

# A SEARCH FOR THE HIGGS BOSON PRODUCED IN ASSOCIATION WITH TOP QUARKS IN MULTILEPTON FINAL STATES AT ATLAS

Chris Lester

A DISSERTATION  
in  
Physics and Astronomy

Presented to the Faculties of The University of Pennsylvania  
in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy  
2014

Joseph Kroll, Professor, Physics  
Supervisor of Dissertation

A.T. Charlie Johnson, Professor, Physics  
Graduate Group Chairperson

## Dissertation Committee

Randall Kamien, Professor, Physics

I. Joseph Kroll, Professor, Physics

Elliot Lipeles, Assistant Professor, Physics

Burt Ovrut, Professor, Physics

Joseph Kroll, Professor, Physics

# A SEARCH FOR THE HIGGS BOSON PRODUCED IN ASSOCIATION WITH TOP QUARKS IN MULTILEPTON FINAL STATES AT ATLAS

COPYRIGHT  
2014  
Chris Lester

All rights reserved.

---

## Acknowledgements

---

30 Acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements ac-  
31 knowledgements acknowledgements acknowledgements acknowledgements acknowledgements acknowl-  
32 edgements acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements  
33 acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements  
34 acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements  
35 acknowledgements acknowledgements.

36 Acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements ac-  
37 knowledgements acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements  
38 acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements  
39 acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements  
40 acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements  
41 acknowledgements acknowledgements.

42 Acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements ac-  
43 knowledgements acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements  
44 acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements  
45 acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements  
46 acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements  
47 acknowledgements acknowledgements.

48

## ABSTRACT

49

A SEARCH FOR THE HIGGS BOSON PRODUCED IN ASSOCIATION WITH TOP  
50 QUARKS IN MULTILEPTON FINAL STATES AT ATLAS

51

Chris Lester

52

Joseph Kroll

53

Abstract abstract abstract abstract abstract abstract abstract abstract abstract  
54 abstract abstract abstract abstract abstract abstract abstract abstract abstract  
55 abstract abstract abstract abstract abstract abstract abstract abstract abstract  
56 abstract abstract abstract abstract abstract abstract abstract abstract abstract  
57 abstract abstract abstract abstract abstract abstract abstract abstract abstract  
58 abstract abstract abstract abstract abstract abstract abstract abstract abstract  
59 abstract abstract abstract abstract abstract abstract abstract abstract abstract  
60 abstract abstract abstract abstract abstract abstract abstract abstract abstract  
61 abstract abstract abstract abstract abstract abstract abstract abstract abstract.

---

# Contents

---

63	<b>Acknowledgements</b>	iii
64	<b>Abstract</b>	iv
65	<b>Contents</b>	v
66	<b>Preface</b>	viii
67	<b>1 Introduction</b>	1
68	<b>2 Theoretical Background</b>	2
69	2.1 The Standard Model . . . . .	2
70	2.1.1 The Standard Model Structure . . . . .	2
71	2.1.2 Electroweak Symmetry Breaking and the Higgs . . . . .	3
72	2.1.3 The Standard Model Parameters . . . . .	4
73	2.2 Collider Physics and the Higgs . . . . .	4
74	2.2.1 Higgs Discovery at the LHC . . . . .	7
75	2.2.2 $t\bar{t}H$ Production . . . . .	8
76	2.3 Conclusion . . . . .	10
77	<b>3 The Large Hadron Collider and the ATLAS Experiment</b>	11
78	3.1 The Large Hadron Collider . . . . .	11
79	3.1.1 The Accelerator Complex . . . . .	11
80	3.1.2 Beam Parameters and Collisions . . . . .	12
81	<b>4 Electrons</b>	14

82	4.1	Electrons at Hadron Colliders . . . . .	14
83	4.2	Reconstruction of Electron at ATLAS . . . . .	14
84	4.3	Identification of Electrons at ATLAS . . . . .	14
85	4.3.1	Pile-up . . . . .	14
86	4.3.2	Trigger vs. Offline . . . . .	14
87	4.3.3	2011 Menu . . . . .	14
88	4.3.4	2012 Menu . . . . .	14
89	4.3.5	Electron Likelihood . . . . .	14
90	4.4	Measurement of Electron Efficiency at ATLAS . . . . .	14
91	4.4.1	Techniques . . . . .	14
92	4.4.2	Issues . . . . .	14
93	<b>5</b>	<b>Search for the TTH Decay in the Multilepton Channel</b>	<b>15</b>
94	5.1	Signal Characteristics . . . . .	15
95	5.2	Background Overview . . . . .	16
96	5.3	Analysis Strategy . . . . .	17
97	<b>6</b>	<b>Dataset and Simulation</b>	<b>18</b>
98	6.1	Data . . . . .	18
99	6.1.1	The 2012 Dataset . . . . .	18
100	6.2	Simulation . . . . .	19
101	6.2.1	Signal Simulation . . . . .	20
102	6.2.2	Background Simulation . . . . .	20
103	<b>7</b>	<b>Object and Event Selection</b>	<b>22</b>
104	7.1	2l Same-Charge Signal Region . . . . .	22
105	7.2	3l Signal Region . . . . .	23
106	7.3	4l Signal Region . . . . .	24
107	7.4	Electron Selection . . . . .	24
108	7.5	Muon Selection . . . . .	25
109	7.6	Jet and b-Tagged Jet Selection . . . . .	25
110	7.7	Tau Selection . . . . .	26
111	7.8	Object Summary and Overlap . . . . .	26
112	7.9	Optimization . . . . .	26

113	<b>8 Background Estimation</b>	<b>28</b>
114	8.1 Vector Boson ( $W^\pm$ , $Z$ ) production in association with top quarks: $t\bar{t}V$ , $tZ$ . . . . .	28
115	8.1.1 $t\bar{t}Z$ Validation Region . . . . .	29
116	8.2 Di-boson Background Estimation: $W^\pm Z$ , ZZ . . . . .	31
117	8.3 Fake Lepton Backgrounds: $t\bar{t}$ . . . . .	31
118	8.4 Charge-Misidentification Background . . . . .	31
119	<b>9 Conclusions</b>	<b>32</b>
120	9.1 Higgs Results in Review . . . . .	32
121	9.2 Prospects for Future . . . . .	32
122	<b>Bibliography</b>	<b>33</b>

---

123

## Preface

---

124 This is the preface. Blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah  
125 blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah blah blah  
126 blah. Blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah  
127 blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah blah  
128 blah blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah blah blah.  
129 Blah blah blah blah blah. Blah blah blah blah blah.

130 Blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah blah  
131 blah. Blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah blah blah  
132 blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah  
133 blah blah blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah blah  
134 blah. Blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah blah  
135 blah blah.

136

Chris Lester  
CERN, Fall 2014

<sup>137</sup>

## CHAPTER 1

<sup>138</sup>

# Introduction

<sup>139</sup> Here is a citation [?].

140

## CHAPTER 2

141

# Theoretical Background

142 The Standard Model of particle physics (SM) is an extraordinarily successful description of the funda-  
143 mental constituents of matter and their interactions. Experiments over the past 50 years have verified  
144 the extremely precise prediction of the SM. This success has culminated most recently in the discovery  
145 of the Higgs Boson. This chapter provides a brief introduction to the structure of the SM and how  
146 scientists are able to test it using hadron collider but focuses primarily on the physics of the Higgs  
147 boson and its decays to top quarks. Particular attention is given to the importance of a measurement  
148 of the rate at which Higgs Bosons are produced in association of top quarks in the context of testing  
149 the predictions the SM.

150 **2.1 The Standard Model**

151 **2.1.1 The Standard Model Structure**

152 The Standard Model (SM) [?, ?, ?, ?] is an example of a quantum field theory that describes the  
153 interactions of all of the known fundamental particles. Particles are understood to be excitations of  
154 the more fundamental object of the theory, the field. The dynamics and interactions of the fields are  
155 derived from the Standard Model Lagrangian, which is constructed to be symmetric under transfor-  
156 mations of the group  $SU(3) \times SU(2) \times U(1)$ .  $SU(3)$  is the group for the color,  $SU(2)$  is the group for  
157 weak iso spin, and  $U(1)$  is the group for weak hyper-charge.

158 Gauging the symmetries demands the introduction of 8 massless gluons, or the boson (full integer  
159 spin) carriers of the strong force [?] from the generators  $SU(3)$  symmetry, and the 4 massless bosons,  
160 carriers for the weak and electromagnetic forces from the 3 generators of the  $SU(2)$  and 1 generator  
161 of the  $U(1)$  group. The weak and the electromagnetic forces are considered part of a larger single

162 unified electroweak group  $SU(2) \times U(1)$  and the associated generators mix.

163 The gauge symmetry allows the theory to be re-normalizable [?], meaning that unwanted infinities  
 164 can be absorbed into observables from theory in a way that allows the theory to be able to predict  
 165 physics at multiple energy scales. Singlets of the  $SU(3)$  group are fermions (half-integer spin particles)  
 166 called leptons and do not interact with the strong, whereas doublets of the  $SU(3)$  group are called  
 167 quarks and do interact with the strong force. The SM is a chiral theory: the weak force violates parity,  
 168 as it only couples to left-chiral particles or right-chiral antiparticles. This means that right-chiral and  
 169 left-chiral fermions arise from different fields, which are different representations of the weak isospin  
 170 group.

171 The discovery of particles and new interactions in various experiments is intertwined with the  
 172 development of the theory that spans many decades and is not discussed in detail here.

