% Xe, 27% CO_2 and 3%O_2, with the exception of the areas that have leaks which are filled with a cheaper Ar mixture. The charged particle ionizes the gas and the resultant free electrons then drift to the wire, are amplified and read out. In between the straw tubes are polymer fibers and foils meant to create transition radiation (at the transition of different materials with different indices of refraction.) This transition radiation allows for additional particle identification differentiation between electrons and charged pions for energies of $1GeV < E < 100GeV$. The TRT's resolution is $130\mu m$ along the axis of the straw tube [45].

Due to the magnetic field, a low momentum particle will curve within the inner detector and spiral until it loses all of its energy. The uniform 2T magnetic field inside the inner detector allows for measurement of a particle's momentum and charge's sign assuming the charge is $\pm$ 1e.

2.2.2 Calorimeters

Calorimeters are designed to measure particles' energy. The calorimeters in ATLAS primarily record the energy of electrons, photons and jets, and the energy deposits in the calorimeters are used to calculate the events' missing transverse energy. Designed with high interaction lengths, ATLAS uses two calorimeters to stop most particles by absorption.[1]

If a particle is able to interact with EM forces, it will lose the majority to all of its energy in the EM calorimeter. ATLAS' EM calorimeter is made of liquid argon (LAr) with lead absorbers. It has both a barrel and forward portion allowing for reconstruction up to $|\eta| = 4.9$. The EM barrel calorimeter

[1] The only known particles that regularly pass through calorimeters are muons and neutrinos with the latter not being directly detected in ATLAS.

has a radius of 2.25m and a length of 6.42, and its endcap has the same radius but instead a length of 0.63m. The EM calorimeter has fine granularity with cells ranging from $\Delta\eta$x$\Delta\phi = 0.003$x0.1 to $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ [47].

Absorption in terms of electrons and photons is done via EM showering. Photons and electrons are especially prone to interacting here, starting a cascade of bremsstrahlung and pair production until the particle's energy is spent. The spent energy is recorded by the calorimeter and is calibrated depending on the likely origin particle. The EM calorimeter is optimized for electrons and photons to deposit all of their energy in the LAr in its 25-30 radiation lengths [47].

The remaining hadronic particles are stopped by the Tile calorimeter constructed of iron plates interspersed with scintillation plates with a thickness of about 10 absorption lengths. The calorimeters are composed of both a barrel portion and a forward portion in the endcaps, allowing for reconstruction of $|\eta| \leq 4.9$ [45]. The hadronic calorimeter's barrel portion is at a radius of 2.25 m and has a length of 5.56m. Its endcap sits at a radius of 4.25 and has a length of 2.91 m. The hadronic calorimeter's finest segments are cells sized as $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ [45].

Similar to a particle's interaction in the EM calorimeters, a hadronic particle will induce a hadronic shower. The hadronic shower can consist of electromagnetic components and a hadronic component with strongly interacting particles.

2.2.3 Muon Spectrometer

Since muons are able to penetrate through the detector, a specialized muon detection system is created as the outermost layer to detect muons, called the Muon Spectrometer (MS). ATLAS' muon system is composed of Cathode Strip Chambers (CSC), Thin-gap chambers (TGC) and Monitored drift tubes (MDT) for muon detection as well as Resistive-plate chambers (RPC) which also function as cosmic-ray background rejection.

The MS is similar to the previous detectors in such that it has both a barrel and separate endcap regions allowing for pseudorapitys of $|\eta| < 1$ and $1.05 < |\eta| < 2.7$ respectively [45]. The total MS is about 22 m in diameter and 44 m in length.

The CSC are able to make precision measurements at large pseudorapidities with 70,000 channels at a resolution of $60\mu m$ [45]. There are multiwire proportional chambers inside a gaseous volume. They are orthogonal to beam

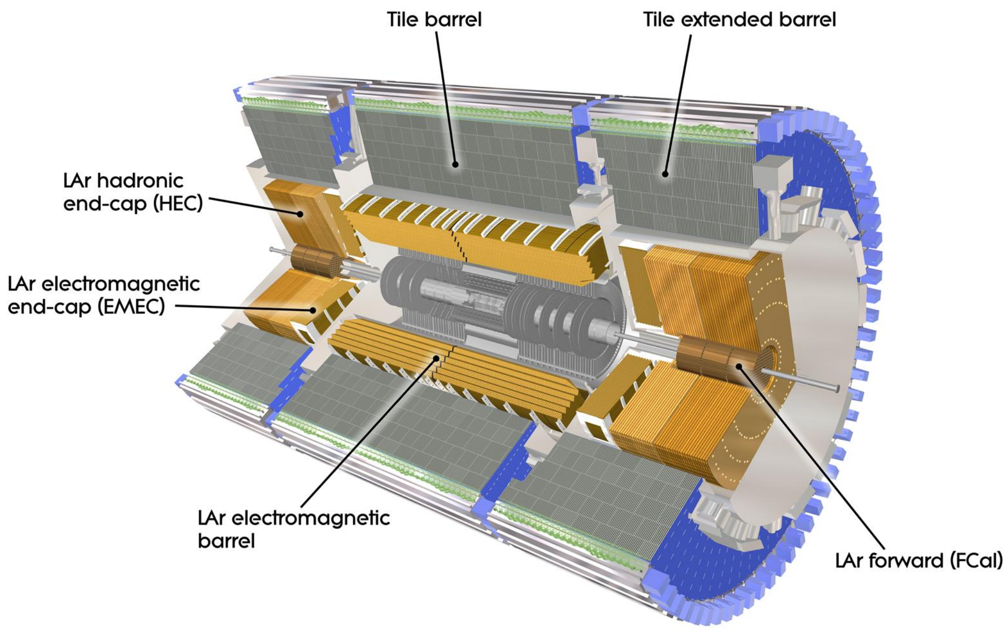


Figure 8: Computer generated view of the ATLAS calorimeters [44].

axis and give a trapezoidal shape to the experiment.

Thin Gap Chambers are designed for triggering and 2nd coordinate measurement at the ends of the detector. They are made of $50\mu m$ diameter gold plated tungsten wires held at a constant positive potential between two grounded cathode planes made of graphite-epoxy mixture. It contains 440,000 channels which provide a resolution of 5-10mm [45].

Monitored Drift Tubes are vital to measure the track curvature. The MDT have 354,240 aluminum tubes each with a 30 mm diameter, a central wire with 50 μm diameter at high voltage, 0.85-6.5 m long and a resolution of $80\mu m$. There is a non-flammable gas mixture at 3 bar inside the tubes that is designed to give a linear relationship between the drift time and drift distance of the signal. The MDT is oriented in the azimuthal direction.

There is a radiation shield used to protect ATLAS from non-collision backgrounds. The Resistive Plate Chambers also are able to help reject cosmic rays from outside sources, and are used for triggering and 2nd coordinate measurements in the central region with a spatial resolution of 5-10mm.

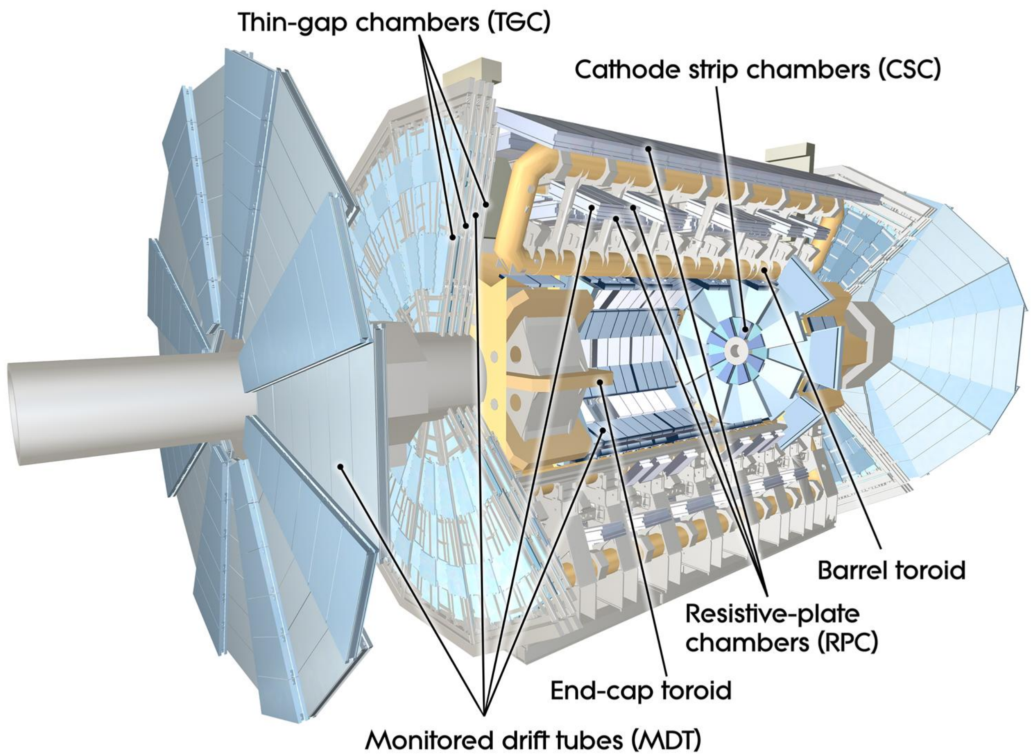


Figure 9: Computer generated view of the ATLAS muon spectrometer [44].

ATLAS has additional toroid magnets to bend these energetic muons into the MDTs. The magnetic field ranges from 2.5Tm in the barrel up to 6Tm in the endcaps.

2.2.4 Trigger Levels

The number of collisions that happen in a second at the LHC is much too large for every single event to be read out and stored. To solve these problem of too much data and too small bandwith, a trigger system is implemented to reduce the number of events to manageable rates. The trigger system has multiple levels, each which reduces the rate while providing more precise (and therefore time consuming) determination of event properties. The triggers are optimized to both keep the recorded event rate low and to capture the most interesting events; the dominant reactions are minimum bias and not very interesting. Minimum bias events are not interesting because they are from soft interactions at relatively low transverse momentum and do not include particles from the hard scattering vertex.

The first layer of the trigger is hardware-based, meaning the decision on whether to keep the event information or throw it away is purely based on pre-programmed decisions made by logic physically present in the readout circuity in less than two microseconds [48]. It keeps 100,000 events of 40 million bunch crossings per second. This first layer is called Level 1.

Level 2 is composed of processors inside the ATLAS cavern. They analyze the event in greater detail than the Level 1 trigger using software algorithms that look at the region of interest, the area in the detector seeded by Level 1 that has the 'interesting' signal instead of the whole event [48]. While this decision is being made the full event information is collected into buffers ready to be transferred out if the event is one of a few thousand per seconds that pass a Level 2 selection.

The final decision is software based, and is commonly referred to as the High Level Trigger (HLT). The full event information has already been saved out to a large computer farm. The entire event's information is used in the algorithms to make a decision on whether to keep an event or not. This final decision passes about 200 events per second to a data storage system [48].

There are two potential problems that the trigger can create. First, if a new physics process is not similar to the final states being triggered then there will be no chance of seeing the data. Additionally, at the lowest level, Level 1, there are so many events it is possible for the datastream to fill temporarily thereby blocking further triggering. If there are too many triggered events in a short span of time, it is possible to miss events that would otherwise be triggered due to this buffering time, the equivalent of five bunch crossings. Since an event that is rejected is lost forever, it is important to maintain high trigger efficiencies at all levels.

2.3 Reconstruction and Particle Identification

Particles passing through the ATLAS experiment leave only energy behind. From this energy, its distribution in the detector and other properties, physicists create algorithms to identify which particle most likely left the track or energy deposit, its momentum, energy and point of production if possible. The following sections describe reconstruction algorithms and particle identification specific to each type of particle.

2.3.1 Electrons

Electrons are reconstructed via the energy deposits in the EM calorimeter combined with the tracks they leave behind in the ID. Starting from a seed deposit (clusters) in the EM calorimeter, a sliding-window algorithm is used to combine associated nearby clusters into a region-of-interest (ROI) with a cone size $\Delta R = 0.3$ [49]. This is only done for seeded clusters which pass loose shower shape requirements.

The next step in identifying electrons is to reconstruct and associate tracks to the seed EM clusters, also known as pattern recognition and track fitting. First the clusters are tested with a pion hypothesis for energy loss. If there are insufficient tracks (7) in the inner detector associated with the cluster it is retried using an electron hypothesis for energy loss with the requirements that the tracks must fall within an EM cluster ROI and have a track seed $p_T > 1$ GeV. Following this, the track candidates are fitted using a global χ^2 fitter. If the track candidate fails the χ^2 fitter due to high energy loss in the pion hypothesis, it is retried with the electron hypothesis [49].

A track candidate is considered loosely matched to the EM cluster depending on matching geometries and the number of hits inside different parts of the inner detector. Success can come from multiple scenarios. One case is if the track has at least 4 silicon hits which can be connected from the primary vertex to the EM calorimeter. If there are less than 4 silicon hits, a track can still be matched to an EM cluster if the last measurement point can be extrapolated to the EM cluster [49].

Once a track candidate has been loosely matched to an EM cluster, it is refitted using an optimized electron track fitter. This is now considered an electron candidate. The electron candidate can have multiple tracks associated to an EM cluster and needs tighter requirements.

Now that an electron candidate has been identified, the candidate is evaluated using a set of variables to determine the likelihood of the electron candidate originating from background objects (hadronic jets, photon conversions, semi-leptonic heavy flavor decays etc.) as opposed to signal objects from more 'interesting' specific physics processes. The candidate's properties (track properties, matching track to cluster, longitudinal and transverse shapes of the EM showers) are organized into different quality categories dependent on the background rejection. This cut based selection defines tight, medium and loose quality electrons with tight having the least background contamination but lowest efficiency, and loose having the largest background

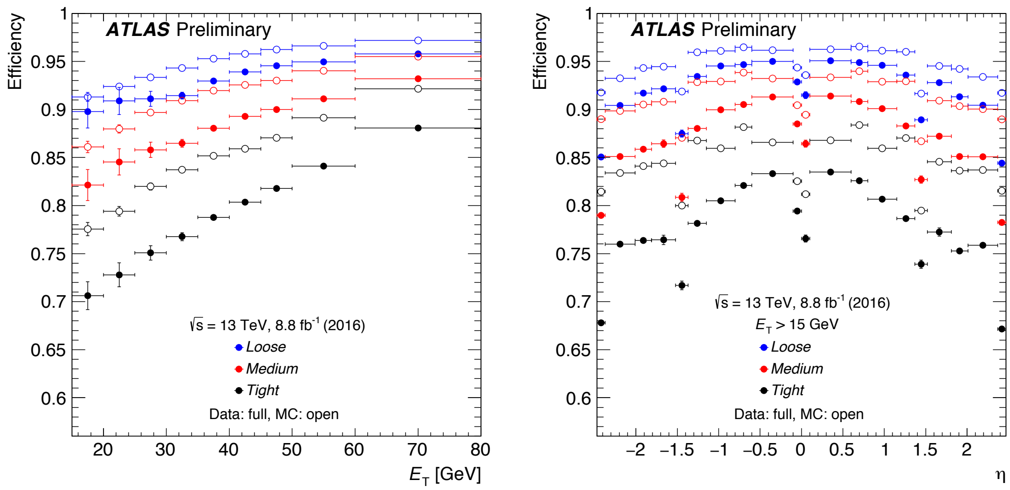


Figure 10: Electron identification efficiency as a function of E_T (left) and η (right) for MC (open circles) and data (full circles). The data efficiencies are taken from $Z \to ee$ measurements in 8.8fb^{-1} of data in 2016 [50].

contamination but highest efficiency for real electrons.

An alternative identification, called likelihood identification (LH), is the currently preferred method. Using signal and background probability density functions (PDFs) of discriminating variables, each electron candidate is assigned a probability of being a signal or background electron. A requirement is applied on the combined discriminant at various efficiences; different values of that discriminant determine the requirements that define loose, medium and tight likelihood identifications of electrons. The efficiency as functions of E_T and η can be found in Figure 10 for the various LHs in data and MC [49].

2.3.2 Muons

Muon reconstruction, like electron reconstruction, also relies on two detector systems and common to both reconstruction methods is the inner detector. There are multiple reconstruction techniques for muons depending on the specific sub detectors used. This leads to different reconstruction efficiencies for the different types of muons.

The inner detector reconstruction of a charged particle is similar to the electron case, discussed in section 2.3.1. Independently, the MS, described in 2.2.3 reconstructs charged particles. Since the toroidal magnets give different bending planes in different areas of the MS, each sub-detector has a different reconstruction technique.

The MDT searches for hits that, when combined, have a path along the bending plane while the RPC and TGC hits are orthogonal to the bending plane. The CSC combines hits in the η and ϕ detector planes. The connected hits in each detector are considered segments. The MDT segments are reconstructed by performing a straight-line fit to the hits found in each layer. Segments are then fed into an algorithm to search for muons. This algorithm starts from the center of the MDT detector and uses any track segments found there as seeds. These are then extended to test if they match with segments in the inner and outer layers. A track requires at least two matching segments in geometric alignment with requirements on hit multiplicity and fit quality, except in the barrel-endcap where a single segment of high quality can count as a track itself [51].

To allow for close-by muons, the same segment can be used to reconstruct multiple muon track candidates. It is later dealt with during overlap removal where it can either be given the best assignment or can be shared among two tracks. The quality of tracks are determined using a χ^2 fit [51].

The first muon discussed is a combined muon (CB). This muon is reconstructed from a global fit in the MS and ID. A muon candidate is reconstructed in the MS and extrapolated to find a matching track in the ID [51]. Complementary to this, the inside-out reconstruction taking the ID track and then finding a matching MS track, is sometimes used.

Segment-tagged (ST) muons are reconstructed by an inside-out algorithm, and they rely heavily on the ID. The tracks in the MS come from either MDT or CSC chambers and often only cross one layer of the MS chambers due to either low p_T or their position is in a low acceptance area of the detector.

Calorimeter-tagged muons (CT) are optimized to find muons that geometrically pass outside of the MS ($|\eta| < 0.1$). This is done by matching a track in the ID with an energy deposit in the calorimeter compatible with a minimum-ionizing particle [51].

The final type of muons are extrapolated muons (ME). These muons are geometrically outside of the ID ($2.5 \leq |\eta| < 2.7$) and instead are fully reconstructed in the MS. The tracks in the MS must produce hits in at least two layers of the MS, and in the forward region must hit at least three layers [51].

After all of the muon types are reconstructed, an overlap removal is applied to any muons that share an ID track. The CB muons are given priority, and afterwards the ST and then CT muons. For ME muons, which don't have any hits in the ID, preference is given to tracks with better fit quality and more hits.

2.3.3 Jets

Jets are the result of a Lorentz boosted quark or gluon's hadronization. Proper reconstruction of jets in this analysis is extremely important because they are in the final state of the signal, $H \to b\bar{b}$.

There are multiple jet reconstruction techniques dependent upon the detectors used and the size of the jet. There are track jets which are composed exclusively from the ID, and calorimeter based jets which can be reconstructed using one of two algorithms: anti-k_T and Cambridge/Aachen. Cambridge/Aachen [52] is shown to be more beneficial for large jets, denoted "fat jets", which can be deconstructed into jet substructure. Jet substructure can give information about the individual jets that merged into the fat jet. This section focuses on the anti-k_T algorithm because it is the algorithm exclusively used by this analysis.

Jets reconstructed with the anti-k_T algorithm [53] are built using calibrated topological clusters from energy deposits in the calorimeters [54] that are calibrated at the hadronic scale as opposed to the electronic scale because jets are composed primarily of hadrons. The size of the jets is described using a radius parameter, which harkens back to the days when cone based reclustering algorithms were used. $\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ where y_i is the rapidity (which is approximately equivalent to the pseudorapidity for high relativity particles) and ϕ_i is the azimuth. Typical sizes for track jets range from $0.2 < \Delta R < 0.4$. The anti$-k_T$ jets are most commonly reconstructed with $\Delta R = 0.4$ in ATLAS, and fat jets are commonly reconstructed with $\Delta R = 1.0$ or 1.2.

Jet algorithms are based upon the distance between particles/pseudojets, d_{ij} defined as

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p})\frac{\Delta R_{ij}^2}{R^2} \tag{25}$$

and the distance from one of those objects to the beam, d_{iB} defined as

$$d_{iB} = k_{ti}^{2p} \tag{26}$$

where k_{ti} is the transverse momentum of an entity and R was defined previously. The algorithm looks for the smallest distance between any pair of clusters [53]. If the minimum is d_{ij} then the two entities are combined, and if the minimum is d_{iB}, the distance between the entity and the beampipe, then the entity is called a jet and removed from consideration for further jets. This process is repeated until the list is depleted of entries.

The main difference between the algorithms originates in the exponential term of Equation 25, p. When $p = 0$ the distance gives the Cambridge/Aachen algorithm, while $p = -1$ denotes the anti-k_t algorithm.

A major benefit of the anti-k_T algorithm is that the resultant jets are regular cone-like shapes. Negative p values give hard radiation priority in the clustering algorithm, meaning that the anti-k_T algorithm clusters hard radiation before soft radiation resulting in a more regular shape compared to its competitors.

2.3.4 B Jet Identification

Simple jet reconstruction is not enough to find the $H \to b\bar{b}$ signal. The jets that define the Higgs system must also be b-tagged, meaning that there is a high chance the jets originated from a b or $\bar{b}$ quark. Heavy flavor quarks, beauty and charm, have different qualities when they hadronize and are identified as a jet as compared to so-called light jets (up, down, strange initiated or due to gluon fragmentation). Top quarks decay before they can hadronize and so are not considered part of heavy flavor. ATLAS combines three different B hadron identification algorithms into a single multi-variate algorithm to identify b-jets [55]. This is also known as b-jet tagging.

The first algorithm type is based on impact parameters. Bottom quarks have a long lifetime ($c\tau \approx 450\mu m$) and therefore the quark can survive quite far from the collision point (primary vertex.) The b-quark will decay from this displaced position. The resultant jet can be traced back to the position where the b-quark decayed, denoted as the secondary vertex. It is secondary compared to the primary vertex. The primary vertex, originating from the hard scatter collision, is identified as the vertex with the highest p_T. The secondary vertex on the other hand, denotes the origin of a b-quark's decay. Continuing through the decay vertex to the z-axis along the beam pipe shows that the jet is inconsistent with the primary vertex. Two quantities are used to quantify this inconsistency: the transverse impact parameter (d_0) and longitudinal impact parameter ($z_0 \sin\theta$). These two quantities are defined using the distance of closest approach in the $r - \phi$ plane of the track and of the track to the primary vertex in the longitudinal plane respectively.

The second algorithm is the Secondary Vertex Finding Algorithm. This algorithm explicitly reconstructs a two track vertex that is significantly displaced from the primary vertex and looks for qualities present in b-decays. The additional properties are used for rejecting long lived particles, photon conversions or hadronic interactions with the detector material by using invariant mass calculations.

The final input algorithm is a decay chain multi-vertex reconstruction algorithm [56]. It looks to reconstruct the weak b and c flavored hadron decays inside the jet explicitly.

The qualities computed by these three algorithms are used as inputs to a Boosted Decision Tree (BDT), a multivariate analysis (MVA), to better identify b-jets than any of the algorithms individually. The algorithm, MV2, can be trained on different proportions of c-flavored jets contaminating the b-flavored jets in order to enhance the charm rejection. It is further optimized to achieve a large light-flavored jet rejection. Various b-jet efficiencies are tested in this analysis and the optimal algorithm was found to be MV2c10 which indicates that the fractional c-flavored contamination the algorithm is trained on is 10%.

2.3.5 Missing Transverse Energy

Neutrinos are the most elusive particles of the SM with an extremely small cross section due to their interaction which is exclusively via the weak force. Neutrino detection necessitates extremely large, dense detectors, and therefore ATLAS does not directly detect them.

ATLAS calculates the neutrino's contribution to the event by looking for the missing energy in the event.[2] Since the neutrino is not measured, the rest of the event's energy is unbalanced and the neutrino's energy can be approximated by a subtraction of the energy of the physics objects present in the event with geometrical considerations. This is limited because ATLAS can only reconstruct the transverse energy of the neutrino.

There are two separate calculations of the missing transverse energy (E_T^{miss}/MET) utilized in this thesis. The first is at the trigger (MET) and the second is at the reconstruction level (E_T^{miss}).

[2]The missing transverse energy can also potentially identify dark matter which, similar to neutrinos, is not expected to interact within the ATLAS experiment's detectors.

There are triggers based upon a completely calorimetric calculation. In this scenario, all of the energy deposited in the calorimeters' clusters, $\vec{p_T}^i$, are inputs to the MET calculation as seen in Equation 27.

$$M\vec{E}T = -\Sigma\vec{p}_{T,i} \tag{27}$$

At analysis level, $E_T^{miss}$is calculated using the objects identified as distinct types of particles in the final state and an additional term composed of soft signals, shown in Equation 28.

$$\vec{E}_T^{miss} = -\sum_{i\in(hard)} \vec{p}_{T,i} - \sum_{j\in(soft)} \vec{p}_{T,j} \tag{28}$$

The hard term is the composed of the transverse momentum of the reconstructed physics objects in the event [57]. The physics objects in this analysis include jets (j) and leptons (ℓ). To avoid double-counting of geometrically nearby tracks or calorimeter clusters, objects are identified in the following order of priority: electrons, photons, hadronically decaying τ leptons, μ, jets. Once the hard term has been created using the physics objects, the remaining tracks in the ID, which are associated with the primary hard scatter collision vertex and are not associated to any physics object, are combined to make the soft term.

3 Signal and Background modeling

To search for any physics phenomenon one must first remove the reducible backgrounds as far as possible and understand the irreducible backgrounds. Reducible backgrounds are events that have artificially similar final states due to misidentification of physics objects (particles) in the final state. Irreducible backgrounds are events that share the same final state with the signal that is being searched for.

The final state this analysis looks for is a charged electron or muon and at least two jets that should originate from b-quarks. Physics processes that share the same final state, mimicking WH, are top pair, vector boson plus multiple jets, diboson production, single top and multijet production. The dominant backgrounds in this analysis are $t\bar{t}$ and $W + jets$. This chapter will discuss the modeling of the signal and the backgrounds, separated by whether the process is modelled by Monte Carlo (MC) samples or a data-driven technique.

3.1 Monte Carlo samples

All samples that are modeled by MC undergo the same processing. The generators are made using theoretical and previous measurements as inputs to create descriptions of collisions and the resultant particles. The events are simulated with multiple collisions per bunch crossing and pile-up effects due to this are accounted for by overlaying minimum bias events. The results then go through a simulation (GEANT-4 [58]) that applies detector effects replicating passage through the ATLAS detector. Finally the MC events are reconstructed using standard ATLAS reconstruction techniques. This is done to properly compare the events recorded in the detector to the events the MC provides.

Each MC sample will be detailed with the generator used, describing the physics used to simulate the events, the cross sections multiplied by the branching ratio that gives the relative contribution of the sample as well as the number of events in MC to indicate the statistical power of the MC samples.

3.1.1 VH Signal

The leptonic final state of $WH \rightarrow \ell\nu b\bar{b}$ has only one charged lepton. The ZH modes have small contributions due to possible misidentifications of the leptons, so the VH signal sample contains multiple sources. Even though this analysis reconstructs $W \rightarrow e$ or $W \rightarrow \mu$, the $W \rightarrow \tau$ process also contributes when the τ's subsequent decay contains an e or μ.

The signal is created via two separate initial partonic collision types in the LHC: quark-quark and gluon-gluon. Separate samples are generated for each case.

Table 5 contains the following information for each contributing signal process: generator used, cross section times the branching ratio($\sigma\times$BR in femtobarns), and the number of MC events in the sample used. The signals originating from quark interactions are modeled using POWHEG MINLO + PYTHIA 8 [59] with NNPDF3 , a set of parton distribution functions (PDF) based on leading order, next-to-leading order and next-to-next-to-leading order QCD theory including electroweak corrections [60]. POWHEG is a generator that matches next-to-leading order (NLO) QCD computations with parton shower simulations [61], and MINLO stands for Multi-scale improved NLO which describes factorization and renormalization scales at NLO in perturbation theory [62]. Signal samples with gluon-gluon origins use POWHEG + PYTHIA 8 with the CT10 tune, the standard PDF setup used in ATLAS for jet event simulations [63].

