Bringing the Higgs Boson to Rest

by

Koos van Nieuwkoop
B.Sc., University of the Fraser Valley, 2010

THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
Master of Science
IN THE
Department of Physics
Faculty of Science

© Koos van Nieuwkoop
SIMON FRASER UNIVERSITY
Spring 2013

All rights reserved. However, in accordance with the Copyright Act of Canada, this work may be reproduced, without authorization, under the conditions for “Fair Dealing.” Therefore, limited reproduction of this work for the purposes of private study, research, criticism, review, and news reporting is likely to be in accordance with the law, particularly if cited appropriately.
APPROVAL

Name: Koos van Nieuwkoop

Degree: Master of Science

Title of thesis: Bringing the Higgs Boson to Rest

Examining Committee: Dr. Eldon Emberly, Associate Professor (Chair)

Dr. Bernd Stelzer, Senior Supervisor
Assistant Professor

Dr. Michel Vetterli, Supervisor
Professor

Dr. Dugan O'Neil, Supervisor
Associate Professor

Dr. Sarah Johnson, Internal Examiner
Senior Lecturer

Date Approved: January 31, 2013
The author, whose copyright is declared on the title page of this work, has granted to Simon Fraser University the right to lend this thesis, project or extended essay to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users.

The author has further granted permission to Simon Fraser University to keep or make a digital copy for use in its circulating collection (currently available to the public at the “Institutional Repository” link of the SFU Library website (www.lib.sfu.ca) at http://summit/sfu.ca and, without changing the content, to translate the thesis/project or extended essays, if technically possible, to any medium or format for the purpose of preservation of the digital work.

The author has further agreed that permission for multiple copying of this work for scholarly purposes may be granted by either the author or the Dean of Graduate Studies.

It is understood that copying or publication of this work for financial gain shall not be allowed without the author’s written permission.

Permission for public performance, or limited permission for private scholarly use, of any multimedia materials forming part of this work, may have been granted by the author. This information may be found on the separately catalogued multimedia material and in the signed Partial Copyright Licence.

While licensing SFU to permit the above uses, the author retains copyright in the thesis, project or extended essays, including the right to change the work for subsequent purposes, including editing and publishing the work in whole or in part, and licensing other parties, as the author may desire.

The original Partial Copyright Licence attesting to these terms, and signed by this author, may be found in the original bound copy of this work, retained in the Simon Fraser University Archive.

Simon Fraser University Library
Burnaby, British Columbia, Canada

revised Fall 2011
Abstract

Within the Standard Model of particle physics, the Higgs boson can decay to a pair of $W$ bosons which decay leptonically. Despite its relatively large branching ratio, this is a challenging channel in which to search for the Higgs boson since we cannot detect neutrinos directly with the ATLAS detector. The matrix element method is a first principles approach that allows for better separation of signal and background by responding to subtle differences in the measured event kinematics. A straightforward implementation of the method is to assume that the Higgs boson is produced at rest in the transverse plane. However, this is often not the case due to next-to-leading order effects like initial state radiation. In order to improve the sensitivity of the matrix element analysis, we developed an estimator for the transverse momentum of the Higgs boson that allows us to boost it into its transverse rest frame. Using a regression tree algorithm to estimate the transverse kinematics of the Higgs boson on an event-by-event basis, we observe a 15% improvement in sensitivity. The application of this technique to Higgs boson property measurements, such as the determination of its spin, is also explored.
Opgedragen aan mijn familie:
Ma & Pa, Henk, Derek-John en William
Acknowledgements

Without the patience and guidance of my supervisor Bernd Stelzer, this thesis would not have materialised. Bernd, you gave me the opportunity to work in Data Quality and get involved in the mad rush to find the Higgs boson. Thank you for the many enlightening and clarifying discussions, but also for the encouragement along the way.

Doug Schouten, thanks for your many ideas and insights that were vital to completing this thesis. I appreciate the time you took to explain physics and computing concepts.

Thanks to Sarah Johnson, Dugan O’Neil and Mike Vetterli, faculty at SFU, for your input and feedback as examining committee. Also thanks to all SFU postdocs and students in high energy physics, Michelle Boudreau, Noel Dawe, Jennifer Godfrey, Jamie Horton, Jiri Kvita, Michele Petteni, David Shinkaruk, Andres Tanasijczuk, and Michel Trottier-McDonald for the enlightening discussions, physics help, and of course, sound advice on dating.

Finally, thanks to my parents, family, and friends for your support in reaching this milestone.
Contents

Approval ii
Abstract iii
Dedication iv
Acknowledgements v
Contents vi
List of Tables ix
List of Figures x

1 The Standard Model 1
  1.1 The Standard Model of Particle Physics ...................... 1
  1.2 The Higgs Boson ........................................... 5

2 ATLAS and the LHC 10
  2.1 The Large Hadron Collider .................................. 10
  2.2 The ATLAS Detector ........................................ 13
    2.2.1 The ATLAS Coordinate System ......................... 14
    2.2.2 The Inner Detector ................................... 15
    2.2.3 The Calorimeter ...................................... 17
    2.2.4 The Muon Spectrometer ................................. 19
    2.2.5 The ATLAS Trigger System ......................... 20
CONTENTS

2.2.6 Data Quality ............................................. 21
2.3 Simulated Data ............................................. 25

3 Higgs Boson Kinematics ..................................... 28
  3.1 Conservation of Angular Momentum .................. 29
  3.2 $H \rightarrow WW$ Production at the LHC .............. 31
  3.3 The Transverse Plane .................................. 34
  3.4 Kinematic Observables .................................. 36

4 $H \rightarrow WW$ Analysis ................................... 41
  4.1 Event Selection .......................................... 41
    4.1.1 Triggers .......................................... 41
    4.1.2 Event Selection .................................. 42
  4.2 Monte Carlo Modelling ................................ 47
  4.3 Background Estimation ................................ 48
  4.4 Results ................................................. 49

5 The Matrix Element Technique ............................. 52
  5.1 The Matrix Element Discriminant .................... 52
  5.2 Calculation of Event Probability Densities ........... 53
  5.3 Variable Transformation ................................ 55
  5.4 Transfer Functions .................................... 56
  5.5 EPD Control Regions .................................. 59
  5.6 Results ................................................. 61

6 Boosted Regression Trees .................................. 62
  6.1 Regression Trees ....................................... 62
  6.2 Training Regression Trees ............................. 63
  6.3 Boosting Regression Trees ............................. 65

7 Higgs Boson Transverse Rest Frame ....................... 67
  7.1 Implementation ........................................ 67
    7.1.1 Application to the ME Calculation ............... 69
  7.2 Validation ............................................. 69
### CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.2.1</td>
<td>Estimator Performance in $p_T^H$</td>
<td>69</td>
</tr>
<tr>
<td>7.2.2</td>
<td>Estimator Performance in $\phi^H$</td>
<td>72</td>
</tr>
<tr>
<td>7.2.3</td>
<td>Pileup Immunity</td>
<td>74</td>
</tr>
<tr>
<td>7.2.4</td>
<td>Regressor Overtraining</td>
<td>75</td>
</tr>
<tr>
<td>7.2.5</td>
<td>Impact of Boost Estimator on EPD Distributions</td>
<td>76</td>
</tr>
<tr>
<td>8</td>
<td>Measurement of the Higgs Boson Spin</td>
<td>79</td>
</tr>
<tr>
<td>8.1</td>
<td>Current Implementation</td>
<td>79</td>
</tr>
<tr>
<td>8.2</td>
<td>Performance Comparison</td>
<td>80</td>
</tr>
<tr>
<td>9</td>
<td>Conclusions</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>Bibliography</td>
<td>91</td>
</tr>
</tbody>
</table>
List of Tables

1.1 Spin-1 force carriers ................................................................. 2
1.2 Spin-\(\frac{1}{2}\) fermions (quarks) ................................................... 3
1.3 Spin-\(\frac{1}{2}\) fermions (leptons) .................................................... 3

4.1 Monte Carlo samples used in \(H \rightarrow WW\) analysis ................. 48

7.1 Monte Carlo samples used to generate regression estimator .......... 68
7.2 Performance comparison of boost strategies ............................... 78
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Coupling constants</td>
<td>5</td>
</tr>
<tr>
<td>1.2</td>
<td>Feynman diagrams for the dominant Higgs boson production processes</td>
<td>6</td>
</tr>
<tr>
<td>1.3</td>
<td>Branching ratios for the Standard Model Higgs boson</td>
<td>7</td>
</tr>
<tr>
<td>1.4</td>
<td>95% confidence level upper limits on the signal strength</td>
<td>8</td>
</tr>
<tr>
<td>1.5</td>
<td>Significance of observed excess as a function of the Higgs boson mass</td>
<td>9</td>
</tr>
<tr>
<td>2.1</td>
<td>Location of the four LHC experiments</td>
<td>11</td>
</tr>
<tr>
<td>2.2</td>
<td>Schematic view of the LHC accelerator</td>
<td>12</td>
</tr>
<tr>
<td>2.3</td>
<td>Total integrated luminosity for 2012</td>
<td>13</td>
</tr>
<tr>
<td>2.4</td>
<td>Schematic view of the ATLAS detector</td>
<td>14</td>
</tr>
<tr>
<td>2.5</td>
<td>Schematic view of the inner detector</td>
<td>16</td>
</tr>
<tr>
<td>2.6</td>
<td>Schematic view of the calorimeter</td>
<td>18</td>
</tr>
<tr>
<td>2.7</td>
<td>Schematic view of the muon spectrometer</td>
<td>20</td>
</tr>
<tr>
<td>2.8</td>
<td>Online Histogram Presenter</td>
<td>22</td>
</tr>
<tr>
<td>2.9</td>
<td>Data Quality Monitoring Display</td>
<td>22</td>
</tr>
<tr>
<td>2.10</td>
<td>$K_S^0 \rightarrow \pi^+\pi^-$ mass distribution</td>
<td>24</td>
</tr>
<tr>
<td>2.11</td>
<td>$K_S^0 \rightarrow \pi^+\pi^-$ rate monitor</td>
<td>24</td>
</tr>
<tr>
<td>2.12</td>
<td>Express stream trigger monitor</td>
<td>25</td>
</tr>
<tr>
<td>3.1</td>
<td>Feynman diagrams for Higgs boson production and SM $WW$ background</td>
<td>28</td>
</tr>
<tr>
<td>3.2</td>
<td>Possible spin configurations for $H \rightarrow WW$</td>
<td>29</td>
</tr>
<tr>
<td>3.3</td>
<td>Possible spin configurations for $H \rightarrow WW \rightarrow \ell\nu\ell\nu$</td>
<td>30</td>
</tr>
<tr>
<td>3.4</td>
<td>Opening angle between charged leptons for signal and background</td>
<td>32</td>
</tr>
<tr>
<td>3.5</td>
<td>Mass of dilepton system for signal and background</td>
<td>32</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.6</td>
<td>CT10 parton distribution function</td>
<td>33</td>
</tr>
<tr>
<td>3.7</td>
<td>Boost along z-axis for WW and Higgs boson events</td>
<td>34</td>
</tr>
<tr>
<td>3.8</td>
<td>Higgs boson $p_T$ distribution</td>
<td>35</td>
</tr>
<tr>
<td>3.9</td>
<td>Initial state radiation</td>
<td>35</td>
</tr>
<tr>
<td>3.10</td>
<td>Leading lepton $p_T$ distribution</td>
<td>36</td>
</tr>
<tr>
<td>3.11</td>
<td>Subleading lepton $p_T$ distribution</td>
<td>37</td>
</tr>
<tr>
<td>3.12</td>
<td>Calorimeter $E_T^{\text{miss}}$ distribution</td>
<td>38</td>
</tr>
<tr>
<td>3.13</td>
<td>Tracker $E_T^{\text{miss}}$ distribution</td>
<td>39</td>
</tr>
<tr>
<td>3.14</td>
<td>$m_T$ distribution</td>
<td>40</td>
</tr>
<tr>
<td>4.1</td>
<td>$E_{T,\text{rel}}^{\text{miss}}$ distribution</td>
<td>44</td>
</tr>
<tr>
<td>4.2</td>
<td>Number of jets distribution</td>
<td>45</td>
</tr>
<tr>
<td>4.3</td>
<td>$p_T^{\ell\ell}$ distribution</td>
<td>45</td>
</tr>
<tr>
<td>4.4</td>
<td>$\Delta\phi$ distribution</td>
<td>46</td>
</tr>
<tr>
<td>4.5</td>
<td>$m_{\ell\ell}$ distribution</td>
<td>46</td>
</tr>
<tr>
<td>4.6</td>
<td>Candidate $H \rightarrow WW \rightarrow e\nu\mu\nu$ event</td>
<td>47</td>
</tr>
<tr>
<td>4.7</td>
<td>$m_T$ distribution</td>
<td>50</td>
</tr>
<tr>
<td>4.8</td>
<td>$m_T$ distribution for data with predicted background subtracted</td>
<td>51</td>
</tr>
<tr>
<td>4.9</td>
<td>Significance of observed excess in $H \rightarrow WW \rightarrow \ell\nu\ell\nu$</td>
<td>51</td>
</tr>
<tr>
<td>5.1</td>
<td>EPD distributions normalised to unit area</td>
<td>56</td>
</tr>
<tr>
<td>5.2</td>
<td>$E_{\text{parton}}$ vs $E_{\text{jet}}$ distribution for $b$-jet events</td>
<td>57</td>
</tr>
<tr>
<td>5.3</td>
<td>Transfer function validation</td>
<td>58</td>
</tr>
<tr>
<td>5.4</td>
<td>EPD distribution in WW control region</td>
<td>59</td>
</tr>
<tr>
<td>5.5</td>
<td>EPD distribution in $tt\bar{t}$ control region</td>
<td>60</td>
</tr>
<tr>
<td>5.6</td>
<td>EPD distribution in $Z/\gamma^*$ control region</td>
<td>60</td>
</tr>
<tr>
<td>5.7</td>
<td>EPD distribution in signal region</td>
<td>61</td>
</tr>
<tr>
<td>6.1</td>
<td>Sample regression tree</td>
<td>63</td>
</tr>
<tr>
<td>6.2</td>
<td>Sample regression estimator</td>
<td>64</td>
</tr>
<tr>
<td>7.1</td>
<td>2D plot of truth and estimated $p_T$</td>
<td>70</td>
</tr>
<tr>
<td>7.2</td>
<td>Truth and estimated $p_T$ distribution</td>
<td>70</td>
</tr>
<tr>
<td>7.3</td>
<td>Non-uniformity of training sample</td>
<td>71</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