173 So far, 3 separate generations of both quarks and leptons have been discovered, differing only by  
 174 mass. The gluon and the 4 electroweak bosons have also been discovered ( $W^+$ ,  $W^-$ ,  $Z^0$ , and  $\gamma$ ) <sup>1</sup>. The  
 175 reason for this 3-fold replication is not known.

### 176 2.1.2 Electroweak Symmetry Breaking and the Higgs

177 Despite the simple structure of theory, the discovery of massive fundamental particles creates two sets  
 178 of problems both related to  $SU(2) \times U(1)$ . First, the force-carrying bosons must enter the theory  
 179 without mass or the symmetries will be explicitly broken in the Lagrangian. Second, adding fermion  
 180 masses to theory in an ad-hoc way allows the right-chiral and left-chiral fermions to mix. Since they  
 181 possess different quantum numbers, as different representations of the weak-isospin group, this too  
 182 breaks gauge invariance.

183 To solve these problems, spontaneous electro-weak symmetry breaking (EWSB) is introduced via  
 184 the Brout-Englert-Higgs mechanism [?, ?, ?]. A massive scalar field in an electro-weak doublet is  
 185 added to the theory with 4 new degrees of freedom and a potential which includes a quartic self-  
 186 interaction term. Each fermion field interacts with the scalar field via a different Yukawa coupling,  
 187 which unites the left and right chiral fields of a single particle type. This field explicitly preserves all  
 188 of the symmetries, but the minimum of the potential does not occur when the expectation of the field  
 189 is zero. The field eventually falls to a state, where it acquires a vacuum-expectation value. However,  
 190 a non-zero field must therefore point in a particular direction of weak-isospin space, breaking the  
 191 symmetry.

---

<sup>1</sup>The actual particles observed to be produced and propagate may be linear combinations of the underlying generators (for the bosons) or gauge representations (for the fermions).

192        The consequences of this spontaneous symmetry breaking are tremendous. First, the universe  
 193    is filled with a field with a non-zero expectation value. The theory can be expanded around this  
 194    new value and 3 of the degrees of freedom can be interpreted as the longitudinal polarizations of  
 195    the  $W^+$ ,  $W^-$ , and  $Z^0$ , while the 4th remains a scalar field, called the Higgs field with an associated  
 196    particle called the Higgs particle or "Higgs" The weak bosons acquire a mass via their longitudinal  
 197    polarizations and the Yukawa couplings of the scalar field to the fermions now behave like a mass  
 198    term at the this new minimum.

199    **2.1.3 The Standard Model Parameters**

200    Although the SM structure is set by specifying the gauge groups, the EWSB mechanism, and ac-  
 201    knowledging the 3-fold replication of the fermions, confronting the SM with experiment requires the  
 202    measurement of  $17^2$  free parameters, which are unconstrained from the theory. These free parameters  
 203    include the fermion masses (from the Yukawa coupling), the force coupling constants, the angles and  
 204    phase of the mixing between quarks, and constants from the Higgs and electroweak sector<sup>3</sup>.

205    Experiments have provided a number of measurements of the parameters of the SM[?]. Prior to the  
 206    discovery of the Higgs boson (discussed later) the Higgs mass, however, was the only fully unconstrained  
 207    parameter, although its value could be inferred via its involvement in loop corrections on the top mass  
 208    ( $M_t$ ) and the W mass ( $M_W$ ). The GFitter collaboration assembles all relevant electroweak observable  
 209    measurements into a statistical model and then allows certain measurements to float within their  
 210    uncertainty to allow for a fit among multiple correlated measurements [?]. Figure 2.1  
 211    shows the fitted constraints on 4 key SM parameters ( $M_H$ ,  $M_W$ ,  $M_t$ ,  $\sin^2\theta_w$ ) with actual measurements  
 212    overlaid. The addition to the fit of the measured Higgs mass from the ATLAS and CMS collaborations  
 213    creates a small tension, as the other observables prefer the mass to be much lower ( $\sim 80$  GeV/ $c^2$ ).  
 214    The tension in the combined electroweak fit (including the Higgs) is not statistically significant with  
 215    a  $p$ -value of 0.07.

216    **2.2 Collider Physics and the Higgs**

217    To test the theory, physicists accelerate particles to extremely high energies and force them to interact  
 218    through collisions. Typically, the particles accelerated are electrons or protons, since they are stable.

---

<sup>2</sup>There are additional parameters from neutrino mass terms and mixing but it is unclear how to include these into the Standard Model, since it does not predict right-chiral neutrinos

<sup>3</sup> The electroweak sector includes parameters like mass of the  $W^\pm$  and  $Z^0$  bosons, the weak mixing angle,  $\sin^2\theta_w$ , the fermi constant  $G_F$ , and Higgs Mass and vacuum expectation value. These parameters however are not wholly independent. As discussed above, it is only necessary theoretically to specify the two parameters relevant to the Higgs potential and the two coupling associated with the gauge groups

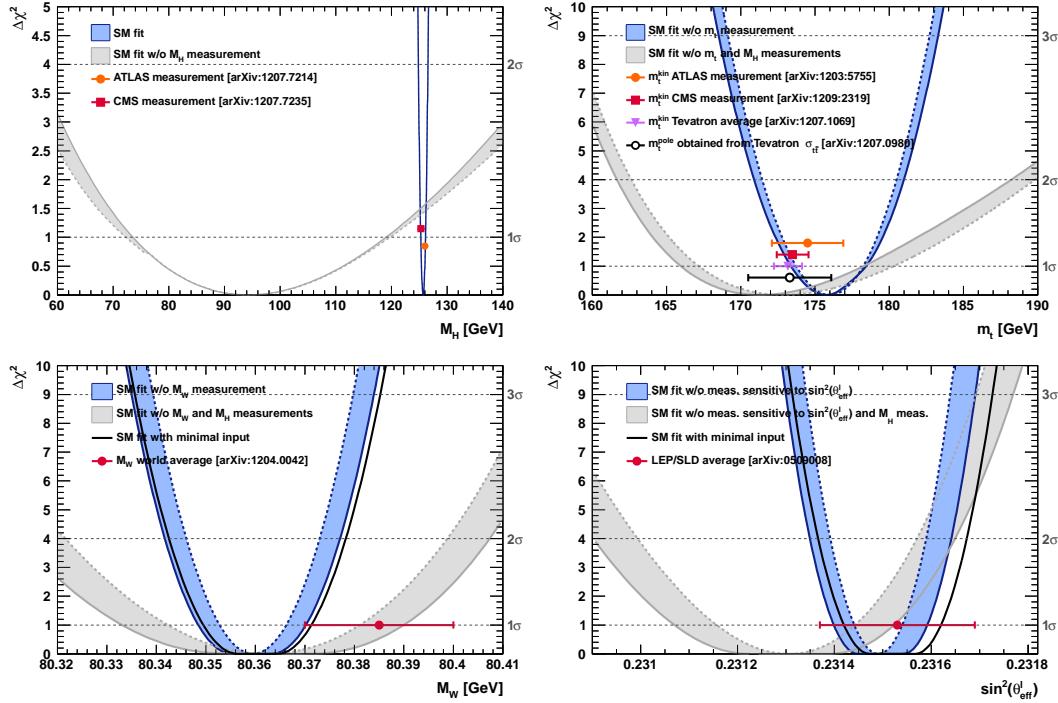


Figure 2.1:  $\chi^2$  as a function of the Higgs mass (top left), the top quark mass (top right), the  $W$  boson mass (bottom left) and the effective weak mixing angle (bottom right) for the combined SM fit from the GFitter group. The data points placed along  $\chi^2 = 1$  represent direct measurements of the respective observable and their  $\pm 1\sigma$  uncertainties. The grey (blue) bands show the results when excluding (including) the new  $M_H$  measurements from (in) the fits.

219 Electron-positron collider machines have a rich history of discovery and measurement in particle  
 220 physics. The advantage of electron accelerators is that the colliding element is itself a fundamental  
 221 particle. However, due synchrotron radiation, curvature of the beam line becomes problematic for  
 222 high energy beams. Hadron colliders, specifically, proton-proton and proton-anti-proton colliders can  
 223 be accelerated in rings without large losses due to synchrotron radiation, but the actual colliding  
 224 objects at high energies are the constituent quarks and gluons. This complicates analysis because the  
 225 initial state of the system is not discernible on a per-collision basis and momentum of hard scatter  
 226 system is unknown along the beam direction.

227 Collider physics rely on form-factor descriptions of the colliding hadrons that describe the fraction  
 228 of momentum carried by the hadrons constituent 'partons'. These are called parton-distribution  
 229 functions, seen in Figure 2.2, and are factorized and integrated through the theoretical calculations  
 230 of various collision processes [?].