Process	Generator	σx BR[fb]	N_{events}
$qq \rightarrow ZH \rightarrow \nu\nu b\bar{b}$	POWHEG MINLO + PYTHIA 8 (NNPDF3)	89.08	2M
$qq \rightarrow WH \rightarrow \ell^{+}\nu b\bar{b}$	POWHEG MINLO + PYTHIA 8 (NNPDF3)	164.6	1M
$qq \rightarrow WH \rightarrow \ell^{-}\nu b\bar{b}$	POWHEG MINLO + PYTHIA 8 (NNPDF3)	104.5	2M
$qq \rightarrow ZH \rightarrow \ell\ell b\bar{b}$	POWHEG MINLO + PYTHIA 8 (NNPDF3)	44.84	3M
$gg \rightarrow ZH \rightarrow \nu_e\bar{\nu_e}b\bar{b}$	POWHEG + PYTHIA 8 (CT10)	4.761	0.1 M
$gg \rightarrow ZH \rightarrow \nu_\mu\bar{\nu_\mu}b\bar{b}$	POWHEG + PYTHIA 8 (CT10)	4.761	0.1 M
$gg \rightarrow ZH \rightarrow \nu_\tau\bar{\nu_\tau}b\bar{b}$	POWHEG + PYTHIA 8 (CT10)	4.761	0.1 M
$gg \rightarrow ZH \rightarrow \ell^{-}\ell^{+}b\bar{b}$	POWHEG + PYTHIA 8 (CT10)	7.228	0.7 M

Table 5: Details of the generators used for simulating the signal VH processes, their contribution through the cross section times branching ratio value as well as the number of events in each MC sample.

3.1.2 V+jets

One of the major backgrounds is a W boson produced in association with jets ($W + jets$). The MC to describe this background, as well as its counterpart with a Z boson, is split into many modes based upon which vector boson is produced (Z or W) and the flavor of the jets. Additionally, the MC samples are split by the vector boson's transverse momentum, p_T^V and the sum of the jet's p_T, H_T, at generator level in order to ensure that statistical fluctuations are small compared to data. All of the samples in this process are generated with SHERPA 2.2.1, the Simulation of High-Energy Reactions of PArticles which does NLO generation [64]. The SHERPA generator uses the 5-flavor scheme with massless b quarks and c quarks, however, c and b quarks must have mass which is taken into account during the parton shower. Further details can be found in Table 6.

Process	Generator	σxBR[pb]	N_{events}
$Z \to \nu\nu$ + jets	SHERPA 2.2.1	1914	146M
$W \to \ell\nu$ + jets	SHERPA 2.2.1	20080	410M
$Z\gamma$ + jets $\to \ell\ell$	SHERPA 2.2.1	2107 ($m_{\ell\ell} > 40$GeV)	171M

Table 6: Details of the generators used for simulating the $V + jets$ processes, their contribution through the cross section times branching ratio value as well as the number of events in MC used in the samples.

The selection criteria used to define the flavor of a sample are detailed in Table 7. The heavy flavor (HF) contribution coming from the two heaviest quarks include $V + bb$, $V + bc$, $V + bl$, $V + cc$, $V + cl$ where l stands for the three remaining quarks up, down and strange. The flavors stand for the two jets with the highest p_T in the sample.

A filter for b quarks is applied (BFilter) and subsequently a filter is applied for c quarks (CFilterBVeto) at the same time as a veto on events which passed the b quark filter. This filter has slightly different requirements, detailed in Table 7, for the Z and W bosons. The remaining events are collected into a single group (CVetoBVeto.)

3.1.3 Diboson

As with the V+jets samples, the diboson backgrounds are also generated using SHERPA [64]. For the processes of WZ and ZZ (collectively called

Filter	Description
BFilter	at least 1 b-hadron with $p_T > 0$ GeV and $\|\eta\| < 4$
CFilterBVeto (Z)	at least 1 c-hadron with $p_T > 4$ GeV and $\|\eta\| < 3$ veto events that pass the BFilter
CFilterBVeto(W)	at least 1 c-hadron with $p_T > 4$ GeV and $\|\eta\| < 4$ matched by $\Delta R < 0.5$ to a truth jet ($AntiKt4TruthJets$) with $p_T > 15$ GeV and $\|\eta\| < 3$ veto events that pass the BFilter
CVetoBVeto	veto events that pass the BFilter or the CFilterBVeto

Table 7: Details of the jet flavor selection criteria in $V + jets$ samples.

VZ), the version used is 2.2.1 while version 2.1.1 is used for the WW process due to operating constraints. The WH-like signal state can arise from many combinations, for example $Z \to b\bar{b}$ and $W \to \ell\nu$ or $W \to q\bar{q}$ and a misidentification of $Z \to \ell\ell$. Details about the diboson samples can be found in Table 8.

Process	Generator	σxBR[pb]	N_{events}
WW	SHERPA 2.1.1	49.74	14.4M
WZ	SHERPA 2.2.1	21.65	28.6M
ZZ	SHERPA 2.2.1	6.53	10.8M

Table 8: Details of the generators used for simulating the diboson processes, their contribution through the cross section times branching ratio value as well as the number of events in MC used in the samples.

3.1.4 $t\bar{t}$

Top pair production, otherwise known as $t\bar{t}$, is generated with POWHEG [61] using NLO calculations and is interfaced with PYTHIA 8.210 [59] using the A14 set of parameters. This sample requires that at least one of the top decay's resulting W bosons decays leptonically. More details regarding $t\bar{t}$'s cross section and statistics used can be found in Table 9.

3.1.5 Single Top

Single top production is generated independently for three scenarios depending on the location of the vector boson in the Feynman diagrams shown in Figure 11: the s-channel, the t-channel and the Wt-channel. The single top samples are produced with the POWHEG generator [61] and CT10

Process	Generator	σxBR[pb]	N_{events}
$t\bar{t}$	POWHEG + PYTHIA 8	831.76	60M

Table 9: Details of the generator used for simulating the top pair process, its cross section times branching ratio value and the number of events in MC used in the samples.

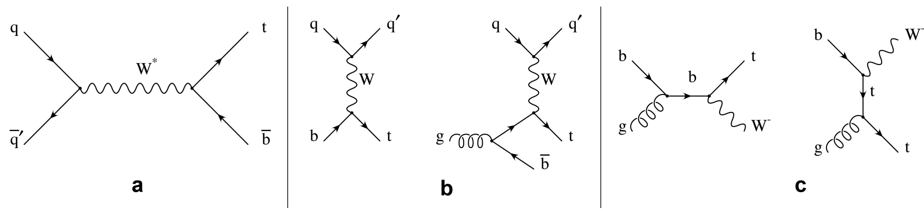


Figure 11: Feynman diagrams of the three single top processes: s-channel (a), t-channel (b) and Wt-channel (c).

tune [63], however, they are interfaced with PYTHIA 6 [59] unlike in $t\bar{t}$. To handle τ decays, PYTHIA 6 is interfaced with TAUOLA [65].

The single top s- and t-channels require a leptonic decay of the W boson. The Wt-channel is generated inclusively, meaning they include all of the final states of the Wt-channel in the same sample.

Further details can be found in Table 10.

Process	Generator	σxBR[pb]	N_{events}
s channel	POWHEG + PYTHIA 6	3.31	2 M
t channel	POWHEG + PYTHIA 6	66.51	10M
Wt channel	POWHEG + PYTHIA 6	68.00	10M

Table 10: Details of the generators used for simulating the single top processes, their contribution through the cross section times branching ratio value as well as the number of events in MC used in the samples.

3.2 Data Driven sample: Multijet

Sometimes there is insufficient information to generate a MC sample that accurately replicates the sample in data. For these types of samples, a data driven approach is chosen to model the background. This is done for one background in this analysis: multijet.

The multijet background is composed of a few different types of events. The main sources of multijet backgrounds are heavy flavor semi-leptonic decays and the geometric overlap of charged and neutral pion decays in jets that are not expected to be well isolated but pass the isolation requirements used in the selection. By only considering the high p_T^W region, the multijet background is significantly reduced. It is, however, still a non-negligible background in the signal region. Additionally, even though the lepton selection and isolation in the main analysis is good at rejecting non-isolated backgronds, the large QCD production cross sections means that the multijet background is significant enough to warrant modeling in the 1 Lepton VH mode.

First the template method that is used in the main analysis will be discussed. Next, the cross-check fake factor method will be discussed.

3.2.1 Template Method

The template method is used to estimate the multijet contribution in the signal region. That method uses a multijet enriched control region to create a distribution that is used in the signal region in the final fit. The multijet enriched control region is taken from data by inverting the lepton's (e/μ) isolation definition. The loose VH signal lepton isolation definition used to separate the 0, 1 and 2 Lepton modes is denoted *IsLooseTrackIso* which has a flat signal efficiency versus lepton p_T of around 99%.

For the 1 Lepton mode, additional tight isolation requirements are made on the signal electron or muon corresponding to around 95% signal efficiency with respect to the loose lepton isolation. The rate of selecting a misidentified election ranges from 3.5-1% for low to high E_T. The tight isolation is evaluated by comparing the transverse momentum of the lepton to the transverse momentum in a cone of radius = 0.2 around the direction of the lepton in the calorimeter (tracker) for electrons (muons). The lepton candidates in the multijet sample satisfy the *IsLooseTrackIso* requirement but fails the tight isolation requirement. A summary of the requirements can be found in Table B.

The template is created using the transverse mass of the W boson, m_T^W, because it has the best separation between the multijet and electroweak backgrounds. To correct for biases from inverting the lepton isolation, event-by-event extrapolation factors dependent on p_T^ℓ, η and in the electron sub-mode E_T^{miss}, are applied to the multijet template. These factors are derived

	Isolated Region	Inverted Isolation Region
e	*IsLooseOnlyIso*	*IsLooseOnlyIso*
	Calorimeter Iso < 3.5 GeV	*Calorimeter Iso* > 3.5
μ	*IsLooseTrackOnlyIso*	*IsLooseTrackOnlyIso*
	Tracker Iso < 1.25 GeV	*Tracker Iso* > 1.25 GeV

Table 11: Summary of the lepton isolation requirements for the tightly isolated and inverted isolation regions. Definitions of the isolation criteria can be found in Appendix B

from a control region with a single jet requirement and a b-tagged jet veto. They are calculated as the ratio between events with an isolated and an inverted isolated lepton.

In the inverted isolation region, the electroweak background is considered contamination and therefore it is subtracted from the data used to model the multijet background to achieve higher purity of the multijet background. The electroweak contamination is modeled by MC events also present in the inverted isolation region.

A normalization scale factor is extracted using a likelihood fit with the templates of the combined top backgrounds, combined $V + jets$ background and the multijet template to data in the m_T^W distribution. The subtraction process results in the final template when combined with the normalization factor. The multijet template is created separately depending on the lepton flavor (e/μ) and jet multiplicity (2/3). The m_T^W distribution before subtraction can found in Figure 12 for various regions. The template is sensitive to the normalization of the MC samples, and so the dominant backgrounds' normalizations ($t\bar{t}$ and $W + jets$) in the signal region are simultaneously extracted with the multi-jet estimates in both the electron and muon modes. The resultant normalizations of the MC backgrounds can be found in Table 12.

Region	Top ($t\bar{t}$+single top)	$W + jets$
2 b-tags, 2 jets	1.01	1.39
2 b-tags, 3 jets	0.98	1.35

Table 12: Normalizations of Electroweak contamination of multijet templates in the isolated lepton region.

The distributions for the kinematic variables m_T^W and E_T^{miss}can be found

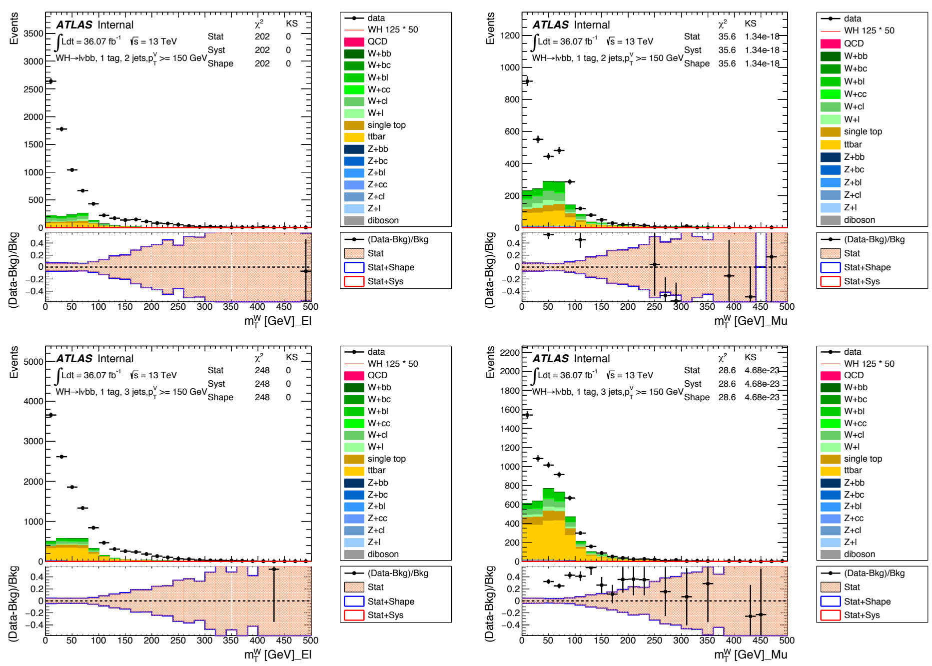


Figure 12: The m_T^W distributions before subtraction in the inverted lepton isolation region, high p_T^W region are shown in different modes: 2 signal jets in upper left is 2 jets in e mode, in upper right is 2 jets in μ mode, in lower left is 3 jets in e mode and in lower right is 3 jets in μ mode.

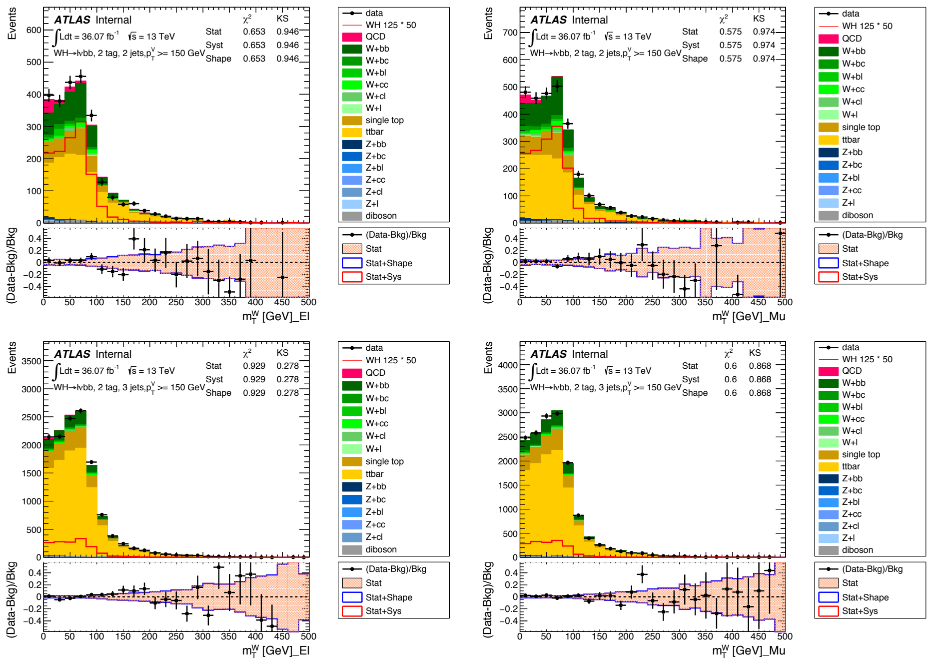


Figure 13: Postfit m_T^W distributions with multijet background using the template method. The upper left is in 2 jets e region, right is in 2 jets μ region, lower left is in 3 jets e region, right is in 3 jets μ region.

in Figures 13 and 14 respectively. The multijet scaling factors and their uncertainties, multijet background composition percentage and normailzation uncertainties from the aforementioned fitting process can all be found in Table 13.

This template method is the limited by the data sample size in the inverted lepton isolation region. In order to increase the statistics, the requisite 2 b-tags in the final state is loosened to 1 b-tag. The difference in extrapolating between 1 b-tag and 2 b-tags is evaluated as a systematic uncertainty and is discussed further in Section 6.2.

3.2.2 Fake Factor Method

The fake factor (FF) method has success in estimating backgrounds from photons and light jets that come from the multijet background. In this method, the leptons that arise from these photons and light jets are denoted

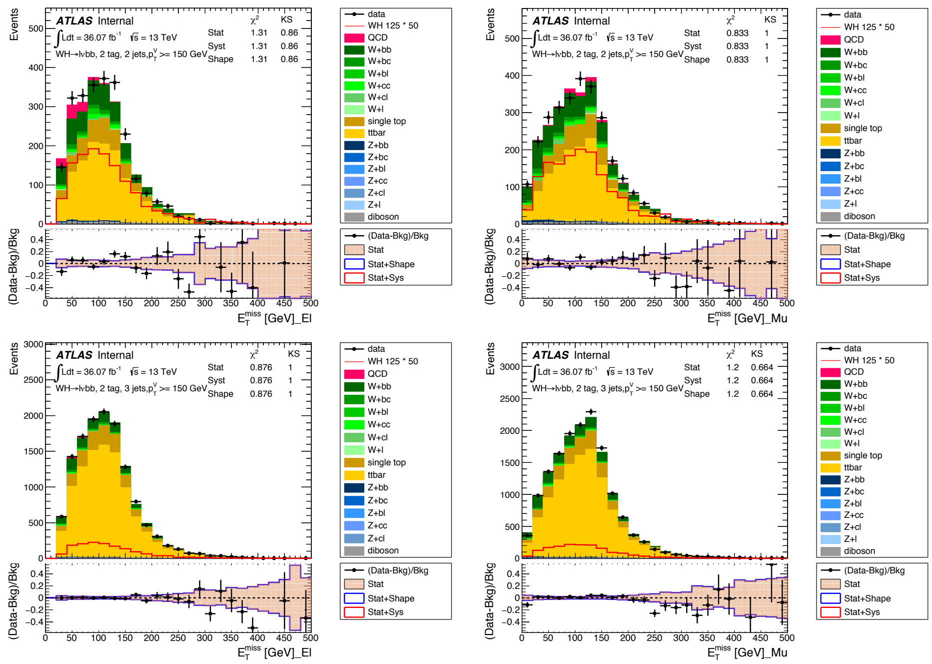


Figure 14: Postfit E_T^{miss}distributions with multijet background using the template method. The upper left is in 2 jets e region, right is in 2 jets μ region, lower left is in 3 jets e region, right is in 3 jets μ region.

Region	MJ Scaling Factor	MJ Fractions (%)	MJ norm. uncertainties
2 tags 2 jets e	$1.91^{+1.18}_{-1.13} \times 10^{-2}$	$4.53^{+2.68}_{-2.79}$	-62%/+60%
2 tags 2 jets μ	$5.62^{+4.35}_{-5.62} \times 10^{-2}$	$3.16^{+2.45}_{-3.16}$	-100%/+76%
2 tags 3 jets e	$0.87^{+1.70}_{-0.87} \times 10^{-2}$	$0.56^{+1.09}_{-0.56}$	-100%/+194%
2 tags 3 jets μ	$2.60^{13.61}_{-2.60} \times 10^{-2}$	$0.38^{+1.99}_{-0.38}$	-100%/+520%

Table 13: Summary of the template method's resultant multijet scaling factors, fractions and their associated total uncertainties as well as their normalization uncertainties.

"fake" leptons. The method relies on extrapolating multiple "fake factors" f based on the lepton's p_T and η.

Equation 29 shows the generic form of the FF. The numerator is made of the number of events in the region which are equivalent to the 1 Lepton mode signal lepton requirement. The denominator is made of the number of events passing the inverted lepton isolation requirements. Table 14 summarizes the exact numerical numerator and denominator requirements for electrons and muons. The inversion of the lepton isolation is done in order to obtain a high purity of the multijet background and also guarantees orthogonality with the signal region. The purity of the multijet background is further achieved by removing the EW contamination via subtraction of the MC simulated samples. The EW background contamination in the sample is significant, therefore, it is important to properly normalize the EW MC contribution. This was found to be best done by fitting the region of 150 $< E_T^{miss} <$ 250 GeV to data and normalizing the main backgrounds, $t\bar{t}$ and $W + HF$.

$$f = \frac{N^{Numerator}}{N^{Denominator}} \tag{29}$$

lepton	iso	denominator	numerator
e	calorimeter	$>$ 3.5 GeV	$<$ 3.5 GeV
μ	tracker	$>$ 1.25 GeV	$<$ 1.25 GeV

Table 14: Isolation requirement between electrons and muons in the numerator and denominators of the fake factor in Equation 29

The FF, f, is calculated using a high statistics dijet sample. An assumption is made that f is constant between the dijet sample selection and the signal selection. This is done by setting the dijet sample's FF equal to the signal's FF. This allows the number of signal events in the numerator which has signal isolation (Equation 30) to be found.

$$N_{WH}^{Numerator} = f \times N_{WH}^{Denominator} = \frac{N_{dijet}^{Numerator}}{N_{dijet}^{Denominator}} \times N_{WH}^{Denominator} \tag{30}$$

The FF is calculated independently for regions with different numbers of jets, and for different bins of the electron or muon's p_T and η. Significant improvement in modeling the electroweak contamination is seen by fitting

	2 jets		3 jets	
	1 b-tag	2 b-tag	1 b-tag	2 b-tag
Signal passing E_T^{miss}> 30 GeV	90%	91%	93%	91%
Multijet passing E_T^{miss}> 30 GeV	33%	33%	42%	42%

Table 15: Expected electron signal and multijet acceptance rates after E_T^{miss}> 30 GeV requirement for a slightly earlier version of the analysis.

the E_T^{miss}distribution, therefore the electron mode is additionally modeled in bins of E_T^{miss}.

The electron mode has an additional requirement of E_T^{miss}> 30 GeV compared to the muon mode. The expected signal and multijet acceptance rates with this additional requirement are listed in Table 15. For a significant reduction in the multijet background there is less than 10% loss in signal.

An additional consideration is whether the multijet backgrounds produced from light jets and from photons have different shape distributions. This was investigated by comparing the electron σ_{d_0} distributions because non-prompt electrons are expected to have the wider σ_{d_0} distribution. No difference in the width of this distribution was observed before and after b-tagging was applied, indicating that the electrons are non-prompt and it is not necessary to consider them separately.

The template method shows better data/prediction agreement for all backgrounds combined, when compared to the fake factor method, and so is used as the default method. The difference between the two methods is used as an experimental uncertainty. Further experimental uncertainties are discussed in Section 6.2.

4 Analysis Selection

The final state of the analysis contains an electron or a muon and missing transverse energy to infer the W boson, and two b-tagged jets to infer the Higgs boson. Different requirements are made to increase the sensitivity of the signal to the expected backgrounds.

This chapter is split into three sections, the pre-selection, the overlap removal and the signal selection.

4.1 Pre-selection

This analysis searches for the specific decay pattern $VH \to \ell\nu b\bar{b}$, one of three VH modes analyzed by ATLAS. There is a common selection designed to preserve orthogonality between the three VH modes. Each mode is given a nickname corresponding to the number of charged leptons identified in the pre-selection: 0, 1 and 2 leptons.

The pre-selection also fulfills a physical constraint: by only selecting events that contain a single lepton that fulfills loose critera (detailed in 4.3) and at least two $R = 0.4$ anti-k_T jets, the amount of data is slimmed down to provide a more manageable data set. The data has the following additional requirements: the event must pass the good run selection that lists luminosity blocks when the detector was running at near to full capacity, the event has a primary vertex, and the event must be a clean event meaning no Tile, LAr, SCT or Core errors. Jet cleaning is applied before any further selections. This good run list, which in this analysis requires the IBL to be working well, gives a data set corresponding to a luminosity of $36.1 fb^{-1}$ from the years 2015 and 2016.

4.1.1 Trigger Requirement

The event is first required to have passed either one of the electron triggers or, if there is a reconstructed muon present, the missing transverse energy (MET) trigger.

In this analysis, the leptonic triggers are the most obvious triggers to utilize. The leptonic HLT can select events based on multiple quantities: lepton p_T, isolation and LH quality of the reconstructed the lepton.

The naïve selection of triggers is the lowest unprescaled single lepton trigger for electrons or muons. A prescaled trigger does not record every

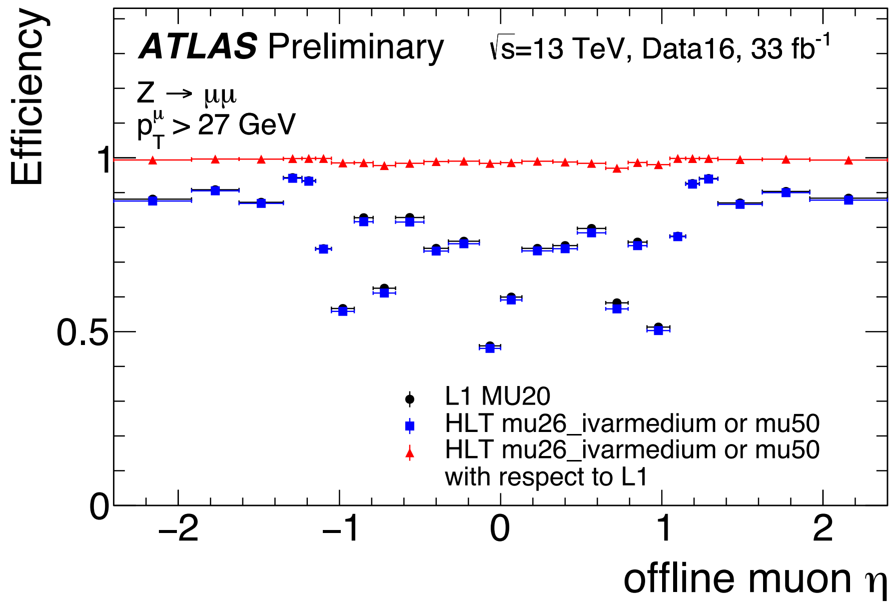


Figure 15: Muon trigger efficiency as a function of the offline reconstructed muon eta [43].

event due to frequency limits. The unprescaled triggers are found to be optimal in the electron case. The triggers are tuned to be optimally efficient with as low a p_T threshold as possible without creating dead time from too high a trigger rate. They are listed in Table 16.

The muon chambers in the ATLAS experiment have gaps in η coverage causing some inefficiency in triggering on muons (see Figure 15.) It is possible to recover some of the inefficiency by using the E_T^{miss}trigger.

The MET trigger is computed from the negative sum of the clusters in the calorimeters. In contrast, the offline E_T^{miss}object is computed from the negative sum of the p_T of physics objects, thus the μ in the final state is accounted for in the offline E_T^{miss}but not the online MET trigger, which means the online MET is actually equivalent to the offline p_T^W.

A p_T^W requirement of 200 GeV gives 100% efficiency with the MET trigger for events including a μ. The selected p_T^W requirement of 150 GeV is selected to achieve good sensitivity to the Higgs boson. Little efficiency is lost by lowering the p_T^W requirement to 150 GeV, as seen by the efficiency curve for the MET trigger in Figure 16.