7.4 Resolution of $p_T$ estimate ........................................ 72
7.5 Resolution of $p_T$ estimate binned in truth $p_T$ ............... 73
7.6 Resolution of $\phi$ estimate ........................................ 73
7.7 Resolution of $\phi$ estimate binned in truth $p_T$ ............... 74
7.8 Resolution of $p_T$ estimate in different pileup conditions ....... 74
7.9 Overtraining check ..................................................... 75
7.10 EPD distribution before and after Lorentz boost ............... 76
8.1 Effect of flattening $p_x$ distribution ......................... 81
8.2 $p_T$ distribution for spin-0 and spin-2 Higgs boson .......... 81
8.3 Resolution of $p_T$ estimate for spin-0 and spin-2 Higgs boson ... 82
8.4 Dilepton $p_T$ distribution for spin-0 and spin-2 Higgs boson ...... 82
8.5 Resolution of dilepton $p_T$ estimate for spin-0 and spin-2 Higgs boson 83
8.6 Dilepton $\Delta\phi$ distribution for spin-0 and spin-2 Higgs boson .... 83
8.7 Resolution of dilepton $\Delta\phi$ estimate for spin-0 and spin-2 Higgs boson 84
8.8 Dilepton $\Delta\psi$ distribution for spin-0 and spin-2 Higgs boson .... 84
8.9 Resolution of dilepton $\Delta\psi$ estimate for spin-0 and spin-2 Higgs boson 85
8.10 Dilepton $p_T$ distribution for spin-0 and spin-2 Higgs boson ...... 86
8.11 Resolution of dilepton $p_T$ estimate for spin-0 and spin-2 Higgs boson . 86
8.12 $\Delta\phi$ distribution for estimated Higgs boson $p_T < 10$ GeV ....... 87
8.13 $\Delta\phi$ distribution for estimated Higgs boson $p_T$ between 20 and 30 GeV 87
8.14 $\Delta\phi$ distribution for estimated Higgs boson $p_T > 50$ GeV .......... 88
Chapter 1

The Standard Model

A myriad of particles were discovered during the last bygone century. Nobel Prize winner Willis Lamb wittily remarked, "the finder of a new elementary particle used to be rewarded by a Nobel Prize, but such a discovery now ought to be punished by a 10,000 dollar fine." Despite the initial confusion around the "particle zoo", physicists developed a coherent theory using a handful of fundamental particles to explain their experimental discoveries. To date, all known particles are described remarkably well by the Standard Model (SM) of particle physics. Ironically, the developments in the SM led to numerous Nobel Prizes after Lamb. In order to explain the origin of mass, it also predicts the presence of the legendary Higgs boson. ATLAS’s recent announcement of a new particle with a mass of approximately 126 GeV [1] is consistent with the SM Higgs boson and further augments the credibility of the theory.

1.1 The Standard Model of Particle Physics

The SM describes the fundamental forces of nature along with the elementary particles on which they act. It is derived from quantum field theory, an amalgamation of quantum mechanics and special relativity. While gravity is not yet included in the SM framework, electromagnetic, weak, and strong interactions are described. Electromagnetic forces are responsible for binding electrons to nuclei inside atoms. It also plays a role in holding atoms together to form molecules. Properties of solids, liquids and gases can be attributed to this force. Nuclear reactions, like the conversion of
CHAPTER 1. THE STANDARD MODEL

hydrogen to helium in the sun, are possible because of the weak interaction. As the name suggests, it is the weakest of the fundamental forces. The strong force is powerful enough to overcome the repulsive electric force between protons inside an atomic nucleus. Without it, atomic nuclei would not be stable. At a more fundamental level, it also holds quarks together inside hadrons such as protons and neutrons.

From a theoretical point of view, fundamental forces act by the exchange of bosons. To date, all exchange particles are measured to be spin-1 vector bosons which are summarised in table 1.1 along with their corresponding interactions. These forces act on two types of spin-$\frac{1}{2}$ fermions known as quarks (table 1.2) and leptons (table 1.3). It is postulated that gravity is mediated by a spin-2 boson called the graviton. Since gravity is extremely weak compared to the other forces, it is expected to have no measurable consequences in particle physics experiments [2].

<table>
<thead>
<tr>
<th>Boson</th>
<th>Force</th>
<th>Electric charge (e)</th>
<th>Mass (GeV)</th>
<th>Charge</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma$ (photon)</td>
<td>Electromagnetic</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>$g$ (gluon)</td>
<td>Strong</td>
<td>0</td>
<td>0</td>
<td>C</td>
</tr>
<tr>
<td>$W^\pm$</td>
<td>Weak</td>
<td>±1</td>
<td>80.4</td>
<td>E, W</td>
</tr>
<tr>
<td>$Z$</td>
<td>Weak</td>
<td>0</td>
<td>91.2</td>
<td>W</td>
</tr>
</tbody>
</table>

Table 1.1: Spin-1 force carriers in the SM [3].

The fundamental forces originate from various types of charge which determine via which forces a particle can interact. They are the electric (E), weak (W), and color (C) charges for the electromagnetic, weak, and strong interactions respectively. For example, as indicated in table 1.3, electrons carry electric and weak charge which implies they can interact by means of the electric and weak interactions but not by the strong interaction. In addition, note that gluons themselves carry color charge (table 1.1) which implies gluons can interact with other gluons, an effect commonly termed as "self-coupling".

Each particle listed in tables 1.1 through 1.3 has a corresponding antiparticle which has the same mass but opposite charge and inverted internal quantum numbers. Some particles, such as the $Z$ and $\gamma$, are their own antiparticle.

Fermions appear in three generations. The second and third generation fermions are copies of their first generation counterparts but with larger masses. Since these


<table>
<thead>
<tr>
<th>Generation</th>
<th>Quark</th>
<th>Electric charge (e)</th>
<th>Mass (GeV)</th>
<th>Charge</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>$u$</td>
<td>$+\frac{2}{3}$</td>
<td>$(2.3^{+0.7}_{-0.5}) \times 10^{-3}$</td>
<td>E, W, C</td>
</tr>
<tr>
<td></td>
<td>$d$</td>
<td>$-\frac{1}{3}$</td>
<td>$(4.8^{+0.7}_{-0.3}) \times 10^{-3}$</td>
<td>E, W, C</td>
</tr>
<tr>
<td>2nd</td>
<td>$c$</td>
<td>$+\frac{2}{3}$</td>
<td>$1.28 \pm 0.03$</td>
<td>E, W, C</td>
</tr>
<tr>
<td></td>
<td>$s$</td>
<td>$-\frac{1}{3}$</td>
<td>$(95 \pm 5) \times 10^{-3}$</td>
<td>E, W, C</td>
</tr>
<tr>
<td>3rd</td>
<td>$t$</td>
<td>$+\frac{2}{3}$</td>
<td>$174 \pm 1$</td>
<td>E, W, C</td>
</tr>
<tr>
<td></td>
<td>$b$</td>
<td>$-\frac{1}{3}$</td>
<td>$4.18 \pm 0.03$</td>
<td>E, W, C</td>
</tr>
</tbody>
</table>

Table 1.2: Spin-$\frac{1}{2}$ fermions (quarks) in the SM [3].

<table>
<thead>
<tr>
<th>Generation</th>
<th>Lepton</th>
<th>Electric charge (e)</th>
<th>Mass (GeV)</th>
<th>Charge</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>$e$</td>
<td>-1</td>
<td>$0.511 \times 10^{-3}$</td>
<td>E, W</td>
</tr>
<tr>
<td></td>
<td>$\nu_e$</td>
<td>0</td>
<td>$&lt; 2 \times 10^{-9}$</td>
<td>W</td>
</tr>
<tr>
<td>2nd</td>
<td>$\mu$</td>
<td>-1</td>
<td>$0.106$</td>
<td>E, W</td>
</tr>
<tr>
<td></td>
<td>$\nu_\mu$</td>
<td>0</td>
<td>$&lt; 0.19 \times 10^{-3}$</td>
<td>W</td>
</tr>
<tr>
<td>3rd</td>
<td>$\tau$</td>
<td>-1</td>
<td>$1.78$</td>
<td>E, W</td>
</tr>
<tr>
<td></td>
<td>$\nu_\tau$</td>
<td>0</td>
<td>$&lt; 18.2 \times 10^{-3}$</td>
<td>W</td>
</tr>
</tbody>
</table>

Table 1.3: Spin-$\frac{1}{2}$ fermions (leptons) in the SM [3].
more massive particles quickly decay to lighter particles, they are typically only observed in high-energy experiments or cosmic rays.

Even though table 1.2 lists the masses for single quarks, these are not observed in nature. This is due to the property of confinement in quantum chromodynamics (QCD), the theory of the strong force. Confinement demands that all observed states have a net color charge of zero. Since quarks carry color charges, they form bound states. The existence of quarks can be deduced from deep inelastic scattering of leptons with protons. Due to its short lifetime, the top quark can decay before forming a bound state via the strong force. Consequently, the quoted mass for the top quark in table 1.2 is determined directly from experiment. The other quark masses are derived from cross section measurements using the MS renormalisation scheme [3].

Neutrinos are leptons without electric charge that can only interact through the weak interaction. As a result, they are extremely hard to detect, one of the reasons why the subject of this thesis requires study as will later become evident. Each second, billions of neutrinos pass through our bodies without interacting. Even though upper limits on neutrino masses are small, experiments confirm that they do have mass [3].

Even though the electromagnetic and weak interactions are listed as separate entities, they were unified by Glashow, Salam, and Weinberg into the electroweak theory. This achievement hints that all fundamental forces can be unified at some energy scale known as the grand unified theory (GUT) scale, an idea still being pursued today. Some theories beyond the SM (like the Minimal Supersymmetric Extension of the SM [4]) predict that the coupling strengths of the electromagnetic, weak, and strong interactions (represented by coupling constants\(^1\) \(\alpha_e\), \(\alpha_w\), and \(\alpha_s\) respectively) converge to a common value at extremely high energies that are currently far from accessible to any experiment. By extrapolating the functional form of these coupling constants, unification is expected to occur on the order of \(10^{15}\) GeV (figure 1.1) [5].

\(^1\)The term "coupling constant" is misleading since \(\alpha\) depends on energy. Hence, it is sometimes called a "running coupling constant".
Figure 1.1: The plot on the left depicts the three coupling constants of the SM as a function of energy. Some theories beyond the SM (like the Minimal Supersymmetric Extension of the SM) predict that the coupling constants merge to a common value at extremely high energies as shown in the plot on the right [6].

1.2 The Higgs Boson

Despite the elegance of the SM, the origin of mass has remained ambiguous for many years. From a group theory point of view, the electromagnetic and weak interactions are unified using a non-Abelian (or non-commutative) gauge theory derived from the $SU(2) \times U(1)$ symmetry group. $U(1)$ refers to a set of transformations represented by complex scalars and $SU(2)$ refer to a set of transformations represented by $2 \times 2$ unitary matrices. However, the resulting theory has four massless gauge bosons which is in disagreement with experimental observations. By introducing a new scalar field, the Higgs field, three of these gauge bosons gain mass to become the $W^+, W^−$, and $Z$, while the last one remains massless to take the place of the $γ$.

The symmetry of the $SU(2) \times U(1)$ symmetry group is broken by means of the Higgs Mechanism. It suggests that the Higgs field permeates all of space. Particles like photons and gluons do not interact with this field, and move through space at the speed of light. In contrast, particles such as the $W$ and $Z$ interact with the Higgs field and acquire mass. The only way we can probe the existence of the Higgs field is through its excitation, the Higgs boson. Unfortunately, the SM does not predict the
mass of the Higgs boson and as a result, it is necessary to scan a large mass range.

The Higgs boson can be produced in several ways but the dominant production mode at high-energy proton-proton colliders is gluon fusion. This is because the gluon parton distribution function is large for momentum fractions sufficient to make a Higgs boson of mass 125 GeV (refer to figure 3.6). Since the Higgs boson is expected to couple to massive particles, it is produced through a top quark loop. Vector boson fusion contributes with a lower cross section\(^2\) but is also interesting because it leads to a distinct signature with two forward jets\(^3\) and probes the exclusive coupling of the Higgs boson to the weak vector bosons. Both production mechanisms are represented by Feynman diagrams in figure 1.2.

The Higgs boson is extremely unstable and after production quickly decays into less massive particles. For a Higgs boson mass of 126 GeV, the SM predicts a mean lifetime on the order of \(10^{-22}\) seconds. Figure 1.3 summarises the branching ratios of the various decay modes. Note that the WW channel has a relatively large branching ratio over a large mass range, also at \(\sim 126\) GeV where a new particle was observed in 2012. This makes WW a viable channel for discovering the Higgs boson and measuring its properties.

\(^2\)Cross section quantifies the likelihood of an interaction taking place.
\(^3\)See section 2.2.3 for description of jets.
Figure 1.3: Branching ratios for SM Higgs boson decay. The observed new particle with a mass at $\sim 126$ GeV is indicated by the orange dotted vertical line [7].

ATLAS’s observation paper on the Higgs-like boson [1] is based on the analyses of $ZZ$, $\gamma\gamma$, and $WW$ channels using 2011 and 2012 data and the $b\bar{b}$ and $\tau^+\tau^-$ channels using only 2011 data. The results are based on 4.8 fb$^{-1}$ of 7 TeV data ATLAS collected in 2011, along with 5.8 fb$^{-1}$ of 8 TeV data collected in 2012. The observation has a significance of 5.9 sigma which corresponds to a local $p_0$ value of $1.7 \times 10^{-9}$. The local $p_0$ value is defined as the probability that a background fluctuation equals or exceeds the excess observed in data. Using all of the above mentioned decay channels on 2011 and 2012 data, CMS has observed a similar excess with a significance of 5.0 sigma [8].

Figure 1.4 shows the upper limits on the Higgs boson production cross section as published in the ATLAS paper. Even though the SM does not predict the Higgs boson mass, it does predict a cross section once a mass is assumed. The solid line in figure 1.4 represents the observed 95% confidence level upper limit on the signal strength $\mu$, where $\mu$ is the Higgs boson production cross section normalised to the SM prediction. All SM Higgs boson mass hypotheses with $\mu$ below 1 are excluded. The dotted line represents the expected limit on $\mu$ based on the background only.
Figure 1.4: Observed and expected 95% confidence level upper limits on the signal strength $\mu$. The SM Higgs boson is excluded for masses where $\mu$ is below 1. The departure of the observed limit above the expected limit at around 126 GeV corresponds to the newly discovered particle [1].

hypothesis. Note the distinct departure of the observed limit above the expected limit around 126 GeV. The green and yellow bands represent the $1\sigma$ and $2\sigma$ uncertainties above and below the expected limit.

The significance of the excess is shown in figure 1.5. The solid line represents the observed local $p_0$, and the dashed represents the expected local $p_0$ for a SM Higgs boson as a function of mass.
Figure 1.5: Observed and expected local $p_0$ value (left $y$-axis) and significance (right $y$-axis) for the SM model Higgs boson as a function of its mass [1].
Chapter 2

ATLAS and the LHC

The Large Hadron Collider (LHC) is a proton-proton collider located near Geneva, Switzerland. It is the home to ATLAS, one of the two general purpose detectors that records LHC collisions.