231 At the Large Hadron Collider or LHC (discussed in detail in Chapter 3 a proton-proton colluder,

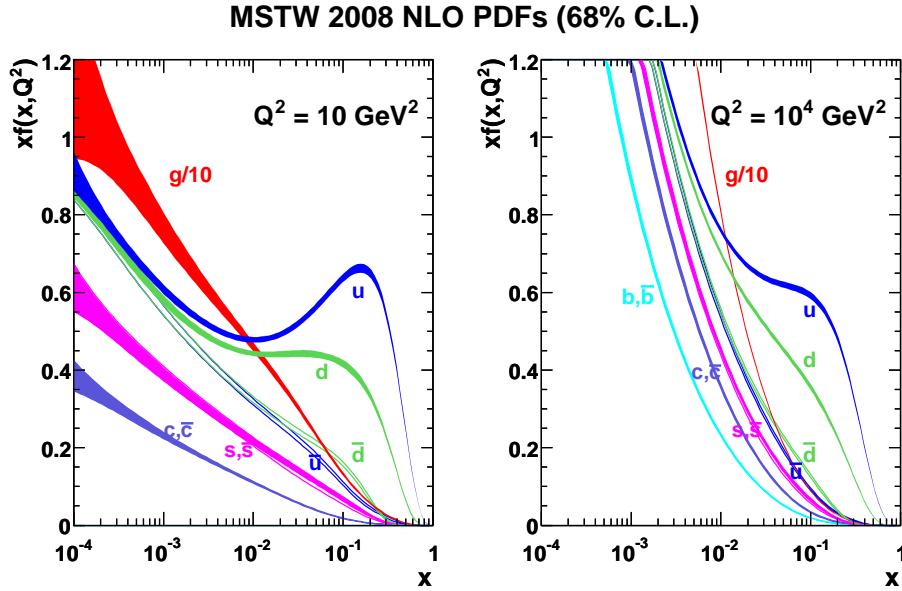


Figure 2.2: Proton Parton Distribution Functions (PDFs) from the MSTW Collaboration at  $Q^2 = 10 \text{ GeV}^2$  and  $Q^2 = 10^4 \text{ GeV}^2$

the types of initial states are quark-quark, quark-gluon, and gluon-gluon. Gluon collisions dominate overall, due to the large number of gluons inside the proton, though the relative importance of different initial states changes with the energy scale of the collision and the type of final state sought after.

A prime motivation for the construction of the Large Hadron Collider was the discovery or exclusion of the Higgs boson[?]. LEP and the Tevatron excluded large swaths of possible Higgs boson masses, especially below  $114 \text{ GeV}/c^2$  and the unitarity of certain diagrams including the  $WWWW$  vertex required the mass to be below about 1 TeV, a range that was achievable at the LHC with high luminosity running [?].

Despite the ubiquity of Higgs couplings, Higgs boson production at the LHC is a low rate process. Because it couples to fermions proportional to mass and because the colliding particles must be stable and therefore light, production of the Higgs must occur through virtual states.

The Higgs boson can be produced through collision at the LHC via 4 mechanisms: gluon-fusion (ggF), vector-boson fusion (VBF), Higgsstrahlung (VH), and production in association with top quarks ( $t\bar{t}H$ ). The diagrams are shown in Figure 2.3 and the production cross-sections as a function of Higgs mass for the 8 TeV LHC proton-proton running are shown in Figure 2.4 [?]. The largest production cross-section is via the gluon fusion channel at 20 pb, which proceeds through a fermion loop that is dominated by the top quark, because of its large Yukawa coupling to the Higgs.

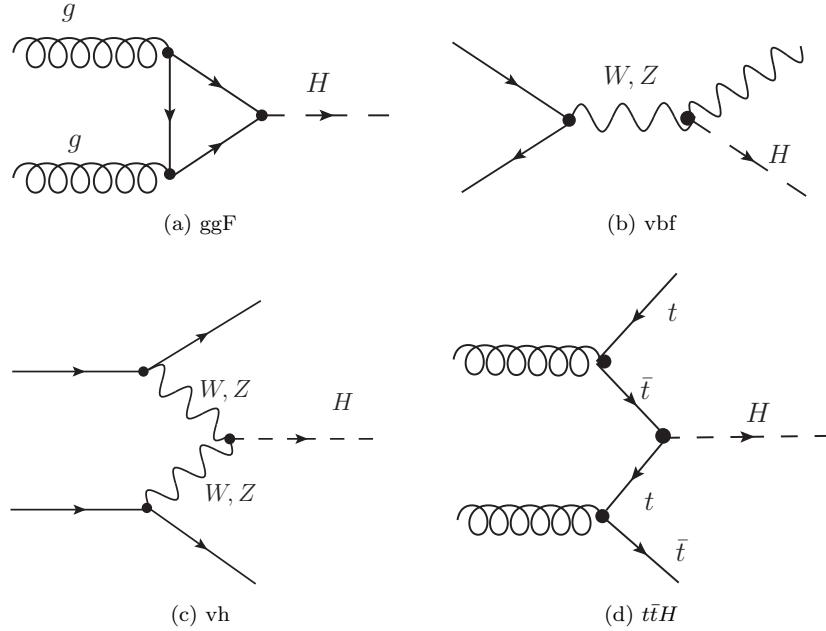


Figure 2.3: Dominant Higgs production modes at the LHC

Because the Higgs' couples to every massive particle, it has a rich set of decays also seen in Figure 2.4, especially for  $m_H = 125$ . Studies of Higgs properties at hadron colliders offers many tests of the Standard Model and ample room for searches for new physics. These tests specifically can verify the link between Yukawa coupling and the particles mass and further constrain details of EWSB by examining Higgs coupling to the weak bosons.

### 2.2.1 Higgs Discovery at the LHC

In 2012 both ATLAS and CMS announced the discovery of a new boson consistent with the Higgs by examining the results of Higgs searches in a number of decay channels ( $H \rightarrow W^+W^-$ ,  $H \rightarrow Z^0Z^0$ , and  $H \rightarrow \gamma\gamma$ ) in the 2011 dataset at  $\sqrt{s} = 7$  TeV and part of the 2012 dataset at  $\sqrt{s} = 8$  TeV. By 2013 and 2014, both experiments have updated and/or finalized their results for the full 2011 and 2012 datasets [?, ?]. I will focus on the ATLAS results in the following. The ATLAS measurements constrain both the Higgs mass[?] and spin[?], as well as provide constraints to the Higgs couplings to different particles.

Figure 2.5 show the results of the searches in all of the measurement channels as well as constraints on the SM Higgs coupling parameters in an example fit, where the couplings to the top-quark, bottom-

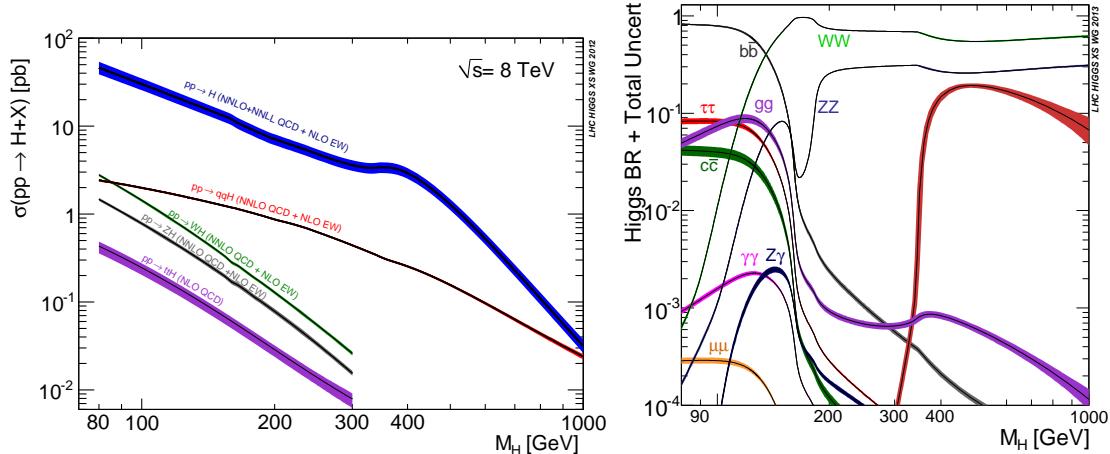


Figure 2.4: 8 TeV LHC Higgs production cross-sections (left) and decay branching fractions

quark, W,Z, and  $\tau$  are allowed to fluctuate independently. These rely on measurements binned in different production and decay channels. They are dominated by higher statistics results in the gluon-fusion production modes, but measurements in the VH and VBF modes are close to SM sensitivity.

The combined results show basic agreement with the SM with much room for improvement with the addition of new production and decay modes and higher statistics. The coupling constraints are particularly strong for the W and Z, which are the most sensitive decay channels, and top quark due to the dominance of the top Yukawa in the ggF loop.

### 2.2.2 $t\bar{t}H$ Production

Notably absent thus far in the SM are searches for the Higgs in the  $t\bar{t}H$  production channel, due to the lack of statistics. Searches are underway and initial results are close to SM sensitivity for ATLAS and CMS. More will be discussed about particular searches later.  $t\bar{t}H$  production depends on the top Yukawa coupling at tree level. Comparison of the gluon-fusion and the  $t\bar{t}H$  modes would allow for disentangling the effects of new particles in the gluon-fusion loop[?]. For instance, generic models would allow for the introduction of new colored particles into the loop. The simplest of these models would be the addition of a new generation of quarks. Fourth generation quarks, which obtain mass from a Higgs Yukawa coupling are already largely excluded due to their enormous effects on the Higgs production cross-section[?]. Other exotic scenarios allow for new colored particles in the gluon-fusion loop, which are not entirely constrained by present measurements[?, ?, ?]. These include, for instance, Supersymmetric models involving the stop quark.

