

# The Higgs boson with the ATLAS experiment at the LHC: Discovery, measurement, and searches for new physics

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.

# The Higgs boson with the ATLAS experiment at the LHC: Discovery, measurement, and searches for new physics

## ABSTRACT

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

# Contents

0	INTRODUCTION	I
I	Preliminaries	2
1	THE STANDARD MODEL AND BEYOND: A THEORETICAL OVERVIEW	3
1.1	The Standard Model of Particle Physics . . . . .	3
1.2	Electroweak Symmetry Breaking and the Higgs . . . . .	3
1.3	Higgs Boson Production and Decay . . . . .	3
1.4	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 I 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	Signal topology . . . . .	7
3.3	Background processes . . . . .	8
3.4	Isolating an $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ signal . . . . .	10
3.5	Background reduction in same-flavor final states . . . . .	14
3.6	Parameters of interest and statistical treatment . . . . .	19
4	THE DISCOVERY OF THE HIGGS BOSON AND THE ROLE OF THE $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ CHANNEL	23
5	OBSERVATION OF VECTOR BOSON FUSION PRODUCTION OF $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$	24
6	COMBINED RUN I $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ RESULTS	25

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	26
7	SEARCH OVERVIEW	27
8	SEARCH FOR HIGGS PAIR PRODUCTION IN BOOSTED FINAL STATES	28
9	RESULTS WITH RUN 2 2015 DATASET	29
IV	Looking ahead	30
10	CONCLUSION	31
	REFERENCES	32

# Listing of figures

3.1	A cartoon of the WW final state. Momenta are represented with thin arrows, spins with thick arrows. <sup>1</sup> . . . . .	7
3.2	Feynman diagram for Standard Model WW production . . . . .	8
3.3	Feynman diagrams for top pair production (left) and $Wt$ production (right) . . . . .	9
3.4	An example Feynman diagram of $W$ +jets production . . . . .	10
3.5	An example Feynman diagram of $Z$ +jets production . . . . .	10
3.6	An illustration of the unique analysis signal regions <sup>1</sup> . . . . .	12
3.7	A comparison of the subleading lepton $p_T$ spectrum between VBF $H \rightarrow WW^*$ production and $t\bar{t}$ background . . . . .	12
3.8	A graphical illustration of the $E_{T,rel}^{miss}$ calculation . . . . .	13
3.9	Predicted backgrounds (compared with data) as a function of $n_j$ (a and b) and $n_b$ (c) . . . .	14
3.10	An event display of a $Z/\gamma^* +$ jets event illustrating the effect of pileup interactions . . . .	16
3.11	The RMS of different missing transverse momentum definitions as a function of the average number of interactions per bunch crossing . . . . .	16
3.12	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
3.13	Comparison of $f_{recoil}$ distributions for $Z/\gamma^* +$ jets, $H \rightarrow WW^*$ , and other backgrounds with real neutrinos. . . . .	19
3.14	Signal significance as a function of cut value in the ggF $H \rightarrow WW^*$ with $n_j = 0$ . . . . .	20

# Listing of tables

3.1	A summary of backgrounds to the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ signal . . . . .	II
-----	--	----

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

- I.1 THE STANDARD MODEL OF PARTICLE PHYSICS
- I.2 ELECTROWEAK SYMMETRY BREAKING AND THE HIGGS
- I.3 HIGGS BOSON PRODUCTION AND DECAY
- I.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 will present 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 present in broad terms how the search and measurement are undertaken, before going into details on the specific sub-categories within the broader analysis.

First, the topology of the signal final state and corresponding backgrounds are presented. Next, an overview of the variables used to reduce the backgrounds and enhance the signal is given. These will be described in general, while specific values of selection cuts and background estimation will be provided in subsequent chapters. Finally, the parameters of interest in the search and measurement will be defined, and a brief overview of the statistical treatment of the final Higgs candidates is shown.



**Figure 3.1:** A cartoon of the  $WW$  final state. Momenta are represented with thin arrows, spins with thick arrows.<sup>1</sup>

### 3.2 SIGNAL TOPOLOGY

The analysis presented here and in subsequent chapters is the study of the Higgs boson in the  $WW$  final state, where each  $W$  boson subsequently decays into a charged lepton and a neutrino. In its simplest form, the final state will then consist 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  $\tau$  leptons, only leptonic decays of the  $\tau$  are considered, leading to additional neutrinos in the final state but still giving 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  $E_T^{\text{miss}}$ ). Final states where both of the charged leptons are electrons or muons are referred to as the “same flavor” final states, while those with one electron and one muon are referred to as “different flavor”.