2.1 The Large Hadron Collider

The LHC is operated in a 26.7 km circular underground tunnel located at CERN. It accelerates protons to nearly the speed of light for the four main experiments along the LHC ring. ATLAS and CMS are general purpose detectors designed for a wide range of physics. ALICE is designed to study quark-gluon plasmas, while LHCb is designed to study charge-parity (CP) violating $B$ meson decays. Figure 2.1 shows the physical location of the four experiments.

Constructing the LHC was challenging from a technical point of view. A high quality vacuum must be maintained throughout the entire circumference of the beam pipe. In addition, a complex system of dipole magnets keeps the beam in its orbit, while quadrupole magnets focus the beam. Since many of the magnets are super-conducting, the LHC requires a cryogenic system which maintains temperatures of only a few degrees above absolute zero.

Protons pass through several stages before colliding inside one of the detectors at the LHC. Hydrogen atoms are ionised into protons which are accelerated to 50 MeV using a linear accelerator (LINAC2). Radio frequency (RF) cavities are used to
increase the energy of collections of protons referred to as bunches. These pass into the proton synchrotron booster (PSB) ring where they are further accelerated to 1.4 GeV. A circular accelerator has the advantage that the same RF cavities can be used many times to accelerate the same particle, a feature not available in linear accelerators. Next, the protons are further accelerated to 25 GeV in the proton synchrotron (PS) and then to 450 GeV in the super proton synchrotron (SPS). Finally, while already travelling at 99.999% of the speed of light, they are transferred into the LHC ring, both in the clockwise and counterclockwise direction. Here they are brought to half the center of mass (CM) energy of the LHC. From 2010 to 2011, data were collected at a CM energy of 7 TeV, while in 2012 it was raised to 8 TeV. After the upgrade starting in 2013, the LHC is expected to operate close to its design energy of 14 TeV. Figure 2.2 illustrates the sequence of accelerators used for proton-proton collisions.

The LHC is sometimes called a "high luminosity" collider. Luminosity is defined as the number of particles per unit time and area, and has units \([\text{s} \text{cm}^{-2}]\) or \([\text{b}]\).\(^1\) Multiplying this by the cross section of a specific process gives the rate at which that process occurs. Since Higgs boson production has a very small cross section, we need a high luminosity to be able to observe it. When the luminosity is integrated

\(^1\)One barn (b) is defined as \(10^{-28} \text{m}^2\).
Figure 2.2: Schematic view of the components of the LHC accelerator chain required for proton-proton collisions.
over time, we get a measure of the total data collected. Figure 2.3 shows the total integrated luminosity for proton collision data recorded at 8 TeV in 2012.

![Figure 2.3: Total integrated luminosity collected in 2012 [9].](image)

### 2.2 The ATLAS Detector

ATLAS stands for *A* *T*oroidal *L*HC *A*parrant. At 22 m high and 44 m long, it is the largest volume particle detector ever built for a collider experiment. As a general purpose detector, it must be able to measure a wide variety of particles and physics quantities in order to satisfy the many physics analyses. Various technologies are used to determine the identity, momentum and energy of particles. Consequently, the detector consists of three distinct layers: the inner detector, the calorimeter, and the muon spectrometer as shown in figure 2.4. A comprehensive description of the detector can be found in the ATLAS Technical Design Report [10] and in reference [11].

One objective in the design of the detector is to maximise the acceptance. The acceptance depends on the underlying process and provides the fraction of collision events that result in particles passing through the instrumented region of ATLAS
where they can be detected. This also improves the resolution of the missing transverse energy determination which is described in section 3.4. As a result, ATLAS subdetectors are typically divided into a central barrel region and an endcap region.

### 2.2.1 The ATLAS Coordinate System

Although various coordinate systems are used for ATLAS analyses, a common choice is to take the beam interaction point as the origin. In Cartesian coordinates, the $z$-axis is parallel with the beam, the $x$-axis points towards the center of the LHC ring, and the $y$-axis points up. Alternatively, spherical coordinates can also be used where:

\begin{align}
  r &= \sqrt{x^2 + y^2 + z^2} \\
  \phi &= \arctan\left(\frac{y}{x}\right)
\end{align}
\[ \theta = \arctan\left( \frac{\sqrt{x^2 + y^2}}{z} \right) \]  

(2.3)

Since particles are often Lorentz boosted along the \( z \)-axis, particle physicists often use pseudorapidity instead of \( \theta \). Differences in pseudorapidity are Lorentz invariant.

\[ \eta = -\ln(\tan \frac{\theta}{2}) \]  

(2.4)

The angular separation between two objects is commonly defined as:

\[ \Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2} \]  

(2.5)

Often we are interested in quantities as measured in the plane transverse to the beam. Energy \( (E) \) and momentum \( (\vec{p}) \) in the transverse plane are respectively defined as:

\[ E_T = E \sin \theta \]  

(2.6)

\[ p_T = |\vec{p}| \sin \theta = \sqrt{p_x^2 + p_y^2} \]  

(2.7)

### 2.2.2 The Inner Detector

The inner detector (ID) is designed to measure the trajectories and momentum of charged particles in the region \( |\eta| < 2.5 \) and pinpoint the location of primary and secondary vertices. Each component of the ID measures spatial points called hits that follow the trajectories of charged particles. These hits are then reconstructed into continuous paths called tracks. A superconducting solenoid generates a 2 Tesla magnetic field that bends the path of charged particles in the \( r-\phi \) plane as they pass through the ID. Based on the curvature of the track, it is possible to calculate the momentum of the particle. Figure 2.5 shows the relative position of the three components of the ID.

The silicon pixel detector is located closest to the beam line and provides measurements with the highest spatial resolution. Particles that pass through the silicon
pixels form electron-hole pairs. The pairs are separated by a 150 V bias voltage and the corresponding signal is then amplified. A pair is interpreted as a hit if the signal exceeds a preset threshold. Both the barrel and endcap regions are arranged in three layers and together consist of approximately 80 million readout channels. Since the pixels are so close to the interaction point, they must be able to withstand large amounts of radiation.

Surrounding the pixel detector is the semiconductor tracker (SCT) which consists of 4 double layered silicon strips which makes it easier to cover larger areas. The layers within each strip are oriented at different angles (stereo) such that the location of the $z$-coordinate can be determined. Approximately 6 million readout channels make up the SCT.

The outermost layer of the ID is the transition radiation tracker (TRT), a gas filled detector that uses drift tubes (also called straws) as sense wires. As charged particles pass through the TRT, they ionise the gas inside it. The resulting electrons and charged ions drift to the sense wires over a period of time commonly referred to as the drift time. Incorporating the drift time into the track reconstruction algorithm significantly improves the spatial resolution of the TRT. In contrast to the pixel
Besides aiding in the tracking of particles, the TRT also plays an important role in particle identification. When highly relativistic particles pass between media of different indices of refraction, they emit transition radiation which can be converted into an electric signal. Electrons are typically more relativistic than more massive particles and hence emit more transition radiation. Radiators are placed in between individual tubes causing the emission of transition radiation photons that ionise the gas in the TRT. Drift tube hits can be distinguished from transition radiation by their longer drift time and larger signals. Despite having less than 300,000 straws, the TRT still provides a significant contribution to the momentum resolution since it measures particles over a longer path than the silicon detectors.

2.2.3 The Calorimeter

The calorimeter is designed to measure the energy of particles by completely stopping them. Incoming high-energy particles interact with dense materials in the calorimeter. This causes them to produce multiple secondary particles, each carrying a fraction of the incoming energy. This process repeats itself until the incoming energy is absorbed and results in a cascade of particles called a shower.

In order to reduce the amount of energy lost by escaping particles, it is important that the calorimeter is deep enough. At the same time, there are constraints on its physical size. ATLAS makes use of sampling calorimeters which implies that layers of active material used to measure energy depositions are separated by dense materials that initiate showers. Even though this makes it possible to reduce the size of the calorimeter, a fraction of the energy is not measured and so the total shower energy must be calibrated.

Since different types of particles interact differently, the calorimeter is divided into an electromagnetic calorimeter (EMCAL) and a hadronic calorimeter (HCAL) as illustrated in figure 2.6. As the name implies, the EMCAL is designed to measure the energy of particles that primarily interact via the electromagnetic interaction such as electrons and photons. Liquid argon (LAr) is used as the active material and lead is used as the absorber. The thicknesses of the absorbers are optimised as a function
of $|\eta|$. Both the magnitude and location of energy depositions are precisely measured.

![Schematic view of the calorimeter](image)

Figure 2.6: Schematic view of the calorimeter [7].

For energies above $\sim 10$ MeV, the dominant energy loss mechanism for electrons (and positrons) is bremsstrahlung which occurs when the electron is deflected by another charged particle and radiates a photon. $X_0$, the radiation length, is defined as the distance over which an electron or positron loses an average of $1 - e^{-1}$ of its energy. The EMCAL is approximately $25X_0$ thick.

The HCAL surrounds the EMCAL and is designed to measure particles that interact primarily via the strong force, namely, hadrons. Note that hadrons also deposit energy in the EMCAL. In the barrel, scintillating tiles, which act as the active material, are separated by steel absorbers. The endcap and forward regions (see figure 2.6) use the same LAr technology as used in the EMCAL but with copper and tungsten absorbers.

Hadronic showers are more challenging to interpret than electromagnetic showers since there are more interactions that can occur. An incoming hadron will eventually interact strongly with a nucleus which initiates a shower of secondary hadrons that continues until there is no longer sufficient energy to produce secondary particles.
Hadronic showers are characterised by \( \lambda \), the interaction length, which is analogous to radiation length for electromagnetic showers. The hadronic calorimeter is approximately 11\( \lambda \) thick. When ATLAS data is prepared for physics analyses, hadron showers are reconstructed into objects known as jets. Since the LHC is a hadron collider, jets are the most commonly observed object in ATLAS.

### 2.2.4 The Muon Spectrometer

Unlike electrons, muons typically penetrate ATLAS’s calorimeters. For low-energy charged particles, ionisation is the main energy loss mechanism. This contribution drops as energy increases until the particle becomes relativistic. The energy loss due to ionisation increases logarithmically as a function of energy for relativistic particles. Particles that fall in the regime where the energy loss due to ionisation is minimised are referred to as minimum ionising particles. Muons are approximately minimum ionising in ATLAS.

Energy loss due to bremsstrahlung increases as particle energy increases and is inversely proportional to mass squared. Due to their relatively large mass when compared to electrons, few muons will lose energy via bremsstrahlung.

As illustrated in figure 2.7, the muon spectrometer (MS) is by far the largest volume subdetector. Since very few particles besides muons are expected to penetrate the detector and leave hits in the MS, it is key in identifying muons and measuring their momentum. Toroidal magnets provide a magnetic field which varies between 2 and 8 Tesla. In contrast to the inner detector, the particles are bent in the \( r-z \) plane.

In principle, the methods used to track muons in the central region are similar to that of the TRT. It is made up of monitored drift tubes (MDT) which are metal, gas-filled tubes that surround a central wire. By using the drift time of electrons and ions produced by the muon passing through, it is possible to infer position measurements of the muon’s trajectory. Forward regions are equipped with cathode strip chambers (CSC) which are multiwire proportional chambers with the cathode divided into strips. Triggering is provided by resistive plate chambers (RPC) in the barrel and thin gap chambers (TGC) in the endcap regions.
2.2.5 The ATLAS Trigger System

At design luminosity, there will be 40 million bunch crossings inside ATLAS every second. Recording all these collision data is impossible as this would result in unmanageable amounts of data. ATLAS uses a three level trigger system to remove uninteresting events to a more manageable $\sim 400$ interactions per second.

The level 1 (L1) trigger reduces the event rate to $\sim 100,000$ Hz using coarse signals from the calorimeter and muon spectrometer. In addition, it also defines "regions of interest" which are subsets of the event that contain full granularity detector signals of particle candidates. Regions of interest are not used in the L1 trigger’s decision, but are passed to the level 2 (L2) trigger which further reduces the event rate to 2,000 Hz. The L2 trigger bases its decision on the full granularity of the ATLAS detector in the regions of interest. By means of a processor farm, it is possible to filter many events simultaneously. Finally, the event filter (EF) is a software based trigger that uses a computing cluster to reconstruct and classify events. The EF reduces the event rate to $\sim 400$ Hz.
2.2.6 Data Quality

It is essential that the quality of the collected data meets certain criteria. Data quality monitoring (DQM) takes place in two stages, first online in the ATLAS control room during the data taking, and then offline after the data have been stored. As part of my service-work for ATLAS, I developed new tools that are used both online and offline.

In the online environment, ATLAS Global Monitoring (AGM) produces data quality histograms using the express stream, a subset of the incoming ATLAS data based on specific triggers. A computing cluster of 48 cores fully reconstructs these events into physics objects using ATHENA [12]. ATHENA is a software framework used by ATLAS for a wide variety of purposes such as reading data, event simulation, and event reconstruction. Quality control histograms are generated based on these events and analysed by the online data quality (DQ) shifter in the ATLAS control room. Incoming data are monitored in real-time, making it possible to spot and analyse problems as soon as they arise. A more thorough data quality check is performed offline where the collected ATLAS dataset is scoured for anomalies.

Online DQ shifters use the Online Histogram Presenter (OHP) to monitor the data quality histograms as shown in figure 2.8. However, this software framework is no longer being maintained and so it was necessary to move to a new framework. One of my service-work tasks was to transition from OHP to the Data Quality Monitoring Display (DQMD). As illustrated in figure 2.9, this framework is already being used by other subsystems.

An advantage of the new framework is that it is easy to implement DQ algorithms. DQ algorithms are automated histogram checks that aid the DQ shifter in identifying problems. If a histogram fails a DQ check, it is marked by a red flag which alerts the shifter of a potential problem. In the example illustrated, at least one of the criteria encoded in the DQ algorithms for the tile monitoring is not satisfied.

From a technical point of view, online DQ histograms are stored in the Information Service (IS) and are periodically updated as data are collected. Configuration files are used to organise the histograms into categories and subcategories. For example, when a button in figure 2.9 is clicked, the configuration file is consulted to determine
Figure 2.8: The Online Histogram Presenter is a tool used by the online shifter to monitor DQ histograms.

Figure 2.9: The Data Quality Monitoring Display is replacing the Online Histogram Presenter since the latter is no longer maintained.
which subcategories are displayed. These may either contain further subcategories or DQ histograms. When a histogram name is clicked, it is accessed from the IS and displayed.

The above mentioned configuration files also specify which DQ algorithm is used for each histogram. The DQAgent reads the configuration files and applies the correct DQ algorithm to each histogram in the IS. On completion, it attaches a status flag to the histogram. The DQAgent repeatedly evaluates DQ algorithms while ATLAS is taking data and updates flags accordingly.

