

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

I. Joseph Kroll, Professor, Physics
Supervisor of Dissertation

Elliot Lipeles, Associate Professor, Physics
Graduate Group Chairperson

Dissertation Committee

Justin Khoury, Associate Professor, Physics

I. Joseph Kroll, Professor, Physics

Burt Ovrut, Professor, Physics

Evelyn Thompson, Associate Professor, Physics

I. 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	Acknowledgements	iii
64	Abstract	iv
65	Contents	v
66	Preface	ix
67	1 Introduction	1
68	2 Theoretical Background	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	3 The Large Hadron Collider and the ATLAS Experiment	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	3.2 The ATLAS Experiment	13
82	3.2.1 Detector Coordinate System	14

83	3.2.2	The Inner Detector	15
84	3.2.3	The Calorimeter	17
85	3.2.4	The Muon Spectrometer	20
86	3.2.5	The Trigger System	21
87	3.2.6	Reconstruction: Jets, Muons and Electrons	22
88	3.2.6.1	Tracks and Clusters	22
89	3.2.6.2	Electrons	22
90	3.2.6.3	Muons	23
91	3.2.6.4	Jets	23
92	3.2.6.5	B-Tagged Jets	24
93	4	Electrons	25
94	4.1	Electrons at Hadron Colliders	25
95	4.2	Reconstruction of Electron at ATLAS	26
96	4.3	Identification of Electrons at ATLAS	27
97	4.3.1	Pile-up	27
98	4.3.2	Trigger vs. Offline	27
99	4.3.3	2011 Menu	27
100	4.3.4	2012 Menu	27
101	4.3.5	Electron Likelihood	27
102	4.4	Measurement of Electron Efficiency at ATLAS	27
103	4.4.1	Techniques	27
104	4.4.2	Issues	27
105	5	$t\bar{t}H$ Analysis Summary	28
106	5.1	Signal Characteristics	28
107	5.2	Background Overview	29
108	5.3	Analysis Strategy	30
109	6	Dataset and Simulation	31
110	6.1	Data	31
111	6.1.1	The 2012 Dataset	31
112	6.2	Simulation	32
113	6.2.1	Signal Simulation	33

114	6.2.2 Background Simulation	33
115	7 Object and Event Selection	35
116	7.1 2l Same-Charge Signal Region	35
117	7.2 3l Signal Region	36
118	7.3 4l Signal Region	37
119	7.4 Electron Selection	37
120	7.5 Muon Selection	38
121	7.6 Jet and b-Tagged Jet Selection	38
122	7.7 Tau Selection	39
123	7.8 Object Summary and Overlap	39
124	7.9 Optimization	39
125	8 Background Estimation	41
126	8.1 Vector Boson (W^\pm, Z) production in association with top quarks: $t\bar{t}V, tZ$	41
127	8.1.1 $t\bar{t}Z$ Validation Region	44
128	8.2 Di-boson Background Estimation: $W^\pm Z, ZZ$	44
129	8.2.1 $W^\pm Z$ Uncertainty	45
130	8.2.2 ZZ Uncertainty	47
131	8.3 Charge-Misidentification Background	49
132	8.3.1 Likelihood Method	50
133	8.3.2 Results	52
134	8.3.3 Systematic and Statistical Uncertainties	52
135	8.4 Fake Lepton Backgrounds	54
136	8.4.1 2l SS Fakes	55
137	8.4.2 3l Fakes	59
138	8.4.3 4l Fakes	63
139	9 Summary of Systematic Uncertainties	65
140	9.1 Systematic Uncertainties on Signal Cross-section and Acceptance	65
141	9.2 Experimental and Detector Systematic Uncertainties	68
142	9.2.1 Lepton Identification, Energy Scale, and Trigger	68
143	9.2.2 Lepton Isolation and Impact Parameter	69
144	9.2.3 Jet Energy	71

145	9.2.4 B-Tagged Jet Efficiency	71
146	9.2.5 Summary	72
147	9.3 Summary of Background and Signal Normalization Uncertainties	72
148	10 Results and Statistical Model	74
149	10.1 Results in Signal Regions	74
150	10.2 Statistical Model	74
151	10.2.1 The Likelihood	74
152	10.2.2 Test Statistic and Profile Likelihood	75
153	10.2.3 CL_s Method	75
154	10.2.4 Exclusion Limits	76
155	10.2.5 μ Measurements	76
156	10.2.6 Nuissance Parameter Impact on the Signal Strength	77
157	11 Conclusions	78
158	11.1 Higgs Results in Review	78
159	11.2 Prospects for Future	78
160	Bibliography	79

Preface

162 This is the preface. Blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah
163 blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah blah blah
164 blah. Blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah
165 blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah blah
166 blah blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah blah blah.
167 Blah blah blah blah blah. Blah blah blah blah blah.

168 Blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah
169 blah. Blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah blah blah
170 blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah
171 blah blah blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah blah
172 blah. Blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah blah
173 blah blah.

175

CHAPTER 1

176

Introduction

CHAPTER 2

Theoretical Background

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

187 **2.1 The Standard Model**

188 **2.1.1 The Standard Model Structure**

189 The Standard Model (SM) [1, 2, 3, 4] is an example of a quantum field theory that describes the
 190 interactions of all of the known fundamental particles. Particles are understood to be excitations of
 191 the more fundamental object of the theory, the field. The dynamics and interactions of the fields are
 192 derived from the Standard Model Lagrangian, which is constructed to be symmetric under transfor-
 193 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
 194 weak iso spin, and $U(1)$ is the group for weak hyper-charge.

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

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

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

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

210 So far, 3 separate generations of both quarks and leptons have been discovered, differing only by
211 mass. The gluon and the 4 electroweak bosons have also been discovered (W^+ , W^- , Z^0 , and γ) ¹. The
212 reason for this 3-fold replication is not known.

213 2.1.2 Electroweak Symmetry Breaking and the Higgs

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

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

1The 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).

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

236 2.1.3 The Standard Model Parameters

237 Although the SM structure is set by specifying the gauge groups, the EWSB mechanism, and ac-
 238 knowledging the 3-fold replication of the fermions, confronting the SM with experiment requires the
 239 measurement of 17² free parameters, which are unconstrained from the theory. These free parameters
 240 include the fermion masses (from the Yukawa coupling), the force coupling constants, the angles and
 241 phase of the mixing between quarks, and constants from the Higgs and electroweak sector³.

242 Experiments have provided a number of measurements of the parameters of the SM[10]. Prior
 243 to the discovery of the Higgs boson (discussed later) the Higgs mass, however, was the only fully
 244 unconstrained parameter, although its value could be inferred via its involvement in loop corrections
 245 on the top mass (M_t) and the W mass (M_W). The GFitter collaboration assembles all relevant
 246 electroweak observable measurements into a statistical model and then allows certain measurements to
 247 float within their uncertainty to allow for a fit among multiple correlated measurements measurements
 248 [11]. Figure 2.1 shows the fitted constraints on 4 key SM parameters (M_H , M_W , M_t , $\sin^2\theta_w$) with
 249 actual measurements overlaid. The addition to the fit of the measured Higgs mass from the ATLAS
 250 and CMS collaborations creates a small tension, as the other observables prefer the mass to be much
 251 lower (~ 80 GeV/ c^2). The tension in the combined electroweak fit (including the Higgs) is not
 252 statistically significant with a p -value of 0.07.

253 2.2 Collider Physics and the Higgs

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

²There 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

³ 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

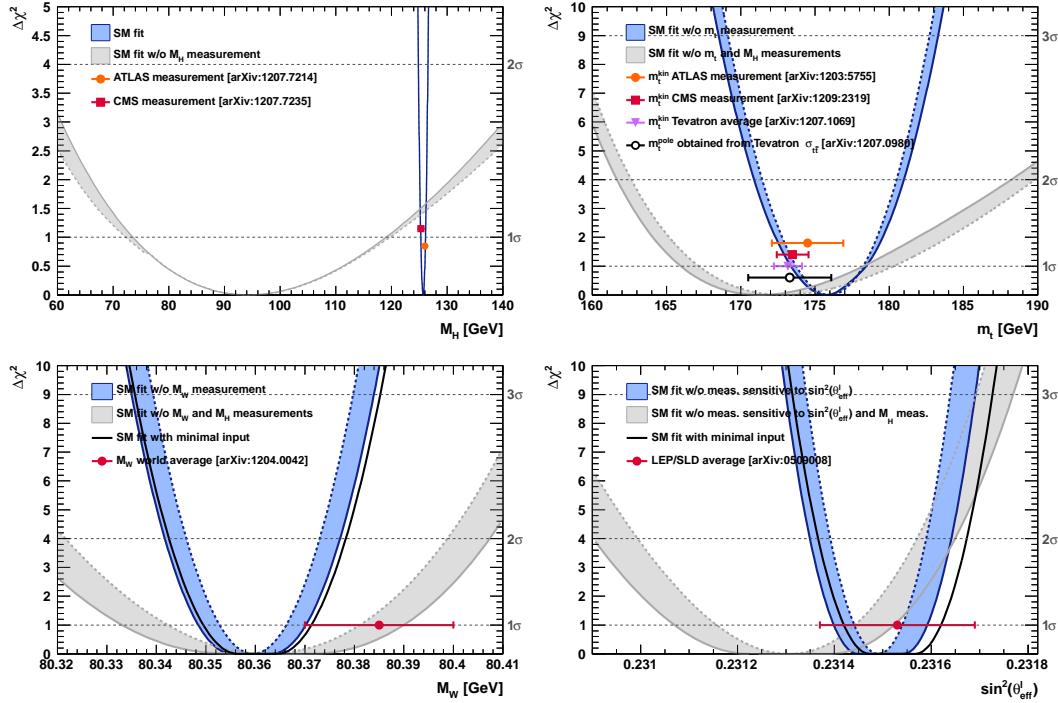


Figure 2.1: χ^2 as a function of the Higgs mass (top left), the top quark mass (top right), the W boson mass (bottom left) and the effective weak mixing angle (bottom right) for the combined SM fit from the GFitter group. The data points placed along $\chi^2 = 1$ represent direct measurements of the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

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

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

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

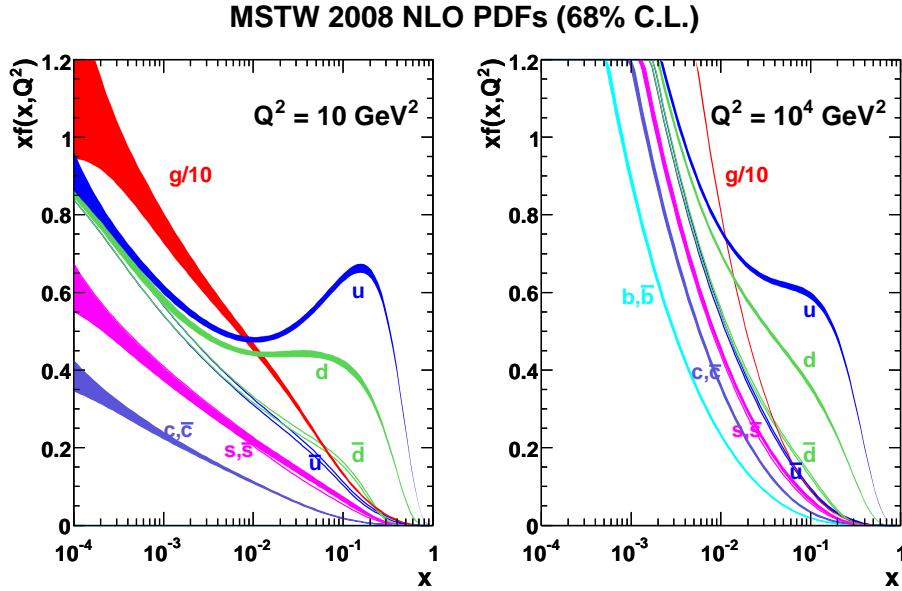


Figure 2.2: Proton Parton Distribution Functions (PDFs) form the MSTW Collaboration at $Q^2 = 10 \text{ GeV}^2$ and $Q^2 = 10^4 \text{ GeV}^2$

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.

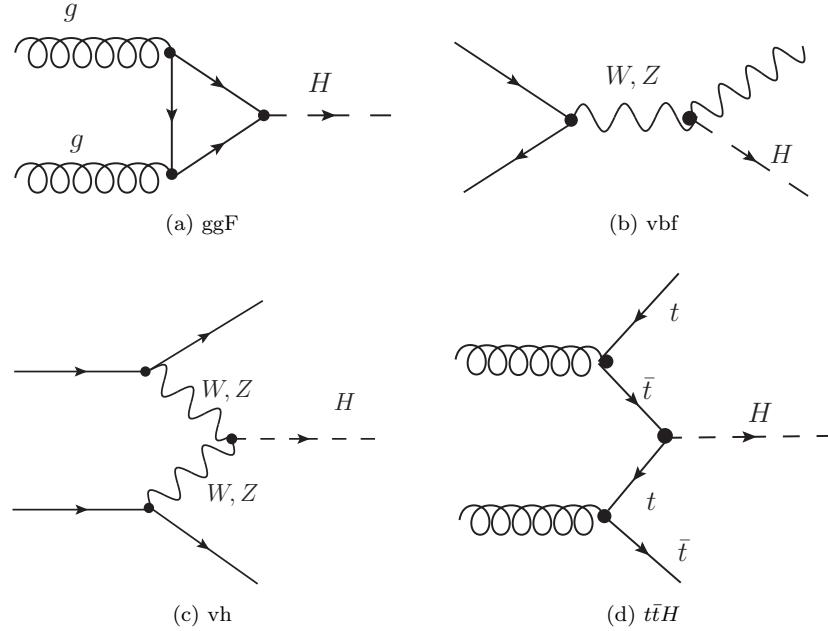


Figure 2.3: Dominant Higgs production modes at the LHC

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-

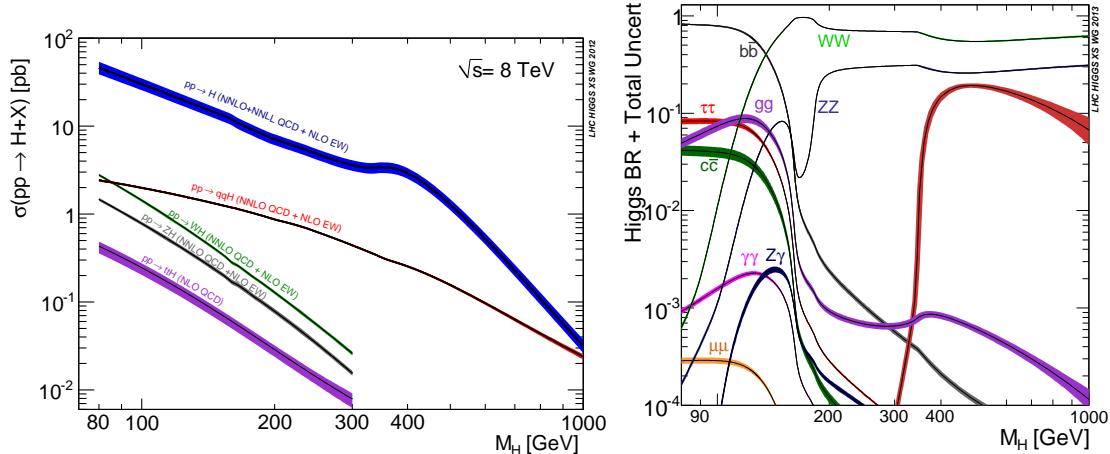


Figure 2.4: 8 TeV LHC Higgs production cross-sections (left) and decay branching fractions

quark, W,Z, and τ 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.

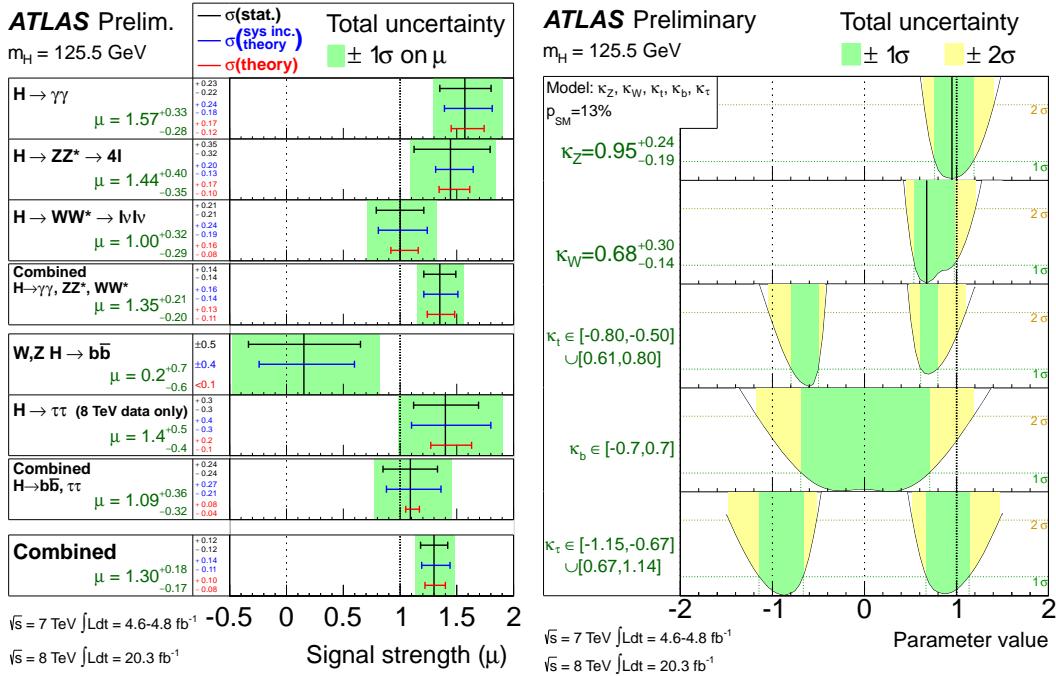


Figure 2.5: ATLAS Higgs combination results for all SM measurement channels as ratios of the measured to SM production cross-sections (left) and extracted Higgs coupling constraint scale-factors for a combined fit to the measurement channels, where the W, Z , top-quark, b -quark, and τ couplings are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

320 Aside from the loop effects, measurement of the $t\bar{t}H$ production cross-section would provide a
 321 precise measurement of the top Yukawa coupling. When compared with the measured top quark mass,
 322 this tests the Higgs mass generation properties for up-type quarks. Despite similar uncertainties on
 323 the overall production cross-sections for $t\bar{t}H$ and the gluon-fusion modes, both of which depend on the
 324 top Yukawa, most of these uncertainties would cancel for $t\bar{t}H$ if normalized to the topologically similar
 325 $t\bar{t}Z$. Finally, the uniqueness of the experimental signature means that searches for $t\bar{t}$ signatures can
 326 be performed for a variety of Higgs decays ($\gamma\gamma, b\bar{b}, WW, ZZ$, and $\tau\bar{\tau}$) with roughly similar degrees of
 327 sensitivity (within a factor of 10)[19].

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

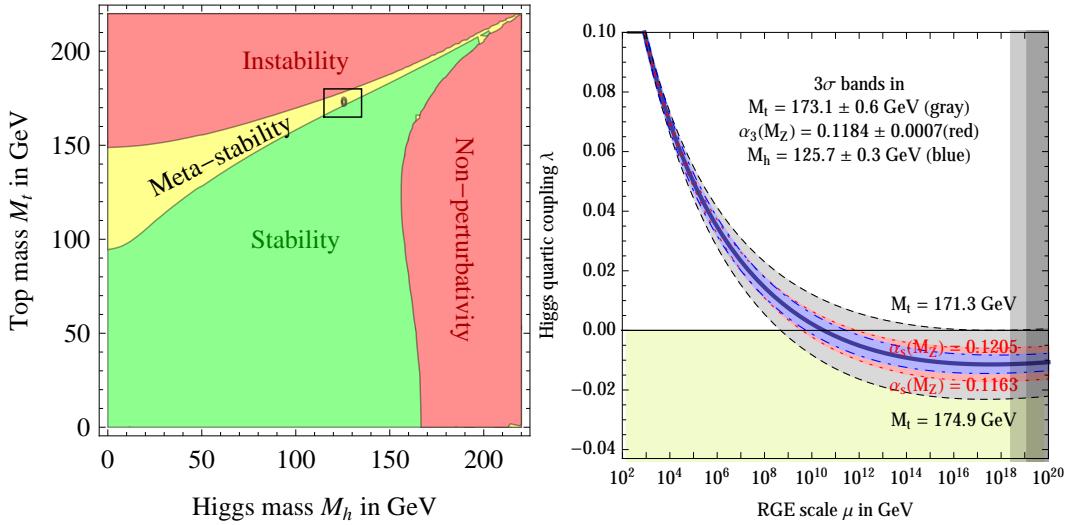


Figure 2.6: RGE for the running of the SM parameter, λ for the Higgs self-coupling term with present values and uncertainty bands for M_H and M_t (left). The two-dimensional plot colored (right) shows regions for which the SM is stable, unstable and metastable based on this RGE

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 island⁴.

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.

⁴The RGE assumed that there is no new physics at all energy scales

351

CHAPTER 3

352

The Large Hadron Collider and the ATLAS Experiment

353

354

3.1 The Large Hadron Collider

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

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

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

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

369

3.1.1 The Accelerator Complex

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

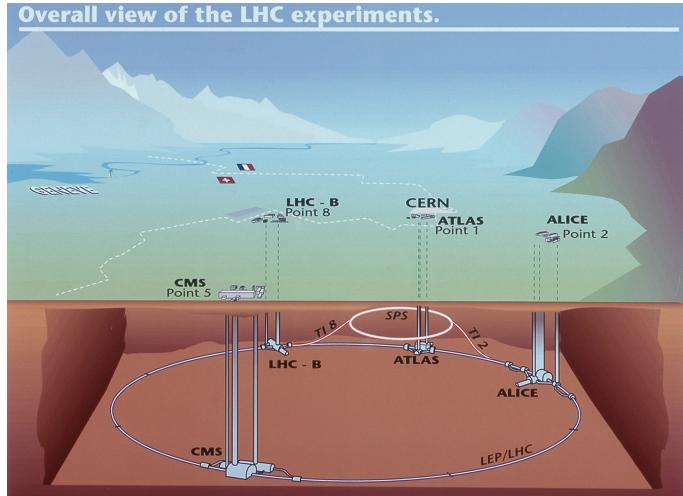


Figure 3.1: Diagram of the Large Hadron collider and location of the 4 main experiments (ATLAS, CMS, LHCb, and ALICE). Around the ring. The diagram also shows the location of the SPS, the final booster ring the accelerator complex that accelerates the protons to 450 GeV before injection into the LHC.

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

376 3.1.2 Beam Parameters and Collisions

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

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

391 3.2 The ATLAS Experiment

392 This section provides a brief overview of the ATLAS experiment. The ATLAS detector is centered on
393 one of the LHC collision points, located 100m underground. Through the combination of a number of
394 subsystems, it designed to identify the particle content arising from these collisions, measure energy
395 and momentum of these particles, and make fast decisions about the content of each collision, in
396 order to save a small fraction of measured collision events for offline study.