Period	e	μ	E_T^{miss}
2015	HLT_e24_lhmedium_L1EM20VH OR HLT_e60_lhmedium OR HLT_e120_lhloose	HLT_mu20_iloose OR HLT_mu40	HLT_xe70
2016: A	HLT_e26_lhtight_nod0_ivarloose OR HLT_e60_lhmedium_nod0 OR HLT_e60_medium OR HLT_e140_lhloose_nod0	HLT_mu24_iloose (data) OR HLT_mu24_iloose_L1MU15(MC) HLT_mu40	HLT_xe90_mht_L1XE50
2016 : B-D3	HLT_e26_lhtight_nod0_ivarloose OR HLT_e60_lhmedium_nod0 OR HLT_e60_medium OR HLT_e140_lhloose_nod0	HLT_mu24_ivarmedium OR HLT_mu50	HLT_xe90_mht_L1XE50
2016: D4-E3	HLT_e26_lhtight_nod0_ivarloose OR HLT_e60_lhmedium_nod0 OR HLT_e60_medium OR HLT_e140_lhloose_nod0	HLT_mu24_ivarmedium OR HLT_mu50	HLT_xe110_mht_L1XE50
2016: $\geq$ F1	HLT_e26_lhtight_nod0_ivarloose OR HLT_e60_lhmedium_nod0 OR HLT_e60_medium OR HLT_e140_lhloose_nod0	HLT_mu26_ivarmedium OR HLT_mu50	HLT_xe110_mht_L1XE50

Table 16: Triggers used in data-taking periods during the 2015 and 2016 pp runs.

Comparison of signal sensitivity is shown in Table 17 for each individual trigger as well as combined triggers scenario. The combined scenario does not show a significant improvement over the MET trigger only scenario so the latter scenario was chosen for simplicity.

Region	μ only	E_T^{miss}only	μ and E_T^{miss}
2 jets	1.061 ± 0.002	1.178 ± 0.001	1.175 ± 0.001
3 jets	0.389 ± 0.001	0.449 ± 0.001	0.452 ± 0.001

Table 17: The statistical sensitivity of the signal ($\frac{S}{\sqrt{B}}$) in the μ sub-mode with different trigger scenarios.

The muon sub-mode does not require trigger matching since the neutrino is not explicitly reconstructed and therefore not able to be matched to the phi direction of the E_T^{miss}. An offline muon p_T of 25 GeV is required.

4.2 Overlap Removal

There are tracks and hits in the ATLAS detector that can be attributed to multiple physics objects. In order to decide which physics objects to keep, a common overlap removal procedure has been created within the ATLAS experiment for most analyses.

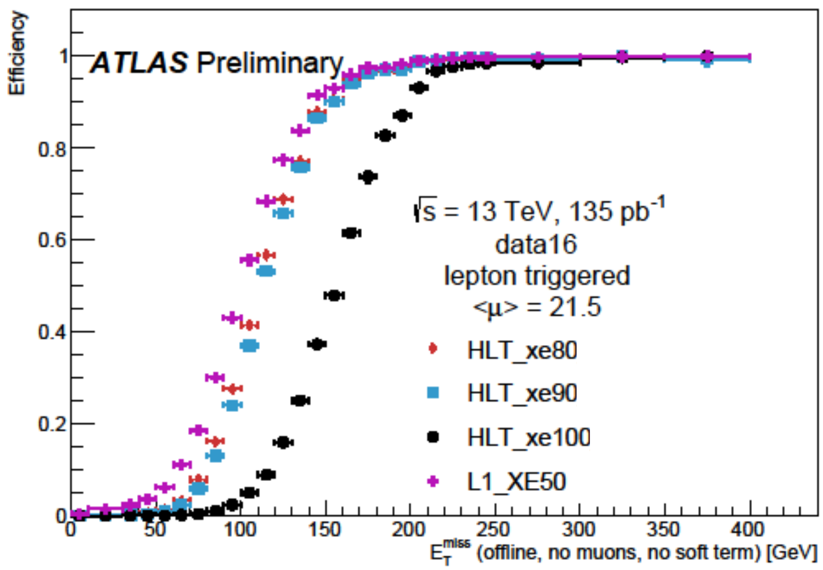


Figure 16: MET Trigger efficiency for multiple triggers, as a function of the offline E_T^{miss}without leptonic contributions or soft terms [43].

All of the reconstructed physics objects passing pre-selection are evaluated and removed if close to another object in the order outlined in Table 18. This prescription is applied after the pre-selection and before the following signal selection.

4.3 Signal Selection

In addition to the pre-selection requirements, the final analysis selection uses more restrictive requirements resulting in an improved signal-to-background ratio. The additional requirements are described in the following sections.

Step	Remove	Accept	Condition
1	τ	e	if $\Delta R(\tau, e) < 0.2$
2	τ	μ	if $\Delta R(\tau, \mu) < 0.2$
			*the τ is not removed if its $p_T > 50$ GeV
			and the μ is not a combined μ
3	e	combined μ	if ID track is shared between e and μ
4	calo-tagged μ	e	if ID track is shared
5	jet	e	if $\Delta R(\text{jet}, e) < 0.2)$
6	e	jet	if $\Delta R(\text{jet}, e) <min(0.4, 0.04 + 10\text{GeV}/p_T^e)$
7	jet	μ	if ΔR(jet,$\mu < 0.2$) or
			if the muon's ID track is associated with the
			jet and there are less than three ID tracks
			associated with the jet that have
			$p_T > 500$ MeV, or if $(p_T^{\mu}/p_T^{jet} > 0.5)$
			and the muon's p_T divided by the
			aforementioned jet associated tracks
			have a ratio less than 0.7
8	μ	jet	if $\Delta R(\text{jet}, \mu) <min(0.4, 0.04 + 10\text{GeV}/p_T^{\mu})$
9	jet	τ	if $\Delta R(\text{jet}, \tau) < 0.2$
10	fat jet	e	if $\Delta R(\text{fat jet}, e) < 1.2)$

Table 18: Overlap Removal condition table detailing the physics objects that are rejected and kept depending on certain conditions in a specific order.

4.3.1 Lepton Selection

The leptons used in this analysis are separated into two categories: VH-loose and WH-tight leptons. Leptons must be VH-loose during pre-selection and must pass the WH-tight selection to be considered a signal lepton.

VH-loose electrons

VH-loose electrons, used for orthogonality between the 0, 1 and 2 Lepton modes, are required to pass a loose likelihood identification detailed in Reference [49]. The momentum and geometrical criterion require good separation from backgrounds and require containment within the ID. Additional selections involving impact parameters are made to reduce non-prompt and pile-up tracks. These criteria are summarized below:

- $p_T^e > 7$ GeV
- $|\eta| < 2.47$
- $|\Delta z_0 \sin\theta| < 0.5$mm
- impact parameter significance $< 5\sigma$
- track isolation tuned to 99% efficiency for prompt electrons

A close geometrical match between the reconstructed electron and the online electron that the trigger identified, and an offline requirement on the p_T of the electron of $1.05 * p_T$ of the trigger threshold to ensure trigger efficiency measurements are applied in an internally consistent manner.

WH-tight electron

The WH-tight electron selection requires, in addition to VH-loose criteria, medium likelihood identification and a calorimeter-based isolation that is tighter than the previous isolation. This isolation corresponds to a 95% signal efficiency.

The 1 Lepton mode has non-negligible multijet backgrounds. To suppress the multijet constribution in the electron mode, which has larger contamination than the muon mode, a 30 GeV requirement is placed on the E_T^{miss}exclusively in the electron mode. No corresponding requirement is made in events with a muon.

VH-loose muon

VH-loose muons, like electrons, are used for orthogonality between the 0, 1 and 2 Lepton modes. They are required to satisfy at least loose muon quality. Once again they have a minimum momentum requirement, but do not have a geometrical component. There are also impact parameter considerations required in order to reject pile-up and cosmic muon backgrounds. The selection for VH-loose muons is:

- $p_T^{\mu} > 7$ GeV
- impact parameter singificance $< 3\sigma$
- $|\Delta z_0 \sin\theta| < 0.5$mm
- inner tracker isolation tuned to 99% efficiency for prompt μ

WH-tight muon

The WH-tight muon selection applies two additional criteria in addition to the VH-loose muon requirements. First, the muon p_T must be greater than 25 GeV. Second is a tighter track based isolation criteria that gives 95% signal efficiency and requires the muon to be within $|\eta| < 2.5$.

4.3.2 Jet Selection

There are two signal regions used in the 1 Lepton mode. One has two $R = 0.4$ anti-k_T jets present in the event while the other has three jets.

All jets used in this analysis are required to pass basic jet cleaning algorithms and have a minimum p_T of 20 GeV. The leading jet must have a p_T of at least 45 GeV. Pileup jet contributions are reduced by using a likelihood-based discriminant labeled the Jet Vertex Tagger (JVT.) The JVT takes information about the primary vertex, jet p_T and track p_T of jets whose total transverse momentum is between 20 and 60 GeV and within $|\eta| < 2.4$, and decides if it is a pile-up jet or not. This gives an average efficiency of 92 % for signal jets (equivalently pileup rejection rate) [66].

Depending on the location in η of the jet, it is classified as either a signal (< 2.5) or forward ($2.4 \leq |\eta| < 4.5$) jet. Only signal jets are used to reconstruct the Higgs boson's $b\bar{b}$ decay, meaning there must be at least two signal jets.

B-tagging

Since the final state of the signal includes a pair of b quarks, two of the signal jets are required to be b-tagged. Studies showed that the best sensitivity was achieved at a working point of 70% efficiency with the MV2c10 algorithm which is described in Section 2.3.4. The MV2c10 discriminant used has a charm rejection rate of 9.19, τ-tagged jet rejection rate of 38.5 and light flavored jet rejection rate of 335 [67]. Rejection rate is defined as the inverse of the number of truly non-b-quark initialized jets that are mistakenly tagged. A third b-tagged jet requirement was tested and shown to decrease sensitivity, dictating the exactly two b-tagged jet selection. The labeling of the jets as b_1 and b_2 is done in order of decreasing p_T.

Truth Tagging

The aforementioned b-tagging algorithm is not used on all samples. To increase the statistical precision of the selected events with light flavored $V + jets$, the jets are truth tagged. For these samples, all events with two or three jets are kept, but the event is weighted by the expected probability for a jet with a certain flavor to be identified as a b-tagged jet. Truth jets are reconstructed from stable, charged particles found in the stimulated event's true particle information using the anti-k_T algorithm with a maximum cone size of $R_{max} = 0.3$. This is done for the following samples: $V+cc, V+cl, V+ll$ and WW. All other samples, including the data, are b-tagged. The truth labeling process proceeds as follows:

1. b-jet if a weakly decaying b-hadron is found within $\Delta R < R_{max}$
2. c-jet if a weakly decaying c-hadron is found within $\Delta R < R_{max}$
3. τ-jet if a τ lepton is found within $\Delta R < R_{max}$
   - If there are more than one jet within the ΔR range, the closet jet is used.
4. All remaining jets are labeled as light jets

4.3.3 Additional Selections

Additionally, a control region is designed to help constrain the modeling of the W boson plus heavy flavor ($W + HF$) background processes. This region is designed to be primarily $W + HF$ background with negligible signal contamination. Control regions are orthogonal to signal regions, so none of the events used to constrain backgrounds are also used in determining the signal.

The $W + HF$ control region requires that $m_{bb} < 75$ GeV, reconstructed from the two b-tagged jets, to reduce signal contamination. Additionally, a selection is made on the reconstructed top mass, m_{top}. The top mass is calculated taken as the minimum invariant mass of the 4-vector of the W

boson (with the assumption that all of the E_T^{miss}in the event originated from its leptonic decay) and one of the signal jets. Selecting $m_{top} > 225$ GeV significantly reduces the $t\bar{t}$ and single t contributions allowing the single bin of the control region to constrain the $W + HF$ normalization. The full selection criteria is summarized in Table 19.

Variable	Requirement
Trigger	lowest unprescaled electron triggers
	lowest unprescaled E_T^{miss}triggers for muon events
Number of WH Tight Leptons	1
E_T^{miss}	$>$ 30 GeV in electron mode
Number of Signal Jets	2 or 3
Number of Forward Jets	0 or 1
Number of B-tagged Jets	2
m_{top}	$<$ 225 GeV
m_{bb}	$>$ 75 GeV

Table 19: Summary of selection criteria in WH analysis.

5 Multivariate Analysis

Multivariate analysis is a statistical technique that encodes information from multiple distributions simultaneously. The MVA used in this analysis enhances the signal purity and background rejection while maximizing the yield by weighting each event instead of rejecting events and thereby losing information that can potentially help distinguish signal from background.

The MVA's input variables are combined to create a single one-dimensional discriminant variable that represents how signal-like the event's characteristics are. The MVA is first trained to identify signal-like event characteristics on a subset of the available events in the signal and background MC samples. In order to not bias the result, the MVA is trained using a 2-fold cross validation training. This means that one training is done using the even (odd) numbered MC events and the quality tested with the odd (even) numbered MC events. This is done to ensure statistical independence between the training and testing samples.

An additional MVA was created for validation purposes. With VH designated as a background, this separate MVA is designed to separate the diboson $VZ \rightarrow Vb\bar{b}$ process as signal from the remaining processes as a cross-check. All available MC and data-driven background samples are used in building each MVA, with either the VH or VZ MC designated as signal.

The distribution of the MVA output discriminant is used in the final fit across all signal and control regions concurrently in both the 1 Lepton mode and the VZ cross-check.

5.1 Boosted Decision Tree

The specific MVA used in this analysis is a Boosted Decision Tree (BDT), similar to the MVA used in Run 1 [68]. Figure 17 is a simplified example of a binary tree structured classifier, otherwise known as a decision tree. The branching requirements, c_i, are not known a priori and must be determined before using the tree. This determination, called training, starts at the root node with a binary decision determined by requirements placed on the input variable (xi, xj and xk in the mentioned figure) that gives the best separation between signal and background. This is repeated for each node until the separation has reached a pre-specificed minimum number of events [69].

The criteria by which the requirements on the input variables are determined can be assessed by multiple types of selection criteria. Three examples are shown below in Equation 31: Gini Index, Statistical Significance, and Misclassification error respectively [69]. The p stands for the purity of the sample, zero is for fully background, one for fully signal and 0.5 for perfectly mixed. If the event is categorized as background, it is given a value of -1, and if it's signal it is given a value of 1.

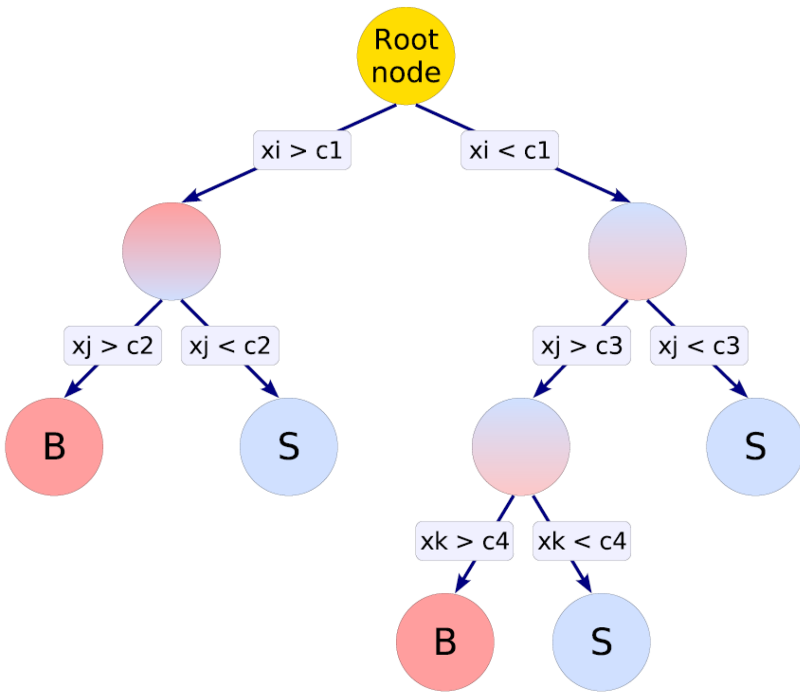


Figure 17: A simplified example of a decision tree algorithm [69].

$$\begin{gathered} p(1 - p) \\ \frac{S}{\sqrt{S+B}} \\ 1 - \mathtt{max}(p, 1 - p) \end{gathered} \tag{31}$$

Once the last node is reached, that 'leaf' node is classified by the majority composition of the events. If the composition of the events on the leaf node is completely pure, then it's likely that the decision tree is overtrained. Overtraining is when noise is being picked up instead of the underlying function because there are too few degrees of freedom with too many paramters and too few data points [69]. The impact of overtraining can be observed by comparing the training and testing samples, which if inconsistent, indicates the tree has picked up on the statistical fluctuations (noise) and needs to be reduced (also called pruned). Pruning is done to remove statistically insignificant nodes to eliminate the overtraining of the tree.

Unless the leaf node has a $p = 0$ or $p = 1$ purity, there are events which are mis-identified. The two possible misidentifications are either background incorrectly identified as signal or vice-versa. Errors made in the decision tree are reweighted in following trees in an iterative process called boosting. The first tree will always be the best possible tree. The second tree corrects the errors of the first tree by giving higher weights to events that had been classified incorrectly and lower weights to the events that had been classified correctly. A third tree reweights the events to correct errors in the first two trees, and so on. The reweighting is done according to Equation 32, where the frequency of the error is defined as f^i_{err} = misclassified events/all events, and the classifier, $C^i(x)$ is of the i-th order tree in the boosting model. This reweighting process is called boosting [69].

$$y(x) = \sum_{i}^{N_{classifier}} \log\left(\frac{1 - f^i_{err}}{f^i_{err}}\right) C^i(x) \tag{32}$$

Combining the trees together gives a 'forest' by averaging different trees together. This helps rectify BDTs' weaknesses at recognizing linear correlations and also helps to remove any statistical fluctuations it may have identified as a signal function [69]. The combination of reweighted BDTs results is the final BDT used to search for the VH or VZ signal.

It is superior to stop adding variables to the tree rather than remove variables later for BDTs [69]. With no obvious signs of overtraining in Figure 18, the BDT is considered to be well behaved, to be well understood and most importantly, to perform well. It shows the separation of signal from background in both the training and testing samples with the signal clustered around a BDT value of 0.4.

To quantify how well a BDT is performing, a goodness of fit Kolmogorov-Smirnov (KS) test is used to evaluate the maximum distance between the trained sample and test sample. Smaller distances indicate that the trained sample represents the test sample well. This maximum distance is defined

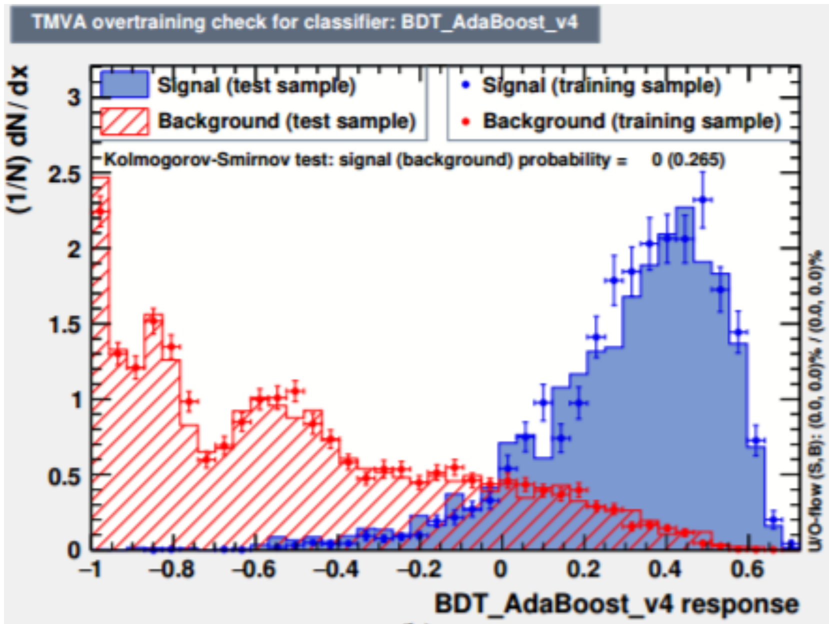


Figure 18: Performance of the BDT in the 2 tags 2 jets region. The background is in red, the signal in blue. The training samples are dots while the test samples are histograms.

by the KS statistic (D_n in Equation 33) which is the supremum[3] of the set of distances between sets of the distribution functions, $F(x)$ [70].

$$D_n = \sup_x |F_n(x) - F(x)| \tag{33}$$

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I_{[-\infty,x]}(X_i) \tag{34}$$

The input variables for the BDT are listed below. The list was started with the variables m_{bb} and $\Delta R(j,j)$, as suggested from Run 1's BDT [68]. Then each candidate variable was added individually and the tree retrained. The variable that gave the best improvement in significance was added to the list. This process was repeated until all of the variables had been studied and no further significant improvement was seen when adding any additional variable. This constituted the final, optimized MVA.

[3]The supremum is also known as the least upper bound.

The following list shows the input variables used for both the two and three jet selections:

- p_T^W
- E_T^{miss}
- $p_T^{b_1}$
- $p_T^{b_2}$
- m_{bb}
- $\Delta R(b_1, b_2)$
- $\Delta\phi(W, bb)$
- $\min[\Delta\phi(\ell, b)]$
- m_T^W
- m_{top}
- $|\Delta Y(W, H)|$

The following two input variables are exclusively used in 3-jet events:

- p_T^j
- m_{bbj}

The W boson p_T is the vector sum of the $E_T^{miss}$and p_T^ℓ in MeV. The first b-tagged jet is denoted as b_1 and the second as b_2. The third jet, if present, is simply j. The mass of the dijet system composed of the two b-jets is m_{bb}, and if there is a third jet, the trijet system mass is m_{bbj}. When the two b-jets' 4-vectors are summed, they represent the Higgs boson, and are used in the rapidity distance calculation between the Higgs boson and W boson to give $|\Delta Y(W, H)|$. The transverse mass of the W boson is calculated using Equation 35. The top mass is that of the reconstructed leptonically decaying top, selecting from the two leading jets, the jet that minimizes m_{top} through Equation 36.

$$m_T^W = \sqrt{2p_T^\ell E_T^{miss}(1 - \cos(\Delta\phi(\ell, E_T^{miss})))} \tag{35}$$

$$\begin{gathered} p_Z^\nu = \frac{1}{2(p_T^\ell)^2}\left[p_z^\ell X^2 \pm E_\ell\sqrt{X^2 - 4(p_T^\ell)^2(E_T^{miss})^2}\right] \\ X = m_W^2 + 2p_x^\ell E_x^{miss} + 2p_y^\ell E_y^{miss} \end{gathered} \tag{36}$$

The signal and backgrounds' input variable distributions for the two jet final state are shown in Figure 20. The correlation between the input variables for both the signal and background samples can be seen in Figure 19.

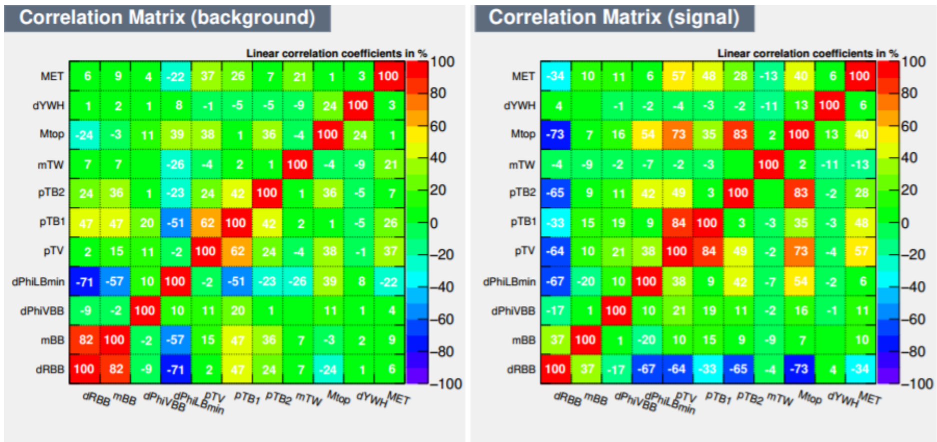


Figure 19: The correlation matrices for the background (left) and signal (right) BDT input variables.

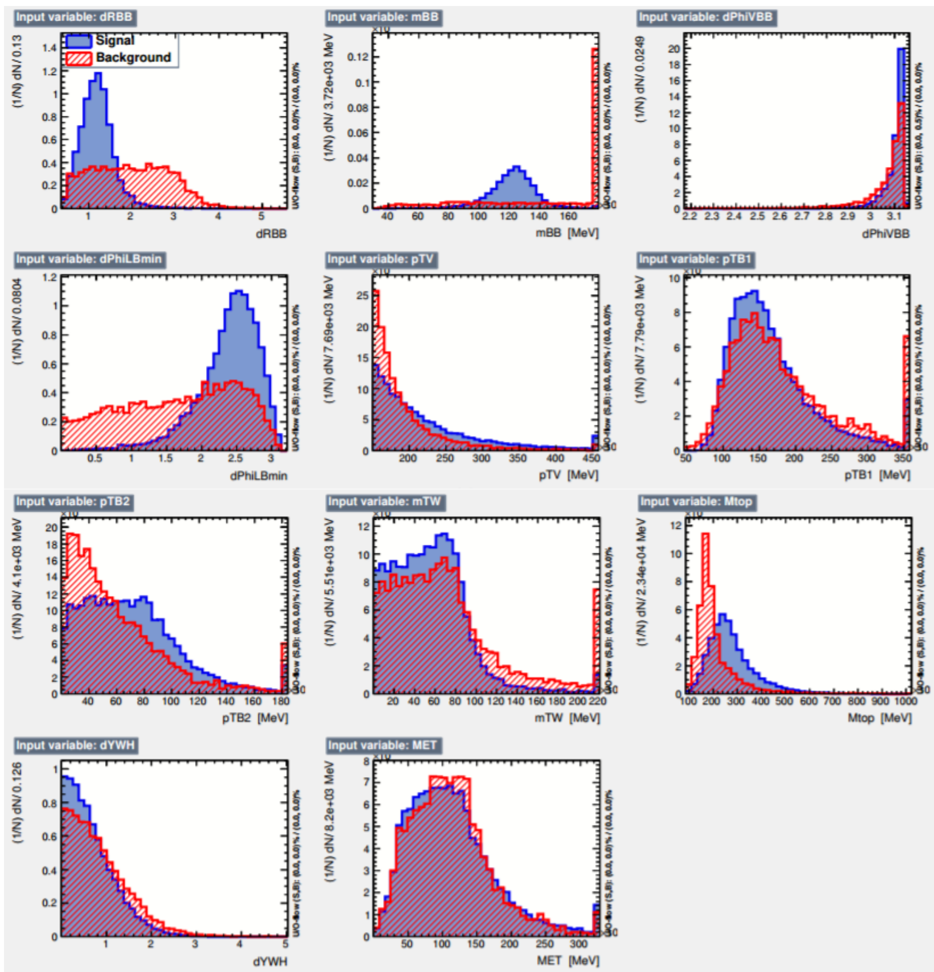


Figure 20: BDT input variable distributions for signal (blue) and backgrounds (red) for the 2 tags 2 jets region.

6 Fitting

The analysis utilizes a maximum profile likelihood estimator to extract normalizations of the backgrounds and the strength of the Higgs boson signal from the data. The statistical fitting procedure which incorporates the Poisson nature of statistical fluctuations in event yields and systematic uncertainties as nuisance parameters is described in Section 6.1.

The combination of all sources of uncertainties: statistical, theoretical and systematic, affect the confidence level on the central value measurement. Uncertainties quantify the confidence analyzers have in their prediction, or in the data recorded, or in their reconstruction of physics objects. The uncertainties used in this analysis are introduced in Section 6.2.

6.1 Maximum Profile Likelihood

The Roostats framework [71] is used to perform the statistical analysis. The fitting is done as a binned likelihood function ($\mathcal{L}$). The likelihood function is constructed as a Poisson probability summed over the bins of the input distributions. All input distributions are fitted simultaneously. The background yields and expected signal are fitted to the data events. The fit takes into account the effects of the systematic uncertainties by the way of nuisance parameters (NP) [72].