One probe to measure the quality of data is by monitoring the rate at which resonance particles are produced. This idea forms the basis of the second component of my service-work. Since we expect production rates to be proportional to the luminosity, any deviation from a smooth non-increasing function would indicate a problem. The following processes are monitored in online and offline DQ:

- $K_S^0 \rightarrow \pi^+\pi^-$
- $J/\psi \rightarrow \mu^+\mu^-$
- $\Upsilon \rightarrow \mu^+\mu^-$
- $Z \rightarrow \mu^+\mu^-$
- $Z \rightarrow e^+e^-$

ATHENA is used to reconstruct ATLAS events into physical objects such as pions, muons, and electrons. A simple sideband subtraction method is used to count the number of resonance particles produced in a fixed time interval. First, the Lorentz invariant mass (see equation 3.2) is calculated for the resonance particle from its decay products. The invariant mass distribution is plotted as shown for $K_S^0 \rightarrow \pi^+\pi^-$ in figure 2.10. Next, the number of events in the background region (shaded red) is subtracted from the signal region (shaded green). This process is repeated every 40 lumiblocks, where one lumiblock is approximately 2 minutes of data taking. The total count every 40 lumiblocks is recorded in a rate monitoring histogram as shown in figure 2.11.
Figure 2.10: Invariant mass distribution of $K^0_S \rightarrow \pi^+\pi^-$. The event count is calculated by subtracting background (B) from signal (S) (Plot shown is for ATLAS run 214523).

Figure 2.11: $K^0_S \rightarrow \pi^+\pi^-$ particle count as a function of luminosity (Plot shown is for ATLAS run 214523).
For rate monitoring, a simple algorithm checks to make sure that the slope of a straight line fit to the data is negative. Note that the first and last bin in the figure (circled in orange) do not follow the general trend of the data. This is because these bins typically cover less than 40 luminosity blocks and consequently are ignored by the algorithm. Note that not all anomalies will be flagged by this algorithm. DQ shifters are advised to scan the rate monitoring histograms for irregularities.

The last component of my service-work involved constructing a monitor that ensures that all the triggers of the express stream are firing. Two histograms are used, with each bin corresponding to one trigger as illustrated in figure 2.12. Each time a trigger fires, its corresponding bin is incremented by one. An algorithm checks to make sure that all the bins are filled with at least one entry.

Figure 2.12: One of the two histograms used to monitor the operation of the express stream triggers (Plot shown is for ATLAS run 209109).

2.3 Simulated Data

In order to search ATLAS data for new physics, they are compared to simulated data commonly referred to as Monte Carlo (MC) simulation. MC modelling is a challenging
and time-consuming procedure that simulates the production and subsequent decay of particles within the detector. The ATLAS simulation infrastructure is described in detail in reference [13].

The first stage of MC simulation is typically referred to as event generation. For a specific process, a software package, known as the generator, is used to simulate the interaction of two incoming particles and the resulting decay products. It is responsible for prompt decays, such as that of the massive gauge bosons, but not for the decay of particles that pass through any part of the detector. For example, consider SM $WW$ production where the $W$ bosons decay into leptons as represented by the Feynman diagram on the right in figure 3.1. The generator simulates the interaction starting with the incoming quarks and ending with the production of the two neutrinos and two charged leptons. Note that the generator does not require a description of the detector.

A record is maintained describing all the particles produced by the generator. The kinematics of particles at generator level are commonly referred to as truth kinematics. In contrast, reconstructed kinematics are quantities as measured after the event has passed through the entire detector simulation. A useful feature of MC simulation is that it is possible to compare reconstructed quantities with truth quantities.

The constituent quarks and gluons of the proton are called partons. In order to generate an interesting event, the colliding partons must have large energies. Consequently, they radiate gluons (QCD radiation) in the same way high-energy electric charges emit photons. However, unlike photons which are chargeless, the radiated gluons carry color charge and so can radiate further gluons. This results in a spray of partons called a parton shower. Furthermore, when two quarks are separated in a high-energy collision, the potential energy in the color field grows until enough energy is present to create a particle-antiparticle pair out of the vacuum. This process repeats until the original energy is dissipated, producing multiple hadrons. Hadronisation is the process of forming hadrons out of quarks. Due to color confinement only color-neutral combinations of quarks that form physical hadrons are observed. As the energy scale increases, the strong interaction coupling constant decreases (refer to figure 1.1). As a result, the probability of hadronisation increases with decreasing quark energy. Both parton showering and hadronisation are modelled using the

Secondly, the response of the detector is simulated. The kinematics of each particle produced by the generator are passed to the detector simulation. For the MC event to be representative of ATLAS data, the simulation must be fully aware of ATLAS’s geometry and its response to particles interacting with each detector component. The GEANT4 [17] environment uses a full description of the ATLAS detector’s composition along with a knowledge of the relevant physics processes to predict how particles interact in the detector. For example, the ionisation that results when a charged particle passes through the tracker is simulated along with the resulting analog hits. Similarly, electromagnetic and hadronic showers are simulated using information on how particles interact with the active and absorbent materials used in the calorimeters. This stage requires enormous amounts of computing power and typically takes several minutes for a single event.

GEANT4 also simulates the conversion of analog hits to digital readouts. ATHENA is used to reconstruct both MC simulation and ATLAS data into physics objects which can be used in analyses. The output from the simulation is formatted identically to that of ATLAS’s data acquisition system (DAQ) in order to facilitate data and MC comparisons.

For each bunch crossing, multiple interactions may take place, an effect commonly referred to as in-time pileup. In addition, the colliding parts of the two protons that provide sufficient CM energy to produce an interesting event are referred to as the hard scatter. Other interactions within the same collision which do not result in interesting events are called the underlying event. Both pileup and the underlying event are incorporated at digitisation in order to save CPU time.

Each of the above described steps is performed by modularised software packages in order to make it possible to run multiple processes at once. Large-scale production of MC samples is executed on the Worldwide LHC Computing Grid, a global collaboration of computing centers. Simon Fraser University hosts one of the LHC’s computing clusters.
Chapter 3

Higgs Boson Kinematics

Despite the relatively large branching fraction of $H \rightarrow WW$ events, they must com-
pete with irreducible background from $WW$ production of the SM. Even though
both processes have identical signatures as shown in figure 3.1, it is still possible to
distinguish them from each other using subtle differences in their event kinematics.

Figure 3.1: Representative Feynman diagrams for gluon fusion Higgs boson produc-
tion with subsequent decay to $W$ bosons (left) and the dominant SM background
from $WW$ production (right).
CHAPTER 3. HIGGS BOSON KINEMATICS

3.1 Conservation of Angular Momentum

The SM Higgs boson is a spin-0 particle and conservation of angular momentum demands that the total spin of its decay products is also zero. Quantum mechanics only allows us to measure the total spin and its projection along one axis which we will call the $z$-axis. Since $W$ bosons are massive and have a spin of 1, there are three possible values for the $z$-component of their spin, -1, 0 and +1. Consequently, there are exactly three spin configurations where the total spin is conserved. They are summarised in figure 3.2. The long thin arrows represent the direction of the particle’s momentum, while the short thick arrows denote the $z$-component of the spin. We define helicity, $\lambda$, as the normalised projection of the particle’s spin onto its direction of motion:

$$\lambda = \frac{\mathbf{s} \cdot \mathbf{p}}{|s||p|}$$

(3.1)

where $\mathbf{s}$ and $\mathbf{p}$ are the particle’s spin and momentum vector respectively. Note that a pair of $W$’s originating from Higgs boson decay always share the same helicity state.

Figure 3.2: The $z$-component of the spin can be aligned with the direction of motion in three different ways, parallel, antiparallel, or perpendicular. Each corresponds to one helicity state.

Under the assumption that neutrinos are massless, the underlying structure of the weak interaction requires that neutrinos have negative helicity while antineutrinos...
have positive helicity\(^1\). This implies that neutrinos can only travel opposite to their spin, and vice versa for antineutrinos. As a result, each of the spin configurations in figure 3.2 can only decay in one way as shown in figure 3.3. Note that each scenario favors both charged leptons moving in the same direction. Consequently, we expect small opening angles between measured leptons (\(\Delta \psi\)) from Higgs boson events. Small opening angles also lead to a small dilepton mass, \(m_{\ell\ell}\). Consider the mass of the dilepton system in terms of the energies (\(E_{\ell_i}\)) and momenta (\(p_{\ell_i}\)) of the two leptons.

\[
m_{\ell\ell}^2 = (E_{\ell_1} + E_{\ell_2})^2 - (p_{\ell_1} + p_{\ell_2})^2
\]

Figure 3.3: The three possible spin configurations for the decay of the \(H \rightarrow WW\) system. Notice that the charged leptons preferentially travel in the same direction.

\(^1\)This is a consequence of the V-A structure of the weak interaction. The weak interaction only acts on left-handed particles and right-handed antiparticles [5].
### CHAPTER 3. HIGGS BOSON KINEMATICS

\[ m_{\ell\ell}^2 = (E_{\ell_1}^2 - \vec{p}_{\ell_1}^2) + (E_{\ell_2}^2 - \vec{p}_{\ell_2}^2) + 2E_{\ell_1}E_{\ell_2} - 2|\vec{p}_{\ell_1}| |\vec{p}_{\ell_2}| \cos(\Delta \psi) \]  

(3.3)

Assuming the masses of the leptons are much smaller than their energy:

\[ m_{\ell\ell}^2 \approx 2E_{\ell_1}E_{\ell_2}(1 - \cos(\Delta \psi)) \]  

(3.4)

Note that \( m_{\ell\ell} \) will be small when \( \Delta \psi \approx 0 \).

Recall that the dominant background is SM \( WW \) production originating from quark and antiquark pairs. Quarks are spin-\( \frac{1}{2} \) particles and so the \( z \)-component of the \( WW \) system’s spin can be -1, 0, or +1. Notice that this allows for states where the two \( W \)’s have different helicities which is suppressed for Higgs boson events. As a result, there are significant differences between the kinematic distributions for SM \( WW \) and 125 GeV Higgs boson production. Figure 3.4 illustrates the opening angle between the charged leptons in the transverse plane (\( \Delta \phi \)) while figure 3.5 shows the mass of the dilepton system (\( m_{\ell\ell} \)). Note that these distributions, as well as the ones that follow in this chapter, have been normalised to unit area. In the ATLAS experiment, Higgs boson events are produced at a much lower rate than SM \( WW \) production. In addition, the selection criteria described in section 4.1.2 before the analysis is divided into same and opposite flavor channels are applied to generate the plots in this section.

### 3.2 \( H \rightarrow WW \) Production at the LHC

As mentioned previously, the LHC is a proton-proton collider, which has implications on the \( H \rightarrow WW \) analysis. The constituent parts of the proton that may interact are valence quarks, sea quarks, and gluons, and are collectively called partons. Protons consist of three valence quarks, two up quarks and one down quark. Sea quarks are quark-antiquark pairs that are spontaneously produced and annihilated inside the proton. Finally, quarks are bound together inside the proton by gluons.

The fractions of the proton’s total momentum carried by the two colliding partons are denoted by \( x_1 \) and \( x_2 \). The distribution of the momentum fraction is called the parton distribution function (PDF). Figure 3.6 shows the PDFs for a momentum
CHAPTER 3. HIGGS BOSON KINEMATICS

Figure 3.4: Spin conservation favors a small opening angle between charged leptons for Higgs boson events in contrast to SM WW events. The distributions shown are plotted using reconstructed quantities in MC.

Figure 3.5: Since Higgs boson events tend to have a small dilepton opening angle ($\Delta \psi$) when compared to WW events, they also tend to have a smaller $m_{\ell\ell}$. The distributions shown are plotted using reconstructed quantities in MC.
transfer of 125 GeV. Recall from figure 3.1 that SM $WW$ events are predominantly produced via quark-antiquark annihilation. The incoming quark can be a sea or valence quark, while the antiquark can only be a sea quark. Since valence quarks tend to carry more of the proton’s momentum than sea quarks, we expect to see $WW$ events that are boosted along the beam axis. In contrast, Higgs boson events are produced by two gluons with identical PDFs. Of course, this does not imply that the system cannot be boosted in the $z$-direction, but on average we expect $WW$ events to be boosted more than Higgs boson events. Figure 3.7 shows the absolute difference between $x_1$ and $x_2$ for Higgs boson and $WW$ events.

![Figure 3.6: CT10 PDF for the proton at a momentum transfer of 125 GeV [18].](image)

Figure 3.6: CT10 PDF for the proton at a momentum transfer of 125 GeV [18].
Figure 3.7: \( WW \) events tend to be more boosted along the beam axis than Higgs boson events and so the absolute difference between momentum fractions \( x_1 \) and \( x_2 \) is larger for \( WW \) events. The distributions shown are plotted using truth quantities in MC.

### 3.3 The Transverse Plane

The Feynman diagram for SM \( WW \) production shown in figure 3.1 is a leading order diagram. This is the simplest diagram with the smallest number of vertices that takes initial state particles to final state particles. A higher order diagram, like the Feynman diagram for gluon fusion shown in figure 3.1, has additional vertices to accommodate loop corrections or radiation of particles. At lowest order, the Higgs boson is produced at rest in the transverse plane. However, the Higgs boson typically carries non-zero transverse momentum (figure 3.8) due to radiative corrections as illustrated in figure 3.9. An additional complicating factor is that the QCD evolution equations\(^2\) predict that gluons have a higher probability of radiating particles than quarks. As a result, Higgs boson events have a larger transverse boost on average than \( WW \) events.

Unfortunately, this boost affects the kinematics of the Higgs boson’s decay prod-

\(^2\)The QCD evolution equations are commonly referred to as DGLAP equations after the names of their authors.
Figure 3.8: Higgs boson $p_T$ distribution at a center of mass energy of 8 TeV for events with no high-energy reconstructed jets.

Figure 3.9: Initial state radiation imparts a transverse boost on the Higgs boson.
ucts. Since kinematic distributions are instrumental in searching for the Higgs boson, it is important to take its boost into account.

### 3.4 Kinematic Observables

The leading lepton is defined as the lepton with the largest $p_T$, while the subleading lepton has the second largest $p_T$. The leptons originating from Higgs boson decay are measured with high accuracy. Figures 3.10 and 3.11 show the $p_T$ distributions for the leading and subleading leptons under the assumption that $m_H = 125$ GeV. Charged leptons originating from Higgs boson decay tend to have a smaller $p_T$ because the assumed mass usually forces at least one of the $W$’s to be off its mass-shell. As a result, the $WW$ system from a Higgs boson event will have a significantly smaller mass than one from SM $WW$. Consequently, less energy will be available to impart momentum to the leptons.

![p_T Distribution of Leading Lepton](image)

**Figure 3.10:** Leading lepton $p_T$ distribution for Higgs boson and SM $WW$ production. The distributions shown are plotted using reconstructed quantities in MC.

Even though ATLAS cannot detect neutrinos directly, the missing transverse energy, $E_T^{\text{miss}}$, provides an indication of the dineutrino system’s momentum in the transverse plane. By conservation of momentum, we expect the net transverse momentum
Figure 3.11: Subleading lepton $p_T$ distribution for Higgs boson and SM WW production. The distributions shown are plotted using reconstructed quantities in MC.