397 ATLAS, shown in Figure 3.2, possesses cylindrical symmetry around the beam pipe. It weights
398 7000 tons, has a diameter of roughly 25m and length of 46m. ATLAS was designed to be a multi-
399 purpose hermetic, particle detector, able to identify many types of particles, and designed to provide
400 a snapshot of the entire collision event. The detector sub-systems form concentric rings around the
401 beam-line at increasing distance. From closest to the beam outward, they are:

- 402 • **Inner Detector:** The inner detector (ID)[33, 34] is immersed in a solenoidal magnetic field[35]
403 and provides measurements of charge particle tracks, through three subsystems: the Pixel
404 Detector[36, 37], the Semi-Conductor Tracker (SCT)[?, ?], and Transition Radiation Tracker
405 [?, ?, ?]. and is capable of making particle
- 406 • **Calorimeter:** The calorimeters measures the energy of particles that participate in the elec-
407 tromagnetic (photons, electrons) and hadronic interactions (pions, protons, neutrons, etc.).
408 The hermeticity of the calorimeters allows for missing transverse energy measurements. The
409 calorimeter is composed of the liquid argon electromagnetic calorimeter (LAr)[38] and the
410 hadronic tile calorimeter[39], the liquid argon hadronic endcap calorimeter, and the forward
411 calorimeters

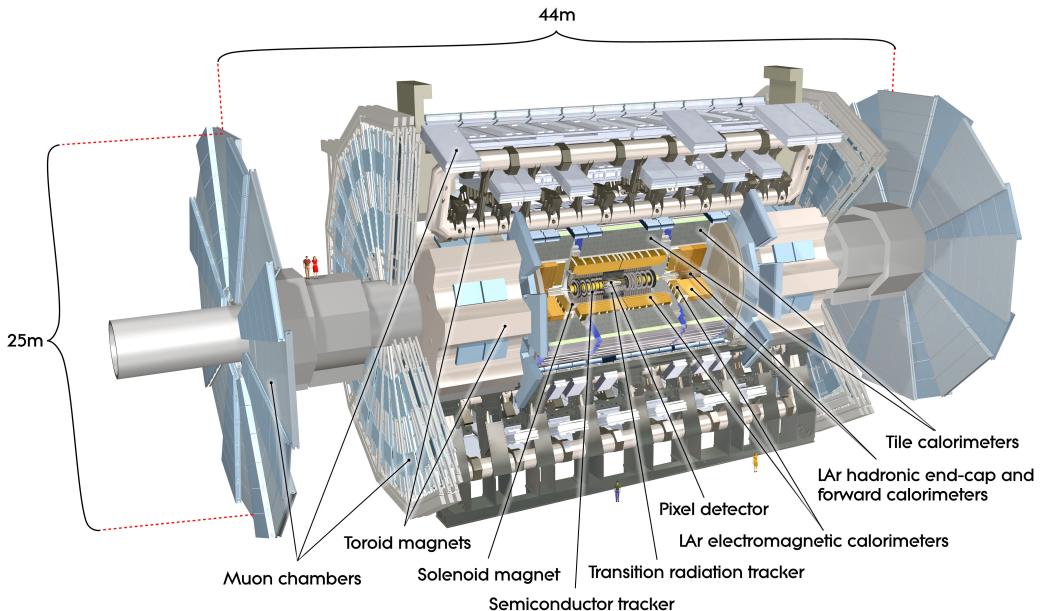


Figure 3.2: Diagram of the ATLAS detector and subsystems

- **Muon Systems** The muon systems(MS)[40] form the outermost detector systems and measure the momentum of minimum ionizing muon tracks, as all other particles are stopped by the calorimeters. The muon systems are immersed in a toroidal magnetic field [35] and are composed of 4 different sub-systems for triggering and tracking measurements [41, 42, 43].
- **Triggering Systems:** The trigger and data aquisition systems[44, 45] read out data from the detector through a three-tiered hardware and software decision making framework to record the most interesting physical processes for a broad physics analysis program.

These systems are discussed in depth in the following sections.

3.2.1 Detector Coordinate System

ATLAS uses a right-handed coordinate system centered at the nominal proton interaction point. The beam line defines the z -axis, and the $x - y$ plane is perpendicular to the beam line. The $x - y$ plane is referred to as the transverse plane. The transverse plane holds special significance in reporting measurements, because the initial momentum of the hard collision system is 0 along the transverse plane in the laboratory rest frame. The momentum of the colliding proton-proton system is also 0 along the z -axis but the colliding partons may have vastly different momenta. Thus, momentum of

427 the hard colliding system along the z -axis differs collision to collision. Particle momenta measured
 428 along the transverse plane is called transverse momenta, and labeled p_T .

429 Because ATLAS possesses nominal cylindrical symmetry, cylindrical and polar coordinates are
 430 used to describe particle trajectories and detector positions. The radial coordinate, r , describes
 431 transverse distances from the beam line. An azimuthal angle, ϕ , describes angles around the z -axis,
 432 and a polar coordinate θ describes angles away from the z -axis. The polar angle is often expressed in
 433 terms of pseudo-rapidity, defined as $\eta = -\ln(\tan(\theta/2))$. Distances in $\eta - \phi$ space are often used to
 434 describe the proximity of objects in the detector, $\Delta R = \sqrt{\eta^2 + \phi^2}$.

435 The 'Barrel' and 'Endcap' are classifications that are often used to label the position of sub-
 436 detectors. Barrel subdetectors occupy positions more central to the detector at $|\eta|$ values roughly less
 437 than 1-2, while the endcap calorimeters extend farther in $|\eta|$. The barrel-endcap transition region
 438 contains detector services and the orientation of the detector elements may change.

439 3.2.2 The Inner Detector

440 The ID makes measurement of the position of charged particles as they move through the detectors 3
 441 sub-systems (ID,TRT,SCT). The individual position measurements can be strung together to form a
 442 particle track. The entire ID is immersed in a 2T solenoidal magnetic field allowing for measurements
 443 of particle momenta through curvature of the tracks. The ID is contained with a radius of 1.15 m and
 444 has a total length of 7m, allowing for particle tracking out to $|\eta| < 2.5$. Figures 3.3 and 3.4 show the
 445 placement of the ID sub-systems in the $R - \phi$ and $R - z$ planes.

446 The Pixel detector has 80 million silicon read out channels (pixels) and is closest to the interaction
 447 point with the finest granularity. As charged particles traverse the silicon, they create electron-hole
 448 pairs, which subsequently drift in an electric field and can be captured and registered as a current
 449 pulse. The Pixel detector has three concentric layers of pixels in the barrel (to $|\eta| < 1.9$) and three
 450 endcap disks on each side of the barrel (to $|\eta| < 2.5$). The closest barrel layer to the beam pipe is
 451 called the b-layer. The pixels provide excellent hit resolution ($R - \phi$ accuracy of 10 μm and $z(R)$
 452 accuracy of 115 μm in the barrel (endcap)).

453 The SCT uses similar silicon technology to the pixels, but the each SCT layer contains a double
 454 layer of silicon strips, which are much longer in length than width. The double layers are inclined
 455 slightly with respect to each other so that these 1D sensors have 2D resolution for coincident hits of
 456 580 μm in $z(R)$ for the barrel(endcap) and 17 μm in $R - \phi$. The SCT has 4 million read out channels
 457 and it comprises 4 barrel layers and 9 endcap layers with coverage to $|\eta| < 2.5$.

458 The TRT is comprised of straw drift tubes filled with a gas mixture, containing mostly Xenon

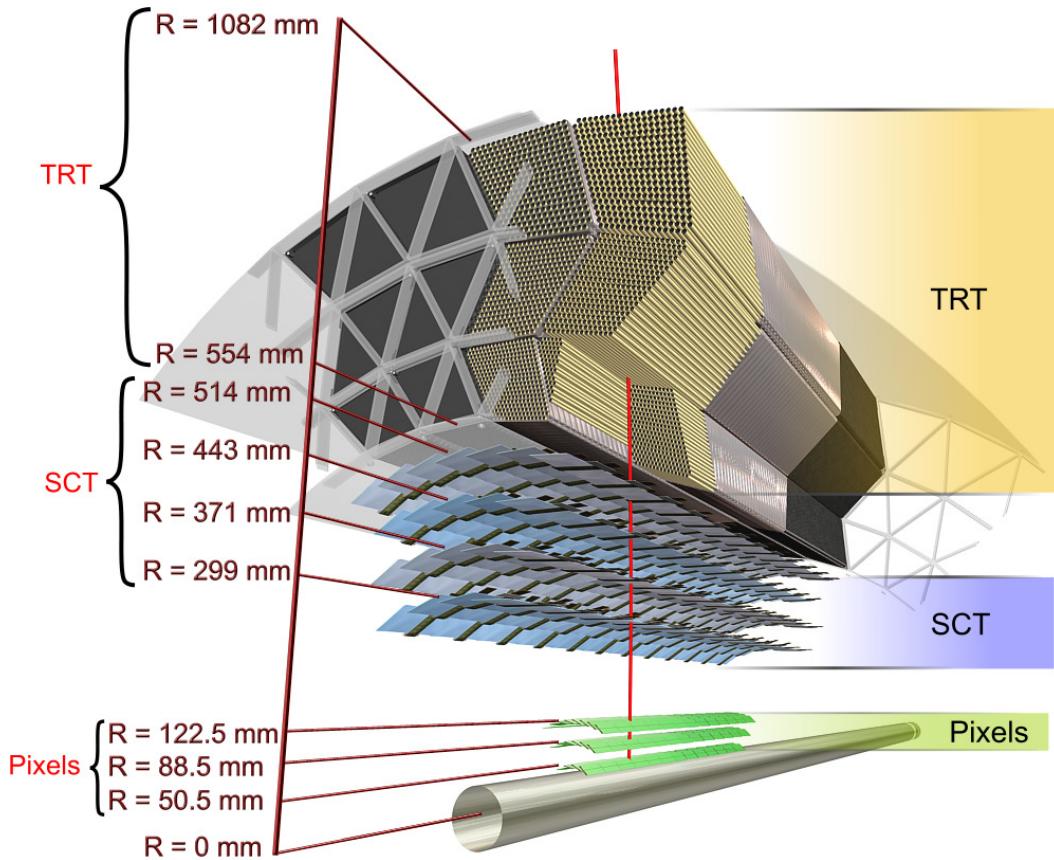


Figure 3.3: Diagram of the ATLAS ID in the $R - \phi$ plane showing the barrel view of the Pixel, SCT, and TRT detectors.

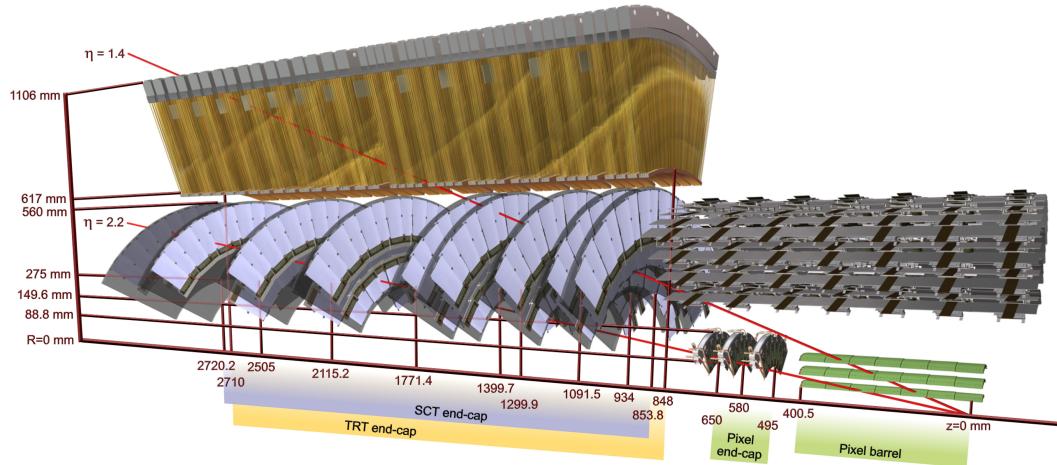


Figure 3.4: Diagram of the ATLAS ID in the $R - z$ plane showing the endcap view of the Pixel, SCT, and TRT detectors. Only one side of the encap is shown.

459 gas. Particles traversing the straws ionize the gas, and the liberated electrons drift to a wire at the
460 center of the straw, which has an applied voltage, and induce a signal on the wire. The TRT has
461 $\sim 300,000$ straws. The barrel straws are arranged cylindrically along the z direction out to $\sim \eta < 1$
462 and the endcap straws point radially outward in the R direction. For this reason, the barrel(endcap)
463 straws provide no measurement in the $R(z)$ directions. The drift tubes provide individual position
464 measurements with resolutions of $\sim 130 \mu$. Each particle track has on average a large number, 35,
465 hits.

466 The TRT is unique in that it also provides particle identification measurements via transition
467 radiation. Charged particles emit transition radiation, when traversing a boundary between materials
468 of different dielectric constants. The volume between the straws is, therefore filled with a radiator
469 material, a polymer foild or foam, to provide this boundary condition. Transition radiation photons
470 are emitted in the direction of the particle trajectory in the keV range and cause a much larger signal
471 amplitude within the straw. Hits that cause a signal at a higher threshold are thus indicative of
472 transition radiation. The probability for emission transition radiation depends on the relativistic γ
473 of the tranversing particle. Because eletrons are much lighter than any other charged particle, their
474 γ -factors tend to be high enough in the GeV range to induce transition radiations, as opposed to
475 pions, muons and other particles.

476 Combined tracking of particles through the 3 sub-detectors results track momentum measurements
477 from 500 MeV, the mimnium energy need to leave the ID due to the magnetic field, and a few TeV.
478 The track p_T resolution is roughly $0.05\% \cdot p_T + 1\%$.

479 3.2.3 The Calorimeter

480 The ATLAS calorimeters measure the energy of electron, photons and hadrons with $|\eta| < 4.5$. They
481 induce a particle shower via electro-magnetic and nuclear interactions with the detector material and
482 contain enough radiation lengths to ensure that all or most of the shower energy remains contained.
483 Muons are minimum ionizing particle that do not partipate in the strong interaction and therefore
484 pass through the ATLAS calorimeters leaving relatively little energy behind. ATLAS calorimeters are
485 sampling calorimeters meaning that the active material of the detector only measures a small fraction
486 of the energy produced by the shower and the overall shower energy is inferred from this fractional
487 measurement. The rest of the material is inactive,heavy dense material, designed to induce showers.
488 The calorimetry system is grossly divided longitudinally (radially) into electro-magnetic(EM) and
489 then hadronic segments, operated with different technologies. Figure 3.5 diagrams the layout of the
490 calorimeter system.

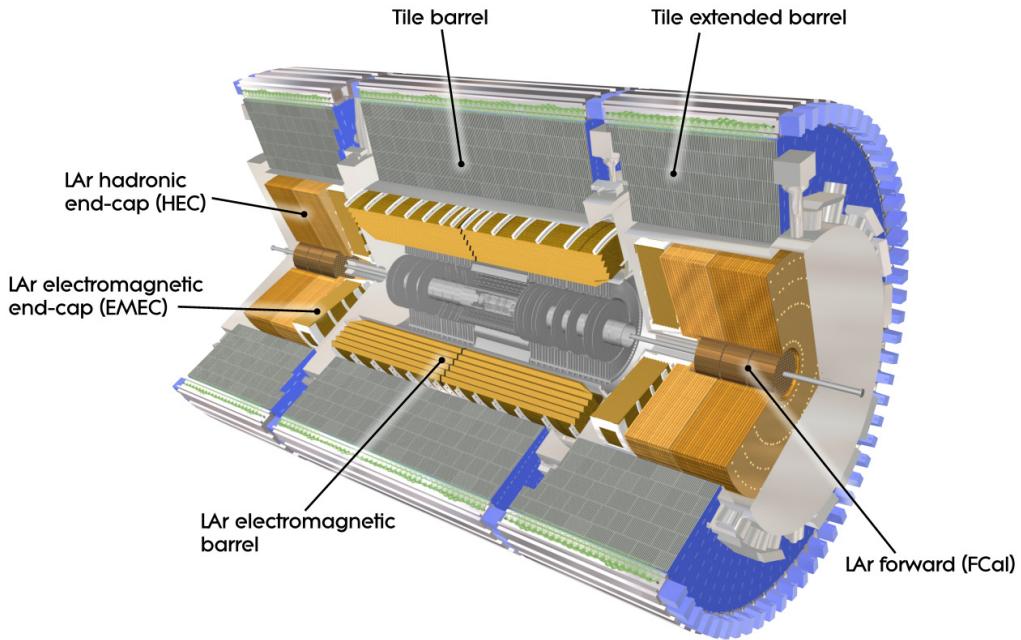


Figure 3.5: Diagram of the ATLAS calorimeters

491 The EM calorimeter, which is located directly outside of the solenoid magnet but within the
 492 same cryostat, has an accordian design with lead absorber and liquid argon (LAr) active material. The
 493 accordian design ensures uniform coverage in ϕ . The barrel and endcap LAr extend to $|\eta| < 2.47$. The
 494 LAr provides highly granular measurements in $\eta - \phi$ with 4 longitudinal segments, totalling $\sim 25\text{--}35$
 495 radiation lengths with the exception of the barrel/endcap transition region ($1.37 < |\eta| < 1.52$). The
 496 geometry of the barrel LAr calorimeter can be seen in Figure ???. The first longitudinal segment
 497 is called the pre-sampler, composed of a thin layer of active liquid argon, designed to detect early
 498 particle showers. The second segment is the most highly granular segment called the 'strips', as it is
 499 composed of thin strips. The strips have a size of $0.025/8 \times 0.1$ in $\eta - \phi$ in the barrel with similar
 500 sizes in the endcap designed to be able to resolve single and double particle showers. This resolution
 501 is particularly useful in distinguishing $\pi^0 \rightarrow \gamma\gamma$ signatures from electron and photon signature. The
 502 bulk of the radiation lengths and therefore the primary energy measurement come from the the third
 503 layer⁵. Each cell in this layer is 0.025×0.025 in $\eta - \phi$. The final layer is coarser in thinner and designed
 504 to estimate energy leaking out of the EM calorimeter. The forward EM calorimeters extend the η
 505 range and use the same technology, but are not used in this analysis. The energy resolution of the

⁵this layer is actually called 'layer 2', since the pre-sampler is referred to as 'layer 0'

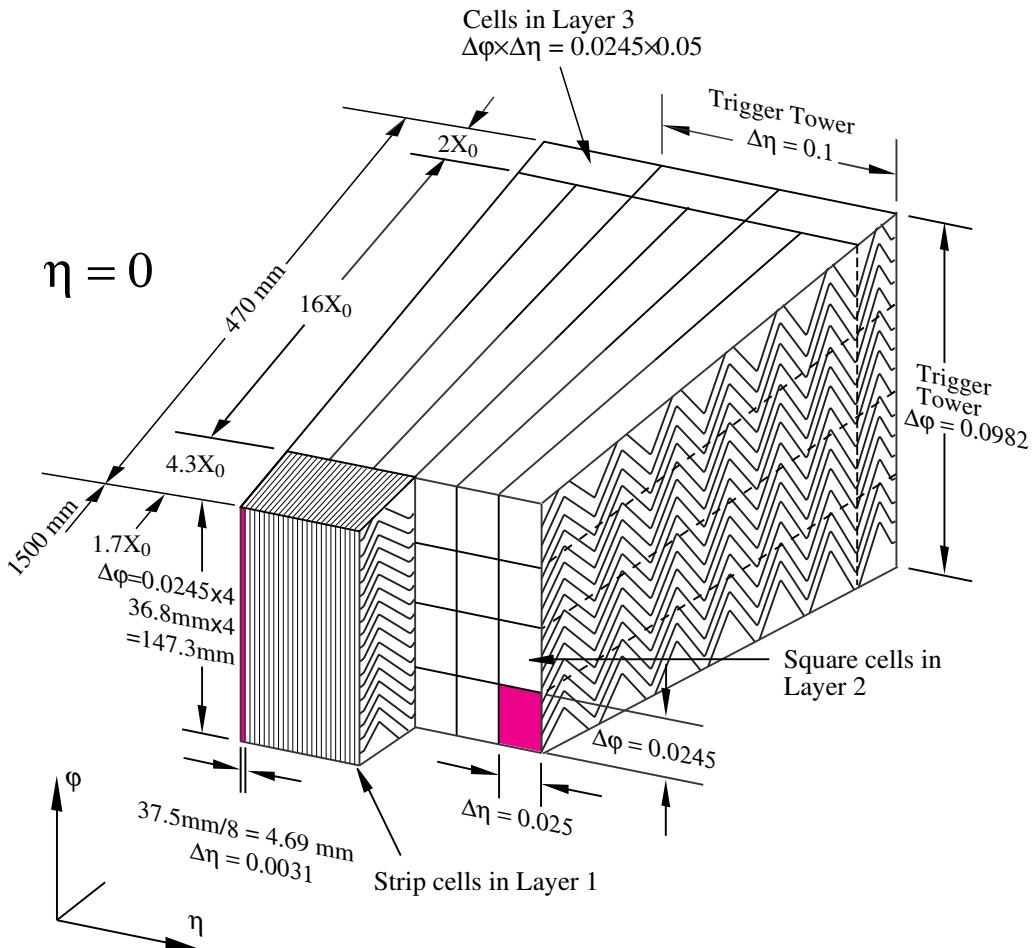


Figure 3.6: Diagram of the ATLAS LAr EM calorimeter showing the longitudinal segmentation and the $\eta - \phi$ cells for the central barrel region

506 EM calorimeters is $\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$, measured in test beam data and confirmed in collision
507 data.

508 The hadronic calorimeter is located directly behind the EM calorimeter and composed of tiles
509 of iron absorber and plastic scintillator in the barrel ($|\eta| < 1.6$), called the TileCal, and copper-
510 liquid argon in the endcap ($1.5 < |\eta| < 3.2$), called the HEC. The calorimeters contain $\sim 10\text{-}19$
511 hadronic interactions lengths with multiple longitudinal segments to contain showers induced by
512 the nuclear interaction of hadronic particles. The energy resolution of the hadronic calorimeters is
513 $\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$. The intrinsic resolution of hadron calorimeters is much worse than electro-
514 magnetic calorimeters, because much of the inelasticity of the nuclear interactions.