The NPs are represented as θ in the profile likelihood ratio, Λ_μ, in Equation 37.

$$\Lambda_\mu = \frac{\mathcal{L}(\mu, \hat{\hat{\theta}}_\mu)}{\mathcal{L}(\hat{\mu}, \hat{\theta})} \tag{37}$$

The ratio is dependent on the signal strength μ. The signal strength can be understood as the measured signal contribution divided by the expected signal contribution as in Equation 38.

$$\mu = \frac{\text{signal}_{\text{measured}}}{\text{signal}_{\text{SM}}} \tag{38}$$

The profile likelihood ratio is also dependent on the conditional maximum likelihood estimate (CMLE) $\hat{\hat{\theta}}_\mu$ which is the value of θ that maximizes the likelihood function when the signal strength μ is fixed. Since there are many

NPs the CMLE is a multidimensional function. Because the CMLE depends on μ, the likelihood is no longer dependent on θ [72].

The fitting procedure defines q_μ as the test statistic (see Equation 39). The maximum profile likelihood estimates the parameters that maximize the likelihood function and equivalently minimize the test statistic.

$$q_\mu = -2\ln\Lambda_\mu \tag{39}$$

The best fit is defined in reference to $\sigma \times BR(H \to b\bar{b})$ where the cross-section is the relevant SM predicted production process, σ_{WH} or σ_{VH} and $BR(H \to b\bar{b})$ is the branching ratio of the SM's prediction for the higgs boson's decay to $b\bar{b}$ and the relevant vector boson's leptonic decay. The branching ratio is 58% for $H \to b\bar{b}$ and 21% for $W \to e\nu$ and $W \to \mu\nu$ combined. Signal cross-contamination from the ZH channels is negligible.

The systematic uncertainties are discussed in detail in Section 6.2. They are utilized in the fit to provide constraints within which the fit is allowed to vary. All uncertainties but the normalizations are constrained by Gaussain probability density functions (PDFs). The normalizations are constrained by log-normal PDFs to prevent them from going negative.

The statistically limited nature of the backgrounds' MC sample size is represented as a single NP per bin using the Beeston-Barlow technique [70]. Additionally, the yields of signal and background events in each bin are dependent on μ and θ. The likelihood, $\mathcal{L}(\mu, \theta)$, incorporates every NP by adding each as a penalty term. As each NP deviates from its nominal value the penalty term decreases the overall test statistic.

6.1.1 VH BDT Analysis

Four regions are used in the fit. They are differentiated by the number of jets present, 2 or 3, and the $W + HF$ signal versus control regions. The signal regions (SR) use the binned BDT distribution as input and the control regions (CR) only use the total yield as input. While the multijet background is calculated separately for the electron and muon events, the events are combined for the fitting procedure. The results are discussed in Section 7.1.

The data has sufficient statistical power to constrain one of the largest backgrounds, $W + HF$. These regions only contain a single bin so the fitted yield is able to provide the predominant background's, $W + HF$, normalizations in the 2 jets and 3 jets regions independently. These normalizations can be fractionally quite large but are found to be consistent with unity

Process	Scale Factor
W+HF 2-jet	1.21 ± 0.14
W+HF 3-jet	1.36 ± 0.14
Z+HF 2-jet	1.30 ± 0.11
Z+HF 3-jet	1.20 ± 0.09
$t\bar{t}$	0.90 ± 0.08

Table 20: Scale factors obtained from the global BDT fit to the 13 TeV data for background normalizations. The error includes statistical and systematic uncertainties.

within two standard deviations given the uncertainties on each. The largest contribution to these uncertainties are theoretical uncertainties on the process' cross sections. Table 20 shows the resultant scale factors and their uncertainties. Also included in this table is the normalization factor for the $t\bar{t}$ background. The $t\bar{t}$ background is primarily composed of two types of events: dileptonic decays where at least one lepton (often a τ lepton) is undetected, and a semileptonic decay ($W \rightarrow q\bar{q}'b$) with at least one undetected light or c-jet.

6.1.2 VH Dijet Mass Analysis

The m_{bb} variables is used as the main input for the dijet-mass cross check analysis instead of the BDT. The $W+HF$ signal and control regions are combined, leaving the four fitted regions differing in the number of jets present, 2 or 3, and the W boson's transverse momentum ($150 < p_T^W < 200$ GeV or $p_T^W > 200$ GeV). The results are discussed in Section 7.1.1.

6.1.3 VZ BDT Analysis

The Diboson cross-check analysis (VZ) is fitted quite similarly to the nominal VH analysis. The VZ signal searches for a Z boson that has decayed to a pair of b quarks instead of the Higgs boson. The Z boson is produced in association with another vector boson, which in the 1 Lepton mode is predominantly W. While WZ and ZZ are targeted, the WW process can also create the same final state. However the WW process is treated as a background and is constrained within its uncertainty.

A different BDT from the nominal VH BDT is used. The VZ specific

BDT is trained to search for a VZ signal and counts the VH MC as a background with a 50% uncertainty on the production cross section. Because the VZ process has a similar signature as the VH result, it provides important validation.

The major differences between the VZ and VH signals are as follows:

- The cross section of VZ is about nine times larger than VH at the SM Higgs' mass.
- The m_{bb} distribution peaks around the Z boson's mass of 91 GeV instead of the Higgs boson's mass of 125 GeV.
- The p_T^{bb} spectrum is softer for VZ.

As a result of these differences, their BDTs are able to provide good separation between the two signals with only a weak correlation of less than 1% in the results.

Similar to the nominal VH analysis, the VZ BDT is fit with a signal strength, μ in Equation 38, relative to its SM prediction with the same binned likelihood function. The same fitting regions are used in the VZ BDT analysis as in the VH analysis. The results are discussed in Section 7.1.2.

6.1.4 VH Combination

The three independent production and decay modes of VH are combined in the same analysis and cross-checks as described by the previous three subsections. The most significant difference is the inclusion of additional fitting regions. All of the regions used are outlined in Table 21.

The three lepton modes can also be decorrelated so as to find the best fit in the individual mode's signal regions. This has a different result compared to running the analysis on separate modes because the backgrounds from the other modes can constrain a selected mode's backgrounds when decorrelated but not when not present in the likelihood fit.

		Categories			
		75 GeV $< p_T^V <$ 150 GeV		$p_T^V >$ 150 GeV	
Mode	SR/CR	2jets	3 jets	2 jets	3 jets
0 lepton	SR	-	-	BDT	BDT
1 Lepton	SR	-	-	BDT	BDT
2 lepton	SR	BDT	BDT	BDT	BDT
1 Lepton	$W + HF$ CR	-	-	Yield	Yield
2 lepton	$e\mu$ CR	m_{bb}	m_{bb}	Yield	m_{bb}

Table 21: Distribution variables used in the VH combination likelihood fit for the BDT analysis.

6.1.5 VH BDT Analysis Combined with Run 1's Analysis

A final statistical analysis was done to combine the VH BDT analyses from both Run 1 and Run 2 datasets. Run 1 data was collected at center of mass energies of 7 and 8 TeV, in comparison to Run 2's 13 TeV. The Run 1 analysis is kept the same. The correlation and compatibility of the Run 2 analysis with the Run 1 analysis were investigated. The experimental systematic uncertainties between the two analyses were studied exclusively for the dominant uncertainties.

The b-tagging uncertainties are assumed to have a negligible correlation due to the improvements in b-tagging from both the inclusion of a new detector and updated b-tagging algorithm. Studies showed that the affect of correlating the leading systematic results in less than a 5% change in the signal strength and a negligible impact on its uncertainty. Different correlation schemes for jet energy scale uncertainties were tested, but only the b-jet specific energy scale uncertainty is correlated across the different center-of-mass energies' datasets.

There are many potential correlations in the modeling of the background processes, but the impact is limited because each systematic uncertainty is merely a fraction of the total uncertainty. It is difficult to directly compare the correlations due to changes in center-of-mass energies, MC generators, object and event selection and the software that does simulation, reconstruction and analysis. The correlations' maximum potential effect is evaluated by testing a χ^2 combination of the two signal strengths obtained from the datasets and varied by linear correlation coefficients meant to represent the degree of correlation between the two measurements. These coefficients varied the $t\bar{t}$, $Z + HF$, and $W + HF$ normalizations and systematic shape variations across the two datasets from uncorrelated to fully correlated. The impact was found to be less than 1% of the signal strength, and less than 4% for the signal strength uncertainty. Because the correlations present such a small impact they are assumed to be zero in this analysis. The only exceptions are for the theoretical uncertainties on the Higgs boson signal in the cross-section, branching ratio and p_T^V dependent NLO EW correction which are assumed to be 100% correlated.

6.2 Systematic Uncertainties

There are many sources of error to take into consideration. The results of the analysis are quoted with separate statistical error from the data collected and systematic uncertainties. The full break down of all of the errors in this analysis can be found in Table 22. Figure 6.2 shows the uncertainties ranked in order of most important to least as well as their impact as NPs.

This section will first discuss the theoretical uncertainties from MC background samples, then the data-driven multijet sample and the MC signal sample before discussing the experimental sources of uncertainty.

6.2.1 Theoretical Uncertainties

Backgrounds that are estimated via MC samples have similar sources of systematic uncertainty. These modeling uncertainties can generally be placed in one of the following three categories: normalization, acceptance differences that cause different normalizations in different regions of the same background, and shape differences. The uncertainties can be derived in one of two ways: either a comparison of the nominal sample with a secondary sample from a different generator done with RIVET or by comparing to data in control regions [73]. These comparisons can identify possible data/MC discrepancies or MC mismodelings. Alternative samples for all MC samples are considered. Table 23 shows the nominal and alternate generators used for each sample. The following discussion is split by the physics process.

Acceptance uncertainties are calculated using Equation 40 to provide internally consistent cross checks.

$$\frac{Acceptance[Category_A(nominalMC)]/Acceptance[Category_A(alternativeMC)]}{Acceptance[Category_B(nominalMC)]/Acceptance[Category_B(alternativeMC))]} \quad (40)$$

Nuisance Parameter	Error
Total	$+0.690/-0.604$
Data Statistics	$+0.400/-0.386$
Full Systematics	$+0.562/-0.465$
Floating normalizations	$+0.087/-0.077$
All normalizations	$+0.126/-0.110$
All but normalizations	$+0.531/-0.110$
Jets MET	$+0.152/-0.429$
b-tagging	$+0.162/-0.138$
Leptons	$+0.010/-0.005$
Luminosity	$+0.057/-0.021$
Diboson	$+0.090/-0.051$
$Z + jets$	$+0.004/-0.005$
$W + jets$	$+0.178/-0.170$
$t\bar{t}$ modeling	$+0.074/-0.048$
Single t modeling	$+0.179/-0.152$
multijet modeling	$+0.052/-0.051$
signal systematics	$+0.311/-0.117$
MC statistiscs	$+0.255/-0.253$

Table 22: A break down of the errors on the fitted data. The errors listed are one standard deviation from the central value.

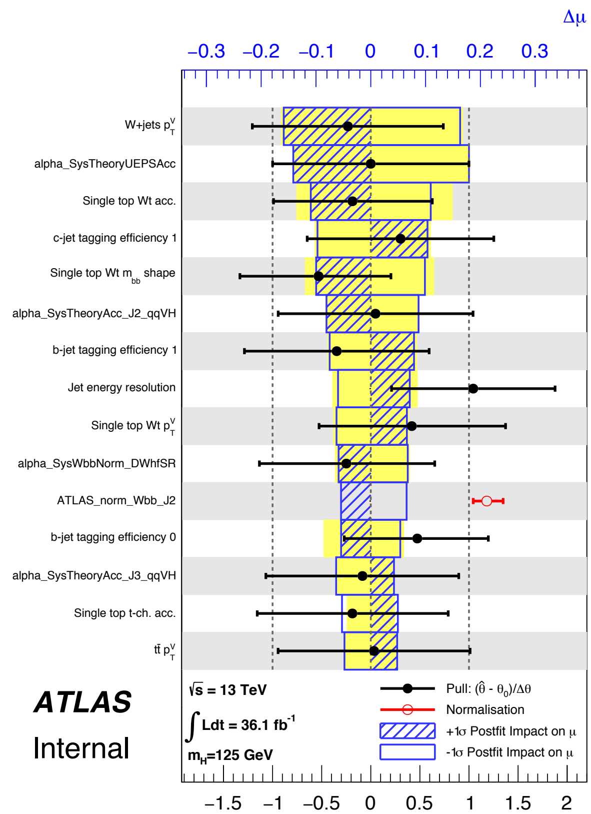


Figure 21: Nuisance parameters impact ranked in order of importance.

Process	Nominal	Systematic
$V + jets$	Sherpa 2.2.1	MADGRAPH 5 with PYTHIA 8
$t\bar{t}$	POWHEG + PYTHIA 8	
non-all-had $t\bar{t}$		POWHEG + PYTHIA 8
dilepton $t\bar{t}$		POWHEG + PYTHIA 8
non-all-had $t\bar{t}$ radUp		POWHEG + PYTHIA 8
non-all-had $t\bar{t}$ radDo		POWHEG + PYTHIA 8
non-all-had $t\bar{t}$		POWHEG + HERWIG 7
dilepton $t\bar{t}$		POWHEG + HERWIG 7
non-all-had $t\bar{t}$		aMC@NLO+PYTHIA 8.2
dilepton $t\bar{t}$		aMC@NLO+PYTHIA 8.2
Single top Wt incl t	POWHEG+PYTHIA 6	POWHEG+HERWIG++
Single top Wt incl $\bar{t}$	POWHEG+PYTHIA 6	POWHEG+HERWIG++
Single top DS Wt incl t		POWHEG+PYTHIA 6
Single top DS Wt incl $\bar{t}$		POWHEG + PYTHIA 6
Single top Wt incl t radHigh		POWHEG + PYTHIA 6
Single top Wt incl $\bar{t}$ radHigh		POWHEG + PYTHIA 6
Single top Wt incl t radLow		POWHEG + PYTHIA 6
Single top Wt incl $\bar{t}$ radLow		POWHEG + PYTHIA 6
Single top Wt incl $t + \bar{t}$		MADGRAPH5_aMC@NLO+HERWIG++
Single top t-chan t	POWHEG+PYTHIA 6	POWHEG+HERWIG++
Single top t-chan $\bar{t}$	POWHEG+PYTHIA 6	POWHEG+HERWIG++
Single top t-chan t radLow		POWHEG+PYTHIA 6
Single top t-chan $\bar{t}$ radLow		POWHEG+PYTHIA 6
Single top t-chan t radHigh		POWHEG+PYTHIA 6
Single top t-chan $\bar{t}$ radHigh		POWHEG+PYTHIA 6
Single top t-chan $t + l\bar{t}$		MADGRAPH5_aMC@NLO+HERWIG++
Single top s-chan t	POWHEG+PYTHIA 6	
Single top s-chan $\bar{t}$	POWHEG+PYTHIA 6	
Single top s-chan t radHigh		POWHEG+PYTHIA 6
Single top s-chan $\bar{t}$ radHigh		POWHEG+PYTHIA 6
Single top s-chan t radLow		POWHEG+PYTHIA 6
Single top s-chan $\bar{t}$ radLow		POWHEG+PYTHIA 6
WplvWmqq	SHERPA 2.1	
WpqqWmlv	SHERPA 2.1	
WlvWqq		POWHEG+PYTHIA
Wlv Wqq		POWHEG+HERWIG++
WlvZqq	SHERPA 2.2.1	POWHEG+PYTHIA
		POWHEG+HERWIG++
WqqZll	SHERPA 2.2.1	POWHEG+PYTHIA
		POWHEG+HERWIG++
WqqZvv	SHERPA 2.2.1	POWHEG+PYTHIA
		POWHEG+HERWIG++
ZqqZll	SHERPA 2.2.1	POWHEG+PYTHIA
		POWHEG+HERWIG++
ZqqZvv	SHERPA 2.2.1	POWHEG+PYTHIA
		POWHEG+HERWIG++

Table 23: Nominal and alternate samples used to derive uncertainties for MC backgrounds. Abbreviations as follows: "incl" for inclusive, "chan" for channel, Wp for positive W boson and Wm negative W boson.

$V + jets$ **Backgrounds**

The acceptance uncertainties for $V + jets$ are calculated for the following comparison regions using Equation 40:

- events with 2 jets and those with 3 jets
- $W + HF$ control region versus $W + HF$ signal region (for $W + HF$ events)

This helps to understand the correlation between the different regions of phase space. The differences are parameterized as NPs.

Additionally the nominal $V + jets$ samples are compared to the same generator samples with different factorization and renormalization scales used in their generation. The $Z + jets$ background in the analysis mode is negligible and so no $Z + jets$ specific uncertainties are derived here. The final normalization systematic uncertainty for $V + jets$ is obtained by summing the uncertainties in quadrature from the following variants:

- envelope of the effect of varying factorization and renormalization scales
- sum in quadrature of half of the variation obtained from samples with different resummantation and matrix element-parton shower merging
- reweighting to alternative PDF compared to nominal
- differences between SHERPA and MADGRAPH samples

The shape uncertainty must also be considered. This is done for the two most important physical variables used in the BDT: $m_{b\bar{b}}$ and p_T^V. The $W + jets$ uncertainties are derived from shape comparisons of SHERPA samples with different factorization and renormalization scales, and with the MADGRAPH 5 sample because it is difficult to isolate a non-phase space dependent high purity sample. This isolation is less difficult for $Z + jets$ samples so their uncertainties are derived from comparing the MC to data.

The $W + HF$ background has multiple contributing processes. Their relative composition is quantified in Table 24 and the relative acceptances quantified by nuisance parameters are detailed in Table 25.

Process	2 jets		3 jets	
	W+HF CR	W+HF SR	W+HF CR	W+HF SR
Wcc	0.04	0.09	0.05	0.07
Wbl	0.02	0.06	0.02	0.04
Wbc	0.05	0.12	0.05	0.13
Wbb	0.88	0.73	0.88	0.76

Table 24: The fractional contribution of sub-processes to the $W + HF$ yields in the analysis regions.

Category	Nuisance Parameter	Uncertainty	Applied to
$W + bc/W + bb$	SysWbcWbbRatio	13.4%	$W + bc$ events
$W + bl/W + bb$	SysWblWbbRatio	4.8%	$W + bl$ events
$W + cc/W + bb$	SysWccWbbRatio	1.2%	$W + cc$ events

Table 25: Uncertainties on the relative acceptance of subprocesses in the $W + HF$ contribution.

Top Pair Backgrounds

There are many alternative samples used to derive systematic uncertainties on the $t\bar{t}$ sample. They are listed in Table 23 and separated to study three variations: parton shower uncertainty, matrix element/generator uncertainty and radiation uncertainties. They are filtered at generator level using truth information to select events where "non-all-had" has at least one W boson decaying leptonically and the "dilepton" filter selects events where both W bosons decay leptonically. The radUp and radDo stand for high and low radiation respectively. There are samples which are interfaced with HERWIG 7 instead of PYTHIA 8 for parton shower, hadronization, underlying event and multiple particle interactions in order to study the difference in the parton shower model. Changing the generator to MADGRAPH5_aMC@NLO but keeping PYTHIA 8.2 for the remaining processing isolates any systematic effects due to the hard scattering generation. Finally, the low and high radiation samples alter the renormalization and factorization scale by a factor of two in opposite directions.

The systematic uncertainties are derived from the variation in acceptance for different analysis categories using Equation 40 while making sure the alternative MC samples are all normalized to the same NNLO cross section. This helps to control the correlations between different phase spaces via NPs.

Systematic	CR 2 jets	SR 2 jets	CR 3 jets	SR 3 jets
norm_ttbar	floating normalization			
SysttbarNorm_J2	0.8%	0.8%	-	-
SysttbarNorm_DWhfCR_L1	0.9%	-	0.9%	-

Table 26: Systematic uncertainties for the top pair production samples based on the 2 jet regions and the $W + HF$ control regions.

The NPs are parametrized with a Gaussian PDF which has the uncertainties quantified by acceptance uncertainties. The final uncertainties are given by a sum in quadrature of the uncertainties, taking into account the difference in yield due to the altered radiation samples.

Shape uncertainties are obtained by normalizing all of the variations to the same area and comparing the shapes for important physical variables in the BDT, p_T^V and $m_{b\bar{b}}$, using a linear fit for the former and exponential fits for the latter. The largest functional form fitted through the variation, different generator, is symmetrized and used. Table 26 summarizes the $t\bar{t}$ normalization systematics' effects.

Single Top Backgrounds

The single top background contains three different production mechanisms differing by the location of the vector boson in the Feynman diagram: s-channel, t-channel and Wt-channel. Each process is produced separately as well as a second sample for each corresponding anti-top production. They are generated for a quark mass of 172.5 GeV and with cross sections obtained from NNLO calculations. For each nominal sample the generator POWHEG is used, and PYTHIA 6 is used for the parton shower, underlying event and multiple parton interactions with PHOTOS for modeling the QED final-state radiation and TAUOLA for τ decay. Systematic samples are produced while varying the generator (MADGRAPH5_aMC@NLO) and the reconstruction algorithm is exchanged with HERWIG++ to obtain alternative matrix element schemes and parton showers. Additionally high and low radiation samples are generated for each channel with the renormalization and factorization factors doubled or halved respectively. The Wt process has an additional systematic sample which alters the matrix element scheme to use the 'diagram subtraction' (DS) instead of 'diagram removal' giving an estimate of single top interfence effects.

The overall normalization uncertainties account for renormalization and factorization scale variations, α_s and errors on the PDFs. They are summed

Production	Uncertainty	Value	Nuisance Parameter
s-channel	overall normalization	0.2%	stopsNorm
t-channel	overall normalization	1.2%	stoptNorm
t-channel	2 jet region acceptance	7.6%	stoptAcc
t-channel	3 jet region acceptance	20%	stoptAcc (correlated with above two)
t-channel	p_T^V shape	shape	StoptPTV
t-channel	m_{bb} shape	shape	StoptMBB
Wt channel	p_T^V shape	shape	StopWtPTV
Wt channel	m_{bb} shape	shape	StopWtMBB

Table 27: Normalization, acceptance and shape uncertainties for the single top processes and their associated nuisance parameter.

in quadrature to obtained uncertainties of 4.4% (t-channel), 4.6% (s-channel) and 6.2% (Wt-channel). The normalization uncertainties were obtained in both 2 and 3 jets regions. Because the uncertainties vary in the same direction the uncertainties are considered correlated.

Shape uncertainties were derived once again for the two most important physical variables in the BDT, m_{bb} and p_T^V, by comparing the acceptances and shapes of the nominal to alternative samples with linear fits. The final numbers were obtained by summing all variations in quadrature. The normalization and shape uncertainties are listed in Table 27 with the corresponding NPs.

Diboson Backrounds

Diboson backgrounds consist of two vector bosons decaying both hadronically and leptonically. While contributions exist from WW, WZ and ZZ events, the signal-like final state of $Z \to b\bar{b} + W \to \ell\nu$ is a significant contribution. Additional misidentification errors can cause processes like $W \to \ell\nu + W \to q\bar{q}$ to also contribute via a mistagged jet or a missed lepton for example.

The three diboson processes are generated and reconstructed with SHERPA 2.2.1. Once again, alternative samples are created with a different generator (POWHEG) and reconstruction algorithms (PYTHIA + HERWIG++) for comparison purposes. They are listed by process in Table 23. It is important that separate samples are created for the WW process depending on the sign of the W boson because there is an asymmetry in the W boson's cross section dependent on its charge that must be configured independently for SHERPA 2.2.1 samples. No low or high radiation samples were created because the SHERPA 2.2.1 samples were able to be reweighted to vary the

Nuisance Parameter	2 jets	3 jets
SysVZ_UEPS_Acc	0.2%	0.2%
SysVZ_UEPS_32JR	-	1.0%
SysVZ_UEPS_VPT	shape only	
SysVZ_UEPS_MBB	shape only	
SysVZ_QCDscale_J2	0.3%	0.3%
SysVZ_QCDscale_J3	1.3%	1.7%
SysVZ_QCDscale_JVeto	0.2%	0.2%
SysVZ_QCDscale_VPT	shape only	
SysVZ_QCDscale_mbb	shape only	

Table 28: VZ background uncertainties due to altering the renormalization, factorization and QCD scales.

renormalization and factorization scales.

Cross section uncertainties for each process are obtained by comparing the SHERPA and POWHEG samples as well as varying the renormalization and factorization scales. The resultant uncertainties are applied exclusively to their respective process: 25% (WW), 26% (WZ) and 20% (ZZ).

Shape uncertainties are obtained from the same dominant kinematic distributions in the BDT: m_{bb} and p_T^V with the distributions normalized to a unit area. The largest variation is taken as the shape systematic, in this case the PS comparison. This is done for each process except WW because of its extremely small contribution in the signal regions. The p_T^V is able to be described by a functional form but its shape uncertainty is within the MC statistical uncertainty and is therefore not applied. On the other hand, the m_{bb} shape does not have a valid simple functional form, therefore the shape uncertainty used in the analysis is parameterized via a binned histogram after smoothing.

The uncertainties originating from the high and low radiation rescaling can be seen in Table 28 as three independent NPs: overall acceptance, relative migration between 2 and 3 jets regions and vetoing events with 4 jets. Uncertainties originating from altering the QCD scale can be found in the same table with "QCDscale" in the NP name.

Multijet Backgrounds

The multijet background is a data-driven background. As such, the sources of systematic uncertainties are very different than backgrounds obtained from MC. Because the multijet background is estimated separately for events with electrons and events with muons, the two are treated as uncorrelated. The template can be affected in two ways: altering the shape of the template and altering its normalization.

The shape uncertainties considered are listed below. They are considered individually and created by changing the definition of the multijet control region and comparing to the nominal region using linear fits.

- Changing the number of b-tagged jets in the multijet control region (2 vs 1.)
- Including more stringent isolation requirements.
- Using a different electron trigger, thereby altering isolation requirements of events at trigger level.
- Removing the extrapolation factors based on p_T^{ℓ}, η, and if present, E_T^{miss}.
- Varying normalization of the dominant backgrounds in the multijet control region before subtraction.

The next and final sources of systematic uncertainty for the multijet background listed below alter the normalization. These variations are added in quadrature to obtain the quantitative differences.

- Different normalization from a fit with the E_T^{miss}distribution instead of the m_T^W distribution.
- Including events with an electron and $E_T^{miss}< 30$ GeV, which significantly increases the number of events.

Signal

The signal also has uncertainties associated with the MC used to produce it beyond the fact it was unverified with an experiment previously.

Since this process has to this point been theoretical, the samples are all normalized to the current best theoretical predictions done at NNLO in QCD and NLO in EW. This takes into account photon-induced contributions on the order of 5% for WH and 1% for ZH. Unaccounted higher order EW effects are accounted for with the following equation:

$$\Delta_{EW} = max\{1\%, \delta_{EW}^2, \Delta_\gamma\} \tag{41}$$

Process	Nominal Sample	Alternative sample
	qq-initiated	
$qq \to W^+H \to \ell^+\nu b\bar{b}$	POWHEG+MiNLO+PYTHIA 8	POWHEG+MiNLO+HERWIG 7
$qq \to W^-H \to \ell^-\nu b\bar{b}$	POWHEG+MiNLO+PYTHIA 8	POWHEG+MiNLO+HERWIG 7
$qq \to WH \to \ell\nu b\bar{b}$		MADGRAPH5_aMC@NLO+PYTHIA 8
$qq \to ZH \to \ell\ell b\bar{b}$	POWHEG+MiNLO+PYTHIA 8	MADGRAPH5_aMC@NLO+PYTHIA 8
$qq \to ZH \to \ell\ell b\bar{b}$		POWHEG+MiNLO+HERWIG 7
$qq \to ZH \to \nu\nu b\bar{b}$	POWHEG+MiNLO+PYTHIA 8	MADGRAPH5_aMC@NLO+PYTHIA 8
$qq \to ZH \to \nu\nu b\bar{b}$		POWHEG+MiNLO+HERWIG 7
	gg-initiated	
$gg \to ZH \to \ell\ell b\bar{b}$	POWHEG+PYTHIA 8	
$gg \to ZH \to \nu\nu b\bar{b}$	POWHEG+PYTHIA 8	

Table 29: Nominal and alternative VH samples used to assess systematic uncertainties for the signal processes.

where 1% is the maximum of the generic size of the neglected NNLO EW effects, the δ_{EW} stands for the relative NLO EW correction, and Δ_γ is the uncertainty of the photon induced cross section relative to the total VH cross section.