$E^\text{miss}_T$ can be based on information from the calorimeter or the tracker. Calorimeter $E^\text{miss}_T$ is calculated from energy deposits in the calorimeter and muons reconstructed in the muon spectrometer. It is defined in Cartesian coordinates as:

$$E^\text{miss}_x(y) = E^\text{miss, e}_x(y) + E^\text{miss, \gamma}_x(y) + E^\text{miss, \tau}_x(y) + E^\text{miss, jets}_x(y) + E^\text{miss, calo-\mu}_x(y) + E^\text{miss, CellOut-Eflow}_x(y) + E^\text{miss, \mu}_x(y)$$  \hspace{1cm} (3.5)

The first five terms correspond to physics objects measured in the calorimeter, namely electrons, photons, hadronically decaying taus, jets, and muons. Each term is calculated as the negative sum of calorimeter cell energies inside that particular physics object. Note that energy deposits are also calibrated differently for different physics objects. $E^\text{miss, CellOut-Eflow}_x(y)$ corresponds to the negative sum of energy deposits in the calorimeter not corresponding to a reconstructed physics object. Since muons are minimum ionising, $E^\text{miss, calo-\mu}_x(y)$ is typically only a fraction of the total muon energy. As a result, $E^\text{miss, \mu}_x(y)$, the negative sum of the reconstructed muon energy inferred from momentum measurements in the muon spectrometer, is added to the equation. Full
details of the missing transverse energy reconstruction can be found in reference [19].

Track $E_T^{\text{miss}}$ is defined as the negative $p_T$ sum of all tracks in the inner detector that have the following properties:

- $p_T > 500$ MeV as measured with respect to primary vertex
- $|\eta| < 2.5$
- $|d_0| < 1.5$ mm, where $d_0$ is the transverse distance to the primary vertex
- $|z_0| < 1.5$ mm, where $z_0$ is the longitudinal distance to the primary vertex
- at least one hit in silicon pixel detector
- at least six hits in the SCT

Figures 3.12 and 3.13 show the $E_T^{\text{miss}}$ distributions as measured by the calorimeter and the tracker respectively.

![Calorimeter Missing Transverse Energy Distribution](image)

Figure 3.12: Calorimeter $E_T^{\text{miss}}$ distribution for Higgs boson and SM $WW$ production. The distributions shown are plotted using reconstructed quantities in MC.

A powerful kinematic observable used to search for the Higgs boson is its invariant mass. However, in the $H \rightarrow WW \rightarrow \ell\nu\ell\nu$ channel this variable cannot be reconstructed since the neutrinos escape detection. Instead, we define the Higgs boson
transverse mass $m_T$, which is correlated to the invariant mass and uses calorimeter $E_T^{\text{miss}}$ to estimate the transverse momenta of neutrinos.

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

(3.6)

where $E_T^{\ell\ell} = \sqrt{|\vec{p}_T^{\ell\ell}|^2 + m_T^{2\ell\ell}}$.

From figure 3.14 it appears easy to apply a cut on $m_T$ that would provide us with a signal rich sample. However, recall these distributions have been normalised to unit area and that the expected background is many times larger than the expected signal.
Figure 3.14: Transverse mass ($m_T$) distribution for Higgs boson and SM $WW$ production. The distributions shown are plotted using reconstructed quantities in MC.
Chapter 4

$H \rightarrow WW$ Analysis

The $H \rightarrow WW \rightarrow ℓνℓν$ search channel is expected to provide good sensitivity for Higgs boson masses between 120 GeV and 200 GeV and is one of the channels presented in the ATLAS discovery paper for a Higgs-like boson [1]. The detailed results of this channel are also summarised in a conference note published in August 2012 [20]. As this analysis forms the basis for later chapters of this thesis, important aspects are summarised here.

4.1 Event Selection

4.1.1 Triggers

The $H \rightarrow WW \rightarrow ℓνℓν$ analysis uses the inclusive single-electron and single-muon triggers that require that the $p_T$ of the lepton exceeds 24 GeV. In addition, the lepton must be isolated which means that the scalar sum of $p_T$ of all charged particle tracks within a cone of $\Delta R < 0.2$ of the lepton’s direction, excluding the lepton itself, is smaller than 0.12 $p_T$ for muons and 0.10 $p_T$ for electrons. The muon trigger only accepts events with $|\eta| < 2.4$ due to the geometry of the detector.
4.1.2 Event Selection

Data quality criteria are applied in order to reduce backgrounds not originating from the collision such as cosmic ray muons, beam-related backgrounds, and noise from the calorimeters. Events must have a primary vertex with at least three associated tracks with $p_T > 400 \text{ MeV}$. Candidate $H \rightarrow WW \rightarrow \ell\nu\ell\nu$ events (where $\ell$ can be either an electron or a muon) are required to satisfy the following criteria:

- must have exactly two leptons with opposite charge
- leading lepton has $p_T > 25 \text{ GeV}$
- subleading lepton has $p_T > 15 \text{ GeV}$
- muons must have $|\eta| < 2.5$
- electrons must have $|\eta| < 2.47$
- electrons in $1.37 < |\eta| < 1.52$, the region where the barrel and end-cap calorimeters meet, are excluded

A tight set of identification criteria are applied to electrons using information from both the tracker and the calorimeter. Thanks to the TRT and the fine segmentation of the calorimeter, the electron performance has not degraded after the increase in luminosity in 2012 which resulted in more pileup. Muon candidates are selected by matching tracks reconstructed in the inner detector and the muon spectrometer. Hits are required in all three components of the inner detector in order to reject muons from pions and kaons that decay as they pass through.

At least one of the leptons must match a triggering object. Next, both leptons are required to be isolated in order to suppress jets that pass the lepton criteria. This means that the scalar sum of $p_T$ of all charged tracks and energy deposits in $\Delta R < 0.3$ of the lepton’s direction, excluding the lepton itself, is smaller than $x p_T$. $x$ is some fraction that varies from 0.12 to 0.20 and depends on the $p_T$ of the lepton, whether the energy is measured in the tracker or the calorimeter, and whether the lepton is an electron or a muon.
At this point, the analysis is divided into two parts, a same flavor ($ee$ and $\mu\mu$) and an opposite flavor ($e\mu$ and $\mu e$) channel. 8 TeV results are only published for the opposite flavor analysis as it is much easier to reduce the Drell-Yan\footnote{Drell-Yan events are produced when a quark and antiquark pair annihilate to produce a photon or $Z$ boson which then decays into oppositely charged leptons with the same flavor.} background in this channel. For the opposite flavor analysis, the invariant mass of the dilepton system, $m_{\ell\ell}$, is required to be above 10 GeV.

From equation 3.5, we define:

\[
\vec{E}_{T}^{\text{miss}} = (E_{x}^{\text{miss}}, E_{y}^{\text{miss}}) \tag{4.1}
\]

Taking $\Delta\phi_{\text{min}}$ to be the minimum azimuthal angle between $\vec{E}_{T}^{\text{miss}}$ and either lepton or any reconstructed jet with $p_T > 25$ GeV, the relative missing transverse energy is defined as:

\[
E_{T,\text{rel}}^{\text{miss}} = \begin{cases} 
E_{T}^{\text{miss}}, & \Delta\phi_{\text{min}} \geq \pi/2 \\
E_{T}^{\text{miss}} \sin \Delta\phi_{\text{min}}, & \Delta\phi_{\text{min}} < \pi/2 
\end{cases} \tag{4.2}
\]

Applying a cut on the missing energy helps reject Drell-Yan and multi-jet events. $E_{T,\text{rel}}^{\text{miss}}$ is favored over $E_{T}^{\text{miss}}$ since it rejects more events with mismeasured objects. In such events the direction of the mismeasured object and $\vec{E}_{T}^{\text{miss}}$ are related. We demand $E_{T,\text{rel}}^{\text{miss}} > 25$ GeV. Figure 4.1 shows the $E_{T,\text{rel}}^{\text{miss}}$ after applying the selection mentioned above except the cut on $E_{T}^{\text{miss}}$.

The anti-$k_t$ algorithm \cite{21} with $R = 0.4$ is used for jet reconstruction. It clusters energy depositions by giving preference to large energy deposits, particularly those that are close together in $\eta - \phi$ space. The parameter $R$ represents the size of the jet in $\eta - \phi$ space. Only jets that satisfy $p_T > 25$ GeV and $|\eta| < 4.5$ are included.

With the increase in luminosity from 2011 to 2012, the average number of interactions per bunch crossing has approximately doubled to 20. In order to reduce the impact of pileup, we demand $p_T > 30$ GeV in the forward region where $2.5 < |\eta| < 4.5$. For the central region, we define the jet vertex fraction (JVF) as the $p_T$ sum of charged tracks coming from the primary vertex divided by the $p_T$ sum of all charged tracks. For $|\eta| < 2.5$, we demand $|\text{JVF}| > 0.5$ which makes the jet count approximately pileup independent.
Figure 4.1: $E_{T,\text{rel}}^{\text{miss}}$ distribution for opposite flavor events satisfying all the selection cuts up to but not including the cut on $E_{T,\text{rel}}^{\text{miss}}$. The superimposed red line indicates the expected signal for a SM Higgs boson with $m_H = 125$ GeV. The diagonal grey lines represent the total uncertainty on the predicted background [20].

As apparent in figure 4.2, the data composition changes significantly as the number of jets changes. The current analysis focuses on the 0-jet and 1-jet channels where the dominant Higgs boson production mode is gluon fusion. The contribution from vector boson fusion is not significant in these channels, but dominates for jet multiplicities of two and larger. Also notice the shift from a $WW$ dominated background in the 0-jet channel to a $t\bar{t}$ dominated background in the 2-jet channel.

As mentioned in chapter 3, both the mass of the dilepton system and the opening angle between the charged leptons for Higgs boson events are expected to be small. As a result, demanding $m_{\ell\ell} < 50$ GeV and $\Delta\phi < 1.8$ enhances the signal purity. In addition, $p_T^{\ell\ell}$, the transverse momentum of the dilepton system, is required to be greater than 30 GeV.

Figures 4.3 through 4.5 depict kinematic distributions after the above described event selection. Note how the selection defines a signal enriched region.

Figure 4.6 shows the ATLAS event display for a $H \rightarrow WW \rightarrow e\nu\mu\nu$ candidate event that satisfies the selection cuts for the 0-jet channel. Note the characteristic
CHAPTER 4. $H \rightarrow WW$ ANALYSIS

Figure 4.2: Number of jets in the opposite flavor channel after the selection described in the text up to and including the cut on $E_{T,\text{rel}}^{\text{miss}}$. The superimposed red line indicates the expected signal for a SM Higgs boson with $m_H = 125$ GeV. The diagonal grey lines represent the total uncertainty on the predicted background [20].

Figure 4.3: $p_T^{\ell\ell}$ distribution after full event selection for the 0-jet analysis. The expected signal for a SM Higgs boson with $m_H = 125$ GeV is shown in red. The diagonal grey lines represent the total uncertainty on the predicted background [20].
CHAPTER 4. $H \rightarrow WW$ ANALYSIS

Figure 4.4: $\Delta \phi$ distribution after full event selection for the 0-jet analysis. The expected signal for a SM Higgs boson with $m_H = 125$ GeV is shown in red. The diagonal grey lines represent the total uncertainty on the predicted background [20].

Figure 4.5: $m_{\ell\ell}$ distribution after full event selection for the 0-jet analysis. The expected signal for a SM Higgs boson with $m_H = 125$ GeV is shown in red. The diagonal grey lines represent the total uncertainty on the predicted background [20].
small opening angle between the two leptons. This event has the following kinematics:

- electron $p_T$ is 33 GeV
- muon $p_T$ is 29 GeV
- $E_{T,rel}^{miss}$ is 35 GeV
- $m_T$ is 94 GeV

Figure 4.6: ATLAS event display for a candidate $H \rightarrow WW \rightarrow e\nu\mu\nu$ event that satisfies all the 0-jet channel selection cuts. The electron is shown in green, the muon in red, and the $E_{T}^{miss}$ in purple [20].

4.2 Monte Carlo Modelling

Table 4.1 summarises the MC samples used in the $H \rightarrow WW$ analysis. For the majority of processes, different programs are used to model different parts of the event simulation. The first name listed in the table is the generator used to simulate
the hard scatter while the second name is the program used to model the parton showering, hadronisation, and the underlying event. Whenever Herwig [16] is used for the parton showering and hadronisation, the underlying event is modelled using Jimmy [15].

<table>
<thead>
<tr>
<th>Process</th>
<th>Generator</th>
</tr>
</thead>
<tbody>
<tr>
<td>$ggF$</td>
<td>PowHeg [22] + Pythia8 [23]</td>
</tr>
<tr>
<td>VBF</td>
<td>PowHeg [24] + Pythia8</td>
</tr>
<tr>
<td>$WH/ZH$</td>
<td>Pythia8</td>
</tr>
<tr>
<td>$q\bar{q}/g \to WW$</td>
<td>MC@NLO [25] + Herwig [16]</td>
</tr>
<tr>
<td>$gg \to WW$</td>
<td>gg2WW [26] + Herwig</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>MC@NLO + Herwig</td>
</tr>
<tr>
<td>$tW/tb$</td>
<td>MC@NLO + Herwig</td>
</tr>
<tr>
<td>$tqb$</td>
<td>AcerMC [27] + Pythia [14]</td>
</tr>
<tr>
<td>inclusive $W$</td>
<td>Alpgen [28] + Herwig</td>
</tr>
<tr>
<td>inclusive $Z/\gamma^*$</td>
<td>Alpgen + Herwig</td>
</tr>
<tr>
<td>$Z^<em>Z^</em> \to 4l$</td>
<td>PowHeg + Pythia8</td>
</tr>
<tr>
<td>$W(Z/\gamma^<em>)(m_{Z/\gamma^</em>} &gt; 7 \text{ GeV })$</td>
<td>PowHeg + Pythia8</td>
</tr>
<tr>
<td>$W(Z/\gamma^<em>)(m_{Z/\gamma^</em>} &lt; 7 \text{ GeV })$</td>
<td>Madgraph [29, 30] + Pythia</td>
</tr>
<tr>
<td>$W\gamma^*$</td>
<td>Madgraph [31] + Pythia</td>
</tr>
<tr>
<td>$W\gamma$</td>
<td>Alpgen + Herwig</td>
</tr>
</tbody>
</table>

Table 4.1: Monte Carlo samples used in $H \to WW$ analysis.

For the PowHeg and MC@NLO samples, the CT10 [32] PDF is used while for Alpgen, MadGraph, and Pythia8, CTEQ6L1 [33] is used.

### 4.3 Background Estimation

The main sources of SM background that result in two isolated, high-$p_T$ leptons and missing transverse energy are $WW$ and top quark pair production ($t\bar{t}$). SM $WW$, $t\bar{t}$, and single top production are modelled using MC and normalised to the observed rates in data control regions that are rich in the process in question. The Drell-Yan
background and diboson processes besides $WW$ production are fully estimated using simulation.