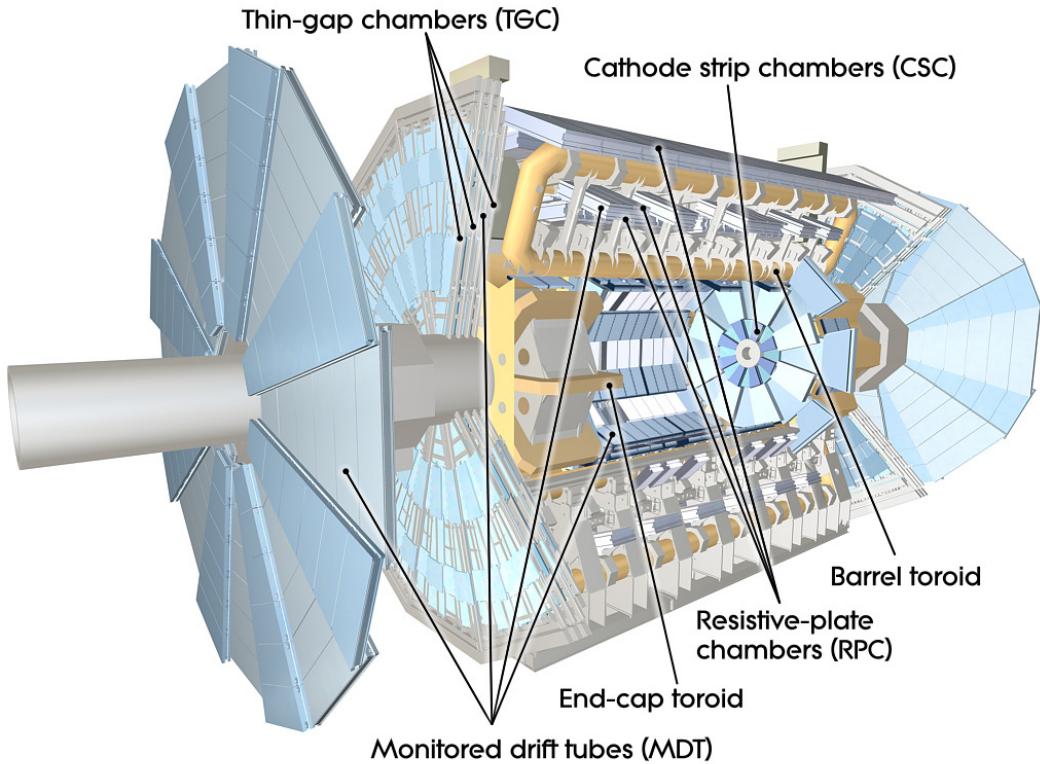


Figure 3.7: Diagram of the ATLAS muon system

515 3.2.4 The Muon Spectrometer

516 The muon spectrometer (MS) measure the trajectory of particles outside of the calorimeters, using
 517 multiple different technologies. All charged particles except for muons are stopped by the calorimeter,
 518 and therefore the majority of partilces in the MS are muons, wiht the exception of rare cases of
 519 hadronic punch-through. Particle momentum spectroscopy is made possible by an air-core toroidal
 520 magnet system, imbedded in the MS in the barrel ($|\eta| < 1.4$), and two smaller end cap toroids that
 521 provide fields out to $|\eta| < 2.7$.

522 In the barrel region, the muon chambers are arranged in three cylindrical layers around the beam,
 523 while the in the endcap-regions the the layers are arranged perpendicular to the beam in wheels. The
 524 arrangement is depicted in Figure 3.7.

525 The chambers in the barrel and most of the endcap are constructed from Monitored Drift Tubes
 526 (MDTs) with an Argon gas mixture. Each chamber contains 3-7 drift tubes and provide hit resolutions
 527 of $80 \mu\text{m}$ per tube and $35 \mu\text{m}$ per chamber in the bending plane. For $|\eta| > 2.0$, Cathode Strip
 528 Chambers (CSCs) are used, primarily to handle the higher incident particle flux. They are composed

529 of cathod strips crossed with anode wires in the gas mixture, but use similar drift technology as the
530 MDTs and have resolutions in the bending plane $40 \mu\text{m}$ per chamber.

531 Separate chambers, called Thin Gap Chambers (TGCs), used in the endcaps, and Resistive Plate
532 Chambers (RPCs), used in the barrel, provide less precise hit information but within a much quicker
533 time window, and are therefore used for triggering, as the CSCs and MDTs are too slow for the level-1
534 trigger.

535 **3.2.5 The Trigger System**

536 The ATLAS trigger system is designed to make quick decisions about individual particle collisions
537 to reduce the enormous collision rate of 20 MHz to a much more managable 400 Hz to be stored for
538 offline analysis. Saving the full ATLAS datastream would require space for 40 TB of raw data per
539 second, but, more importantly, most of these collisions result in the unintersting inlastic break-up of
540 the colliding protons. To select out collisions to allow for a diverse physics program, ATLAS devotes
541 a large portion of the bandwidth to general purpose single lepton triggers. The presence of leptons in
542 the event indicates the presence of the weak or electro-magnetic interaction and therefore occurs at
543 many order of magnitude less frequently then interactions involving the strong intereaction. Moreover,
544 many interesting physics signatures that are analyzable by ATLAS involve leptonic final stats. The
545 remaining bandwith is allocated to jet, missing energy, tau, and unbiased supporting triggers.

546 The ATLAS trigger system is composed of 3 levels: level-1 (L1), level (L2), and the event filter(EF).
547 The first level is hardware only trigger that reduces the input 20 MHz rate to $\sim 75 \text{ kHz}$, or 1 out of
548 every 250 events. The available buffering means that the decisions need to be made within $2.5 \mu\text{s}$,
549 which would not be possible with software. The L1 selection is based on calorimeter clustering and
550 tracking finding in the MS for small areas of the detector called regions-of-interst (ROIs). It selects
551 ROIs with high p_T muon candidates.

552 The second and third stages L2 and EF are software based. The L2 algorithms perform more
553 detailed object reconstruction for leptons, jets and photons inside of the ROIs provided by L1, by
554 perforeng tracking and in depth calorimeter clustering algorithms. The decisions are made within 50
555 ms per event and pass 1 out of every 15 events to the EF. At the EF, events undergo full reconstruction
556 using similar but faster versions of the algorithms used offline. The EF makes decisions on the
557 presenece of fully id-objects in the event, multi-object events and event topological quantities within
558 4s to reduce the L2 output by a factor of 10. The events that pass this stage are then written to tape
559 for offline study.

560 **3.2.6 Reconstruction: Jets, Muons and Electrons**

561 Physicsts analyze the collision event as a collection of identified objects, expressed a momentum 4-
 562 vectors. These objects arise from the final state particles in the event, which can be combined and
 563 counted to infer properties of the hard scatter. The particles that make detectable signatures are
 564 those that are stable enough to pass through the detector: muons, electrons, photons, and quarks and
 565 gluons. Jets and b-tagged jets, muons and electrons, are used in the $t\bar{t}H$ analysis to define our search
 566 regions and to separate the Higgs signal from backgrounds. Other analyses may used photons, taus
 567 and missing energy⁶, but these are not discussed in depth here. Figure 3.8 shows an $R - \phi$ schematic
 568 of the interaction of various particle signatures in the ATLAS detector.

569 **3.2.6.1 Tracks and Clusters**

570 The process of converting the disparate detector signatures and signals into a unified 4-momentum
 571 description of individual objects is called Reconstruction. The basic components of reconstruction are
 572 sensor measurments, or hits, in trackers (ID, MS) and energy measurments in the calorimeter. Hits
 573 in the ID and MS undergo pattern recognition, which identifies hits that belong to a single track, and
 574 fitting, which fits a curve to the track to assess the particle trajectory. Charged particle trajectories
 575 are generally helical in a magnetic field, but the fitting algorithm takes into more detailed information
 576 about energy loss to material along the tracks length. The result of the fitting is an estimation of the
 577 particles momentum 3-vector. ATLAS track reconstruction is discussed in depth here []. Electron,
 578 photon, and hadronic particles leave clustered deposits of energy in the EM and hadronic calorimeters
 579 from their showers. Electron and photon showers are primarily contained with the EM calorimeter,
 580 while hadronic showers are deposit most of their energy in the hadronic calorimeters. The process
 581 of assocaiting individual read-out cells of energy in the calorimeter to clusters of ennergy from the
 582 showers of individual particles is called clustering. ATLAS clustering algorithms are discussed in
 583 depth here ???. From these basic, peices, tracks and clusters, more complex objects can be created.

584 **3.2.6.2 Electrons**

585 Electrons leave both a track in the ID and a narrow and isolated cluster of energy in the EM calorime-
 586 ter, $\Delta R < 0.1$. Electron reconstruction proceeds using a sliding window algorithm, which scans a
 587 fixed size rectangle in $\eta - \phi$ space over the EM calorimeter cells to find relative maxima of enery in
 588 the window[46]. These maxima seed the clustering algorithms. Because electrons are light, they both

⁶missing energy is the presence of momentum imbalance in the transverse plane of the calorimeter due to escaping neutrinos

589 lose energy to the material gradually via scattering and more catastrophically through the emission of
590 a high energy photon, through interaction with the nuclear. This process is called bremmstrahlung.
591 Tracks for electrons are reconstructed uniquely because they must include the hypothesis that the
592 electron loss energy through bremmstrahlung. Generally, the emitted photon is contained within the
593 same energy cluster and therefore the sliding window algorithm is always wider in the direction of
594 bending, ϕ . A single track is then matched to the cluster within certain minimum matching require-
595 ments in η, ϕ , and p_T . Electrons are distinguished from photon conversion,s which also have a track,
596 by association with conversion vertices, found with a dedicated algorith.

597 Electron have many lever arms for further identification to suppress backgrounds from fake sources.
598 The narrowness of the shower shape, quality of track, and presence of transition radiation are used
599 by cut-based and multivaraiate algorithms are used by indeitification algorithms. This is dicussed in
600 depth in Chatper ???. Electrons are reliably reconstructed and identified with energies above 7 GeV.

601 3.2.6.3 Muons

602 Muons are reconstructed from a combination of ID and MS tracks when possible. The two tracks must
603 meet matching criteria to ensure they are from the same particle and the muon momentum 3-vector
604 comes from the combined ID,MS fit. Muons leave little energy in the calorimeters and are generally
605 isolation from other particle, when produced from electro-weak bosons. Identification algoritms make
606 requirments on the number of tracking hits in the ID and MS and the quality of the matching of the
607 two tracks. Muons are reliably reconstructed and identified with energies above 5 GeV. More about
608 muon reconstruction and identification can be found here [47].

609 3.2.6.4 Jets

610 Quarks and gluons are colored objects that cannot exist alone. When emitted, they undergo a process
611 called hadronization, in which the convert into 'jets' of colorless hadrons that emerge collimated from
612 the interaction point. The majority of these hadrons are charged and neutral pions, though other
613 hadrons are often present. Jets are reconstructed using conglomerations of calorimeter energy clusters
614 chosen via anti- k_t algorithm, with a radius of $\Delta R < 0.4$ [48]. The algorithm has been shown to
615 be infrared safe, meaning the jet quantities are not sensitive to low energy, small angle radiative
616 divergences. Jets at ATLAS are reconstructed from 10 GeV, calibration of the energy scale and
617 resolution are only available for energies greater than 20-25 GeV.

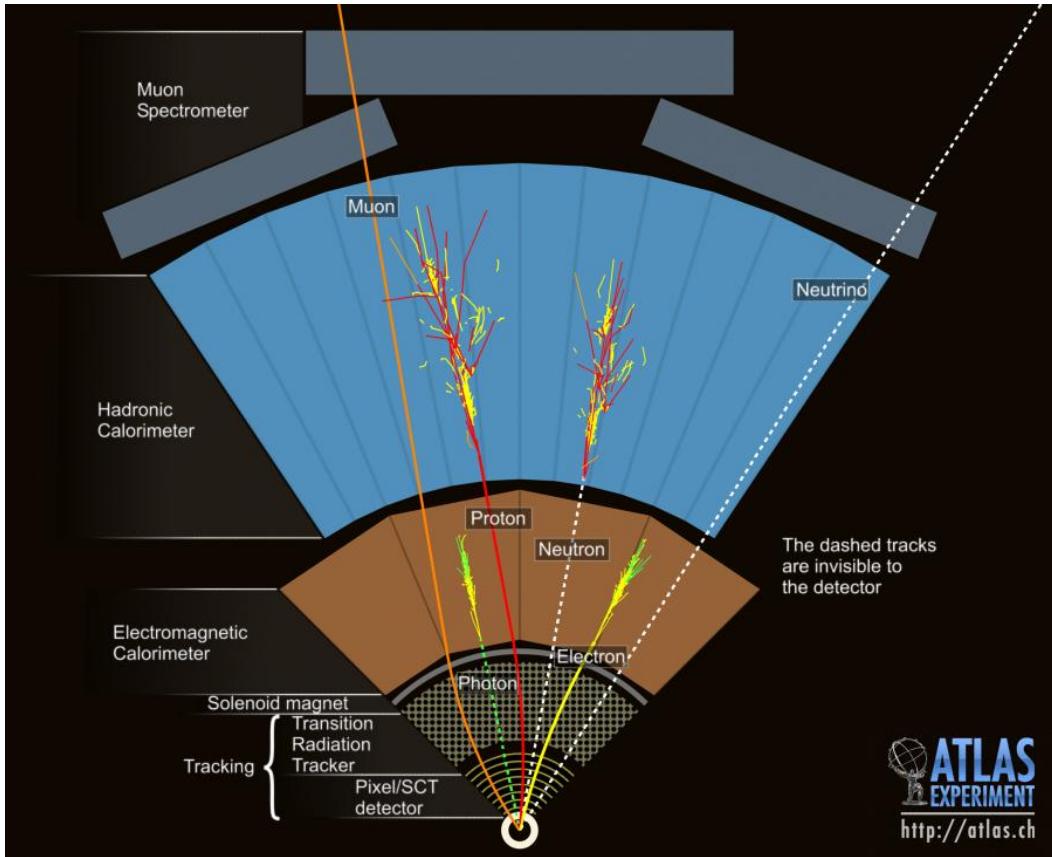


Figure 3.8: $R - \phi$ schematic of the ATLAS detector and various particle signatures

618 3.2.6.5 B-Tagged Jets

619 Generally, the flavor of the initiating quark is not known from the reconstructed jet, although gluon
 620 initiated jets and quark initiated jets have slightly different properties. B-quark jets, however, are
 621 unique in that the long life-time of the produced b-mesons allow for measureable decays in flight.
 622 This property is used to tag b-quark initiated jets. This analysis used the MV1 tagging algorithm [?],
 623 which is a neural network based algorithm that looks for secondary displaced decay vertices inside
 624 the event and takes into account jet track parameters and energy flow with respect to these vertices.
 625 B-quark jets often involve b-meson decays to leptons, especially muons, which can be used to tag an
 626 orthogonal b-jet sample for studying tagging efficiencies.

CHAPTER 4

Electrons

629 4.1 Electrons at Hadron Colliders

630 High energy electron signatures are important elements of searches and measurements at hadron col-
631 liders. The overwhelming majority of collisions that deposit energy in the detectors are the result of
632 strong-force mediate interactions of the constituent partons. These collisions result in the production
633 of high energy jets in the detector. Figure XX shows the cross-sections of various processes as a
634 function of the center of mass energy of the collision. Physics involving the electroweak interaction
635 or even strong production of massive states occur many orders of magnitude less frequently than the
636 total inelastic cross-section. Interesting physics signatures, both standard model and beyond, often
637 involve the production of light leptons as a result of the decay of massive particles. Choosing events
638 that have high energy electrons or muons targets events that contain electroweak vertices and dramat-
639 ically reduce the background from the more copiously produced strong physics. Electron and muon
640 energy and momenta are also relatively well-measured compared to jets. This allows for the use of
641 well-resolved kinematic shapes used to discriminate the signatures of different processes in analyses.

642 At ATLAS, the primary datasets for most analyses are collected with electron and muon triggers.
643 Electron triggers are particularly important, because the muon trigger system has a 20% smaller
644 acceptance than the electrons. The challenge in identifying electrons is distinguishing the production of
645 electrons from direct production of W and Z decays from electrons produced in the more copiously
646 produced b-meson decays, fake-electron signatures from rare jet fragmentations into charged and
647 neutral pions, and photon conversions in the inner detector. The identification of electrons, the
648 precise measurement of the identification efficiency, and the measurement of the rate of fake electron
649 signatures lead are often the most important and challenging pieces of an analysis. The following
650 sections discuss the reconstruction and identification of electrons for the primary electron trigger and

651 offline physics analyses as well as the measurement of the electron identification efficiency in 2012.
652 Because I had a major role in these projects, I will at times discuss their historical evolution and not
653 simply focus of the particular measurement relevant to the $t\bar{t}H$ analysis.

654 **4.2 Reconstruction of Electron at ATLAS**

655 High-energy electrons are charged particles that produced both a track in the inner detector and
656 cluster of energy in the calorimeter. The electron reconstruction algorithm is discussed in detail in
657 XXXX. Electrons are seeded by their clusters of energy in the liquid argon calorimeter (LAr). A
658 sliding window algorithm uses rectangular regions in ϕ and η of a fixed number of calorimeter cells
659 and translates in both directions to maximize the amount of energy in window. For the barrel and
660 endcap calorimeters ($|\eta| < 2.47$, the windows are 3x7 and 5x5 cles respectively ($\eta \times \phi$). They are
661 wider in phi to allow for energy smearing in phi due to bremsstrahlung and the curvature of electron
662 track in the magnetic field. The exact cell sizes are give in section XXXX. Windows containing at
663 least $E_T < 2.5$ GeV of energy in the second layer of the calorimeter.

664 To distinguish from photons, which undergo the same seeding algorithm (albeit with a different
665 window size), electron energy cluster must match to a reconstructed track within a window of 0.050.10
666 ($\eta \times \phi$) and with loosely matching calorimeter energy and track momentum $E/p \pm 10$. Photons, however,
667 often convert in the inner detector material. To resolve the ambiguity, electrons that match to tracks
668 that are already identified as part of a conversion vertex during the previously run conversion re-
669 construction are removed from the electron candidate container.

670 Material in front of the calorimeter makes electron reconstruction difficult due to electron energy
671 loss through bremsstrahlung. Standard track fitting is accomplished via a least squares fit to linear-
672 lized helical model with accounting for the accumulated scattering of pion or muon-like particle. This
673 sort of fitting fails to adequately describe the change in curvature from abrupt energy losses during
674 bremsstrahlung that occur for the much lighter electron. This results in non-insignificant inefficiencies
675 of matching tracks to electron clusters when the standard tracking algorithms are used for electron
676 candidates, especially lower energy electron candidates, whose tracks bend more in the magnetic field.
677 Moreover, as shown in Figure XXXX the material profile of the inner detector has a strong eta de-
678 pendence due to the TRT and SCT services. This results in large eta dependencies in the resolution
679 of electron track parameters.

680 For reconstruction of data collected in 2012, a new fitting technique was introduced that incorpo-
681 rated bremsstrahlung energy losses through a Gaussian Sum Filter (GSF) XXXX. Figures XXXXa and
682 XXXXb show the improvement in the electron reconstruction efficiency from 2011 to 2012, achieved

683 through the introduction of this filter and the improvement in a the track impact parameter resolution
684 for electron tracks. Because the fitting is cpu intensive, it is not run online during data collection.

685 Electrons are recalculated

686 **4.3 Identification of Electrons at ATLAS**

687 **4.3.1 Pile-up**

688 **4.3.2 Trigger vs. Offline**

689 **4.3.3 2011 Menu**

690 **4.3.4 2012 Menu**

691 **4.3.5 Electron Likelihood**

692 **4.4 Measurement of Electron Efficiency at ATLAS**

693 **4.4.1 Techniques**

694 **4.4.2 Issues**

695

CHAPTER 5

696

$t\bar{t}H$ Analysis Summary

697 This chapter provides an overview of the set of analyses searching for the Standard Model (SM)
 698 production of the Higgs boson in association with top quarks in multi-lepton final states with multiple
 699 jets (including b-quark tagged jets). Searches in $t\bar{t}H$ final states with 2 same-charge, 3 and 4 light
 700 leptons (e, μ) are discussed in depth. These final states target specifically Higgs decays to vector
 701 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
 702 final states targeting the $H \rightarrow b\bar{b}$ [49], $H \rightarrow \gamma\gamma$ [50], and $H \rightarrow \tau\tau$ decay modes.

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

708 Detailed description of the event and selection section are provided in Chapter 7, background
 709 modelling in Chapter 8, the effect of systematic errors and the statistical analysis in Chapter 9 and
 710 final results in Chapter 10.

711 **5.1 Signal Characteristics**

712 $t\bar{t}H$ can be observed in a number of different final states related to the Higgs boson and the top
 713 quark decay modes.

714 Three Higgs boson decays are relevant for this analysis: W^+W^- , $\tau^+\tau^-$ and ZZ . The top and
 715 anti-top quarks decay in $W^\pm b$. Each W^\pm boson decays either leptonically ($l=e^\pm, \mu^\pm, \tau^\pm$) with missing
 716 energy or hadronically. Table 5.1 provides the fractional contribution of the main Higgs decay modes
 717 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%

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

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

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

730 5.2 Background Overview

731 Background processes can be sorted into two categories:

- 732 • Events with a non prompt or a fake lepton selected as prompt lepton. These processes cannot
 733 lead to a final state compatible with the signal signature without a misreconstructed object. This
 734 category includes events with a prompt lepton but with misreconstructed charge⁷ and events
 735 with jets that "fake" leptons. These processes are rejected with tight object isolation and
 736 identification criteria, requiring a large jet multiplicity, and veto-ing events consistent with a
 737 leptonically decaying Z boson.

⁷Charge 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

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

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

745 **5.3 Analysis Strategy**

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

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

754 The channels are fed into a poisson model

755

CHAPTER 6

756

Dataset and Simulation

757 6.1 Data

758 6.1.1 The 2012 Dataset

759 The LHC successfully produced datasets for physics studies in 2010, 2011 and 2012. The 2012 proton-
760 proton dataset was delivered with collisions with a CME of 8 TeV with bunch collisions every 50 ns and
761 reached a total integrated luminosity of around 20 fb^{-1} [51]. Figure 6.1 shows the accumulation of this
762 dataset over time. Despite doubling the bunch spacing (thereby halving the bunch collision frequency),
763 the lumonisty neared the design lumonisty due to unexpected improvements in the transverse beam
764 profile[52]. This increased the amout of pile-up, or number of collisions per bunch crossing and in
765 general collision events were busier due to these multiple interactions. Figure 6.2 shows the average
766 number of interaction per bunch crossing for the 2011 and 2012 datasets. The 2012 dataset shows an
767 average of 20-25 interactions.

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

771 The datasets used in the analysis were collected with the primary electron (EF_e24vhi_medium1
772 — EF_e60_medium1) and muon triggers (EF_24i.tight — EF_36.tight). The electron triggers
773 require a electron with at least 25 GeV of calorimeter energy, passing the medium identification
774 requirement and loose tracking isolation. Above 60GeV, the isolation requirement is dropped and the
775 identification is loosened slightly. Similarly, the muon trigger requires a good inner detector track and
776 matching hits in the muon spectrometer, as well as loose tracking isolation, which also is dropped
777 about 36 GeVThe data sample must contain either a primary muon or primary electron trigger.

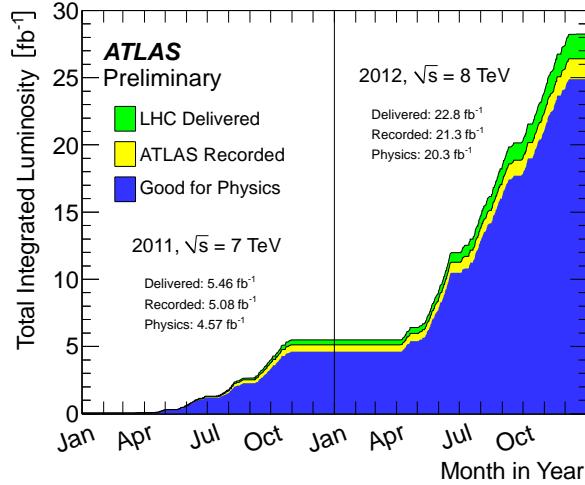


Figure 6.1: Plot showing the accumulation of the integrated luminosity delivered to the ATLAS experiment over 2011 and 2012. The rough size of the usable, physics ready dataset for 2012 is 20 fb^{-1} and is the dataset used for the following analysis.