The final state leptons will also 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 3.1. Because the neutrino has a left handed helicity and the anti-neutrino has a right handed helicity, 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 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.

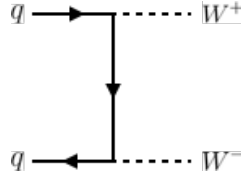


While the basic final state consists of two leptons and  $E_T^{\text{miss}}$ , there can be additional objects as well depending on the production mode of the Higgs. As described in detail in Chapter 1, if the Higgs is produced via vector boson fusion production, there will be two additional forward jets in the event. Even in gluon fusion, one or more jets can be produced through initial state radiation from the incoming gluons. The analysis is separated into different signal regions depending on the number of hard jets reconstructed in the final state as well.

### 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. The details of background estimation and specific cuts are left for later sections. Table 3.1 summarizes the different processes.

#### 3.3.1 STANDARD MODEL WW PRODUCTION



**Figure 3.2:** Feynman diagram for 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.

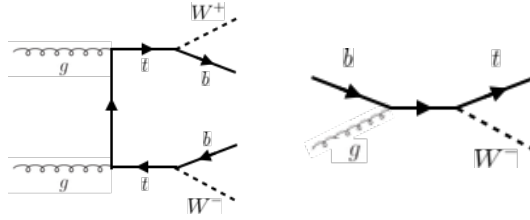


Figure 3.3: Feynman diagrams for top pair production (left) and  $Wt$  production (right)

### 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.

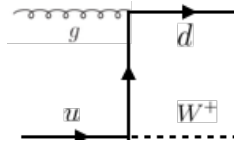


Figure 3.4: An example Feynman diagram of  $W^+$  jets production

### 3.3.4 $Z/\gamma^*$ +JETS BACKGROUND

Production of a  $Z/\gamma^*$  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  $E_T^{\text{miss}}$  can be used to reduce the background.

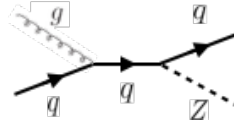


Figure 3.5: An example Feynman diagram of  $Z$ +jets production

### 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.

## 3.4 ISOLATING AN $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ SIGNAL

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 a combinatorically large number of potential signal regions. Additionally, signal regions can be optimized separately to be sensitive to the distinct production modes of the Higgs. Gluon fusion, vector

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

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.

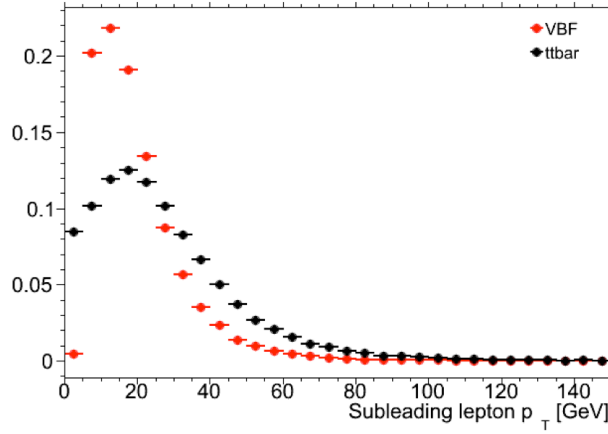
#### 3.4.1 EVENT PRE-SELECTION

Before being sorted into the distinct signal regions, basic cuts are applied on the reconstructed objects in the event to select Higgs-like event candidates. First, two oppositely charged leptons are required. 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 3.7 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.

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,  $E_T^{\text{miss}}$  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



**Figure 3.6:** An illustration of the unique analysis signal regions<sup>1</sup>



**Figure 3.7:** A comparison of the subleading lepton  $p_T$  spectrum between VBF  $H \rightarrow WW^*$  production and  $t\bar{t}$  background

$E_T^{\text{miss}}$  spectrum than backgrounds, especially if those backgrounds did not contain neutrinos. One additional consideration when using  $E_T^{\text{miss}}$  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  $E_T^{\text{miss}}$  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  $E_T^{\text{miss}}$  vector is within  $\pi/2$  radians in the transverse plane, the  $E_T^{\text{miss}}$  is projected away from this object. Otherwise, the normal  $E_T^{\text{miss}}$  vector is used. Figure 3.8 shows a graphical illustration of this concept.