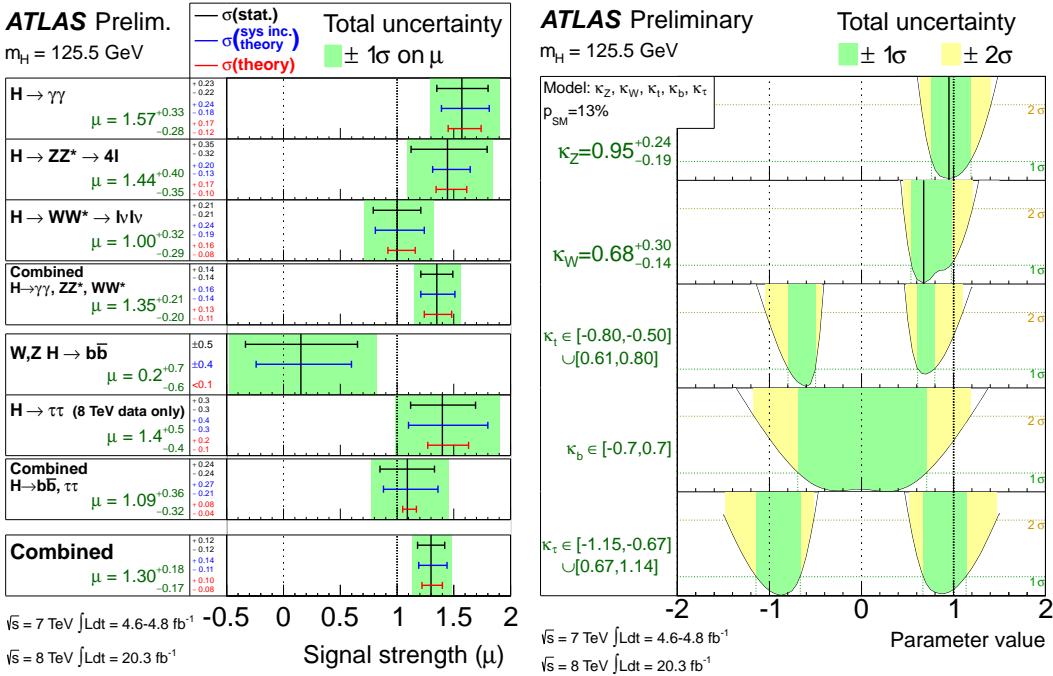


Figure 2.5: ATLAS Higgs combination results for all SM measurement channels as ratios of the measured to SM production cross-sections (left) and extracted Higgs coupling constraint scale-factors for a combined fit to the measurement channels, where the  $W, Z$ , top-quark,  $b$ -quark, and  $\tau$  couplings are allowed to float. The p-value of this particular model is 0.13 and in agreement with SM expectations

283      Aside from the loop effects, measurement of the  $t\bar{t}H$  production cross-section would provide a  
 284      precise measurement of the top Yukawa coupling. When compared with the measured top quark mass,  
 285      this tests the Higgs mass generation properties for up-type quarks. Despite similar uncertainties on  
 286      the overall production cross-sections for  $t\bar{t}H$  and the gluon-fusion modes, both of which depend on the  
 287      top Yukawa, most of these uncertainties would cancel for  $t\bar{t}H$  if normalized to the topologically similar  
 288       $t\bar{t}Z$ . Finally, the uniqueness of the experimental signature means that searches for  $t\bar{t}$  signatures can  
 289      be performed for a variety of Higgs decays ( $\gamma\gamma, b\bar{b}, WW, ZZ$ , and  $\tau\bar{\tau}$  with roughly similar degrees of  
 290      sensitivity (within a factor of 10)[?].

291      It is important to note the importance of the top Yukawa coupling to the overall structure of the  
 292      SM, due to its enormous size compared to other couplings. The top Yukawa is for instance 350000x as  
 293      large as the electron Yukawa coupling. Because of this the top Yukawa coupling, along with the Higgs  
 294      mass, is one of the most important pieces of the renormalization group equations (RGE) responsible  
 295      for the running of the parameter that determines the Higgs self-coupling  $\lambda$ . If this parameter runs  
 296      negative, then the potential responsible for the entire mechanism of EWSB no longer has a minimum  
 297      and becomes unbounded, resulting in instability in the universe [?]. Metastability occurs when the

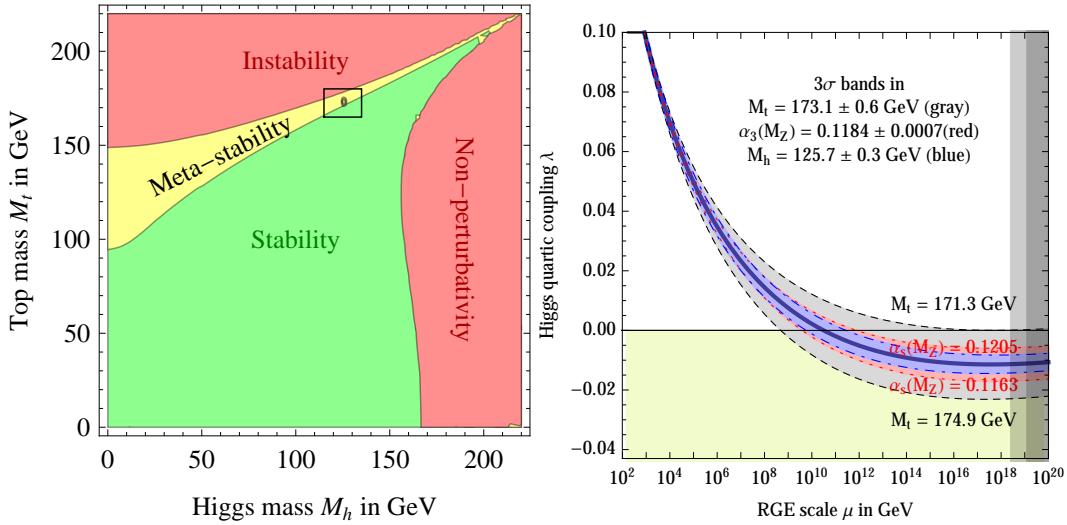


Figure 2.6: RGE for the running of the SM parameter,  $\lambda$  for the Higgs self-coupling term with present values and uncertainty bands for  $M_H$  and  $M_t$  (left). The two-dimensional plot colored (right) shows regions for which the SM is stable, unstable and metastable based on this RGE

shape of the potential allows for a false local minimum. Figure 2.6 shows the running of this parameter, the regions for which the universe is stable, unstable and metastable. Current measurements suggest that universe lies in a metastable island<sup>4</sup>.

### 2.3 Conclusion

The Standard Model, despite its success in providing a unified description of fundamental particles and interactions into single theory, has its flaws. These have been discussed in depth elsewhere, but include issues like the description of massive neutrinos, the failure to include gravity, and the unnaturalness of large quantum corrections to Higgs parameters. For these reasons, it seems the SM might be a lower energy approximation to a more fundamental theory. The discovery of the Higgs boson at the LHC provided a stunning verification of one of the fundamental aspects of the theory but at the same time offers new area to search for glimpses of something grander. The production of samples of Higgs bosons allows for a rich array of new tests of the Standard Model, which is now finally over-constrained by experiment. Searches for the  $t\bar{t}H$  production, one category of which is the topic of this thesis, provide tree-level access to a central parameter of the theory, the top Yukawa coupling, as well as access a variety of Higgs decays, which will eventually provide a rigorous new test of the SM.

<sup>4</sup>The RGE assumed that there is no new physics at all energy scales

314

## CHAPTER 3

315

# The Large Hadron Collider and the ATLAS Experiment

316

### 317 3.1 The Large Hadron Collider

318 Production of a sufficient number of high energy collisions to adequately explore particle physics at  
319 the electro-weak scale required the development of one of the most complex machines ever built, the  
320 Large Hadron Collider or LHC.

321 The LHC is the world's highest energy particle accelerator and is located 100m underneath the  
322 Franco-Swiss border at the European Organization for Nuclear Research (CERN) in a 26.7 km tunnel.

323 The technology involved in the development of the LHC and very briefly touched upon in this  
324 chapter is an enormous achievement in its own right and is documented in detail here [?, ?, ?].

325 The LHC is a circular machine capable of accelerating beams of protons and colliding them at  
326 center of mass energies up to  $\sqrt{s} = 14\text{TeV}$  at 4 collision sites around the ring, where 4 experiments  
327 are housed (ATLAS[?], CMS[?], LHCb[?], and ALICE[?]). Figure 3.1 is a diagram of the layout of the  
328 LHC and its experiments[?]. The LHC also operates in modes with beams of heavy ions. The LHC  
329 is composed of thousands of super-conducting Niobium-Titanium magnets, cooled to  $2.7^\circ \text{ C}$  with  
330 liquid Helium, which steer and focus the particle beams, and a superconducting resonant-frequency  
331 (RF) cavity, which boosts the beam to higher energies.

#### 332 3.1.1 The Accelerator Complex

333 The accelerator complex is a progressive series of machines with the LHC as the final stage. Protons are  
334 obtained from hydrogen atoms and are accelerated to 50 MeV using the Linac2, before being injected  
335 into the Proton-Synchrotron Booster (PSB). In the PSB the protons are accelerated to energies of 1.4

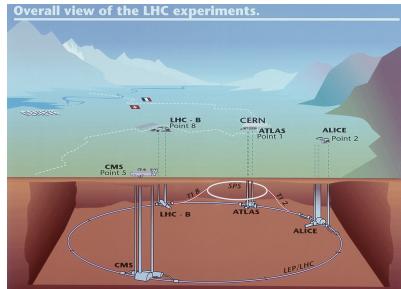


Figure 3.1: Diagram of the Large Hadron collider and location of the 4 main experiments (ATLAS, CMS, LHCb, and ALICE). Around the ring. The diagram also shows the location of the SPS, the final booster ring the accelerator complex that accelerates the protons to 450 GeV before injection into the LHC.

336 GeV for injection in to the Proton-Synchrotron (PS). The PS accelerates the protons to 25 GeV and  
 337 dumps bunches into the Super Proton Synchrotron (SPS), where they are accelerated to 450 GeV  
 338 and finally dumped into the LHC.

