

Observation and measurement of Higgs
boson decays to WW^* at $\sqrt{s} = 7$ and 8 TeV
and search for Higgs pair production in the
 $HH \rightarrow b\bar{b}b\bar{b}$ channel at $\sqrt{s} = 13$ TeV with the
ATLAS detector at the LHC

A DISSERTATION PRESENTED
BY
TOMO LAZOVICH
TO
THE DEPARTMENT OF PHYSICS

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN THE SUBJECT OF
PHYSICS

HARVARD UNIVERSITY
CAMBRIDGE, MASSACHUSETTS
MAY 2016

©2014 – TOMO LAZOVICH
ALL RIGHTS RESERVED.

Observation and measurement of Higgs boson decays to WW^* at
 $\sqrt{s} = 7$ and 8 TeV and search for Higgs pair production in the
 $HH \rightarrow b\bar{b}b\bar{b}$ channel at $\sqrt{s} = 13$ TeV with the ATLAS detector at
the LHC

ABSTRACT

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

Contents

o INTRODUCTION	I
I Preliminaries	2
I THE STANDARD MODEL AND BEYOND: A THEORETICAL OVERVIEW	3
I.I The Standard Model of Particle Physics	3
I.2 Electroweak Symmetry Breaking and the Higgs	3
I.3 Higgs Boson Production and Decay	3
I.4 Physics Beyond the Standard Model	3
2 THE ATLAS DETECTOR AND THE LARGE HADRON COLLIDER	4
2.1 The Large Hadron Collider	4
2.2 The ATLAS Detector	4
II Observation and measurement of Higgs boson decays to WW^* with the ATLAS detector in LHC Run 1 at $\sqrt{s} = 7$ and 8 TeV	5
3 $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ ANALYSIS STRATEGY	6
3.1 Introduction	6
3.2 The $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ signal in ATLAS	7
3.3 Background processes	8
3.4 Shared signal region selection requirements	12
3.5 Background reduction in same-flavor final states	14
3.6 Parameters of interest and statistical treatment	20
4 THE DISCOVERY OF THE HIGGS BOSON AND THE ROLE OF THE $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ CHANNEL	26
5 OBSERVATION OF VECTOR BOSON FUSION PRODUCTION OF $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$	27
5.1 Introduction	27
5.2 Data and simulation samples	28
5.3 Object selection	32

5.4	Analysis selection	35
5.5	Background estimation	44
5.6	Systematic uncertainties	55
5.7	Results	59
6	COMBINED RUN I $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ RESULTS	64
6.1	Introduction	64
6.2	Results of dedicated gluon fusion $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ search	65
6.3	Signal strength measurements in ggF and VBF production	67
6.4	Measurement of Higgs couplings to vector bosons and fermions	70
6.5	Higgs production cross section measurement	71
6.6	Conclusion	72
III	Search for Higgs pair production in the $HH \rightarrow b\bar{b}b\bar{b}$ channel in LHC Run 2 at $\sqrt{s} = 13$ TeV	73
7	SEARCH OVERVIEW	74
8	SEARCH FOR HIGGS PAIR PRODUCTION IN BOOSTED FINAL STATES	75
9	RESULTS WITH RUN 2 2015 DATASET	76
IV	Looking ahead	77
10	CONCLUSION	78
	REFERENCES	80

Listing of figures

3.1	Branching ratios for a WW system. q refers to quarks. ℓ can be either an electron or muon, and the leptonic branching ratios of the τ are included. For example, the $\ell\nu qq$ final state includes one W decaying to $e\nu$, $\mu\nu$, or $\tau\nu$. τ_h refer to hadronic decays of the τ	8
3.2	Feynman diagram for Standard Model WW production	9
3.3	Feynman diagrams for top pair production (left) and Wt production (right)	10
3.4	An example Feynman diagram of $W + \text{jets}$ production	10
3.5	An example Feynman diagram of $Z + \text{jets}$ production	11
3.6	An illustration of the unique analysis signal regions ¹³	12
3.7	A graphical illustration of the $E_{\text{T},\text{rel}}^{\text{miss}}$ calculation	14
3.8	Predicted backgrounds (compared with data) as a function of n_j (a and b) and n_b (c)	15
3.9	An event display of a $Z/\gamma^* + \text{jets}$ event illustrating the effect of pileup interactions	16
3.10	The RMS of different missing transverse momentum definitions as a function of the average number of interactions per bunch crossing	17
3.11	The difference between the true and reconstructed values of the missing transverse momentum (a) and m_{T} (b) in a gluon fusion signal sample	18
3.12	Comparison of f_{recoil} distributions for $Z/\gamma^* + \text{jets}$, $H \rightarrow WW^*$, and other backgrounds with real neutrinos.	19
3.13	Signal significance as a function of required value for f_{recoil} and $p_{\text{T},\text{rel}}^{\text{miss}(\text{trk})}$ in the ggF $H \rightarrow WW^*$ with $n_j = 0$	20
5.1	A comparison of the subleading lepton p_{T} spectrum between VBF $H \rightarrow WW^*$ production and $t\bar{t}$ background	28
5.2	Leading jet η in VBF $H \rightarrow WW^*$ (red) and $t\bar{t}$ (black)	37
5.3	Distributions of (a) m_{jj} , (b) Δy_{jj} , (c) $C_{\ell 1}$, and (d) $\Sigma m_{\ell j}$, for the VBF analysis. The top panels compare simulation and data, while the bottom panels show normalized distributions for all background processes and signal ¹³	39
5.4	A cartoon of the WW final state. Momenta are represented with thin arrows, spins with thick arrows. ¹³	40
5.5	Event display of a VBF candidate event ¹³	42
5.6	Distributions of $m_{\ell\ell}$ (top left), $\Delta\phi_{\ell\ell}$ (top right), and m_{T} (bottom) for the VBF BDT analysis. These are plotted after all of the BDT pre-training selection cuts ¹³	43

5.7	Distributions of m_{jj} (top left), Δy_{jj} (top right), $\sum C_\ell$ (bottom left), and $\Sigma m_{\ell j}$ (bottom right) for the VBF BDT analysis. These are plotted after all of the BDT pre-training selection cuts ¹³ .	44
5.8	Distributions of O_{BDT} for the VBF signal and associated backgrounds after the VBF pre-training selection ¹³ .	45
5.9	Distributions of m_{jj} (a) and O_{BDT} (b) in the VBF $n_b = 1$ top CR ¹³ .	48
5.10	Comparison of m_{jj} shape in a same flavor $Z \rightarrow \ell\ell$ control region and the VBF cut-based signal region.	49
5.11	General illustration of the ABCD region definitions for $Z/\gamma^* \rightarrow \ell\ell$ background estimation.	50
5.12	Distribution of $m_{\text{T}2}$ in the WW validation region of the VBF analysis ¹³ .	52
5.13	Extrapolation factors for the $W + \text{jets}$ estimate derived for muons (a) and electrons (b) as a function of lepton p_T ¹³ .	54
5.14	Background composition in final VBF signal region ¹³ .	54
5.15	Variations in the top background extrapolation factor in the cut-based analysis due to PDF uncertainties, binned in m_{T} .	57
5.16	Variations in the top background extrapolation factor in the cut-based analysis due to QCD scale uncertainties, binned in m_{T} .	57
5.17	Postfit distributions in the cut-based VBF analysis. Panel (a) shows the one-dimensional m_{T} distribution, while (b) shows the data candidates split into the bins of m_{T} and m_{jj} used in the final fit ¹³ .	61
5.18	Postfit distributions in the BDT VBF analysis ¹³ .	62
5.19	Overlap between cut-based and BDT VBF signal region candidates in the m_{jj} - m_{T} plane.	62
6.1	Post-fit m_{T} distribution in the $n_j \leq 1$ regions ¹³ .	66
6.2	Best fit signal strength $\hat{\mu}$ as a function of hypothesized m_H ¹³ .	68
6.3	Local p_0 as a function of m_H ¹³ .	68
6.4	Likelihood as a function of $\mu_{\text{VBF}}/\mu_{\text{ggF}}$ ¹³ .	69
6.5	Likelihood scan as a function of μ_{VBF} and μ_{ggF} ¹³ .	70
6.6	Likelihood scan as a function of κ_F and κ_V ¹³ .	71

Listing of tables

3.1 A summary of backgrounds to the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ signal	ii
5.1 Single lepton triggers used for electrons and muons. A logical “or” of the triggers listed for each lepton type is taken. Units are in GeV, and the i denotes an isolation requirement in the trigger.	29
5.2 Di-lepton triggers used for different flavor combinations. The two thresholds listed refer to leading and sub-leading leptons, respectively. The di-muon trigger only requires a single lepton at level-i.	29
5.3 Trigger efficiency for signal events and relative gain of adding a dilepton trigger on top of the single lepton trigger selection. The first lepton is the leading, while the second is the sub-leading. Efficiencies shown here are for the ggF signal in the $n_j = 0$ category but are comparable for the VBF signal.	30
5.4 Monte Carlo samples used to model the signal and background processes ¹³	31
5.5 p_T dependent isolation requirements for muons. Muons are required to have the amount of calorimeter or track based cone sums be less than this fraction of their p_T	33
5.6 p_T dependent requirements for electrons. Electrons are required to have the amount of calorimeter or track based cone sums be less than this fraction of their E_T	34
5.7 Event selection for the $n_j \geq 2$ VBF analysis in the 8 TeV cut-based analysis ¹³	41
5.8 Top normalization factors computed at each stage of the cut-based selection. Uncertainties are statistical only.	47
5.9 Top normalization factors computed for each bin of O_{BDT} . Uncertainties are statistical only.	47
5.10 $Z/\gamma^* \rightarrow \tau\tau$ correction factors for the VBF cut-based analysis. Uncertainties are statistical only.	50
5.11 $Z/\gamma^* \rightarrow \ell\ell$ normalization factors for cut-based and BDT analyses. Uncertainties are statistical only.	51
5.12 Systematic uncertainties for various processes in the BDT analysis, given in units of % change in yield. Values are given for the most sensitive BDT bin (bin 3), except where noted with a *, in which case the uncertainty affect the normalization in all BDT bins. Empty entries indicate that the uncertainty is negligible or not applicable to this background.	56
5.13 Systematic uncertainties for various processes in the cut-based VBF analysis, given in units of % change in yield. Values are given for the low m_{jj} signal region.	56

5.14	Composition of the post-fit uncertainties (in %) on the total signal (N_{sig}), total background (N_{bkg}), and individual background yields in the VBF analysis ¹³	59
5.15	Event selection for the VBF BDT analysis. The event yields in (a) are shown after the pre-selection and the additional requirements applied before the BDT classification (see text). The event yields in (b) are given in bins in O_{BDT} after the classification ¹³	60
6.1	All signal regions definitions input into final statistical fit ¹³	65
6.2	Post-fit yields in the different ggF and VBF dedicated signal regions ¹³	66

THIS IS THE DEDICATION.

Acknowledgments

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

0

Introduction

Part I

Preliminaries

1

The Standard Model and beyond: a theoretical overview

1.1 THE STANDARD MODEL OF PARTICLE PHYSICS

1.2 ELECTROWEAK SYMMETRY BREAKING AND THE HIGGS

1.3 HIGGS BOSON PRODUCTION AND DECAY

1.4 PHYSICS BEYOND THE STANDARD MODEL

This is some random quote to start off the chapter.