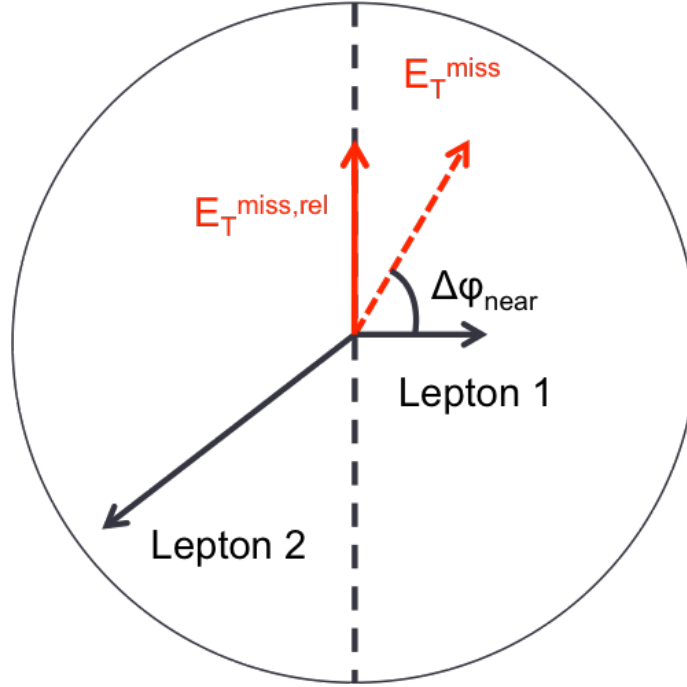


Figure 3.8: A graphical illustration of the  $E_{T,\text{rel}}^{\text{miss}}$  calculation

Once both the lepton and  $E_T^{\text{miss}}$  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 hard 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 3.9 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 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.

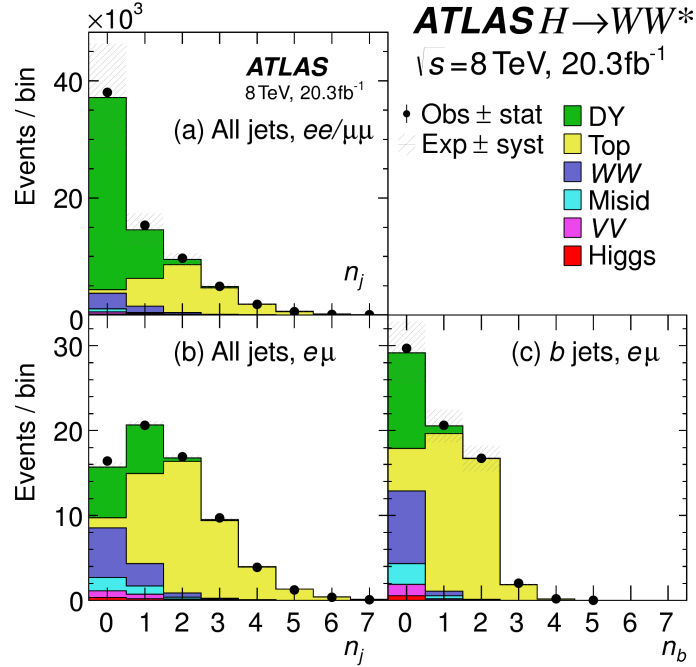


Figure 3.9: Predicted backgrounds (compared with data) as a function of  $n_j$  (a and b) and  $n_b$  (c)

### 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  $Z/\gamma^*$  decays to same flavor leptons. Because real neutrinos are absent in the  $Z/\gamma^*$  decays to  $ee$  and  $\mu\mu$ , a cut on

$E_T^{\text{miss}}$  should largely reduce the background. However, as this section will demonstrate, with increasing pileup conditions the resolution of the calorimeter-based  $E_T^{\text{miss}}$  degrades greatly. Therefore, two new variables for  $Z/\gamma^*$  background reduction are constructed and described in this section.

### 3.5.1 PILEUP AND $E_T^{\text{miss}}$ 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  $E_T^{\text{miss}}$  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  $E_T^{\text{miss}}$  resolution degrades, meaning that the  $E_T^{\text{miss}}$  resolution is particularly sensitive to LHC instantaneous luminosity conditions.

Figure 3.10 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 3.11 shows the RMS of the  $E_T^{\text{miss}}$  distribution in  $Z \rightarrow \mu\mu$  events (where there are no real neutrinos) 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 efficacy of a cut on  $E_T^{\text{miss}}$  to reduce the  $Z/\gamma^*$  background.