### 339 3.1.2 Beam Parameters and Collisions

340 For the physics studied at the ATLAS experiment, the two most important parameters of the collisions  
 341 provided by the LHC are the center of mass energy and instantaneous luminosity. High center of mass  
 342 energies are necessary for the production of new high mass particles, and because the constituents of  
 343 the actual collisions are the partons of the proton, the CME of the collisions must in general be much  
 344 higher than the mass of the particles needed to be produced. The

345 The instantaneous luminosity of the collisions,  $\mathcal{L}$ , is a measure of the collision rate. The integrated  
 346 luminosity over time is a measure of the size of the dataset and when multiplied by the cross-section of a  
 347 particular process gives the total number of expected events produced for that process. Instantaneous  
 348 luminosity depends on the number of colliding bunches of protons, the intensity of those bunches, the  
 349 revolution frequency, and the nomralized transverse spread of the beam in momentum and position  
 350 phase space, called the emmitance, and the transverse beam size. The LHC has the option for colliding  
 351 beams with 2808 bunches of protons, each with around  $10^{11}$  protons, at a rate of one bunch collision  
 352 every 25 ns, or 40 MHz. These correspond to a design luminosity of around  $10^{34} \text{ cm}^2 \text{ s}^{-1}$  or  $10 \text{ nb}^{-1}$

353 s<sup>-1</sup>

354

## CHAPTER 4

---

355

# Electrons

---

356 **4.1 Electrons at Hadron Colliders**

357 **4.2 Reconstruction of Electron at ATLAS**

358 **4.3 Identification of Electrons at ATLAS**

359 **4.3.1 Pile-up**

360 **4.3.2 Trigger vs. Offline**

361 **4.3.3 2011 Menu**

362 **4.3.4 2012 Menu**

363 **4.3.5 Electron Likelihood**

364 **4.4 Measurement of Electron Efficiency at ATLAS**

365 **4.4.1 Techniques**

366 **4.4.2 Issues**

## CHAPTER 5

---

# Search for the TTH Decay in the Multilepton Channel

---

370 This chapter provides an overview of the set of analyses searching for the Standard Model (SM)  
 371 production of the Higgs boson in association with top quarks in multi-lepton final states with multiple  
 372 jets (including b-quark tagged jets). Searches in  $t\bar{t}H$  final states with 2 same-charge, 3 and 4 light  
 373 leptons ( $e, \mu$ ) are discussed in depth. These final states target specifically Higgs decays to vector  
 374 bosons,  $H \rightarrow W^\pm W^\pm$  and  $H \rightarrow Z^\pm Z^\pm$  and form a complement to searches for  $t\bar{t}H$  production in  
 375 final states targeting the  $H \rightarrow b\bar{b}$  [?],  $H \rightarrow \gamma\gamma$  [?], and  $H \rightarrow \tau\tau$  decay modes.

376 Based on SM production cross-sections, observation lies just outside the sensitivity of the Run I  
 377 dataset, even when combining all searches. The analyses provide an opportunity to constrain for the  
 378 first time the  $t\bar{t}H$  production mode with limits reasonably close to the actual production rate. As  
 379 such the analysis is optimized to overall sensitivity to the  $t\bar{t}H$  production rather than individual decay  
 380 modes, which would be more useful for constraining Higgs couplings.

381 Detailed description of the event and objection section are provided in Chapter 7, background  
 382 modelling in Chapter 8, the effect of systematic errors and the statistical analysis in Chapter ?? and  
 383 final results in Chapter ??.

384 **5.1 Signal Characteristics**

385  $t\bar{t}H$  can be observed in a number of different final states related to the Higgs boson and the top  
 386 quark decay modes.

387 Three Higgs boson decays are relevant for this analysis:  $W^+W^-$ ,  $\tau^+\tau^-$  and  $ZZ$ . The top and  
 388 anti-top quarks decay in  $W^\pm b$ . Each  $W^\pm$  boson decays either leptonically ( $l=e^\pm, \mu^\pm, \tau^\pm$ ) with missing

389 energy or hadronically. Table 5.1 provides the fractional contribution of the main Higgs decay modes  
390 at the generator level to  $t\bar{t}H$  search channels. These numbers will be modified by lepton acceptances.

Table 5.1: Contributions of the main Higgs decay modes to the 3 multilepton  $t\bar{t}H$  signatures at generation level.

Signature	$H \rightarrow WW$	$H \rightarrow \tau\tau$	$H \rightarrow ZZ$
Same-sign	100%	–	–
3 leptons	71%	20%	9%
4 leptons	53%	30%	17%

391 All modes are generally dominated by the  $WW$  signature, though the 3l and 4l channels possess  
392 some contribution from the  $\tau\tau$  and  $ZZ$  decays.

393 The signal is expected to be characterized by the presence of 2 b-quark jets from the top quark  
394 decays, leptons from vector boson and tau decays, a high jet multiplicity, and missing energy. In  
395 general, the number of leptons is anti-correlated with the number of jets, since a vector boson can  
396 either decay leptonically or hadronically. For  $H \rightarrow W^+W^-$ , the light quark multiplicity,  $N_q$ , and the  
397 number of leptons,  $N_l$ , follow this relation:  $2N_l + N_q + N_b = 10$ .

- 398 • In the same-sign channel, the  $t\bar{t}H$  final state contains 6 quarks. These events are then characterised by a large jet multiplicity.
- 400 • In the 3 lepton channel, the  $t\bar{t}H$  final state contains 4 quarks from the hard scatter.
- 401 • In the 4 lepton channel, the  $t\bar{t}H$  final state contains a small number of light quarks, 0 ( $H \rightarrow$   
402  $W^+W^-$  case), 2 or 4 ( $H \rightarrow ZZ$  case).

## 403 5.2 Background Overview

404 Background processes can be sorted into two categories:

- 405 • Events with a non prompt or a fake lepton selected as prompt lepton. These processes cannot lead to a final state compatible with the signal signature without a misreconstructed object. This category includes events with a prompt lepton but with misreconstructed charge<sup>5</sup> and events

---

5Charge mis-identification is almost exclusively a phenomenon of electrons at our energy scales. While it is possible for both electrons and muons to have an extremely straight track, whose direction of curvature is difficult to measure, this happens only at extremely high momentum. Electrons on the other hand may interact with the detector material, resulting in bremsstrahlung. The bremsstrahlung photon may subsequently convert resulting in 2 additional tracks near the original electron. This process, called a trident process, may cause a mismatching of the electron cluster with the original electron track and thus a charge-mis identification and happens also at low energies

408       with jets that "fake" leptons. These processes are rejected with tight object isolation and  
409       identification criteria, requiring a large jet multiplicity, and veto-ing events consistent with a  
410       leptonically decaying Z boson.

411       The main backgrounds of this sort are:  $t\bar{t}$  and  $Z+jets$ . Data-driven techniques are used to  
412       control some of these processes. Their importance varies depending on the channel.

- 413       • Events which can lead to the same final state as the signal (irreducible backgrounds). The  
414       main background of this category are:  $t\bar{t}V$ ,  $W^\pm Z$ , and  $ZZ$ . They are modeled using the  
415       Monte Carlo simulations. In general, these backgrounds are combatted with jet and b-tagged  
416       jet requirements. Although the jet multiplicity of  $t\bar{t}V$  is high, the multiplicity of  $t\bar{t}H$  events is  
417       still higher.

418       **5.3 Analysis Strategy**

419       The analysis search is conducted in 3 channels, based on counting of fully identified leptons: 2 SS  
420       leptons, 3 leptons, and 4 leptons. The lepton counting occurs before additional object cuts are made in  
421       each individual channel to ensure orthogonality. The division into lepton channels rather than channels  
422       targeting specific decay modes allows channels with different sensitivities to be considered separately. We  
423       further divide the 2l SS into sub channels based on the number of jets and flavor of the leptons and  
424       the 4l channel into subchannels enriched and depleted in OS leptons arising from Z decays.

425       The channels are fed into a posson model

426

## CHAPTER 6

427

# Dataset and Simulation

## 428 6.1 Data

### 429 6.1.1 The 2012 Dataset

430 The LHC successfully produced datasets for physics studies in 2010, 2011 and 2012. The 2012 proton-  
431 proton dataset was delivered with collisions with a CME of 8 TeV with bunch collisions every 50 ns and  
432 reached a total integrated luminosity of around  $20 \text{ fb}^{-1}$  [?]. Figure 6.1 shows the accumulation of this  
433 dataset over time. Despite doubling the bunch spacing (thereby halving the bunch collision frequency),  
434 the luminosity neared the design luminosity due to unexpected improvements in the transverse beam  
435 profile[?]. This increased the amount of pile-up, or number of collisions per bunch crossing and in  
436 general collision events were busier due to these multiple interactions. Figure 6.2 shows the average  
437 number of interaction per bunch crossing for the 2011 and 2012 datasets. The 2012 dataset shows an  
438 average of 20-25 interactions.

439 The  $t\bar{t}H$  analysis uses the entire 2012 ATLAS dataset, collected from April to December. The size  
440 of the dataset corresponds to  $20.3 \text{ fb}^{-1}$ , after passing data quality requirements, ensuring the proper  
441 operation of the tracking, calorimeter and muon subsystems.

442 The datasets used in the analysis were collected with the primary electron (EF\_E24VHL\_MEDIUM1  
443 — EF\_E60\_MEDIUM1) and muon triggers ( EF\_24I\_TIGHT — EF\_36\_TIGHT). The electron  
444 triggers require an electron with at least 25 GeV of calorimeter energy, passing the medium identification  
445 requirement and loose tracking isolation. Above 60 GeV, the isolation requirement is dropped and the  
446 identification is loosened slightly. Similarly, the muon trigger requires a good inner detector track and  
447 matching hits in the muon spectrometer, as well as loose tracking isolation, which also is dropped  
448 about 36 GeV. The data sample must contain either a primary muon or primary electron trigger.