A data-driven estimate is used for the $W$+jets background which is very different from the other backgrounds. It is generated by selecting data events where both leptons satisfy loose identification and isolation criteria, and exactly one satisfies tight criteria. Most events in this sample come from $W$+jets events where the jets generate an object that is reconstructed as a lepton. In this control region, approximately 90% of events in the electron channel, and 80% of events in the muon channel are $W$+jet events.

Next, an inclusive dijet data sample is used to determine the ratio of leptons satisfying tight criteria to those satisfying loose criteria. The number of events in the control region is scaled by this ratio. Finally, the contamination of $W\gamma$ and $W\gamma^*/WZ^*(\gamma)$ events is removed by subtraction using MC samples.

For the top quark control sample, the selection criteria described in section 4.1.2 are applied up until, but not including, the cut on jet multiplicity. The number of background events from top quark production is normalised to the number of events that pass these criteria in data. As illustrated in figure 4.2, this region is dominated by top quark events. The contribution from the $W$+jets background is estimated from data, while the other backgrounds are estimated using MC. The normalisation factor for this control region is $1.11 \pm 0.06$ (stat).

The $WW$ control region is defined by applying all the signal selection criteria, removing the $\Delta\phi$ cut, and replacing the $m_{\ell\ell}$ cut with $m_{\ell\ell} > 80$ GeV. The normalisation factor for the $WW$ control region is $1.06 \pm 0.06$ (stat).

### 4.4 Results

The $H \rightarrow WW$ analysis described in the previously mentioned conference note uses transverse mass in order to test for the presence of a signal (figure 4.7). Using 2012 data, a broad excess of events is observed for $m_H < 150$ GeV which is consistent with the excess observed in the $\gamma\gamma$ and $ZZ$ channel. The excess has a significance of 3.2 standard deviations at 120 GeV.

Figure 4.8 shows the $m_T$ distribution in data after subtracting the predicted back-
CHAPTER 4.  $H \rightarrow WW$ ANALYSIS

Figure 4.7: $m_T$ distribution after event selection for the 0-jet analysis. The expected signal for a SM Higgs boson with $m_H = 125$ GeV is shown in red. The diagonal grey lines represent the total uncertainty on the predicted background [20].

$\sqrt{s} = 8$ TeV, $\int L dt = 5.8$ fb$^{-1}$

$H \rightarrow WW^* \rightarrow \ell \nu \ell' \nu'$ channel does not provide accuracy for a mass measurement. The local $p_0$ and significance of the excess is shown in figure 4.9. For full details of the analysis, refer to the conference note [20].
Figure 4.8: $m_T$ distribution in data with the predicted background subtracted. The predicted signal for $m_H = 125$ GeV is overlaid in red. The error bars represent the statistical error in the data and subtracted background. Note that this plot also includes the 1-jet contribution [20].

Figure 4.9: Observed and expected local $p_0$ value (left $y$-axis) and significance (right $y$-axis) for SM $H \rightarrow WW \rightarrow l\nu\ell\nu$ as a function of Higgs boson mass [20].
Chapter 5

The Matrix Element Technique

Since its proposal in 1964, the Higgs boson has been notoriously difficult to detect. Multivariate techniques are powerful tools that typically provide better sensitivity than simple cut based methods. One example is the matrix element (ME) method [34] which has been used successfully in the measurement of single top quark production at CDF [35] and D0 [36].

5.1 The Matrix Element Discriminant

In the ME method [37], each event is assigned a numerical value which indicates how signal-like or background-like it is. This value is referred to as the event probability discriminant (EPD):

\[
EPD(m_H) = \frac{P_H(m_H)}{P_H(m_H) + \sum_i \beta_i P_{bkg_i}}
\]  

(5.1)

where \(P_H(m_H)\) is the probability density of a candidate event being consistent with the Higgs boson hypothesis of mass \(m_H\) and \(P_{bkg_i}\) corresponds to the probability density of the event being consistent with the \(i\)th background hypothesis. In order to account for the relative contributions of the backgrounds, their probabilities are weighted with \(\beta_i\) which depends on the acceptance and cross section for that process. Since \(\beta_i\) is small for small backgrounds, it is a reasonable approximation to include only large backgrounds. In the 0-jet analysis, the expression for EPD can be simplified.
to:

\[ EPD(m_H) = \frac{P_H(m_H)}{P_H(m_H) + P_{WW}} \]  \hspace{1cm} (5.2)

Each probability density in equations 5.1 and 5.2 is calculated using the matrix element for that particular process. At truth level these can be calculated exactly, but at reconstructed level we do not have access to the neutrino four-momenta and hence integrate over these unknowns.

Since the matrix element takes the hypothesised Higgs boson mass as input, the EPD must be reevaluated for each mass assumption. As a result, the ME method is fairly CPU intensive for the WW channel. However, unlike most other multivariate techniques that employ machine learning algorithms on simulated data, the ME method makes use of our theoretical knowledge for each physics process.

### 5.2 Calculation of Event Probability Densities

For a process \( \alpha \), the event probability is defined as

\[ P_\alpha = \frac{1}{\sigma} \sum_{ij} \int \frac{d\sigma_{ij}}{dy} T(z, y) dy \]  \hspace{1cm} (5.3)

where:

- \( \sigma \) is the cross section for that process
- \( i \) and \( j \) represent all possible combinations of initial state partons
- \( \frac{d\sigma_{ij}}{dy} \) is the differential cross section for a given combination of incoming partons
- \( z \) is the four-vector of the reconstructed final state
- \( y \) is the four-vector of the truth final state
- \( T(z, y) \) is the transfer function which relates measured quantities (\( z \)) to truth level quantities (\( y \))
Recall that the LHC collides proton constituents. As a result, the colliding partons do not carry the full beam energy but rather a fraction of it and so we include PDFs:

\[ P_\alpha = \frac{1}{\sigma} \sum_{i,j} \int f_i(x_i, Q^2) f_j(x_j, Q^2) \frac{d\sigma_{ij}}{dy} T(z, y) dx_i dx_j dy \]  

(5.4)

where:

- \( x_i \) is the momentum fraction for the \( i \)th parton
- \( f_i(x_i, Q^2) \) is the PDF for parton type \( i \)

Based on Fermi’s golden rule, the differential cross section for two incoming particles decaying to \( n \) final state particles is [3]:

\[ d\sigma_{ij} = \frac{(2\pi)^4 |M_{ij}|^2}{4 \sqrt{(q_1 \cdot q_2)^2 - m_{q_1}^2 m_{q_2}^2}} d\Phi_n(q_1 + q_2; p_1, p_2, ..., p_n) \]  

(5.5)

where:

- \( M_{ij} \) is the Lorentz invariant matrix element
- \( q_i \) and \( m_{q_i} \) are the four-momenta and masses of the incoming particles
- \( p_i \) are the four-momenta of the outgoing particles
- \( d\Phi_n \) is the phase-space for \( n \) outgoing particles and is given by [3]:

\[ d\Phi_n(q_1 + q_2; p_1, p_2, ..., p_n) = \delta^4 \left( q_1 + q_2 - \sum_{i=1}^{n} p_i \right) \prod_{i=1}^{n} \frac{d^3p_i}{(2\pi)^3 2E_i} \]  

(5.6)

where:

- \( E_i \) are the energies of the outgoing particles
- \( \delta^4 \) is a four dimensional delta function that enforces energy-momentum conservation
Since the incoming partons must have large momenta in order to produce an interesting event, we assume their masses are negligible. In addition, we can write $q_1$ and $q_2$ in terms of the momentum fraction:

$$q_1 = (0, 0, x_1 E_{beam}, x_1 E_{beam})$$  \hspace{1cm} (5.7)$$

$$q_2 = (0, 0, -x_2 E_{beam}, x_2 E_{beam})$$  \hspace{1cm} (5.8)$$

where $E_{beam}$ is the beam energy. Then the denominator of equation 5.5 can be rewritten as:

$$4 \sqrt{(q_1 \cdot q_2)^2 - m_{q_1}^2 m_{q_2}^2} = 8 x_1 x_2 E_{beam}^2$$  \hspace{1cm} (5.9)$$

We assume that the leptons are well measured and have delta function resolution. In addition, conservation of energy and momentum allows us to eliminate two neutrino unknowns. Further assuming that the neutrinos are massless reduces the number of neutrino unknowns to four. At leading order, equation 5.6 simplifies to:

$$d\Phi_4 = \frac{1}{(2(2\pi)^3)^4} \frac{d\rho_x^\nu_1 d\rho_y^\nu_1 d\rho_x^\nu_2 d\rho_y^\nu_2}{E_{\ell_1} E_{\ell_2} E_{\nu_1} E_{\nu_2}}  \hspace{1cm} (5.10)$$

Thus, in the absence of any jets, equation 5.4 becomes:

$$P_\alpha = \frac{1}{\sigma} \sum_{i,j} \int \frac{|M_{ij}|^2 f_i(x_i, Q^2) f_j(x_j, Q^2) d\rho_x^\nu_1 d\rho_y^\nu_1 d\rho_x^\nu_2 d\rho_y^\nu_2}{8 \cdot 16(2\pi)^8 x_i x_j E_{beam}^2 E_{\ell_1} E_{\ell_2} E_{\nu_1} E_{\nu_2}}  \hspace{1cm} (5.11)$$

Equation 5.11 returns the probability that a single event is consistent with process $\alpha$. Of course, background events may have signal-like characteristics and as a result score a high probability of being a Higgs boson event. However, the EPD distribution should peak close to 0 for background events, and close to 1 for signal events as illustrated in figure 5.1.

### 5.3 Variable Transformation

The presence of a four-dimensional integral in equation 5.11 makes the matrix element calculation CPU intensive. Rather than performing a brute force integration, time can be saved by transforming the integration variables. Under the Higgs boson
hypothesis, the integration variables used are $m_H$, $m_{W^+}$, $m_{W^-}$, and $\eta_H$, where $\eta_H$ is the pseudorapidity of the Higgs boson. When calculating SM $WW$ MEs, $m_{W^+}$, $m_{W^-}$, $x_1$, and $x_2$ are used as integration variables. The details of the constraint equations which make this transformation possible can be found in reference [38].

5.4 Transfer Functions

The transfer function is designed to model the detector resolution (and acceptance) for a specific observed quantity. For a given truth value $y$, the transfer function $T(z, y)$ returns the probability of measuring reconstructed value $z$. Jet angles and lepton kinematics are well measured by the ATLAS detector and so we directly input these values into the probability calculation. This is equivalent to setting $T(z, y)$ to a delta function centered at the measured value. In contrast, the resolution in jet energy is not negligible and a suitable transfer function must be derived. Neutrino momenta are integrated over using a flat transfer function.

Jet energies at ATLAS are already scaled by the jet energy scale (JES), but it is
still necessary to apply further corrections. JES corrections demand that the means of the jet energies and parton level energies are the same, but this does not account for shape discrepancies in the difference between parton and observed energies. Also, out-of-cone corrections and missing energy from neutrinos are ignored.

A double Gaussian function is used to describe the jet response in MC simulation:

\[
T_{\text{jet}}(E_{\text{parton}}, E_{\text{jet}}) = \frac{1}{\sqrt{2\pi}(p_2 + p_3 p_5)} \left( e^{-\frac{(\delta E - p_1)^2}{2p_2^2}} + p_3 e^{-\frac{(\delta E - p_4)^2}{2p_5^2}} \right)
\]

(5.12)

where:

- \(\delta E = E_{\text{parton}} - E_{\text{jet}}\)
- \(p_i = a_i + b_i E_{\text{parton}}\), with \(a_i\) and \(b_i\) representing 10 parameters that define \(T_{\text{jet}}(E_{\text{parton}}, E_{\text{jet}})\)

Figure 5.2: 2D plot of \(E_{\text{parton}}\) vs \(E_{\text{jet}}\) from MC (left) and as predicted by the transfer function convoluted with \(E_{\text{parton}}\) (right) [37].

For \(b\)-jets, the 10 parameters are derived from a MC \(t\bar{t}\) sample, while for gluon jets we use a QCD dijet sample. Partons are only matched to jets if \(\Delta R\), the angle between the parton and jet, is smaller than 0.4. The matched events are fitted to equation 5.12, using a maximum likelihood fit to derive the 10 parameters. Figure 5.3 shows the distribution of \(\delta E\) in various bins of parton level energy with fits as predicted by the derived transfer function. Note the asymmetric tails originating from high-energy partons with underestimated reconstructed energy. This implies that large amounts of jet energy have escaped detection.
Figure 5.3: The transfer function is validated by comparing $\delta E$ for $b$-jets matched to partons (black) to the corresponding predictions of the transfer function (red). The agreement is good over a wide range of $E_{\text{parton}}$ [37].
5.5 EPD Control Regions

Before the ME technique is applied to the Higgs boson search, it must first be validated in data control regions where very little signal is expected. Data and MC are compared to ensure that the EPD prediction from MC properly models the data.

The $WW$ control region shown in figure 5.4 is prepared by applying the cuts outlined in section 4.1.2 and requiring $m_{\ell\ell} > 80$ GeV. The $t\bar{t}$ control region is selected using 1-jet events$^1$ and requiring the jet to be $b$-tagged (figure 5.5). $b$-tagged jets are identified as being consistent with jets that originate from $b$ and $c$ quarks using the event topology of decay products originating from $b$- and $c$-flavored hadrons [39]. The $Z/\gamma^*$ control region is produced by applying the selection cuts from section 4.1.2 but with $p_T^{\ell\ell} < 30$ GeV.

![EPD distribution in the $WW$ control region. The diagonal grey lines represent the total uncertainty on the predicted background$^3$ [37].](image)

Figure 5.4: EPD distribution in the $WW$ control region. The diagonal grey lines represent the total uncertainty on the predicted background$^3$ [37].

Overall, good agreement between data and MC is observed in the control regions indicating the method is robust.

---

$^1$Refer to reference [20] for details on 1-jet selection criteria.

$^3$Plots shown are for 7 TeV as 8 TeV ME results have not been published yet.
CHAPTER 5. THE MATRIX ELEMENT TECHNIQUE

Figure 5.5: EPD distribution in the $t\bar{t}$ control region. The diagonal grey lines represent the total uncertainty on the predicted background [37].

Figure 5.6: EPD distribution in the $Z/\gamma^*$ control region. The diagonal grey lines represent the total uncertainty on the predicted background [37].
5.6 Results

Figure 5.7 shows the 0-jet EPD distribution for 7 TeV MC and data in the signal region. All selection criteria in chapter 4 are applied to the ME analysis except the $\Delta \phi$ cut. With $4.7 \text{ fb}^{-1}$ there is no significant evidence for a SM Higgs boson and so limits are set. Using only 2011 data, the SM Higgs boson is excluded for $128 \text{ GeV} < m_H < 263 \text{ GeV}$ while the expected exclusion is $129 \text{ GeV} < m_H < 253 \text{ GeV}$.

