

# 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 . . . . .	31
116	8.2 Di-boson Background Estimation: $W^\pm Z, ZZ$ . . . . .	31
117	8.2.1 $W^\pm Z$ Uncertainty . . . . .	32
118	8.2.2 $ZZ$ Uncertainty . . . . .	34
119	8.3 Charge-Misidentification Background . . . . .	36
120	8.3.1 Likelihood Method . . . . .	37
121	8.3.2 Results . . . . .	39
122	8.3.3 Systematic and Statistical Uncertainties . . . . .	39
123	8.4 Fake Lepton Backgrounds . . . . .	41
124	8.4.1 2l SS Fakes . . . . .	42
125	8.4.2 3l Fakes . . . . .	46
126	8.4.3 4l Fakes . . . . .	50
127	<b>9 Systematic Uncertainties and Statistical Formalism</b>	<b>52</b>
128	9.1 Systematic Uncertainties on Signal Cross-section and Acceptance . . . . .	52
129	9.2 Summary of Background Systematic Uncertainties . . . . .	54
130	9.3 Summary of Experimental and Detector Systematic Uncertainties . . . . .	54
131	<b>10 Conclusions</b>	<b>55</b>
132	10.1 Higgs Results in Review . . . . .	55
133	10.2 Prospects for Future . . . . .	55
134	<b>Bibliography</b>	<b>56</b>

---

## Preface

---

136 This is the preface. Blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah  
137 blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah blah blah  
138 blah. Blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah  
139 blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah blah  
140 blah blah blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah blah blah.  
141 Blah blah blah blah blah blah. Blah blah blah blah blah blah.

142 Blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah blah  
143 blah. Blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah blah blah  
144 blah blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah  
145 blah blah blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah blah blah  
146 blah. Blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah blah  
147 blah blah blah.

Chris Lester  
CERN, Fall 2014

149

## CHAPTER 1

---

150

# Introduction

---

151

## CHAPTER 2

152

# Theoretical Background

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

161 **2.1 The Standard Model**

162 **2.1.1 The Standard Model Structure**

163 The Standard Model (SM) [1, 2, 3, 4] is an example of a quantum field theory that describes the  
164 interactions of all of the known fundamental particles. Particles are understood to be excitations of  
165 the more fundamental object of the theory, the field. The dynamics and interactions of the fields are  
166 derived from the Standard Model Lagrangian, which is constructed to be symmetric under transfor-  
167 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  
168 weak iso spin, and  $U(1)$  is the group for weak hyper-charge.

169 Gauging the symmetries demands the introduction of 8 massless gluons, or the boson (full integer  
170 spin) carriers of the strong force [5] from the generators  $SU(3)$  symmetry, and the 4 massless bosons,  
171 carriers for the weak and electromagnetic forces from the 3 generators of the  $SU(2)$  and 1 generator  
172 of the  $U(1)$  group. The weak and the electromagnetic forces are considered part of a larger single

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

174 The gauge symmetry allows the theory to be re-normalizable [6], meaning that unwanted infinities  
 175 can be absorbed into observables from theory in a way that allows the theory to be able to predict  
 176 physics at multiple energy scales. Singlets of the  $SU(3)$  group are fermions (half-integer spin particles)  
 177 called leptons and do not interact with the strong, whereas doublets of the  $SU(3)$  group are called  
 178 quarks and do interact with the strong force. The SM is a chiral theory: the weak force violates parity,  
 179 as it only couples to left-chiral particles or right-chiral antiparticles. This means that right-chiral and  
 180 left-chiral fermions arise from different fields, which are different representations of the weak isospin  
 181 group.

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

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

### 187 2.1.2 Electroweak Symmetry Breaking and the Higgs

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

194 To solve these problems, spontaneous electro-weak symmetry breaking (EWSB) is introduced via  
 195 the Brout-Englert-Higgs mechanism [7, 8, 9]. A massive scalar field in an electro-weak doublet is  
 196 added to the theory with 4 new degrees of freedom and a potential which includes a quartic self-  
 197 interaction term. Each fermion field interacts with the scalar field via a different Yukawa coupling,  
 198 which unites the left and right chiral fields of a single particle type. This field explicitly preserves all  
 199 of the symmetries, but the minimum of the potential does not occur when the expectation of the field  
 200 is zero. The field eventually falls to a state, where it acquires a vacuum-expectation value. However,  
 201 a non-zero field must therefore point in a particular direction of weak-isospin space, breaking the  
 202 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).

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

### 210      2.1.3    The Standard Model Parameters

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

216      Experiments have provided a number of measurements of the parameters of the SM[10]. Prior  
 217      to the discovery of the Higgs boson (discussed later) the Higgs mass, however, was the only fully  
 218      unconstrained parameter, although its value could be inferred via its involvement in loop corrections  
 219      on the top mass ( $M_t$ ) and the W mass ( $M_W$ ). The GFitter collaboration assembles all relevant  
 220      electroweak observable measurements into a statistical model and then allows certain measurements to  
 221      float within their uncertainty to allow for a fit among multiple correlated measurements measurements  
 222      [11]. Figure 2.1 shows the fitted constraints on 4 key SM parameters ( $M_H$ ,  $M_W$ ,  $M_t$ ,  $\sin^2\theta_w$ ) with  
 223      actual measurements overlaid. The addition to the fit of the measured Higgs mass from the ATLAS  
 224      and CMS collaborations creates a small tension, as the other observables prefer the mass to be much  
 225      lower ( $\sim 80$  GeV/ $c^2$ ). The tension in the combined electroweak fit (including the Higgs) is not  
 226      statistically significant with a  $p$ -value of 0.07.

## 227      2.2    Collider Physics and the Higgs

228      To test the theory, physicists accelerate particles to extremely high energies and force them to interact  
 229      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

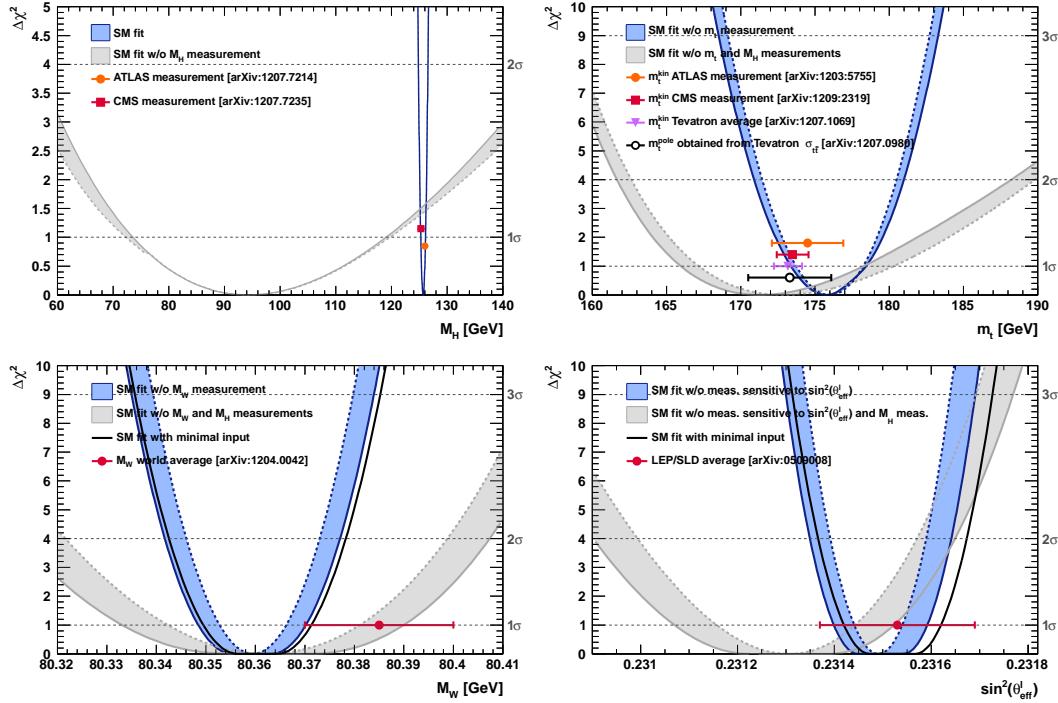


Figure 2.1:  $\chi^2$  as a function of the Higgs mass (top left), the top quark mass (top right), the  $W$  boson mass (bottom left) and the effective weak mixing angle (bottom right) for the combined SM fit from the GFitter group. The data points placed along  $\chi^2 = 1$  represent direct measurements 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.

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

238 Collider physics rely on form-factor descriptions of the colliding hadrons that describe the fraction  
 239 of momentum carried by the hadrons constituent ‘partons’. These are called parton-distribution  
 240 functions, seen in Figure 2.2, and are factorized and integrated through the theoretical calculations  
 241 of various collision processes [12].

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

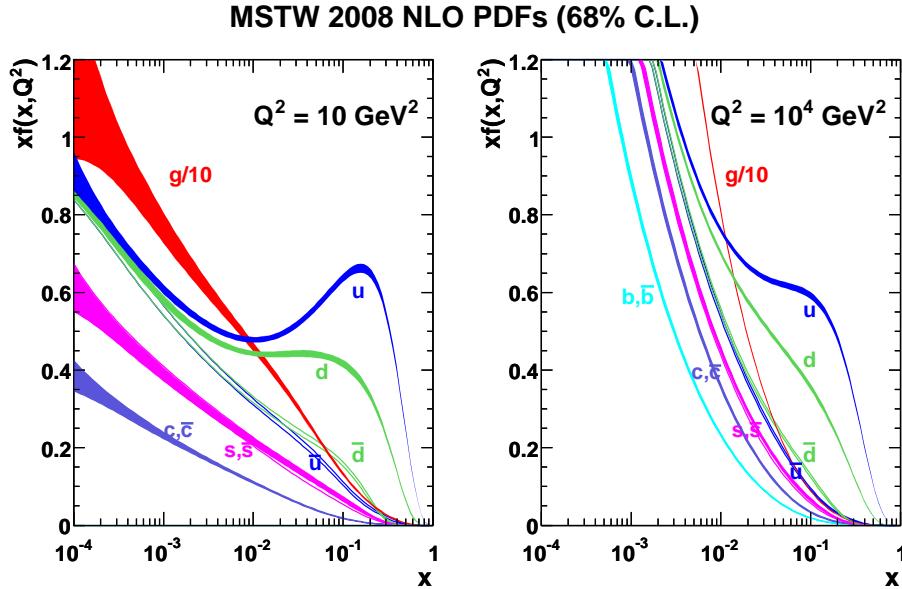


Figure 2.2: Proton Parton Distribution Functions (PDFs) form the MSTW Collaboration at  $Q^2 = 10 \text{ GeV}^2$  and  $Q^2 = 10^4 \text{ GeV}^2$

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[13]. 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 [10].

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 [14]. 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.

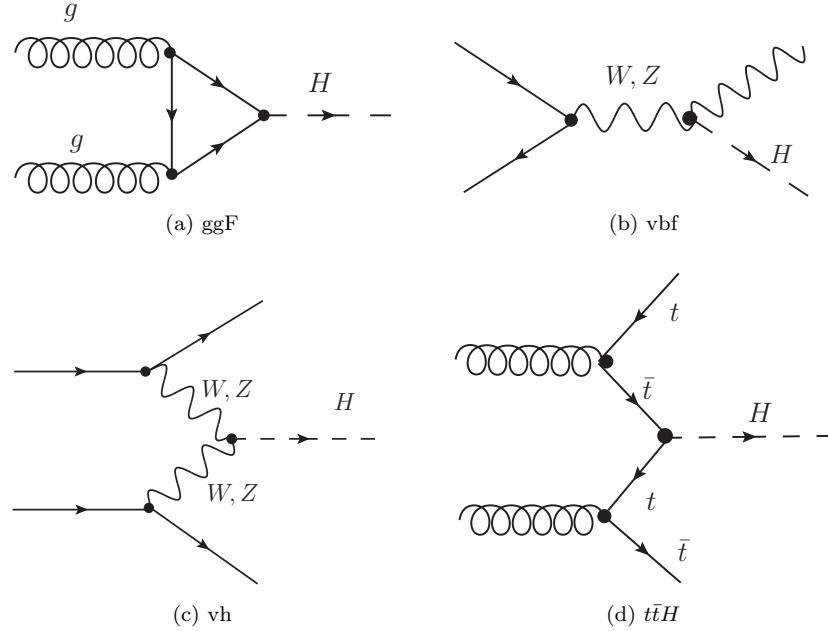


Figure 2.3: Dominant Higgs production modes at the LHC

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 [15, 16]. I will focus on the ATLAS results in the following. The ATLAS measurements constrain both the Higgs mass[17] and spin[18], 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-

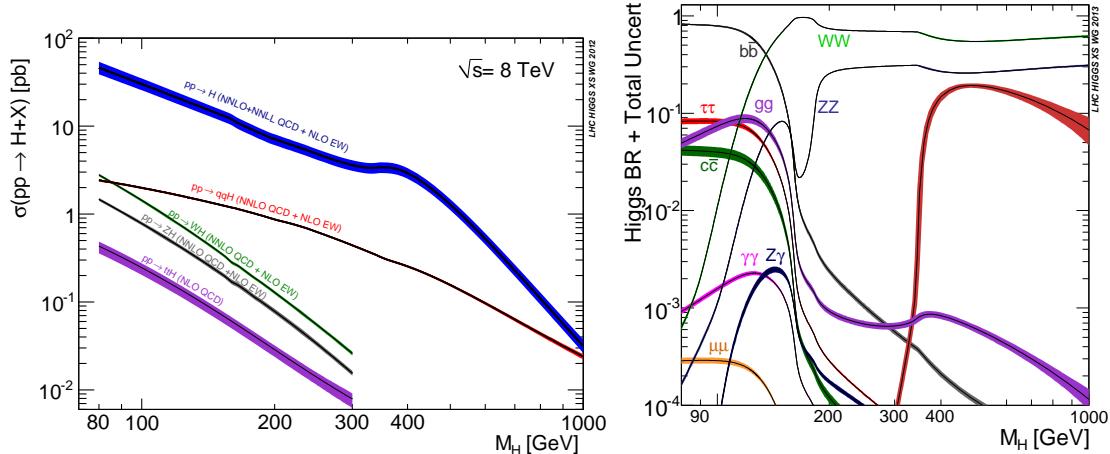


Figure 2.4: 8 TeV LHC Higgs production cross-sections (left) and decay branching fractions

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[19]. 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[20]. Other exotic scenarios allow for new colored particles in the gluon-fusion loop, which are not entirely constrained by present measurements[21, 22, 23]. These include, for instance, Supersymmetric models involving the stop quark.

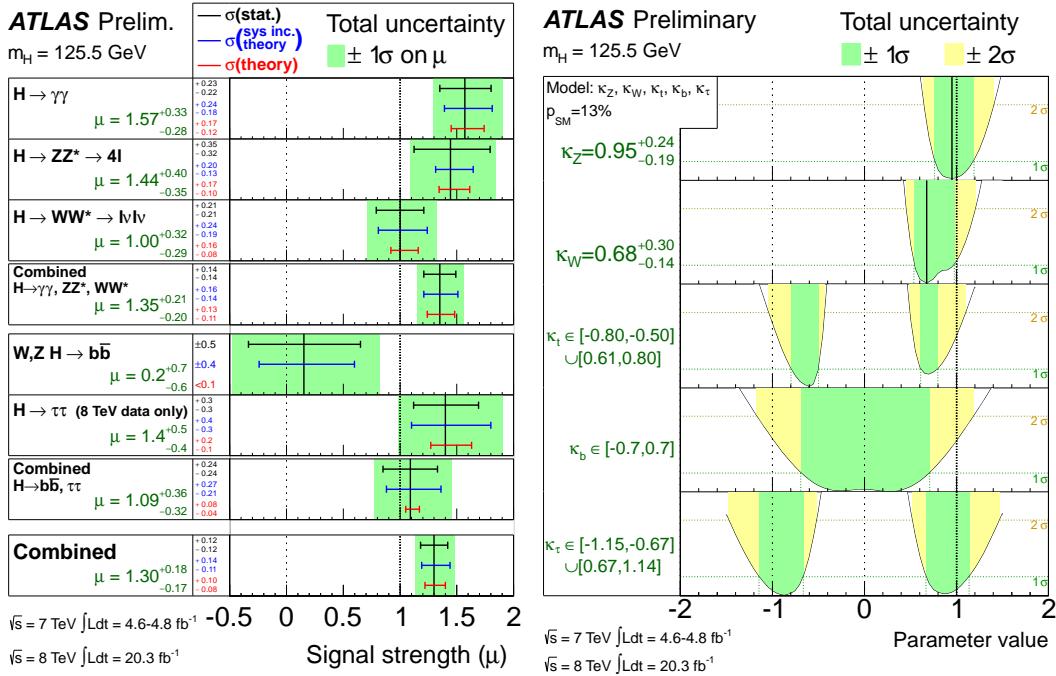


Figure 2.5: ATLAS Higgs combination results for all SM measurement channels as ratios of the measured to SM production cross-sections (left) and extracted Higgs coupling constraint scale-factors for a combined fit to the measurement channels, where the  $W, Z$ , top-quark,  $b$ -quark, and  $\tau$  couplings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

294     Aside from the loop effects, measurement of the  $t\bar{t}H$  production cross-section would provide a  
295     precise measurement of the top Yukawa coupling. When compared with the measured top quark mass,  
296     this tests the Higgs mass generation properties for up-type quarks. Despite similar uncertainties on  
297     the overall production cross-sections for  $t\bar{t}H$  and the gluon-fusion modes, both of which depend on the  
298     top Yukawa, most of these uncertainties would cancel for  $t\bar{t}H$  if normalized to the topologically similar  
299      $t\bar{t}Z$ . Finally, the uniqueness of the experimental signature means that searches for  $t\bar{t}$  signatures can  
300     be performed for a variety of Higgs decays ( $\gamma\gamma, b\bar{b}, WW, ZZ$ , and  $\tau\bar{\tau}$  with roughly similar degrees of  
301     sensitivity (within a factor of 10)[19].