Firstname lastname

2

The ATLAS detector and the Large Hadron Collider

2.1 THE LARGE HADRON COLLIDER

2.2 THE ATLAS DETECTOR

Part II

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

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

Wernher von Braun

3

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

3.1 INTRODUCTION

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

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

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

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

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

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

Figure 3.1 shows the relative branching fractions for the $H \rightarrow WW^*$ process, calculated from the Particle Data Group values for the W and τ branching ratios²². The largest branching ratio is both W bosons decaying to quark pairs at 45.44%. The next largest is one W decaying leptonically and the other decaying to quarks, a branching ratio of 34.18%. In all cases, ℓ denotes either an electron or muon, and the leptonic branching ratios of the τ are included. For example, the $\ell\nu qq$ final state includes one W decaying to $e\nu$, $\mu\nu$, or $\tau\nu$. In the case of the $W \rightarrow \tau\nu$ decay, the τ lepton then decays to an electron or

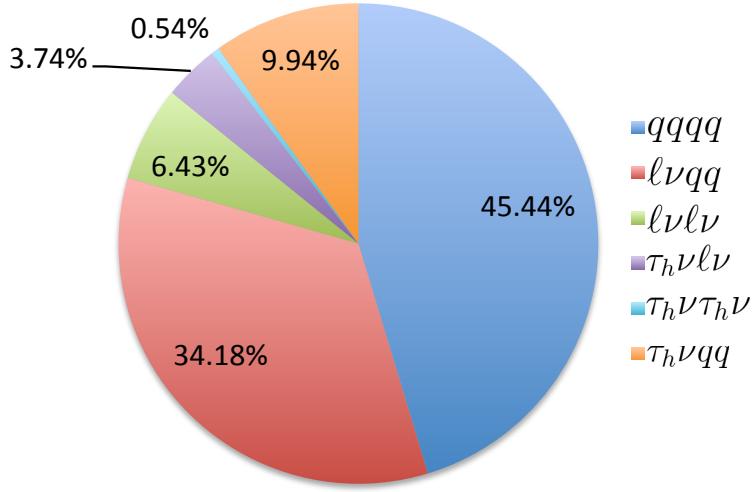


Figure 3.1: Branching ratios for a WW system. q refers to quarks. ℓ can be either an electron or muon, and the leptonic branching ratios of the τ are included. For example, the $\ell\nu qq$ final state includes one W decaying to $e\nu$, $\mu\nu$, or $\tau\nu$. τ_h refer to hadronic decays of the τ .

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

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

3.3 BACKGROUND PROCESSES

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

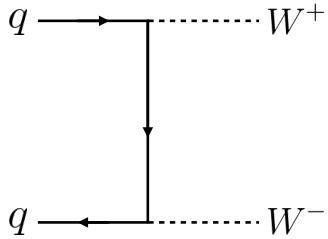


Figure 3.2: Feynman diagram for Standard Model WW production

3.3.1 STANDARD MODEL WW PRODUCTION

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

3.3.2 TOP QUARK PRODUCTION

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

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

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

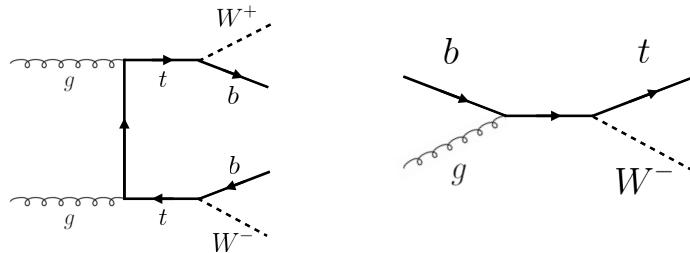


Figure 3.3: Feynman diagrams for top pair production (left) and Wt production (right)

3.3.3 W +JETS BACKGROUND

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

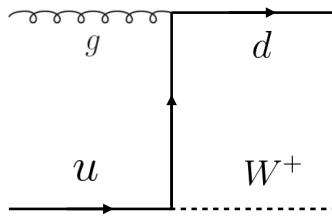


Figure 3.4: An example Feynman diagram of W +jets production

3.3.4 Z/γ^* +JETS BACKGROUND

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

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

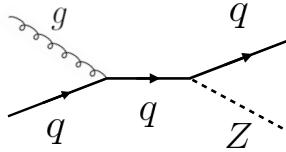


Figure 3.5: An example Feynman diagram of $Z + \text{jets}$ production

3.3.5 OTHER (SUBDOMINANT) BACKGROUNDS

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

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

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

3.4 SHARED SIGNAL REGION SELECTION REQUIREMENTS

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

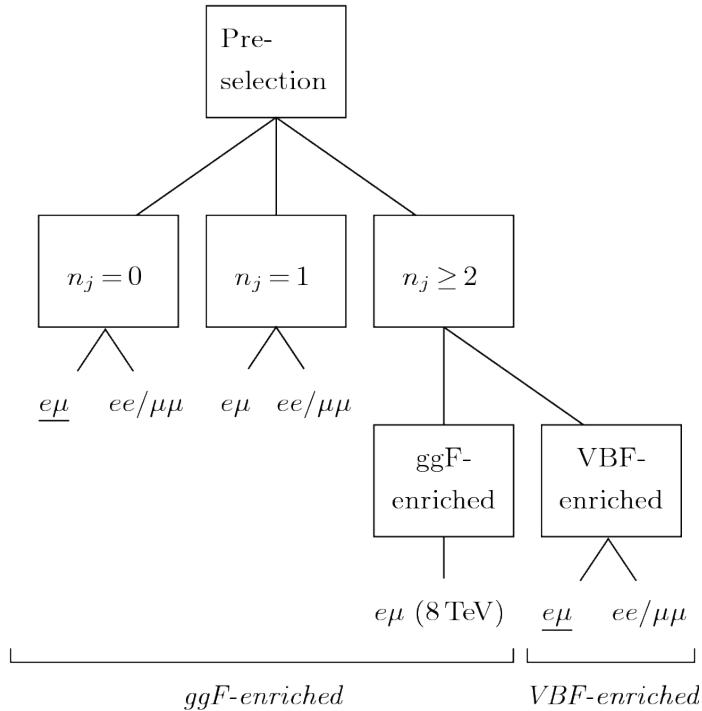


Figure 3.6: An illustration of the unique analysis signal regions ¹³

3.4.1 EVENT PRE-SELECTION

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

quired.

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

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

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

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

3.4.2 JET MULTIPLICITY

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

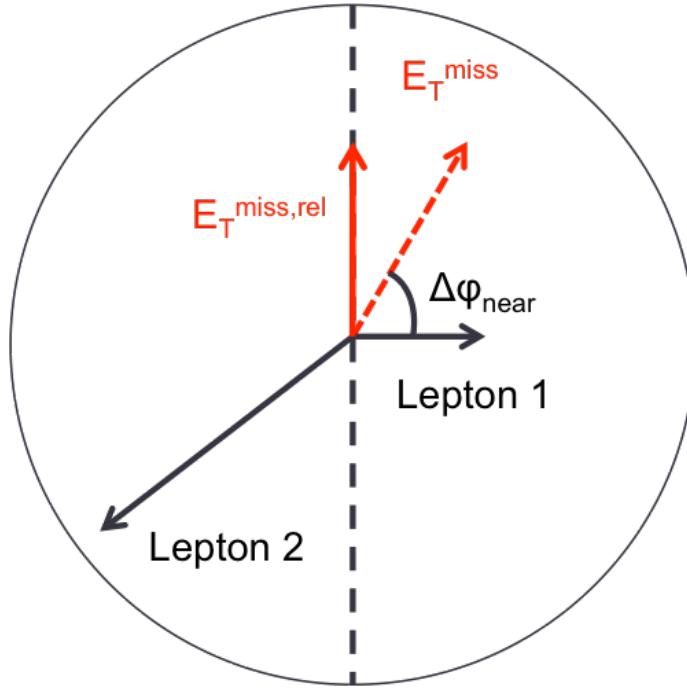


Figure 3.7: A graphical illustration of the $E_{T,\text{rel}}^{\text{miss}}$ calculation

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

3.5 BACKGROUND REDUCTION IN SAME-FLAVOR FINAL STATES

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

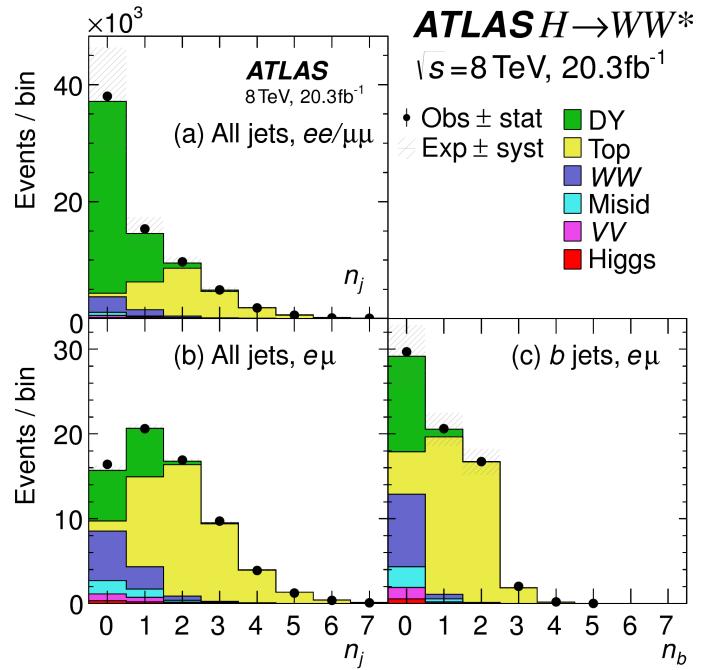


Figure 3.8: Predicted backgrounds (compared with data) as a function of n_j (a and b) and n_b (c)

3.5.1 PILEUP AND E_T^{miss} RESOLUTION

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

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

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

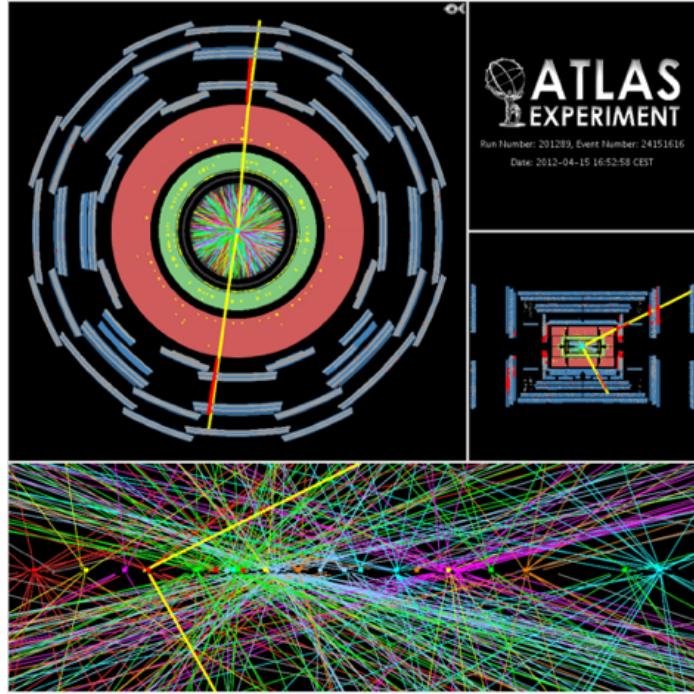


Figure 3.9: An event display of a Z/γ^* + jets event illustrating the effect of pileup interactions

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

3.5.2 TRACK-BASED DEFINITIONS OF MISSING TRANSVERSE MOMENTUM

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

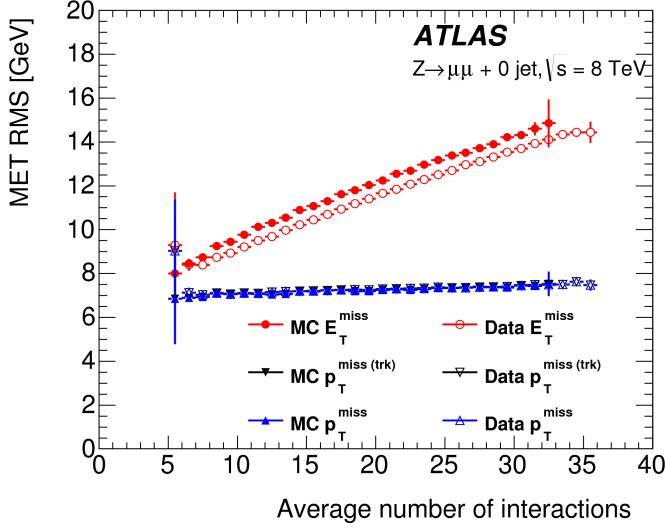


Figure 3.10: The RMS of different missing transverse momentum definitions as a function of the average number of interactions per bunch crossing

tion 3.2.