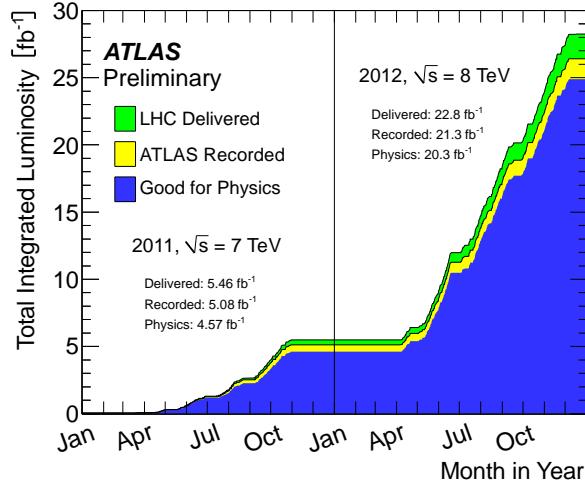


Figure 6.1: Plot showing the accumulation of the integrated luminosity delivered to the ATLAS experiment over 2011 and 2012. The rough size of the usable, physics ready dataset for 2012 is  $20 \text{ fb}^{-1}$  and is the dataset used for the following analysis.

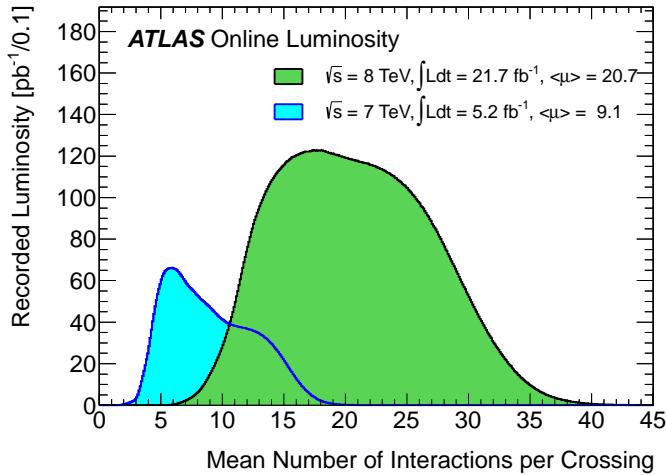


Figure 6.2: The average number of interactions per bunch-crossing for the 2012 and 2011 LHC proton-proton dataset. Most of these interactions are uninteresting but leave energetic signatures in particle detectors called pile-up which interfere with measurements

## 449 6.2 Simulation

450 Simulation samples based on are used to determine the overall event selection acceptance and efficiency  
 451 and for investigations not directly involved in the final result. The simulated samples are created using  
 452 parton distribution function (PDF) and model using Monte Carlo (MC) techniques the hard parton  
 453 scatter, underlying event activity and parton showering and hadronization. The samples are then

Table 6.1: Monte Carlo samples used for signal description.

Process	Generator	Cross-section [fb]	$\mathcal{L}$ [fb $^{-1}$ ]	Detector simulation
$t\bar{t}H \rightarrow \text{allhad} + H$	PowHel+Pythia8	59.09	2146.5	Full
$t\bar{t}H \rightarrow \text{ljets} + H$	PowHel+Pythia8	56.63	2238.9	Full
$t\bar{t}H \rightarrow ll + H$	PowHel+Pythia8	13.58	9332.0	Full

454 passed through a full ATLAS detector simulation[?] based on GEANT4 [?]. Small corrections are  
 455 then applied to the overall efficiencies to re-scale object identification efficiencies, energy scales, and  
 456 the pile-up, discussed in depth later.

### 457 6.2.1 Signal Simulation

458 The signal Monte Carlo samples are described in Table 6.1. These large samples are generated with  
 459 inclusive Higgs boson decays, with branching fractions set to the LHC Higgs Cross Section Working  
 460 Group (Yellow Report) recommendation for  $m_H = 125$  GeV [?]. The matrix element calculation is  
 461 performed at next-to-leading order (NLO); we use  $t\bar{t}H$  Les Houches event format files provided by  
 462 the authors of the PowHel software [?], decayed and showered with Pythia8[?]. The CT10[?] parton  
 463 distribution function is used for matrix element generation. The inclusive cross section (129.3 fb at  
 464  $m_H = 125$  GeV) is also obtained from the Yellow Report [?].

### 465 6.2.2 Background Simulation

466 The background simulations used for this analysis are listed in Table ???. In general, the Alpgen[?],  
 467 MadGraph[?], and AcerMC[?] samples use the CTEQ6L1[?] parton distribution function, while the  
 468 Powheg[?], Sherpa[?], are generated with the CT10 PDF. The exception is the MadGraph  $t\bar{t}t\bar{t}$  sample,  
 469 which is generated with the MSTW2008 PDF[?]. The highest order calculations available are used  
 470 for cross sections.

Table 6.2: Monte Carlo samples used for background description. Unless otherwise specified MadGraph samples use Pythia 6 for showering and Alpgen samples use Herwig+Jimmy.

Process	Generator	Detector simulation
$t\bar{t}W^\pm, t\bar{t}Z$	MadGraph	Full
$tZ$	MadGraph	AF2

Table 6.2: Monte Carlo samples used for background description. Unless otherwise specified MadGraph samples use Pythia 6 for showering and Alpgen samples use Herwig+Jimmy.

Process	Generator	Detector simulation
$t\bar{t}t\bar{t}$	MadGraph	Full
$t\bar{t}W^\pm W^\pm$	Madgraph+Pythia8	AF2
$t\bar{t}$	Powheg+Pythia6	Full/AF2
single top tchan	AcerMC+Pythia6	Full
single top schan $\rightarrow l$	Powheg+Pythia6	Full
single top $W^\pm t$	Powheg+Pythia6	Full
$W\gamma^*$	MadGraph	Full
$W\gamma + 4p_T$	Alpgen	Full
$W^+W^-$	Sherpa	Full
$W^\pm Z$	Sherpa	Full
Same-sign WW	Madgraph+Pythia8	AF2
$ZZ \rightarrow$	Powheg+Pythia8,gg2ZZ+Herwig	Full
$Z\gamma^*$	Sherpa	Full
$Z+jets$	Sherpa	Full
ggF_H(125)	Powheg+Pythia8	Full

471

## CHAPTER 7

472

# Object and Event Selection

473 As stated in Chapter 5, the analysis is divided into 3 signal regions based on lepton counting: 2  
474 same-charge leptons, 3 leptons and 4 leptons. The lepton counting occurs for fully identified leptons  
475 with full overlap removal with transverse momentum over 10 GeV to ensure orthogonality. Lepton  
476 selections are tightened afterward within each region.

477 The cuts for each signal region are provided in Table ?? and the object selections are detailed in  
478 the following sections. The selections are based on optimizations of the region sensitivity performed  
479 using MC (event for data driven backgrounds) and adhoc values for normalization systematic errors.  
480 The optimziation is detailed in Section ?? . All signal regions are comprised of three basic requirements:  
481 the presence of b-tagged jets, the presence of additional light jets, and a veto of same flavor opposite  
482 sign leptons with an invariant mass within the Z window. Additional requirements on the invariant  
483 mass of the leptons, the missing transverse energy in the event, and the total object energy ( $H_T$ )  
484 proved to have negligible additional benefit at our level of statistics.

485 **7.1 2l Same-Charge Signal Region**

486 The 2 lepton same-charge signal region (2l SS) requires two leptons similar charge. The signal is  
487 symmetric in charge but the background from  $t\bar{t}$  di-lepton production is overwhelming, necessitating  
488 the same-sign cut. Requiring only two leptons allows the extra 2 W bosons in the event to decay  
489 hadronically, resulting in on average 4 additional light jets plus 2 additional b-quark jets from the top  
490 decays.

491 A leading lepton with transverse momentum of at least 25 GeV/ $c$  that matches to a trigger and a  
492 subleading lepton of at least 20 GeV/ $c$ , a b-tagged jet, and at least 4 jets in total are required.

493 In order to suppress non-prompt backgrounds, the lepton isolation criteria for tracking and

Table 7.1: Selections in the 2l SS, 3l and 4l Signal Regions

Signal Region	2l SS	3l	4l
Trigger Matched Lepton	Yes	Yes	Yes
$N_l^6$	=2	=3	=4
Lepton Charge Sum	+2 or -2	+1 or -1	0
Lepton Momentum ( $\text{GeV}/c^7$ )	$p_{T0} > 25$ $p_{T1} > 20$	$p_{T0} > 10$ $p_{T1,2} > 25, 20$	$p_{T0} > 25$ $p_{T1} > 20$ $p_{T2,3} > 10$
Jet Counting	$N_b \geq 1, N_{Jet} = 4$	$N_b \geq 1, N_{Jet} \geq 4$ or $N_b \geq 2, N_{Jet} = 3$	$N_b \geq 1, N_{Jet} \geq 2$
Mass Variables ( $\text{GeV}/c^2$ )	$ M_{ee} - M_Z  < 10$	$ M_{SFOS} - M_Z  < 10$	$M_{SFOS} > 10$ $150 < M_{4l} < 500$ $ M_{SFOS} - M_Z  < 10$
Sub-channels	$2 (N_{Jet} = 4, N_{Jet} \geq 5)$ x 3(ee,e $\mu$ , $\mu\mu$ )	none	2 (No SFOS leps, SFOS leps)

494 calorimeter are tightened from less than 10% of the lepton momentum to 5%. To suppress charge mis-  
 495 indentification, the electron is required to be extremely central ( $|\eta| < 1.37$ ) to avoid the material-rich  
 496 regions of the detector. Additionally,  $ee$  events with a lepton pair invariant mass within 10  $\text{GeV}/c^2$  of  
 497 the Z pole are removed.