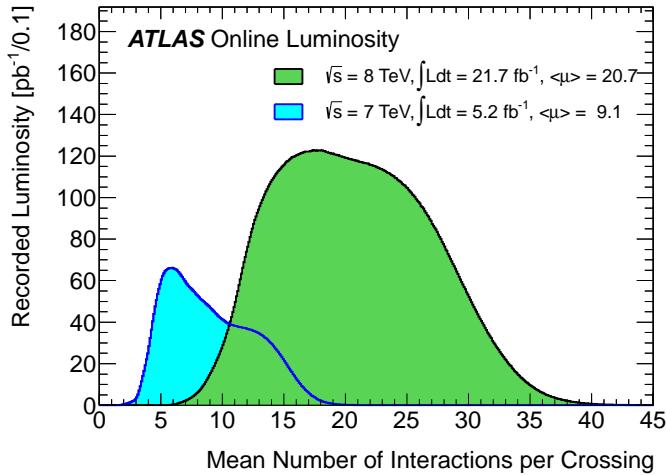


Figure 6.2: The average number of interactions per bunch-crossing for the 2012 and 2011 LHC proton-proton dataset. Most of these interactions are uninteresting but leave energetic signatures in particle detectors called pile-up which interfere with measurements

778 6.2 Simulation

779 Simulation samples based on are used to determine the overall event selection acceptance and efficiency
 780 and for investigations not directly involved in the final result. The simulated samples are created using
 781 parton distribution function (PDF) and model using Monte Carlo (MC) techniques the hard parton
 782 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

783 passed through a full ATLAS detector simulation[53] based on GEANT4 [54]. Small corrections are
 784 then applied to the overall efficiencies to re-scale object identification efficiencies, energy scales, and
 785 the pile-up, discussed in depth later.

786 6.2.1 Signal Simulation

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

795 The signal Monte Carlo samples are summarized in Table 6.1. These large samples are generated with
 796 inclusive Higgs boson decays, with branching fractions set to the LHC Higgs Cross Section Working
 797 Group (Yellow Report) recommendation for $m_H = 125$ GeV [55]. The inclusive cross section (129.3
 798 fb at $m_H = 125$ GeV) is also obtained from the Yellow Report [55].

799 6.2.2 Background Simulation

800 The background simulations used for this analysis are listed in Table 6.2. In general, the Alpgen[56],
 801 MadGraph[57], and AcerMC[58] samples use the CTEQ6L1[59] parton distribution function, while
 802 the Powheg[60], Sherpa[61], are generated with the CT10 PDF. The exception is the MadGraph $t\bar{t}t\bar{t}$
 803 sample, which is generated with the MSTW2008 PDF[62]. The highest order calculations available
 804 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

805

CHAPTER 7

806

Object and Event Selection

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

811 The cuts for each signal region are provided in Table 7.1 and the object selections are detailed in
812 the following selections. The selections are based on optimizations of the region sensitivity performed
813 using MC (event for data driven backgrounds) and adhoc values for normalization systematic errors.

814 All signal regions are comprised of three basic requirements: the presence of b-tagged jets, the presence
815 of additional light jets, and a veto of same flavor opposite sign leptons with an invariant mass within
816 the Z window. Additional requirements on the invariant mass of the leptons, the missing transverse
817 energy in the event, and the total object energy (H_T) proved to have negligible additional benefit at
818 our level of statistics.

819 **7.1 2l Same-Charge Signal Region**

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

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

827 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^8	=2	=3	=4
Lepton Charge Sum	+2 or -2	+1 or -1	0
Lepton Momentum (GeV/c^9)	$p_{T0} > 25$ $p_{T1} > 20$	$p_{T0} > 10$ $p_{T1,2} > 25, 20$	$p_{T0} > 25$ $p_{T1} > 20$ $p_{T2,3} > 10$
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 (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 μ , $\mu\mu$)	none	2 (No SFOS leps, SFOS leps)

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

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

835 7.2 3l Signal Region

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

- 838 • lep0: the lepton that is opposite in charge to the other two leptons
- 839 • lep1: the lepton that is closer in ΔR to lep0
- 840 • lep2: the lepton that is farther in ΔR from lep1

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

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

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

851 7.3 4l Signal Region

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

858 The four-lepton invariant mass is required to be between 100 and 500 GeV. This choice of mass
 859 window suppresses background from the on-shell $Z \rightarrow 4\ell$ peak and exploits the high-mass differences
 860 between the signal and the dominant $t\bar{t}Z$ background. Events containing an OS-SF lepton pair
 861 within 10 GeV of the Z boson mass are discarded. This Z-veto procedure greatly reduces background
 862 contributions from ZZ production as well as $t\bar{t}Z$ and while it also affects the signal by vetoing
 863 $H \rightarrow ZZ^*$, $Z \rightarrow \ell^+\ell^-$, these events constitute a small amount of the total expected signal. Finally,
 864 selected events are required to have at least two jets, at least one of which must be tagged as a b-quark
 865 initiated jet.

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

870 7.4 Electron Selection

871 The electrons are reconstructed by a standard algorithm of the experiment [46] and the electron
 872 cluster is required to be fiducial to the barrel or endcap calorimeters: $|\eta_{\text{cluster}}| < 2.47$. Electrons in
 873 the transition region, $1.37 < |\eta_{\text{cluster}}| < 1.52$, are vetoed. Electron reconstruction and identification
 874 is discussed in depth in Chapter ?? Electrons must pass the the VERYTIGHT likelihood identification
 875 criteria.

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

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

888 The electron selection is provided in Table 7.2.

889 7.5 Muon Selection

890 Muons used in the analysis are formed by matching reconstructed inner detector tracks with either
 891 a complete track or a track-segment reconstructed in the muon spectrometer (MS). The muons must
 892 satisfy $|\eta| < 2.5$. The muon track are required to be a good quality combined fit of inner detector
 893 hits and muon spectrometer segments, unless the muon is not fiducial to the inner detector, $|\eta| > 2.4$.
 894 Muons with inner detector tracks are further required to pass standard inner detector track hit
 895 requirements [47].

896 As with electrons, muons are required to be isolated from additional tracking or calorimeter energy:
 897 $ptcone20/P_t < 0.1$, $Etcone20/E_T < 0.1$) A cell-based $Etcone20/P_T$ relative isolation variable is used.
 898 A pile-up energy subtraction based on the number of reconstructed vertices in the event is applied.
 899 The subtraction is derived from a Z boson control sample.

900 The muons must also originate from the primary vertex and have impact parameter requirements,
 901 $d0$ significance < 3 , and $z0 \sin \theta < 0.1$ cm, similar to the electrons.

902 The muon selection is provided in Table 7.2.

903 7.6 Jet and b-Tagged Jet Selection

904 Jets are reconstructed in the calorimeter using the anti- k_t [48] algorithm with a distance parameter
 905 of 0.4 using locally calibrated topologically clusters as input (LC Jets).

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

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

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

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

918 **7.7 Tau Selection**

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

921 **7.8 Object Summary and Overlap**

922 **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	$E_{\text{T}}\text{Cone}20/p_T, p_{\text{T}}\text{Cone}20/p_T < 0.05$	
Jet overlap removal	$\Delta R > 0.3$	
$ d_0^{\text{sig}} $	$< 4\sigma$	
$z_0 \sin\theta$	$< 1 \text{ cm}$	
Muons		
p_T	$> 10 \text{ GeV}$	
$ \eta $	< 2.5	
ID	Tight	
Jet overlap removal	$\Delta R > 0.04 + 10 \text{ GeV}/p_T$	
Isolation	$E_{\text{T}}\text{Cone}20/p_T, p_{\text{T}}\text{Cone}20/p_T < 0.1$	< 0.05 for 2 leptons
$ d_0^{\text{sig}} $	$< 3\sigma$	
z_0	$< 1 \text{ cm}$	
Taus		
p_T	$> 25 \text{ GeV}$	
ID	Medium BDT	
Isolation		
Jet overlap removal	$\Delta R > 0.3$	
e/μ 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.

923

CHAPTER 8

924

Background Estimation

925 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
926 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)
927 and $t\bar{t}$ production with a jet misidentified as a lepton (fakes). The 2l SS regions possess a unique
928 background of charge misidentification from Z and top events. The methods for estimating these
929 backgrounds are discussed in this chapter. Monte Carlo simulation is used for the prompt $t\bar{t}V$ and
930 VV contributions. Systematic uncertainties on the overall normalization of these backgrounds in the
931 signal region are provided from theoretical studies and past ATLAS analyses and are verified in
932 data-based validation regions. The non-prompt backgrounds from $t\bar{t}$ jet-misidentification and charge-
933 misidentification are estimated using data-driven methods.

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

939 **8.1 Vector Boson (W^\pm, Z) production in association with top quarks:**

940 $t\bar{t}V, tZ$

941 This section describes the estimation and $t\bar{t}V$ productions. Production of top quarks plus vector
942 boson is an important background in all multilepton channels. A large part of the $t\bar{t}V$ component,
943 arising from on-shell $Z \rightarrow \ell\ell$, can be removed via a Z mass veto on like-flavour, opposite sign leptons.
944 However the $Z \rightarrow \tau\tau$ and γ^* components remain. The $t\bar{t}W^\pm$ and tZ processes generally require extra
945 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				
	≥ 5 jets		4 jets		3 leptons		Z enriched		
	$e^\pm e^\pm$	$e^\pm \mu^\pm$	$\mu^\pm \mu^\pm$	$e^\pm e^\pm$	$e^\pm \mu^\pm$	$\mu^\pm \mu^\pm$	$e^\pm e^\pm$	Z depleted	
$t\bar{t}H$	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	2.34 \pm 0.04	0.19 \pm 0.01	
$t\bar{t}V$	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

(μ_F) and renormalisation μ_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 6. $t\bar{t}W^\pm$ 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 ℓ 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 [63]: 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 [64]: calculate the PDF uncertainty using the MSTW2008nlo [62] PDF for the central value and then the final PDF uncertainty envelope is derived from three PDF error sets each with different α_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. [65]. 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.

967 **8.1.1 $t\bar{t}Z$ Validation Region**

968 Unlike $t\bar{t}W^\pm$, a $t\bar{t}Z$ validation region can be obtained by simply inverting the veto on same-flavor
969 opposite sign lepton pairs near the Z pole in the 3 lepton signal region. This region thus requires 3
970 leptons (with momentum and identification cuts discussed in Chapter 7, at least one opposite sign,
971 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
972 jet or exactly 3 jet and 2 or more b-tagged jets. The resulting region has low statistics and is not
973 used as a control region but is instead used as a validation to demonstrate that the normalization
974 uncertainty, discussed above, is properly evaluated.

975 The region defined by this is predicted to be 67% $t\bar{t}Z$, 17% WZ , and 13% tZ . We predict
976 19.3 ± 0.5 events and observe 28, giving a observed-to-predicted ratio of $1.45 \pm 0.27 \pm 0.03$ (where the
977 errors are from data and simulation statistics, respectively). Given the large errors, the region is still
978 in agreement with the predictions to within $1-1.5 \sigma$. Distributions of various variables are shown in
979 Fig. 8.1.

980 **8.2 Di-boson Background Estimation: $W^\pm Z, ZZ$**

981 $W^\pm Z$ and ZZ di-boson production with additional and b-tagged jets constitute small contributions
982 to the 3- and 4-lepton channels respectively. In the 3-lepton case $W^\pm Z$ comprises ~ 1 event of \sim
983 10 total background events while the ZZ contribution accounts for approximately 10% of the total
984 background in the 4-lepton channel. Because of the small size of these contributions, each of the
985 above processes can be assigned a non-aggressive uncertainty based on similar previous analyses with
986 ATLAS and cross-checked with data validation regions and MC truth studies. We assign an overall
987 50% error on both the $W^\pm Z$ 3-lepton signal region contribution and the ZZ 4-lepton signal region
988 contribution. The details of this error assignment are discussed below.

989 Both $W^\pm Z$ and ZZ production have been studied by ATLAS [66][67] but neither process has
990 been investigated thoroughly in association with multiple jets and b-quark jets. However, both $W + b$
991 [68] and $Z + b$ [69] production in 7 TeV data have been shown to agree with MC models to within
992 20-30%. A single W produced in association with b-tagged jets possesses a similar topology to the
993 $W^\pm Z + b$ process at a different energy scale and has been shown to be dominated by charm mis-tags
994 and b-jets from gluon splitting and multiple parton interaction. The $W + b$ analysis unfortunately
995 uses Alpgen MC with Herwig PS modeling and only provides results to 1 additional jet and therefore
996 is not directly applicable to this $t\bar{t}H$ analysis (where $W^\pm Z$ is modeled using Sherpa with massive
997 c and b quarks). $Z + b$ production originates from slightly different diagrams than $ZZ + b$ however

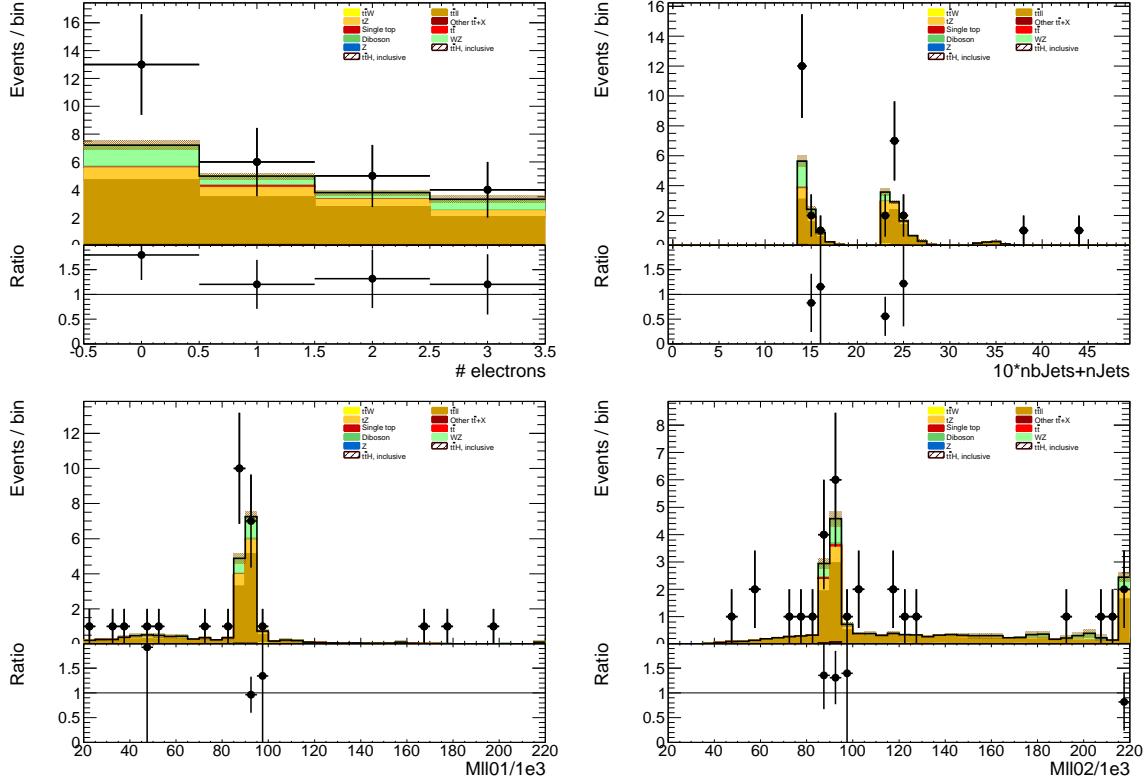


Figure 8.1: Data/MC comparison plots for $t\bar{t}Z$ control region A (≥ 4 jets, ≥ 1 b -tag and 3 jets, ≥ 2 b -tag). In all plots, the rightmost bin contains any overflows. Top left: number of electrons. Top right: 10^* the number of b -tags + the total number of jets. Middle left: the invariant mass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

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

1000 In the following two sections the uncertainty assignments for each of these two di-boson processes
 1001 will be reviewed in turn.

1002 8.2.1 $W^\pm Z$ Uncertainty

1003 The $t\bar{t}H$ analyses has two validation regions to test the Sherpa agreement with data for $W^\pm Z$: one
 1004 inclusive 3 lepton region, using the three-lepton channel object and p_T cuts; and a $W^\pm Z + b$ region
 1005 with 1 b -tagged jet, fewer than 4 jets (to remove $t\bar{t}V$), and a requirement that at least one same-flavor
 1006 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 ¹⁰. 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.

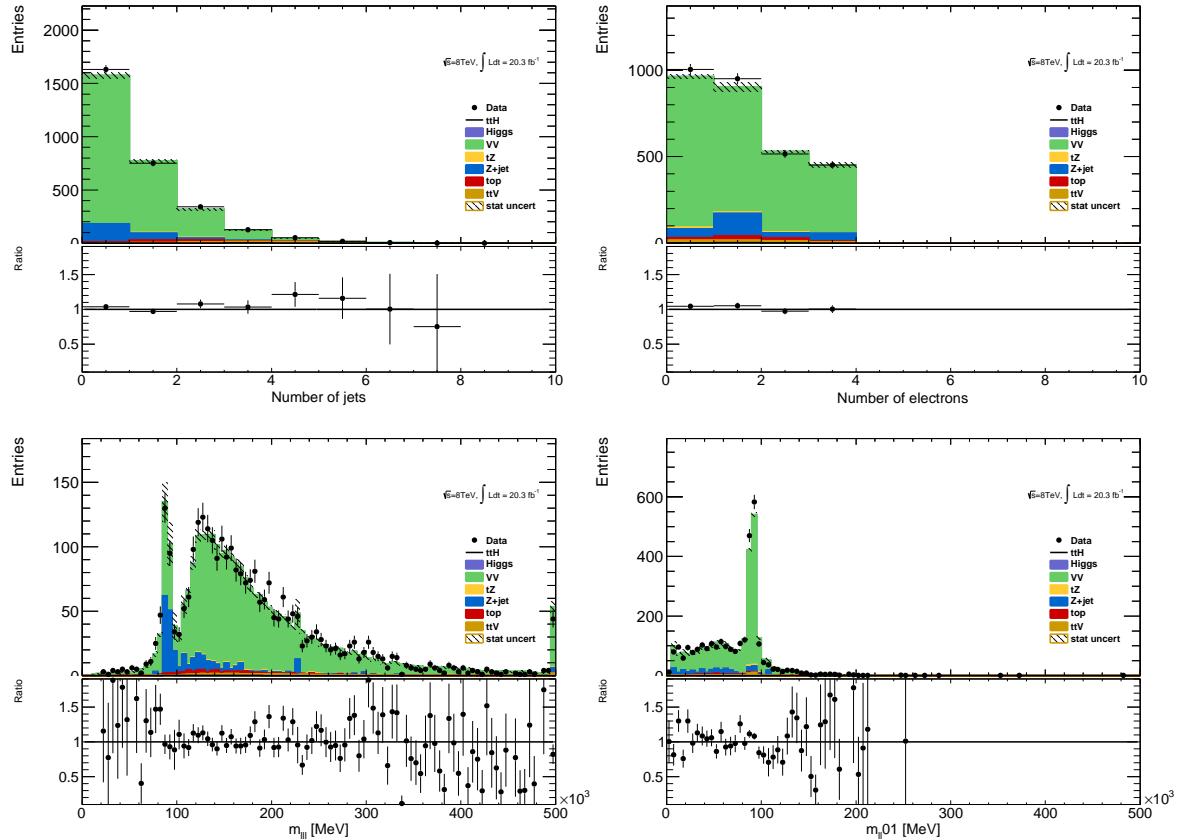


Figure 8.2: Inclusive 3 lepton $W^\pm Z$ validation region using the $t\bar{t}H$ lepton identification and momentum cuts: mass, number of jet and flavor variables

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

¹⁰the fakes are taken directly from MC

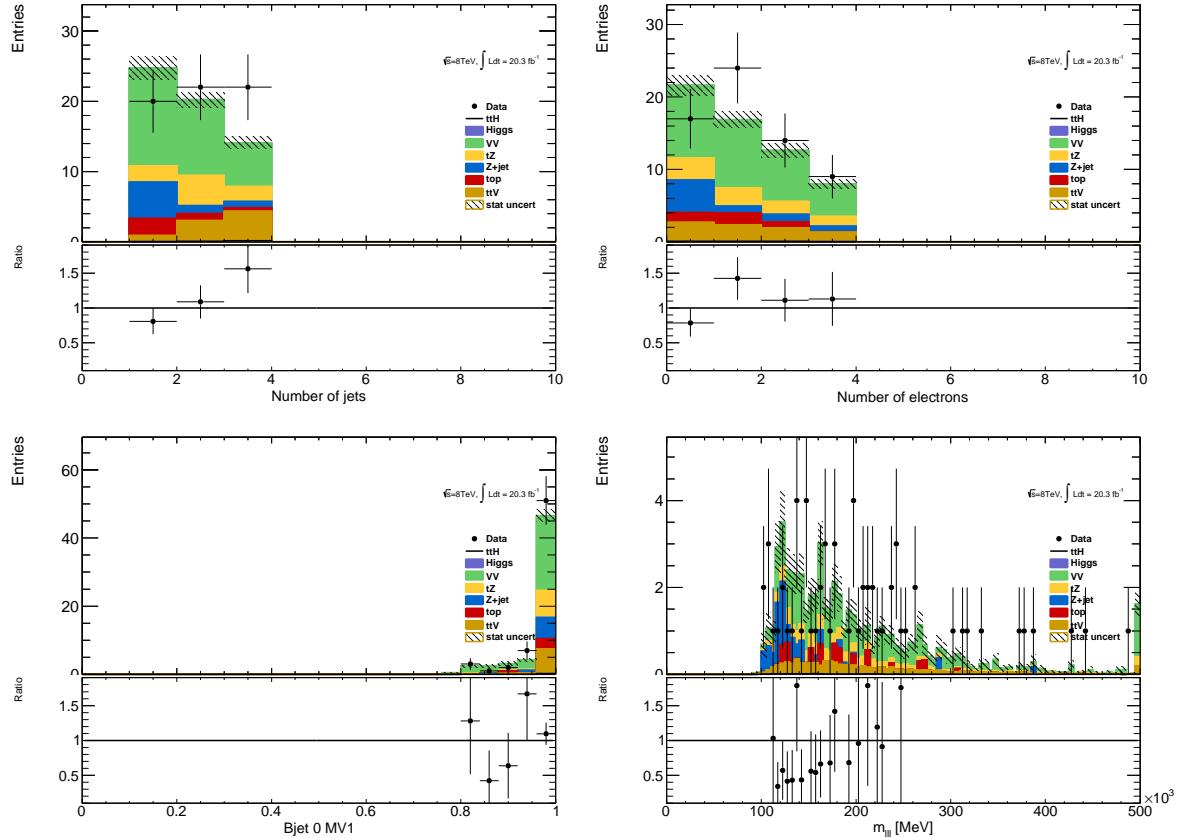


Figure 8.3: $W^\pm Z + b$ validation region: NJet, NElec, BJet MV1 and Mass Variables

1018 in the 3-jet bin. Importantly, also the tagged jet composition is also similar to the composition in the
 1019 $V + b$ analysis, already measured by ATLAS and discussed above.