302     It is important to note the importance of the top Yukawa coupling to the overall structure of the  
303     SM, due to its enormous size compared to other couplings. The top Yukawa is for instance 350000x as  
304     large as the electron Yukawa coupling. Because of this the top Yukawa coupling, along with the Higgs  
305     mass, is one of the most important pieces of the renormalization group equations (RGE) responsible  
306     for the running of the parameter that determines the Higgs self-coupling  $\lambda$ . If this parameter runs  
307     negative, then the potential responsible for the entire mechanism of EWSB no longer has a minimum  
308     and becomes unbounded, resulting in instability in the universe [24]. Metastability occurs when the

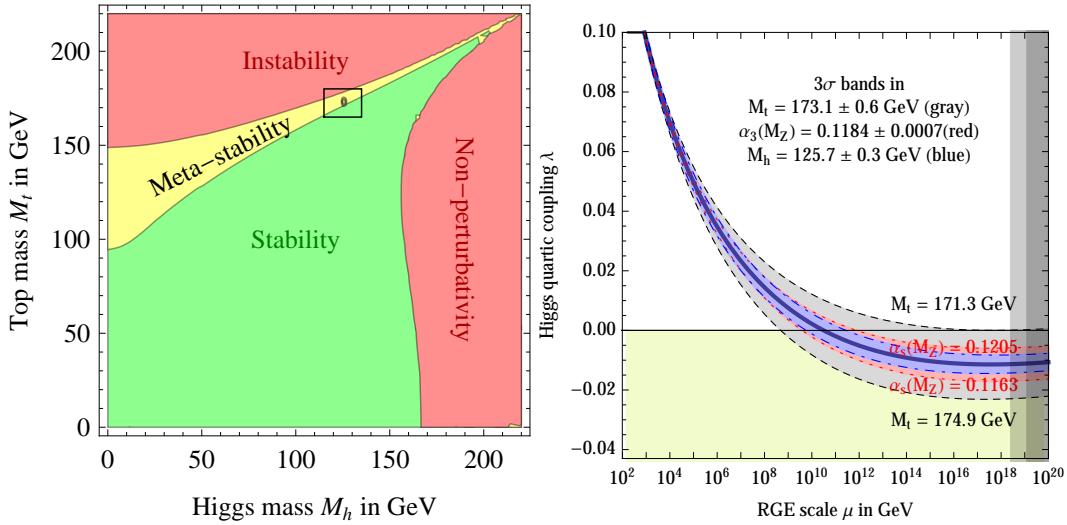


Figure 2.6: RGE for the running of the SM parameter,  $\lambda$  for the Higgs self-coupling term with present values and uncertainty bands for  $M_H$  and  $M_t$  (left). The two-dimensional plot colored (right) shows regions for which the SM is stable, unstable and metastable based on this RGE

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

325

## CHAPTER 3

326

# The Large Hadron Collider and the ATLAS Experiment

327

### 328 3.1 The Large Hadron Collider

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

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

334 The technology involved in the development of the LHC and very briefly touched upon in this  
335 chapter is an enormous achievement in its own right and is documented in detail here [25, 26, 27].

336 The LHC is a circular machine capable of accelerating beams of protons and colliding them at  
337 center of mass energies up to  $\sqrt{s} = 14\text{TeV}$  at 4 collision sites around the ring, where 4 experiments are  
338 housed (ATLAS[28], CMS[29], LHCb[30], and ALICE[31]). Figure 3.1 is a diagram of the layout of the  
339 LHC and its experiments[32]. The LHC also operates in modes with beams of heavy ions. The LHC  
340 is composed of thousands of super-conducting Niobium-Titanium magnets, cooled to  $2.7^\circ \text{ C}$  with  
341 liquid Helium, which steer and focus the particle beams, and a superconducting resonant-frequency  
342 (RF) cavity, which boosts the beam to higher energies.

#### 343 3.1.1 The Accelerator Complex

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

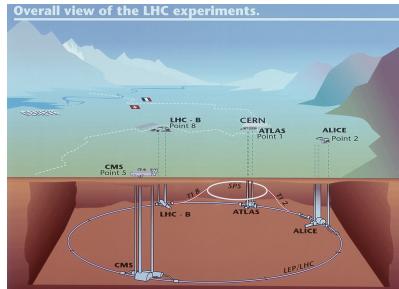


Figure 3.1: Diagram of the Large Hadron collider and location of the 4 main experiments (ATLAS, CMS, LHCb, and ALICE). Around the ring. The diagram also shows the location of the SPS, the final booster ring the accelerator complex that accelerates the protons to 450 GeV before injection into the LHC.

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

### 350 3.1.2 Beam Parameters and Collisions

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

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

364 s<sup>-1</sup>

## CHAPTER 4

# Electrons

367 **4.1 Electrons at Hadron Colliders**

368 **4.2 Reconstruction of Electron at ATLAS**

369 **4.3 Identification of Electrons at ATLAS**

370 **4.3.1 Pile-up**

371 **4.3.2 Trigger vs. Offline**

372 **4.3.3 2011 Menu**

373 **4.3.4 2012 Menu**

374 **4.3.5 Electron Likelihood**

375 **4.4 Measurement of Electron Efficiency at ATLAS**

376 **4.4.1 Techniques**

377 **4.4.2 Issues**

## CHAPTER 5

---

# Search for the TTH Decay in the Multilepton Channel

---

381 This chapter provides an overview of the set of analyses searching for the Standard Model (SM)  
 382 production of the Higgs boson in association with top quarks in multi-lepton final states with multiple  
 383 jets (including b-quark tagged jets). Searches in  $t\bar{t}H$  final states with 2 same-charge, 3 and 4 light  
 384 leptons ( $e, \mu$ ) are discussed in depth. These final states target specifically Higgs decays to vector  
 385 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  
 386 final states targeting the  $H \rightarrow b\bar{b}$  [33],  $H \rightarrow \gamma\gamma$ [34], and  $H \rightarrow \tau\tau$  decay modes.

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

392 Detailed description of the event and selection section are provided in Chapter 7, background  
 393 modelling in Chapter 8, the effect of systematic errors and the statistical analysis in Chapter 9 and  
 394 final results in Chapter ??.

395 **5.1 Signal Characteristics**

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

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

400 energy or hadronically. Table 5.1 provides the fractional contribution of the main Higgs decay modes  
 401 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%

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

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

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

## 414 5.2 Background Overview

415 Background processes can be sorted into two categories:

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

---

<sup>5</sup>Charge 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

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

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

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

429       **5.3 Analysis Strategy**

430       ADD SOMETHING HERE FOR HOW TO CALCULATE A CROSS-SECTION

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

438       The channels are fed into a posson model

## CHAPTER 6

---

# Dataset and Simulation

---

## 441 6.1 Data

### 442 6.1.1 The 2012 Dataset

443 The LHC successfully produced datasets for physics studies in 2010, 2011 and 2012. The 2012 proton-  
 444 proton dataset was delivered with collisions with a CME of 8 TeV with bunch collisions every 50 ns and  
 445 reached a total integrated luminosity of around  $20 \text{ fb}^{-1}$  [35]. Figure 6.1 shows the accumulation of this  
 446 dataset over time. Despite doubling the bunch spacing (thereby halving the bunch collision frequency),  
 447 the lumonisty neared the design lumonisty due to unexpected improvements in the transverse beam  
 448 profile[36]. This increased the amout of pile-up, or number of collisions per bunch crossing and in  
 449 general collision events were busier due to these multiple interactions. Figure 6.2 shows the average  
 450 number of interaction per bunch crossing for the 2011 and 2012 datasets. The 2012 dataset shows an  
 451 average of 20-25 interactions.

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

455 The datasets used in the analysis were collected with the primary electron (EF\_e24vhi\_medium1  
 456 — EF\_e60\_medium1) and muon triggers ( EF\_24i.tight — EF\_36.tight). The electron triggers  
 457 require a electron with at least 25 GeV of calorimeter energy, passing the medium identification  
 458 requirement and loose tracking isolation. Above 60GeV, the isolation requirement is dropped and the  
 459 identification is loosened slightly. Similarly, the muon trigger requires a good inner detector track and  
 460 matching hits in the muon spectrometer, as well as loose tracking isolation, which also is dropped  
 461 about 36 GeVThe data sample must contain either a primary muon or primary electron trigger.

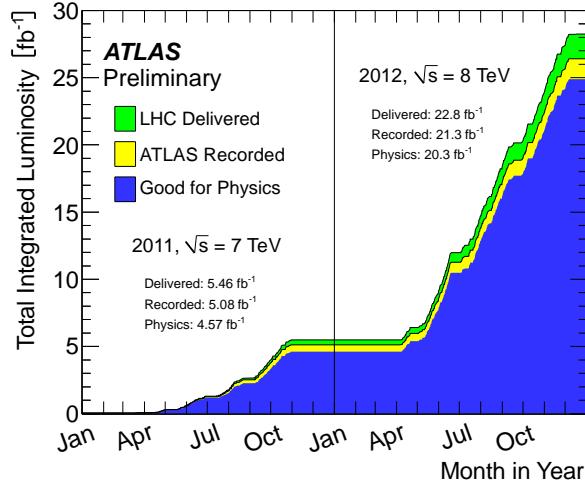


Figure 6.1: Plot showing the accumulation of the integrated luminosity delivered to the ATLAS experiment over 2011 and 2012. The rough size of the usable, physics ready dataset for 2012 is  $20 \text{ fb}^{-1}$  and is the dataset used for the following analysis.

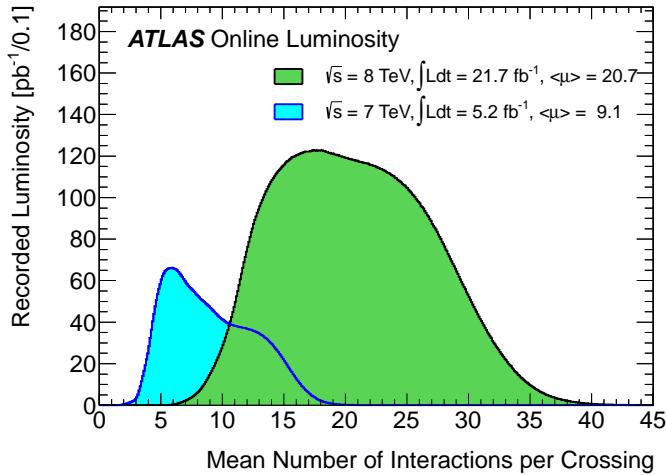


Figure 6.2: The average number of interactions per bunch-crossing for the 2012 and 2011 LHC proton-proton dataset. Most of these interactions are uninteresting but leave energetic signatures in particle detectors called pile-up which interfere with measurements

## 462 6.2 Simulation

463 Simulation samples based on are used to determine the overall event selection acceptance and efficiency  
 464 and for investigations not directly involved in the final result. The simulated samples are created using  
 465 parton distribution function (PDF) and model using Monte Carlo (MC) techniques the hard parton  
 466 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

467 passed through a full ATLAS detector simulation[37] based on GEANT4 [38]. Small corrections are  
468 then applied to the overall efficiencies to re-scale object identification efficiencies, energy scales, and  
469 the pile-up, discussed in depth later.

### 470 6.2.1 Signal Simulation

471 The  $t\bar{t}H$  production is modelled using matrix elements obtained from the HELAC-Oneloop package [?]  
472 that corresponds to the next-to-leading order (NLO) QCD accuracy. POWHEG BOX [?, ?, ?] serves  
473 as an interface to the parton shower Monte Carlo programs. The samples created using this approach  
474 are referred to as PowHel samples. CT10NLO PDF sets are used and the factorisation ( $\mu_F$ ) and  
475 renormalization ( $\mu_R$ ) scales are set to  $\mu_0 = \mu_F = \mu_R = m_t + m_H/2$ . Pile-up and the underlying  
476 events are simulated by Pythia 8.1 [?] with the CTEQ61L set of parton distribution functions and  
477 AU2 underlying event tune. The Higgs boson mass is set to 125 GeV and the Top quark mass is set  
478 to 172.5 GeV.

479 The signal Monte Carlo samples are summarized in Table 6.1. These large samples are generated with  
480 inclusive Higgs boson decays, with branching fractions set to the LHC Higgs Cross Section Working  
481 Group (Yellow Report) recommendation for  $m_H = 125$  GeV [39]. The inclusive cross section (129.3  
482 fb at  $m_H = 125$  GeV) is also obtained from the Yellow Report [39].

### 483 6.2.2 Background Simulation

484 The background simulations used for this analysis are listed in Table 6.2. In general, the Alpgen[40],  
485 MadGraph[41], and AcerMC[42] samples use the CTEQ6L1[43] parton distribution function, while  
486 the Powheg[44], Sherpa[45], are generated with the CT10 PDF. The exception is the MadGraph  $t\bar{t}t\bar{t}$   
487 sample, which is generated with the MSTW2008 PDF[46]. The highest order calculations available  
488 are used 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
$t\bar{t}t\bar{t}$	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

489           

## CHAPTER 7

490           

---

# Object and Event Selection

---

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

495       The cuts for each signal region are provided in Table 7.1 and the object selections are detailed in  
496       the following selections. The selections are based on optimizations of the region sensitivity performed  
497       using MC (event for data driven backgrounds) and adhoc values for normalization systematic errors.  
498       All signal regions are comprised of three basic requirements: the presence of b-tagged jets, the presence  
499       of additional light jets, and a veto of same flavor opposite sign leptons with an invariant mass within  
500       the Z window. Additional requirements on the invariant mass of the leptons, the missing transverse  
501       energy in the event, and the total object energy ( $H_T$ ) proved to have negligible additional benefit at  
502       our level of statistics.

503       

### 7.1 2l Same-Charge Signal Region

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

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

511       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)

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

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

## 519 7.2 3l Signal Region

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

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

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

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

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

### 535      7.3 4l Signal Region

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

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

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

### 554      7.4 Electron Selection

555        The electrons are reconstructed by a standard algorithm of the experiment [47] and the electron  
 556        cluster is required to be fiducial to the barrel or endcap calorimeters:  $|\eta_{\text{cluster}}| < 2.47$ . Electrons in  
 557        the transition region,  $1.37 < |\eta_{\text{cluster}}| < 1.52$ , are vetoed. Electron reconstruction and identification  
 558        is discussed in depth in Chapter 4. Electrons must pass the the VERYTIGHT likelihood identification  
 559        criteria.

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

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

572 The electron selection is provided in Table 7.2.

## 573 7.5 Muon Selection

574 Muons used in the analysis are formed by matching reconstructed inner detector tracks with either  
 575 a complete track or a track-segment reconstructed in the muon spectrometer (MS). The muons must  
 576 satisfy  $|\eta| < 2.5$ . The muon track are required to be a good quality combined fit of inner detector  
 577 hits and muon spectrometer segments, unless the muon is not fiducial to the inner detector,  $|\eta| > 2.4$ .  
 578 Muons with inner detector tracks are further required to pass standard inner detector track hit  
 579 requirements [48].

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

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

586 The muon selection is provided in Table 7.2.

## 587 7.6 Jet and b-Tagged Jet Selection

588 Jets are reconstructed in the calorimeter using the anti- $k_t$  [49] algorithm with a distance parameter  
 589 of 0.4 using locally calibrated topologically clusters as input (LC Jets).

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

594         $p_T$  and  $\eta$  cuts are tuned based on the sensitivity to  $t\bar{t}H$ .