The signal samples are normalized to the two contributing cross sections, $\sigma_{W^+H} = 0.84$ pb and $\sigma_{W^-H} = 0.53$ pb which give a total production cross section of 1.37 ± 0.04 pb. The normalization uncertainties take into account these cross sections and the Higgs boson's predicted branching ratio of 58% at a mass of 125 GeV with 1.7% uncertainty on the LHC Higgs working group's recommendation.

Systematic samples were created with an alternative matrix elements (MADGRAPH5_aMC@NLO) and alternative parton shower model (HERWIG 7) to study the generators. The systematic samples can be seen in Table 6.2.1. The renormalization and factorization scale factors were compared at multiplicatives of three unlike the background samples which were compared at two. The PDFs and α_s used have a confidence level of 68% corresponding to an uncertainty of 1.9%. The QCD scale uncertainty was 0.7%. When comparing the nominal to alternative samples, they were normalized to the same cross section in order to isolate the acceptance uncertainties.

Shape systematics were once again performed on the two most important physical distributions in the BDT: m_{bb} and p_T^V. The difference between the nominal and the MADGRAPH was parameterized by a second order polynomial for the former and a linear fit for the latter. The up and down tune for each variation was fitted separately for events with two jets and those with three. The p_T^V shape deviations were small and similar, so the value of 12.1%

Nuisance Parameter	2 jets	3 jets
ATLAS_UEPS_VH_hbb	15.9%	16.3%
ATLAS_UEPS_VH_hbb_32JR	2.4%	3.8 %
ATLAS_UEPS_VH_hbb_VPT	shape only	
ATLAS_UEPS_VH_hbb_MBB	shape only	
QCDscale_VH_ANA_hbb_J2	0.3%	0.3%
QCDscale_VH_ANA_hbb_J3	1.3 %	1.7%
QCDscale_VH_ANA_hbb_JVeto	1.6%	0.2%
QCDscale_VH_ANA_hbb_VPT	shape only	
QCDscale_VH_ANA_hbb_MBB	shape only	
pdf_HIGGS_VH_ANA_hbb	1.5%	1.5%
pdf_VH_ANA_hbb_VPT	shape only	
pdf_VH_ANA_hbb_MBB	shape only	

Table 30: Systematic uncertainties for signal processes.

was used for both. Because the m_{bb} distributions were complex, the largest deviation was used as the shape uncertainty.

$W + jets$	
W+ll normalization	32%
$W + c$l normalization	37%
$W + bb$ normalization	Floating (2-jet, 3-jet)
$W + b$l-to-$W + bb$ ratio	23%
$W + bc$-to-$W + bb$ ratio	30%
$W + cc$-to-$W + bb$ ratio	30%
$W + HF$ CR to SR ratio	10%
m_{bb}, p_T^V	S
$t\bar{t}$	
$t\bar{t}$ normalization	Floating
$W + HF$ CR to SR ratio	25%
m_{bb}, p_T^V	S
Single top quark	
Cross section (s-channel)	4.6%
Cross section (t-channel)	4.4%
Cross section (Wt)	6.2%
Acceptance, 2-jet	17% (t-channel), 35% (Wt)
Acceptance, 3-jet	20% (t-channel), 41% (Wt)
m_{bb}, p_T^V	S (t-channel, Wt)
Z+jets	
Z+ll normalization	18%
$Z + c$l normalization	23%
$Z + bb$ normalization	Floating (2-jet, 3-jet)
$Z + bc$-to-$Z + bb$ ratio	30-40%
$Z + cc$-to-$Z + bb$ ratio	20-25%
m_{bb},p_T^V	S
ZZ	
Normalization	20%
Acceptance from scale variations (var.)	10-18% (Stewart-Tackmann jet binning method)
m_{bb}, p_T^V, from scale var.	S (correlated with WZ uncertainties)
m_{bb}, p_T^V, from PS/UE var.	S (correlated with WZ uncertainties)
m_{bb}, p_T^V, from matrix-element var.	S (correlated with WZ uncertainties)
WZ	
Normalization	26%
Acceptance from scale var.	12-21% (Stewart-Tackmann jet binning method)
Acceptance from PS/UE var. for 2 or more jets	3.9%
Acceptance from PS/UE var. for 3 jets	
Acceptance from m_{bb}, p_T^V, from scale var.	S (correlated with ZZ uncertainties)
m_{bb}, p_T^V, from PS/UE var.	S (correlated with ZZ uncertainties)
m_{bb}, p_T^V, from matrix-element var.	S (correlated with ZZ uncertainties)
WW	
Normalization	25%

Table 31: Systematic uncertainties of the MC backgrounds. S is listed when only a shape uncertainty is used. The regions where normalizations are allowed to float independently are listed in parentheses. “PS/UE” stands for parton shower / underlying event.

Multijet Sample	
Normalization	60-100% (2 jets), 100-400% (3 jets)
BDT template	Shape Uncertainty

Table 32: The systematic uncertainties for the multijet sample. S is listed when only a shape uncertainty is used. The regions where normalizations are allowed to float independently are listed in parentheses.

6.2.2 Experimental Uncertainties

Detector Based

Experimental uncertainties originate from the capabilities of ATLAS' detectors and the measurements made with them. The three dominant experimental uncertainties are: flavor-tagging simulation-to-data efficiency correction factors, jet energy scale corrections and jet energy resolution.

The flavor-tagging simulation-to-data efficiency relies on multiple correction factors dependent on the flavor, p_T and $|\eta|$ of the jet. These correction factors are treated as independent uncorrelated components. There are three components for b-tagged jets with a combined uncertainty of 2%, three components for c-tagged jets with a combined uncertainty of 10% and five components for light tagged jets with a combined uncertainty of 30%. Additional uncertainties associated with b-jet efficiency calibrations are necessary for jets with $p_T > 300$ GeV due to the misidentification of hadronically decaying τ leptons.

Jet energy scale and resolution's uncertainties are dependent on measurements in data. While not dominant, there are uncertainties from the reconstruction, identification, isolation, trigger efficiencies, energy scale and resolution of muons and electrons. These are also based on the data, but they are relatively small. The missing transverse energy is calculated based on the jets and leptons present in the event, so uncertainties associated with them are propagated to the E_T^{miss}. The E_T^{miss}also has a soft term computed from tracks, so the scale, resolution and efficiency of the tracks are additional incorporated uncertainties. Finally, when an event is triggered by a E_T^{miss}based trigger, there is an additional trigger scale factor to account for simulation-to-data statistical uncertainties.

The LHC's delivered luminosity has a combined uncertainty of 3.2% for the dataset used in this analysis. It comes from luminosity scale calibrations using $x-y$ beam separation scans performed in 2015 (2.1%) and 2016 (3.4%).

There is an additional 9% uncertainty on the average number of interactions per bunch crossing due to rescaling for simulation-data agreement.

7 Results

This chapter looks at signal over background yields in comparison to the standard model yields predicted by the MC discussed in Sections 3.1.1 and 3.1.3, as well as a signal sensitivity calculation of $\log(S/B)$. First, the results from the 1 Lepton WH search are shown, as well as the 1 Lepton specific validations of the dijet mass and the VZ diboson cross-checks. Next the results are quoted for the combination of all associated VH production seen in Run 2 as well as its corresponding cross-checks, the dijet mass and VZ validations. The decorrelated 1 Lepton signal strength and WH production signal strength are also discussed. Next the signal strengths are converted into production cross-sections for the WH, ZH, and VH cross sections. Finally the Run 1 and Run 2 combination in ATLAS is quoted to provide context for future physics analyses.

7.1 $WH \to b\bar{b}$

After the full event selection is made in the 1 Lepton mode, the BDT is trained for the VH signal on half of the MC events. The resultant BDT is then applied to the full 36.1 fb^{-1} of data, all background samples and the simulated signal events.

The fitting procedure to extract the signal first fits the output BDT by normalizing the backgrounds, then refines the fit further by setting the systematic uncertainties according to the values of the floating normalizations and nuisance parameters obtained during the fit.

Normalizing the backgrounds in the CR allow better constraints to be made on the $W + HF$ background in the SR. The kinematic distributions in the SR can be seen in Figures 22 and 23. All resultant plots have the signal scaled to the best signal strength value. The kinematic distributions show good data-prediction agreement indicating the BDT describes the data well. The obtained fitted signal strength is

$$\mu = 1.35^{+0.40}_{-0.39}(\text{stat.})^{+0.56}_{-0.47}(\text{syst.})$$

The resultant BDT histograms are shown in Figure 24, and the corresponding post-likelihood yields can be found in Table 33. When the Poisson probability has been converted to it's gaussian equivalent significance, the expected significance is 1.80 standard deviations and the observed significance is 2.30 standard deviations.

The BDT discriminant is gathered into bins of signal over background inclusive of both signal regions. Figure 25 shows the event yields plotted in the form of $\log(S/B)$ in order to best distinguish the signal from the background.

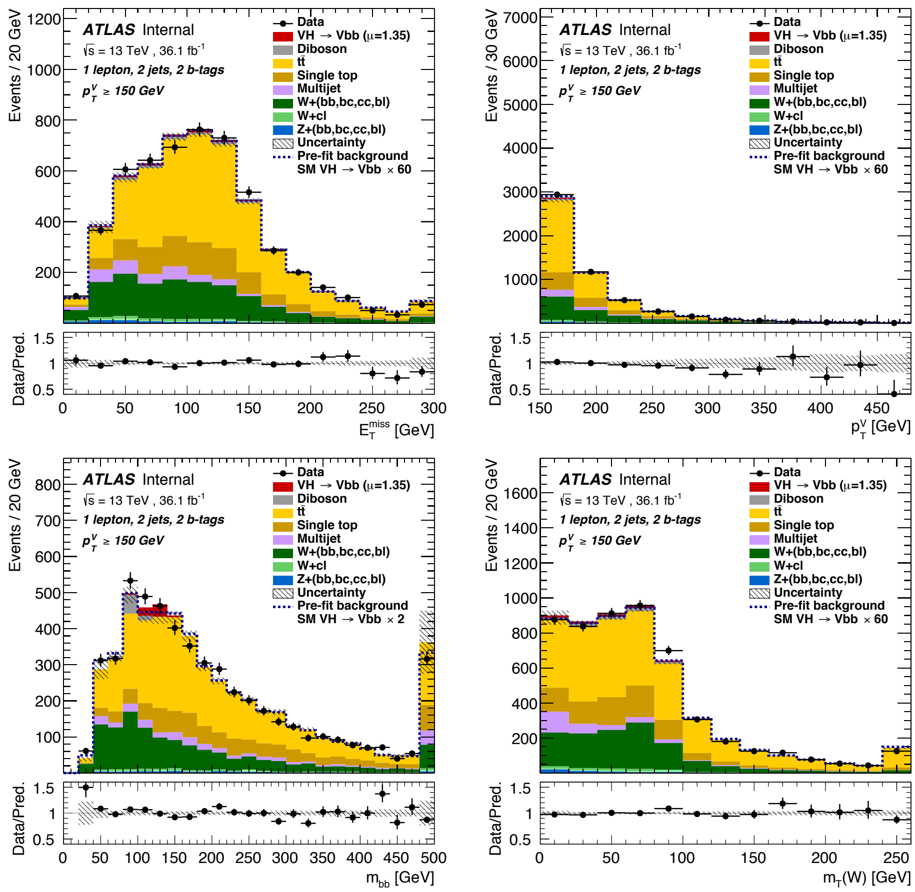


Figure 22: Fitted input variables, E_T^{miss}, p_T^W, m_{bb} and m_T^W shown in the 1 Lepton mode and high p_T^W regime to illustrate the good agreement between data and prediction in events with 2 jets. The VH signal has been scaled by the best fit value in the histogram.

Process	2 jets CR	3 jets CR	2 jets SR	3 jets SR
Z+$\ell\ell$	< 1	< 1	< 1	-
Z+$c\ell$	-	< 1	2.2±0.4	4.3± 0.6
Z+HF	5.1±0.9	15.4±1.5	62.9±3.0	154.6±4.6
W+$\ell\ell$	1.1±0.2	2.9±0.7	23.2±4.2	49.2±5.7
W+$c\ell$	2.9±0.9	10.6±1.6	75.1±4.8	135.8±5.9
W+HF	195.8±7.6	453.5±10.0	1052.6±20.5	2480.4±783.0
Single top	10.0±2.4	43.4±4.5	863.8±18.5	3926.7±41.0
$t\bar{t}$	24.0±3.1	116.2±6.8	2869.0±34.2	22395.1±95.8
Diboson	5.0±1.3	11.9±1.5	79.4±4.3	147.9±5.8
Multijet (e)	9.6±0.6	2.4±0.1	119.9±1.9	29.0±0.5
Multijet (μ)	7.1±1.4	3.5±0.8	139±6.3	86.3±3.9
Total bkg	260.5±3.3	660.0±2.0	5287.9±0.9	29409.3±0.4
Signal (fit)	< 1	< 1	51.1±0.3	55.5±0.3
Data	302	790	5307	28168
S/B	0.004	< 0.002	0.010	0.002
% Difference from prediction	15.5	19.5	-0.6	-4.4

Table 33: Fitted yields of the BDT fit for the 1 Lepton analysis with VH signal in the control and signal regions.

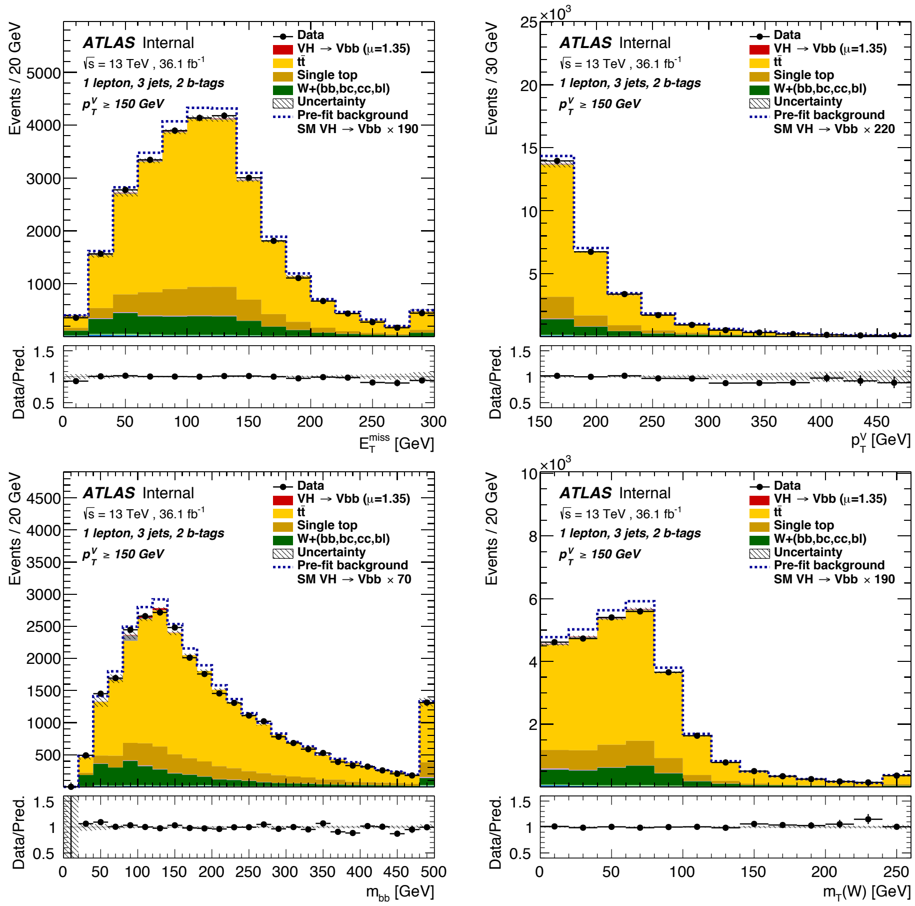


Figure 23: Fitted input variables, E_T^{miss}, p_T^W, m_{bb} and m_T^W shown in the 1 Lepton mode and high p_T^W regime to illustrate the good agreement between data and prediction in events with 3 jets. The VH signal has been scaled by the best fit value in the histogram.

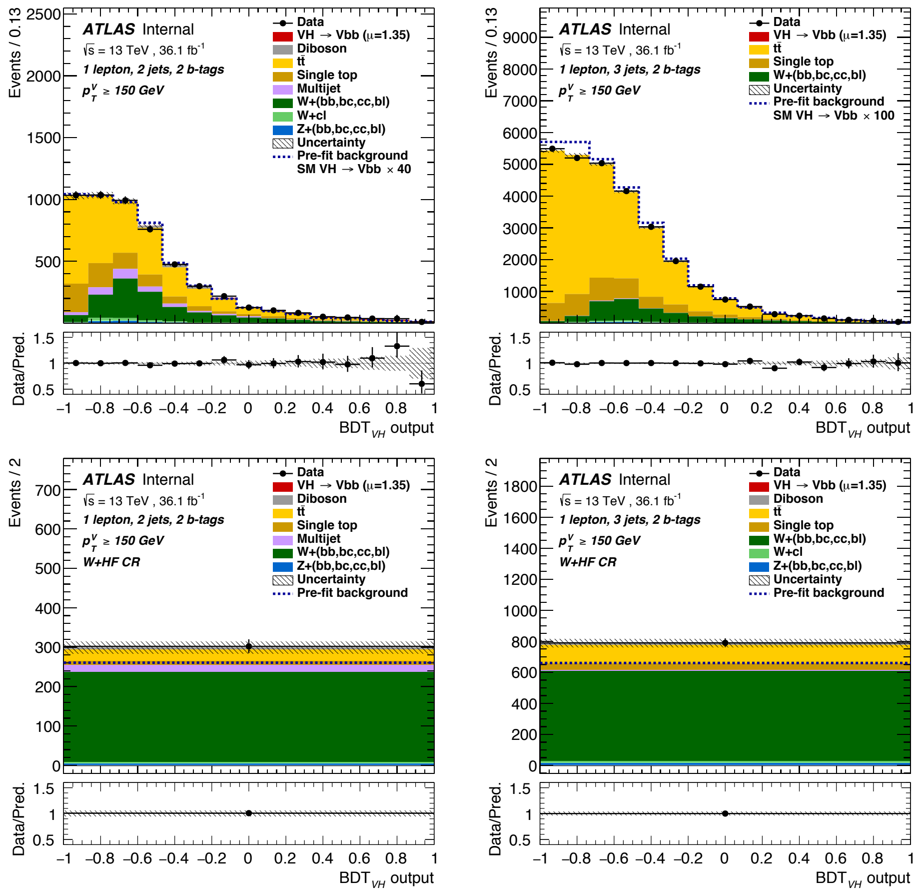


Figure 24: Postfit plots of the Signal (top) and Control (bottom) Regions in 2 (left) and 3 (right) jets regions of the BDT binned variable in the 1 Lepton mode with VH as the extracted signal. The red line is the signal multiplied to be an arbitrary scale factor to show the discerning shape difference.

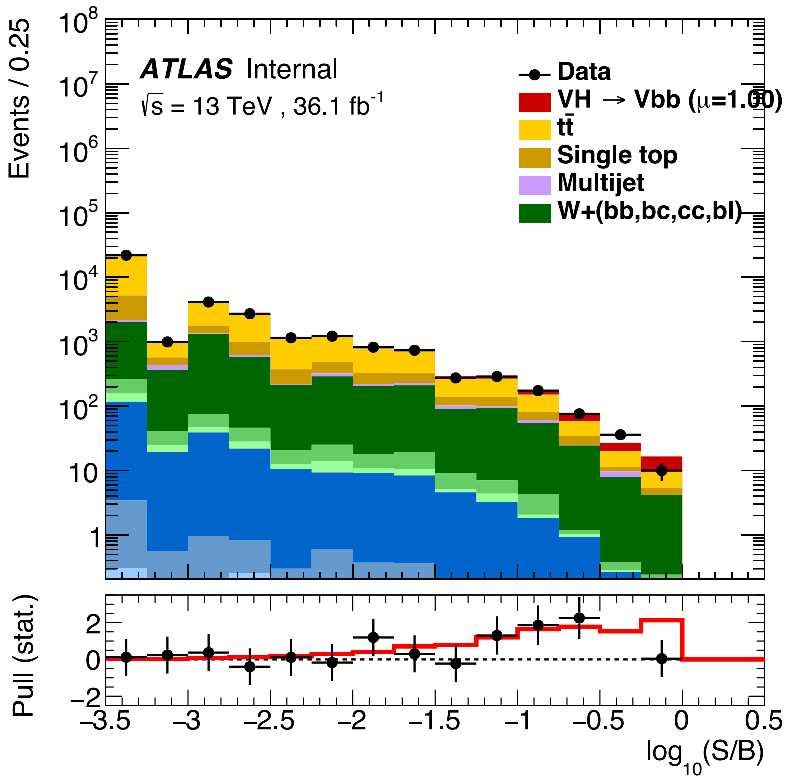


Figure 25: The combined signal regions event yields shown via log(S/B) for the 1 Lepton mode with VH rescaled to the signal strength extracted from data. In the bottom ratio plot, the pull (residual divided by its uncertainty) is shown compared to the background-only prediction with statistical uncertainties only.

7.1.1 Dijet Mass Validation

One way to verify this analysis is to perform a simplistic cut and count search. In this case, the important discriminant is m_{bb} from which the Higgs boson is inferred. A cut and count analysis is usually not as precise as a multivariate algorithm because the fit relies on a single variable distribution, as opposed to using multiple variables to distinguish signal from background on an event-by-event basis. However, the dijet mass validation is useful as a check on the central value obtained by a MVA.

The postfit kinematic distributions of the characteristic variables, E_T^{miss}, p_T^W, m_{bb} and m_T^W, show good data-prediction agreement indicating the dijet mass analysis result describes the data well. The distributions can be found in Figures 26 and 27, and the accompanying yields can be found in Table 34.

Compared to the previous BDT-based analysis, the dijet mass analysis includes an additional requirement that the reconstructed transverse mass of the W boson be less than 120 GeV to suppress $t\bar{t}$ background. The dijet mass analysis has two p_T^W regimes, $150 \leq p_T^W < 200$ GeV and ≥ 200 GeV, which each have an additional requirement on the distance between the two b-tagged jets, $\Delta R(b_1, b_2)$, to be less than 1.8 and 1.2 for the two p_T^W respectively in order to reduce multijet backgrounds. No dedicated $W + HF$ control region is created in the dijet mass analysis since the m_{bb} spectrum already provides sufficient information to constrain the $W + HF$ background. This m_{bb} spectrum is seen in Figure 28 for signal regions defined by p_T^W and the number of signal jets.

To isolate the likely VH contribution, predicted backgrounds excluding VZ and VH are subtracted from data. The VZ is not subtracted due to the proximity of the $Z \to b\bar{b}$ process at a mass of 91 GeV to the $H \to b\bar{b}$ process at a mass of 125 GeV. The resultant plot, Figure 29, shows a fitted signal strength of

$$\mu = 1.39^{+0.48}_{-0.46}(\text{stat.})^{+0.63}_{-0.52}(\text{syst.})$$

and the agreement of the fit can be seen by the naked eye best in this figure.

The dijet mass analysis result has an extracted expected significance of 1.49 standard deviations and a measured significance of 2.01 standard deviations. The log(S/B) plot without any subtractions can be found in Figure 30.

The difference in the central value between the dijet mass analysis and the VH MVA analysis discussed previous is only 3%. This is well within

Process	[150,200), 2 jets	[200, ∞), 2 jets	[150,200), 3 jets	[200, ∞), 3 jets
Z+ $\ell\ell$	< 1	< 1	-	< 1
Z+ $c\ell$	< 1	< 1	1.8± 0.3	< 1
Z+HF	18.8±1.7	5.0±0.7	46.9±2.8	15.4±1.1
W+ $\ell\ell$	4.7±0.9	3.7±1.1	15.1±4.8	4.6±0.6
W+ $c\ell$	20.2±2.8	8.0±1.5	31.2±2.8	18.9±2.0
W+HF	493.3±14.5	241.3±7.5	1046.1±17.6	567.1±12.7
Single top	183.7±9.1	39.5±4.4	795.2±4.4	208.3±9.4
$t\bar{t}$	804.4±18.0	108.2±6.6	4466.3±42.8	752.6±19.5
Diboson	41.4±3.1	33.4±2.7	63.6±3.8	45.1±3.1
Multijet e	13.5±0.5	10.6±0.4	16.3±0.6	15.9±0.5
Multijet μ	47.0±3.7	11.4±2.1	34.9±3.3	8.9±2.1
Total bkgd	1627.8±1.6	461.3±2.6	6516.7±0.8	1637.3±1.6
Signal (fit)	22.5±0.2	19.6±0.2	22.0±0.2	18.9±0.1
Data	1774	495	6397	1635
S/B	0.014	0.042	0.003	0.012
% Difference from prediction	7.5	2.9	-2.2	-1.3

Table 34: Yields for dijet mass analysis in the 1 Lepton mode.

experimental uncertainties and helps validate the MVA analysis.

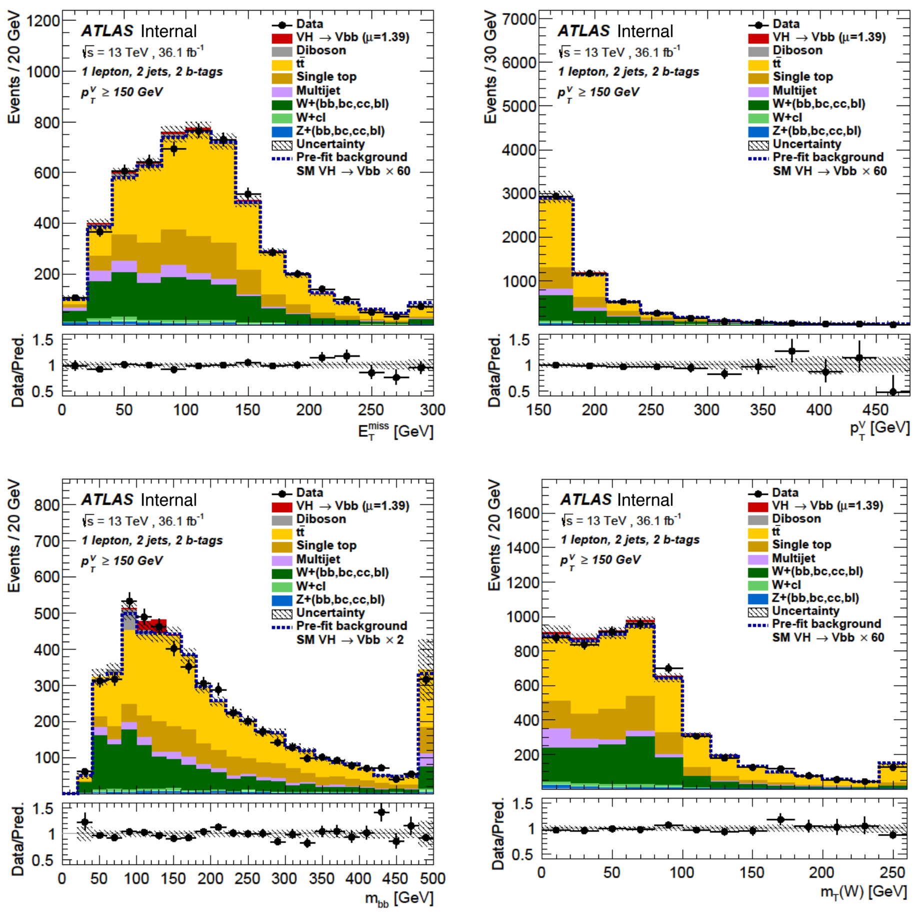


Figure 26: Fitted input variables, E_T^{miss}, p_T^W, m_{bb} and m_T^W shown in the dijet mass analysis and high p_T^W regime to illustrate the good agreement between data and prediction in events with 2 jets. The VH signal has been scaled by the best fit value in the histogram.