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

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

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

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

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

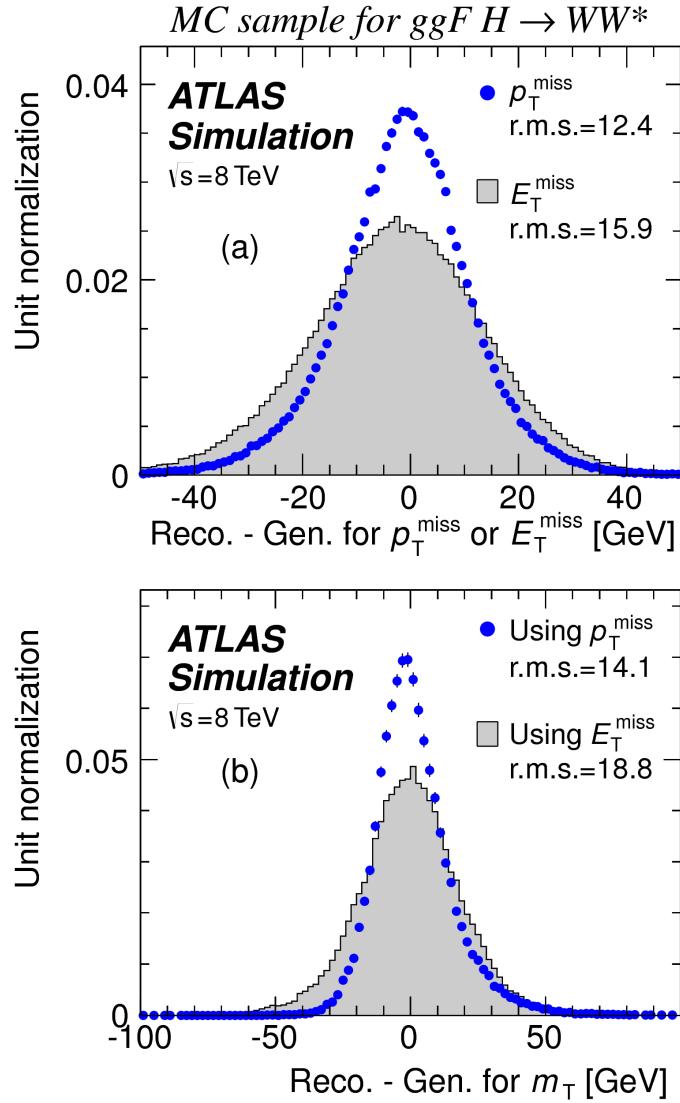


Figure 3.11: The difference between the true and reconstructed values of the missing transverse momentum (a) and m_T (b) in a gluon fusion signal sample

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

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

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

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

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

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

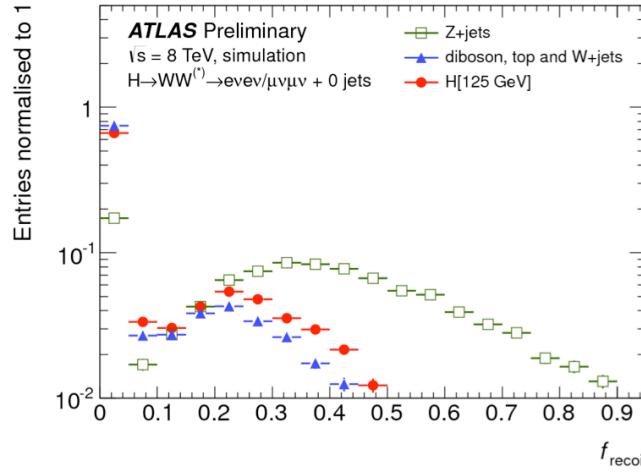


Figure 3.12: Comparison of f_{recoil} distributions for $Z/\gamma^* + \text{jets}$, $H \rightarrow WW^*$, and other backgrounds with real neutrinos.

3.5.4 OPTIMIZING BACKGROUND REDUCTION SELECTION REQUIREMENTS

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

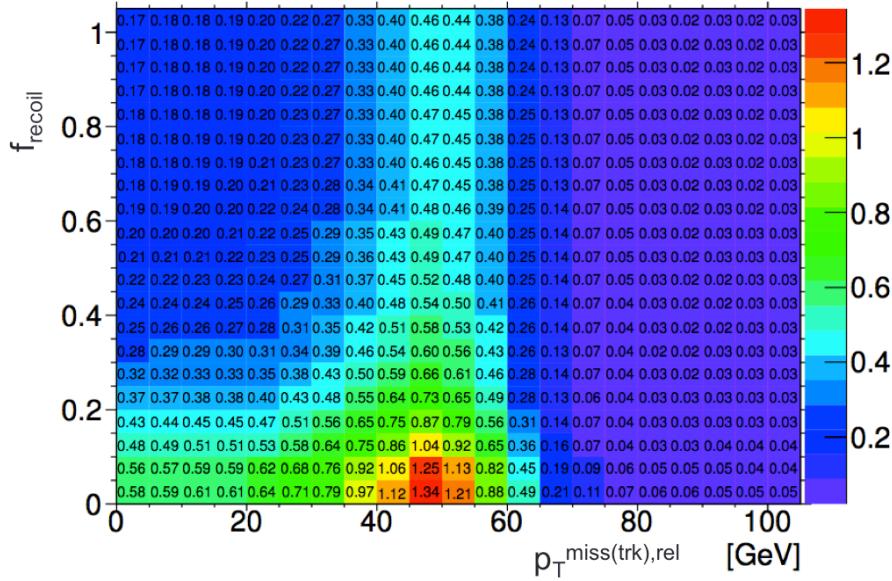


Figure 3.13: Signal significance as a function of required value for f_{recoil} and $p_{T,\text{rel}}^{\text{miss(trk)}}$ in the ggF $H \rightarrow WW^*$ with $n_j = 0$

3.6 PARAMETERS OF INTEREST AND STATISTICAL TREATMENT

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

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

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

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

3.6.1 TRANSVERSE MASS

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

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

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

3.6.2 STATISTICAL TREATMENT^{*}

LIKELIHOOD FUNCTION

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

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

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

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

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

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

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

*Many thanks to Aaron Armbruster, whose thesis⁶ inspired parts of this section.

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

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

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

in equation 3.10.

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

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

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

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

TEST STATISTIC

To distinguish whether the data match a background only or background and signal hypothesis, a test statistic must be used. The $H \rightarrow WW^*$ analysis used the profile likelihood technique¹⁵. The first step in formulating this test statistic is to define the profile likelihood ratio, shown in equation 3.11.

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

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

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

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

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

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

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

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

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

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

Marcel Proust

4

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

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

Richard Feynman

5

Observation of Vector Boson Fusion

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

5.1 INTRODUCTION

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

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

5.2 DATA AND SIMULATION SAMPLES

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

5.2.1 TRIGGERS

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

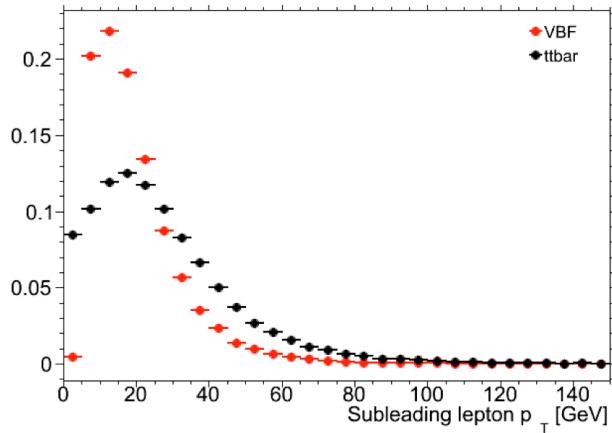


Figure 5.1: A comparison of the subleading lepton p_T spectrum between VBF $H \rightarrow WW^*$ production and $t\bar{t}$ background

As discussed in Chapter 2, there are multiple levels in the ATLAS trigger system, and there are different p_T thresholds imposed for the leptons at each level. Additionally, some triggers have a loose selection on the isolation of the lepton (looser than that applied offline in the analysis object selection). Table 5.1 shows the thresholds used for single lepton triggers, while table 5.2 shows the thresholds coming from di-lepton triggers. The single lepton trigger efficiency for muons that pass the analysis object selection is 70% for muons in the barrel region ($|\eta| < 1.05$) and 90% in the endcap region. The electron trigger efficiency increases with electron p_T but the average is approximately 90%. These efficiencies are measured by combined performance and trigger signature groups^{11,12}.

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

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

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

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

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

5.2.2 MONTE CARLO SAMPLES

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

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

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

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

Table 5.4 shows the MC generators used for the signal and background processes, as well as their cross sections. In order to include corrections up to next-to-leading order (NLO) in the QCD coupling constant α_s , the `POWHEG`²¹ generator is often used. In some cases, only leading order generators like `ACERMC`⁹ and `GG2VV`¹⁸ are available for the process in question. If the process requires good modeling for very high parton multiplicities, the `SHERPA`¹⁷ and `ALPGEN`²⁰ generators are used to provide merged calculations for five or fewer additional partons. These matrix element level calculations must then be additionally matched to models of the underlying event, hadronization, and parton shower. There are four possible generators for this: `SHERPA`, `PYTHIA 6`²⁶, `PYTHIA 8`²⁷, or `HERWIG`¹⁴ + `JIMMY`¹⁰. The simulation additionally requires an input parton distribution function (PDF). The `CT10`¹⁶ PDFs are used for `SHERPA` and `POWHEG` simulated samples, while `CTEQ6L1`²³ is used for `ALPGEN + HERWIG` and `ACERMC` simulations. The Drell-Yan samples are reweighted to the `MRST`²⁵ PDFs, as these are found to give the best agreement between data and simulation.

Once the basic hard scattering process is simulated, it must be passed through a detector simulation and additional pile-up events must be overlaid. The pile-up events are modeled with `PYTHIA 8`, and the ATLAS detector is simulated with `GEANT4`²⁴. Because of the unique phase space of the $H \rightarrow WW^*$ analysis, events are sometimes filtered at generator level to allow for more efficient generation of relevant events. The efficiency of the trigger in MC simulation does not always match the measured efficiency in data, so trigger scale factors are applied to correct the MC efficiency to the data. These are derived by the

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

Table 5.4: Monte Carlo samples used to model the signal and background processes ¹³.

combined performance groups^{11,12}.

5.3 OBJECT SELECTION

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

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

5.3.1 MUONS

The analysis uses combined muon candidates, where a track in the Inner Detector has been matched to a standalone track in the Muon Spectrometer. The track parameters are combined statistically in the muon reconstruction algorithm⁵. The muons are required to be within $|\eta| < 2.5$ and have a $p_T > 10$ GeV. To reduce backgrounds coming from mis-reconstructed leptons, there are requirements on the impact parameter of the muon relative to the primary vertex. The transverse impact parameter d_0 is required to be small relative to its estimated uncertainty, the exact cut value being $d_0/\sigma_{d_0} < 3$. The longitudinal impact parameter z_0 must satisfy $|z_0 \sin \theta| < 1$ mm.