### 3.5.2 TRACK-BASED DEFINITIONS OF MISSING TRANSVERSE MOMENTUM

Because the increasing number of secondary proton-proton interactions degrades calorimeter-based  $E_T^{\text{miss}}$  resolution, a new variable using only contributions from the primary interaction vertex is necessary to further reduce the  $Z/\gamma^*$  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



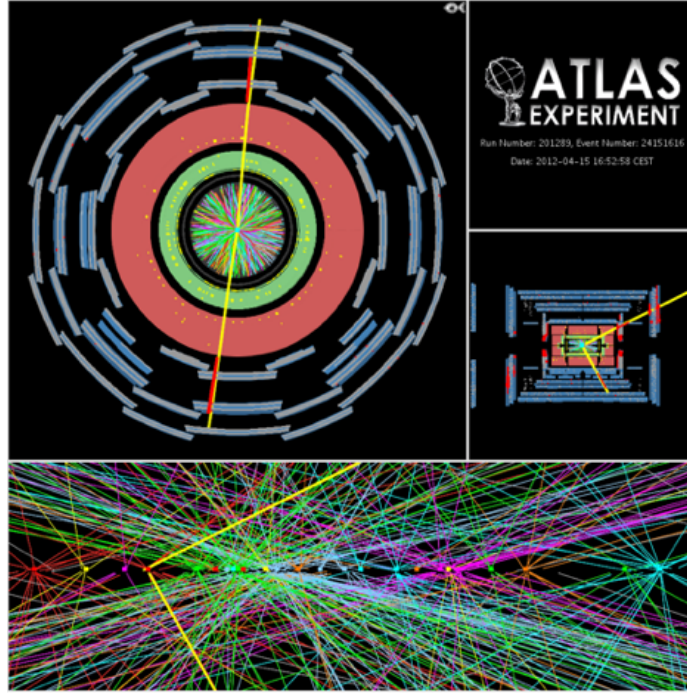


Figure 3.10: An event display of a  $Z/\gamma^* + \text{jets}$  event illustrating the effect of pileup interactions

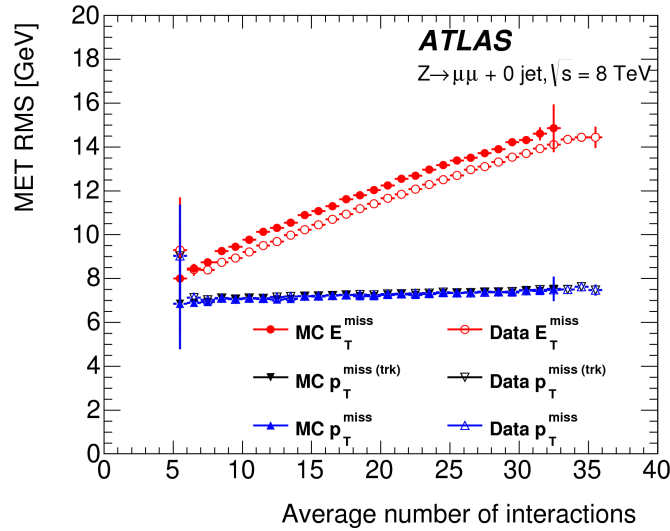


Figure 3.11: The RMS of different missing transverse momentum definitions as a function of the average number of interactions per bunch crossing

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,  $p_T^{\text{miss(trk)}}$ , is the vectorial sum of the  $p_T$  of all of the tracks in the event and the selected leptons (excluding the tracks asso-

ciated with the selected leptons to avoid double counting). This is defined in equation 3.2.

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

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,  $p_T^{\text{miss}}$ , 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.