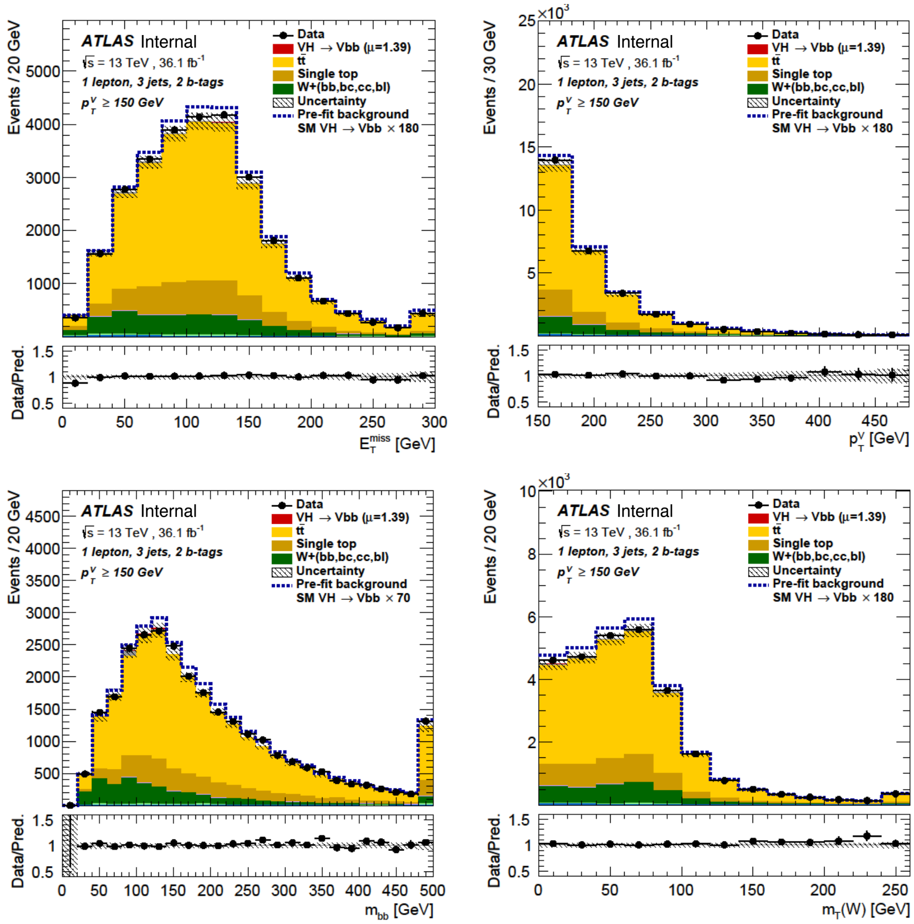


Figure 27: Fitted input variables, E_T^{miss}, p_T^W, m_{bb} and m_T^W shown in the dijet mass analysis and high p_T^W regime to illustrate the good agreement between data and prediction in events with 3 jets. The VH signal has been scaled by the best fit value in the histogram.

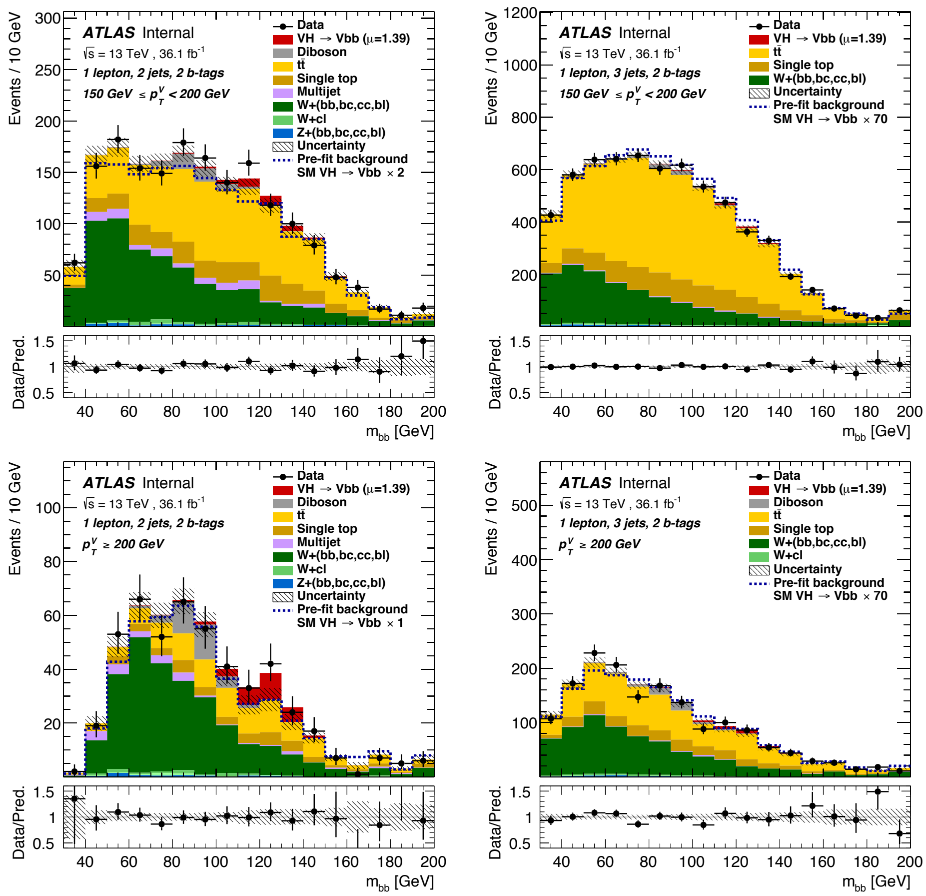


Figure 28: Postfit plots of the signal regions in the 2 (3) jets and $150 < p_T^V <$ 200 GeV ($p_T^V > 200 GeV$) for the left, right, top and bottom plots. This cross-check is a cut and count analysis that searches for the Higgs mass in the m_{bb} spectrum.

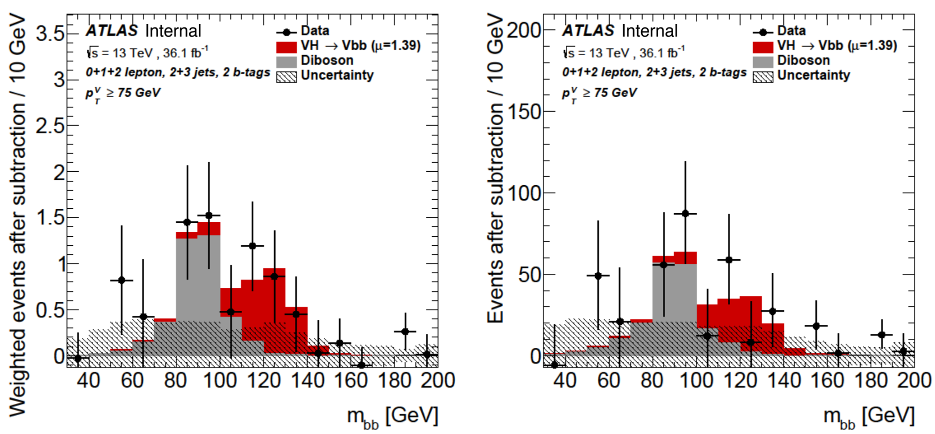


Figure 29: Data/prediction plots after all background processes have been subtracted excluding the diboson processes. The left shows the samples that have been weighted by their relative Higgs signal to background contributions, and the right shows the plot before reweighting. The VH signal has been scaled to the best fit value, $\mu = 1.39$. The uncertainty band incorporates both statistical and systematical uncertainties.

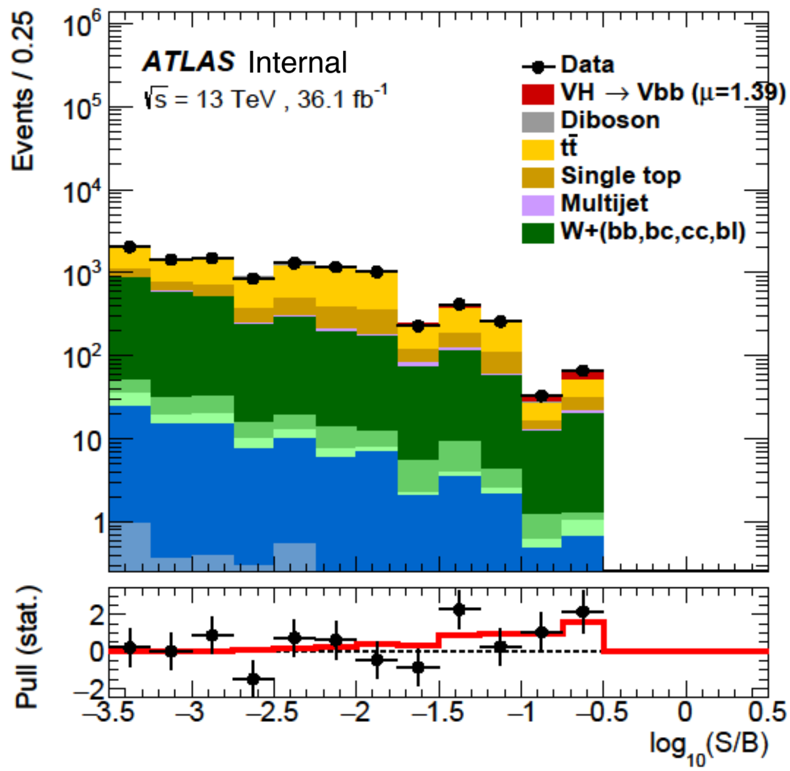


Figure 30: The combined signal regions event yields shown via log(S/B) for the 1 Lepton dijet mass analysis with VH rescaled to the signal strength extracted from data. In the bottom ratio plot, the pull (residual divided by its uncertainity) is shown compared to the background-only prediction with statistical uncertainties only.

7.1.2 Diboson Validation

To validate the BDT's power to distinguish signal from background, a cross-check analysis is performed using dibosons, VZ, as the signal instead of VH. The diboson signal requires the Z boson to decay into a pair of $b\bar{b}$ quarks similar to the Higgs boson. The main phenomenological difference is the reconstructed mass profile centered around the Z boson's mass of 91 GeV instead of the Higgs boson's mass of 125 GeV. The VZ BDT is trained independently of the VH BDT. Due to lower VZ MC statistics available than VH MC statistics, the VZ BDT is more statistical limited. This results in wider VZ BDT bins.

The output BDT is once again allowed to have floating normalizations for the backgrounds taking into account the systematic uncertainties according to the floating normalization and nuisance parameters obtained as part of the signal strength fit.

The performance of the VZ BDT is checked by looking at the characteristic variables, E_T^{miss}, p_T^W, m_{bb} and $\Delta R(b_1, b_2)$, in the signal and control regions. As seen in Figures 31 and Figure 32, there is good data-prediction agreement. The VZ signal is multiplied by the fitted signal strength in the Figures. A signal strength completely consistent with the SM would have a signal strength of $\mu = 1.00$. The resultant fitted signal strength is consistent with the SM predication:

$$\mu = 1.03 \pm 0.23(\text{stat.})^{+0.54}_{-0.55}(\text{syst.})$$

The VZ BDT distributions can be seen in Figure 33. The translated expected significance is 1.79 standard deviations while the observed is 1.69 standard deviations. Figure 34 shows the event yields in the log(S/B) plot, while the yields in the BDTs are detailed in Table 35.

Process	2 jets CR	3 jets CR	2 jets SR	3 jets SR
$Z + \ell\ell$	< 1	< 1	< 1	-
$Z + c\ell$	-	< 1	2.2±0.4	4.2±0.6
$Z + HF$	5.1±0.9	15.4±1.5	63.3±3.0	154.7±4.6
$W + \ell\ell$	1.1±0.2	2.9±0.7	23.3±4.2	49.2±5.7
$W + c\ell$	2.9±0.9	10.6±1.6	75.1±4.8	135.8±5.9
$W + HF$	195.7±7.3	453.5±10.0	1051.9±20.5	2479.4±28.0
Single top	9.9±2.4	43.4±4.5	863.8±18.5	3926.7±41.0
$t\bar{t}$	24.0±3.1	116.2±6.8	2869.0±34.2	22395.1±95.8
VH	0.3±0.02	1.0±0.03	51.1±0.3	55.5±0.3
Multijet e	9.6±0.6	2.4±0.1	119.9±1.9	29.0±0.5
Multijet μ	7.1±1.4	3.5±0.8	138.2±6.3	86.3±3.9
Total bkgd	256.1±3.3	651.1±2.0	5262.5±0.9	29331.2±0.4
Signal (fit)	5.0±1.0	12.1±1.5	79.4±4.3	147.9±5.8
Data	302	790	5307	28168
S/B	0.020	0.019	0.015	0.005
% Difference from prediction	15.7	19.1	-0.7	-4.4

Table 35: Yields for VZ analysis for the 1 Lepton mode.

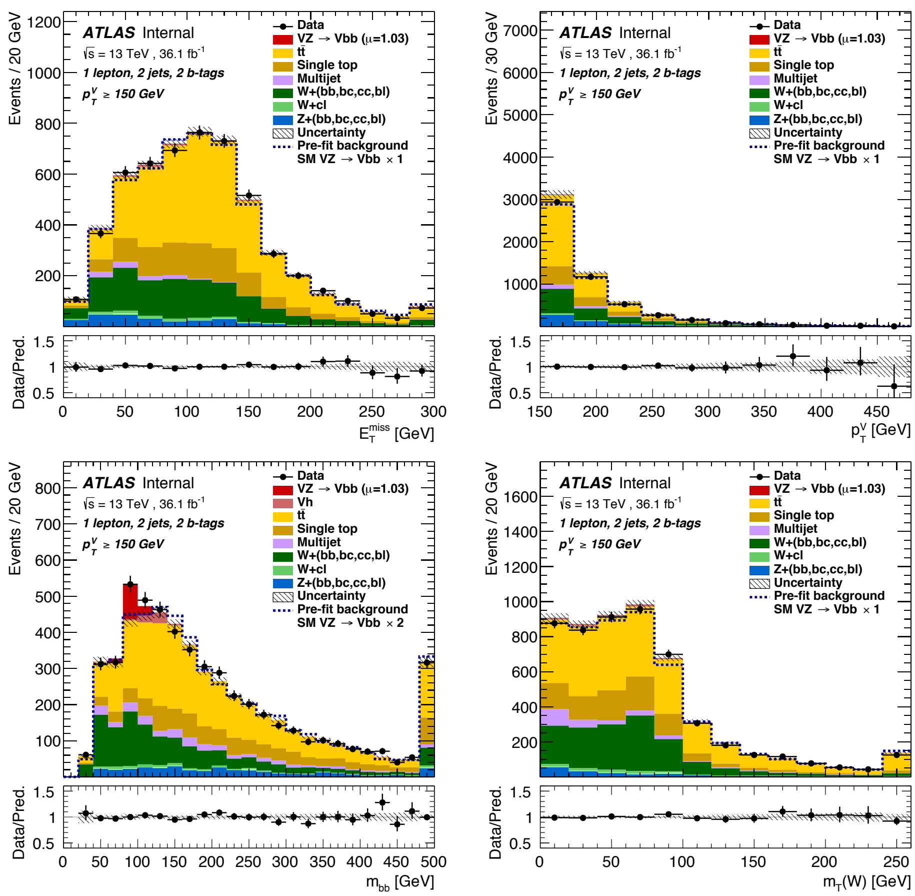


Figure 31: Fitted input variables, E_T^{miss}, p_T^W, m_{bb} and m_T^W shown in the 1 Lepton mode and high p_T^W regime to illustrate the good agreement between data and prediction in events with 2 jets. The VZ signal has been scaled by the best fit value in the histogram.

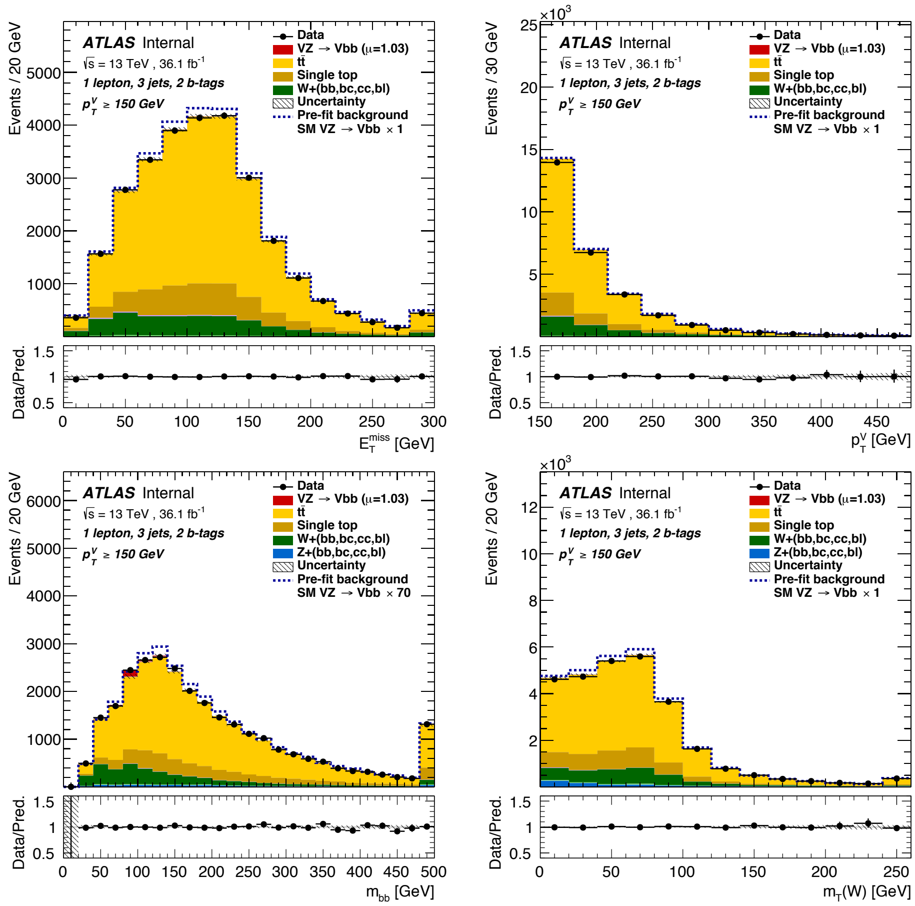


Figure 32: Fitted input variables, E_T^{miss}, p_T^W, m_{bb} and m_T^W shown in the 1 Lepton mode and high p_T^W regime to illustrate the good agreement between data and prediction in events with 3 jets. The VZ signal has been scaled by the best fit value in the histogram.

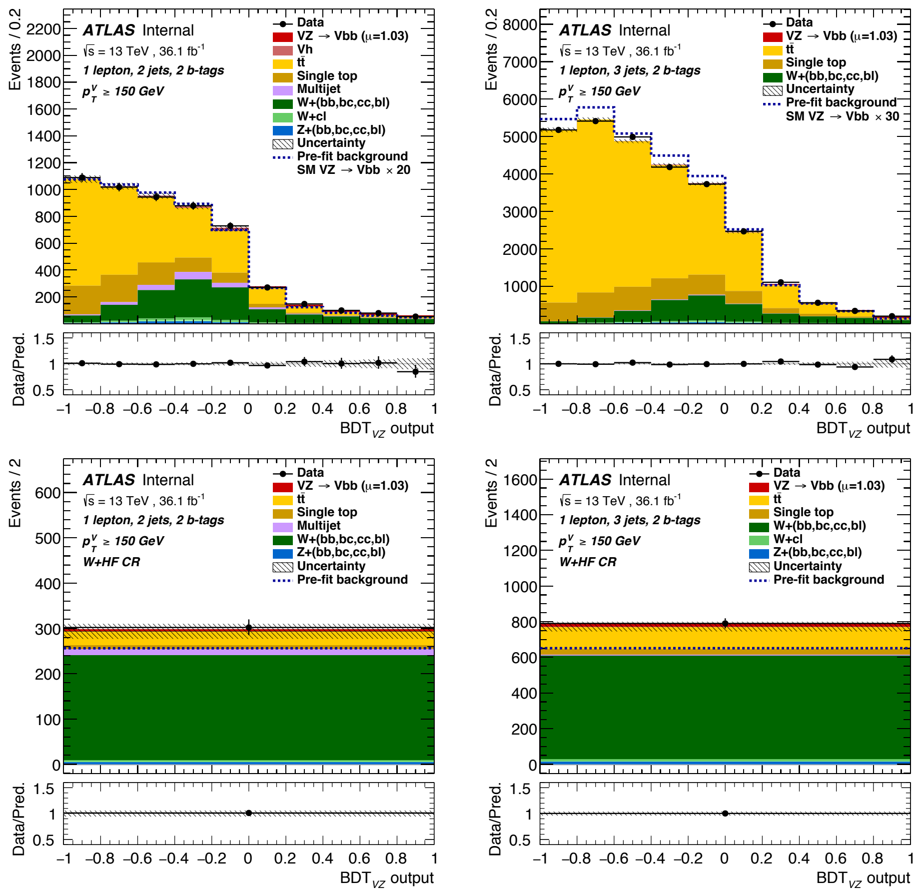


Figure 33: The BDT discriminant in the signal and control regions of the 1 Lepton mode with $VZ \to Vb\bar{b}$ as the extracted signal.

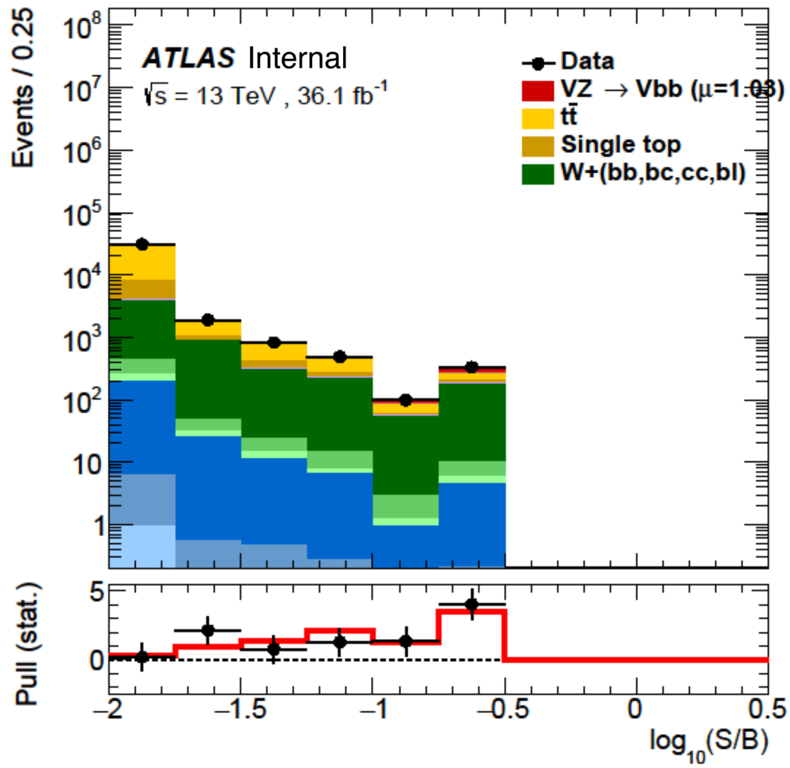


Figure 34: The combined signal regions event yields shown via log(S/B) for the 1 Lepton mode with $VZ \to Vb\bar{b}$ rescaled to the signal strength extracted from data. In the bottom ratio plot, the pull (residual divided by its uncertainity) is shown compared to the background-only prediction with statistical uncertainties only.

7.2 VH Combination

Since the goal of the over-arching analysis was to find evidence of the $H \rightarrow b\bar{b}$ decay, the results from all three VH production modes searches were combined. This dissertation detailed the 1 Lepton mode which searched for the mode $WH \rightarrow \ell\nu b\bar{b}$. The result shown at the European Physical Society Conference on High Energy Physics 2017 and submitted to be published in the Journal of High Energy Physics combines the three modes, 0, 1 and 2 leptons, including both WH and ZH leptonic decays in its combined BDT [74]. The following results summarize the combined VH analysis.

This combined fitting leads to slightly different best fit values of the 1 Lepton mode, as it is further constrained by requirement of consistency with the two other modes. Additionally, when looking at the final states to which WH contributes the associated W boson production is present in not just 1 Lepton, but is a non-negligible contributor to the 0 Lepton mode. The 0 Lepton mode search is characterized by a large E_T^{miss}contribution from a $Z \rightarrow \nu\nu$ decay, however, studies showed that up to 20% of the total VH signal in that mode actually originated from WH events where no lepton was reconstructed. Therefore when extracting the WH production cross-section from data, both modes must be taken into consideration.

The simultaneous fit of all three modes has good data-background agreement as shown in Figure 35 with the VH signal scaled to the final combined fitted value. Each individual mode's expected and observed significances converted from the Poisson probabilities can be found in Table 36. The three individual modes' BDTs for the combined fit are trained and applied in the same fashion as the previous section and can be found in Figure 36. The separate BDTs are all combined into a single BDT of which the log(S/B) plot is shown in Figure 37. The yields composition of the four most significant bins in the signal over background plot can be found in Table 37.

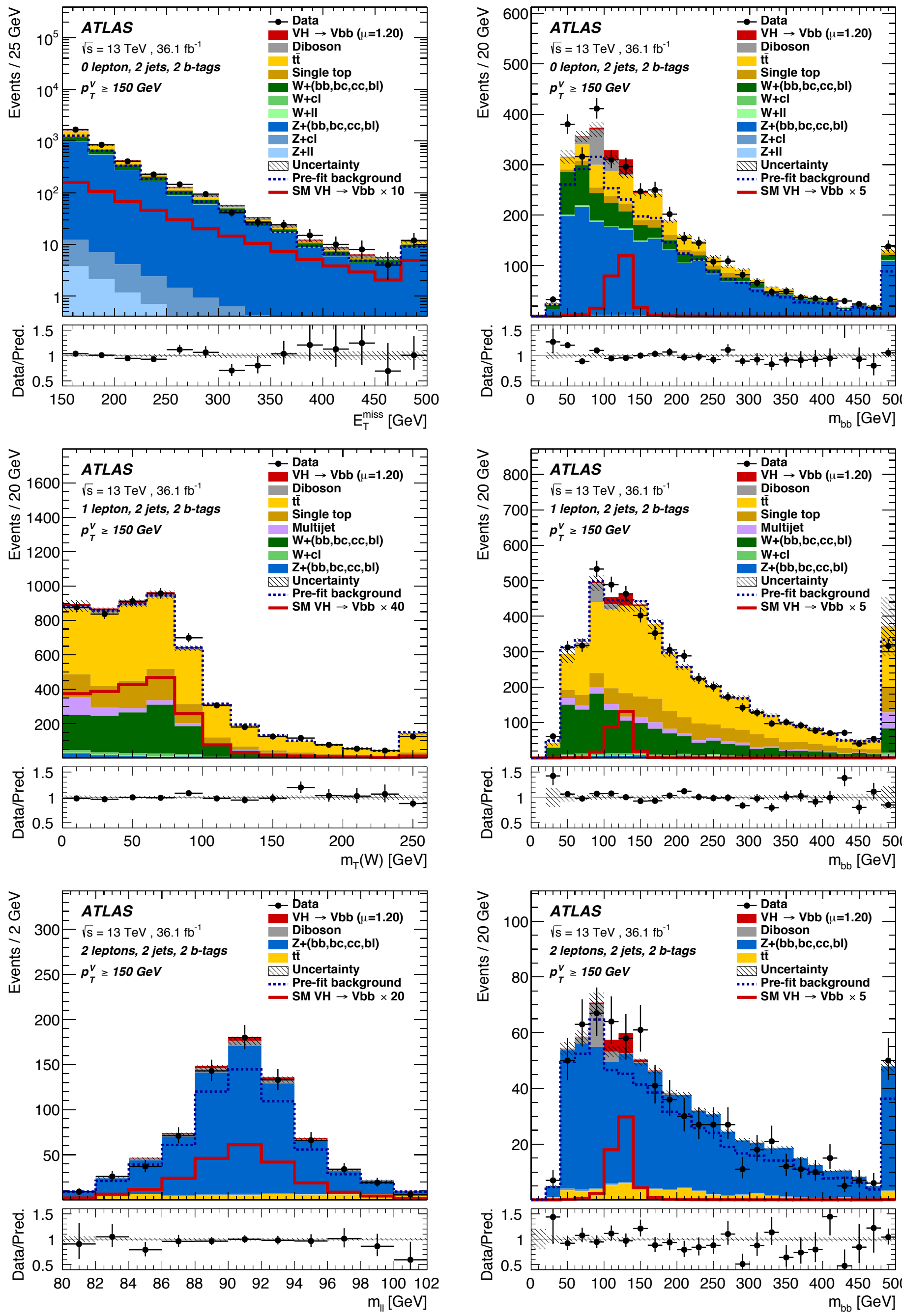


Figure 35: Kinematic distributions for 0, 1, 2 modes in simultaneous fit with VH as signal. Shows the inferred leptonic vector boson on the left and the Higgs boson on the right, for 0, 1, and 2 modes in descending order.