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

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

602        **7.7 Tau Selection**

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

605        **7.8 Object Summary and Overlap**

606        **7.9 Optimization**

Parameter	Values	Remarks
Electrons		
$p_T$	$> 10 \text{ GeV}$	
$ \eta $	$< 2.47$ veto crack	
ID	Very Tight Likelihood	$< 1.37$ for 2l SS channel
Isolation	$E_{\text{T}}\text{Cone}20/p_T, p_T\text{Cone}20/p_T < 0.05$	
Jet overlap removal	$\Delta R > 0.3$	
$ d_0^{\text{sig}} $	$< 4\sigma$	
$z_0 \sin\theta$	$< 1 \text{ cm}$	
Muons		
$p_T$	$> 10 \text{ GeV}$	
$ \eta $	$< 2.5$	
ID	Tight	
Jet overlap removal	$\Delta R > 0.04 + 10 \text{ GeV}/p_T$	
Isolation	$E_{\text{T}}\text{Cone}20/p_T, p_T\text{Cone}20/p_T < 0.1$	$< 0.05$ for 2 leptons
$ d_0^{\text{sig}} $	$< 3\sigma$	
$z_0$	$< 1 \text{ cm}$	
Taus		
$p_T$	$> 25 \text{ GeV}$	
ID	Medium BDT	
Isolation		
Jet overlap removal	$\Delta R > 0.3$	
$e/\mu$ vetoes	Medium electron veto	
Jets		
$p_T$	$> 25 \text{ GeV}$	
$ \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

---

609 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  
 610 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)  
 611 and  $t\bar{t}$  production with a jet misidentified as a lepton (fakes). The 2l SS regions possess a unique  
 612 background of charge misidentification from Z and top events. The methods for estimating these  
 613 backgrounds are discussed in this chapter. Monte Carlo simulation is used for the prompt  $t\bar{t}V$  and  
 614 VV contributions. Systematic uncertainties on the overall normalization of these backgrounds in the  
 615 signal region are provided from theoretical studies and past ATLAS analyses and are verified in  
 616 data-based validation regions. The non-prompt backgrounds from  $t\bar{t}$ jet-misidentification and charge-  
 617 misidentification are estimated using data-driven methods.

620 For reference, Table 8.1 provides a summary of the  $t\bar{t}H$  signal and background expectation for  
 621 each of the signal regions, including the data-driven estimates discussed in this section. For each  
 622 region, the background contribution exceeds the size of the signal.

623 **8.1 Vector Boson ( $W^\pm, Z$ ) production in association with top quarks:**

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

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

Table 8.1: Expected number of signal and background events in 2l SS, 3l and 4l signal regions. For data-driven backgrounds, monte-carlo only numbers are given for reference. The expected sensitivity  $\frac{s}{\sqrt{s+b}}$  is provided for data-driven, mc only and for the inclusion of ad-hoc systematic uncertainties (20% for  $t\bar{t}V$  and 30% for  $t\bar{t}\text{fake}$ ).

	Same-sign				4 leptons				
	$\geq 5$ jets		4 jets		3 leptons		Z depleted		
	$e^\pm e^\pm$	$e^\pm \mu^\pm$	$\mu^\pm \mu^\pm$	$e^\pm e^\pm$	$e^\pm \mu^\pm$	$\mu^\pm \mu^\pm$	Z enriched	Z depleted	
$tH$	0.73 $\pm$ 0.03	2.13 $\pm$ 0.05	1.41 $\pm$ 0.04	0.44 $\pm$ 0.02	1.16 $\pm$ 0.03	0.74 $\pm$ 0.03	0.19 $\pm$ 0.01	0.03 $\pm$ 0.00	
$t\bar{t}V$	2.60 $\pm$ 0.13	7.42 $\pm$ 0.17	5.01 $\pm$ 0.16	3.05 $\pm$ 0.13	8.39 $\pm$ 0.24	5.79 $\pm$ 0.20	7.21 $\pm$ 0.24	0.74 $\pm$ 0.05	
$tZ$							0.71 $\pm$ 0.03	incl. in $t\bar{t}V$	
$VV$	0.48 $\pm$ 0.25	0.37 $\pm$ 0.23	0.68 $\pm$ 0.30	0.77 $\pm$ 0.27	1.93 $\pm$ 0.80	0.54 $\pm$ 0.30	0.89 $\pm$ 0.25	0.08 $\pm$ 0.01	
$t^{\pm}, tX$ (MC)	1.31 $\pm$ 0.67	2.55 $\pm$ 0.84	1.76 $\pm$ 0.67	4.99 $\pm$ 1.19	8.19 $\pm$ 1.41	3.70 $\pm$ 1.03	2.46 $\pm$ 0.19	0.00 $\pm$ 0.00	
$Z + \text{jets}$ (MC)	0.16 $\pm$ 0.16	0.28 $\pm$ 0.20	0.12 $\pm$ 0.12	1.37 $\pm$ 0.78	0	0.23 $\pm$ 0.23	0	0.00 $\pm$ 0.00	
fake leptons (DD)	2.31 $\pm$ 0.97	3.87 $\pm$ 1.01	1.24 $\pm$ 0.41	3.43 $\pm$ 1.38	6.82 $\pm$ 1.63	2.38 $\pm$ 0.78	2.62 $\pm$ 0.51	$(1.1 \pm 0.6) \cdot 10^{-3}$	
Q misid (DD)	1.10 $\pm$ 0.09	0.85 $\pm$ 0.08	—	1.82 $\pm$ 0.11	1.39 $\pm$ 0.08	—	—	$(0.09 \pm 0.03) \cdot 10^{-3}$	
Tot Background (fake MC)	4.56 $\pm$ 1.17	10.62 $\pm$ 1.54	7.57 $\pm$ 1.31	10.18 $\pm$ 2.43	18.51 $\pm$ 2.54	10.26 $\pm$ 1.82	11.27 $\pm$ 0.40	0.83 $\pm$ 0.07	
Tot Background (fake DD)	6.49 $\pm$ 1.04	12.51 $\pm$ 1.04	6.93 $\pm$ 0.52	9.07 $\pm$ 1.42	18.53 $\pm$ 1.83	8.71 $\pm$ 0.88	11.43 $\pm$ 0.62	0.831 $\pm$ 0.075	
$s/\sqrt{b}$ (fake MC)	0.34	0.65	0.51	0.14	0.27	0.23	0.70	0.21	
$s/\sqrt{b}$ (fake DD)	0.29	0.60	0.54	0.15	0.27	0.25	0.69	0.21	
$s/\sqrt{b} \oplus 0.3\text{fake}(MC) \oplus 0.2\text{t}tV$	0.33	0.58	0.47	0.12	0.22	0.21	0.63	0.207	
$s/\sqrt{b} \oplus 0.3\text{fake}(DD) \oplus 0.2\text{t}tV$	0.27	0.53	0.50	0.14	0.23	0.23	0.62	0.207	

Table 8.2: NLO cross section and theoretical uncertainty calculations derived from MadGraph5\_aMC@NLO.

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
$tZ$	160.0	+4	-4	+7	-7	8.0
$tZ$	76.0	+5	-4	+7	-7	8.6

( $\mu_F$ ) 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 [50]: 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 [51]: calculate the PDF uncertainty using the MSTW2008nlo [46] 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 8.2.

The  $tZ$  process is normalized to NLO based on the calculation in Ref. [52]. 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.

---

**651 8.1.1  $t\bar{t}Z$  Validation Region**

652 Unlike  $t\bar{t}W^\pm$ , a  $t\bar{t}Z$  validation region can be obtained by simply inverting the veto on same-flavor  
 653 opposite sign lepton pairs near the Z pole in the 3 lepton signal region. This region thus requires 3  
 654 leptons (with momentum and identification cuts discussed in Chapter 7, at least one opposite sign,  
 655 same-flavor pair of leptons within  $10 \text{ GeV}/c^2$  of the Z mass, and either 4 jets and at least 1 b-tagged  
 656 jet or exactly 3 jet and 2 or more b-tagged jets. The resulting region has low statistics and is not  
 657 used as a control region but is instead used as a validation to demonstrate that the normalization  
 658 uncertainty, discussed above, is properly evaluated.

659 The region defined by this is predicted to be 67%  $t\bar{t}Z$ , 17%  $WZ$ , and 13%  $tZ$ . We predict  
 660  $19.3 \pm 0.5$  events and observe 28, giving a observed-to-predicted ratio of  $1.45 \pm 0.27 \pm 0.03$  (where the  
 661 errors are from data and simulation statistics, respectively). Given the large errors, the region is still  
 662 in agreement with the predictions to within  $1-1.5 \sigma$ . Distributions of various variables are shown in  
 663 Fig. 8.1.

**664 8.2 Di-boson Background Estimation:  $W^\pm Z, ZZ$** 

665  $W^\pm Z$  and  $ZZ$  di-boson production with additional and b-tagged jets constitute small contributions  
 666 to the 3- and 4-lepton channels respectively. In the 3-lepton case  $W^\pm Z$  comprises  $\sim 1$  event of  $\sim$   
 667 10 total background events while the  $ZZ$  contribution accounts for approximately 10% of the total  
 668 background in the 4-lepton channel. Because of the small size of these contributions, each of the  
 669 above processes can be assigned a non-aggressive uncertainty based on similar previous analyses with  
 670 ATLAS and cross-checked with data validation regions and MC truth studies. We assign an overall  
 671 50% error on both the  $W^\pm Z$  3-lepton signal region contribution and the  $ZZ$  4-lepton signal region  
 672 contribution. The details of this error assignment are discussed below.

673 Both  $W^\pm Z$  and  $ZZ$  production have been studied by ATLAS [53][54] but neither process has  
 674 been investigated thoroughly in association with multiple jets and b-quark jets. However, both  $W + b$   
 675 [55] and  $Z + b$  [56] production in 7 TeV data have been shown to agree with MC models to within  
 676 20-30%. A single  $W$  produced in association with b-tagged jets possesses a similar topology to the  
 677  $W^\pm Z + b$  process at a different energy scale and has been shown to be dominated by charm mis-tags  
 678 and b-jets from gluon splitting and multiple parton interaction. The  $W + b$  analysis unfortunately  
 679 uses Alpgen MC with Herwig PS modeling and only provides results to 1 additional jet and therefore  
 680 is not directly applicable to this  $t\bar{t}H$  analysis (where  $W^\pm Z$  is modeled using Sherpa with massive  
 681  $c$  and  $b$  quarks).  $Z + b$  production originates from slightly different diagrams than  $ZZ + b$  however

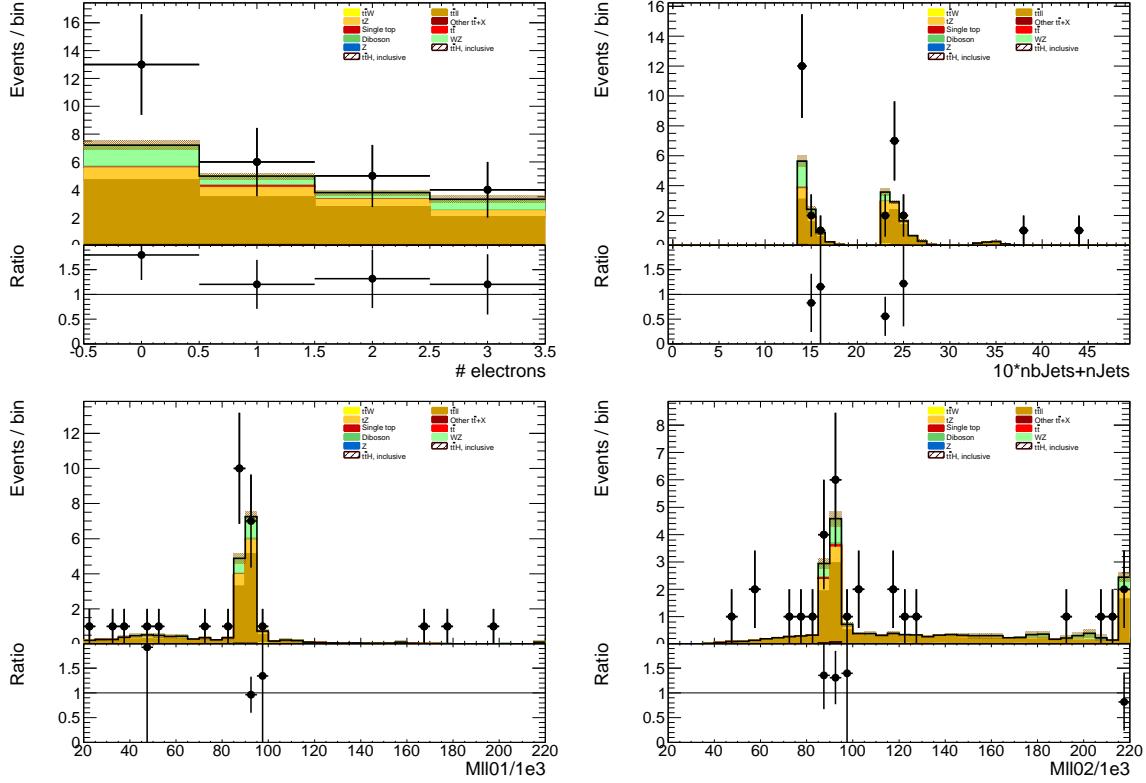


Figure 8.1: Data/MC comparison plots for  $t\bar{t}Z$  control region A ( $\geq 4$  jets,  $\geq 1$   $b$ -tag and 3 jets,  $\geq 2$   $b$ -tag). In all plots, the rightmost bin contains any overflows. Top left: number of electrons. Top right:  $10^*$ the number of  $b$ -tags + the total number of jets.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

682 the sources of the  $b$ -tags are similar and the analysis above provides results with Sherpa MC with an  
 683 agreement of  $\sim 30\%$ .

684 In the following two sections the uncertainty assignments for each of these two di-boson processes  
 685 will be reviewed in turn.

### 686 8.2.1 $W^\pm Z$ Uncertainty

687 The  $t\bar{t}H$  analyses has two validation regions to test the Sherpa agreement with data for  $W^\pm Z$ : one  
 688 inclusive 3 lepton region, using the three-lepton channel object and  $p_T$  cuts; and a  $W^\pm Z + b$  region  
 689 with 1  $b$ -tagged jet, fewer than 4 jets (to remove  $t\bar{t}V$ ), and a requirement that at least one same-flavor  
 690 opposite sign pair have an invariant mass within  $10 \text{ GeV}/c^2$  of the  $Z$  mass. Figure 8.2 shows kinematic

variables for the inclusive region <sup>8</sup>. The NJet spectrum shows good agreement within statistics across the full spectrum, giving confidence about the Sherpa high NJet SR extrapolation. Figure 8.3 shows NJet spectrum for the  $W^\pm Z + b$  validation region with good agreement in the 1 and 2 jet bins, but a slight data-MC discrepancy in the 3 jet bin. The region has low stats and around  $\sim 60\%$  purity.

We assign a conservative 50% systematic error to cover MC modeling based on these distributions and the agreement seen in similar  $W + b$  and  $Z + b$  analyses and use the MC central value for the final  $W^\pm Z$  in the SR.

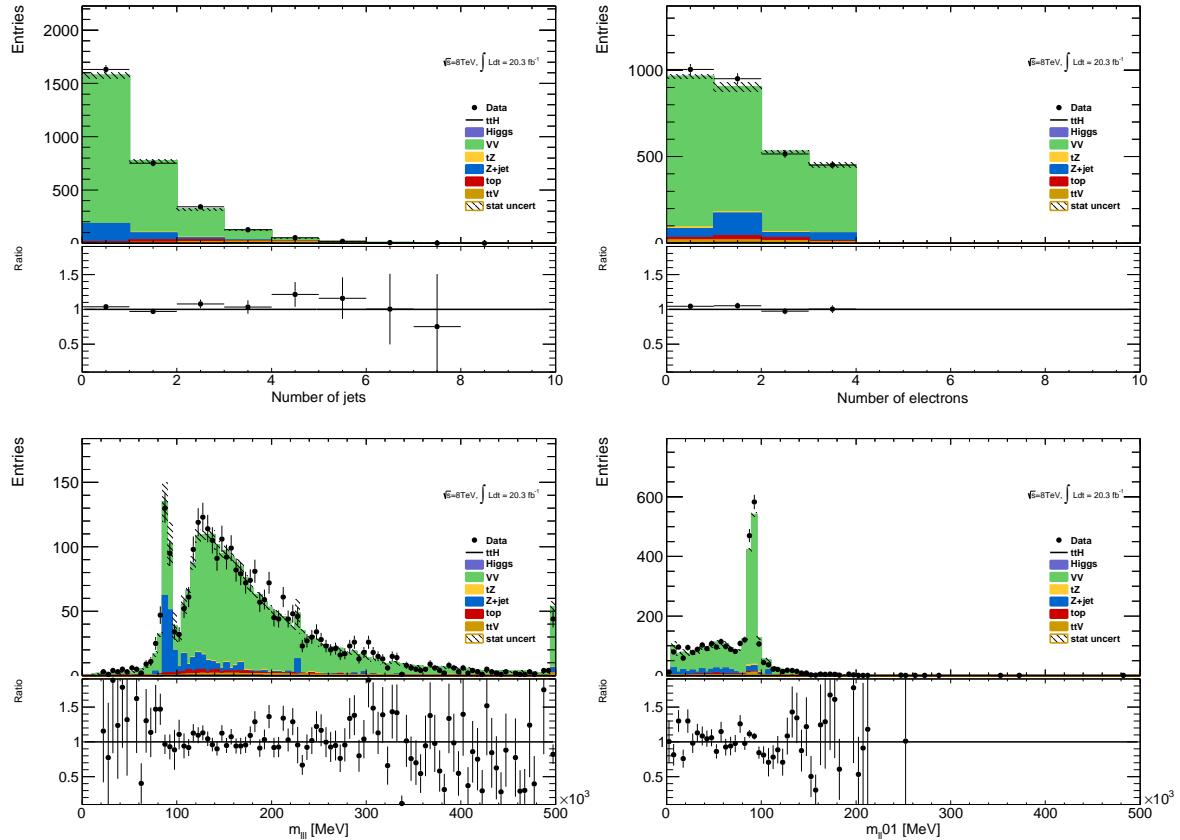


Figure 8.2: Inclusive 3 lepton  $W^\pm Z$  validation region using the  $t\bar{t}H$  lepton identification and momentum cuts: mass, number of jet and flavor variables