Figure 5.7: EPD distribution in the signal region for the 0-jet analysis. The superimposed red line indicates the expected signal for a SM Higgs boson with $m_H = 125 \text{ GeV}$. The diagonal grey lines represent the total uncertainty on the predicted background [37].
Chapter 6

Boosted Regression Trees

Boosted decision trees (BDTs) have seen wide use in high-energy physics. One of the methods used by both CDF [35] and D0 [36] for the single top quark discovery is a BDT. A sequence of cuts is used to divide a sample of events into two categories typically labelled as signal and background. A boosted regression tree generalises this concept to divide a sample into an arbitrary number of categories. This makes it possible to construct an estimator of a quantity that is not directly measurable but is correlated to a set of measured quantities.

6.1 Regression Trees

The Toolkit for Multivariate Data Analysis with ROOT (TMVA) [40] is especially designed for applications in high-energy physics. Several multivariate methods are available in TMVA such as neural nets. We chose to use the boosted regression tree because of its excellent "out of the box" performance along with its short training time. In addition, regression trees are relatively easy to visualise as simple 2-dimensional trees.

A regression tree is an estimator with a tree-like structure as shown in figure 6.1. For each set of input variables \((X_1, X_2, \ldots)\), the estimator returns one value for the output variable \((Y)\) which is derived from truth information in MC samples. Figure 6.2 illustrates a possible distribution for \(Y\) based on two input variables. At each node in the tree, a binary decision is made using a cut on one input variable. The
same input variable may show up at multiple nodes, while others may not show up at all depending on how correlated they are with the output variable. Each terminal node corresponds to one possible output value of the estimator.

6.2 Training Regression Trees

The process of defining the cuts at each node is referred to as the training of the regression tree. For each node, the variance is defined as:

\[ \sigma^2 = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \]

(6.1)

where:

- \( N \) is the number of events in that node
• $y$ is the truth value of the variable being estimated defined for each event

• $\hat{y}$ is the mean over all values of $y$ in that node

For a given cut on a variable, we define:

$$\Delta \sigma^2 = \frac{\sigma^2_{\text{parent}} - (\sigma^2_{\text{left}} + \sigma^2_{\text{right}})}{\sigma^2_{\text{parent}}} \quad (6.2)$$

At the first iteration (root node), the complete training sample is split into two subsets such that $\Delta \sigma^2$ is maximised. The optimal variable and corresponding cut is determined by trial-and-error. For each variable, a preset number of cuts are tested at equally spaced intervals in the variable’s range and the combination with the largest value for $\Delta \sigma^2$ is chosen. Of course the number of cuts can be increased at the cost of computing time, but no significant increase in performance is observed by setting it above 30. Also note that this procedure automatically ignores variables which are not highly correlated to the quantity being estimated. The node splitting procedure is repeated iteratively until a stop criterion is met. For instance, a limit can be set on the depth of the tree or the number of nodes. The output variable $Y$ for each node is equal to $\hat{y}$ as calculated for that node.
Simple regression trees tend to be sensitive to statistical fluctuations. For instance, for two input variables with similar separation power, a statistical fluctuation in the training sample may cause the training algorithm to choose a different variable at a specific node. In this scenario, the entire tree below this node is changed, potentially resulting in a different regressor response. This effect may be remedied by means of boosting.

6.3 Boosting Regression Trees

Boosting extends the idea of a single tree to multiple trees commonly referred to as a forest. Subsequent trees are generated by applying the same training procedure, but with reweighted (boosted) training events. Events that have $y$ far removed from $\hat{y}$ are assigned larger weights than those that have $y$ close to $\hat{y}$. After training a given number of trees, the estimator is taken to be the weighted average of all trees in the forest. This technique significantly improves the performance of regression trees and makes them less sensitive to statistical fluctuations. Unfortunately, the forest of weighted trees does not have the intuitive interpretation of a single regression tree.

Several boosting methods are available in TMVA, but gradient boosting provides the best performance for our purposes. The regressor is expressed as the weighted sum of many trees:

$$Y = \sum_{m=0}^{M} \beta_m f(\vec{x}, \vec{w}_m)$$  \hspace{1cm} (6.3)

where:

- $f(\vec{x}, \vec{w}_m)$ represents one regression tree
- $\vec{x}$ is the set of input variables
- $\vec{w}_m$ is the set of event weights for the $m$th regression tree
- $\beta_m$ is the weight of the $m$th regression tree
- $M$ is the number of trees in the forest
The boosting procedure optimises parameters $\beta_m$ and $\vec{w}_m$ such that the deviation between the regressor $Y$ and the truth value $y$ is minimised for the training sample. The deviation is measured with a loss function, $L(Y, y)$. Once the loss function is defined, the algorithm for calculating $\beta_m$ and $\vec{w}_m$ is fully determined. For some loss functions, such as the exponential loss function, it is possible to derive an analytic method for determining the parameters.

The loss function we use is called the Huber loss function which has the advantage that it reduces the impact of outliers on the regression estimate:

$$L(Y, y) = \begin{cases} \frac{1}{2}(y - Y)^2, & |y - Y| \leq \delta \\ \delta(|y - Y| - \delta/2), & |y - Y| > \delta \end{cases}$$

(6.4)

As the name implies, the minimisation of the loss function is performed by calculating its gradient. Repeating this iteratively yields the forest of regression trees that minimises the loss function.

In order to reduce the effect of statistical fluctuations, a different, randomly chosen subset of the training sample is used for each tree. The fraction of events used in each training is referred to as the bagging fraction. In relation, the shrinkage is defined as the learning rate for the boosting algorithm.
Chapter 7

Higgs Boson Transverse Rest Frame

As mentioned in chapter 3, the Higgs boson typically carries transverse momentum due to next-to-leading order effects. In order to account for this we attempt to boost the Higgs boson into its transverse rest frame using a Lorentz transformation. Since the neutrinos pass through ATLAS undetected, we do not have direct access to the Higgs boson four-momenta. Using a boosted regression tree, we estimate the transverse momentum of the Higgs boson based on dilepton kinematics and missing transverse energy.

7.1 Implementation

We use TMVA to train a regression tree that estimates Higgs boson momentum along the $x$-axis ($p^H_x$) and along the $y$-axis ($p^H_y$). Thanks to the symmetry of the ATLAS detector in the transverse plane, the same estimator can be used for both. This allows us to double our sample size since each event includes two data points. This is advantageous since we are required to divide our sample into two statistically independent subsets, one for training and one for testing the regression tree.

The training and testing samples are generated by combining gluon-fusion Higgs boson and SM $WW$ MC events. Signal events are used at two mass hypotheses in order to increase statistics. Table 7.1 summarises the samples used.

The estimator is only used under the Higgs boson hypothesis. SM $WW$ events are primarily produced by quark fusion and consequently undergo significantly less


Table 7.1: MC samples used for training and testing the Higgs boson $p_T$ estimator.

<table>
<thead>
<tr>
<th>Process</th>
<th>Generator</th>
<th>Mass (GeV)</th>
<th>Number of Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>$gg \rightarrow H \rightarrow WW$</td>
<td>PowHeg [22] + Pythia8 [23]</td>
<td>125</td>
<td>43838</td>
</tr>
<tr>
<td>$gg \rightarrow H \rightarrow WW$</td>
<td>PowHeg + Pythia8</td>
<td>130</td>
<td>70822</td>
</tr>
<tr>
<td>$qq \rightarrow WW$</td>
<td>MC@NLO [25]</td>
<td></td>
<td>141474</td>
</tr>
<tr>
<td>$gg \rightarrow WW$</td>
<td>gg2WW [26]</td>
<td></td>
<td>21396</td>
</tr>
</tbody>
</table>

initial state radiation than Higgs boson events. As a result, SM $WW$ events have less momentum in the transverse plane and so are less affected by the boost.

The discriminating variables used by the estimator can be classified into three categories:

1. Dilepton $p_T^H \sim$ Electrons and muons are well-measured by the ATLAS detector and so their kinematics provide valuable information to the training.

2. Dineutrino $E_T^{miss}$ \sim Since neutrinos cannot be measured directly, we infer the dineutrino momentum by training on both the calorimeter and track-based $E_T^{miss}$. Calorimeter $E_T^{miss}$ has better resolution at high Higgs boson $p_T$, while track $E_T^{miss}$ performs better at low $p_T$. Due to a number of factors such as dead regions in the detector, limited coverage of the detector, and various sources of noise, the resolution of $E_T^{miss}$ is significantly worse than that of leptons. Consequently, the $E_T^{miss}$ is the limiting factor in our estimator’s resolution.

3. Pileup independence $\sim$ By training on the number of primary vertices and the number of interactions per bunch crossing it is possible to reduce the impact of pileup on the regression estimator.

Gradient boosting is used to grow a forest of 300 trees, each of which is allowed a maximum of 15 nodes and a maximum depth of 3. In order to determine which variable and corresponding cut provides the largest $\Delta \sigma^2$, 30 cuts are tested for each variable. The gradient bagging fraction is set to 50%, and the shrinkage is set to 0.1.
7.1.1 Application to the ME Calculation

The four dimensional delta function in the phase-space term (equation 5.6) enforces energy and momentum conservation. In the transverse plane, it demands:

\[ p_{H}^{x,y} = 0 = p_{\ell_1}^{x,y} + p_{\ell_2}^{x,y} + p_{\nu_1}^{x,y} + p_{\nu_2}^{x,y} \]  

(7.1)

Notice that this assumes that the Higgs boson is produced at rest. In order to apply our boost estimator, we replace this constraint by:

\[ p_{H}^{x,y} = p_{\text{estimate}}^{x,y} = p_{\ell_1}^{x,y} + p_{\ell_2}^{x,y} + p_{\nu_1}^{x,y} + p_{\nu_2}^{x,y} \]  

(7.2)

7.2 Validation

Since statistical fluctuations in the training sample affect the growth of a tree, a statistically independent sample is used for testing. MC events with even event numbers are used for training, while odd numbered events are used for testing. Note that all performance plots that follow were made after applying the selection cuts described in section 4.1.2 before the analysis is divided into same and opposite flavor channels.

7.2.1 Estimator Performance in $p_T^H$

In order to consider the resolution in magnitude and direction of the regression estimator separately, we consider the performance in $p_T$ and $\phi$ as defined in equations 2.7 and 2.2 respectively. Figures 7.1 and 7.2 compare the truth and estimated $p_T^H$ for both Higgs boson events and $WW$ background events.

On average the regressor underestimates the truth $p_T^H$. This effect can be traced back to the training method as described by equation 6.1. Recall that the estimator is trained on truth $p_x$ which is a Gaussian distribution centered at zero. The training algorithm specifies cuts on the input variables that divide the truth $p_x$ distribution into a number of end-nodes. Each end-node has a corresponding $p_x$ estimate which is calculated to be the mean of all events in that node. Consider an end-node with an estimate close to zero. We expect the $p_x$ distribution inside that end-node to be symmetric since we sample events from a distribution that is symmetric around zero.
Figure 7.1: Estimated vs truth $p_T$ for 125 GeV Higgs boson (left) and SM WW production (right).

Figure 7.2: Truth (blue) and estimated (red) $p_T$ distribution for 125 GeV Higgs boson (left) and SM WW production (right).
(green shaded area in figure 7.3). In contrast, the truth $p_x$ distribution for an end-node with a large positive estimate will not be symmetric as it is dominated by events taken from the tail of a Gaussian (red shaded area in figure 7.3). The mean of this distribution will typically be shifted to the left causing the regressor to underestimate $p_x$.

![Figure 7.3: The non-uniformity of the training sample introduces a bias in the estimator.](image)

In order to compensate for the bias, it is possible to assign larger event weights to high-$p_x$ events such that the $p_x$ distribution is flat. Events are reweighted such that each 2 GeV bin in truth $p_x$ contains the same number of effective events. This strategy improves the performance of the Higgs boson $p_x$ estimate at high $p_x$ at the cost of lowering the performance at low $p_x$. The impact of this reweighting is summarised in section 7.2.5.

Additionally, the Gaussian training distribution causes a larger part of the regression tree to focus on events with low $|p_x|$ since the majority of events have $|p_x|$ close to zero. However, on average, these result in small boosts which do not have a large impact on the Higgs boson event kinematics. The events that require large boosts (and affect the kinematic distributions the most) are in the tail of the $p_x$ distribution. Unfortunately, the higher the $|p_x|$, the lower the statistics. One could imagine an optimal training sample where the density of events increases with increasing $|p_x|$.

The resolution of the estimator is very similar for Higgs boson and $WW$ processes as demonstrated in figure 7.4. The asymmetry in these distributions is caused by the
previously mentioned high-$p_T$ events. Note that the estimator is trained on signal and background, but it is also possible to only train on Higgs boson events. As expected, this improves the resolution of the estimator for signal events but worsens the resolution for background events. When only training on signal events, the width of the resolution of background events for the testing sample widens by approximately 4 GeV.

![Resolution of $p_T$ estimate for 125 GeV Higgs boson (left) and SM $WW$ production (right).](image)

Figure 7.4: Resolution of $p_T$ estimate for 125 GeV Higgs boson (left) and SM $WW$ production (right).

Figure 7.5 confirms that the tail in the $p_T$ resolution originates from high-$p_T$ events. Each colored distribution corresponds to a specific range of truth $p_T$. As truth $p_T$ increases, the more likely it is to underestimate the $p_T$.

### 7.2.2 Estimator Performance in $\phi^H$

Of course, the magnitude of the boost is useless without a direction. Figure 7.6 illustrates the resolution of the $\phi$ coordinate which takes the shape of a Gaussian. The tails of the distribution are primarily made up of low-$p_T$ events since it is much more difficult to separate these events from background noise. Fortunately, we only apply small boosts to these events, so even if the estimated direction is antiparallel, the effect on the kinematics is not large. In contrast, the resolution improves for higher $p_T$ events as shown in figure 7.7.
CHAPTER 7. HIGGS BOSON TRANSVERSE REST FRAME

Figure 7.5: Resolution of $p_T$ estimate for 125 GeV Higgs boson (left) and SM $WW$ production (right) binned in truth $p_T$.

Figure 7.6: Resolution of $\phi$ estimate for 125 GeV Higgs boson (left) and SM $WW$ production (right).
CHAPTER 7. HIGGS BOSON TRANSVERSE REST FRAME

7.2.3 Pileup Immunity

Recall that pileup can contaminate the hard scatter with additional soft interactions. By including the average number of interactions per bunch crossing and the number of primary vertices in the training, it is possible to reduce the Higgs boson \( p_T \) estimator’s dependence on pileup. Figure 7.8 shows the resolution for three very different ranges of \( \mu \), the average number of interactions per bunch crossing. The resolution under different pileup conditions is very similar.

Figure 7.7: Resolution of \( \phi \) estimate for 125 GeV Higgs boson (left) and SM WW production (right) binned in truth \( p_T \).