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

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

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

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

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

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

5.3.2 ELECTRONS

Electrons are identified by matching reconstructed clusters in the electromagnetic calorimeter with tracks in the inner detector. The electrons are identified using a likelihood based method^{4,2} which takes into account the shower shapes in the calorimeter, the matching of tracks to clusters, and the amount of transition radiation in the TRT. The electrons are required to have $|\eta| < 2.47$, and candidates in the transition region between the barrel and endcap ($1.37 < |\eta| < 1.52$) are excluded. As the muons, the electrons are required to have transverse impact parameter significance < 3 , while in the longitudinal direction they must have $|z_0 \sin \theta| < 0.4$ mm. Some electron requirements also vary with electron E_T , and these requirements are summarized in table 5.6.

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

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

5.3.3 JETS

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

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

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

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

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

5.3.4 OVERLAP REMOVAL

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

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

5.4 ANALYSIS SELECTION

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

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

5.4.1 COMMON PRE-SELECTION

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

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

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

5.4.2 CUT-BASED SELECTION

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

GENERAL BACKGROUND REDUCTION

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

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

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

In the different flavor channels, a cut is made to reduce the contamination from $Z \rightarrow \tau\tau$ decays. The di- τ invariant mass, $m_{\tau\tau}$, is constructed by assuming that the neutrinos from the τ decays were collinear with the leptons². The analysis requires that this mass not be consistent with a Z by requiring $m_{\tau\tau} < m_Z - 25 \text{ GeV}$.

VBF TOPOLOGICAL CUTS

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

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

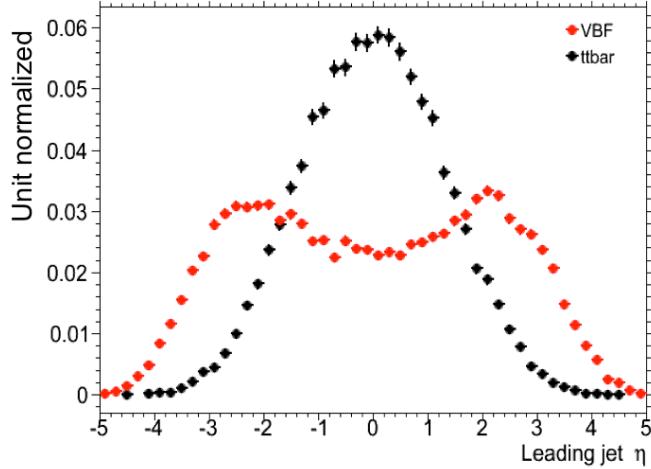


Figure 5.2: Leading jet η in VBF $H \rightarrow WW^*$ (red) and $t\bar{t}$ (black)

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

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

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

equation 5.2.

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

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

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

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

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

HIGGS TOPOLOGICAL CUTS

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

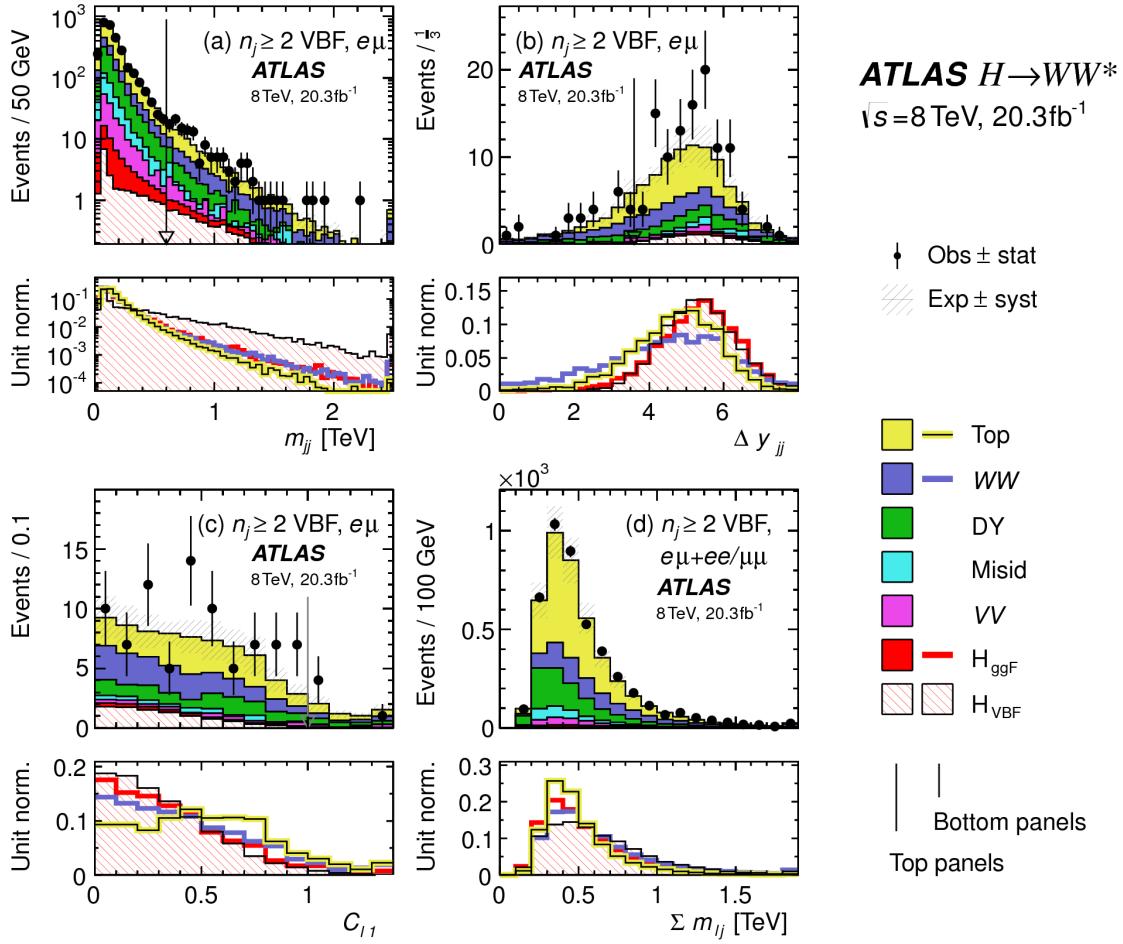


Figure 5.3: Distributions of (a) m_{jj} , (b) Δy_{jj} , (c) $C_{\ell 1}$, and (d) $\sum m_{ij}$, for the VBF analysis. The top panels compare simulation and data, while the bottom panels show normalized distributions for all background processes and signal¹³.

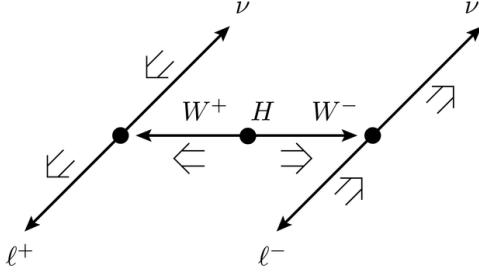


Figure 5.4: A cartoon of the WW final state. Momenta are represented with thin arrows, spins with thick arrows.¹³

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

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

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

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

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

5.4.3 BDT BASED SELECTION

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

Table 5.7: Event selection for the $n_j \geq 2$ VBF analysis in the 8 TeV cut-based analysis⁴³.

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

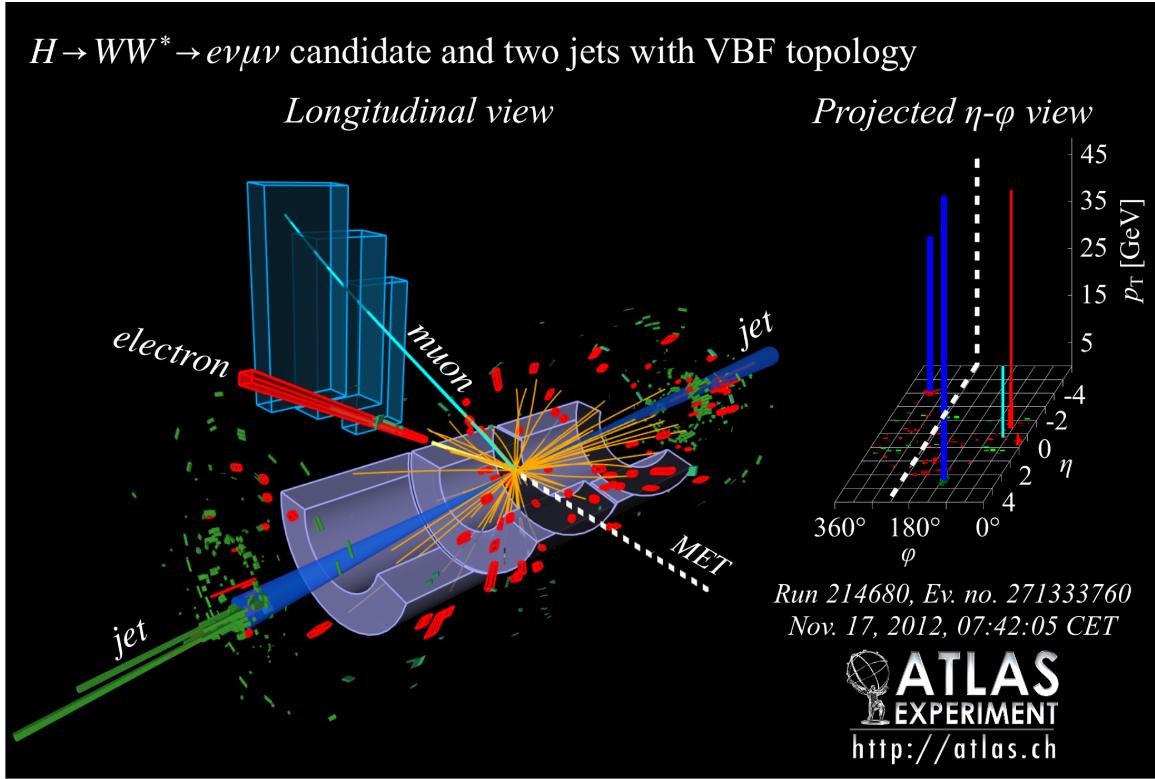


Figure 5.5: Event display of a VBF candidate event ¹³.

$H \rightarrow WW^*$ production.

PRE-TRAINING SELECTION AND BDT INPUTS

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

possible lepton-jet combinations, shown in equation 5.3.

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

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

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

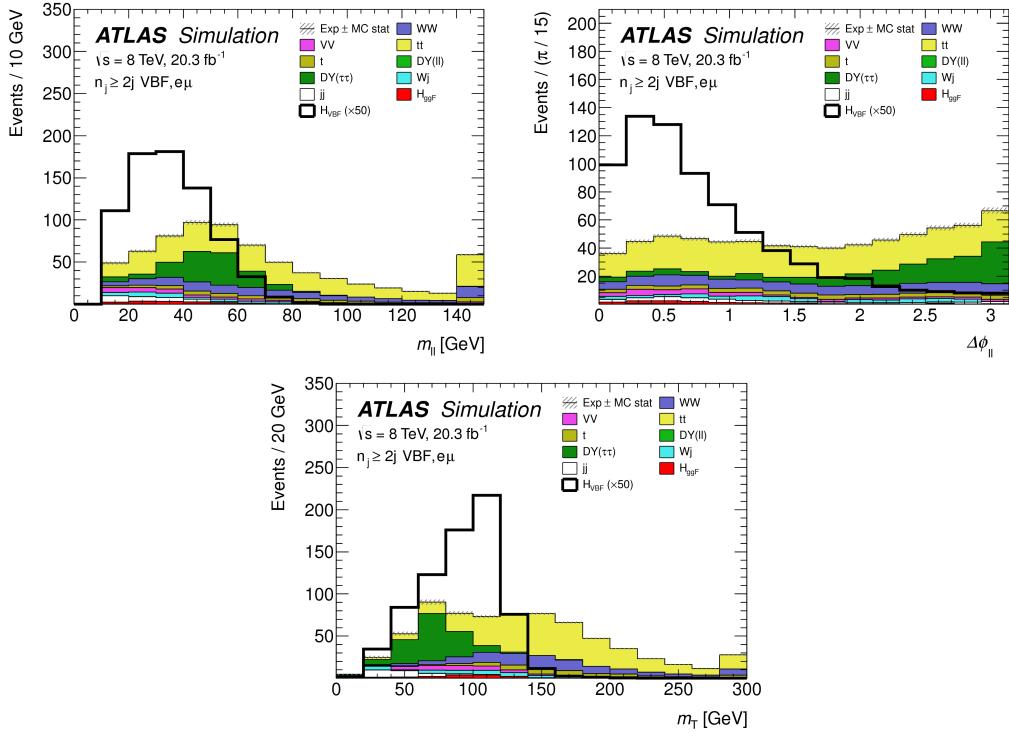


Figure 5.6: Distributions of m_{ll} (top left), $\Delta\phi_{ll}$ (top right), and m_T (bottom) for the VBF BDT analysis. These are plotted after all of the BDT pre-training selection cuts¹³.