A cross-check is undertaken by examining the  $W^\pm Z$  truth origins of the b-jet in the  $W^\pm Z + b$  validation region (VR) and the signal region using the sherpa sample available. Table 8.3 shows these fractions. As expected the charm and b contributions dominate, though there is a small dependence on the number of jets. The composition of the VR is fairly similar to that of the signal region, especially

<sup>8</sup>the fakes are taken directly from MC

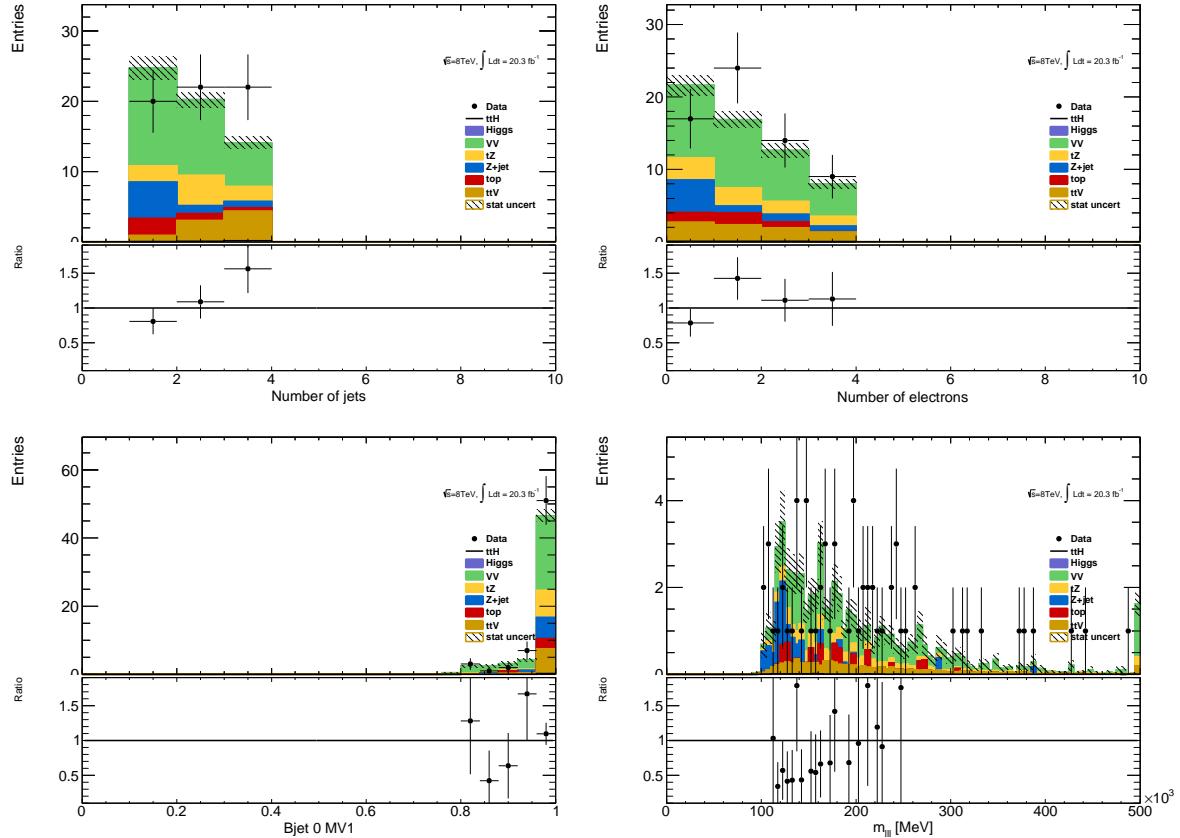


Figure 8.3:  $W^\pm Z + b$  validation region: NJet, NElec, BJet MV1 and Mass Variables

in the 3-jet bin. Importantly, also the tagged jet composition is also similar to the composition in the  $V + b$  analysis, already measured by ATLAS and discussed above.

	Bottom	Charm	Light
$W^\pm Z + b$ VR 1 Jet	$0.25 \pm 0.03$	$0.054 \pm 0.04$	$0.20 \pm 0.03$
$W^\pm Z + b$ VR 2 Jet	$0.34 \pm 0.04$	$0.052 \pm 0.06$	$0.13 \pm 0.03$
$W^\pm Z + b$ VR 3 Jet	$0.40 \pm 0.07$	$0.041 \pm 0.07$	$0.18 \pm 0.04$
$3l$ SR	$0.43 \pm 0.14$	$0.038 \pm 0.17$	$0.18 \pm 0.11$

Table 8.3: Truth Origin of highest energy b-tagged jet in the  $W^\pm Z + b$  VR and  $3l$  SR

### 8.2.2 ZZ Uncertainty

In order to investigate the MC agreement with data in the  $ZZ$  case, two validation regions similar to the  $W^\pm Z$  case are defined. Firstly, a 4 lepton  $ZZ$  region is constructed using the object selections for the 4-lepton channel and requiring exactly two pairs of opposite sign-same flavour leptons with a di-

lepton invariant mass within  $10 \text{ GeV}/c^2$  of the  $Z$  mass. Additionally, the  $ZZ + b$  process is investigated directly using a similar validation region which again requires exactly two  $Z$ -candidate lepton pairs as well as at least 1 b-tagged jet. Some kinematic distributions are shown in Figures 8.4 and 8.5, and particular attention should be paid to the NJet spectrum which shows good data-MC agreement in the high-jet bins, with a slight discrepancy in the 1-jet bin. The agreement for the region with at least 2 jets yields confidence in the NJet MC modelling in this region which lies close to the 4-lepton signal region.

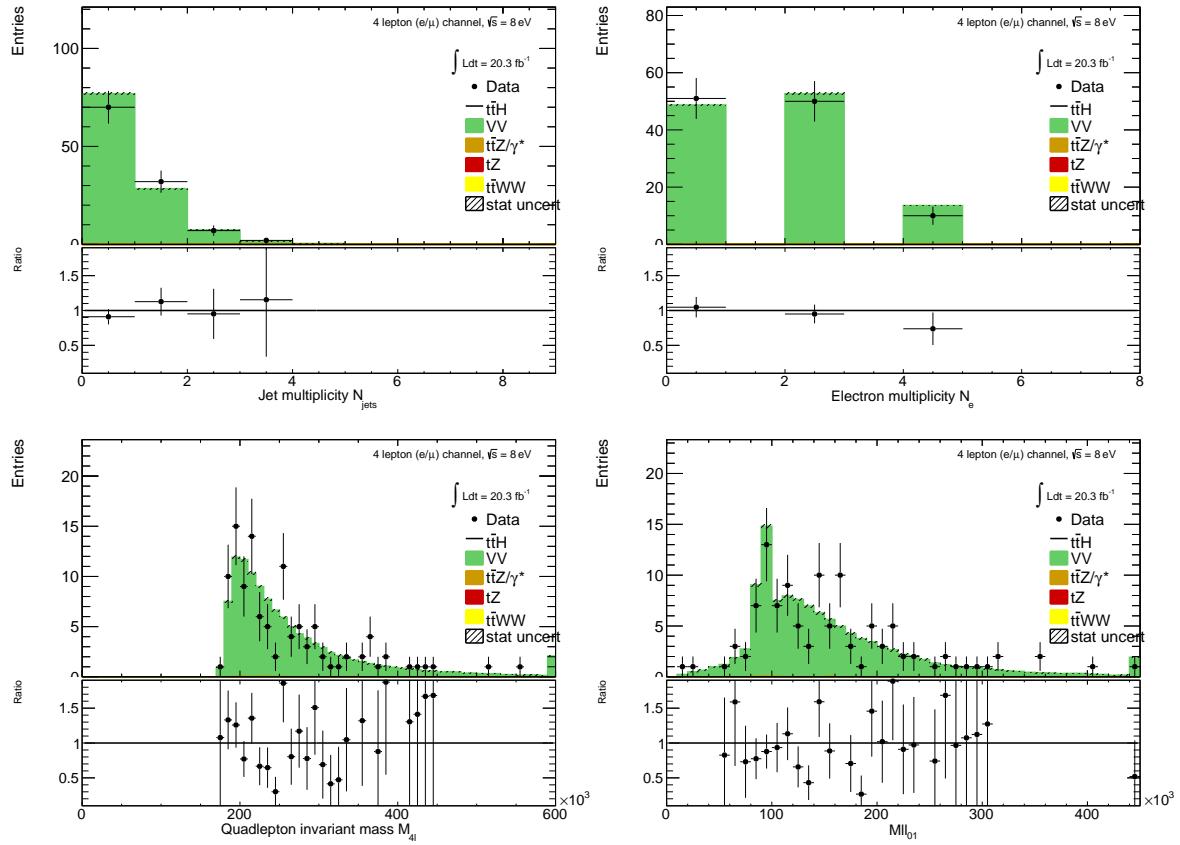


Figure 8.4: Jet-inclusive 4-lepton  $ZZ$  validation region using the  $t\bar{t}H$  lepton identification and momentum cuts

Recall that in the  $W^\pm Z$  case an overall systematic uncertainty of 50% was assigned to cover the MC modeling. Based on the study of the  $ZZ$  and  $ZZ + b$  validation regions and the overall agreement noted with the  $Z + b$  analysis, we expect a similar error to be appropriate in the  $ZZ$  case. A similar truth origin study is undertaken in MC to demonstrate a similar b-jet origin to the  $W^\pm Z$  case. The true origin of the leading (highest energy) b-tagged jet is shown in Table 8.4 for the 4-lepton signal

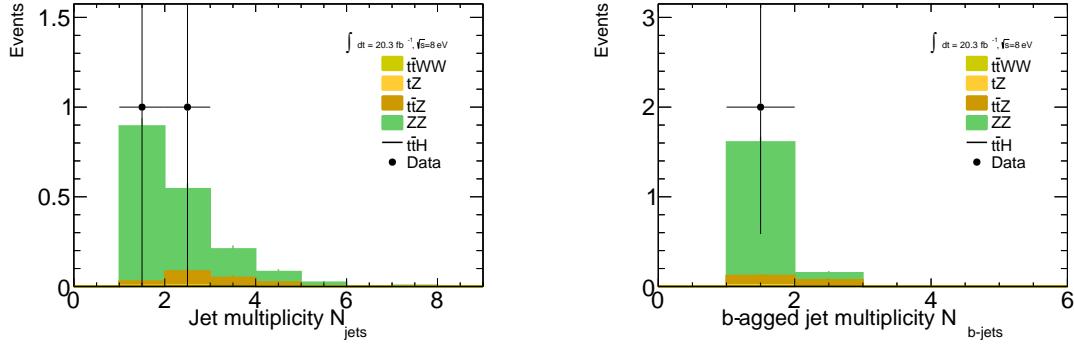


Figure 8.5:  $ZZ + b$  validation region using the  $t\bar{t}H$  lepton identification and momentum cuts

region as well as the  $ZZ + b$  validation region described above divided into jet bins. It can be seen that in case, as it was in the  $W^\pm Z$  case above, the true origin of the b-jet in  $ZZ + b$  is dominated by  $c$  and  $b$ .

	Bottom	Charm	Light
$ZZ + b$ VR 1 Jet	$0.50 \pm 0.02$	$0.21 \pm 0.01$	$0.18 \pm 0.01$
$ZZ + b$ VR 2 Jet	$0.25 \pm 0.02$	$0.12 \pm 0.01$	$0.11 \pm 0.01$
$ZZ + b$ VR 3 Jet	$0.085 \pm 0.014$	$0.040 \pm 0.011$	$0.036 \pm 0.011$
$4l$ SR	$0.020 \pm 0.008$	$0.025 \pm 0.008$	$0.014 \pm 0.005$

Table 8.4: Truth Origin of highest energy b-tagged jet in the  $ZZ + b$  VR and  $4l$  SR

### 8.3 Charge-Misidentification Background

Charge-misidentification contributes to the background for  $2l$  SS case and only for flavor channels, which include electrons. The same-sign require is useful in removing large SM opposite sign backgrounds, but because of their size even small charge mis-identification rates result in contamination in same-sign regions. For the  $2l$  SS signal regions and low NJet control regions, charge-misidentification background arise primarily from  $t\bar{t}\text{di-lepton}$  events with a smaller contribution from leptonic  $Z$  decays.

In general, charge-misidentification can arise in two ways. The first occurs for ultra-high energy electrons and muons, which leave tracks in the detector that are too straight for the fit to determine the direction of curvature with high confidence. This type of charge mis-identification is not a concern to the  $t\bar{t}H$  multi-lepton analysis, as most of the leptons have momentum  $> 150 \text{ GeV}/c$ . The second source of charge misidentification is from 'tridents', which only occurs for electrons, because their low mass allows for high rate bremmstrahlung in the detector material. In some cases, after an electron

736 releases a photon through bremstrahlung, the photon may convert nearby resulting in three electron  
 737 tracks. The reconstruction algorithms may sometimes match the wrong track to the calorimeter  
 738 energy deposit, resulting in a possible charge mis-identification. As discussed in the selection, tight  
 739 track-cluster geometric and energy matching requirements are applied on the electron candidates to  
 740 reduce the overall rate and the electron acceptance is narrowed to ( $|\eta| < 1.37$ ), since most of the  
 741 material is concentrated more forward in the detector. For this reason, muon charge-misidentification  
 742 is considered negligible for this analysis.

743 We estimate the contribution of charge-misidentification events in our 2l SS signal regions and  
 744 relevant control regions by applying a weight per electron in the opposite-sign region with otherwise  
 745 same cuts. The weight is related to the charge-misidentification rates and is estimated using a like-  
 746 lihood method in the opposite-sign and same-sign  $Z \rightarrow ee$  control regions. The rate measured from  
 747 these control regions is binned in electron  $p_T$  and  $\eta$ , to account for dependencies in these variables.  
 748 The method, validations and associated errors are discussed in detail in the following sub-sections.

### 749 8.3.1 Likelihood Method

750 The number of reconstructed same-sign ( $N_{ss}$ ) and opposite sign ( $N_{os}$ )  $Z \rightarrow ee$  events are related  
 751 to number of produced  $Z \rightarrow ee$  opposite sign events ( $N$ ) via factors related to the charge mis-  
 752 identification rate. For a single per-electron charge mis-identification rate ( $\epsilon$ , these quantities are  
 753 related as follows (with the assumption that  $\epsilon$  is very small):

- 754 •  $N^{os} = (1 - 2\epsilon + 2\epsilon^2)N$  opposite-sign events,
- 755 •  $N^{ss} = 2\epsilon(1 - \epsilon)N \simeq 2\epsilon N$  same-sign events,

756 Knowing  $\epsilon$ , the charge-misidentification rate, and supposing we can have a different rate per-  
 757 electron, it is possible to estimate the number of same-sign events from the number of opposite sign  
 758 events.

- 759 •  $N^{ss} = \frac{\epsilon_i + \epsilon_j - 2\epsilon_i\epsilon_j}{1 - \epsilon_i - \epsilon_j + 2\epsilon_i\epsilon_j} N^{os}$  for the  $ee$  channel,
- 760 •  $N^{ss} = \frac{\epsilon}{1 - \epsilon} N^{os}$  for the  $e\mu$  channel,

761 where  $\epsilon_i$  and  $\epsilon_j$  are the charge mis-identification rates for the two different electrons.

762 Although it is impossible from a typical same-sign  $Z \rightarrow ee$  to know which electron's charge  
 763 was mis-identified, we can use a likelihood method over the whole sample to measure charge mis-  
 764 identification rate ( $\epsilon$ ) depends on the electron  $p_T$  and  $\eta$ . The likelihood method assumes that the

765 mis-identification rates of the electron charge are independent for different pseudorapidity regions.  
 766 Therefore, the probability to have a number of same-sign events ( $N_{ss}^{ij}$ ) with electrons in  $|\eta|$  region  $i$   
 767 and  $j$  can be written as a function of the number of events  $N^{ij}$  as follows:

$$N_{ss}^{ij} = N^{ij}(\epsilon_i + \epsilon_j). \quad (8.1)$$

768 If all the same-sign events in the  $Z$  peak are produced by charge mis-identification, then  $N_{ss}^{ij}$  is  
 769 described by a Poisson distribution:

$$f(k, \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad (8.2)$$

770 where  $k$  is the observed number of occurrences of the event, i.e.  $k = N_{ss}^{ij}$ , and  $\lambda$  is the expected number,  
 771 i.e.  $\lambda = N^{ij}(\epsilon_i + \epsilon_j)$ . Thus, the probability for both electrons to produce a charge mis-identification  
 772 is expressed by:

$$P(\epsilon_i, \epsilon_j | N_{ss}^{ij}, N^{ij}) = \frac{[N^{ij}(\epsilon_i + \epsilon_j)]^{N_{ss}^{ij}} e^{-N^{ij}(\epsilon_i + \epsilon_j)}}{N_{ss}^{ij}!}. \quad (8.3)$$

773 The likelihood  $L$  for all the events is obtained by evaluating all the  $|\eta|$  combinations:

$$L(\epsilon | N_{ss}, N) = \prod_{i,j} \frac{[N^{ij}(\epsilon_i + \epsilon_j)]^{N_{ss}^{ij}} e^{-N^{ij}(\epsilon_i + \epsilon_j)}}{N_{ss}^{ij}!}, \quad (8.4)$$

774 where the rates  $\epsilon_i$  and  $\epsilon_j$  can be obtained by minimizing the likelihood function. In this process, the  
 775  $-\ln L$  is used in order to simplify and make easier the minimization. Terms which do not depend on  
 776 the rates  $\epsilon_i$  and  $\epsilon_j$  are removed in this step. This way, the final function to minimize is given by the  
 777 following expression:

$$-\ln L(\epsilon | N_{ss}, N) \approx \sum_{i,j} \ln[N^{ij}(\epsilon_i + \epsilon_j)] N_{ss}^{ij} - N^{ij}(\epsilon_i + \epsilon_j). \quad (8.5)$$

778 The events are selected within the  $Z$  peak and stored –with the electron order by  $|\eta|$ – in two  
 779 triangular matrices: one for the same-sign events  $N_{ss}^{ij}$ , and the other one for all events  $N^{ij}$ . The  
 780 likelihood method takes into account electron pairs with all  $|\eta|$  combinations, which allows to use the  
 781 full available statistics getting therefore lower statistical uncertainties. Moreover, it does not bias the  
 782 kinematical properties of the electrons, compared to other methods like tag-and-probe.