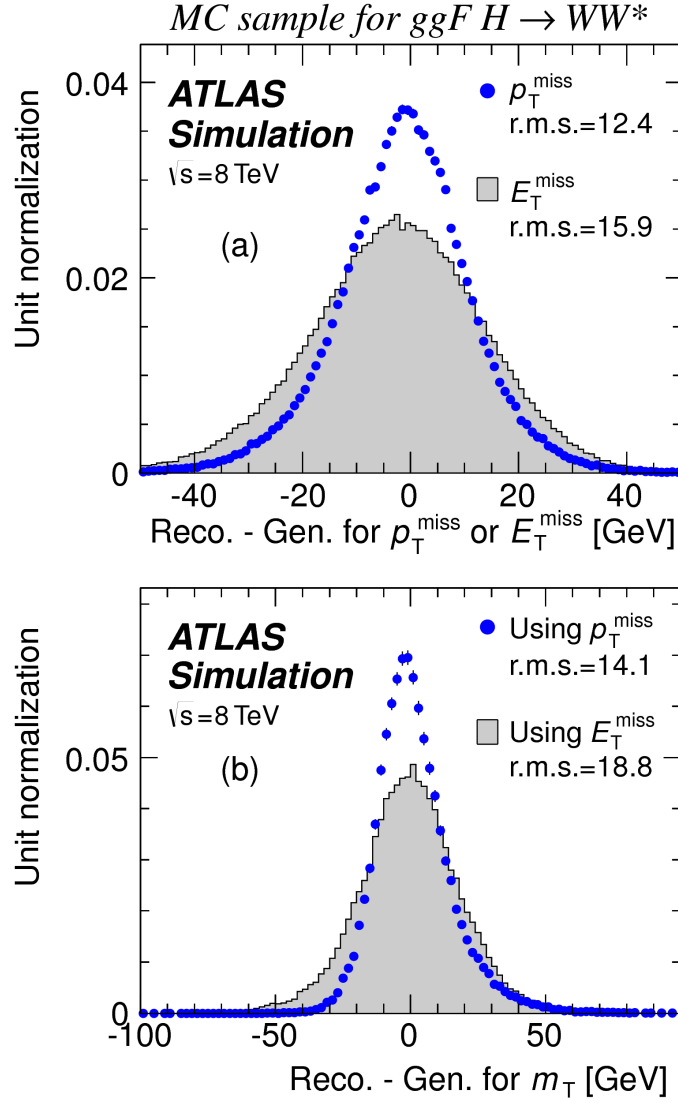
$$\mathbf{p}_T^{\text{miss}} = - \left( \sum_{\text{selected leptons}} \mathbf{p}_T + \sum_{\text{selected jets}} \mathbf{p}_T + \sum_{\text{other tracks}} \mathbf{p}_T \right), \quad (3.3)$$

Figure 3.11 illustrates that these two new variables accomplish their intended purpose. The resolution as a function of mean number of interactions for both  $p_T^{\text{miss(trk)}}$  and  $p_T^{\text{miss}}$  is much flatter compared to the dependence for  $E_T^{\text{miss}}$ .

Figure 3.12a shows the difference between the true and reconstructed values of missing transverse momentum using both the track-based  $p_T^{\text{miss}}$  and calorimeter based  $E_T^{\text{miss}}$ . The RMS of the distribution improves by 3.5 GeV when using  $p_T^{\text{miss}}$ .

### 3.5.3 DISTINGUISHING $Z/\gamma^* + \text{JETS}$ 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 kinematic and topological differences between the  $Z/\gamma^*$  background and  $H \rightarrow WW^*$  signal. Because there are no real neutrinos in the final state (in the case of  $Z/\gamma^* \rightarrow e\bar{e}, \mu\bar{\mu}$  decays), the dilepton system of a  $Z/\gamma^*$  will be balanced with the jets produced in the hard scatter. A new variable,  $f_{\text{recoil}}$ , 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  $f_{\text{recoil}}$  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



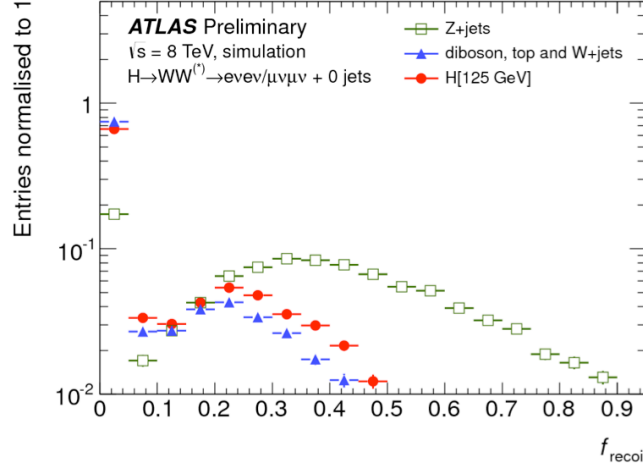
**Figure 3.12:** The difference between the true and reconstructed values of the missing transverse momentum (a) and  $m_T$  (b) in a gluon fusion signal sample

Vertex Fraction (JVF, described in chapter 2). The denominator is the magnitude of the dilepton  $p_T$ .

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

Figure 3.13 shows a shape comparison of the distribution of  $f_{\text{recoil}}$  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 cut on  $f_{\text{recoil}}$  will also reduce the  $Z/\gamma^* +$  jets background while maintaining a good signal efficiency.