498 In order to maintain orthogonality with the  $\tau$  analyses, events with fully identified taus are vetoed.  
 499 For the statistical combination the channel is divided into 6 sub-channels: 2 jets counting bins  
 500 ( $N_{Jet} = 4, N_{Jet} \geq 5$ ) x 3 lepton flavor bins (ee, $\mu\mu$ ,e $\mu$ ). The splitting allows

## 501 7.2 3l Signal Region

502 The 3 lepton channel requires 3 leptons, whose summed charge is either  $-1$  or  $+1$ . The leptons are  
 503 ordered in this way:

- 504 • lep0: the lepton that is opposite in charge to the other two leptons
- 505 • lep1: the lepton that is closer in  $\Delta R$  to lep0
- 506 • lep2: the lepton that is farther in  $\Delta R$  from lep1

507 Since events with a "fake" lepton arise from di-lepton processes,  $t\bar{t}$  and Z+jets, where additional  
 508 jets are mis-identified as the third lepton, lep0 is never the fake lepton. As a result, the transverse  
 509 momentum requirement of lep0 is lower than the other two,  $> 25 \text{ GeV}/c$ . For the additional two  
 510 leptons, one must must match a trigger and have  $p_T > 25 \text{ GeV}/c$  and the other must have  $p_T > 10$   
 511  $\text{GeV}/c$ .

512        The 3l channel further requires at least one b-tagged jets and at least 4 jets in total, or two b-  
 513        tagged jets and exactly 3 jets in total. Additionally, to suppress  $W^\pm Z$  and Z+jet events, events with  
 514        same-flavor opposite sign pairs within  $10 \text{ GeV}/c^2$  of the Z pole are vetoed.

515        Additional cuts, including an  $M_{ll}$  cut, and splittings were investigated but low statistics proved  
 516        to wash out any advantages.

### 517      7.3 4l Signal Region

518        In the four lepton signal region, selected events must have exactly four leptons with a total charge  
 519        of zero. At least one leptons with must be matched to one of the applied single lepton trigger. The  
 520        leading and sub-leading leptons are required to have a  $p_{\text{T}}$  of 25 and 15 GeV respectively. In order to  
 521        suppress background contributions from low-mass resonances and Drell-Yan radiation, all opposite-  
 522        sign-same-flavour (OS-SF) lepton pairs are required to have a dilepton invariant mass of at least 10  
 523        GeV.

524        The four-lepton invariant mass is required to be between 100 and 500 GeV. This choice of mass  
 525        window suppresses background from the on-shell  $Z \rightarrow 4\ell$  peak and exploits the high-mass differences  
 526        between the signal and the dominant  $t\bar{t}Z$  background. Events containing an OS-SF lepton pair  
 527        within 10 GeV of the Z boson mass are discarded. This Z-veto procedure greatly reduces background  
 528        contributions from  $ZZ$  production as well as  $t\bar{t}Z$  and while it also affects the signal by vetoing  
 529         $H \rightarrow ZZ^*$ ,  $Z \rightarrow \ell^+\ell^-$ , these events constitute a small amount of the total expected signal. Finally,  
 530        selected events are required to have at least two jets, at least one of which must be tagged as a b-quark  
 531        initiated jet.

532        The contribution from  $t\bar{t}Z$  comprises approximately 75% of the total background in the inclusive  
 533        signal region. A signal region categorization which factorizes  $t\bar{t}Z$  from the remaining backgrounds is  
 534        thus beneficial. The signal region is accordingly divided into two categories based on the presence of  
 535        OS-SF lepton pairs in the final state.

### 536      7.4 Electron Selection

537        The electrons are reconstructed by a standard algorithm of the experiment [?] and the electron  
 538        cluster is required to be fiducial to the barrel or endcap calorimeters:  $|\eta_{\text{cluster}}| < 2.47$ . Electrons in  
 539        the transition region,  $1.37 < |\eta_{\text{cluster}}| < 1.52$ , are vetoed. Electron reconstruction and identification  
 540        is discussed in depth in Chapter 4. Electrons must pass the the VERYTIGHT likelihood identification  
 541        criteria.

542 In order to reject jets (b-quark jets in particular) misidentified as electrons, electron candidate  
 543 must also be well isolated from additional tracks and calorimeter energy around the electron cluster.  
 544 Both the tracking and calorimeter energy within  $\Delta R = 0.2$  of the electron cluster must be less  
 545 than 5% of the electron transverse momentum:  $ptcone20/P_t < 0.05$  and  $Etcone20/E_T < 0.05$ . All  
 546 quality tracks with momentum greater than 400 MeV contribute to the isolation energy. Calorimeter  
 547 isolation energy is calculated using topological clusters with corrections for energy leaked from the  
 548 electron cluster [?]. Pile-up and underlying event corrections are applied using a median ambient  
 549 energy density correction, developed in [?].

550 The electron track must also match the primary vertex. The longitudinal projection of the track  
 551 along the beam line,  $z0 \sin \theta$ , must be less than 1 cm) and the transverse projection divided by the  
 552 parameter error,  $d0$  significance, must be less than 4. These cuts are used in particular to suppress  
 553 backgrounds from conversions, heavy-flavor jets and electron charge-misidentifications.

554 The electron selection is provided in Table ??.

## 555 7.5 Muon Selection

556 Muons used in the analysis are formed by matching reconstructed inner detector tracks with either  
 557 a complete track or a track-segment reconstructed in the muon spectrometer (MS). The muons must  
 558 satisfy  $|\eta| < 2.5$ . The muon track are required to be a good quality combined fit of inner detector  
 559 hits and muon spectrometer segments, unless the muon is not fiducial to the inner detector,  $|\eta| > 2.4$ .  
 560 Muons with inner detector tracks are further required to pass standard inner detector track hit  
 561 requirements [?].

562 As with electrons, muons are required to be isolated from additional tracking or calorimeter energy:  
 563  $ptcone20/P_t < 0.1$ ,  $Etcone20/E_T < 0.1$ ) A cell-based  $Etcone20/P_T$  relative isolation variable is used.  
 564 A pile-up energy subtraction based on the number of reconstructed vertices in the event is applied.  
 565 The subtraction is derived from a Z boson control sample.

566 The muons must also originate from the primary vertex and have impact parameter requirements,  
 567  $d0$  significance  $< 3$ , and  $z0 \sin \theta < 0.1$  cm, similar to the electrons.

568 The muon selection is provided in Table ??.

## 569 7.6 Jet and b-Tagged Jet Selection

570 Jets are reconstructed in the calorimeter using the anti- $k_t$  [?] algorithm with a distance parameter of  
 571 0.4 using locally calibrated topologically clusters as input (LC Jets).

572        Jets must pass loose quality requirement, ensuring the proper functioning of the calorimeter at the  
573        time of data taking. Jets near a hot Tile cell in data periods B1/B2 are rejected. The local hadronic  
574        calibration is used for the jet energy scale, and ambient energy corrections are applied to account for  
575        energy due to pileup.

576         $p_T$  and  $\eta$  cuts are tuned based on the sensitivity to  $t\bar{t}H$  as explained in section ??.

577        For jets within  $|\eta| < 2.4$  and  $p_T < 50$  GeV, are required to be associated with the primary vertex,  
578        the “jet vertex fraction” (or JVF), which is the fraction of track  $p_T$  associated with the jet that comes  
579        from the primary vertex, must exceed 0.5 (or there must be no track associated to the jet).

580        B-jets are tagged using a Multi-Variate Analysis (MVA) method called MV1 and relying on infor-  
581        mation of the impact parameter and the reconstruction of the displaced vertex of the hadron decay  
582        inside the jet. The output of the tagger is required to be above 0.8119 which corresponds to a 70%  
583        efficient Working Point (WP).

584        **7.7 Tau Selection**

585        The tau selection is important only in the 2l SS channel, due the tau veto to ensure orthogonality  
586        with analyses searching for tau final states for a future combination.

587        **7.8 Object Summary and Overlap**

588        **7.9 Optimization**

Parameter	Values	Remarks
<b>Electrons</b>		
$p_T$	$> 10 \text{ GeV}$	
$ \eta $	$< 2.47$ veto crack	
ID	Very Tight Likelihood	$< 1.37$ for 2l SS channel
Isolation	$E_{\text{T}}\text{Cone}20/p_T, p_{\text{T}}\text{Cone}20/p_T < 0.05$	
Jet overlap removal	$\Delta R > 0.3$	
$ d_0^{\text{sig}} $	$< 4\sigma$	
$z_0 \sin\theta$	$< 1 \text{ cm}$	
<b>Muons</b>		
$p_T$	$> 10 \text{ GeV}$	
$ \eta $	$< 2.5$	
ID	Tight	
Jet overlap removal	$\Delta R > 0.04 + 10 \text{ GeV}/p_T$	
Isolation	$E_{\text{T}}\text{Cone}20/p_T, p_{\text{T}}\text{Cone}20/p_T < 0.1$	$< 0.05$ for 2 leptons
$ d_0^{\text{sig}} $	$< 3\sigma$	
$z_0$	$< 1 \text{ cm}$	
<b>Taus</b>		
$p_T$	$> 25 \text{ GeV}$	
ID	Medium BDT	
Isolation		
Jet overlap removal	$\Delta R > 0.3$	
$e/\mu$ vetoes	Medium electron veto	
<b>Jets</b>		
$p_T$	$> 25 \text{ GeV}$	
$ \eta $	$< 2.5$	
JVF	$\text{JVF} > 0.5$ or no associated track or $p_T > 50 \text{ GeV}$	
b-Tag	MV1 70% operating point	

Table 7.2: Object identification and selection used to define the 5 channels of the multilepton  $t\bar{t}H$  analysis. Some channels use a sub-sample of objects as precised in the *Remarks* column.