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

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

1020 8.2.2 ZZ Uncertainty

1021 In order to investigate the MC agreement with data in the ZZ case, two validation regions similar to
 1022 the $W^\pm Z$ case are defined. Firstly, a 4 lepton ZZ region is constructed using the object selections for
 1023 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.

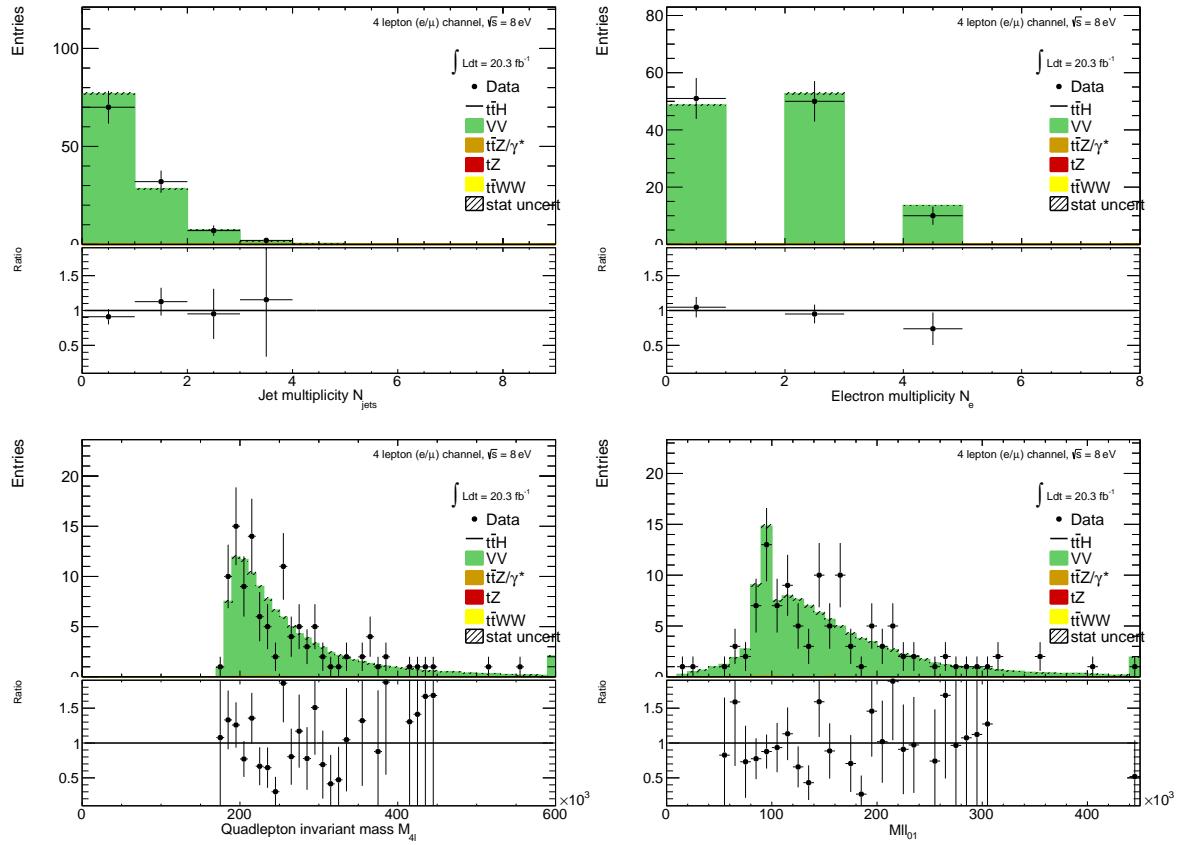


Figure 8.4: Jet-inclusive 4-lepton ZZ validation region using the $t\bar{t}H$ lepton identification and momentum cuts

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

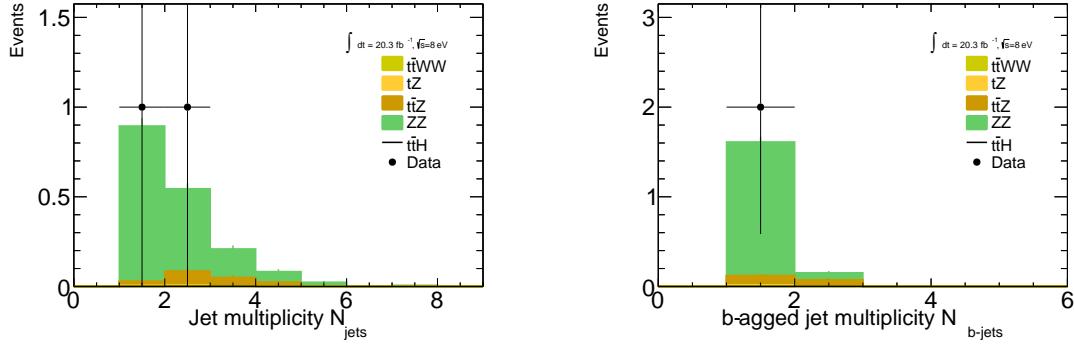


Figure 8.5: $ZZ + b$ validation region using the $t\bar{t}H$ lepton identification and momentum cuts

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 ± 0.02	0.21 ± 0.01	0.18 ± 0.01
$ZZ + b$ VR 2 Jet	0.25 ± 0.02	0.12 ± 0.01	0.11 ± 0.01
$ZZ + b$ VR 3 Jet	0.085 ± 0.014	0.040 ± 0.011	0.036 ± 0.011
$4l$ SR	0.020 ± 0.008	0.025 ± 0.008	0.014 ± 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 $t\bar{t}di$ -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 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

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

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

8.3.1 Likelihood Method

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

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

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

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

where ϵ_i and ϵ_j are the charge mis-identification rates for the two different electrons.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

1104 The likelihood method uses only Z *signal* events. Therefore, background coming from other
 1105 processes where the dilepton invariant mass corresponds to the one of the Z boson needs to be
 1106 subtracted. The background subtraction is done using a simple side-band method. This method
 1107 consists in dividing the Z invariant mass in three regions, i.e. A , B and C , where B is the central
 1108 region corresponding to the Z peak. The number of events is counted in the regions on the sides of
 1109 the peak, i.e. n_A and n_C , and removed from the total number of events in the peak region B , n_B .
 1110 This way, the number of signal events N_Z is given by

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

1111 Once the background has been subtracted, the likelihood method can be applied. MINUIT is used
 1112 for the minimization and MIGRAD to compute the uncertainty on these rates.

1113 8.3.2 Results

1114 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]
 1115 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
 1116 the rates are calculated using $t\bar{t}$ MC, scaled by the data-MC ratio of the rates in the lower p_T bins,
 1117 [90-130] GeV/ c . Figure 8.6 shows the extracted rates in all bins.

1118 To validate the likelihood approach, we apply the full method to the Z MC samples (extracting
 1119 rates via a likelihood fit and applying them to opposite sign events) and compare to the MC predicted
 1120 number of same-sign events. The invariant mass of the Z from our charge mis-identification and
 1121 directly from the MC can be seen on Figure 8.7. In the simulated Z samples, the number of same-
 1122 sign Z events is 5 049 while the estimation is $5\ 031^{+375}_{-365}$. The uncertainties combine both statistical
 1123 systematic uncertainties, which are discussed in depth below. The validation gives compatible results
 1124 within uncertainties.

1125 8.3.3 Systematic and Statistical Uncertainties

1126 Statistical uncertainties dominate the combined uncertainty on the charge mis-identification estimate.
 1127 The statistical uncertainties come primarily from the size of the Z same-sign sample in data and are
 1128 especially large for central, material-poor regions where the charge mis-identification rate is extremely
 1129 low. Additionally systematic uncertainties are included for a comparison between the positron and
 1130 electron rate, the per-bin MC closure test discussed above in the Results section, and for the effect of
 1131 varying the invariant mass window used for the background subtraction for three different cases. The
 1132 high p_T extrapolation induces a statistical error only in the last p_T bin. This bins is essentially irrelevant

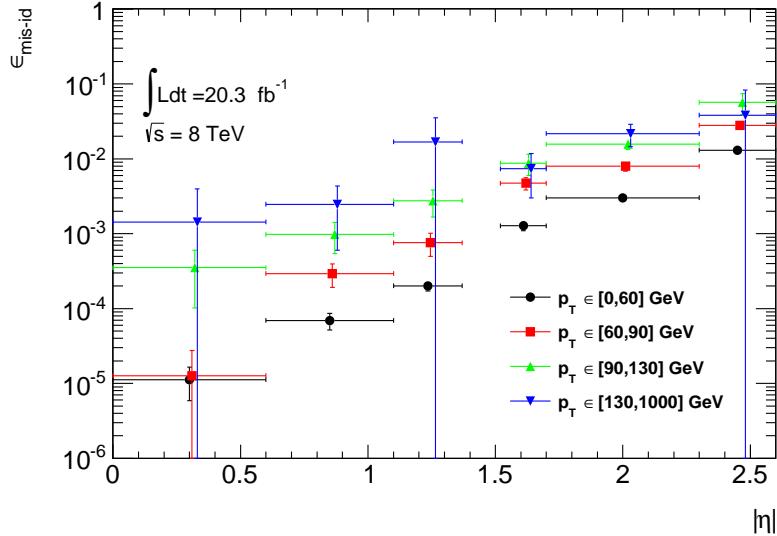


Figure 8.6: Electron charge mis-identification rates measured in data with the likelihood method on Z events (black points, red squares and blue triangles) as a function of $|\eta|$ and parametrized in p_T . The full 2012 dataset has been used to estimate the rates below 130 GeV. Above this value, the 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.

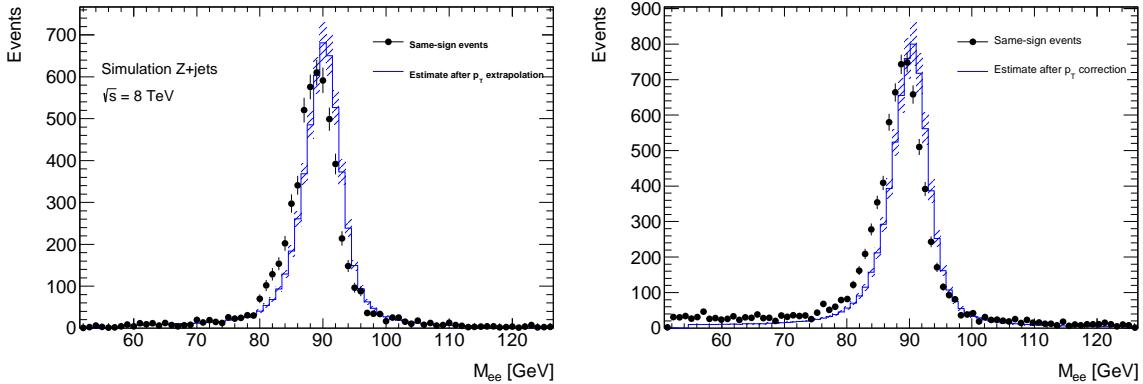


Figure 8.7: Closure test on simulated $Z \rightarrow e^+e^- + \text{jets}$ events (a) and data (b). The black circles show the distribution of same-sign events while the blue histograms show the distribution of the reweighted opposite-sign events together with the statistical and systematic uncertainties.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

1133 to the energy scales considered in this analysis. Figure 8.8 shows the relative bin uncertainties for all
1134 rate bins.

1135 We apply the rates to estimate the charge mis-identification background in the 2l SS signal regions,

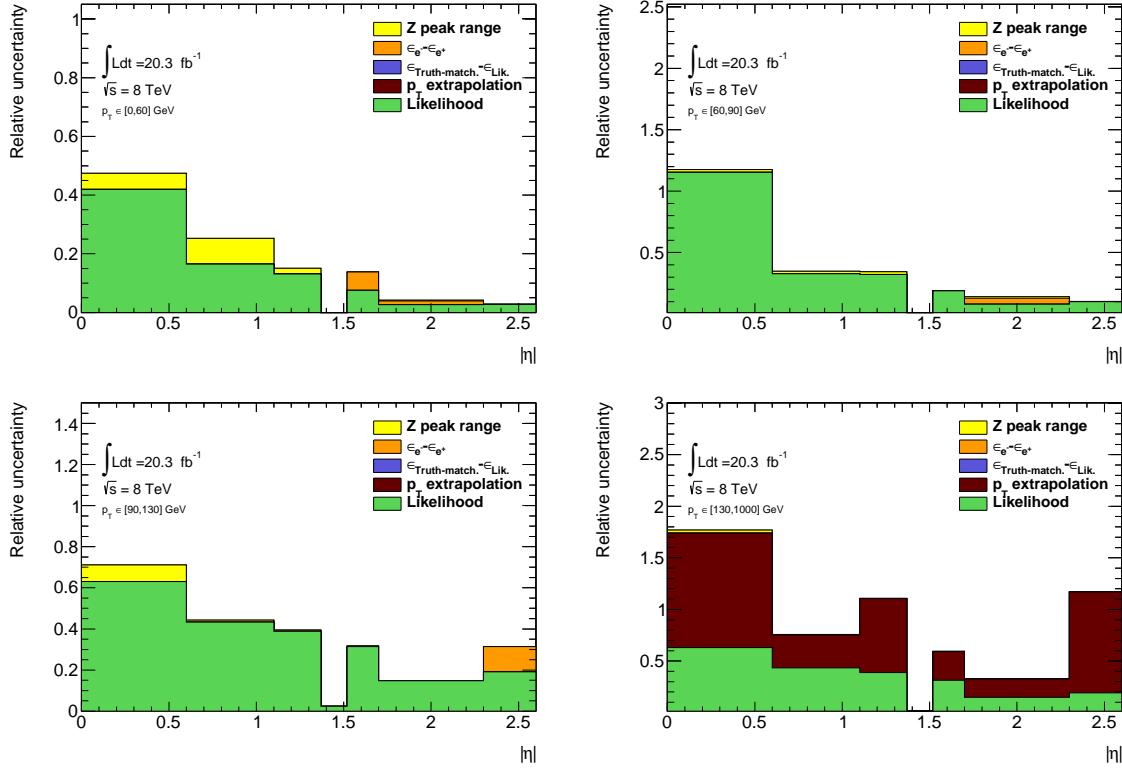


Figure 8.8: Relative systematic uncertainty contributions on the charge mis-identification rate, for different bins in p_T and $|\eta|$. Tight++ electrons have been used to produce this plot.

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}^μ). This region is fake-dominated with small contributions from prompt backgrounds, which are subtracted from the data. The total number of fake events

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

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

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

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

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

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

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

8.4.1 2l SS Fakes

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

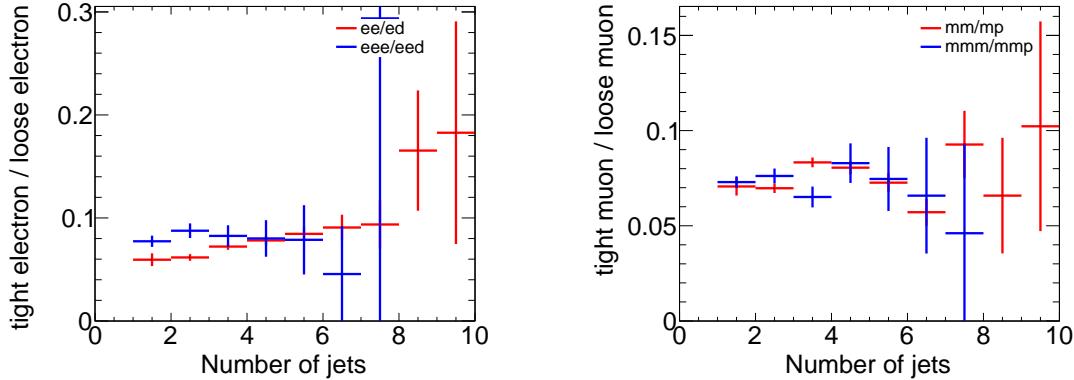


Figure 8.9: Ratios of regions with tight and anti-tight leptons in 2-lepton and 3-lepton channels

- 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, θ_e and θ_μ , 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 $t\bar{t}$ MC is included in the plots for reference, although it is not included in the measurements.

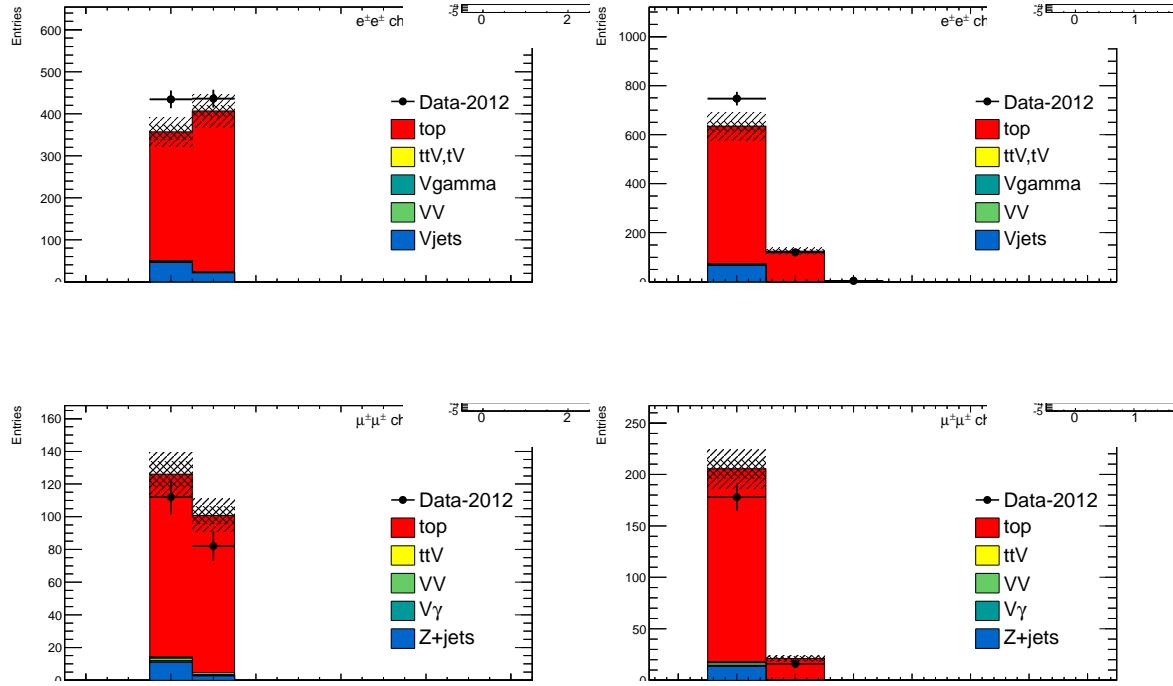


Figure 8.10: 2,3 Jet SS 21 ed (above) and μp control regions with at least one b-tagged jet. This region and the associated tight-id regions are used to determine θ_e and θ_μ . The $t\bar{t}MC$ (red) is used for reference and not used in the actual calculation

Factor	Expected (MC)	Measured (data)
θ_e Rev. Id.	0.0136 ± 0.0062	0.0156 ± 0.0062
θ_μ Rev. p_T	0.078 ± 0.012	0.1156 ± 0.0288

Table 8.5: Expected and measured values of the θ factors.

1193 The transfer factors obtained from the 2 and 3 jet regions are shown in 8.5 with statistical errors
 1194 and propagated systematic errors from the subtraction of relevant backgrounds ($t\bar{t}V$ and charge mis-
 1195 identification). The MC values are just for comparison. An additional systematic error is added by
 1196 comparing the transfer factors, obtained from the low jet control region, to those obtained from the
 1197 higher jet signal regions, using $t\bar{t}MC$. The value of this systematic is about 20 % and can be seen in
 1198 the above Figure 8.9. The overall systematic uncertainties and contribution from each source in all of
 1199 the sub-channels of the signal region are shown in Table 8.5 and the final contribution of fake events
 1200 to the signal region are show in Table 8.1 found at the beginning of the chapter.

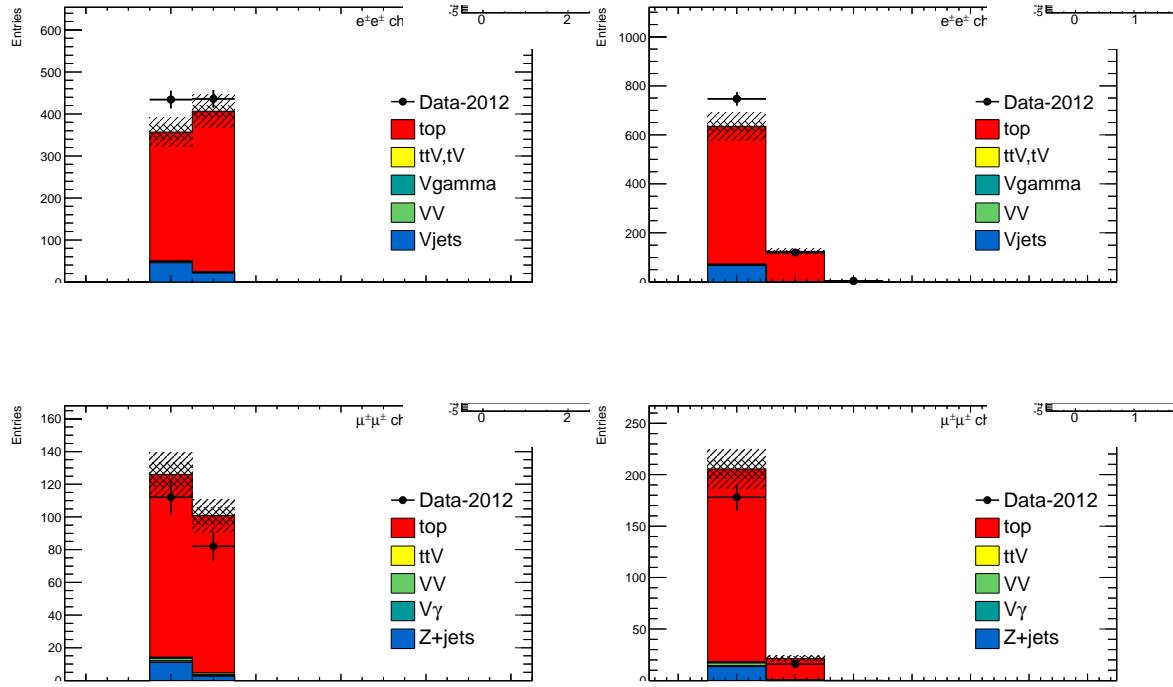


Figure 8.11: 4,5 Jet SS 2l ed (above) and μp control regions with at least one b-tagged jet. After subtraction of prompt and charge mis-identification backgrounds, these regions are scaled by the transfer factors, θ_e and θ_μ to obtain the final number of fake events in the CR. The $t\bar{t}$ MC (red)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

Uncertainties		Channels					
		4 jets		≥ 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 θ_e and θ_μ 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)

1201 8.4.2 3l Fakes

1202 The 3l fake method follows the same general strategy as the 2l SS case. Transfer factors are used
 1203 extrapolate from an anti-tight, fake-rich control region in data into the signal region. However, the
 1204 equivalent low jet control regions are too low in statistics to provide the transfer factors from data
 1205 directly, as above. Instead, the transfer factors are obtained directly from the $t\bar{t}$ simulation and data
 1206 control regions are used to determine the modeling of the identification and isolation variables, used in
 1207 the transfer factor extrapolation. The low jet regions are still employed in a low statistics validation
 1208 of the entire fake procedure.