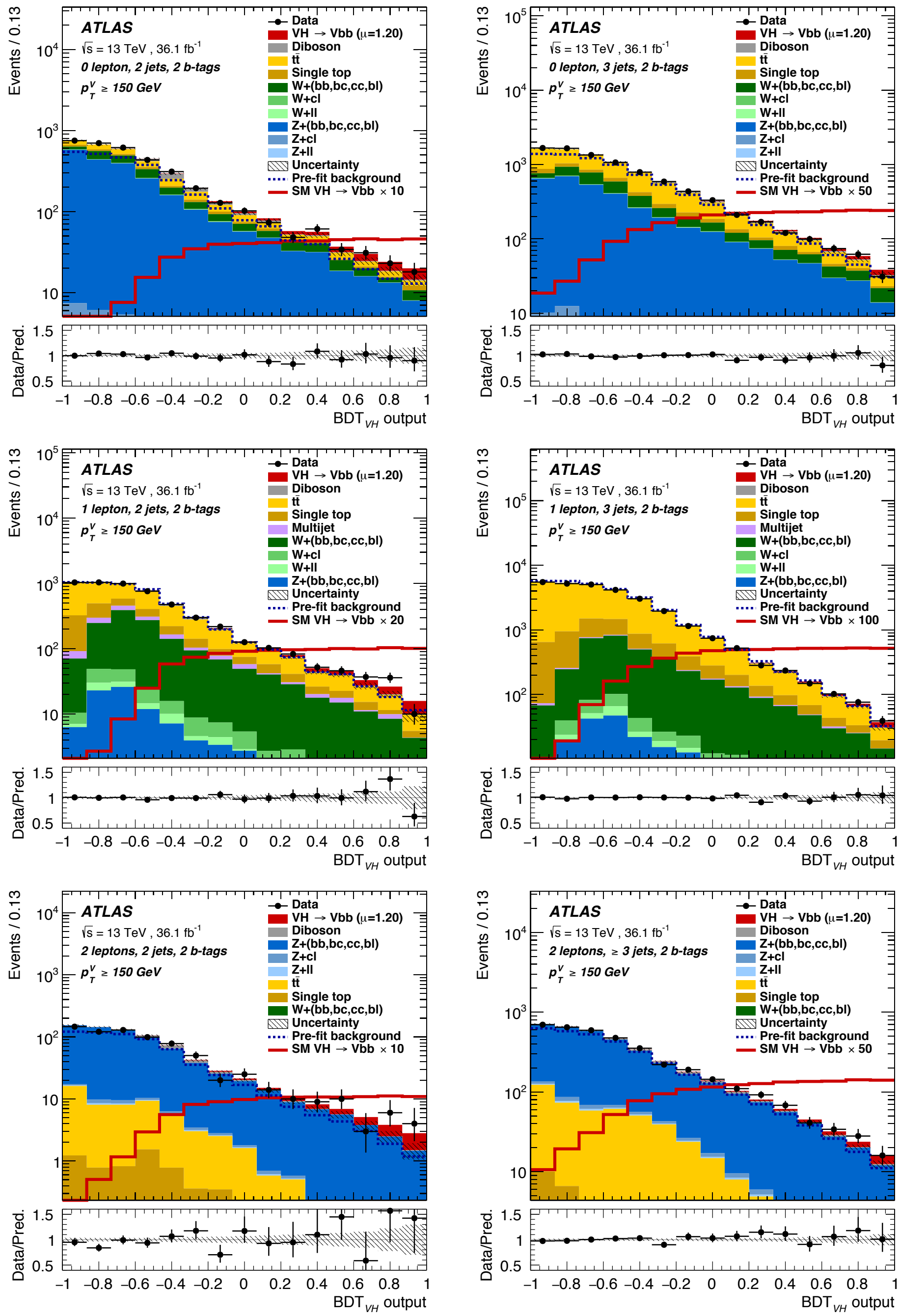


Figure 36: The BDT distributions of the simultaneously fit 0, 1 and 2 Lepton modes with the VH signal scaled to the corresponding best fit value of 1.20. The VH signal is overlayed in red and scaled to be visible.

Mode	Expected Significance	Observed Significance
0 Lepton	1.7	0.5
1 Lepton	1.8	2.3
2 Lepton	1.9	3.6
Combined	3.0	3.5

Table 36: Expected and observed significance results for combined VH search using the full luminosity at $\sqrt{s}$ =13 TeV in both the individual modes and the combined result.

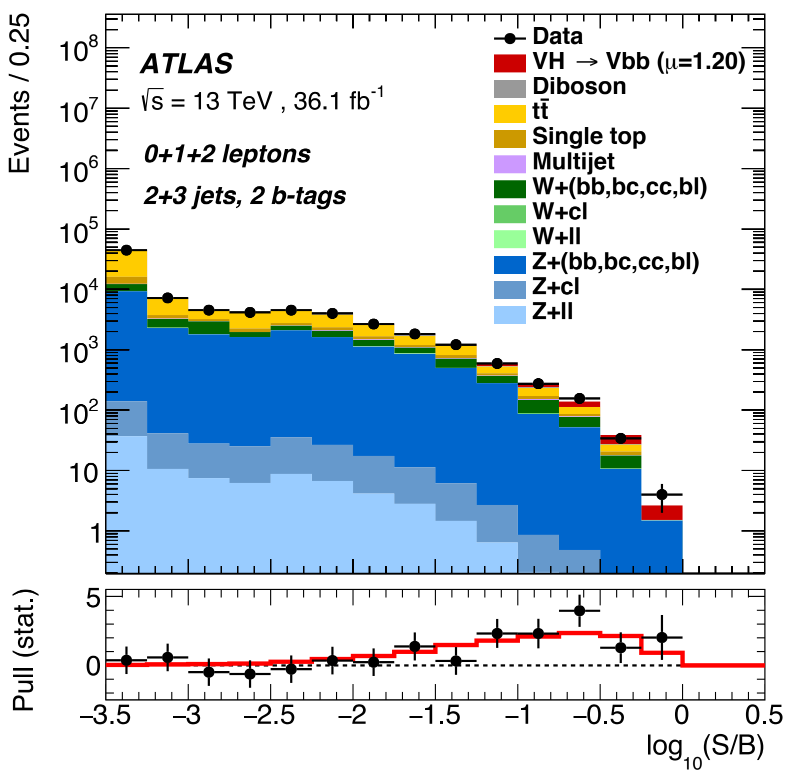


Figure 37: The combined signal regions event yields shown via log(S/B) for the combined VH modes with VH rescaled to the signal strength extracted from data. In the bottom ratio plot, the pull (residual divided by its uncertainity) is shown compared to the background-only prediction with statistical uncertainties only.

Process	Bin 11	Bin 12	Bin 13	Bin 14
Data	274	156	34	4
Signal (fit)	32.4	25.0	11.1	1.1
Total Background	238.3	113.7	27.3	1.5
S/B	0.14	0.22	0.41	0.73
Difference from prediction (%)	1.2	12.5	-11.5	53.8
$Z + \ell\ell$	0.2	0.1	< 0.1	< 0.1
$Z + c\ell$	0.7	0.4	< 0.1	< 0.1
Z+HF	86.1	51.3	10.5	1.5
$W + \ell\ell$	0.20	0.1	< 0.1	-
$W + c\ell$	1.6	0.2	< 0.1	-
W+HF	58.9	24.5	6.9	-
Single top-quark	19.2	7.6	2.9	-
$t\bar{t}$	61.3	25.7	6.2	-
Diboson	4.7	1.7	0.4	< 0.1
Mutli-jet e sub-ch.	0.1	-	-	-
Multij-jet μ sub-ch.	5.2	2.0	< 0.1	-

Table 37: Yields of the most sensitive bins in the signal over background plot of the three combined modes, 0, 1 and 2 leptons, found in Figure 37.

The fitted signal strength of the combined VH modes

$$1.20^{+0.24}_{-0.23}(\text{stat.})^{+0.34}_{-0.28}(\text{stat.})$$

is consistent with the SM prediction. The decorrelated 1 Lepton mode was found to be individually consistent with the SM as well at a signal strength of

$$\mu = 1.43^{+0.40}_{-0.38}(\text{stat.})^{+0.56}_{-0.45}(\text{syst.}).$$

The full set of signal strengths in comparison to the standard model for the combined analysis can be found in Figures 38 and 39, where the μ value is broken down separately by VH mode and by associated vector boson production - WH and ZH. Of particular interest to this analysis is the WH mode because its main contributor is the 1 Lepton mode. The WH production's signal strength was found to be consistent with the SM at

$$\mu = 1.35^{+0.40}_{-0.38}(\text{stat.})^{+0.55}_{-0.45}(\text{syst.}).$$

All combined results are consistent with the SM, and are above 3 standard deviations away from the null hypothesis of zero standard model signal present.

Since the WH and ZH modes were observed at 2.4 and 2.6 standard deviations respectively, the observed signal strength can be translated into cross sections (times the branching ratio for the Higgs decay and inclusive of all leptonic decays of the vector boson.) The cross sections stated below are quoted without theoretical uncertainties in the production cross-sections and branching ratio and are compared to expectation.

$$\begin{aligned}
\sigma(WH) \times \text{BR}(H \to b\bar{b}) &= 1.07^{+0.54}_{-0.47}\text{pb vs. } 0.80\text{pb} \\
\sigma(ZH) \times \text{BR}(H \to b\bar{b}) &= 0.56^{+0.25}_{-0.22}\text{pb vs. } 0.51\text{pb} \\
\sigma(VH) \times \text{BR}(H \to b\bar{b}) &= 1.57^{+0.54}_{-0.47}\text{pb vs. } 1.31\text{pb}
\end{aligned}$$

7.2.1 Dijet Mass Validation

As in Section 7.1.1, the analysis is validated by using the m_{bb} discriminant for a cut and count analysis to achieve a 3-mode simultaneous fitted result. The combined result is

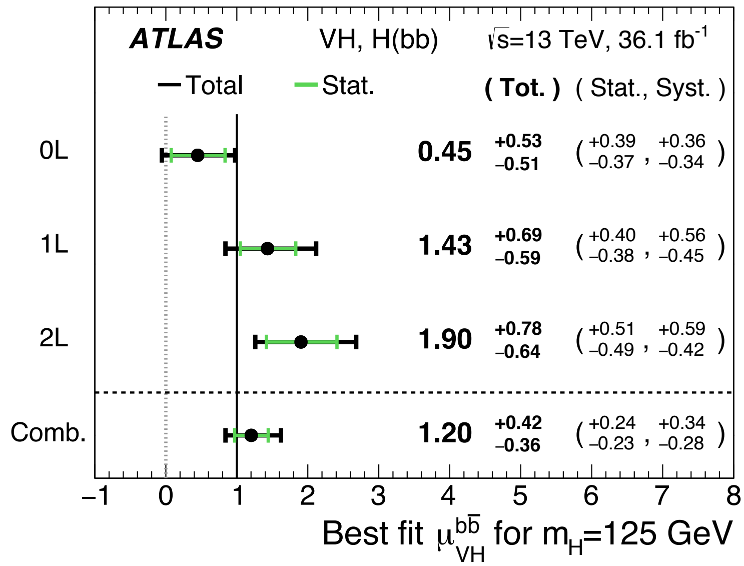


Figure 38: The fitted signal strength compared to the standard model in the combined fit decomposed by lepton analysis mode.

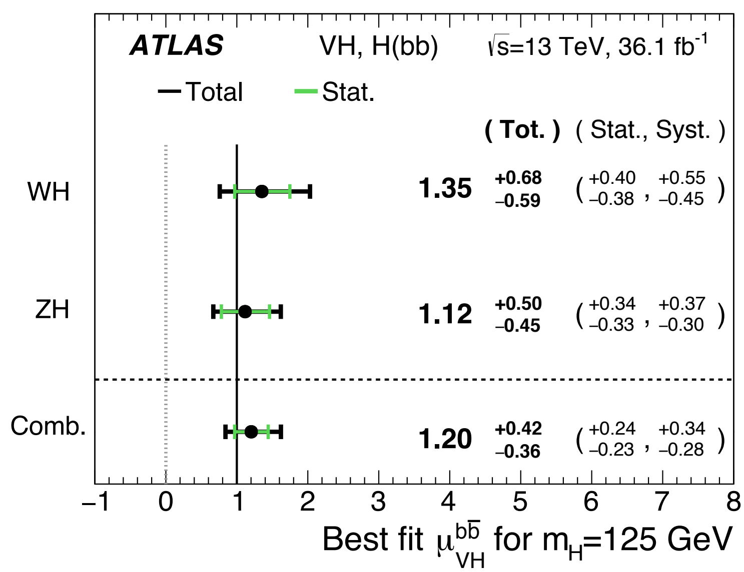


Figure 39: The fitted signal strength compared to the standard model in the combined fit decomposed by the mode of associated vector production for Run 2.

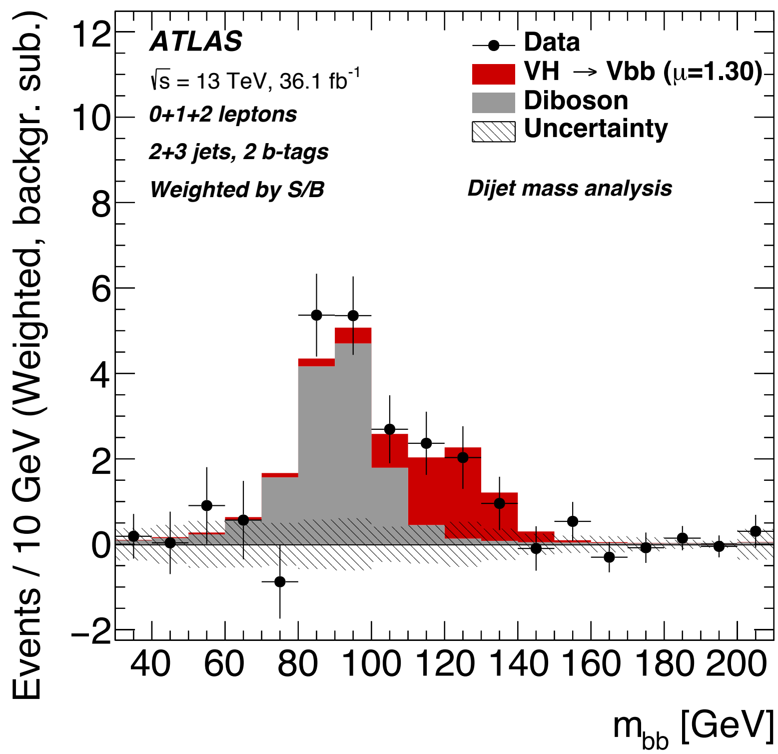


Figure 40: Data/prediction plots after all background processes have been subtracted excluding the diboson processes. The samples that have been weighted by their relative Higgs signal to background contributions, and the VH signal has been scaled to the best fit value, $\mu = 1.20$. The uncertainty band incorporates both statistical and systematical uncertainties.

$$\mu = 1.30^{+0.28}_{-0.27}(\text{stat.})^{+0.37}_{-0.29}(\text{syst.}).$$

A good visualization of the $VH \to b\bar{b}$ signal can be seen in Figure 40 by subtracting the predicated backgrounds from data except for VZ which has a close mass peak.

7.2.2 Diboson Validation

The BDT used for the combined result is validated in the same fashion as the 1 Lepton mode previously. The signal the BDT is trained to identified is switched from VH to the well established VZ with $Z \to b\bar{b}$. This gives a result of

$$\mu = 1.11^{+0.12}_{-0.11}(\text{stat.})^{+0.22}_{-0.19}(\text{syst.})$$

which is consistent with the standard model and indicates the trustworthiness of the BDT. Figures 41 and 42 illustrate the diboson BDT fit in each mode as well as the combined event yield as a function of log(S/B).

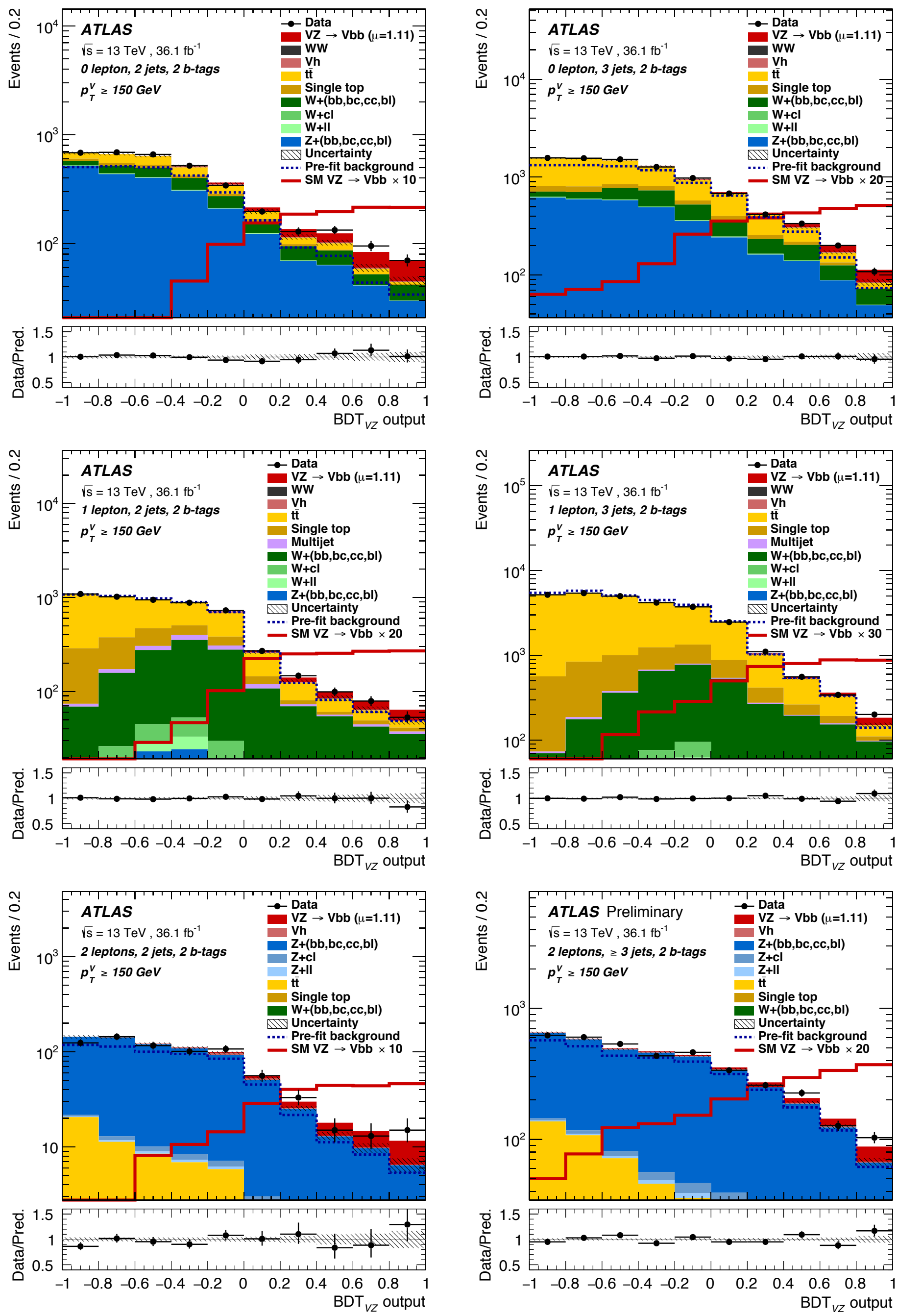


Figure 41: Combined BDT for the VZ cross-check analysis.

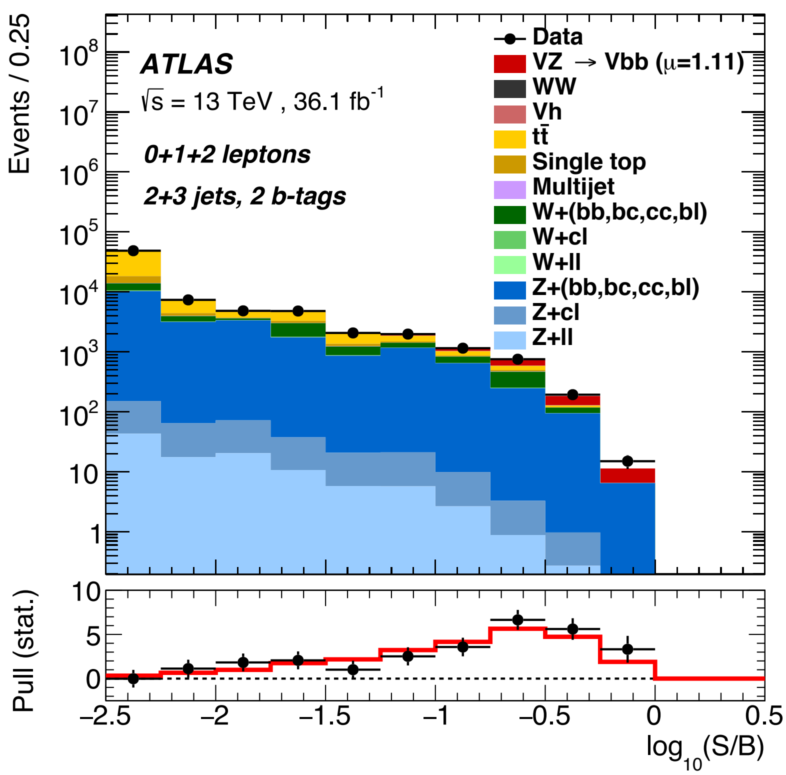


Figure 42: The combined signal regions event yields shown via log(S/B) for the combined Run 2 modes with $VZ \rightarrow Vb\bar{b}$ rescaled to the signal strength extracted from data. In the bottom ratio plot, the pull (residual divided by its uncertainity) is shown compared to the background-only prediction with statistical uncertainties only.

7.3 Run 1 - Run 2 Combination

In order to obtain the best possible result, the Run 2 results are combined with the Run 1 results [68]. The increase in integrated luminosity decreases the statistical uncertainty. Since there are few correlations between Run 1 and Run 2, the total systematic uncertainty also decreases when the two runs are combined. Both Run 1 and Run 2 utilized three Lepton modes and their own BDTs. The Run 1 results, using 4.7 fb^{-1} at $\sqrt{s}$ =7 TeV and 20.3 fb^{-1} at $\sqrt{s}$ =8 TeV, gave signal strengths of

$$\mu = -1.61^{+1.22}_{-1.13}(\text{stat.})^{+0.87}_{-0.92}(\text{syst.})$$

$$\mu = 0.65^{+0.33}_{-0.32}(\text{stat.})^{+0.28}_{-0.24}(\text{syst.})$$

respectively. Combining these results with the $36.1 fb^{-1}$ at $\sqrt{s}$ =13 TeV of Run 2 results gives a signal strength of

$$\mu = 0.90^{+0.18}_{-0.18}(\text{stat.})^{+0.21}_{-0.19}(\text{syst.}).$$

The S/B distribution resulting from combining the full Run 1 and Run 2 data can be found in Figure 43.

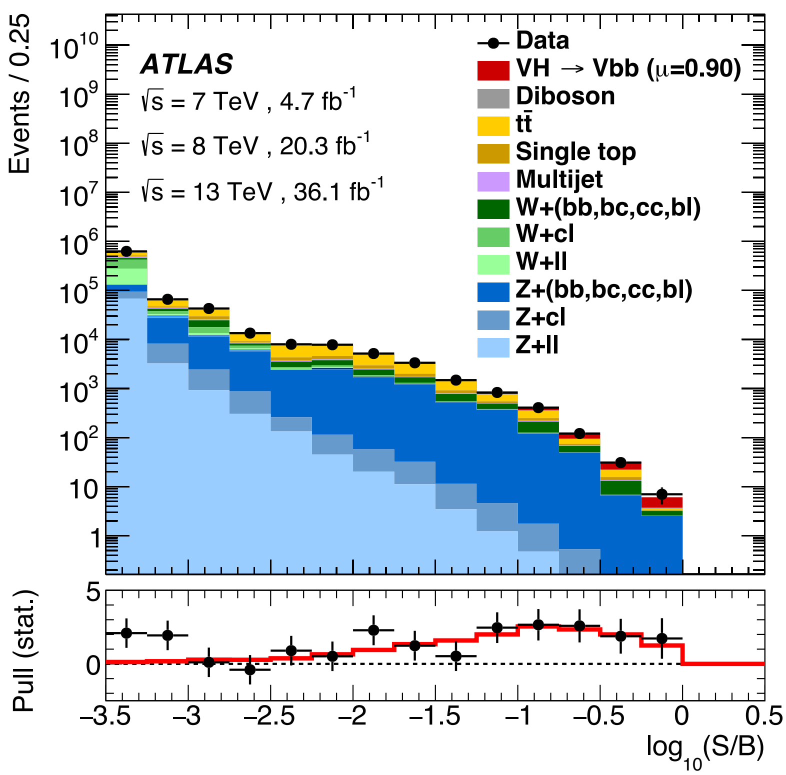


Figure 43: The combined signal regions event yields shown via log(S/B) for combined VH modes with both Run 1 and Run 2 data with VH rescaled to the signal strength extracted from data. In the bottom ratio plot, the pull (residual divided by its uncertainity) is shown compared to the background-only prediction with statistical uncertainties only.

8 Future Outlook

The standard hallmark of discovery of a particle or process in high energy experimental particle physics is typically five standard deviations from the null hypothesis. Evidence is marked at three standard deviations, indicating a higher signal significance above most statistical fluctuations, yet not sufficient enough for cautious physicists to claim a discovery. The evidence standard represents a 1 in 740 chance that the background has fluctuated to give the data seen without a signal present, while the discovery standard represents a 1 in 350 000 000 chance of the same.

Evidence has been shown for $H \to b\bar{b}$ in the associated vector boson production mode [74]. The possibility of discovery within this mode by the end of the current run of the LHC is discussed within this chapter, as well as further properties that are interesting to measure as a consequence of the evidence this analysis provides.

8.1 Projection at End of Run 2

The first question is whether a five standard deviations result can be reasonably expected by the end of Run 2. The answer depends on the uncertainty limits. The smaller the uncertainty, the larger the standard deviations to be quoted on the result.

Expected uncertainties on the signal strength of μ as per our current analysis techniques discussed in this dissertation can be broken down as follows: statistical, signal theory and remaining systematics. The current collected data and understanding of systematic uncertainties give expected best uncertainty limits. These limits were calculated by running a conditional fit (not on actual data) which gives

$$\mu = 1.00^{+0.41}_{-0.36} = 1.00^{+0.24}_{-0.23}(\text{stat.})^{+0.19}_{-0.11}(\text{theory})^{+0.26}_{-0.25}(\text{remaining syst.}).$$

Run 2 of the LHC is scheduled to last mid-way through 2018 with an estimated integrated luminosity of 150 fb^{-1}. Compared to the integrated luminosity used in this analysis, the total statistics collected in Run 2 are anticipated to be approximately five times as much data, corresponding to approximately two times improvement in the statistical uncertainty. The analysis as it currently stands is not predominantly statistically limited. This means improved statistical uncertainty will not significantly reduce the overall signal strength uncertainty.