**Figure 3.13:** Comparison of  $f_{\text{recoil}}$  distributions for  $Z/\gamma^* + \text{jets}$ ,  $H \rightarrow WW^*$ , and other backgrounds with real neutrinos.

### 3.5.4 OPTIMIZING BACKGROUND REDUCTION CUTS

The cuts on  $p_{\text{T}}^{\text{miss}(\text{trk})}$  and  $f_{\text{recoil}}$  used to reduce the  $Z + \text{jets}$  background must be optimized to maximize their efficacy. Figure 3.14 shows an early attempt to optimize the combination of the two cuts in the gluon fusion zero jet bin. Each bin shows the expected signal significance if the  $p_{\text{T,rel}}^{\text{miss}(\text{trk})}$  is required to be greater than the left edge of the bin and the  $f_{\text{recoil}}$  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  $p_{\text{T,rel}}^{\text{miss}(\text{trk})}$  values greater than 45 GeV.

### 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

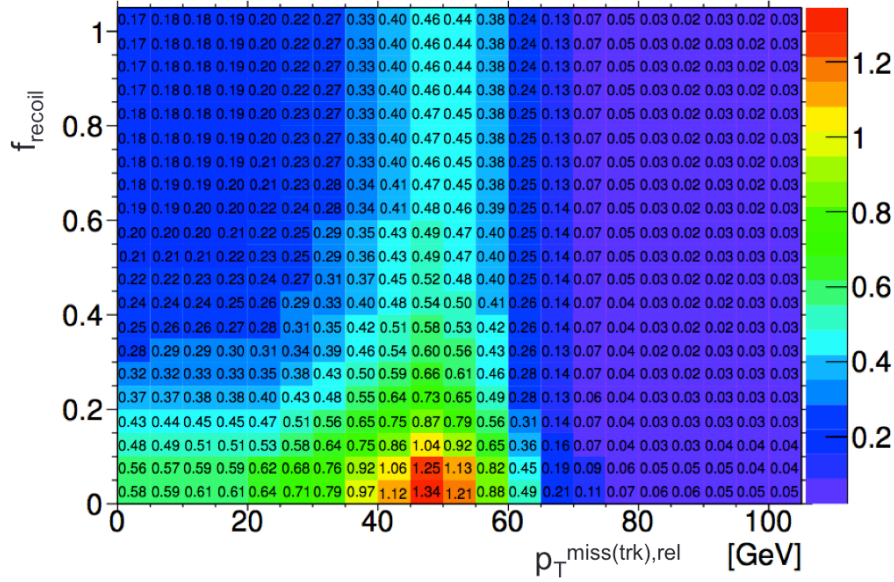


Figure 3.14: Signal significance as a function of cut value in the ggF  $H \rightarrow WW^*$  with  $n_j = 0$

the final state. However, a proxy for the invariant mass using transverse plane information can be defined. 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  $\mu$ . 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 parametrized as a function of  $\mu$ .  $\mu$  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.

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

Here the  $E_T^{\ell\ell}$  and  $p_T^{\ell\ell}$  are the transverse energy and momentum of the dilepton system, while  $p_T^{\text{miss}}$  is a proxy for the transverse momentum of the di-neutrino system. The track-based  $p_T^{\text{miss}}$  is used in the  $m_T$  rather than the calorimeter based  $E_T^{\text{miss}}$  because it has a better resolution on the true transverse mass. Figure 3.12b 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  $p_T^{\text{miss}}$  in the  $m_T$  calculation.

### 3.6.2 STATISTICAL TREATMENT\*

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,  $\mu$  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 back-

---

\*Many thanks to Aaron Armbruster, whose thesis<sup>3</sup> inspired parts of this section.

ground normalization. This is shown in equation 3.8.

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

Here,  $\vartheta$  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,  $N_{\text{CR}}$

# 4

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



# 5

Observation of Vector Boson Fusion

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

# 6

Combined Run 1  $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$  results

## 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] Collaboration, A. (2015). Observation and measurement of higgs boson decays to  $ww^*$  with the atlas detector. *Phys. Rev. D*, 92(012006).



**T**HIS THESIS WAS TYPESET using  $\text{\LaTeX}$ , originally developed by Leslie Lamport and based on Donald Knuth's  $\text{\TeX}$ .

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](https://github.com/asm-products/Dissertate) or from its lead author, Jordan Suchow, at [suchow@post.harvard.edu](mailto:suchow@post.harvard.edu).