1209 Anti-tight electrons and muons are defined in slightly different ways, compared to the 2l SS case:

- 1210 • **anti-tight electron (d):** fails to verify the verytight likelihood operating point, but still verifies
 1211 the veryloose operating point. the isolation selection is released $E_T^{rel} > 0.05$, $p_T^{rel} > 0.05$.
- 1212 • **anti-tight muon(p):** muons must pass identification but the p_T cuts is lowered to $6 \text{ GeV}/c$, the
 1213 overlap removal with jets and isolation cuts are released.

1214 The transfer factors, θ_e and θ_μ , extracted from MC, is defined as the ratio of the number of top ($t\bar{t}$ +
 1215 single-top) events in the signal region, to the number of $t\bar{t}$ events in the anti-tight regions. The factors
 1216 are calculated in separate flavor regions to ensure that the electron jet fakes and muon jet fakes are
 1217 calculated separately. The calculation follows the same for as in Equation 8.8, but now lep0, which
 1218 by construction is almost never a fake is allowed to be either electron or muon in both cases, denoted
 1219 below in Equation 8.12.

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

1220 The MC modeling of the variables involved in the transfer factor can be verified when another
 1221 variable fails. For instance, the MC modeling of the electron isolation variable can be compared to
 1222 data when the particle identification variable fails and vice-versa. The modeling of muon-jet ΔR ,
 1223 involved in the overlap removal, can be compared when either the isolation variable or the p_T fails.
 1224 The comparison of the electron variables in this manner can be seen in Figure 8.13 and the muon
 1225 variables in Figure 8.12. The regions used have the same selection as the signal region with an added
 1226 missing transverse energy requirement, $> 60 \text{ GeV}/c$, to ensure only top fakes. 20% and 21% systematic
 1227 uncertainties are assigned to the muon and electron transfer factors, respectively, to account for data-
 1228 MC discrepancies. This method for evaluating data-MC agreement for individual electron and muon
 1229 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.

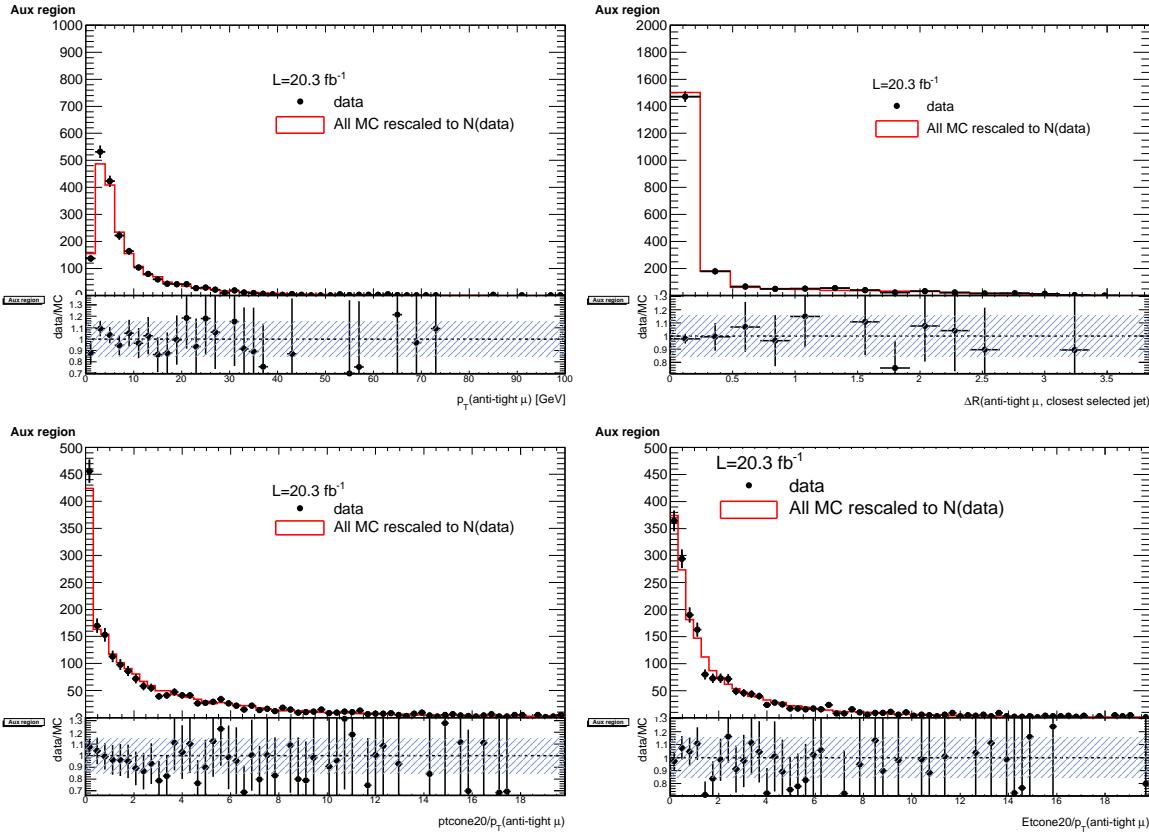


Figure 8.12: Distributions of the properties of the anti-tight muons in data (dots), compared with the total simulation (red line), rescaled to the integral of the data for a shape comparison. The uncertainty on the data distribution is statistical. The number of events for each of them is also presented 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

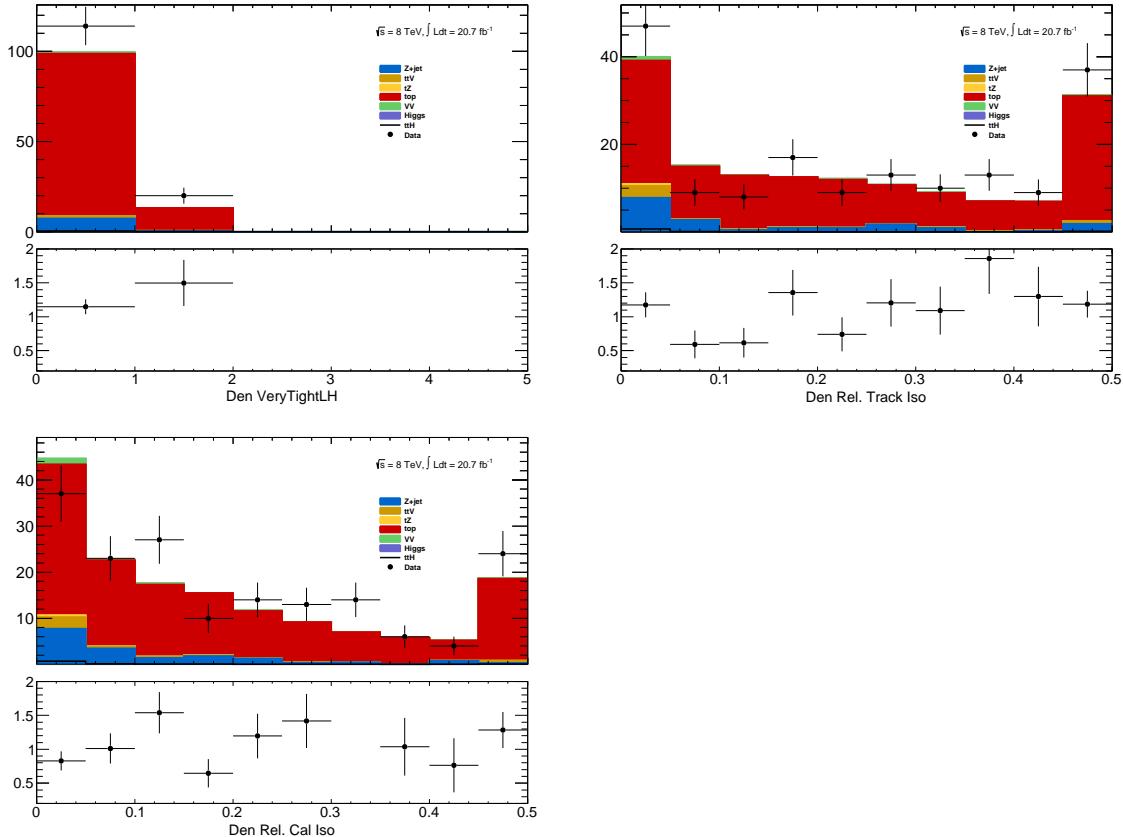


Figure 8.13: Distributions of anti-tight electron variables. The variables presented are, from top left to bottom right, p_T , η , Very Tight Likelihood word value, $\text{ptcone20}/p_T$, $\text{Etcone20}/p_T$. The plotted regions have the same cuts as the signal region, except the anti-tight electron 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.

1239 in the calculation are shown in Table 8.7.

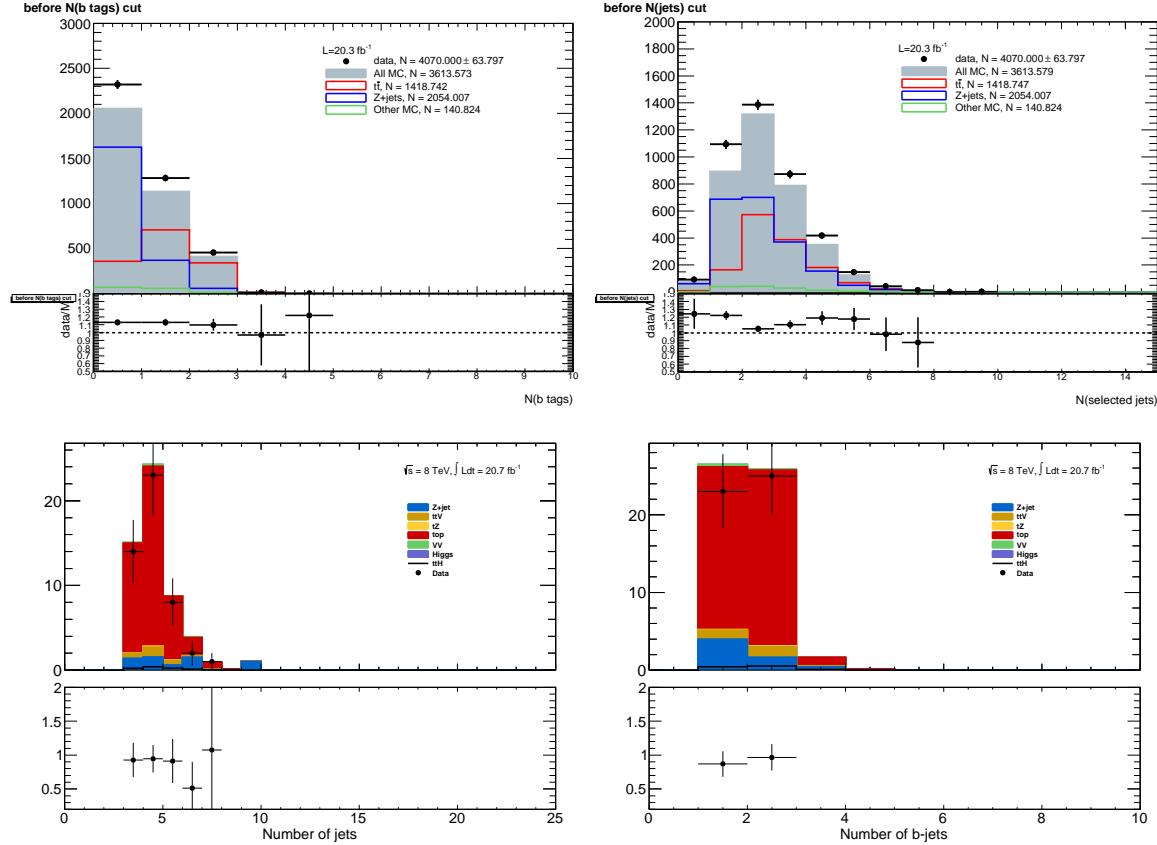


Figure 8.14: Muon-xxp (above) and electron-xxd (below) anti-tight control regions: jet variables. Note: the $t\bar{t}$ and top MC in the plots is used only as comparison, but is not included in the fake measurement

Stage	Muon	Electron
Anti-Tight CR Normalization	364.62 ± 20.02 (5%)	38.2 ± 6.9 (17%)
Transfer factor θ_μ	0.0047 ± 0.0011 (23%)	0.0240 ± 0.0064 (29%)
SR Contribution	1.71 ± 0.42 (25%)	0.91 ± 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

1240 Finally, the low jet region (1,2,3j) is used as a validation for the method, described above. The
 1241 $t\bar{t}$ and single top fakes are estimated using the procedure above, but instead using the lower jet region.
 1242 Similar systematics are assessed. This region with the fake estimate is plotted in Figure 8.15. The
 1243 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.

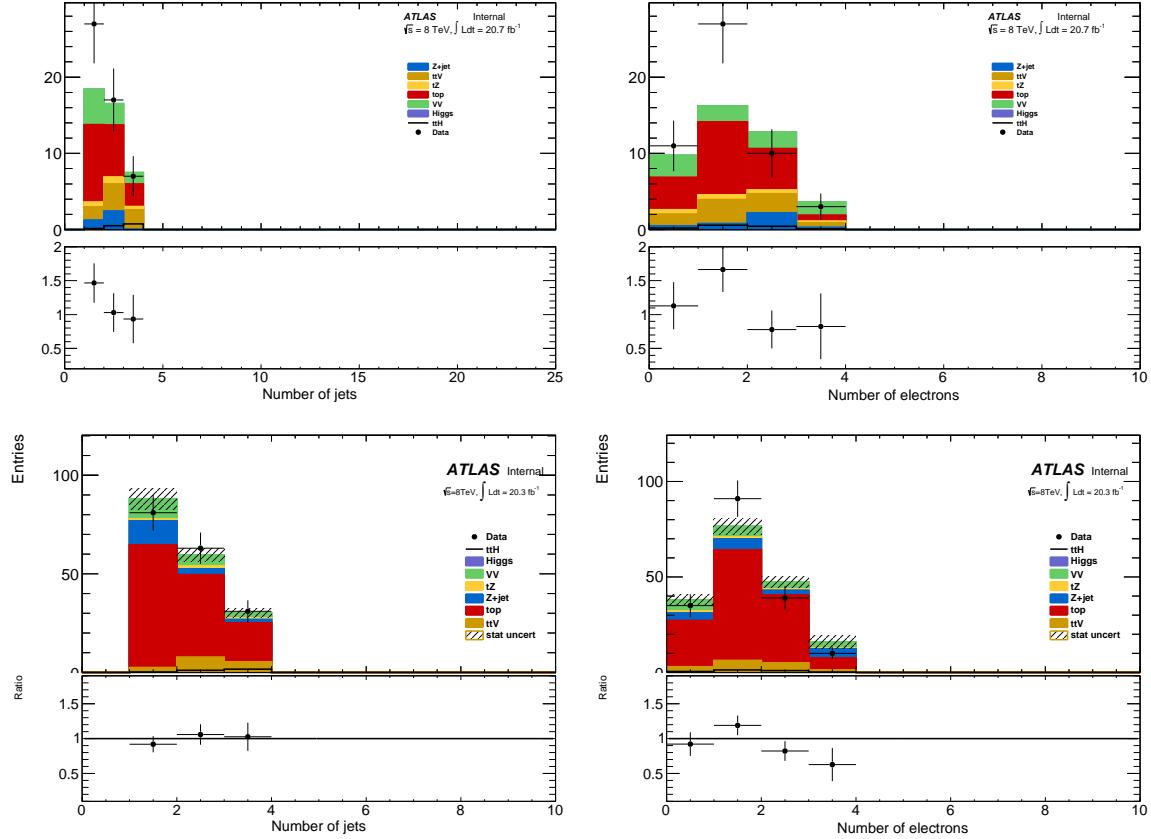


Figure 8.15: 3l fake validation regions for nominal p_T selection (above) and relaxed p_T selection, $> 10 \text{ GeV}/c$, (below). Plotted are the number of jets and the number of electrons in each event. The data and MC ratio below each plot agree with 1 within the statistics of the region and the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

1254 identified and 2 anti-identified leptons are created. These control regions do not have a number of jets
1255 requirement in order to increase statistics. From these control regions, two extrapolations are made.
1256 First, a transfer factor is applied to extrapolate from the anti-tight to tight regions for electrons
1257 and muons. The regions are defined with identifical object identification selection and reversal as
1258 the 3l case, and the same transfer factors can be used. They must be used twice however, because
1259 there are two anti-identified leptons in each event. Second, the jet inclusive regions are extrapolated
1260 into the 2-jet signal region, using as a second extrapolation factor derived from $t\bar{t}$ events. Since, the
1261 majority of fake leptons arise from b-quark initiated jets, the jet spectrum $t\bar{t}$ events with the additional
1262 requirement of 2-btagged jets from data are used as a model for the jet extrapolation. The overall
1263 systematic uncertainty on this measurement arises from the statistics in the control regions and MC
1264 based assessments of non-closure and are 35%-50% depending on the sub-channel.

1265

CHAPTER 9

1266

Summary of Systematic Uncertainties

1267 This chapter summarizes the various systematic uncertainties that enter the measurement of the limit
 1268 of $t\bar{t}H$ multi-lepton analysis, as well as structure of the statistical analysis model used to obtain
 1269 the measurement. The systematic uncertainties arise from three main sources. The first are the
 1270 normalization uncertainties on the background process estimation methods, which are discussed in
 1271 depth in . The second source is the theoretical uncertainties on the $t\bar{t}H$ production cross-section
 1272 and acceptance. The final source are the experimental and detector related systematic uncertainties
 1273 related to event selection efficiencies and measurements and identification of the objects. They
 1274 affect only the non-data driven backgrounds and the $t\bar{t}H$ signal, as simulation is used to model their
 1275 acceptance and efficiency for the analysis selection.

1276 These systematic uncertainties, the estimated background and signal event counts in each of the
 1277 signal regions, and the observed data in each signal region are combined in a statistical fit to an
 1278 analysis model to extract the measurement of interest. We measure per-channel and combined ratios
 1279 of the observed production rate to the theoretically predicted production rate of $t\bar{t}H$, a parameter
 1280 called μ . In the absence of a statistically significant observation, this measurement is translated into a
 1281 upper confidence limit on μ . The details of this procedure are discussed in the following sections and
 1282 the results with the observed data are discussed in Chapter 10

1283 **9.1 Systematic Uncertainties on Signal Cross-section and Acceptance**

1284 The $t\bar{t}H$ signal is simulated with matrix elements at NLO (next-to-leading order) in QCD with Powheg
 1285 and is discussed in Chapter 6.

1286 The production cross section and the Higgs boson decay branching fractions together with their
 1287 theoretical uncertainties from the QCD scale and PDF choice are taken from the NLO theoretical

1288 calculations reported in Ref. [55]. The uncertainty from the QCD scale estimated by varying μ_0 by a
 1289 factor of 2 from the nominal value is $^{+3.8\%}_{-9.3\%}$, while the uncertainty from the PDF set and the value of
 1290 α_S is $\pm 8.1\%$.

1291

1292 The impact of the choice of the QCD scale on the simulation of the $t\bar{t}H$ event selection efficiency
 1293 is estimated in two independent ways.

1294 First, the factorisation and renormalisation scales μ_0 are varied by a factor of 2, as $\mu = 2\mu_0$ and
 1295 $\mu = \mu_0/2$. The effects of these new scales are estimated via the application of event reweighting
 1296 procedures on the nominal simulation using kinematic distributions at parton level. The weights used
 1297 are dependent on the transverse momenta of both the $t\bar{t}H$ system and of the top quark, as described
 1298 in Ref. [70].

1299 Second, the choice of the factorisation and renormalisation scales, dependent on fixed (“static”)
 1300 parameters in the nominal simulation, is tested comparing its prediction with an alternative (“dy-
 1301 namic”), but still physics motivated choice $\mu_0 = (m_T^t m_T^{\bar{t}} m_T^H)^{\frac{1}{3}}$, which depends on kinematic variables.
 1302 This comparison is performed via event reweighting of the nominal static simulation based on weights
 1303 derived as a function of the $t\bar{t}H$ transverse momentum [70]. In order to take the difference between
 1304 the choices of scale as systematic uncertainties, a symmetric envelope around the nominal simulation
 1305 is built applying the weights and also their inverses.

1306 Fig. 9.1 shows the impact of the different choices for the factorization and renormalization scales
 1307 on the jet and b-jet multiplicities in events with 2 SS leptons. Similar variations are seen also in
 1308 the other event categories. In order to not double-count the variations on the total cross section
 1309 the predictions from the different QCD scales are normalised to the same total cross section. That
 1310 means that the observed differences are only coming from the event selection. Significant variations
 1311 on the jet multiplicities can be seen and these translate into different predictions on the signal event
 1312 yields in the signal regions. Such differences, listed in Table 9.1, are taken as theoretical systematic
 1313 uncertainties in addition to the ones affecting the total $t\bar{t}H$ production cross section. The “Static”
 1314 uncertainties come from the variations by a factor of 2 from the nominal scale and they are correlated
 1315 with the uncertainties on the total cross section, which are estimated with the same procedure. The
 1316 “Dynamic” uncertainties come from the difference between the nominal and the alternative dynamic
 1317 scale and are treated as an independent source of theoretical uncertainty.

1318 The uncertainty of the $t\bar{t}H$ event selection due to the PDF sets is estimated comparing the predic-
 1319 tions with three different PDF sets, varying each set within errors and taking the width of the envelope
 1320 as systematic uncertainty. The recommended sets are CT10, MSTW2008nlo68cl and NNPDF21_100. We

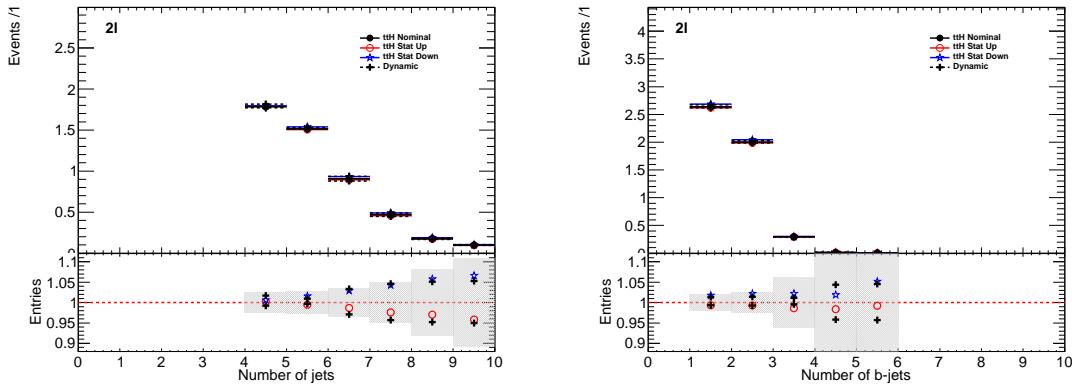


Figure 9.1: Effects on the jet multiplicities in 2 SS lepton $t\bar{H}$ events from different choices of the factorization and renormalization scales. “Static” refers to the variations by a factor of 2 of the nominal μ_0 , while “Dynamic” refers to the alternative choice of μ_0 which depends on the event kinematic. The grey band in the lower panels represents the statistical uncertainty of the nominal sample.