## CHAPTER 8

---

# Background Estimation

---

591 The  $t\bar{t}H$  multi-lepton signal regions discussed in Chapter 5 are contaminated by background contributions at a similar order of magnitude to the signal. The dominant background for each region is  
 592 vector boson production in association with top quarks ( $t\bar{t}V$ ). Sub-dominant but important backgrounds include the production of vector boson pairs in association with jets and b-quark jets (VV)  
 593 and  $t\bar{t}$ production with a jet misidentified as a lepton. The 2l SS regions possesses a unique background  
 594 of charge misidentification from Z and top events. The methods for estimating these backgrounds are  
 595 discussed in this chapter. Monte Carlo simulation is used for the prompt  $t\bar{t}V$  and VV contributions.  
 596 Systematic uncertainties on the overall normalization of these backgrounds in the signal region are  
 597 provided from theoretical studies and past ATLAS analyses and are verified in data-based validation  
 598 regions. The non-prompt backgrounds from  $t\bar{t}$ jet-misidentification and charge-misidentification are  
 599 estimated using data-driven methods.

600 For reference, Table ?? provides a summary of the  $t\bar{t}H$  signal and background expectation for each  
 601 of the signal regions, including the data-driven estimates discussed in this section. For each region,  
 602 the background contribution exceeds the size of the signal.

603 **8.1 Vector Boson ( $W^\pm, Z$ ) production in association with top quarks:**

604  $t\bar{t}V, tZ$

605 This section describes the estimation and  $t\bar{t}V$  productions. Production of top quarks plus vector  
 606 boson is an important background in all multilepton channels. A large part of the  $t\bar{t}V$  component,  
 607 arising from on-shell  $Z \rightarrow \ell\ell$ , can be removed via a  $Z$  mass veto on like-flavour, opposite sign leptons.  
 608 However the  $Z \rightarrow \tau\tau$  and  $\gamma^*$  components remain. The  $t\bar{t}W^\pm$  and  $tZ$  processes generally require extra  
 609 jets to reach the multiplicity of our signal regions. Uncertainties from the choice of the factorization  
 610

Process	$\sigma_{NLO}$ [fb]	Scale Uncertainty [%]		PDF Uncertainty [%]		Total symmetrised uncertainty [%]
$t\bar{t}W^+$	144.9	+10	-11	+7.7	-8.7	13.3
$t\bar{t}W^-$	61.4	+11	-12	+6.3	-8.4	13.6
$t\bar{t}Z$	206.7	+9	-13	+8.0	-9.2	14.0

Table 8.1: NLO cross section and theoretical uncertainty calculations derived from MadGraph5\_aMC@NLO.

( $\mu_F$ ) and renormalisation  $\mu_R$  scales as well as from the PDF sets are considered evaluating their impacts on both the production cross sections and on the event selection efficiencies (particularly resulting from effects on the shape of number of jets spectrum).

Monte Carlo events for these processes are generated with MadGraph 5 and showered with Pythia.  $t\bar{t}W^\pm$  events are generated with up to two extra partons at matrix element level, while for  $t\bar{t}Z$  up to one extra parton at matrix-element level is produced. The  $tZ$  process is simulated without extra partons. The next-to-leading-order (NLO) cross sections are implemented by applying a uniform  $k$ -factor to the leading-order (LO) events for each process. For  $t\bar{t}Z$ , there is a large component of off-shell production, and for the 3 and 4  $\ell$  channels low mass  $\gamma^*/Z \rightarrow \ell\ell$  is an important background after on-shell production is removed with a  $Z$  veto. In this case the  $k$ -factor is determined by comparing LO and NLO cross sections for on-shell  $Z$  production only.

The  $t\bar{t}V$  uncertainties are calculated using the internal QDC scale and PDF reweighting that is available with MadGraph5\_aMC@NLO. The prescription for the scale envelope is taken from [?]: the central value  $\mu = \mu_R = \mu_F = m_t + m_V/2$  and the uncertainty envelope is  $[\mu_0/2, 2\mu_0]$ . The PDF uncertainty prescription used is the recipe from [?]: calculate the PDF uncertainty using the MSTW2008nlo [?] PDF for the central value and then the final PDF uncertainty envelope is derived from three PDF error sets each with different  $\alpha_S$  values (the central value and the upper and lower 90% CL values). The final NLO cross section central values and uncertainties are given in Table 8.1.

The  $tZ$  process is normalized to NLO based on the calculation in Ref. [?]. Here the scales are set to  $\mu_0 = m_t$  and the scale variations are by a factor of four; the scale dependence is found to be quite small.

### 8.1.1 $t\bar{t}Z$ Validation Region

Unlike  $t\bar{t}W^\pm$ , a  $t\bar{t}Z$  validation region can be obtained by simply inverting the veto on same-flavor opposite sign lepton pairs near the  $Z$  pole in the 3 lepton signal region. This region thus requires 3 leptons (with momentum and identification cuts discussed in Chapter 7, at least one opposite sign,

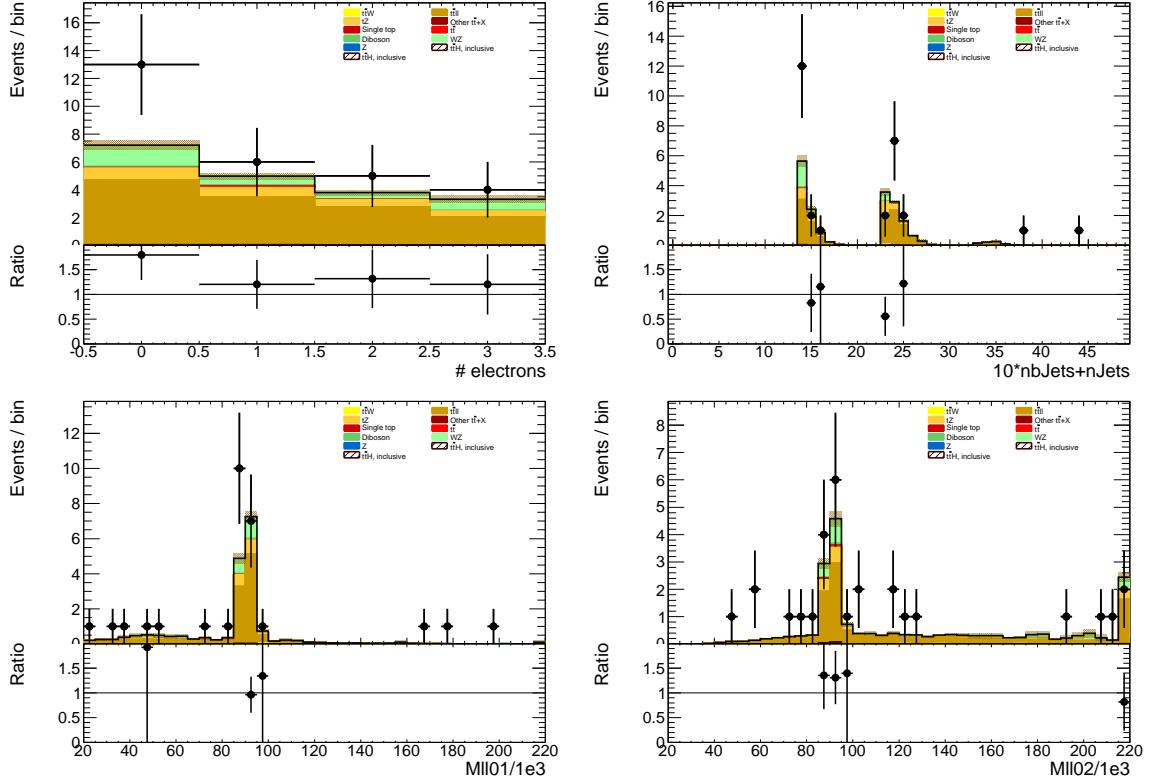


Figure 8.1: Data/MC comparison plots for  $t\bar{t}Z$  control region A ( $\geq 4$  jets,  $\geq 1$   $b$ -tag and 3 jets,  $\geq 2$   $b$ -tag). In all plots, the rightmost bin contains any overflows. Top left: number of electrons. Top right:  $10^*$ the number of  $b$ -tags + the total number of jets. Middle left: the invariant mass of the (0,1) lepton pair (see the text for the definition of the lepton ordering). Middle right: the invariant mass of the (0,2) lepton pair.

637 same-flavor pair of leptons within  $10$  GeV/ $c^2$  of the Z mass, and either 4 jets and at least 1 b-tagged  
 638 jet or exactly 3 jet and 2 or more b-tagged jets. The resulting region has low statistics and is not  
 639 used as a control region but is instead used as a validation to demonstrate that the normalization  
 640 uncertainty, discussed above, is properly evaluated.

641 The region defined by this is predicted to be 67%  $t\bar{t}Z$ , 17%  $WZ$ , and 13%  $tZ$ . We predict  
 642  $19.3 \pm 0.5$  events and observe 28, giving a observed-to-predicted ratio of  $1.45 \pm 0.27 \pm 0.03$  (where the  
 643 errors are from data and simulation statistics, respectively). Given the large errors, the region is still  
 644 in agreement with the predictions to within  $1$ - $1.5$   $\sigma$ . Distributions of various variables are shown in  
 645 Fig. 8.1.

646 **8.2 Di-boson Background Estimation:  $W^\pm Z$  , ZZ**

647 **8.3 Fake Lepton Backgrounds:  $t\bar{t}$**

648 **8.4 Charge-Misidentification Background**

649

## CHAPTER 9

650

# Conclusions

651 **9.1 Higgs Results in Review**

652 **9.2 Prospects for Future**

---

## Bibliography

---