Figure 7.8: The resolution of the Higgs boson \( p_T \) estimate for 125 GeV Higgs boson (left) and SM WW production (right) does not change significantly as pileup conditions vary.
7.2.4 Regressor Overtraining

Given a large enough regression tree, one could imagine the limiting case of as little as one event in each end-node. Even though this tree could potentially provide perfect separation for the training sample, the regression tree will be very dependent on statistical fluctuations, an effect known as overtraining. An overtrained regressor performs significantly better on the training sample than on the testing sample. When determining the sensitivity of the analysis, both the training and testing samples are used and so an overtrained regressor will yield an optimistic result. As a consequence, we need to verify that the resolution distributions for the training and testing samples are similar.

The Kolmogorov-Smirnov test is a non-parametric test that provides a measure of how similar two distributions are. The null hypothesis is that both samples are from the same parent distribution. Full details of the test can be found in reference [41]. Since values above 0.20 are typically deemed acceptable, the measured test statistic of 0.793 for signal events and 0.685 for background events is satisfactory.

As an additional check, we can verify that the performance of the regressor is not significantly better for the training sample than for the testing sample. Figure 7.9 compares the $p_T$ distribution for the regressor when applied to the training and testing samples.

Figure 7.9: For both signal (left) and background (right), the estimated Higgs boson $p_T$ distributions are similar for the testing and training samples. There are no obvious signs of overtraining.
7.2.5 Impact of Boost Estimator on EPD Distributions

The original goal of the boost estimator was to improve the separation between background and signal in the ME analysis. This corresponds to shifting the EPD distribution for signal events closer to 1, and for background closer to 0.

Figure 7.10 depicts the EPD distribution before and after incorporating the Higgs boson transverse momentum. Before applying a Lorentz boost, the EPD has a distinct peak at 0 for signal and background events. This corresponds to events that are kinematically not compatible with the Higgs boson hypothesis since we assume the Higgs boson is produced at rest. However, after applying the Higgs boson $p_T$ estimator, a significant fraction of these events are kinematically compatible with the Higgs boson hypothesis. For instance, note the increase in Higgs boson events with EPD values close to 1. At the same time, Lorentz boosting also increases the number of background events compatible with the Higgs boson hypothesis. Nevertheless, it is easier to distinguish signal from background after incorporating the Higgs boson $p_T$ estimator.

![Figure 7.10: EPD distribution before (left) and after (right) applying a Lorentz boost. Signal is indicated by the black dotted line. Applying the Lorentz boost allows us to recover signal events with EPD scores close to zero.](image)

One example of a kinematic constraint that would result in EPD scores close to zero is that the dilepton transverse momentum is bounded above by half the Higgs
boson mass. Applying conservation of energy to a Higgs boson at produced at rest:

\[ m_H = E_H = (E_{\ell\ell} + E_{\nu\nu}) \quad (7.3) \]

Since momentum is conserved:

\[ \vec{p}_{\ell\ell} = -\vec{p}_{\nu\nu} \quad (7.4) \]

Then equation 7.3 can be rewritten as:

\[ m_H = \sqrt{(p_{\ell\ell}^2 + m_{\ell\ell}^2)} + \sqrt{(p_{\nu\nu}^2 + m_{\nu\nu}^2)} \quad (7.5) \]

\[ m_H \geq 2\sqrt{p_{\ell\ell}^2} \quad (7.6) \]

In the transverse plane:

\[ \frac{m_H}{2} \geq p_{T\ell\ell} \quad (7.7) \]

Any event not satisfying equation 7.7 will be classified as a background event. If the Higgs boson has transverse momentum, the measured \( p_{T\ell\ell} \) may be significantly larger than \( \frac{m_H}{2} \) and that event would also resemble a background event.

A simple and intuitive metric to determine the sensitivity of a search is based on \( \frac{S}{\sqrt{B}} \), where \( S \) represents the number of signal events, and \( B \) the number of background events. It involves determining the cut on EPD such that \( \frac{S}{\sqrt{B}} \) to the right of the cut is maximised.

Table 7.2 summarises the improvements for the various training methods. The best improvement is observed when the Higgs boson boost estimator is trained on signal and background events without applying any event weights to flatten the training distribution. In this scenario, the expected sensitivity improves by \( \sim 15\% \).
Table 7.2: Improvement in $\frac{s}{\sqrt{B}}$ for four different training scenarios. The training sample may be weighted such that the $p_x$ distribution is flat but this seems to have an adverse effect on the performance.
Chapter 8

Measurement of the Higgs Boson Spin

To date, the production rate of the newly observed particle is consistent with the SM Higgs boson. However, there are many other properties to be studied. For example, since the new particle is observed decaying to vector bosons with a net electric charge of zero, we conclude it must be neutral.

Another property that merits study is the spin. According to the Landau-Yang theorem [42], a spin-1 boson cannot decay to a pair of photons. Thus, the two spin hypotheses that are currently being explored are spin-0 and spin-2. As discussed in chapter 3, the spin of a particle will influence the kinematics of its decay products. The four variables that provided the best separation between a spin-0 and spin-2 hypothesis are $\Delta\phi_{\ell\ell}$, $m_T$, $m_{\ell\ell}$, and $p^\ell_T$ [34]. As with the ME calculation, these distributions may change depending on the Higgs boson’s transverse boost.

8.1 Current Implementation

The ATLAS $H \rightarrow WW$ spin analysis currently uses a simplified boost estimator. The transverse components of the dineutrino system are taken to be equal to the calorimeter $E_{T}^{\text{miss}}$. At this point, both $m_{\nu\nu}$ and $p^\nu_{z}$ are unconstrained. Since the median mass of the dineutrino system is 30 GeV in MC simulation, we take $m_{\nu\nu} = 30$ GeV. Then further assuming the Higgs boson mass is 125 GeV, $p^\nu_{z}$ is constrained by:

$$m^2_H = (E_{\ell\ell} + E_{\nu\nu})^2 - (\vec{p}_{\ell\ell} + \vec{p}_{\nu\nu})^2$$

(8.1)
\[ m_H^2 = m_{\ell\ell}^2 + m_{\nu\nu}^2 + 2E_{\ell\ell}E_{\nu\nu} - 2\vec{p}_{\ell\ell} \cdot \vec{E}_{T}^\text{miss} \] (8.2)

\[ m_H^2 = m_{\ell\ell}^2 + m_{\nu\nu}^2 + 2E_{\ell\ell}\sqrt{(E_T^\text{miss})^2 + (p_z^{\nu\nu})^2 + m_{\nu\nu}^2} - 2p_{\ell\ell}^T \cdot \vec{E}_T^\text{miss} - 2p_{\ell\ell}^T p_z^{\nu\nu} \] (8.3)

Of the two resulting solutions for \( p_z^{\nu\nu} \), the one that minimises the angle between the dilepton system and beam direction in the rest frame is chosen. The angle is defined as:

\[ \cos(\theta_{\ell\ell}) = \frac{\vec{p}_{\ell\ell} \cdot \hat{z}}{p_{\ell\ell}} \] (8.4)

This analytic boost estimator provides reasonable performance, but it can be improved using the regression estimator.

### 8.2 Performance Comparison

As mentioned in chapter 7, Higgs boson events with high \( p_x \) are underestimated due to the asymmetric training distribution. Applying event weights such that the \( p_x \) distribution is flat improves the Higgs boson estimate and so is applied in this study. Figure 8.1 compares the Higgs boson \( p_T \) estimator resolution binned in truth Higgs boson \( p_T \) before and after flattening the training distribution. The resolution for high-\( p_T \) events improves at the cost of resolution for low-\( p_T \) events. However, since larger boosts have a larger impact on the kinematic distributions, the trade-off is worthwhile. The estimator is trained using only the PowHeg Higgs boson samples listed in table 7.1.

Figure 8.2 compares the Higgs boson \( p_T \) distribution for spin-0 and spin-2 MC samples with a Higgs boson mass of 125 GeV. Note that the regression estimator follows the truth distribution closer than the analytic method. Figure 8.3 illustrates the resolution of the estimated Higgs boson \( p_T \). Notice that the regression estimate has a significantly smaller RMS and a mean closer to 0 compared to the analytic method.
CHAPTER 8. MEASUREMENT OF THE HIGGS BOSON SPIN

Figure 8.1: Resolution of Higgs boson $p_T$ estimator binned in truth $p_T$ before (left) and after (right) flattening the $p_x$ distribution.

Figure 8.2: Comparison of the truth Higgs boson $p_T$ (blue) with the analytic estimator (green) and the regression estimator (red) for spin-0 (left) and spin-2 (right) Higgs boson hypothesis.
CHAPTER 8. MEASUREMENT OF THE HIGGS BOSON SPIN

82

Figure 8.3: Comparison of Higgs boson $p_T$ resolution for the analytic estimator (green) and the regression estimator (red) for spin-0 (left) and spin-2 (right) Higgs boson hypothesis.

In order to further evaluate the validity of the estimate, we can compare kinematic quantities after boosting into the estimated rest frame of the Higgs boson using a Lorentz transformation. The dilepton $p_T$ is depicted in figure 8.4 along with the resolution in figure 8.5. Notice that the high-$p_T$ tail in the analytic method is not present in the regression estimate. This is the main source of the improvement in the resolution of the regression estimate.

Figure 8.4: Comparison of the truth dilepton $p_T$ (blue) with analytic estimator (green) and the regression estimator (red) for spin-0 (left) and spin-2 (right) Higgs boson hypothesis.
Both the analytic and regression based estimates perform reasonably well when estimating $\Delta \phi$ in the Higgs boson rest frame as shown in figure 8.6. However, as apparent from figure 8.7, the resolution for the regression estimator is better than that of the analytic estimate.

The $\Delta \psi$ distribution as depicted in figures 8.8 and 8.9, is significantly more challenging than any of the previously mentioned transverse variables as it requires knowledge of the $z$-component of the Higgs boson’s boost. Note in particular the poor
Chapter 8. Measurement of the Higgs Boson Spin

Figure 8.7: Comparison of the dilepton opening angle for the analytic estimator (green) and the regression estimator (red) for spin-0 (left) and spin-2 (right) Higgs boson hypothesis. The performance of both boost estimators for the spin-2 Higgs boson. This is because the spin-2 Higgs boson is primarily produced via $q\bar{q}$ annihilation which typically have boosts in the $z$-direction due to the asymmetric PDFs.

Although it is possible to obtain good results for the transverse momentum of the Higgs boson, we do not have sufficient information to obtain a reliable estimate for the $z$-component. This is problematic as we need $p_z$ to determine the boost vector, the vector that specifies the magnitude of the Lorentz boost along each Cartesian
CHAPTER 8. MEASUREMENT OF THE HIGGS BOSON SPIN

Figure 8.9: Comparison of the dilepton opening angle for the analytic estimator (green) and the regression estimator (red) for spin-0 (left) and spin-2 (right) Higgs boson hypothesis.

coordinate:
\[
\vec{b} = -\left( \frac{p_p^H}{E^H}, \frac{p_y^H}{E^H}, 0 \right)
\]  
(8.5)

where energy depends on the momentum in each Cartesian coordinate:
\[
E^2 = m^2 + p_x^2 + p_y^2 + p_z^2
\]
(8.6)

One method to circumvent this problem is to divide the analysis into bins of estimated \( p_T \). This gives us access to events where the Higgs boson is close to its transverse rest frame without requiring any knowledge of \( p_z \). The green distributions in figures 8.10 and 8.11 are generated by only using events (unboosted) that have estimated \( p_T < 10 \) GeV.

Figures 8.12 through 8.14 compare the \( \Delta \phi \) distributions for spin-0 and spin-2 in bins of estimated Higgs boson \( p_T \). Note the significant shift in the \( \Delta \phi \) distributions as the estimated Higgs boson \( p_T \) increases.
Figure 8.10: Comparison of the truth dilepton $p_T$ (blue) with the regression Higgs boson $p_T$ estimator (red) and unboosted events with Higgs boson $p_T < 10$ GeV (green) for spin-0 (left) and spin-2 (right) Higgs boson hypothesis.

Figure 8.11: Comparison of dilepton $p_T$ resolution for the regression estimator (red) and unboosted events with estimated Higgs boson $p_T < 10$ GeV (green) for spin-0 (left) and spin-2 (right) Higgs boson hypothesis.
Figure 8.12: $\Delta \phi$ distribution for spin-0 (red) and spin-2 (blue) Higgs boson hypothesis using only events with estimated Higgs boson $p_T < 10$ GeV.

Figure 8.13: $\Delta \phi$ distribution for spin-0 (red) and spin-2 (blue) Higgs boson hypothesis using only events with estimated Higgs boson $p_T$ between 20 and 30 GeV.
Figure 8.14: $\Delta \phi$ distribution for spin-0 (red) and spin-2 (blue) Higgs boson hypothesis using only events with estimated Higgs boson $p_T > 50$ GeV.
Chapter 9

Conclusions

Since its proposal in 1964, the Higgs boson has evaded detection for nearly half a century. Hence, to particle physicists, 2012 will be remembered as an exciting and noteworthy year. This is the year in which we caught a first glimpse of a new particle with a mass of approximately 126 GeV. To date, the discovery is consistent with the Standard Model Higgs boson. The next step is to measure precisely the properties of this new particle: production rate, branching ratios, and internal quantum numbers.

Thanks to its relatively large branching ratio, $H \rightarrow WW \rightarrow \ell\nu\ell\nu$ is a valuable search channel. However, the reconstruction of the Higgs boson is complicated by the presence of neutrinos that leave ATLAS undetected. As a result, we cannot directly reconstruct the Higgs boson mass. In addition, we have no direct handle on the transverse boost of the Higgs boson which significantly alters the measured kinematics. However, using regression techniques, it is possible to obtain estimates for unknown quantities such as the dineutrino transverse kinematics based on a set of measured variables. This allows us to perform a Lorentz boost into the transverse rest frame of the Higgs boson. The $H \rightarrow WW \rightarrow \ell\nu\ell\nu$ matrix element analysis benefits significantly from the regression boost estimator. We observe up to 15% improvement in sensitivity after applying the boost estimator. The spin measurement in the $H \rightarrow WW \rightarrow \ell\nu\ell\nu$ channel also benefits from the Higgs boson transverse boost estimator. The regression estimator provides better estimates for the Higgs boson transverse momentum than the analytic estimate currently used in the ATLAS spin analysis.
The tools we developed are useful in both the matrix element Higgs boson search and the measurement of the spin of the Higgs boson. In the cases we explored, the purpose of the regression estimator was to determine the kinematics of unmeasured neutrinos. However, regression techniques can provide reasonable estimates of any variable given other variables that are correlated to the one being estimated.
Bibliography


[37] ATLAS Collaboration, “A Search for the Higgs Boson in the $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$ Decay Mode Using a Matrix Element Method and 4.7 fb-1 of Data Collected with the ATLAS Detector at 7 TeV,” 2012.