Table 9.1: Theoretical uncertainties of the signal event yields in the signal regions due to the impact of QCD scale uncertainties on the event selection.

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

Table 9.2: Uncertainties on $t\bar{H}$ acceptance in signal regions due to PDF variation.

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

1321 determine the change in the acceptance due to the PDF sets via the formula 9.1 to disentangle it from
 1322 the change in production cross section.

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

1323 Fig. 9.2 shows the estimated PDF systematic uncertainties as a function of the jet multiplicity in
 1324 $t\bar{H}$ events with at least two leptons. The uncertainties are compatible with the uncertainty on the
 1325 production cross section estimated in Ref. [55] and indicated by the dashed red lines in the lower panel.
 1326 Table 9.2 shows the half-width of the envelope of the acceptance under all eigenvector variations of
 1327 the three PDF sets. No significant dependence on the event topology is observed, so that the PDF
 1328 systematic uncertainty on the $t\bar{H}$ event selection is neglected.

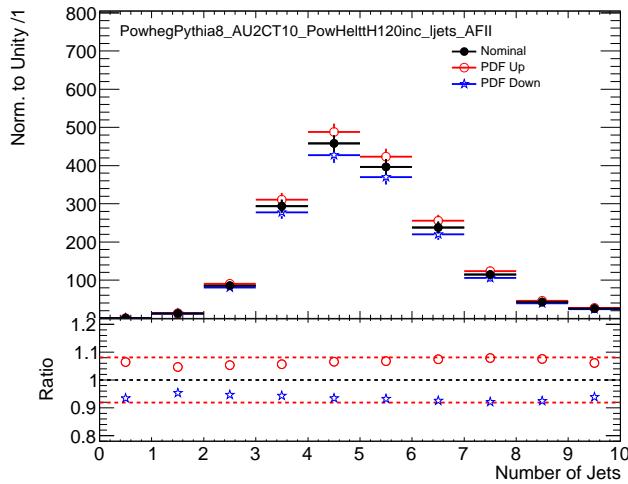


Figure 9.2: PDF systematic uncertainty on the jet multiplicities in $t\bar{t}H$ events with at least 2 leptons. The dashed red lines in the lower panel indicate the systematic uncertainty on the $t\bar{t}H$ production cross section.

1329 9.2 Experimental and Detector Systematic Uncertainties

1330 Experimental and detector systematics uncertainties arise from the efficiency of identifying objects
 1331 and the efficiency of the event selections. These affect only MC models of physics processes, $t\bar{t}V$,
 1332 $t\bar{t}H$, VV . Data-driven backgrounds take into account these effects by construction. We consider
 1333 systematic effects from a number of sources: the lepton and jet energy scale measurements, the lepton
 1334 identification and isolation selections, the efficiency and mis-identification rate associated with tagging
 1335 b-quark jets. Effects due to modelling the energy and objects from additional vertices were studied
 1336 and found to be negligible. The vast majority of the individual detector systematics effects are small.
 1337 The sum total of the systematic effects are comparable to some of the overall normalization and
 1338 cross-section uncertainties on some of the physics processes and is shown in Table ??.

1339 9.2.1 Lepton Identification, Energy Scale, and Trigger

1340 The electron[46] and muon identification efficiencies[71] are measured in data using Z boson and
 1341 J/Ψ control samples. They are shown in Figure 9.3. The uncertainty on the muon efficiencies are
 1342 measured as functions of η and p_T and are generally less 1%. The uncertainty on the electron and
 1343 muon efficiencies are also measured as functions of η and p_T and are at the 1 % level for p_T above
 1344 30 GeV/ c , but become much larger 5-10% for the lower p_T regimes. These translate into sub-% level
 1345 effects on the $t\bar{t}V$ and $t\bar{t}H$ event counts in the signal regions for the muons and \sim % level effects for

1346 the electrons. The effects become more important with increasing numbers of leptons.

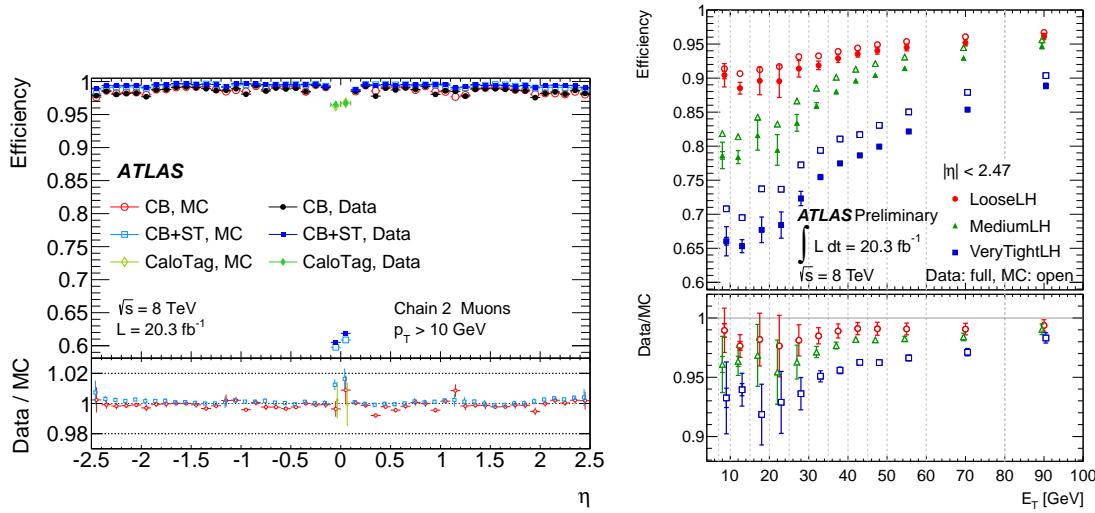


Figure 9.3: Muon (left) and electron(right) identification efficiencies in Data and MC as a function of η and p_T respectively. For electrons, the verytight likelihood operating point is used and for muons the CB+ST (combined+segment tagged) operating point is used

1347 The electron[72] and muon[71] energy scale and resolution are also measured using the Z -boson
 1348 control samples in data. The uncertainties related to the scale and resolution for the leptons affect
 1349 the overall event acceptance through the lepton momentum cuts primary and have negligible impact
 1350 on the event count uncertainties in the signal regions.

1351 The efficiencies for muons and electrons to pass muon[73] and electron triggers[74] have been
 1352 calculated with respect to the offline identification operating points using the Z boson control samples.
 1353 They are in the range of 90% for electron triggers and 70% for muon triggers, owing to gaps in muon
 1354 trigger coverage, and have % level errors. When statistically combined for 2,3,4 and lepton signal
 1355 regions, the overall trigger efficiency is high and the error on the number of expected events is negligible.

1356 9.2.2 Lepton Isolation and Impact Parameter

1357 The isolation and impact parameter selections are specific to this analysis and are discussed in depth
 1358 in Chapter 7. We calculated their combined efficiency with respect to the full lepton identification
 1359 selection using the Z boson control samples and define data-MC scale factors to correct the efficiency

in the simulation. Background are subtracted using shape templates in the di-lepton invariant mass spectrum. The electron template is derived from MC, while the background template is derived from the same-sign control region, with certain object cuts reversed in the electron case. We measure the efficiency scale-factors in bins of lepton momentum. Uncertainties are assigned per-bin to account for the level of statistics and variations caused by the fit parameters. An additional 1% uncertainty envelope is added to both the electron and muon measurements to account for trends observed in the dependence of the data-MC efficiency scale-factor as a function of the number of jets. Stability of the efficiency scale-factor as a function of the number of jets is important for this analysis, because event activity in the low jet Z sample where the efficiency is measured is much different from the high jet signal regions where the efficiencies are applied. The dependence of the scale-factor on the number of jets can be seen Figure 9.4. The isolation scale-factor uncertainties are around 1-3% depending on the particle momentum, but these uncertainties propagate to 2-5 % (some of the largest) effects in the event counts in the signal regions. The uncertainties are more important in the regions with more leptons.

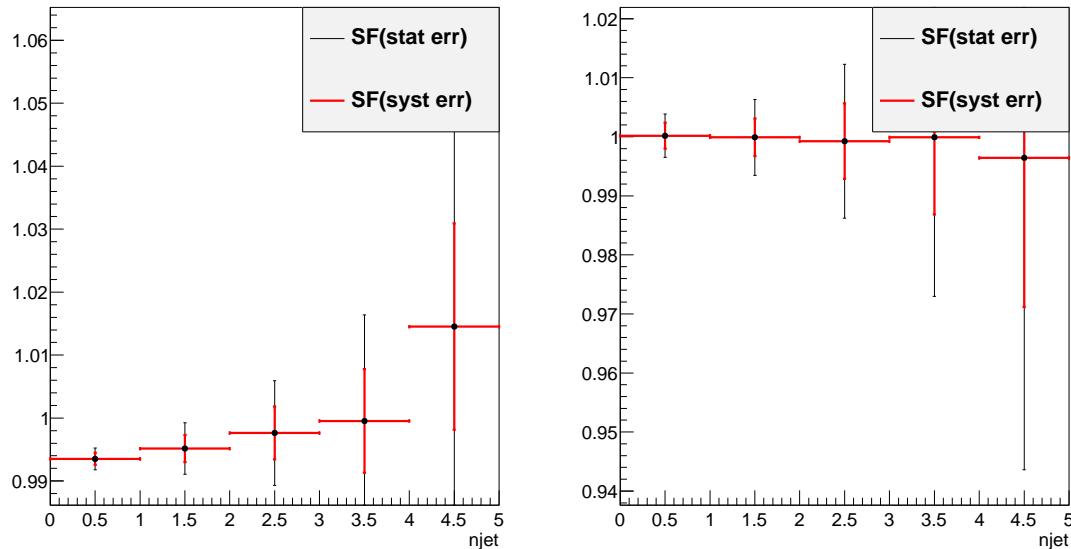


Figure 9.4: Muon (left) and electron(right) isolation efficiency scale-factors from the Z control sample as a function of the number of jets in the event. An additional systematic uncertainty of 1% is added to encompass the variation in the number of jets variable

1374 9.2.3 Jet Energy

1375 The jet energy scale (JES) is calculated using a combination of data-based insitu techniques, where
 1376 jet transverse momentum is balanced with respect to a reference photon or a Z boson, as well as single
 1377 particle test-stand studies[75]. Additional smaller effects are taken into account including the b-quark
 1378 jet specific response, near-by jets, the effects of pile-up and an intercalibration of similar η regions
 1379 using di-jet events. The JES systematic errors arises from numerous sources that are diagonalized
 1380 into eigenvectors so that they can be combined in an uncorrelated way. The combined uncertainty
 1381 is plotted in Figure 9.5 as a function of jet η and p_T and is the range 2-4% for jets used in this
 1382 analysis. The jet energy resolution is calculated in a similar way with slightly larger errors, 10% [76].
 1383 Propagated to the event counts in the signal regions, the combined scale and resolution systematics
 1384 are of non-negligible effects 6-7%.

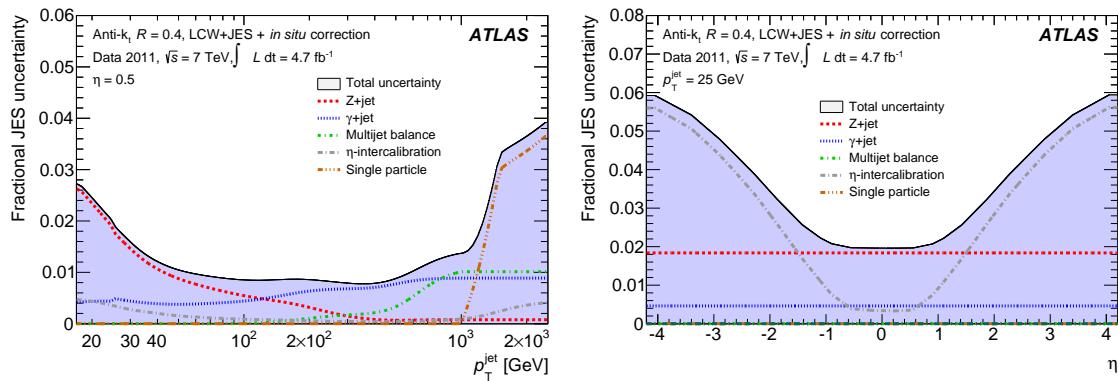


Figure 9.5: JES systematic uncertainties as a function of jet η (for jets $p_T > 25$ GeV) and p_T (for jets $|\eta| < 0.4$). The combined systematic uncertainty is shown with contributions from the largest sources

1385 9.2.4 B-Tagged Jet Efficiency

1386 The b -quark tagging efficiency must be calculated separately for charm, light and b -quarks. ATLAS
 1387 uses three data based control regions: an inclusive jet sample for mis-tagged light quarks[77], the $t\bar{t}$
 1388 sample for b -quarks[78], and a sample of D^* mesons for charm quarks[79]. These efficiencies and rates
 1389 are well-measured in MC and the data-based corrections are small. The data-MC efficiency scale-
 1390 factor shown in Figure 9.6 is close to 1 and has an overall systematic uncertainty of around 5%. The
 1391 uncertainties are applied to the analysis via a number of eigenvectors. Together these uncertainties
 1392 have a 4 % effect in the event expectation in the signal regions.

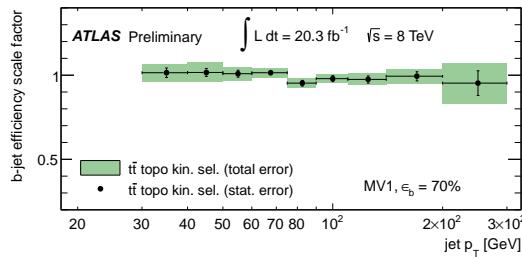


Figure 9.6: b-Tagging data-MC efficiency scalefactors versus jet p_T calculated in the $t\bar{t}$ sample from 2012 data. The uncertainties are combined statistical and systematic.

Total Systematic Uncertainty	2ee4j Down-Up	2ee5jincl Down-Up	2em4j Down-Up	2em5jincl Down-Up
ttH	-4.68 5.84	-8.24 6.14	-5.10 3.50	-5.52 6.40
ttW	-7.20 5.45	-8.72 11.30	-3.63 6.22	-9.72 7.95
ttZ	-9.68 5.07	-5.87 10.98	-4.07 6.16	-8.37 4.99
Total Systematic Uncertainty	2mm4j Down-Up	2mm5jincl Down-Up	3l Down-Up	4l Down-Up
ttH	-5.20 7.51	-7.28 6.75	-5.84 5.59	-6.54 6.54
ttW	-4.54 5.23	-8.63 6.88	6.36 8.16	— —
ttZ	-5.24 8.69	-9.73 8.18	-6.14 6.66	-9.58 6.94

Table 9.3: Sum in quadrature of all the systematic uncertainties on the number of event yields per channel.

9.2.5 Summary

The combined effect of these detector and experimental systematics on the $t\bar{t}V$ and $t\bar{t}H$ is provided in Table 9.3. The effects are smaller than the normalization uncertainties on some of the backgrounds. However, since they effect all processes signal and background. They are dominated by the lepton isolation scale-factor measurements and the electron identification with smaller contributions from the JES and b-tagging efficiencies. These detector systematic uncertainties enter the fit individually and their ranking of influence on the overall measurement uncertainty can be seen in Figure??.

9.3 Summary of Background and Signal Normalization Uncertainties

Tab.9.4 gives the summary of the systematic uncertainties that are included in the analysis for the normalization and acceptance of each process. The relative importance of these uncertainties to the final fit can be seen in Figure.

Type	Description	Name	Uncertainty
Signal (ttH)			
QCD Scale	Cross Section (Dynamic Scale) Analyses Acceptance	QCDscale_XS_ttH QCDscale_Acceptance_ttH	+3.8% –9.3% (Section 9.1) 0.-2.6%
PDF+ α_s	Cross Section Analyses Acceptance	pdf_ttH pdf_Acceptance_ttH	\pm 8.1% Negligible
ttW (Irreducible background)			
QCD Scale	Cross Section (Dynamic Scale) Analyses Acceptance	QCDscale_XS_ttW QCDscale_Acceptance_ttW	\pm 15% (Section 8.1) 0.4-3.5%
PDF+ α_s	Cross Section Analyses Acceptance	pdf_ttW PDF_Acceptance_ttW	\pm 13% 1.1-4.8%
ttZ (Irreducible background)			
QCD Scale	Cross Section (Dynamic Scale) Analyses Acceptance	QCDscale_XS_ttZ QCDscale_Acceptance_ttZ	\pm 12% (Section 8.1) 0.1-3.1%
PDF+ α_s	Cross Section Analyses Acceptance	pdf_ttZ PDF_Acceptance_ttZ	\pm 9% 0.9-2.7%
VV Backgrounds			
Cross Sections	WZ,ZZ Processes	WZxsec,ZZxsec	\pm 50% (Section 8.2)
Data-Driven Backgrounds			
Normalization Uncertainty		Jet Fakes	\pm 30-50% (Section 8.4))
Normalization Uncertainty		Charge MisID	\pm 20-30% (Section 8.3)

Table 9.4: Summary of systematics for processes present in the signal regions in the analysis, with their type, description, name, values and uncertainties, and status of inclusion in the final results.

1404

CHAPTER 10

1405

Results and Statistical Model

1406 10.1 Results in Signal Regions

1407 10.2 Statistical Model

1408 NEED TO DEFINE MU IN ANALYSIS SUMMARY CHAPTER

1409 We use the above results to make two sets of measurements: an upper confidence limit on μ , the
1410 signal strength parameter, and a measurement of μ . These measurements are done for each channel
1411 individually and then combined. The interpretation of the results in the form of a statsitical model
1412 follow the procedure, discussed here [80]. We interpret the results as counting experiments in each
1413 signal region. Therefore agreement in kinematic shapes do not affect the statistical procedure.

1414 10.2.1 The Likelihood

1415 The observed and expected event yields in the signal regions are analyzed using a binned likelihood
1416 function (\mathcal{L}), built from product of Poission models of expected event counts for each bin, where the
1417 bins for our case are the separate signal regions:

$$\mathcal{L} \propto \prod_{i=0}^{N_{ch}} P(N_{obs}^i | \mu \cdot s_{exp}^i + b_{exp}^i) \quad (10.1)$$

1418 where s_{exp}^i is the SM signal expectation in the signal region, b_{exp}^i are the background expectations, i
1419 counts over the signal regions, and P is the Poisson distribution. The signal strength parameter is the
1420 paramter of interest in the model (POI) and acts as a simple scale-factor to the SM $t\bar{t}H$ production
1421 rate and is common to all channels. Setting μ to 0 corresponds to the background only scenario. The
1422 background paramter, b , is a sum over all background processes.

1423 The signal and background expectations , s and b , depend on systematic errors. These are included
 1424 in the likelihood function in the form of a vector nuisance parameters, $\vec{\theta}$, which are constrained to
 1425 fluctuate within Gaussian distributions. These fluctuations affect the background and signal expecta-
 1426 tions by response functions, $\nu(\vec{\theta})$, set by uncertainties measured in the previous section. For instance,
 1427 the $W^\pm Z$ normalization uncertainty is 50% from Section 8.2 and is included in the fit as its own unit
 1428 gaussian, $G(\theta|0, 1)$. The fluctuations of the gaussian, θ_{WZ} scale the background contribution via the
 1429 form, $0.5 \cdot (1 + \theta_{WZ}) \cdot b_{WZ}$. For many of the detector systematics, the uncertainties are two sided
 1430 and are included as piecewise Gaussians. We add correlations to various uncertainties by hand, when
 1431 appropriate. With these nuisance parameters, the likelihood takes this form:

$$\mathcal{L}(\mu, \vec{\theta}) = \left(\prod_{i=0}^{N_{ch}} P(N_{obs}^i; \mu \cdot \nu_s(\vec{\theta}) \cdot s_{exp}^i + \nu_b(\vec{\theta}) \cdot b_{exp}^i) \right) \times \prod_j^{N_\theta} G(\theta_j; 1, 0) \quad (10.2)$$

1432 10.2.2 Test Statistic and Profile Likelihood

1433 Values of μ are tested with the negative log quantity, $q_\mu = -2\ln(\lambda(\mu))$, where $\lambda(\mu)$ is the test statistic.
 1434 $\lambda(\mu)$ is defined as:

$$\lambda(\mu) \equiv \frac{\mathcal{L}(\mu, \hat{\vec{\theta}}_\mu)}{\mathcal{L}(\hat{\mu}, \hat{\vec{\theta}})} \quad (10.3)$$

1435 where $\hat{\vec{\theta}}_\mu$ are values of the nuisance parameter vector that maximize the likelihood for a given value
 1436 of μ and $\hat{\mu}$ and $\hat{\vec{\theta}}$ are the fitted values of signal strength and nuisance parameters that maximize the
 1437 likelihood overall. μ is constrained to be positive.

1438 10.2.3 CL_s Method

1439 Exclusions limits on the signal strength are calculated with the test statistic using a modified fre-
 1440 quentist method, called the CL_s method[81]. CL_s is defined as a ratio of two frequentist quantities.
 1441 The numerator quantifies the probability of finding the observed data given the signal + background
 1442 hypothesis. The denominator quantifies the probability of the data given the background only hy-
 1443 pothesis.

1444 Using the numerator alone has the undesirable property that, if the data fluctuates below the
 1445 expectation, an exclusion limit can be reached that is far beyond the sensitivity of the experiment.
 1446 Normalizing to the background only hypothesis penalizes these low sensitivity cases.

1447 The probability of obtaining an observation as extreme as the data given a particular signal +
 1448 background hypothesis is given by the p-value, p_μ defined as:

$$p_\mu = \int_{q_\mu^{obs}}^{\infty} f(q_\mu) dq_\mu \quad (10.4)$$

1449 and the probability of obtaining an observation as extreme as the data given the background hypothesis
 1450 p_b is:

$$p_b = \int_{q_{\mu=0}^{obs}}^{\infty} f(q_{\mu=0}) dq_{\mu=0} \quad (10.5)$$

1451 where $f(q_{\mu})$ is the distribution of q_{μ} for all possible observations for a given μ and q is defined above.
 1452 Therefore,

$$CL_s = \frac{p_{\mu}}{1 - p_b} \quad (10.6)$$

1453 . A μ value is considered excluded at 95% confidence when CL_s is less than 5%.

1454 10.2.4 Exclusion Limits

1455 Table ?? shows expected exclusion limits for all channels, including the analysis uncertainties cumula-
 1456 tively. It shows the relative importance of the statistical and systematic uncertainties to the analysis
 1457 sensitivity. The observed limits using observed data and predictions can be seen in Figures ??-?? for
 1458 splitting and combining the subchannels and in Table XX by numbers. We expect a combined limit of
 1459 XXX (background only) and XXX (signal injected) and see a limit of XXX. The channel sensitivity
 1460 is dominated by the 2l and 3l channels.