783 The likelihood method can be easily extended to measure the charge mis-identification rates as a  
 784 function of two parameters. In this study, the interest lies not only on the measurement of the rates  
 785 as a function of the pseudorapidity, but also transverse momentum. Thus, the probability to find a  
 786 same-sign event given the rates for each electron is  $(\epsilon_{i,k} + \epsilon_{j,l})$ , where the two indices represent binned  
 787  $|\eta|$ - and  $p_T$ -dependence. Thus, the Eq. 8.5 transforms into

$$-\ln L(\epsilon | N_{ss}, N) \approx \sum_{i,j,k,l} \ln[N^{ij,kl}(\epsilon_{i,k} + \epsilon_{j,l})] N_{ss}^{ij,kl} - N^{ij,kl}(\epsilon_{i,k} + \epsilon_{j,l}). \quad (8.6)$$

788        The likelihood method uses only  $Z$  *signal* events. Therefore, background coming from other  
 789 processes where the dilepton invariant mass corresponds to the one of the  $Z$  boson needs to be  
 790 subtracted. The background subtraction is done using a simple side-band method. This method  
 791 consists in dividing the  $Z$  invariant mass in three regions, i.e.  $A$ ,  $B$  and  $C$ , where  $B$  is the central  
 792 region corresponding to the  $Z$  peak. The number of events is counted in the regions on the sides of  
 793 the peak, i.e.  $n_A$  and  $n_C$ , and removed from the total number of events in the peak region  $B$ ,  $n_B$ .  
 794 This way, the number of signal events  $N_Z$  is given by

$$N_Z = n_B - \frac{n_A + n_C}{2}. \quad (8.7)$$

795        Once the background has been subtracted, the likelihood method can be applied. MINUIT is used  
 796 for the minimization and MIGRAD to compute the uncertainty on these rates.

### 797        8.3.2 Results

798        The charge mis-identification rate is calculated in 7  $|\eta|$  bins [0.0, 0.6, 1.1, 1.37, 1.52, 1.7, 2.3, 2.47]  
 799 by 4  $p_T$ bins [15,60,90,130,1000]. For  $p_T$ bins above 130 GeV/ $c$ , the  $Z$  dataset becomes too small and  
 800 the rates are calculated using  $t\bar{t}$ MC, scaled by the data-MC ratio of the rates in the lower  $p_T$ bins,  
 801 [90-130] GeV/ $c$ . Figure 8.6 shows the extracted rates in all bins.

802        To validate the likelihood approach, we apply the full method to the  $Z$  MC samples (extracting  
 803 rates via a likelihood fit and applying them to opposite sign events) and compare to the MC predicted  
 804 number of same-sign events. The invariant mass of the  $Z$  from our charge mis-identification and  
 805 directly from the MC can be seen on Figure 8.7. In the simulated  $Z$  samples, the number of same-  
 806 sign  $Z$  events is 5 049 while the estimation is  $5 031^{+375}_{-365}$ . The uncertainties combine both statistical  
 807 systematic uncertainties, which are discussed in depth below. The validation gives compatible results  
 808 within uncertainties.

### 809        8.3.3 Systematic and Statistical Uncertainties

810        Statistical uncertainties dominate the combined uncertainty on the charge mis-identification estimate.  
 811 The statistical uncertainties come primarily from the size of the  $Z$  same-sign sample in data and are  
 812 especially large for central, material-poor regions where the charge mis-identification rate is extremely  
 813 low. Additionally systematic uncertainties are included for a comparison between the positron and  
 814 electron rate, the per-bin MC closure test discussed above in the Results section, and for the effect of  
 815 varying the invariant mass window used for the background subtraction for three different cases. The  
 816 high  $p_T$ extrapolation induces a statistical error only in the last  $p_T$ bin. This bins is essentially irrelevant

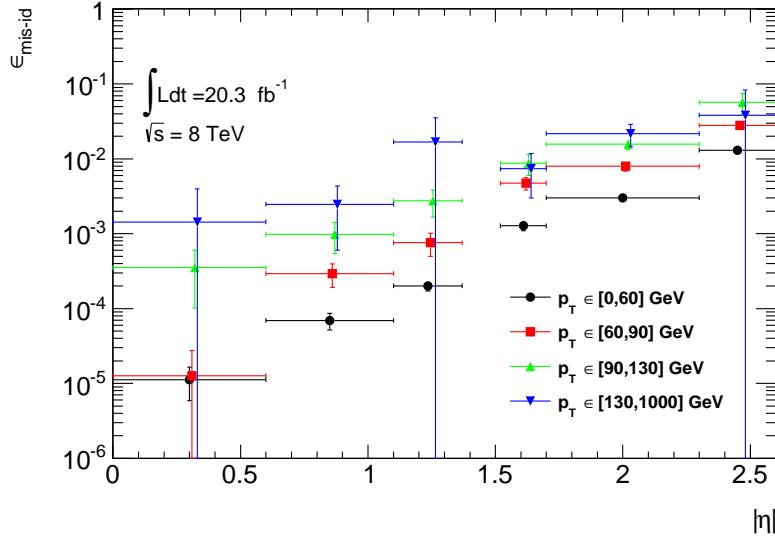


Figure 8.6: Electron charge mis-identification rates measured in data with the likelihood method on  $Z$  events (black points, red squares and blue triangles) as a function of  $|\eta|$  and parametrized in  $p_T$ . The full 2012 dataset has been used to estimate the rates below 130 GeV. Above this value, the charge mis-identification rates have been estimated by extrapolating the rates in the region where the  $p_T \in [90, 130]$  GeV with a  $p_T$  dependent factor extracted from simulated  $t\bar{t}$  events (green triangles). Statistical and systematic uncertainties have been included in this plot.

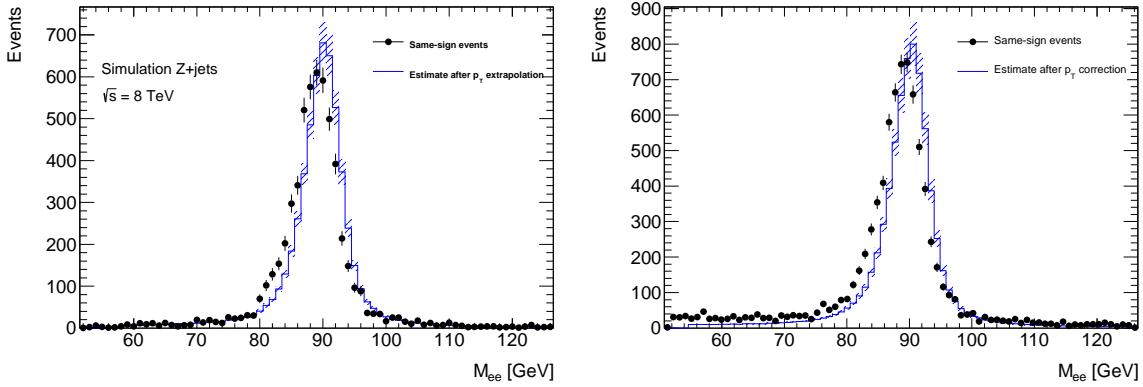


Figure 8.7: Closure test on simulated  $Z \rightarrow e^+e^- + jets$  events (a) and data (b). The black circles show the distribution of same-sign events while the blue histograms show the distribution of the reweighted opposite-sign events together with the statistical and systematic uncertainties. The distributions are not expected to overlay exactly, due to the loss of energy of the trident electrons for the same-sign peak.

to the energy scales considered in this analysis. Figure 8.8 shows the relative bin uncertainties for all rate bins.

We apply the rates to estimate the charge mis-identification background in the 2l SS signal regions,

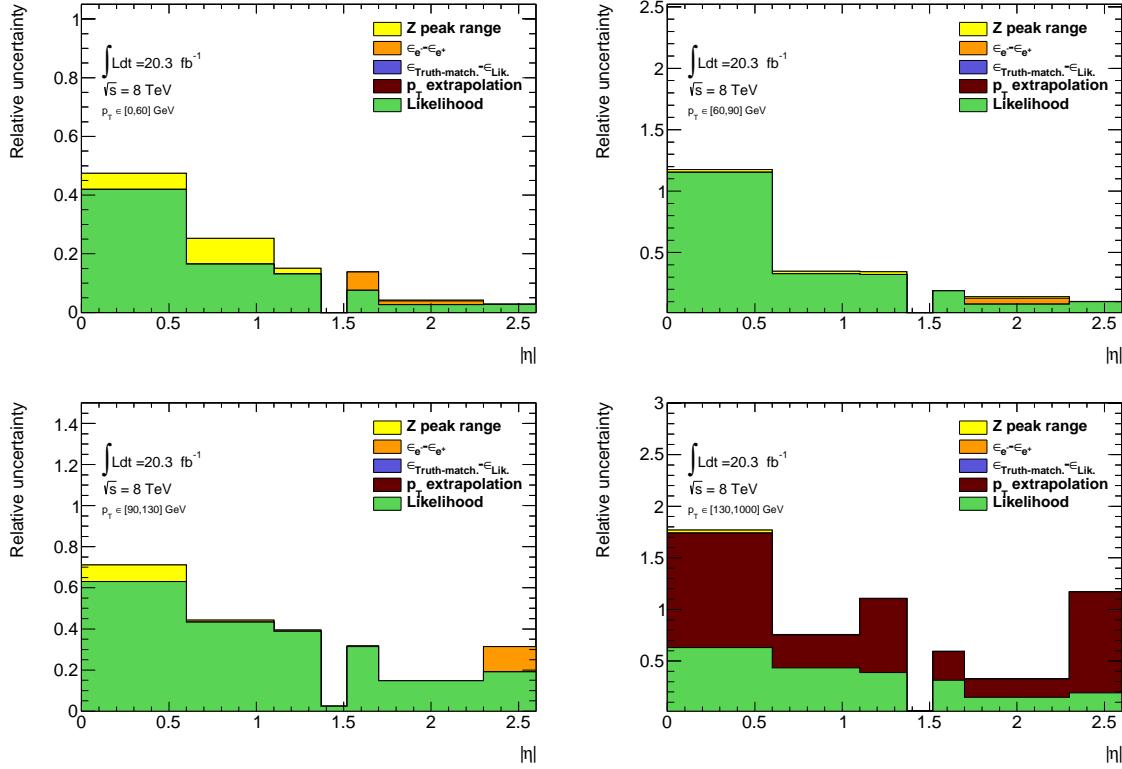


Figure 8.8: Relative systematic uncertainty contributions on the charge mis-identification rate, for different bins in  $p_T$  and  $|\eta|$ . Tight++ electrons have been used to produce this plot.

and find  $\sim 25\%$  contamination in the  $e^\pm e^\pm$  regions and a  $\sim 10\%$  contribution to the  $e^\pm \mu^\pm$  regions with a  $10\%$  systematic error overall. The low overall error can be attributed to the fact that the statistical error is lowest where the bulk of charge misidentifications occur.

#### 8.4 Fake Lepton Backgrounds

Fake Leptons, from the mis-identification of jets as either electrons or muons, primarily arise from  $t\bar{t}$  and single top processes in the 2l SS, 3l and 4l channels. Smaller contributions come from  $Z + \text{jet}$  events. These backgrounds are sub-dominant but important in the 2l SS and 3l channels. They are extremely small in the 4l channels. Truth studies suggest that these mis-identified leptons arise overwhelmingly from b-quark initiated jets. The general method for estimating fakes in all channels is to define a reversed object selection region (usually isolation) for each lepton flavor with otherwise identical signal region selection ( $N_{CR}^e$ ,  $N_{CR}^\mu$ ). This region is fake-dominated with small contributions from prompt backgrounds, which are subtracted from the data. The total number of fake events

832 in this region is then scaled by a transfer factor ( $\theta$ ) to estimate the number of fake events of the  
 833 appropriate flavor in the signal region. The transfer factor is defined in Equations 8.8 and the simple  
 834 formula for determining fakes is defined in Equations 8.9. 'd' refers to anti-identified electrons, and  
 835 'p' refers to anti-identified muons.

$$\theta_e = \frac{N_{ee}}{N_{ed}}, \theta_\mu = \frac{N_{\mu\mu}}{N_{\mu p}} \quad (8.8)$$

$$N_{fake} = \theta_e * N_{CR}^e + \theta_\mu * N_{CR}^\mu \quad (8.9)$$

836 This approach factorizes the background model into two separate measurements.  $N_{CR}$  is sensitive  
 837 to the overall  $t\bar{t}$  production rate, especially in the presence of additional jets from QCD ratio, as well as  
 838 the object-level misidentification of a jet as a lepton. The transfer factor  $\theta$  is sensitive to only the  
 839 object level properties of the mis-identified jet, and in particular only the variables which are reversed  
 840 in the anti-tight identification.

841 The transfer factor is obtained in a different way for each channel, due to unique issues with  
 842 statistics and contamination, but each method relies heavily on the data-based control regions with  
 843 fewer jets. Figure 8.9 shows a truth study of the stability of the transfer factor for the 2l SS and 3l  
 844 cases as a function of the number of jets in the event for events with one-btagged jet. This suggests  
 845 that the regions with fewer jets are a good model of the fakes in the signal regions with more jets and  
 846 is expected because of the homogeneity of origin of the fakes across all jet bins.

847 The details of the methods for each channel are discussed in depth in the following sections. For  
 848 all methods, the overall systematic uncertainty on the normalization of the fake estimate is in the  
 849 range 30%-50% and arise primarily from statistics and the closure on assumptions used to obtain the  
 850 transfer factor.

851 Because these methods do provide a per-object transfer-factor that depends on the properties of  
 852 the faking object, we must use the MC to model the shapes of the fake kinematic distributions in the  
 853 signal regions. This is not an essential issue, since the analysis only considers only the total number  
 854 of events in each signal region in the final measurement of  $t\bar{t}H$  production.

#### 855 8.4.1 2l SS Fakes

856 The 2l SS fake method follows the procedure outlined in general above. We define anti-tight electron  
 857 and muon control regions with reversed particle identification criteria for each signal region, including  
 858 the 6 flavor and jet-counting sub regions. The anti-tight muon and electron criteria are provided  
 859 below:

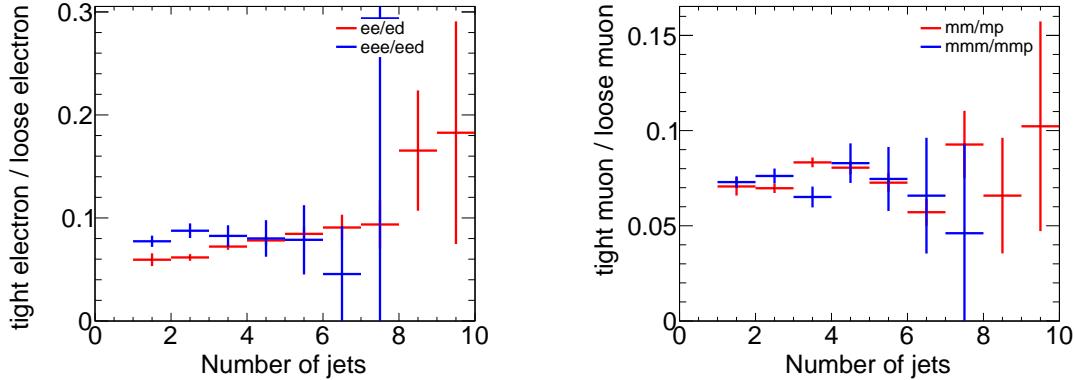


Figure 8.9: Ratios of regions with tight and anti-tight leptons in 2-lepton and 3-lepton channels

- anti-tight electron (d): fails to verify the verytight likelihood operating point, but still verifies the veryloose opoperating point. fails relative tracking and calorimeter isolation,  $E_T^{rel} > 0.05$  and  $p_T^{rel} > 0.05$ .
- anti-tight muon (p:  $6 \text{ GeV}/c < p_T < 10 \text{ GeV}/c$

The electron and muon transfer factors,  $\theta_e$  and  $\theta_\mu$ , are calculated in the region with signal region selection but fewer jets,  $NJet == 2$  or  $NJet == 3$  and are defined as the ratio of the number of events for two fully identified leptons to the number of events with one fully identified lepton and one anti-identified lepton, after the prompt and charge mis-identification backgrounds are subtracted. Only same-flavor channels are used to ensure that muon and electron transfer factors maybe estimated separately: on every region, the prompt and charge-misidentification backgrounds are subtracted from the data.

$$\theta_e = \frac{N_{ee}}{N_{ed}} = \frac{N_{ee}^{Data} - N_{ee}^{\text{Prompt SS}} - N_{ee}^{\text{QMisId}}}{N_{ed}^{Data} - N_{ed}^{\text{Prompt SS}} - N_{ed}^{\text{QMisId MC}}} \quad (8.10)$$

$$\theta_\mu = \frac{N_{\mu\mu}}{N_{\mu p}} = \frac{N_{\mu\mu}^{Data} - N_{\mu\mu}^{\text{Prompt SS}}}{N_{\mu p}^{Data} - N_{\mu p}^{\text{Prompt SS}}} \quad (8.11)$$

The 2,2 jet anti-tight regions used in obtaining the transfer factors are shown in Figure 8.10 and the 4,5 jet anti-tight regions, scaled by the transfer factors to get the fake estimates in the SR are shown in Figure 8.11. The  $t\bar{t}\text{MC}$  is included in the plots for reference, although it is not included in the measurements.