The theoretical uncertainties associated with the signal MC will change as the probed phase space is altered with increasing luminosity. For example, the 2 Lepton high p_T regime will become more important and therefore the overall uncertainty associated with the signal will be decreased because this regime has less theoretical uncertainty.

The expected statistical and systematic precision on the signal strength at the end of Run 2 is

$$\mu = 1.00^{+0.30}_{-0.25} = 1.00^{+0.12}_{-0.11}(\text{stat.})^{+0.17}_{-0.11}(\text{sigth.})^{+0.20}_{-0.19}(\text{remaining syst.}).$$

The expected significances for the current analysis' luminosity and the total run 2 luminosity are: 3.0 standard deviations and 4.4 standard deviations respectively. While the current analysis benefited from an excess of observed events compared to the expectation, resulting in the equivalent of 3.5 standard deviations being observed, this is not constant between runs and cannot be relied upon to occur again. A 4.4σ result would give more credibility than the current 3.5σ result but does not reach the discovery threshold.

Due to the analysis being primarily systematic uncertainty limited, the best method towards a discovery is to find ways to reduce the systematic uncertainties. Without significant advances to the analysis' sensitivity or reduction of the systematic uncertainties a 5σ discovery is unlikely in Run 2.

8.2 Long Term Improvements

In order to reach the five standard deviations benchmark by the end of Run 2, the analysis' significance must be increased and/or the systematic uncertainties must be decreased. There are multiple ideas currently being investigated with the hope that they will have enough significant impact to warrant their inclusion in the next iteration of this analysis at ATLAS. The most promising ideas are discussed below.

One avenue for improving the analysis that has been pursued in similar analyses is the boosted modes. This has been investigated in exotic/non-SM searches as well as in other experiments. Searching for heavier Higgs partners in ATLAS in the $H \to b\bar{b}$ mode has been done at the current center of mass energy with no evidence found [39]. The strategy for these exotic searches can easily transfer to the SM analysis. Alternatively, CMS has published a direct search for SM $H \rightarrow b\bar{b}$ where the boosted jet reconstructs the $b\bar{b}$ pair. The analysis observed a 1.5σ local significance compared to the expected 0.7σ and had a signal strength of $2.3^{+1.8}_{-1.6}$ [75]. This result shows that boosted topologies are promising.

Boosted searches consist of a singular fat jet to infer the Higgs boson instead of two resolved signal jets. A fat jet is denoted as such because its radius parameter, R, is typically greater than or equal to 1.0 compared to the typical ATLAS jet's $R = 0.4$ For higher p_T jets which are close together in ΔR, identifying the decay products through jet substructure in a fat jet can increase the sensitivity. In the final state of this analyis, as the center of mass energy increases, the average Higgs boson's transverse momentum increases and therefore its decay products, the $b\bar{b}$, will be more collimated. These jets are more fully reconstructed as merged b-tagged jets due to the high energy high p_T.

Another idea is to use Particle Flow (PF) jets. PF jets are an alternative to the calorimeter based jets used in this analysis. The algorithm to reconstruct PF jets infers individual particles by combining information from both the tracker (for charged particles) and the calorimeter (for neutral particles). The PF jets exhibit improved energy and angular resolution compared to solely topo-clustered, otherwise known as calorimeter based jets. PF jets outperform track based jet discriminants in reducing jets due to pile-up. By using PF instead of the current topo-clustered jets, the systematic uncertainties should be reduced [76].

Additional signal acceptance could be attained by explicitly looking for τ lepton decays. Searching for the final states $WH \rightarrow \tau\nu b\bar{b}$ or $ZH \rightarrow \tau^+\tau^- b\bar{b}$ with hadronic τ decays gives access to unexplored phase space. Hadronic τ lepton decays occur almost 60% of the time and can be identified accurately with a 60% efficiency. By adding the hadronic τ to the final states, a gain of over 10% in signal acceptance could be achieved.

Probable enhanced sensitivity is likely to come from switching to continuous b-tagging. In the current iteration of the analysis a b-tagging algorithm is applied with a flat 70% efficiency working point resulting in a binary decision of whether the jet is b-tagged or not. In continuous b-tagging weights are applied to allow reconstructed jets to be associated with a higher efficiency the more their traits are aligned with those of b-tagged jets. This extra information should prove valuable as additional input for the BDT algorithm and provide better distinguishing power between b, c and light jets and therefore better signal versus background separation.

Hopefully one or more of these ideas will increase the sensitivity of the analysis and decrease the uncertainties enough to reach a 5σ discovery result before the end of Run 2.

8.3 Additional Measurements

The current analysis can lead the path for interesting measurements of SM properties such as the cross section of the Higgs boson in its favorite decay mode, $H \rightarrow b\bar{b}$. The signal strength was extracted from the current data available at $\sqrt{s} = 13$ TeV and translated into production cross sections and therefore the branching ratio of the decay. These values are interesting because a large contingent of BSM physics relies on a portion of the Higgs' decay being re-assigned to exotic final state signatures. Since the $b\bar{b}$ branching ratio is the theoretical largest at $m_H = 125$ GeV, $\approx 58\%$, it has the largest phase space available to give to BSM processes. More precise measurement of the cross section would be valuable in quantifying the $b\bar{b}$ branching ratio and identifying if it is consistent with SM predictions or not.

Similarly, the cross check validation done for VZ with $Z \rightarrow b\bar{b}$ is also a valuable measurement. Measuring the cross sections of these processes is another good check of SM predictions and these cross sections can also help by providing valuable input to hadronic MC models.

9 Conclusion

A search for the SM Higgs boson produced with an associated W boson and decaying to a pair of $b\bar{b}$ quarks was performed with 36.1 fb^{-1} of data collected from proton-proton collisions at a center of mass energy $\sqrt{s}$ = 13 TeV with the ATLAS detector during Run 2 of the LHC. There is an excess over the SM expected background with an observed significance of 2.30 standard deviations compared to the expected 1.80 standard deviations. The signal strength relative to the SM prediction of a Higgs boson at $m_H = 125$ GeV is $\mu = 1.35^{+0.40}_{-0.39}(\text{stat.})^{+0.56}_{-0.47}(\text{syst.})$.

Additionally, the $H \to b\bar{b}$ decay was searched for in the other two associated vector boson production leptonic modes. Combining the three modes gives the VH result. It was observed at a significance of 3.5 standard deviations compared to the expected 3.0 standard deviations, and at a signal strength of $\mu = 1.20^{+0.24}_{-0.23}(\text{stat.})^{+0.34}_{-0.28}(\text{syst.})$ Like the individual WH search quoted above, the combined search is consistent with the SM prediction of the Yukawa coupling of the Higgs boson to bottom quarks in the SM.

The VH result from Run 2 data was combined with a previous result obtained from Run 1 data (4.7 fb^{-1} at $\sqrt{s} = 7$ TeV and 20.3 fb^{-1} at $\sqrt{s} = 8$ TeV) [68]. The Run 1 + Run 2 combination shows an excess of events over a SM background only hypothesis with a significance of 3.6 standard deviations compared to the expected 4.0 standard deviations. The resultant signal strength is $\mu = 0.90 \pm 0.18(\text{stat.})^{+0.21}_{-0.19}(\text{syst.})$ which is consistent with the expected signal from the SM.

No deviations from the SM have been observed and evidence for the $H \to b\bar{b}$ decay and the production of VH has been obtained.

Acknowledgments

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently.

We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWFW and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS, CEA-DSM/IRFU, France; SRNSF, Georgia; BMBF, HGF, and MPG, Germany; GSRT, Greece; RGC, Hong Kong SAR, China; ISF, I-CORE and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MNiSW and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; MES of Russia and NRC KI, Russian Federation; JINR; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DST/NRF, South Africa; MINECO, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC, United Kingdom; DOE and NSF, United States of America. In addition, individual groups and members have received support from BCKDF, the Canada Council, CANARIE, CRC, Compute Canada, FQRNT, and the Ontario Innovation Trust, Canada; EPLANET, ERC, ERDF, FP7, Horizon 2020 and Marie Skłodowska-Curie Actions, European Union; Investissements d'Avenir Labex and Idex, ANR, Région Auvergne and Fondation Partager le Savoir, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF; BSF, GIF and Minerva, Israel; BRF, Norway; CERCA Programme Generalitat de Catalunya, Generalitat Valenciana, Spain; the Royal Society and Leverhulme Trust, United Kingdom.

The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK) and BNL (USA), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in Ref. [77].

References

[1] J. J. Thomson. Cathode rays. *Philosophical Magazine 44*, 1897.

[2] S. H. Neddermeyer and C. D. Anderson. Note on the nature of cosmic ray particles. *Phys. Rev. 51*, 1936.

[3] M. L. Perl et al. Evidence for anomalous lepton production in $e^+ - e^-$ annihilation. *Phys. Rev. Lett. 35*, 1975.

[4] C. Patrignani et al. (Particle Data Group). *The Review of Particle Physics.* Chinese Phys. C 40, 2017.

[5] R. Davis Jr. Solar neutrinos ii, experimental. *Phys. Rev. Lett. 12*, 1964.

[6] Super-Kamiokande Collaboration. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81(8), 1998.

[7] Daya Bay Collaboration. Observation of electron-antineutrino disappearance at daya bay. *Phys. Rev. Lett.*, 108(17), 2012.

[8] W. Pauli translated by K. Winter. *On the Earlier and More Recent History of the Neutrino (1957).* Cambridge University Press, 1991.

[9] Jr. et al. C. L. Cowan. Detection of the free neutrino: A confirmation. *Science 124*, 1956.

[10] F. Reines and C. L. Cowan Jr. The neutrino. *Nature 178*, 1956.

[11] E. Fermi. Versuch einer theorie der β-strahlen. i (fermi's theorye of beta decay). *American Journal of Physics 36*, 1934.

[12] M. Kobayashi and T. Maskawa. Cp-violation in the renormalizable theory of weak interaction. *Progress of Theoretical Physics*, 1973. .

[13] J. Griffin and M. Griffin. Richard feynman: A life in science. *Viking Press*, 1997.

[14] E. L. Hill. Hamilton's principle and the conservation theorems of mathematical physics. *Rev. Mod. Phys. 23*, 1953.

[15] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett. 13*, 1964.

[16] P. W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett. 13*, 1964.

[17] S. Weinberg. A model of leptons. *Phys. Rev. Lett. 19*, 1967.

[18] UA1 Collaboration. Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s}$ =540 gev. *Phys. Lett. B*, 1983.

[19] UA2 Collaboration. Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the cern *pp* collider. *Physics Lett. B 122*, 1983.

[20] UA1 Collaboration. Experimental observation of lepton pairs of invariant mass around 95 gev/c^2 at the cern sps collider. *Physics Lett. B*, 1983.

[21] UA2 Collaboration. Evidence for $z^0 \to e^+e^-$ at the cern anti-p p collider. *Phys. Lett. B 129*, 1983.

[22] SLAC E-004 Collaboration. High-energy inesastic $e - p$ scattering at 6° and 10°. *Phys. Rev. Lett.*, 23(16), 1969.

[23] S. L. Glashow, J. Iliopoulos, and L. Maiani. Weak interactions with lepton-hadron symmetry. *Phys. Rev. D. 2*, 1970.

[24] G. S. Abrams et al. Discovery of a second narrow resonance in e^+e^- annihilations. *Phys. Rev. Lett.*, 1974.

[25] S. W. Herb et al. . *Phys. Rev. Lett. 39*, 1977.

[26] CDF Collaboration. Observation of top quark production in $\bar{p}p$ collisions with the collider detector at fermilab. *Phys. Rev. Lett.*, 74(14), 1995.

[27] D∅ Collaboration. Search for high mass top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ tev. *Phys. Rev. Lett.*, 74(13), 1995.

[28] DONUT Collaboration. Observation of tau neutrino interactions. *Phys. Lett. B*, 504, 2001.

[29] M. Gell-Mann. Isotopic spin and new unstable particles. *Physical Review 92*, 1953.

[30] ATLAS Collaboration. Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Phys. Lett. B 716*, 2012. arXiv:1207.7214 [hep-ex].

[31] CMS COllaboration. Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc. *Phys. Lett. B 716*, 2012.

[32] Nobel Media. The nobel prize in physics 2013. nobelprize.org. Accessed: 2017.

[33] R. Panek. *The 4 Percent Universe*. Houghton Mifflin Harcourt, 2011.

[34] A. Fowlie. Cmssm, naturalness and the 'fine-tuning price' of the very large hadron collider. *Phys. Rev. D. 90*, 2014.

[35] LHC Higgs Cross Section Working Group. Cross sections. Accessed: 2017.

[36] ALEPH Collaboration DELPHI Collaboration L3 Collaboration OPAL Collaboration SLD Collaboration LEP Electroweak Working Group SLD electroweak heavy flavour groups. Precision electroweak measurements on the z resonance. *Phys. Rept.*, 2006. arXiv:hep-ex/0509008.

[37] M. Drewes. The phenomenology of right handed neutrinos. *International Journal of Modern Physics E. 22*, 2013.

[38] R. Aaij et al. Test of lepton universality using $b^+ \to k^+\ell^+\ell^-$ decays. *Phys. Rev. Lett.*, 2009.

[39] ATLAS Collaboration. Search for heavy resonances decaying to a w or z boson and a higgs boson in final states with leptons and b-jets in 36.1fb^{-1} of pp collision data at $\sqrt{s} = 13$ tev with the atlas detector. ATLAS-CONF-2017-055, 2017.

[40] J. de Blas, J. M. Lizana, and M. Perez-Victoria. Combining searches of z' and w' boson. *JHEP 01*, 2013. arXiv:1211.2229 [hep-ph].

[41] D. Pappadopulo et al. Heavy vector triplets: Bridging theory and data. *JHEP 09*, 2014. arXiv:1402.4431 [hep-ph].

[42] A. Horvath. The lhc experiments and the preaccelerators. wikimedia.org/wiki/File:LHC.svg, 2003.

[43] ATLAS Collaboration. Atlas public results. twiki.cern.ch/twiki/bin/view/AtlasPublic, 2017. Acessed: 2017.

[44] CERN. Atlas website. atlasexperiment.org/photos/index.html, Accessed: 2017.

[45] ATLAS Collaboration. Atlas detector and physics performance technical design report. Technical report, 1999.

[46] ATLAS Collaboration. Atlas insertable b-layer technical design report. Technical report, 2010.

[47] ATLAS Collaboration. Atlas liquid argon calorimeter phase-i upgrade technical design report. Technical report, 2013.

[48] ATLAS Collaboration. Atlas high-level trigger, data acquisition and controls technical design report. Technical report, 2003.

[49] ATLAS Collaboration. Electrons in atlas: from run 1 to run 2. ATL-PHYS-PROC-2015-180, 2015.

[50] ATLAS Collaboration. Electron reconstruction and identification efficiency measurements in 2016 data. ATL-COM-PHYS-2017-260, 2017.

[51] ATLAS Collaboration. Muon reconstruction performance of the atlas detector in proton-proton collision data at $\sqrt{s} = 13$ tev. *Eur. Phys. J. C 76*, 2016. arXiv:1207.7214 [hep-ex].

[52] S. Moretti Y. Dokshitzer, G. Leder and B. Webber. Better jet clustering algorithms. *JHEP*, 9708:001, 1997.

[53] M. Cacciari et al. The anti-k_t jet clustering algorithm. arXiv:0902.1189 [math.FA], .

[54] M. LeBlanc D. Melini B. Nachman M. Nelson J. Alberto A. Pozo M. Swiatlowski and M. Vos. Jet reclustering and close-by effects in atlas run ii. ATL-COM-PHY-2016-1703.

[55] ATLAS Collaboration. Performance of b-jet identification in the atlas experiment. *Journal of Instrumentation*, 11, 2016.

[56] ATLAS Collaboration. Expected performance of the atlas experiment - detector, trigger and physics. arXiv:0901.0512 [hep-ex], 2008.

[57] ATLAS Collaboration. Performance of missing transverse momentum reconstruction in atlas with 2011 proton-proton collisions at $\sqrt{s} = 7$ tev. ATLAS-CONF-2012-101, 2012.

[58] GEANT4 Collaboration. Geant4: A simulation toolkit. *Nucl. Instrum. and Methods A*, 506, 2003.

[59] T. Sjostrand S. Mrenna and P. Skands. Pythia 6.4 physics and manual. *JHEP 0605*, 2006. arXiv:hep-ph/0603175.

[60] R. Ball et al. Parton distributions for the lhc run ii. arXiv:1410.8849v4v4 [hep-ph], .

[61] S. Frixione, P. Nason, and C. Oleari. Matching nlo qcd computations with parton shower simulations: the powheg method. *GEF-TH-21*, 2007. arXiv:0709.2092.

[62] K. Hamilton, P. Nason, and G. Zanderighi. Minlo: Multi-scale improved nlo. arXiv:1206.3562 [cs.OH].

[63] ATLAS Collaboration. Measurement of the underlying event in jet events from 7 tev proton-proton collisions with the atlas detector. *Eur. Phys. J. C 74*, 2014. arXiv:1406.0392 [hep-ex].

[64] Sherpa Team. *Sherpa 2.2.1 Manual*, 2017.

[65] T. Przedzinski e. Richter-Was N. Davidson, G. Nanava and Z. Was. *Universal Interface of TAUOLA*, 2010. arXiv:1002.0543v1.

[66] ATLAS Collaboration. Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ tev using the atlas detector. *Eur. Phys. J. C 76*, 2016. arXiv:1510.03823 [hep-ex].

[67] ATLAS Collaboration. Expected performance of the atlas b-tagging algorithms in run 2. ATL-PHYS-PUB-2015-022, 2015.

[68] ATLAS Collaboration. Search for the $b\bar{b}$ decay of the standard model higgs boson in associated $(w/z)h$ production with the atlas detector. *JHEP*, 2015.

[69] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, and E. von Toerne. *TMVA 4: Toolkit for Multivariate Data Analysis with ROOT (Users Guide).*

[70] I. Narsky and F. C. Porter. *Statistical Analysis Techniques in Particle Physics: Fits, Density Estimation and Supervised Learning.* Wiley-VCH, 2013.

[71] D. Kirkby W. Verkerke. *RooFit Users Manual*, v2.91 edition, 2008.

[72] K. Cranmer. Practical statistics for the lhc. arXiv.org/pdf/1503.07622 [physics.data-an], 2015.

[73] A. Buckley et al. *RIVET user manual.*

[74] ATLAS Collaboration. Evidence for $h \to b\bar{b}$ with the atlas detector. ATLAS-HIGG-2016-29, 2017.

[75] CMS Collaboration. Inclusive search for the standard model higgs boson produced in pp collisions at $\sqrt{s} = 13$ tev using $h \to b\bar{b}$ decays, 2017. cds.cern.ch/record/2266164.

[76] ATLAS Collaboration. Jet reconstruction and performance using particle flow with the atlas detector. *Eur. Phys. J. C*, 2017. arXiv:1703.10485 [hep-ex].

[77] ATLAS Collaboration. Atlas computing acknowledgements 2016-2017. ATL-GEN-PUB-2016-002.

Appendices

A E_T^{miss}vs μ Trigger Studies

The signal sensitivity expected in the lower luminosity conditions of data taking in 2015, with accompanying lower p_T threshold triggers, can be seen in Table A. The benefit of using E_T^{miss}instead of μ triggers in 1 Lepton μ events has diminishing returns as the gain in sensitivity in 2015 is larger than in the total 36.1 fb^{-1} dataset used in this analysis.

Region	μ only	E_T^{miss}only	μ and E_T^{miss}
2 jets	0.317 ± 0.007	0.333 ± 0.009	0.334 ± 0.009
3 jets	0.145 ± 0.002	0.165 ± 0.002	0.166 ± 0.002

Table 38: The statistical sensitivity of the signal ($\frac{S}{\sqrt{B}}$) in the μ sub-mode with different trigger scenarios for 2015 data.

B Lepton Isolation Definitions

All leptons used in the combined analysis (0,1,2 modes) are identified with the "isLooseTrackOnlyIso" isolation working point, which corresponds to an efficiency of 99% in regions binned in p_T and η for both electrons and muons in the tracking detector. The working point of "isLooseOnlyIso" has the same efficiency requirements as above, with the addition of a 99% efficiency in the calorimeters.

A very important component of performing the data-driven modeling of the multijet background in 1 Lepton is inverting the lepton isolation. The 1 Lepton requirements are intended to find isolated electrons or muons that originated from a W boson decay. This motivates the selected leptons to have relatively high p_T and be well isolated. Leptons in multijet backgrounds, on the other hand, are generally non-prompt and not well isolated. The goal of defining an inverted lepton isolation region is to create a sample of as kinematically similar leptons. The further away the lepton's isolation from the signal lepton's isolation is, the more dissimilar the leptons' kinematics. The ideal sample composes of events with a slightly less well isolated lepton.

The lepton isolation for these selections is dependent on the lepton flavor (e vs μ) and is defined by whether the total energy or momentum present

Working Point	Object	Calorimeter Isolation	Tracker Isolation
Nominal Isolation	e	$E_T < 3.50$ GeV	
	μ	-	$p_T < 1.25$ GeV
Inverted Isolation	e	$E_T > 3.50$ GeV and isLooseTrackOnly	-
	μ	-	$p_T > 1.25$ GeV and isLooseTrackOnly

Table 39: Lepton isolation working point definitions for electrons and muons used to define signal region (nominal) and multijet sampling region (inverted.) The selection is based upon the transverse energy or momentum within a cone surrounding the lepton of size $R = 0.2$.

within a small cone of $R = 0.2$ divided by the lepton's p_T is greater than or less than a certain criteria. This is calculated independently in the tracker and calorimeter systems. Relevant isolation selections to this isolation can be found in Table B.

C Extended Systematic Uncertainties

The W+l and W+cl backgrounds both represent less than 1% of the backgrounds and so a single uncertainty is applied for both backgrounds. The floating normalization of W+HF backgrounds, due to its large uncertainty, is constrained by the W+HF control region. Extrapolation uncertainties between 3-to-2-jets regions and the CR-SR are handled by additional nuisance parameters.

Process	Systematic	Region			
		CR 2 jets	SR 2 jets	CR 3 jets	SR 3 jets
W+1	SysWclNorm	32%			
W+cl	SysWlNorm	37%			
W+HF	norm_Wbb	Floating Normalization			
W+HF	SysWbbNorm_J3	-	-	18%	18%
W+HF	SysWbbNorm_DWhfCR_L1	-	10%	-	10%

Table 40: $W + jets$ normalization systematics and their 1σ effect in various analysis regions. C(S)R stands for the $W + HF$ Control (Signal) Regions, and l stands for the light flavored jets, u, d, s combined.

D Postfit Combined VZ Plots

Additional postfit plots for the VZ BDT analysis in 1 Lepton mode can be found in Figures D and D.

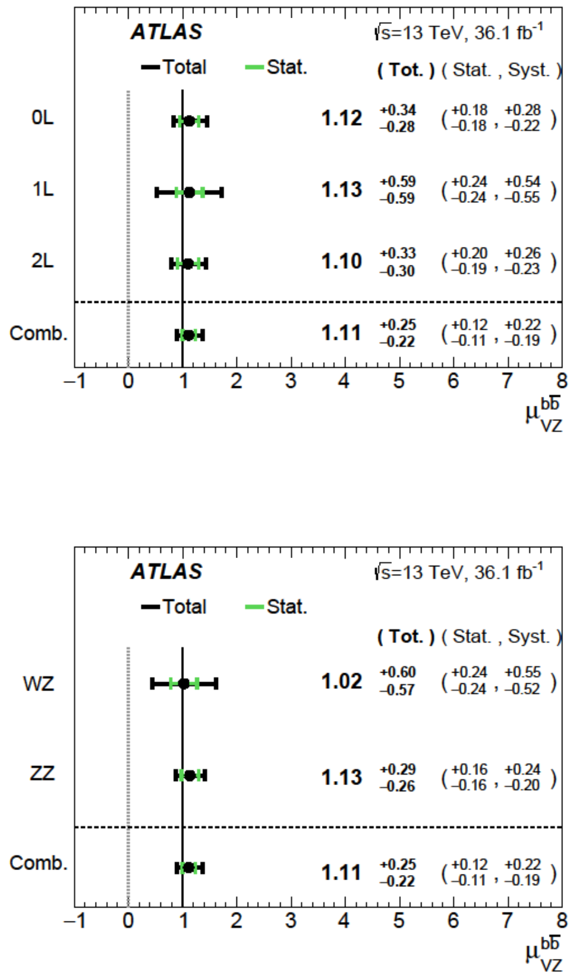


Figure 44: The fitted signal strength compared to the standard model in the 1 Lepton VZ fit decomposed by lepton analysis mode on top, and by the mode of associated vector production below.

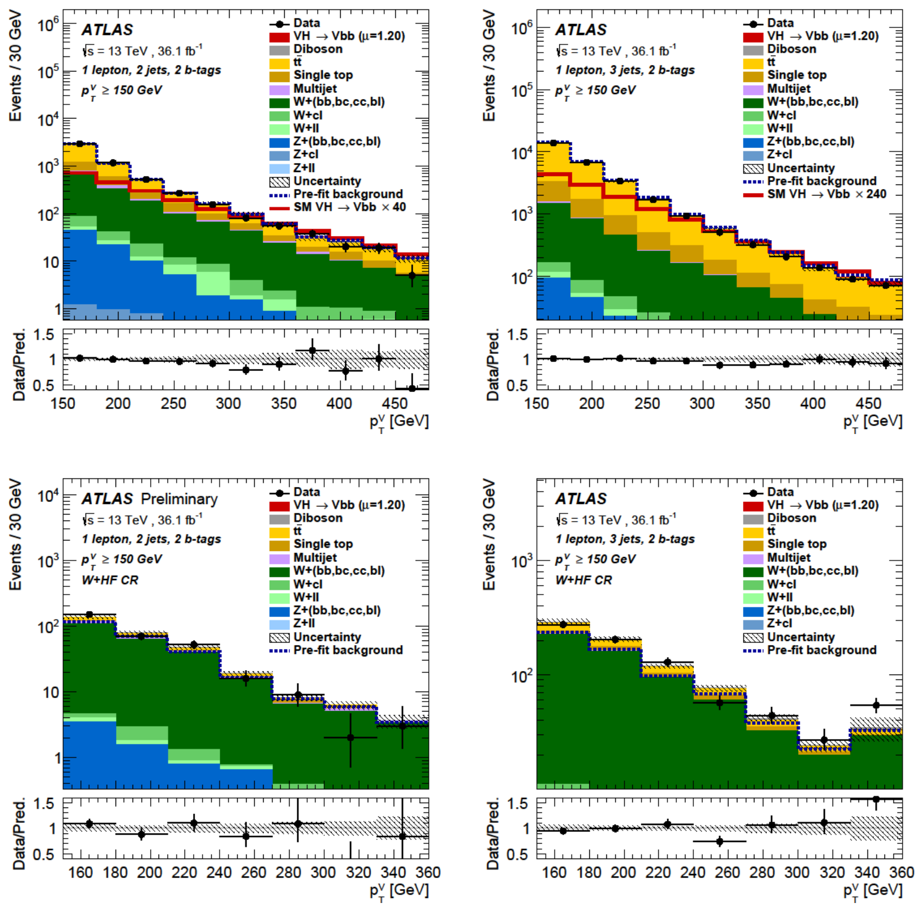


Figure 45: The p_T^V distributions of 1 Lepton mode with the VZ signal scaled to the best fit value and an arbitrarily large scaled signal overlayed in red. On the right are plots in the region with 2 jets, and those on the left have 3 jets. The top two plots show the the signal region while the bottom two show the corresponding control regions.