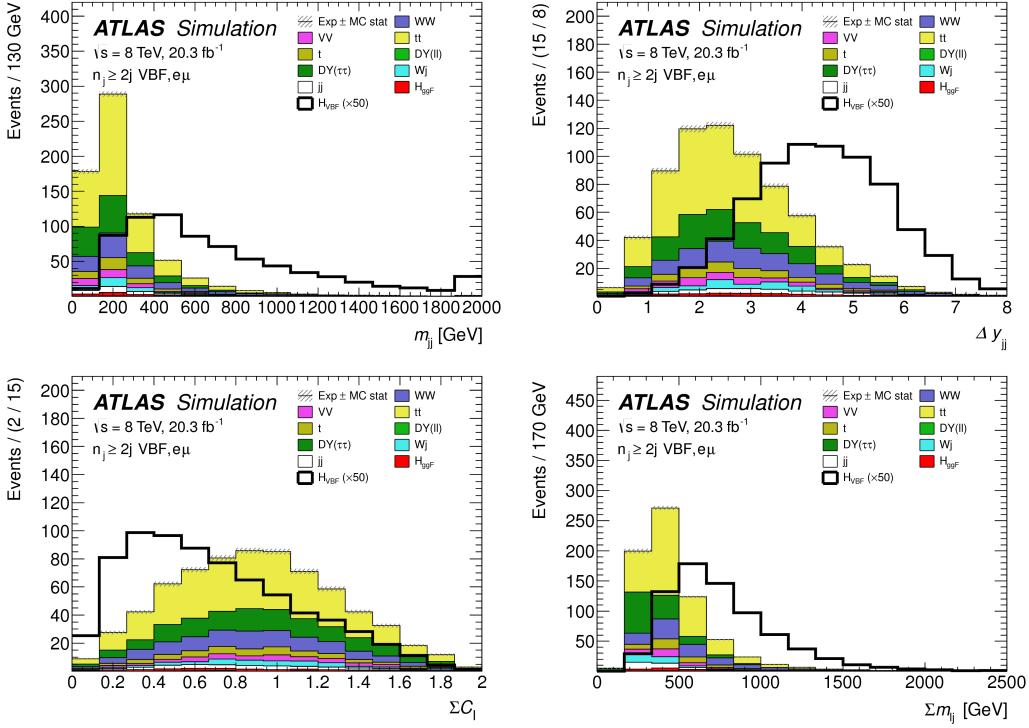


Figure 5.7: Distributions of m_{jj} (top left), Δy_{jj} (top right), $\sum C_\ell$ (bottom left), and $\sum m_{\ell j}$ (bottom right) for the VBF BDT analysis. These are plotted after all of the BDT pre-training selection cuts¹³.

BDT OUTPUT

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

5.5 BACKGROUND ESTIMATION

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

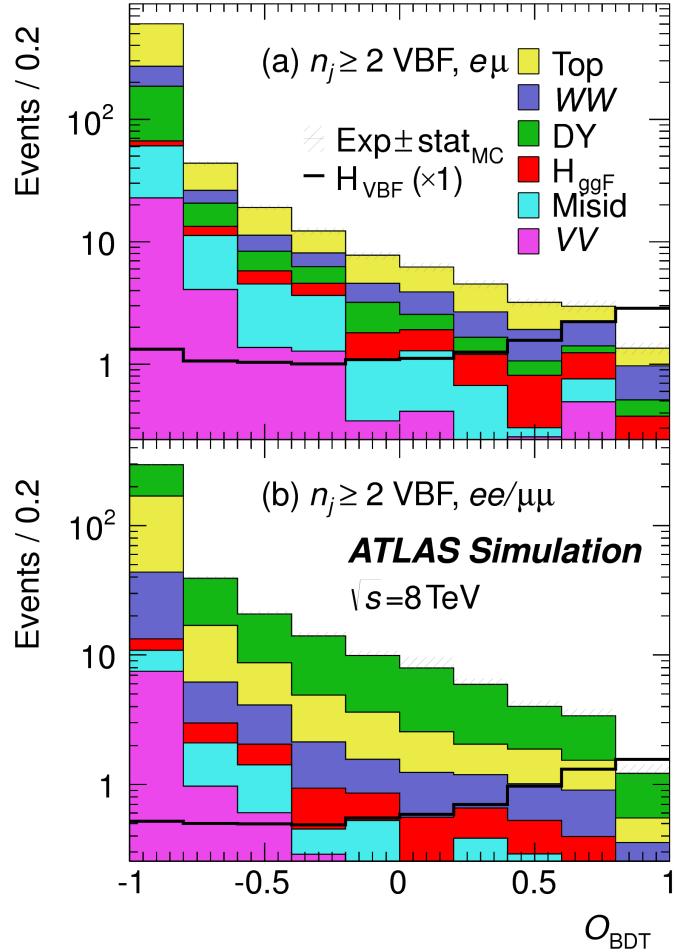


Figure 5.8: Distributions of O_{BDT} for the VBF signal and associated backgrounds after the VBF pre-training selection¹³.

5.5.1 GENERAL STRATEGY

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

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

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

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

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

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

5.5.2 TOP BACKGROUND

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

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

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

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

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

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

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

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

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

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

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

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

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

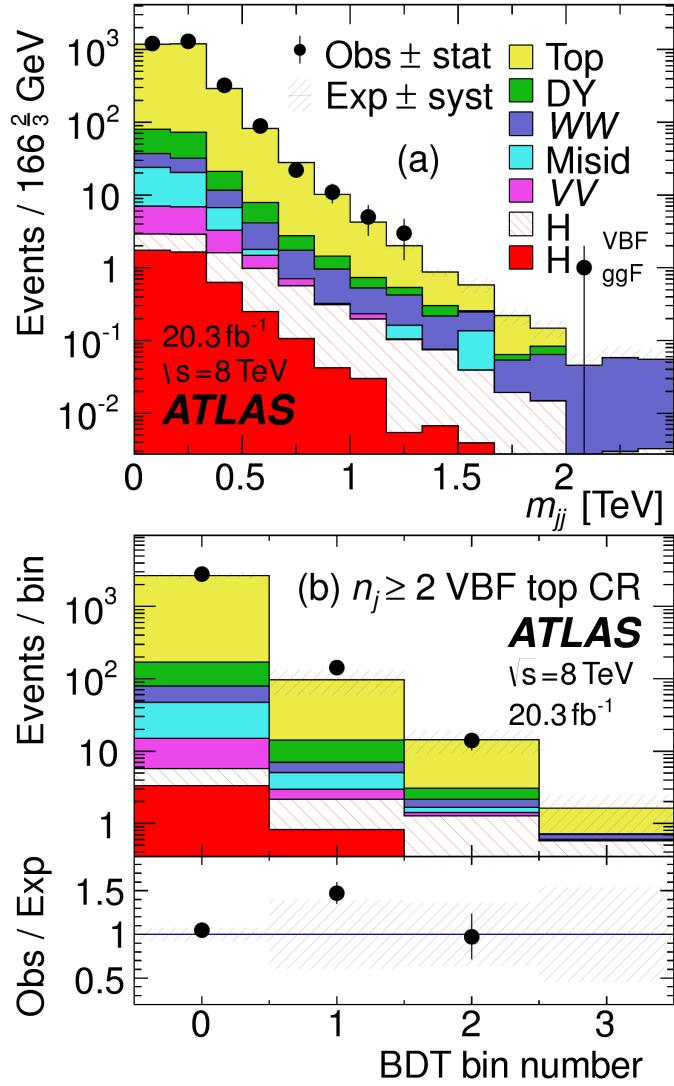


Figure 5.9: Distributions of m_{jj} (a) and O_{BDT} (b) in the VBF $n_b = 1$ top CR¹³.

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

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

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

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

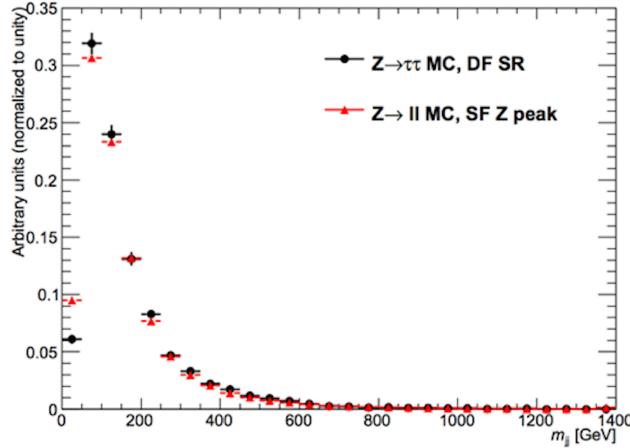


Figure 5.10: Comparison of m_{jj} shape in a same flavor $Z \rightarrow \ell\ell$ control region and the VBF cut-based signal region.

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

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

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

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

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

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

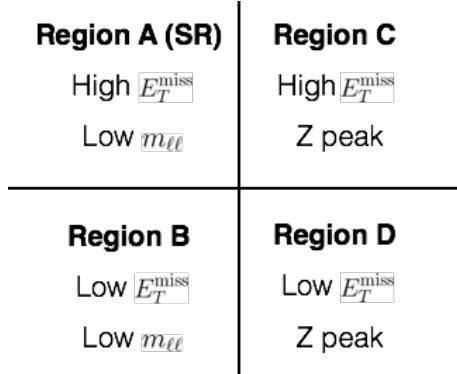


Figure 5.11: General illustration of the ABCD region definitions for $Z/\gamma^* \rightarrow \ell\ell$ background estimation.

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

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

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

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

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

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

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

	$\beta_{\ell\ell}$
BDT Bin 1	1.01 ± 0.15
BDT Bin 2	0.89 ± 0.28
Cut-based	0.81 ± 0.21

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

5.5.5 WW AND OTHER DIBOSON BACKGROUNDS

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

As WW is the largest of these backgrounds and is irreducible, validating the estimate is of particular importance. The validation region is constructed by requiring the pre-selection cuts on leptons and $m_{\ell\ell}$, $n_b = 0$, and $m_T > 100$ GeV. The m_{T2} variable² is an additional discriminant that will isolate the

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

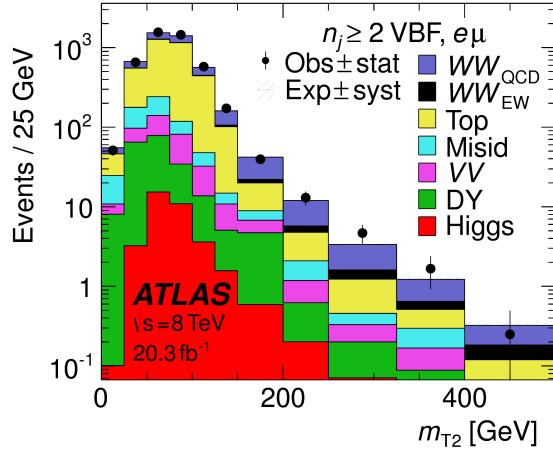


Figure 5.12: Distribution of m_{T2} in the WW validation region of the VBF analysis¹³.