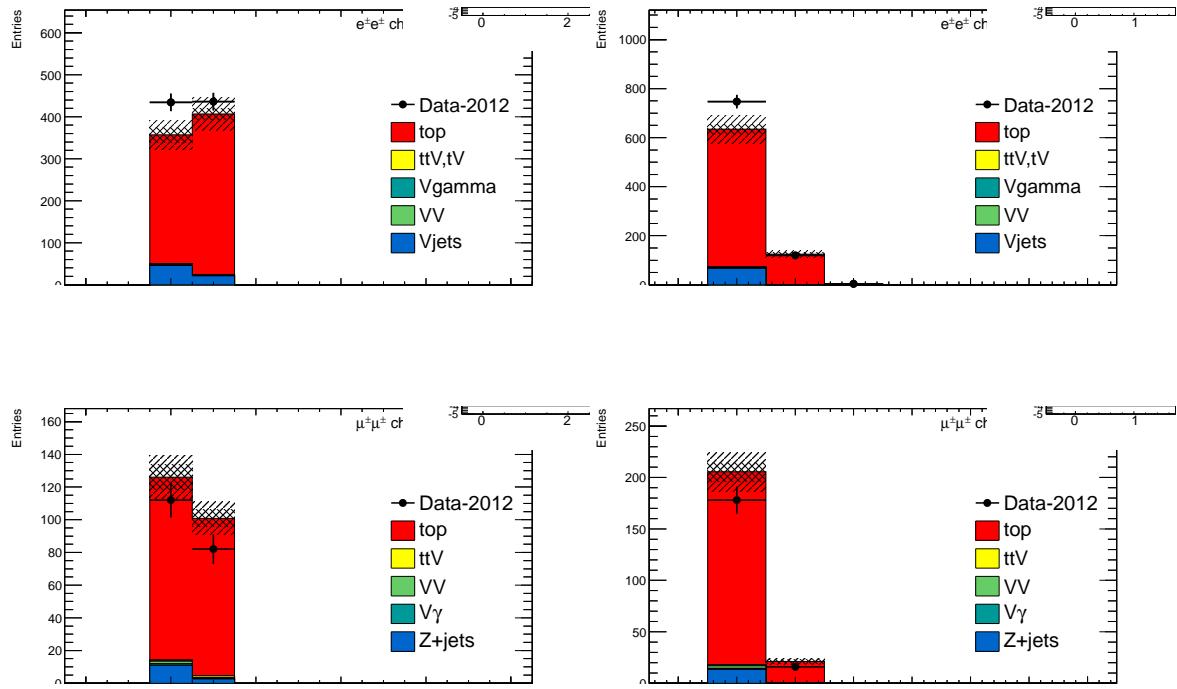


Figure 8.10: 2,3 Jet SS 2l  $ed$  (above) and  $\mu p$  control regions with at least one b-tagged jet. This region and the associated tight-id regions are used to determine  $\theta_e$  and  $\theta_\mu$ . The  $t\bar{t}$ MC (red) is used for reference and not used in the actual calculation

Factor	Expected (MC)	Measured (data)
$\theta_e$ Rev. Id.	$0.0136 \pm 0.0062$	$0.0156 \pm 0.0062$
$\theta_\mu$ Rev. $p_T$	$0.078 \pm 0.012$	$0.1156 \pm 0.0288$

Table 8.5: Expected and measured values of the  $\theta$  factors.

877     The transfer factors obtained from the 2 and 3 jet regions are shown in 8.5 with statistical errors  
 878     and propagated systematic errors from the subtraction of relevant backgrounds ( $t\bar{t}V$  and charge mis-  
 879     identification). The MC values are just for comparison. An additional systematic error is added by  
 880     comparing the transfer factors, obtained from the low jet control region, to those obtained from the  
 881     higher jet signal regions, using  $t\bar{t}$ MC. The value of this systematic is about 20 % and can be seen in  
 882     the above Figure 8.9. The overall systematic uncertainties and contribution from each source in all of  
 883     the sub-channels of the signal region are shown in Table 8.5 and the final contribution of fake events  
 884     to the signal region are show in Table 8.1 found at the beginning of the chapter.

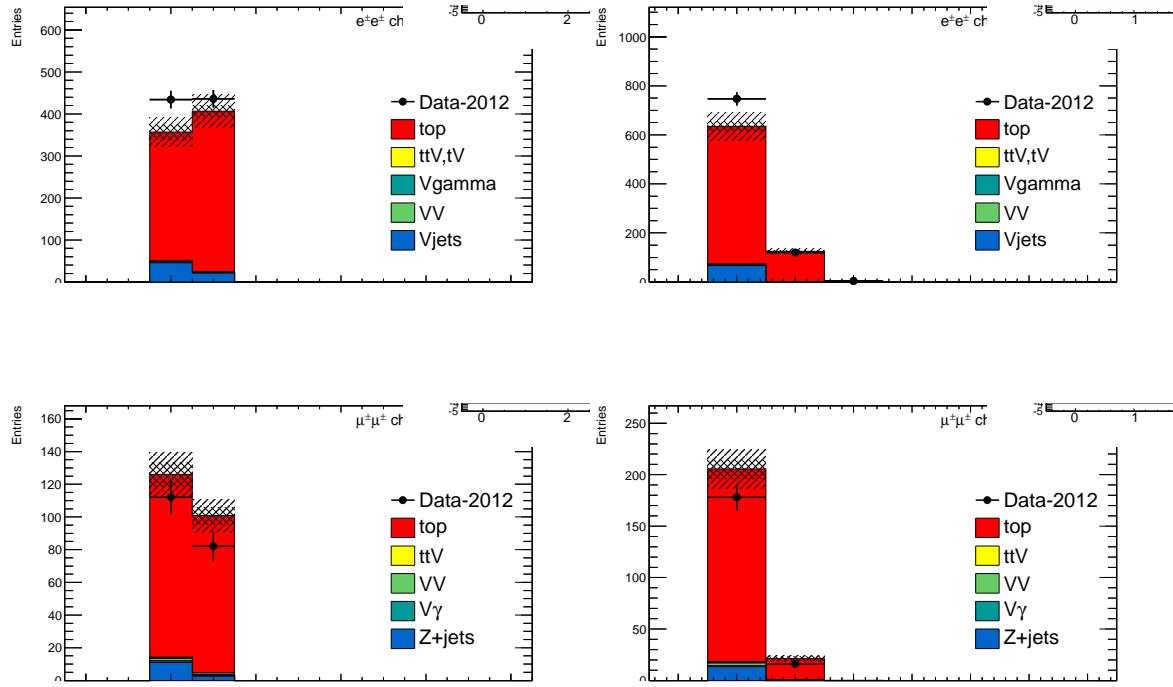


Figure 8.11: 4,5 Jet SS 2l  $ed$  (above) and  $\mu p$  control regions with at least one b-tagged jet. After subtraction of prompt and charge mis-identification backgrounds, these regions are scaled by the transfer factors,  $\theta_e$  and  $\theta_\mu$  to obtain the final number of fake events in the CR. The  $t\bar{t}$ MC (red) is used for reference and not used in the actual calculation

Uncertainties		Channels					
		4 jets		$\geq 5$ jets			
		$e^\pm e^\pm$	$\mu^\pm \mu^\pm$	$e^\pm \mu^\pm$	$e^\pm e^\pm$	$\mu^\pm \mu^\pm$	$e^\pm \mu^\pm$
Statistical	$\Delta\theta_e^{stat}$	39.6	—	14.2	39.6	—	18.5
	$\Delta\theta_\mu^{stat}$	—	24.7	15.8	—	24.7	13.1
	$\Delta N_{\ell anti-\ell}(n \text{ jets})(\text{stat})$	6.8	18.1	21.3	8.4	25.9	22.7
Systematics	$\Delta\theta_e^{syst}$ (closure)	21.8	—	7.8	26.7	—	12.5
	$\Delta\theta_\mu^{syst}$ (closure)	—	23.3	18.4	—	31.2	19.7
	Q Mis Id ( $\ell anti-\ell$ )	2.2	—	1.5	2.4	—	1.5
Total		45.7	38.5 (36.3)	35.7	48.5	47.8 (43.9)	39.6
Systematic	Q Mis Id ( $\ell\ell$ )	24.0	—	8.6	24.0	—	11.3

Table 8.6: Summary of the uncertainties (in %) in  $e^\pm e^\pm$  (reverse Id + reverse isolation method),  $\mu^\pm \mu^\pm$  and  $e^\pm \mu^\pm$ . Statistical uncertainty is split into statistical uncertainties on  $\theta_e$  and  $\theta_\mu$  and uncertainty due to the Control Region size ( $\Delta N_{\ell anti-\ell}(n \text{ jets})$ ). For channels with muons, uncertainties between parenthesis are obtained when anti-tight muons are defined such as  $p_T < 20$  GeV (includes blinded data)

885 **8.4.2 3l Fakes**

886 The 3l fake method follows the same general strategy as the 2l SS case. Transfer factors are used  
 887 extrapolate from an anti-tight, fake-rich control region in data into the signal region. However, the  
 888 equivalent low jet control regions are too low in statistics to provide the transfer factors from data  
 889 directly, as above. Instead, the transfer factors are obtained directly from the  $t\bar{t}$  simulation and data  
 890 control regions are used to determine the modeling of the identification and isolation variables, used in  
 891 the transfer factor extrapolation. The low jet regions are still employed in a low statistics validation  
 892 of the entire fake procedure.

893 Anti-tight electrons and muons are defined in slightly different ways, compared to the 2l SS case:

- 894 • **anti-tight electron (d):** fails to verify the verytight likelihood operating point, but still verifies  
 895 the veryloose operating point. the isolation selection is released  $E_T^{rel} > 0.05$ ,  $p_T^{rel} > 0.05$ .
- 896 • **anti-tight muon(p):** muons must pass identification but the  $p_T$  cuts is lowered to  $6 \text{ GeV}/c$ , the  
 897 overlap removal with jets and isolation cuts are released.

898 The transfer factors,  $\theta_e$  and  $\theta_\mu$ , extracted from MC, is defined as the ratio of the number of top ( $t\bar{t}$ +  
 899 single-top) events in the signal region, to the number of  $t\bar{t}$ events in the anti-tight regions. The factors  
 900 are calculated in separate flavor regions to ensure that the electron jet fakes and muon jet fakes are  
 901 calculated separately. The calculation follows the same for as in Equation ??, but now lep0, which  
 902 by construction is almost never a fake is allowed to be either electron or muon in both cases, denoted  
 903 below in Equation 8.12.

$$\theta_e = \frac{N_{xe\bar{e}}}{N_{x\bar{e}d}}, \theta_\mu = \frac{N_{x\mu\mu}}{N_{x\mu p}} \quad (8.12)$$

904 The MC modeling of the variables involved in the transfer factor can be verified when another  
 905 variable fails. For instance, the MC modeling of the electron isolation variable can be compared to  
 906 data when the particle identification variable fails and vice-versa. The modeling of muon-jet  $\Delta R$ ,  
 907 involved in the overlap removal, can be compared when either the isolation variable or the  $p_T$  fails.  
 908 The comparison of the electron variables in this manner can be seen in Figure 8.13 and the muon  
 909 variables in Figure 8.12. The regions used have the same selection as the signal region with an added  
 910 missing transverse energy requirement,  $> 60 \text{ GeV}/c$ , to ensure only top fakes. 20% and 21% systematic  
 911 uncertainties are assigned to the muon and electron transfer factors, respectively, to account for data-  
 912 MC discrepancies. This method for evaluating data-MC agreement for individual electron and muon  
 913 variables in turn relies on the assumption that these variables are largely un-correlated and that the

transfer factor itself is factorizable into pieces for each variable. This factorized and fully correlated transfer factors have therefore been compared using MC and shown to have differences than the systematic quotes, suggesting that these assumptions are reasonable.

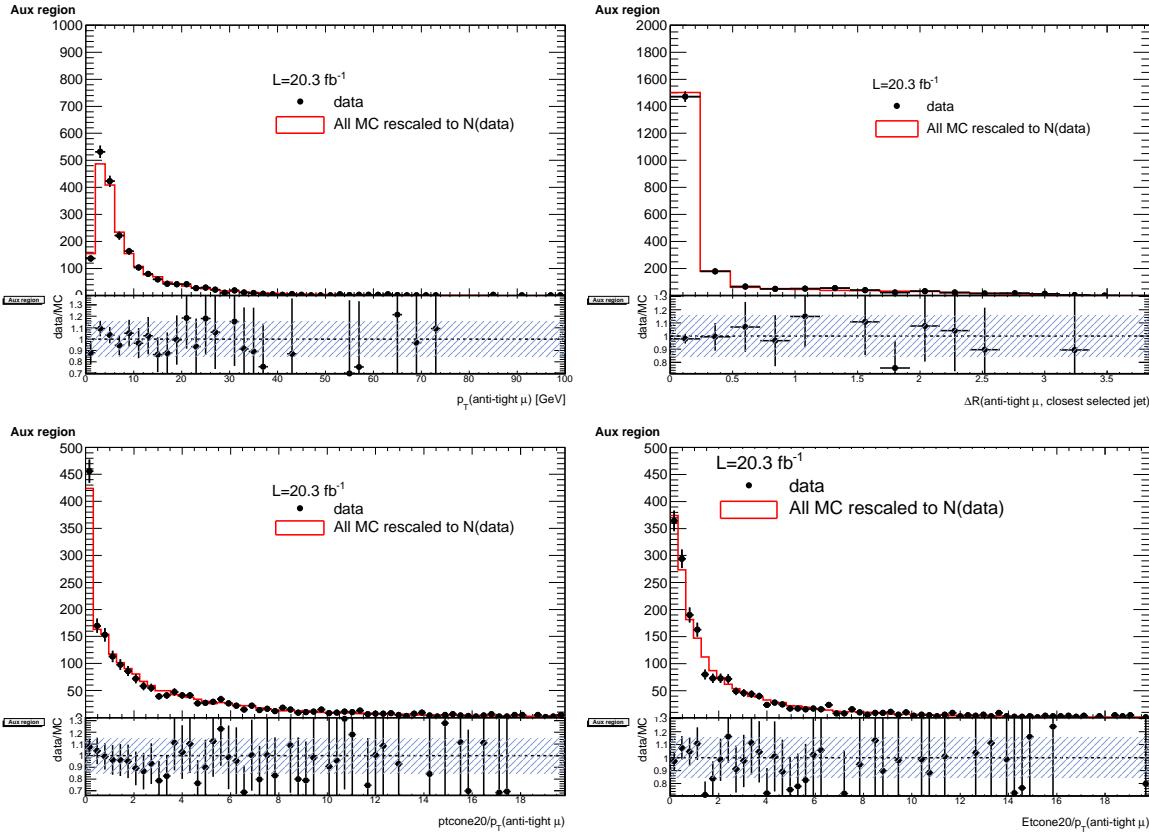


Figure 8.12: Distributions of the properties of the anti-tight muons in data (dots), compared with the total simulation (red line), rescaled to the integral of the data for a shape comparison. The uncertainty on the data distribution is statistical. The number of events for each of them is also presented in the legend. The variables probed are, top:  $p_T$  and  $\Delta R(\mu, \text{closest selected jet})$ ; bottom:  $\text{ptcone20}/p_T$  and  $\text{Etcone20}/p_T$ . The selection is the signal region event selection with one anti-tight muon (failing at least one of the isolation, muon-jet overlap, or  $p_T$  selection criteria) A ratio plot is containing the 20% area around 1, is displayed, demonstrating that a 20% data-MC comparison systematic is sufficient.