Channels		Stat	+Fakes Unc.	+Theory	+ Experimental
2l	2lee	7.44	8.52	8.82	8.94
	2lem	3.46	3.81	4.07	4.18
	2lmm	4.03	4.14	4.47	4.57
	2ltau	8.08	8.92	10.00	10.03
	All	2.16	2.44	2.81	2.90
3l		3.40	3.43	3.59	3.66
4l		15.16	15.16	15.44	15.55
1l2tau		10.41	13.84	14.20	14.22
All		1.68	1.85	2.14	2.22

Table 10.1: 95%CL limits on μ for all channels and combination with cumulative uncertainties.

1461 10.2.5 μ Measurements

1462 In addition to setting a limit on the signal strength, we also fit the best value of the signal strength
 1463 for μ . We do this by minimizing the negative log likelihood value, q_{μ} or conversely maximizing the
 1464 likelihood. The $1-\sigma$ error band is set via a profile likelihood scan, where the value q_{μ} is scanned as
 1465 a function of μ . Values of μ that increase q_{μ} by 1 form the edges of the error band. The fitted
 1466 values of μ with errors are provided in Table XXX for each subchannel fit as well as the combined fit.

1467 10.2.6 Nuissance Parameter Impact on the Signal Strength

1468 Finally, we examine the post-fit impact of the various nuisance parameters on the final fit. We expect
1469 to have measured the various anaylsis uncertainties well and do not expect the fit to have much futher
1470 constraint. For that reason, we expect the pulls of the nuissance parameters to be close to 0 and the
1471 measured uncertainties on those parameters to be consistent with the input uncertainties. Figures
1472 XXXX show.

1473

CHAPTER 11

1474

Conclusions

1475 **11.1 Higgs Results in Review**

1476 **11.2 Prospects for Future**

Bibliography

- 1478 [1] S. L. Glashow, *Partial-symmetries of weak interactions*, Nucl. Phys. **22** (1961) no. 4, 579. [2.1.1](#)
- 1479 [2] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264. [2.1.1](#)
- 1480 [3] A. Salam and J. C. Ward, *Gauge theory of elementary interactions*, Phys. Rev. **136** (1964)
1481 [763–768](#). [2.1.1](#)
- 1482 [4] S. Weinberg, *Non-abelian gauge theories of the strong interactions*, Phys. Rev. Lett. **31** (1973)
1483 [494–497](#). [2.1.1](#)
- 1484 [5] D. J. Gross and F. Wilczek, *Ultraviolet behavior of non-abelian gauge theories*, Phys. Rev. Lett.
1485 **30** (1973) 1343–1346. [2.1.1](#)
- 1486 [6] G. 't Hooft and M. Veltman, *Regularization and renormalization of gauge fields*, Nuclear
1487 Physics B **44** (1972) 189 – 213. [2.1.1](#)
- 1488 [7] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964)
1489 [508](#). [2.1.2](#)
- 1490 [8] P. W. Higgs, *Spontaneous symmetry breakdown without massless bosons*, Phys. Rev. **145** (1966)
1491 [1156](#). [2.1.2](#)
- 1492 [9] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*,
1493 Phys. Rev. Lett. **13** (1964) 321–322. [2.1.2](#)
- 1494 [10] The ALEPH, CDF, DØ, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak
1495 Working Group, the Tevatron Electroweak Working Group, and the SLD electroweak and
1496 heavy flavour groups, *Precision Electroweak Measurements and Constraints on the Standard
1497 Model*, CERN-PH-EP-2010-095 (2010) , arXiv:1012.2367 [hep-ex]. [2.1.3](#), [2.2](#)
- 1498 [11] M. Baak, M. Goebel, J. Haller, A. Hoecker, D. Kennedy, R. Kogler, K. Mnig, M. Schott, and
1499 J. Stelzer, *The electroweak fit of the standard model after the discovery of a new boson at the
1500 LHC*, The European Physical Journal C **72** (2012) no. 11, .
1501 <http://dx.doi.org/10.1140/epjc/s10052-012-2205-9>. [2.1.3](#)
- 1502 [12] J. C. Collins, D. E. Soper, and G. Sterman, *Factorization for short distance hadron-hadron
1503 scattering*, Nuclear Physics B **261** (1985) 104 – 142. [2.2](#)
- 1504 [13] CERN, . CERN, Geneva, 1984. [2.2](#)

- 1505 [14] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and
 1506 R. Tanaka (Eds.), *Handbook of LHC Higgs Cross Sections: 2. Differential Distributions*,
 1507 CERN-2012-002 (CERN, Geneva, 2012) , [arXiv:1201.3084 \[hep-ph\]](https://arxiv.org/abs/1201.3084). **2.2**
- 1508 [15] *Updated coupling measurements of the Higgs boson with the ATLAS detector using up to 25*
 1509 *fb⁻¹ of proton-proton collision data*, Tech. Rep. ATLAS-CONF-2014-009, CERN, Geneva, Mar,
 1510 2014. **2.2.1**
- 1511 [16] CMS Collaboration Collaboration, *Precise determination of the mass of the Higgs boson and*
 1512 *studies of the compatibility of its couplings with the standard model*, Tech. Rep.
 1513 CMS-PAS-HIG-14-009, CERN, Geneva, 2014. **2.2.1**
- 1514 [17] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of the Higgs boson mass from*
 1515 *the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4\ell$ channels with the ATLAS detector using 25 fb⁻¹ of pp*
 1516 *collision data*, [arXiv:1406.3827 \[hep-ex\]](https://arxiv.org/abs/1406.3827). **2.2.1**
- 1517 [18] *Evidence for the spin-0 nature of the Higgs boson using {ATLAS} data*, Physics Letters B **726**
 1518 (2013) no. 13, 120 – 144.
 1519 <http://www.sciencedirect.com/science/article/pii/S0370269313006527>. **2.2.1**
- 1520 [19] S. Dawson, A. Gritsan, H. Logan, J. Qian, C. Tully, et al., *Working Group Report: Higgs*
 1521 *Boson*, [arXiv:1310.8361 \[hep-ex\]](https://arxiv.org/abs/1310.8361). **2.2.2**
- 1522 [20] O. Eberhardt, G. Herbert, H. Lacker, A. Lenz, A. Menzel, et al., *Impact of a Higgs boson at a*
 1523 *mass of 126 GeV on the standard model with three and four fermion generations*,
 1524 Phys.Rev.Lett. **109** (2012) 241802, [arXiv:1209.1101 \[hep-ph\]](https://arxiv.org/abs/1209.1101). **2.2.2**
- 1525 [21] M. Carena, S. Gori, N. R. Shah, C. E. Wagner, and L.-T. Wang, *Light Stops, Light Staus and*
 1526 *the 125 GeV Higgs*, JHEP **1308** (2013) 087, [arXiv:1303.4414](https://arxiv.org/abs/1303.4414). **2.2.2**
- 1527 [22] N. Arkani-Hamed, K. Blum, R. T. D’Agnolo, and J. Fan, *2:1 for Naturalness at the LHC?*,
 1528 JHEP **1301** (2013) 149, [arXiv:1207.4482 \[hep-ph\]](https://arxiv.org/abs/1207.4482). **2.2.2**
- 1529 [23] D. Carmi, A. Falkowski, E. Kuflik, and T. Volansky, *Interpreting LHC Higgs Results from*
 1530 *Natural New Physics Perspective*, JHEP **1207** (2012) 136, [arXiv:1202.3144 \[hep-ph\]](https://arxiv.org/abs/1202.3144). **2.2.2**
- 1531 [24] G. Degrassi, S. Di Vita, J. Elias-Miro, J. R. Espinosa, G. F. Giudice, et al., *Higgs mass and*
 1532 *vacuum stability in the Standard Model at NNLO*, JHEP **1208** (2012) 098, [arXiv:1205.6497](https://arxiv.org/abs/1205.6497)
 1533 [\[hep-ph\]](https://arxiv.org/abs/1205.6497 [hep-ph]). **2.2.2**
- 1534 [25] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) no. 08, S08001. **3.1**
- 1535 [26] T. S. Pettersson and P. Lefevre, *The Large Hadron Collider: conceptual design.*, Tech. Rep.
 1536 CERN-AC-95-05 LHC, CERN, Geneva, Oct, 1995. <https://cdsweb.cern.ch/record/291782>.
 1537 **3.1**
- 1538 [27] T. Linnecar et al., *Hardware and Initial Beam Commissioning of the LHC RF Systems.*
 1539 [oai:cds.cern.ch:1176380](https://cds.cern.ch/record/1176380), Tech. Rep. LHC-PROJECT-Report-1172.
 1540 CERN-LHC-PROJECT-Report-1172, CERN, Geneva, Oct, 2008.
 1541 <https://cdsweb.cern.ch/record/1176380>. **3.1**
- 1542 [28] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3**
 1543 (2008) S08003. **3.1**

- 1544 [29] The CMS Collaboration, *The CMS experiment at the CERN LHC*, Journal of Instrumentation
1545 **3** (2008) no. 08, S08004. [3.1](#)
- 1546 [30] The LHCb Collaboration, *The LHCb Detector at the LHC*, Journal of Instrumentation **3** (2008)
1547 no. 08, S08005. [3.1](#)
- 1548 [31] The ALICE Collaboration, *The ALICE experiment at the CERN LHC*, Journal of
1549 Instrumentation **3** (2008) no. 08, S08002.
1550 <http://stacks.iop.org/1748-0221/3/i=08/a=S08002>. [3.1](#)
- 1551 [32] A. Team, *The four main LHC experiments*, Jun, 1999. [3.1](#)
- 1552 [33] ATLAS Collaboration, *ATLAS inner detector: Technical Design Report 1*. Technical Design
1553 Report ATLAS. CERN, Geneva, 1997. <https://cdsweb.cern.ch/record/331063>. [3.2](#)
- 1554 [34] ATLAS Collaboration, *ATLAS inner detector: Technical Design Report, 2*. Technical Design
1555 Report ATLAS. CERN, Geneva, 1997. <https://cdsweb.cern.ch/record/331064>. [3.2](#)
- 1556 [35] ATLAS Collaboration, *ATLAS magnet system: Technical Design Report, 1*. Technical Design
1557 Report ATLAS. CERN, Geneva, 1997. <https://cdsweb.cern.ch/record/338080>. [3.2](#)
- 1558 [36] ATLAS Collaboration, *ATLAS pixel detector: Technical Design Report*. Technical Design
1559 Report ATLAS. CERN, Geneva, 1998. <https://cdsweb.cern.ch/record/381263>. [3.2](#)
- 1560 [37] ATLAS Collaboration, *ATLAS pixel detector electronics and sensors*, JINST **3** (2008) P07007.
1561 [3.2](#)
- 1562 [38] T. A. Collaboration, *ATLAS liquid argon calorimeter: Technical design report*, .
1563 CERN-LHCC-96-41. [3.2](#)
- 1564 [39] ATLAS Collaboration, *ATLAS tile calorimeter: Technical Design Report*. Technical Design
1565 Report ATLAS. CERN, Geneva, 1996. <https://cdsweb.cern.ch/record/331062>. [3.2](#)
- 1566 [40] ATLAS Collaboration, *ATLAS muon spectrometer: Technical Design Report*. Technical Design
1567 Report ATLAS. CERN, Geneva, 1997. <https://cdsweb.cern.ch/record/331068>. [3.2](#)
- 1568 [41] G. Aielli, A. Aloisio, M. Alviggi, V. Aprodu, V. Bocci, et al., *The RPC first level muon trigger
1569 in the barrel of the ATLAS experiment*, Nucl.Phys.Proc.Suppl. **158** (2006) 11–15. [3.2](#)
- 1570 [42] F. Bauer, U. Bratzler, H. Dietl, H. Kroha, T. Lagouri, et al., *Construction and test of MDT
1571 chambers for the ATLAS muon spectrometer*, Nucl.Instrum.Meth. **A461** (2001) 17–20. [3.2](#)
- 1572 [43] T. Argyropoulos, K. A. Assamagan, B. H. Benedict, V. Chernyatin, E. Cheu, et al., *Cathode
1573 strip chambers in ATLAS: Installation, commissioning and in situ performance*, IEEE
1574 Trans.Nucl.Sci. **56** (2009) 1568–1574. [3.2](#)
- 1575 [44] ATLAS Collaboration, *ATLAS level-1 trigger: Technical Design Report*. Technical Design
1576 Report ATLAS. CERN, Geneva, 1998. <https://cdsweb.cern.ch/record/381429>. [3.2](#)
- 1577 [45] P. Jenni, M. Nessi, M. Nordberg, and K. Smith, *ATLAS high-level trigger, data-acquisition and
1578 controls: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 2003.
1579 <https://cdsweb.cern.ch/record/616089>. [3.2](#)
- 1580 [46] *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton
1581 collision data*, Tech. Rep. ATLAS-CONF-2014-032, CERN, Geneva, Jun, 2014. [3.2.6.2](#), [7.4](#),
1582 [9.2.1](#)

- 1583 [47] *Preliminary results on the muon reconstruction efficiency, momentum resolution, and*
 1584 *momentum scale in ATLAS 2012 pp collision data*, Tech. Rep. ATLAS-CONF-2013-088, CERN,
 1585 Geneva, February, 2013. [3.2.6.3](#), [7.5](#)
- 1586 [48] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti- $k(t)$ jet clustering algorithm*, [JHEP 0804](#)
 1587 [\(2008\) 063](#), [arXiv:0802.1189 \[hep-ph\]](#). [3.2.6.4](#), [7.6](#)
- 1588 [49] ATLAS Collaboration Collaboration, G. Aad et al., *Search for $H \rightarrow \gamma\gamma$ produced in association*
 1589 *with top quarks and constraints on the Yukawa coupling between the top quark and the Higgs*
 1590 *boson using data taken at 7 TeV and 8 TeV with the ATLAS detector*, [arXiv:1409.3122](#)
 1591 [\[hep-ex\]](#). [5](#)
- 1592 [50] *Search for the Standard Model Higgs boson produced in association with top quarks and decaying*
 1593 *to $b\bar{b}$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector at the LHC*, Tech. Rep.
 1594 ATLAS-CONF-2014-011, CERN, Geneva, Mar, 2014. [5](#)
- 1595 [51] ATLAS Collaboration, G. Aad et al., *Improved luminosity determination in pp collisions at \sqrt{s}*
 1596 *= 7 TeV using the ATLAS detector at the LHC*, [Eur.Phys.J. C73 \(2013\) 2518](#), [arXiv:1302.4393](#)
 1597 [\[hep-ex\]](#). [6.1.1](#)
- 1598 [52] CERN, . CERN, Geneva, 2012. [6.1.1](#)
- 1599 [53] ATLAS Collaboration Collaboration, G. Aad et al., *The ATLAS Simulation Infrastructure*,
 1600 [Eur.Phys.J. C70 \(2010\) 823–874](#), [arXiv:1005.4568 \[physics.ins-det\]](#). [6.2](#)
- 1601 [54] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A Simulation toolkit*,
 1602 [Nucl.Instrum.Meth. A506 \(2003\) 250–303](#). [6.2](#)
- 1603 [55] LHC Higgs Cross Section Working Group Collaboration, S. Heinemeyer et al., *Handbook of*
 1604 *LHC Higgs Cross Sections: 3. Higgs Properties*, [arXiv:1307.1347 \[hep-ph\]](#). [6.2.1](#), [9.1](#), [9.1](#)
- 1605 [56] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *ALPGEN, a generator*
 1606 *for hard multiparton processes in hadronic collisions*, [JHEP 0307 \(2003\) 001](#),
 1607 [arXiv:hep-ph/0206293 \[hep-ph\]](#). [6.2.2](#)
- 1608 [57] F. Maltoni and T. Stelzer, *MadEvent: Automatic event generation with MadGraph*, [JHEP 0302](#)
 1609 [\(2003\) 027](#), [arXiv:hep-ph/0208156 \[hep-ph\]](#). [6.2.2](#)
- 1610 [58] B. P. Kersevan and E. Richter-Was, *The Monte Carlo event generator AcerMC versions 2.0 to*
 1611 *3.8 with interfaces to PYTHIA 6.4, HERWIG 6.5 and ARIADNE 4.1*, [Comput.Phys.Commun.](#)
 1612 [184 \(2013\) 919–985](#), [arXiv:hep-ph/0405247 \[hep-ph\]](#). [6.2.2](#)
- 1613 [59] P. M. Nadolsky, H.-L. Lai, Q.-H. Cao, J. Huston, J. Pumplin, et al., *Implications of CTEQ*
 1614 *global analysis for collider observables*, [Phys.Rev. D78 \(2008\) 013004](#), [arXiv:0802.0007 \[hep-ph\]](#).
 1615 [6.2.2](#)
- 1616 [60] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower*
 1617 *simulations: the POWHEG method*, [JHEP 0711 \(2007\) 070](#), [arXiv:0709.2092 \[hep-ph\]](#). [6.2.2](#)
- 1618 [61] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, et al., *Event generation with*
 1619 *SHERPA 1.1*, [JHEP 0902 \(2009\) 007](#), [arXiv:0811.4622 \[hep-ph\]](#). [6.2.2](#)
- 1620 [62] A. Martin, W. Stirling, R. Thorne, and G. Watt, *Parton distributions for the LHC*, [Eur.Phys.J.](#)
 1621 [C63 \(2009\) 189–285](#), [arXiv:0901.0002 \[hep-ph\]](#). [6.2.2](#), [8.1](#)

- [63] M. Garzelli, A. Kardos, C. Papadopoulos, and Z. Trocsanyi, *$t\bar{t}W^\pm$ and $t\bar{t}Z$ Hadroproduction at NLO accuracy in QCD with Parton Shower and Hadronization effects*, JHEP **1211** (2012) 056, arXiv:1208.2665 [hep-ph]. [8.1](#)
- [64] J. M. Campbell and R. K. Ellis, *$t\bar{t}W^{+-}$ production and decay at NLO*, JHEP **1207** (2012) 052, arXiv:1204.5678 [hep-ph]. [8.1](#)
- [65] Campbell, John and Ellis, R. Keith and Röntsch, Raoul, *Single top production in association with a Z boson at the LHC*, Phys.Rev. **D87** (2013) 114006, arXiv:1302.3856 [hep-ph]. [8.1](#)
- [66] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of WZ production in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Eur.Phys.J. **C72** (2012) 2173, arXiv:1208.1390 [hep-ex]. [8.2](#)
- [67] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of ZZ production in pp collisions at $\sqrt{s} = 7$ TeV and limits on anomalous ZZZ and $ZZ\gamma$ couplings with the ATLAS detector*, JHEP **1303** (2013) 128, arXiv:1211.6096 [hep-ex]. [8.2](#)
- [68] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of the cross-section for W boson production in association with b-jets in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, JHEP **1306** (2013) 084, arXiv:1302.2929 [hep-ex]. [8.2](#)
- [69] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of differential production cross-sections for a Z boson in association with b-jets in 7 TeV proton-proton collisions with the ATLAS detector*, arXiv:1407.3643 [hep-ex]. [8.2](#)
- [70] S. Guindon, E. Shabalina, J. Adelman, M. Alhroob, S. Amor dos Santos, A. Basye, J. Bouffard, M. Casolino, I. Connelly, A. Cortes Gonzalez, V. Dao, S. D'Auria, A. Doyle, P. Ferrari, F. Filthaut, R. Goncalo, N. de Groot, S. Henkelmann, V. Jain, A. Juste, G. Kirby, D. Kar, A. Knue, K. Kroeninger, T. Liss, E. Le Menedeu, J. Montejo Berlingen, M. Moreno Llacer, O. Nackenhorst, T. Neep, A. Onofre, M. Owen, M. Pinamonti, Y. Qin, A. Quadt, D. Quilty, C. Schwanenberger, L. Serkin, R. St Denis, J. Thomas-Wilsker, and T. Vazquez-Schroeder, *Search for the Standard Model Higgs boson produced in association with top quarks and decaying to $b\bar{b}$ in pp collisions at $\sqrt{s}=8$ TeV with the ATLAS detector at the LHC*, Tech. Rep. ATL-COM-PHYS-2013-1659, CERN, Geneva, Dec, 2013. The note contains internal documentation of the $t\bar{t}H(bb)$ analysis approved as a preliminary result (ATLAS-CONF-2014-011). [9.1](#)
- [71] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data*, arXiv:1407.3935 [hep-ex]. [9.2.1](#), [9.2.1](#)
- [72] ATLAS Collaboration Collaboration, G. Aad et al., *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, Eur.Phys.J. **C74** (2014) no. 10, 3071, arXiv:1407.5063 [hep-ex]. [9.2.1](#)
- [73] ATLAS Collaboration Collaboration, *Performance of the ATLAS muon trigger in 2011*, Tech. Rep. ATLAS-CONF-2012-099, CERN, Geneva, Jul, 2012. [9.2.1](#)
- [74] ATLAS Collaboration Collaboration, *Performance of the ATLAS Electron and Photon Trigger in p-p Collisions at $\text{sqrt}s = 7$ TeV in 2011*, Tech. Rep. ATLAS-CONF-2012-048, CERN, Geneva, May, 2012. [9.2.1](#)

- 1663 [75] ATLAS Collaboration Collaboration, G. Aad et al., *Jet energy measurement and its systematic*
1664 *uncertainty in proton-proton collisions at $\sqrt{s} = 7 \text{ TeV}$ with the ATLAS detector,*
1665 [arXiv:1406.0076 \[hep-ex\]](https://arxiv.org/abs/1406.0076). **9.2.3**
- 1666 [76] ATLAS Collaboration Collaboration, G. Aad et al., *Jet energy resolution in proton-proton*
1667 *collisions at $\sqrt{s} = 7 \text{ TeV}$ recorded in 2010 with the ATLAS detector,* *Eur.Phys.J.* **C73** (2013)
1668 **2306**, [arXiv:1210.6210 \[hep-ex\]](https://arxiv.org/abs/1210.6210). **9.2.3**
- 1669 [77] ATLAS Collaboration Collaboration, *Measurement of the Mistag Rate with 5 fb^{-1} of Data*
1670 *Collected by the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2012-040, CERN, Geneva, 2012.
1671 **9.2.4**
- 1672 [78] ATLAS Collaboration Collaboration, *Measuring the b-tag efficiency in a top-pair sample with*
1673 *4.7 fb^{-1} of data from the ATLAS detector*, Tech. Rep. ATLAS-CONF-2012-097, CERN,
1674 Geneva, 2012. **9.2.4**
- 1675 [79] ATLAS Collaboration Collaboration, *b-jet tagging calibration on c-jets containing $D^?+$ mesons*,
1676 Tech. Rep. ATLAS-CONF-2012-039, CERN, Geneva, 2012. **9.2.4**
- 1677 [80] Glen Cowan, Kyle Cranmer, Eilam Gross, Ofer Vitells, *Asymptotic formulae for likelihood-based*
1678 *tests of new physics*, *Eur.Phys.J.C* **71** (2011) 1554. **10.2**
- 1679 [81] A. L. Read, *Presentation of search results: the CL_s technique*, *Journal of Physics G: Nuclear*
1680 *and Particle Physics* **28** (2002) no. 10, 2693.
1681 <http://stacks.iop.org/0954-3899/28/i=10/a=313>. **10.2.3**