5.5.6 HIGGS PRODUCTION VIA GLUON-GLUON FUSION

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

5.5.7 BACKGROUNDS WITH MISIDENTIFIED LEPTONS

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

$W + \text{jets}$ BACKGROUND

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

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

QCD MULTIJET BACKGROUND

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

5.5.8 BACKGROUND COMPOSITION IN FINAL SIGNAL REGION

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

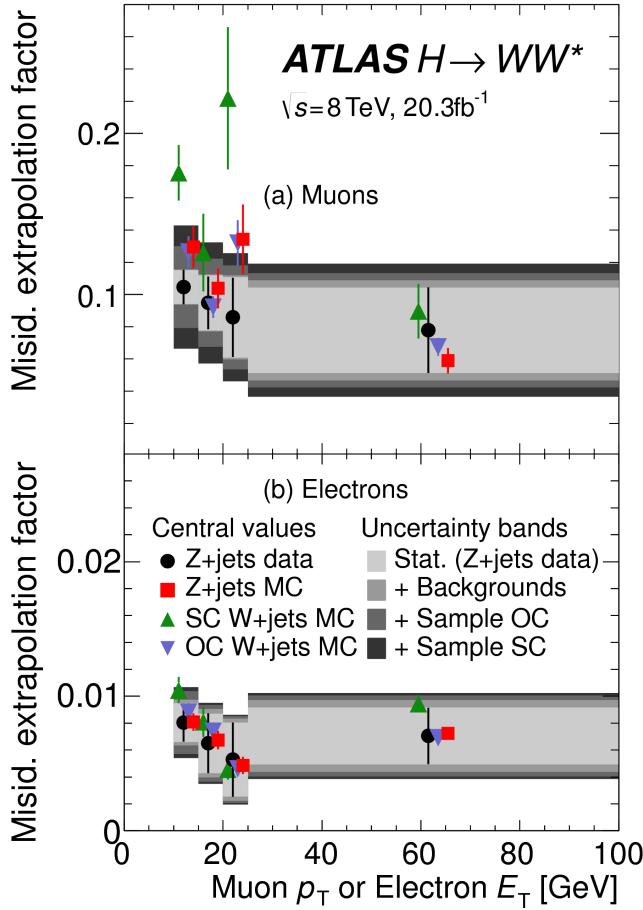


Figure 5.13: Extrapolation factors for the W +jets estimate derived for muons (a) and electrons (b) as a function of lepton p_T ¹³.

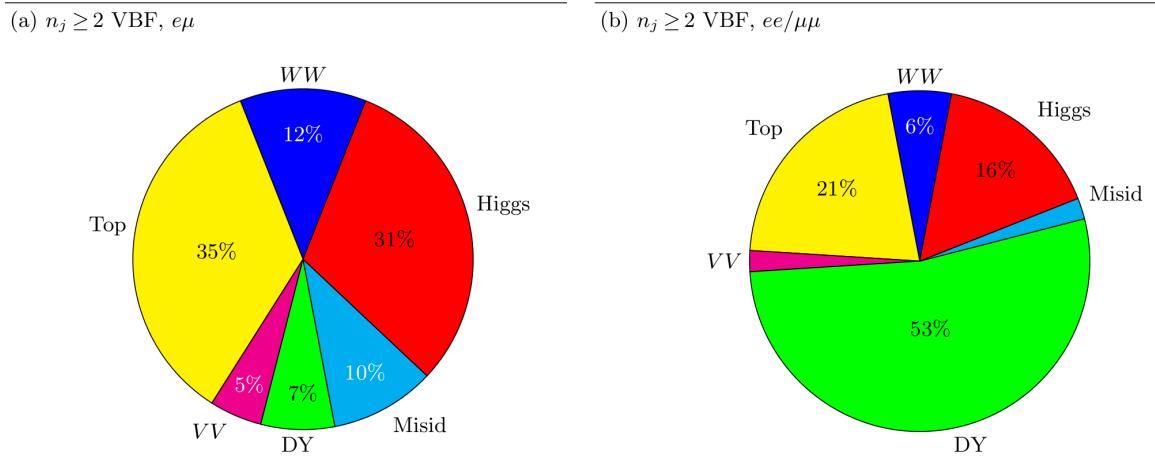


Figure 5.14: Background composition in final VBF signal region¹³.

5.6 SYSTEMATIC UNCERTAINTIES

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

5.6.1 THEORETICAL UNCERTAINTIES

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

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

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

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

uncertainties assigned to the Higgs total production cross section. Then, there are uncertainties assigned based on the fact that the analysis is done in exclusive jet bins and it is possible for signal events to migrate from one bin to the next depending on the presence or absence of jets. These are assigned using the Jet Veto Efficiency (JVE) procedure⁷ for ggF events and the Stewart-Tackmann (ST) method²⁸ for VBF production.

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

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

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

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

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

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

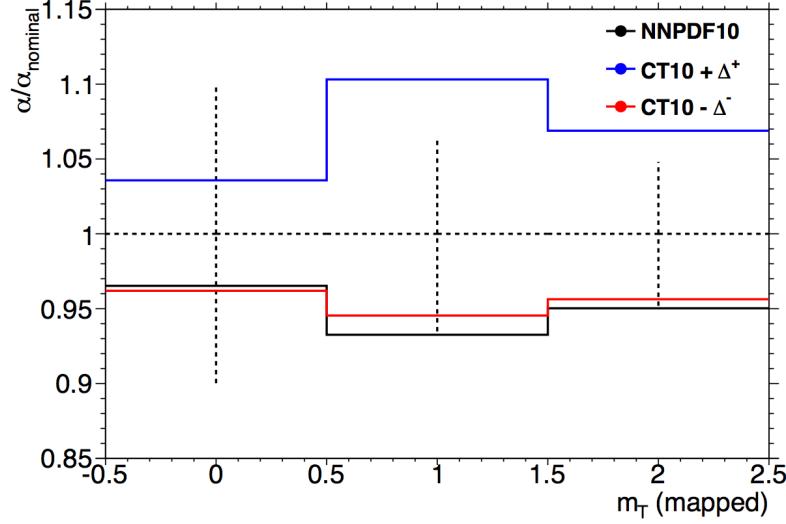


Figure 5.15: Variations in the top background extrapolation factor in the cut-based analysis due to PDF uncertainties, binned in m_T .

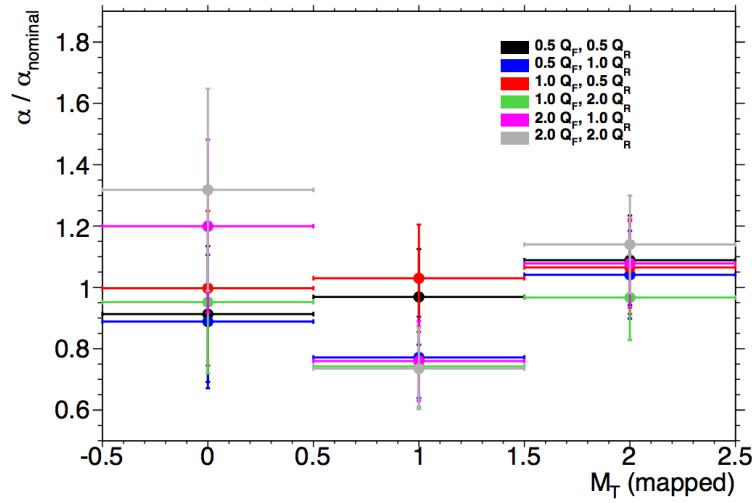


Figure 5.16: Variations in the top background extrapolation factor in the cut-based analysis due to QCD scale uncertainties, binned in m_T .

5.6.2 EXPERIMENTAL UNCERTAINTIES

In this analysis, the theoretical uncertainties end up being the most dominant, but there are some experimental uncertainties that make a contribution as well. The first is the uncertainty on the measured integrated luminosity, which affects backgrounds whose normalization is taken from MC and is measured to be 2.8% in the 8 TeV dataset⁷. The dominant sources of uncertainty overall are uncertainties on the jet energy scale and resolution and the *b*-tagging efficiency. Additional sources include lepton uncertainties on identification, resolution, and trigger efficiency, as well as uncertainties on the missing transverse momentum .

The jet energy scale uncertainty is split into several independent components, including jet-flavor dependent calorimeter response uncertainties, uncertainties on modeling of pile-up interactions, uncertainties on extrapolation from the central to forward detector regions, and MC non-closure³. The uncertainty on energy scale for jets used in this analysis ranges from 1% to 7% depending on the jet p_T and η . The jet energy resolution varies from 5% to 20%, with uncertainties ranging from 2% to 40% (the largest uncertainties occurring at the selection threshold).

The *b*-tagging efficiency is independently measured in data samples enriched in dileptonic decays of $t\bar{t}$ events or in events where a muon is reconstructed in the vicinity of a jet^{1,8}. The efficiencies and their uncertainties are binned in p_T and decomposed into uncorrelated components using an eigenvector method⁷. Uncertainties on the efficiency range from 1% to 7.8%. The uncertainty on the rate of misidentification of *c*-jets as *b*-jets ranges from 6-14%, while the uncertainty on the rate of light jet mis-tagging ranges from 9-19% depending on p_T and η .

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

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

Table 5.14: Composition of the post-fit uncertainties (in %) on the total signal (N_{sig}), total background (N_{bkg}), and individual background yields in the VBF analysis¹³.

5.7 RESULTS

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

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

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

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

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

(a) Before the BDT classification

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

(b) Bins in O_{BDT}

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

Table 5.15: Event selection for the VBF BDT analysis. The event yields in (a) are shown after the preselection and the additional requirements applied before the BDT classification (see text). The event yields in (b) are given in bins in O_{BDT} after the classification¹³.

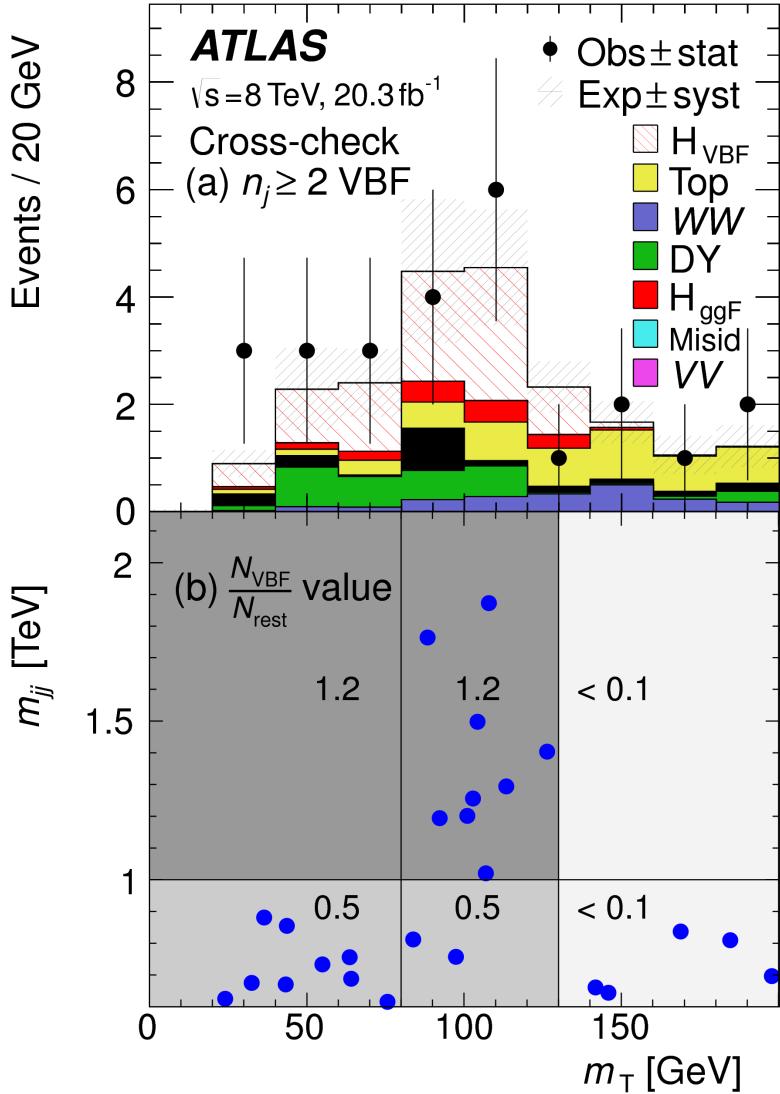


Figure 5.17: Postfit distributions in the cut-based VBF analysis. Panel (a) shows the one-dimensional m_T distribution, while (b) shows the data candidates split into the bins of m_T and m_{jj} used in the final fit¹³.