The electron and muon anti-tight control regions, which are scaled by the transfer factors are shown in Figure 8.14. The prompt MC subtracted data in these regions are scaled by the transfer factors to obtain the overall contribution of fake electron and muon events in the signal region. The systematic uncertainties are split between the statistical error on the transfer factor and normalization of the anti-tight control regions and the data-MC comparisons outlined above. For muon fakes the total systematic uncertainty is 25% and for electrons it is 34%. The numbers and uncertainties involved

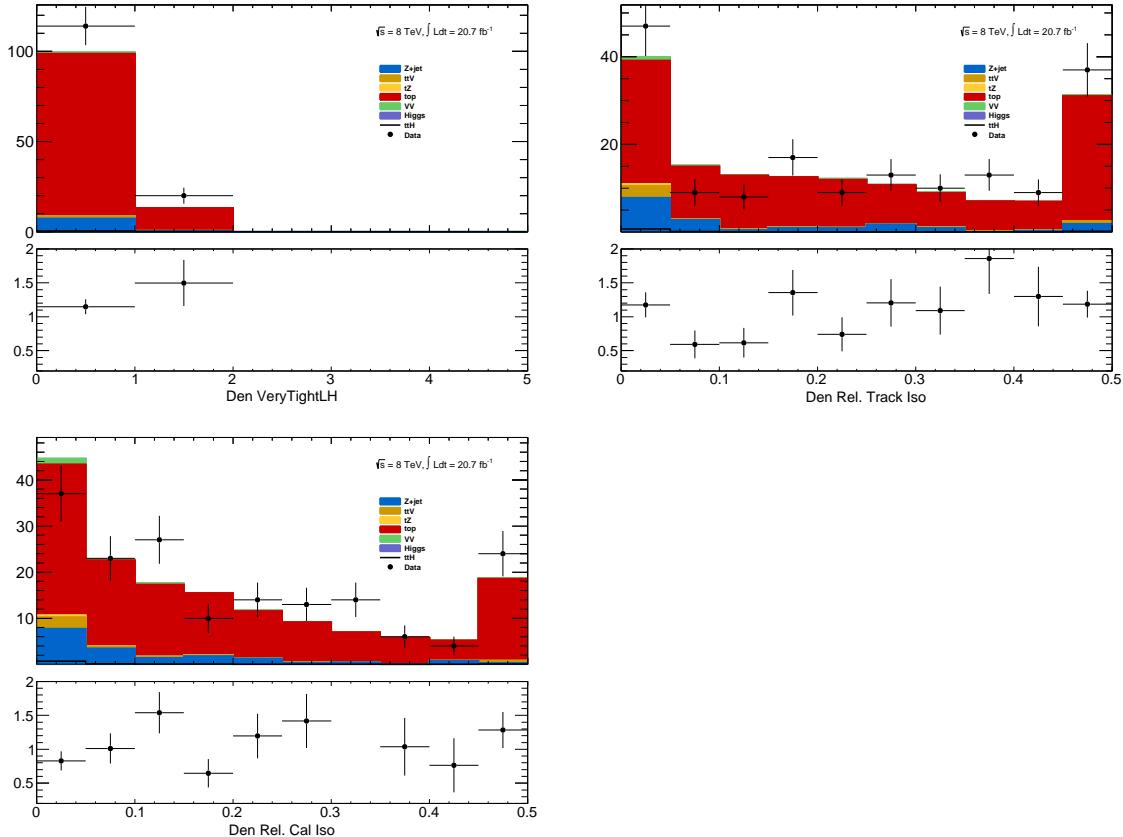


Figure 8.13: Distributions of anti-tight electron variables. The variables presented are, from top left to bottom right,  $p_T$ ,  $\eta$ , Very Tight Likelihood word value,  $\text{ptcone20}/p_T$ ,  $\text{Etcone20}/p_T$ . The plotted regions have the same cuts as the signal region, except the anti-tight electron must fail isolation for the plot of the verytight identification word or fail the verytight identification word for the plots of the isolation. Data (dots) are compared with a stacked histogram of the various simulated samples: top in red, V+jets (blue), VV (purple) and ttV (yellow). The uncertainty on the data distribution is statistical.

923 in the calculation are shown in Table 8.7.

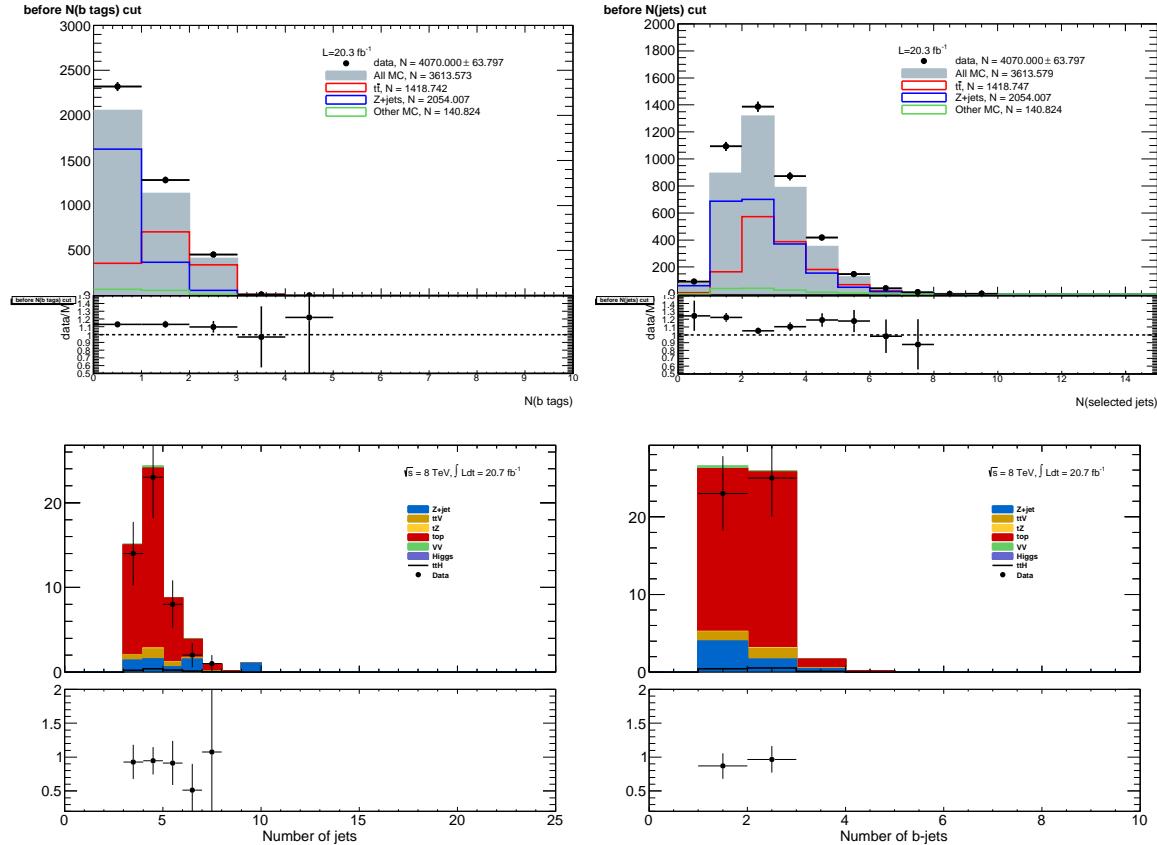


Figure 8.14: Muon-xxp (above) and electron-xxd (below) anti-tight control regions: jet variables. Note: the  $t\bar{t}$  and top MC in the plots is used only as comparison, but is not included in the fake measurement

Stage	Muon	Electron
Anti-Tight CR Normalization	$364.62 \pm 20.02$ (5%)	$38.2 \pm 6.9$ (17%)
Transfer factor $\theta_\mu$	$0.0047 \pm 0.0011$ (23%)	$0.0240 \pm 0.0064$ (29%)
SR Contribution	$1.71 \pm 0.42$ (25%)	$0.91 \pm 0.29$ (34%)

Table 8.7: Summary of regions and inputs to the extraction of the number of  $t\bar{t}$  background events with a fake muon in the SR

924 Finally, the low jet region (1,2,3j) is used as a validation for the method, described above. The  
 925  $t\bar{t}$  and single top fakes are estimated using the procedure above, but instead using the lower jet region.  
 926 Similar systematics are assessed. This region with the fake estimate is plotted in Figure 8.15. The  
 927 agreement of data and summed prediction for the fakes and prompt backgrounds is well within the

systematic and statistical uncertainties. The figure also shows the same region with relaxed  $p_T$ cuts on all leptons to  $10 \text{ GeV}/c^2$ . This increases the purity of fakes in the region as well as the statistics. The data and summed fake and prompt predictions are also well within the statistical and systematic uncertainties.

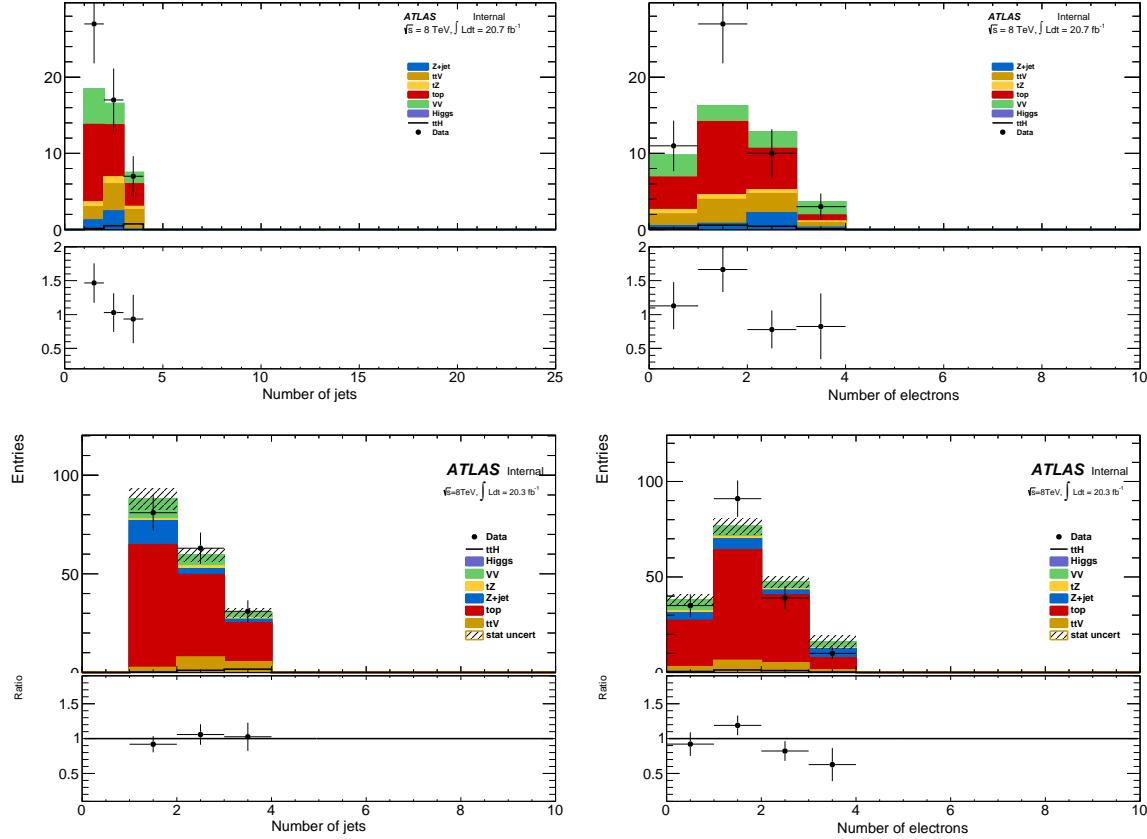


Figure 8.15: 3l fake validation regions for nominal  $p_T$ selection (above) and relaxed  $p_T$ selection,  $> 10 \text{ GeV}/c$ , (below). Plotted are the number of jets and the number of electrons in each event. The data and MC ratio below each plot agree with 1 within the statistics of the region and the overall systematic assigned for the fake component (red)

### 8.4.3 4l Fakes

We will not discuss the 4l fakes in depth, as it is a very small background - at the % level and will have almost no impact on the final result. It is important, however, to carry out the measurement using the data to ensure that this is indeed the case. The fake method used in the the 4l case is similar to the 2l and 3l cases discussed above. All fakes arise from  $t\bar{t}$  and single-top events, where two jets are mis-identified as leptons. To measure the contribution of this background, control regions with 2 fully

938 identified and 2 anti-identified leptons are created. These control regions do not have a number of jets  
939 requirement in order to increase statistics. From these control regions, two extrapolations are made.  
940 First, a transfer factor is applied to extrapolate from the anti-tight to tight regions for electrons  
941 and muons. The regions are defined with identifical object identification selection and reversal as  
942 the 3l case, and the same transfer factors can be used. They must be used twice however, because  
943 there are two anti-identified leptons in each event. Second, the jet inclusive regions are extrapolated  
944 into the 2-jet signal region, using as a second extrapolation factor derived from  $t\bar{t}$ events. Since, the  
945 majority of fake leptons arise from b-quark initiated jets, the jet spectrum  $t\bar{t}$  events with the additional  
946 requirement of 2-btagged jets from data are used as a model for the jet extrapolation. The overall  
947 systematic uncertainty on this measurement arises from the statistics in the control regions and MC  
948 based assessments of non-closure and are 35%-50% depending on the sub-channel.

949

## CHAPTER 9

---

950

# Systematic Uncertainties and Statistical Formalism

---

951

952 This chapter summarizes the various systematic uncertainties that enter the measurement of the limit  
 953 of  $t\bar{t}H$  multi-lepton analysis, as well as structure of the statistical analysis model used to obtain  
 954 the measurement. The systematic uncertainties arise from three main sources. The first are the  
 955 normalization uncertainties on the background process estimation methods, which are discussed in  
 956 depth in . The second source is the theoretical uncertainties on the  $t\bar{t}H$  production cross-section  
 957 and acceptance. The final source are the experimental and detector related systematic uncertainties  
 958 related to event selection efficiencies and measurements and identification of the objects. They  
 959 affect only the non-data driven backgrounds and the  $t\bar{t}H$  signal, as simulation is used to model their  
 960 acceptance and efficiency for the analysis selection.

961 These systematic uncertainties, the estimated background and signal event counts in each of the  
 962 signal regions, and the observed data in each signal region are combined in a statistical fit to an  
 963 analysis model to extract the measurement of interest. We measure per-channel and combined ratios  
 964 of the observed production rate to the theoretically predicted production rate of  $t\bar{t}H$  , a parameter  
 965 called  $\mu$ . In the absence of a statistically significant observation, this measurement is translated into a  
 966 upper confidence limit on  $\mu$ . The details of this procedure are discussed in the following sections and  
 967 the results with the observed data are discussed in Chapter ??

### 968 9.1 Systematic Uncertainties on Signal Cross-section and Acceptance

969 The  $t\bar{t}H$  signal is simulated with matrix elements at NLO (next-to-leading order) in QCD with Powheg  
 970 and is discussed in Chapter ??.

971 The production cross section and the Higgs boson decay branching fractions together with their

972 theoretical uncertainties from the QCD scale and PDF choice are taken from the NLO theoretical  
 973 calculations reported in Ref. [39]. The uncertainty from the QCD scale estimated by varying  $\mu_0$  by a  
 974 factor of 2 from the nominal value is  $^{+3.8\%}_{-9.3\%}$ , while the uncertainty from the PDF set and the value of  
 975  $\alpha_S$  is  $\pm 8.1\%$ .

976

977 The impact of the choice of the QCD scale on the simulation of the  $t\bar{t}H$  event selection efficiency  
 978 is estimated in two independent ways.

979 First, the factorisation and renormalisation scales  $\mu_0$  are varied by a factor of 2, as  $\mu = 2\mu_0$  and  
 980  $\mu = \mu_0/2$ . The effects of these new scales are estimated via the application of event reweighting  
 981 procedures on the nominal simulation using kinematic distributions at parton level. The weights used  
 982 are dependent on the transverse momenta of both the  $t\bar{t}H$  system and of the top quark, as described  
 983 in Ref. [?].

984 Second, the choice of the factorisation and renormalisation scales, dependent on fixed (“static”)  
 985 parameters in the nominal simulation, is tested comparing its prediction with an alternative (“dy-  
 986 namic”), but still physics motivated choice  $\mu_0 = (m_T^t m_T^{\bar{t}} m_T^H)^{1/3}$ , which depends on kinematic variables.  
 987 This comparison is performed via event reweighting of the nominal static simulation based on weights  
 988 derived as a function of the  $t\bar{t}H$  transverse momentum [?]. In order to take the difference between  
 989 the choices of scale as systematic uncertainties, a symmetric envelope around the nominal simulation  
 990 is built applying the weights and also their inverses.