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

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

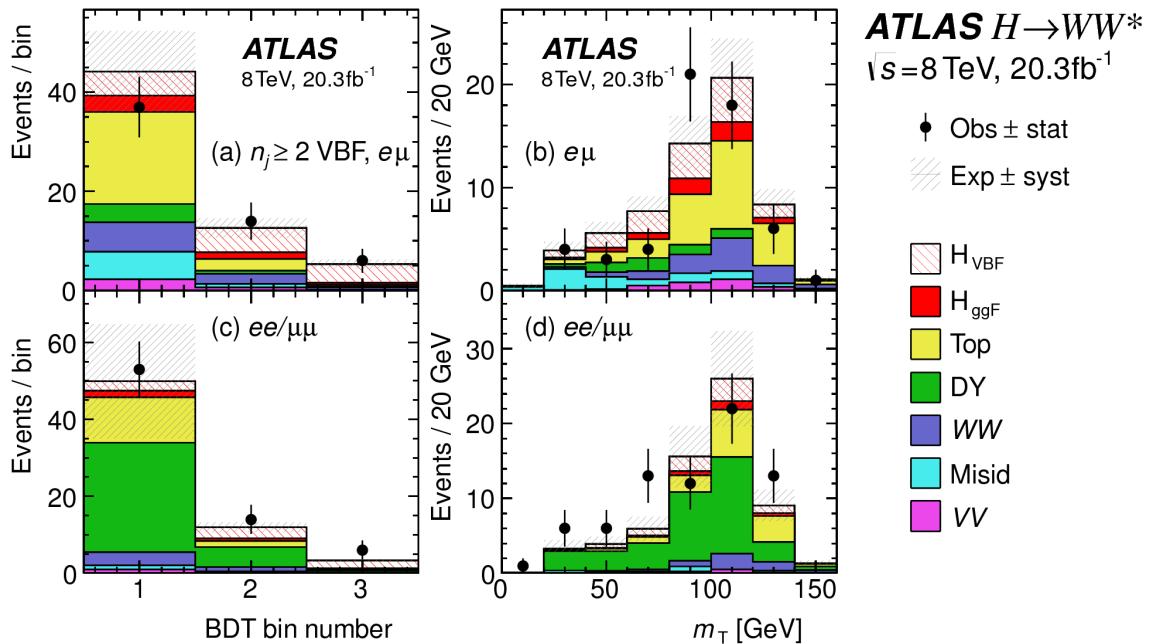


Figure 5.18: Postfit distributions in the BDT VBF analysis¹³.

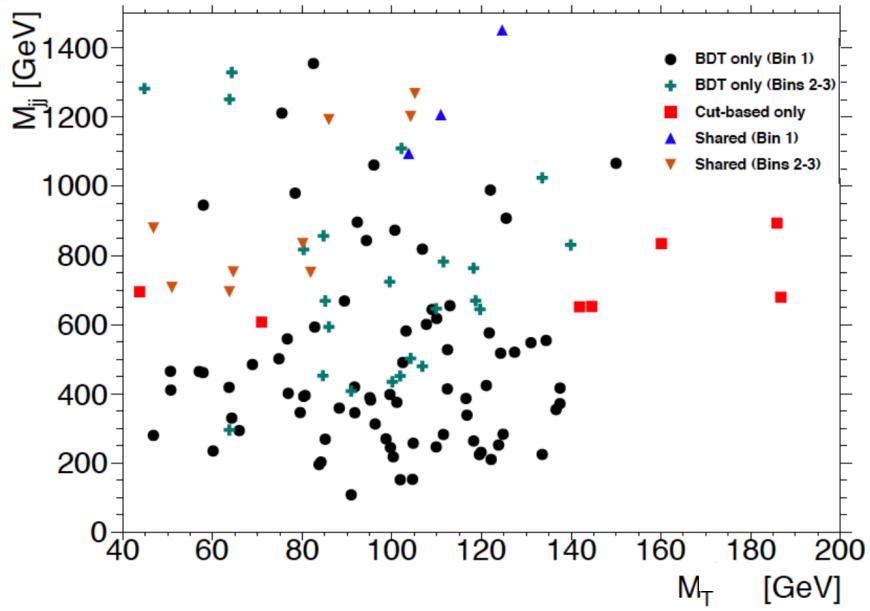


Figure 5.19: Overlap between cut-based and BDT VBF signal region candidates in the m_{jj} - m_T plane.

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

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

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

Chuck Palahniuk

6

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

6.1 INTRODUCTION

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

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

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

Table 6.1: All signal regions definitions input into final statistical fit ¹³.

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

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

The details of the dedicated gluon fusion $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ search are not discussed in this thesis and instead left to more comprehensive sources¹³. However, a brief summary of the results are essential for describing the results of the full analysis and interpreting the results of the dedicated VBF search in this broader context.

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

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

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

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

Table 6.2: Post-fit yields in the different ggF and VBF dedicated signal regions ¹³.

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

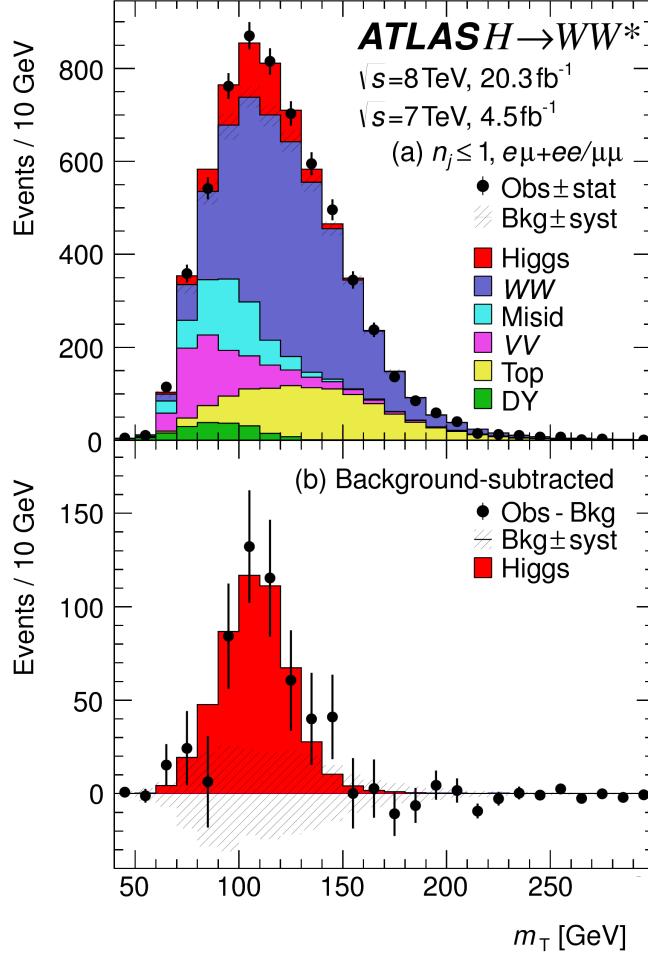


Figure 6.1: Post-fit m_T distribution in the $n_j \leq 1$ regions ¹³.

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

6.3 SIGNAL STRENGTH MEASUREMENTS IN ggF AND VBF PRODUCTION

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

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

Figure 6.2 gives the best fit signal strength $\hat{\mu}$ as a function of hte hypothesized Higgs mass. The value at 125.36 GeV corresponds to the μ quoted in equation 6.1. This value of the Higgs mass is used because it is the most precise mass measurement from ATLAS, a result of the combined $\gamma\gamma$ and ZZ mass measurements³.

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

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

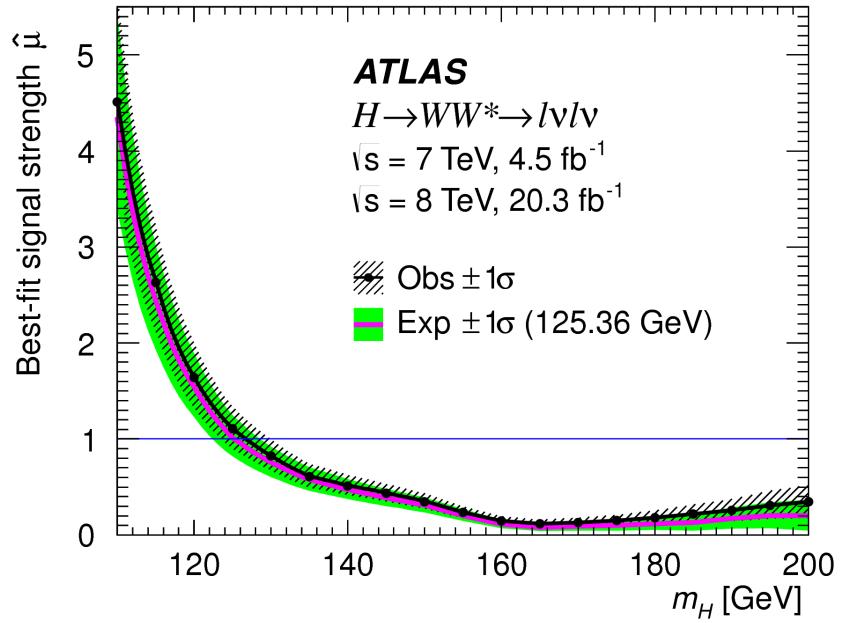


Figure 6.2: Best fit signal strength $\hat{\mu}$ as a function of hypothesized m_H ¹³.

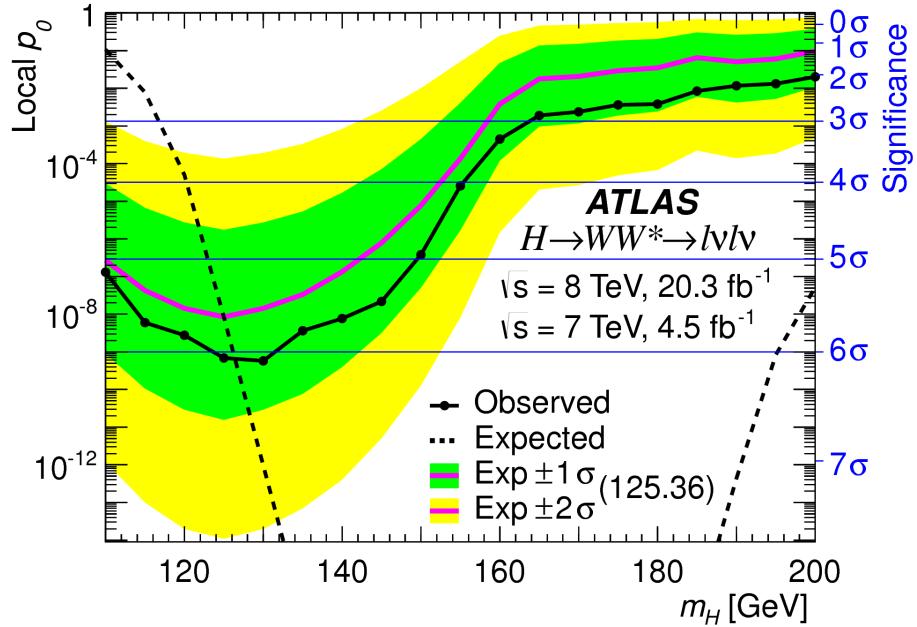


Figure 6.3: Local p_0 as a function of m_H ¹³.

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

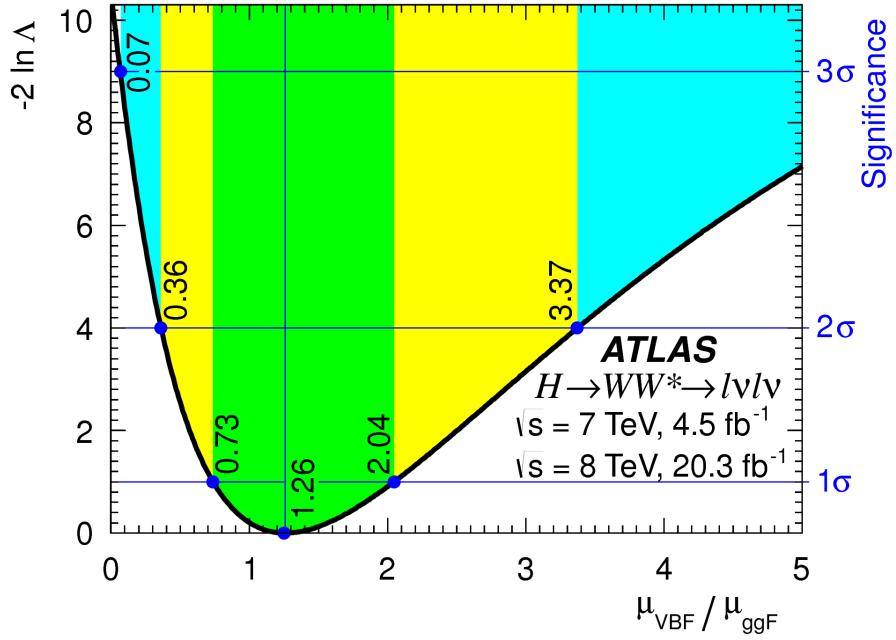


Figure 6.4: Likelihood as a function of $\mu_{\text{VBF}}/\mu_{\text{ggF}}$.¹³

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

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

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

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

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

(stat.) (syst.)

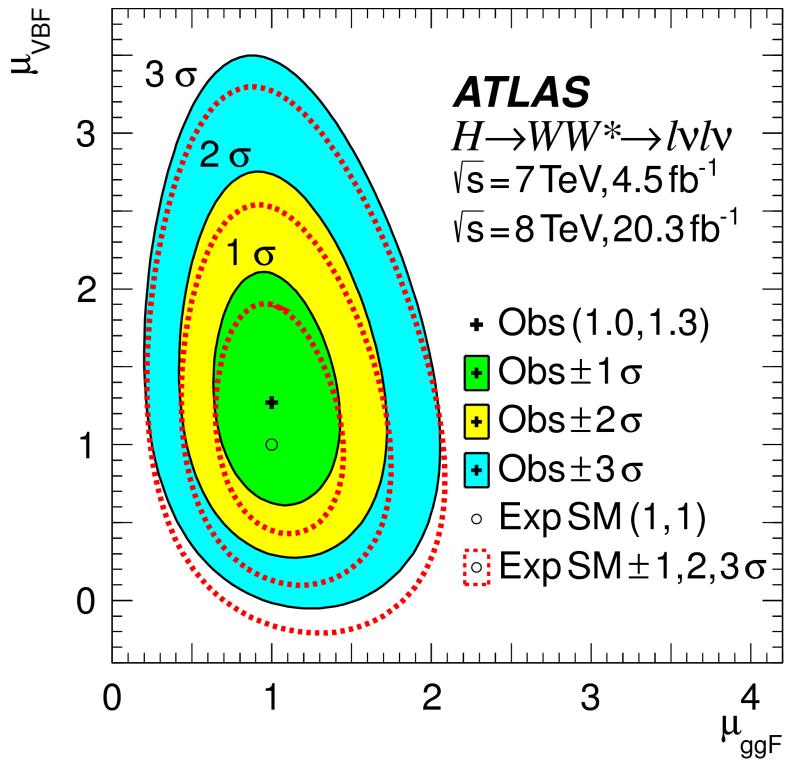


Figure 6.5: Likelihood scan as a function of μ_{VBF} and μ_{ggF} ¹³.

6.4 MEASUREMENT OF HIGGS COUPLINGS TO VECTOR BOSONS AND FERMIONS

Similar to the parameterization of signal strength, the couplings of the Higgs to fermions and bosons can also be parameterized. The parameter of interest in this case is κ , or the ratio of the measured coupling to the standard model expectation. Both the fermion and boson couplings have these so-called scale factors, κ_F for fermions and κ_V for bosons. Gluon fusion production is sensitive to the fermion couplings through the top quark loops in its production, while VBF production is sensitive to the vector boson couplings in its production. Both modes are sensitive to the vector boson couplings in their decays. The signal strengths will have dependence on the coupling scale factors as described in equation 6.4¹⁹.

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

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

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

(stat.) (syst.)

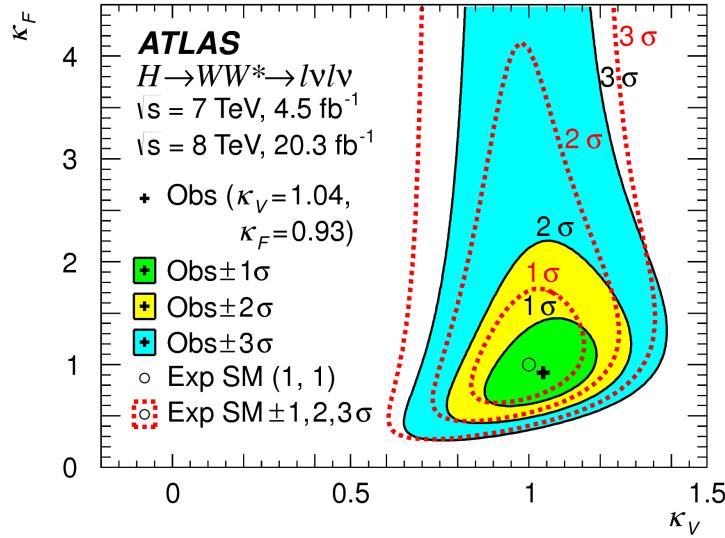


Figure 6.6: Likelihood scan as a function of κ_F and κ_V ¹³.

6.5 HIGGS PRODUCTION CROSS SECTION MEASUREMENT

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

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

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

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

(stat.) (syst.)

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

6.6 CONCLUSION

The combined analysis of the gluon fusion and vector boson fusion processes in $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ in the 7 and 8 TeV datasets has yielded the first discovery level significance for Higgs production in this decay channel. Additionally, precise measurements of the couplings to vector bosons and fermions are given. Finally, signal strengths and cross sections for each production mode are measured.

Part III

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

7

Search overview

8

Search for Higgs pair production in boosted
final states

9

Results with Run 2 2015 dataset

Part IV

Looking ahead

10

Conclusion

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

References

- [1] (2011). Calibrating the b -tag efficiency and mistag rate in 35 pb^{-1} of data with the atlas detector. *ATLAS-CONF-2011-089*.
- [2] (2012). *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*. Technical Report ATLAS-CONF-2012-047, CERN, Geneva.
- [3] (2013). Jet energy scale and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7 \text{ tev}$ with atlas 2011 data. *ATLAS-CONF-2013-004*.
- [4] (2014). *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data*. Technical Report ATLAS-CONF-2014-032, CERN, Geneva.
- [5] Aad, G. et al. (2014). Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton–proton collision data. *Eur. Phys. J.*, C74(11), 3130.
- [6] Armbruster, A. J. (2013). Discovery of a Higgs Boson with the ATLAS detector. CERN-THESIS-2013-047.
- [7] ATLAS Collaboration (2011). Luminosity Determination in pp Collisions at $\sqrt{s} = 7 \text{ TeV}$ Using the ATLAS Detector at the LHC. *Eur. Phys. J.*, C 71, 1630.
- [8] ATLAS Collaboration (2012). Measurement of the b -tag Efficiency in a Sample of Jets Containing Muons with 5 fb^{-1} of Data from the ATLAS Detector. *ATLAS-CONF-2012-043*.
- [9] B. P. Kersevan and E. Richter-Was (2004). The Monte Carlo event generator AcerMC version 2.0 with interfaces to PYTHIA 6.2 and HERWIG 6.5.
- [10] Butterworth, J. M., Forshaw, J. R., & Seymour, M. H. (1996). Multiparton interactions in photoproduction at HERA. *Z. Phys.*, C 72, 637.
- [11] Collaboration, A. (2014). Performance of the ATLAS muon trigger in pp collisions at $\sqrt{s} = 8 \text{ TeV}$. *Eur. Phys. J. C*, (arXiv:1408.3179. CERN-PH-EP-2014-154), 75, 19 p.
- [12] collaboration, A. (2015). Electron trigger performance in 2012 ATLAS data. ATLAS-COM-DAQ-2015-091.
- [13] Collaboration, A. (2015). Observation and measurement of Higgs boson decays to WW^* with the ATLAS detector. *Phys. Rev. D*, 92(012006).

- [14] Corcella, G. et al. (2001). HERWIG 6: An event generator for hadron emission reactions with interfering gluons (including super-symmetric processes) . *JHEP*, 01, 010.
- [15] G. Cowan, K. Cranmer, E. Gross, and O. Vitells (2011). Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J.*, C 71, 1554.
- [16] Gao, J., Guzzi, M., Huston, J., Lai, H.-L., Li, Z., et al. (2014). The CT10 NNLO Global Analysis of QCD. *Phys. Rev.*, D89, 033009.
- [17] Gleisberg, T., Hoeche, S., Krauss, F., Schonherr, M., Schumann, S., et al. (2009). Event generation with SHERPA 1.1. *JHEP*, 0902, 007.
- [18] Kauer, N. & Passarino, G. (2012). Inadequacy of zero-width approximation for a light Higgs boson signal.
- [19] LHC Higgs Cross Section Working Group, Heinemeyer, S., Mariotti, C., Passarino, G., & Tanaka (Eds.), R. (2013). Handbook of LHC Higgs Cross Sections: 3. Higgs Properties.
- [20] Mangano, M. L. et al. (2003). ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 0307, 001.
- [21] Nason, P. (2004). A new method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11, 040.
- [22] Olive, K. A. et al. (2014). Review of Particle Physics. *Chin. Phys.*, C38, 090001.
- [23] P. M. Nadolsky (2008). Implications of CTEQ global analysis for collider observables. *Phys. Rev.*, D 78, 013004.
- [24] S. Agostinelli et al. (2003). GEANT4, a simulation toolkit. *Nucl. Instrum. Meth.*, A 506, 250.
- [25] Sherstnev, A. & Thorne, R. S. (2009). Parton distributions for the LHC. *Eur. Phys. J.*, C 55, 553.
- [26] Sjostrand, T., Mrenna, S., & Skands, P. Z. (2006). PYTHIA 6.4 Physics and Manual. *JHEP*, 0605, 026.
- [27] Sjostrand, T., Mrenna, S., & Skands, P. Z. (2008). A Brief Introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178, 852–867.
- [28] Stewart, I. & Tackmann, F. (2012). Theory uncertainties for Higgs mass and other searches using jet bins. *Phys. Rev.*, D 85, 034011.



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

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