991 Fig. ?? shows the impact of the different choices for the factorization and renormalization scales  
 992 on the jet and b-jet multiplicities in events with 2 SS leptons. Similar variations are seen also in  
 993 the other event categories. In order to not double-count the variations on the total cross section  
 994 the predictions from the different QCD scales are normalised to the same total cross section. That  
 995 means that the observed differences are only coming from the event selection. Significant variations  
 996 on the jet multiplicities can be seen and these translate into different predictions on the signal event  
 997 yields in the signal regions. Such differences, listed in Table ??, are taken as theoretical systematic  
 998 uncertainties in addition to the ones affecting the total  $t\bar{t}H$  production cross section. The “Static”  
 999 uncertainties come from the variations by a factor of 2 from the nominal scale and they are correlated  
 1000 with the uncertainties on the total cross section, which are estimated with the same procedure. The  
 1001 “Dynamic” uncertainties come from the difference between the nominal and the alternative dynamic  
 1002 scale and are treated as an independent source of theoretical uncertainty.

1003 The uncertainty of the  $t\bar{t}H$  event selection due to the PDF sets is estimated comparing the predic-  
 1004 tions with three different PDF sets, varying each set within errors and taking the width of the envelope

QCD scale [%]	2l4jets	2l $\geq$ 5jets	3l	4l
Static	+0.6 -0.0	+2.7 -1.3	+2.3 -0.8	+0.9 -0.2
Dynamic	+1.7 -0.8	+2.0 -2.6	+1.7 -1.1	+0.5 -0.0

Table 9.1: Uncertainties on  $t\bar{t}H$  acceptance in signal regions due to PDF variation.

Sample	2l 4j	2l 5j	3l	4l
$t\bar{t}H$	0.3%	1.0%	0.5%	1.4%

as systematic uncertainty. The recommended sets are CT10, MSTW2008nlo68cl and NNPDF21\_100. We determine the change in the acceptance due to the PDF sets via the formula 9.1 to disentangle it from the change in production cross section.

$$\left( \frac{\text{Reweighted yield in SR}}{\text{Reweighted total number of events}} \right) \left( \frac{\text{Original yield in SR}}{\text{Original total number of events}} \right)^{-1} - 1 \quad (9.1)$$

Fig. ?? shows the estimated PDF systematic uncertainties as a function of the jet multiplicity in  $t\bar{t}H$  events with at least two leptons. The uncertainties are compatible with the uncertainty on the production cross section estimated in Ref. [39] and indicated by the dashed red lines in the lower panel. Table 9.1 shows the half-width of the envelope of the acceptance under all eigenvector variations of the three PDF sets. No significant dependence on the event topology is observed, so that the PDF systematic uncertainty on the  $t\bar{t}H$  event selection is neglected.

## 9.2 Summary of Background Systematic Uncertainties

## 9.3 Summary of Experimental and Detector Systematic Uncertainties

1016

## CHAPTER 10

1017

# Conclusions

1018 **10.1 Higgs Results in Review**

1019 **10.2 Prospects for Future**

---

## Bibliography

---

- 1021 [1] S. L. Glashow, *Partial-symmetries of weak interactions*, Nucl. Phys. **22** (1961) no. 4, 579. [2.1.1](#)
- 1022 [2] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264. [2.1.1](#)
- 1023 [3] A. Salam and J. C. Ward, *Gauge theory of elementary interactions*, Phys. Rev. **136** (1964)  
763–768. [2.1.1](#)
- 1024 [4] S. Weinberg, *Non-abelian gauge theories of the strong interactions*, Phys. Rev. Lett. **31** (1973)  
494–497. [2.1.1](#)
- 1025 [5] D. J. Gross and F. Wilczek, *Ultraviolet behavior of non-abelian gauge theories*, Phys. Rev. Lett.  
**30** (1973) 1343–1346. [2.1.1](#)
- 1026 [6] G. 't Hooft and M. Veltman, *Regularization and renormalization of gauge fields*, Nuclear  
Physics B **44** (1972) 189 – 213. [2.1.1](#)
- 1027 [7] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964)  
508. [2.1.2](#)
- 1028 [8] P. W. Higgs, *Spontaneous symmetry breakdown without massless bosons*, Phys. Rev. **145** (1966)  
1156. [2.1.2](#)
- 1029 [9] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*,  
Phys. Rev. Lett. **13** (1964) 321–322. [2.1.2](#)
- 1030 [10] The ALEPH, CDF, DØ, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak  
Working Group, the Tevatron Electroweak Working Group, and the SLD electroweak and  
heavy flavour groups, *Precision Electroweak Measurements and Constraints on the Standard  
Model*, CERN-PH-EP-2010-095 (2010) , [arXiv:1012.2367 \[hep-ex\]](#). [2.1.3](#), [2.2](#)
- 1031 [11] M. Baak, M. Goebel, J. Haller, A. Hoecker, D. Kennedy, R. Kogler, K. Mnig, M. Schott, and  
J. Stelzer, *The electroweak fit of the standard model after the discovery of a new boson at the  
LHC*, The European Physical Journal C **72** (2012) no. 11, .  
<http://dx.doi.org/10.1140/epjc/s10052-012-2205-9>. [2.1.3](#)
- 1032 [12] J. C. Collins, D. E. Soper, and G. Sterman, *Factorization for short distance hadron-hadron  
scattering*, Nuclear Physics B **261** (1985) 104 – 142. [2.2](#)
- 1033 [13] CERN, . CERN, Geneva, 1984. [2.2](#)

- 1048 [14] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and  
 1049 R. Tanaka (Eds.), *Handbook of LHC Higgs Cross Sections: 2. Differential Distributions*,  
 1050 CERN-2012-002 (CERN, Geneva, 2012) , [arXiv:1201.3084 \[hep-ph\]](https://arxiv.org/abs/1201.3084). **2.2**
- 1051 [15] *Updated coupling measurements of the Higgs boson with the ATLAS detector using up to 25*  
 1052 *fb<sup>-1</sup> of proton-proton collision data*, Tech. Rep. ATLAS-CONF-2014-009, CERN, Geneva, Mar,  
 1053 2014. **2.2.1**
- 1054 [16] CMS Collaboration Collaboration, *Precise determination of the mass of the Higgs boson and*  
 1055 *studies of the compatibility of its couplings with the standard model*, Tech. Rep.  
 1056 CMS-PAS-HIG-14-009, CERN, Geneva, 2014. **2.2.1**
- 1057 [17] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of the Higgs boson mass from*  
 1058 *the  $H \rightarrow \gamma\gamma$  and  $H \rightarrow ZZ^* \rightarrow 4\ell$  channels with the ATLAS detector using 25 fb<sup>-1</sup> of pp*  
 1059 *collision data*, [arXiv:1406.3827 \[hep-ex\]](https://arxiv.org/abs/1406.3827). **2.2.1**
- 1060 [18] *Evidence for the spin-0 nature of the Higgs boson using {ATLAS} data*, Physics Letters B **726**  
 1061 (2013) no. 13, 120 – 144.  
 1062 <http://www.sciencedirect.com/science/article/pii/S0370269313006527>. **2.2.1**
- 1063 [19] S. Dawson, A. Gritsan, H. Logan, J. Qian, C. Tully, et al., *Working Group Report: Higgs*  
 1064 *Boson*, [arXiv:1310.8361 \[hep-ex\]](https://arxiv.org/abs/1310.8361). **2.2.2**
- 1065 [20] O. Eberhardt, G. Herbert, H. Lacker, A. Lenz, A. Menzel, et al., *Impact of a Higgs boson at a*  
 1066 *mass of 126 GeV on the standard model with three and four fermion generations*,  
 1067 Phys.Rev.Lett. **109** (2012) 241802, [arXiv:1209.1101 \[hep-ph\]](https://arxiv.org/abs/1209.1101). **2.2.2**
- 1068 [21] M. Carena, S. Gori, N. R. Shah, C. E. Wagner, and L.-T. Wang, *Light Stops, Light Staus and*  
 1069 *the 125 GeV Higgs*, JHEP **1308** (2013) 087, [arXiv:1303.4414](https://arxiv.org/abs/1303.4414). **2.2.2**
- 1070 [22] N. Arkani-Hamed, K. Blum, R. T. D’Agnolo, and J. Fan, *2:1 for Naturalness at the LHC?*,  
 1071 JHEP **1301** (2013) 149, [arXiv:1207.4482 \[hep-ph\]](https://arxiv.org/abs/1207.4482). **2.2.2**
- 1072 [23] D. Carmi, A. Falkowski, E. Kuflik, and T. Volansky, *Interpreting LHC Higgs Results from*  
 1073 *Natural New Physics Perspective*, JHEP **1207** (2012) 136, [arXiv:1202.3144 \[hep-ph\]](https://arxiv.org/abs/1202.3144). **2.2.2**
- 1074 [24] G. Degrassi, S. Di Vita, J. Elias-Miro, J. R. Espinosa, G. F. Giudice, et al., *Higgs mass and*  
 1075 *vacuum stability in the Standard Model at NNLO*, JHEP **1208** (2012) 098, [arXiv:1205.6497](https://arxiv.org/abs/1205.6497)  
 1076 [\[hep-ph\]](https://arxiv.org/abs/1205.6497 [hep-ph]). **2.2.2**
- 1077 [25] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) no. 08, S08001. **3.1**
- 1078 [26] T. S. Pettersson and P. Lefevre, *The Large Hadron Collider: conceptual design.*, Tech. Rep.  
 1079 CERN-AC-95-05 LHC, CERN, Geneva, Oct, 1995. <https://cdsweb.cern.ch/record/291782>.  
 1080 **3.1**
- 1081 [27] T. Linnecar et al., *Hardware and Initial Beam Commissioning of the LHC RF Systems.*  
 1082 [oai:cds.cern.ch:1176380](https://cds.cern.ch/record/1176380), Tech. Rep. LHC-PROJECT-Report-1172.  
 1083 CERN-LHC-PROJECT-Report-1172, CERN, Geneva, Oct, 2008.  
 1084 <https://cdsweb.cern.ch/record/1176380>. **3.1**
- 1085 [28] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3**  
 1086 (2008) S08003. **3.1**

- 1087 [29] The CMS Collaboration, *The CMS experiment at the CERN LHC*, *Journal of Instrumentation*  
 1088 **3** (2008) no. 08, S08004. [3.1](#)
- 1089 [30] The LHCb Collaboration, *The LHCb Detector at the LHC*, *Journal of Instrumentation* **3** (2008)  
 1090 no. 08, S08005. [3.1](#)
- 1091 [31] The ALICE Collaboration, *The ALICE experiment at the CERN LHC*, *Journal of*  
 1092 *Instrumentation* **3** (2008) no. 08, S08002.  
<http://stacks.iop.org/1748-0221/3/i=08/a=S08002>. [3.1](#)
- 1093 [32] A. Team, *The four main LHC experiments*, Jun, 1999. [3.1](#)
- 1094 [33] ATLAS Collaboration Collaboration, G. Aad et al., *Search for  $H \rightarrow \gamma\gamma$  produced in association*  
 1095 *with top quarks and constraints on the Yukawa coupling between the top quark and the Higgs*  
 1096 *boson using data taken at 7 TeV and 8 TeV with the ATLAS detector*, [arXiv:1409.3122](#)  
 1097 [[hep-ex](#)]. [5](#)
- 1098 [34] *Search for the Standard Model Higgs boson produced in association with top quarks and decaying*  
 1099 *to  $b\bar{b}$  in  $pp$  collisions at  $\sqrt{s} = 8$  TeV with the ATLAS detector at the LHC*, *Tech. Rep.*  
 1100 *ATLAS-CONF-2014-011*, CERN, Geneva, Mar, 2014. [5](#)
- 1101 [35] ATLAS Collaboration, G. Aad et al., *Improved luminosity determination in  $pp$  collisions at  $\sqrt{s}$*   
 1102 *= 7 TeV using the ATLAS detector at the LHC*, *Eur.Phys.J.* **C73** (2013) 2518, [arXiv:1302.4393](#)  
 1103 [[hep-ex](#)]. [6.1.1](#)
- 1104 [36] CERN, . CERN, Geneva, 2012. [6.1.1](#)
- 1105 [37] ATLAS Collaboration Collaboration, G. Aad et al., *The ATLAS Simulation Infrastructure*,  
 1106 *Eur.Phys.J.* **C70** (2010) 823–874, [arXiv:1005.4568](#) [[physics.ins-det](#)]. [6.2](#)
- 1107 [38] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A Simulation toolkit*,  
 1108 *Nucl.Instrum.Meth.* **A506** (2003) 250–303. [6.2](#)
- 1109 [39] LHC Higgs Cross Section Working Group Collaboration, S. Heinemeyer et al., *Handbook of*  
 1110 *LHC Higgs Cross Sections: 3. Higgs Properties*, [arXiv:1307.1347](#) [[hep-ph](#)]. [6.2.1](#), [9.1](#), [9.1](#)
- 1111 [40] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *ALPGEN, a generator*  
 1112 *for hard multiparton processes in hadronic collisions*, *JHEP* **0307** (2003) 001,  
 1113 [arXiv:hep-ph/0206293](#) [[hep-ph](#)]. [6.2.2](#)
- 1114 [41] F. Maltoni and T. Stelzer, *MadEvent: Automatic event generation with MadGraph*, *JHEP* **0302**  
 1115 (2003) 027, [arXiv:hep-ph/0208156](#) [[hep-ph](#)]. [6.2.2](#)
- 1116 [42] B. P. Kersevan and E. Richter-Was, *The Monte Carlo event generator AcerMC versions 2.0 to*  
 1117 *3.8 with interfaces to PYTHIA 6.4, HERWIG 6.5 and ARIADNE 4.1*, *Comput.Phys.Commun.*  
 1118 **184** (2013) 919–985, [arXiv:hep-ph/0405247](#) [[hep-ph](#)]. [6.2.2](#)
- 1119 [43] P. M. Nadolsky, H.-L. Lai, Q.-H. Cao, J. Huston, J. Pumplin, et al., *Implications of CTEQ*  
 1120 *global analysis for collider observables*, *Phys.Rev.* **D78** (2008) 013004, [arXiv:0802.0007](#) [[hep-ph](#)].  
 1121 [6.2.2](#)
- 1122 [44] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower*  
 1123 *simulations: the POWHEG method*, *JHEP* **0711** (2007) 070, [arXiv:0709.2092](#) [[hep-ph](#)]. [6.2.2](#)
- 1124 [45] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, et al., *Event generation with*  
 1125 *SHERPA 1.1*, *JHEP* **0902** (2009) 007, [arXiv:0811.4622](#) [[hep-ph](#)]. [6.2.2](#)

- 1127 [46] A. Martin, W. Stirling, R. Thorne, and G. Watt, *Parton distributions for the LHC*, *Eur.Phys.J.*  
1128 **C63** (2009) 189–285, [arXiv:0901.0002 \[hep-ph\]](#). **6.2.2, 8.1**
- 1129 [47] *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton*  
1130 *collision data*, Tech. Rep. ATLAS-CONF-2014-032, CERN, Geneva, Jun, 2014. **7.4**
- 1131 [48] *Preliminary results on the muon reconstruction efficiency, momentum resolution, and*  
1132 *momentum scale in ATLAS 2012 pp collision data*, Tech. Rep. ATLAS-CONF-2013-088, CERN,  
1133 Geneva, February, 2013. **7.5**
- 1134 [49] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti- $k(t)$  jet clustering algorithm*, *JHEP* **0804**  
1135 (2008) 063, [arXiv:0802.1189 \[hep-ph\]](#). **7.6**
- 1136 [50] M. Garzelli, A. Kardos, C. Papadopoulos, and Z. Trocsanyi,  *$t\bar{t}W^\pm$  and  $t\bar{t}Z$  Hadroproduction at*  
1137 *NLO accuracy in QCD with Parton Shower and Hadronization effects*, *JHEP* **1211** (2012) 056,  
1138 [arXiv:1208.2665 \[hep-ph\]](#). **8.1**
- 1139 [51] J. M. Campbell and R. K. Ellis,  *$t\bar{t}W^{+-}$  production and decay at NLO*, *JHEP* **1207** (2012) 052,  
1140 [arXiv:1204.5678 \[hep-ph\]](#). **8.1**
- 1141 [52] Campbell, John and Ellis, R. Keith and Röntsch, Raoul, *Single top production in association*  
1142 *with a Z boson at the LHC*, *Phys.Rev.* **D87** (2013) 114006, [arXiv:1302.3856 \[hep-ph\]](#). **8.1**
- 1143 [53] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of  $WZ$  production in*  
1144 *proton-proton collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector*, *Eur.Phys.J.* **C72** (2012)  
1145 **2173**, [arXiv:1208.1390 \[hep-ex\]](#). **8.2**
- 1146 [54] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of  $ZZ$  production in pp*  
1147 *collisions at  $\sqrt{s} = 7$  TeV and limits on anomalous  $ZZZ$  and  $ZZ\gamma$  couplings with the ATLAS*  
1148 *detector*, *JHEP* **1303** (2013) 128, [arXiv:1211.6096 \[hep-ex\]](#). **8.2**
- 1149 [55] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of the cross-section for  $W$*   
1150 *boson production in association with b-jets in pp collisions at  $\sqrt{s} = 7$  TeV with the ATLAS*  
1151 *detector*, *JHEP* **1306** (2013) 084, [arXiv:1302.2929 \[hep-ex\]](#). **8.2**
- 1152 [56] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of differential production*  
1153 *cross-sections for a  $Z$  boson in association with b-jets in 7 TeV proton-proton collisions with*  
1154 *the ATLAS detector*, [arXiv:1407.3643 \[hep-ex\]](#). **8.2**