

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	List of Tables	ix
67	List of Figures	xi
68	1 Introduction	1
69	2 Theoretical Background	2
70	2.1 The Standard Model	2
71	2.1.1 The Standard Model Structure	2
72	2.1.2 Electroweak Symmetry Breaking and the Higgs	3
73	2.1.3 The Standard Model Parameters	5
74	2.2 Collider Physics and the Higgs	5
75	2.2.1 Higgs Discovery at the LHC	9
76	2.2.2 The Importance $t\bar{t}H$ Production	9
77	2.3 Conclusion	11
78	3 The Large Hadron Collider and the ATLAS Experiment	13
79	3.1 The Large Hadron Collider	13
80	3.1.1 The Accelerator Complex	13
81	3.1.2 Beam Parameters and Collisions	15

82	3.2	The ATLAS Experiment	15
83	3.2.1	Detector Coordinate System	17
84	3.2.2	The Inner Detector	17
85	3.2.3	The Calorimeter	19
86	3.2.4	The Muon Spectrometer	22
87	3.2.5	The Trigger System	22
88	3.2.6	Reconstruction: Jets, Muons and Electrons	24
89	3.2.6.1	Tracks and Clusters	24
90	3.2.6.2	Electrons	25
91	3.2.6.3	Muons	25
92	3.2.6.4	Jets	25
93	3.2.6.5	B-Tagged Jets	26
94	4	Electrons	27
95	4.1	Electrons at Hadron Colliders	27
96	4.2	Identification of Electrons at ATLAS	28
97	4.2.1	Pile-up and Electron identification	28
98	4.2.2	2011 Menu and Trigger	28
99	4.2.3	2012 Menu and Trigger	28
100	4.2.4	Electron Likelihood	28
101	4.3	Measurement of Electron Efficiency at ATLAS	28
102	4.3.1	Techniques	28
103	4.3.2	Issues	28
104	5	Analysis Summary	29
105	5.1	Signal Characteristics	29
106	5.2	Background Overview	30
107	5.3	Analysis Strategy	31
108	6	Dataset and Simulation	32
109	6.1	Data	32
110	6.1.1	The 2012 Dataset	32
111	6.2	Simulation	33
112	6.2.1	Signal Simulation	34

113	6.2.2	Background Simulation	34
114	7	Object and Event Selection	36
115	7.1	2ℓ Same-Charge Signal Region	36
116	7.2	3ℓ Signal Region	37
117	7.3	4ℓ Signal Region	38
118	7.4	Electron Selection	38
119	7.5	Muon Selection	39
120	7.6	Jet and b-Tagged Jet Selection	39
121	7.7	Object Summary and Overlap	40
122	7.8	Optimization	40
123	8	Background Estimation	42
124	8.1	Vector Boson (W^\pm, Z) production in association with top quarks: $t\bar{t}V, tZ$	42
125	8.1.1	$t\bar{t}Z$ Validation Region	45
126	8.2	Di-boson Background Estimation: $W^\pm Z, ZZ$	45
127	8.2.1	$W^\pm Z$ Uncertainty	46
128	8.2.2	ZZ Uncertainty	48
129	8.3	Charge-Misidentification Background	50
130	8.3.1	Likelihood Method	51
131	8.3.2	Results	53
132	8.3.3	Systematic and Statistical Uncertainties	53
133	8.4	Fake Lepton Backgrounds	55
134	8.4.1	2ℓ SS Fakes	56
135	8.4.2	3ℓ Fakes	60
136	8.4.3	4ℓ Fakes	64
137	9	Summary of Systematic Uncertainties	66
138	9.1	Systematic Uncertainties on Signal Cross-section and Acceptance	66
139	9.2	Experimental and Detector Systematic Uncertainties	69
140	9.2.1	Lepton Identification, Energy Scale, and Trigger	69
141	9.2.2	Lepton Isolation and Impact Parameter	70
142	9.2.3	Jet Energy	72
143	9.2.4	B-Tagged Jet Efficiency	72

144	9.2.5 Summary	73
145	9.3 Summary of Background and Signal Normalization Uncertainties	73
146	10 Results and Statistical Model	75
147	10.1 Results in Signal Regions	75
148	10.2 Statistical Model	75
149	10.2.1 The Likelihood	75
150	10.2.2 Test Statistic and Profile Likelihood	76
151	10.2.3 CL_s Method	76
152	10.2.4 Exclusion Limits	77
153	10.2.5 μ Measurements	77
154	10.2.6 Nuisance Parameter Impact on the Signal Strength	78
155	11 Conclusions	79
156	11.1 Higgs Results in Review	79
157	11.2 Prospects for Future	79
158	Bibliography	80

List of Tables

160	5.1	Contributions of the main Higgs decay modes to the 3 multi-lepton $t\bar{t}H$ signatures at generation level.	30
162	6.1	Monte Carlo samples used for signal description.	34
163	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. $t\bar{t}$, single top, and $Z+jets$ samples are replaced with data-driven estimates for the final result	35
166	7.1	Selections in the 2ℓ SS, 3ℓ and 4ℓ Signal Regions	37
167	7.2	Object identification and selection used to define the 5 channels of the multi-lepton $t\bar{t}H$ analysis. Some channels use a sub-sample of objects as explained in the Remarks column.	41
169	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 $t\bar{V}$ and 30% for $t\bar{t}\text{fake}$).	43
173	8.2	NLO cross section and theoretical uncertainty calculations derived from MadGraph5_aMC@NLO.	44
174	8.3	Truth Origin of highest energy b-tagged jet in the $W^\pm Z + b$ VR and $3l$ SR	48
175	8.4	Truth Origin of highest energy b-tagged jet in the $ZZ + b$ VR and $4l$ SR	50
176	8.5	Expected and measured values of the θ factors.	58
177	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 ($\Delta N_{\ell\text{anti}-\ell}(n \text{ jets})$). For channels with muons, uncertainties between parenthesis are obtained when anti-tight muons are defined such as $p_T < 20$ GeV (includes blinded data)	59
182	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	63
184	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.	68
186	9.2	Uncertainties on $t\bar{t}H$ acceptance in signal regions due to PDF variation.	68
187	9.3	Sum in quadrature of all the systematic uncertainties on the number of event yields per channel.	73

189	9.4	Summary of systematic uncertainties 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.	74
192	10.1	95%CL limits on μ for all channels and combination with cumulative uncertainties.	77

List of Figures

194 2.1	The Standard Model Particle Content	4
195 2.2	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.	6
201 2.3	Proton Parton Distribution Functions (PDFs) form the MSTW Collaboration at $Q^2 = 10$ GeV 2 and $Q^2 = 10^4$ GeV 2	7
203 2.4	Dominant Higgs production modes at the LHC	8
204 2.5	8 TeV LHC Higgs production cross-sections (left) and decay branching fractions	8
205 2.6	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	10
210 2.7	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	11
214 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.	14
218 3.2	Diagram of the ATLAS detector and subsystems	16
219 3.3	Diagram of the ATLAS ID in the $R - \phi$ plane showing the barrel view of the Pixel, SCT, and TRT detectors.	18
221 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 endcap is shown.	18
223 3.5	Diagram of the ATLAS calorimeters	20
224 3.6	Diagram of the ATLAS LAr EM calorimeter showing the longitudinal segmentation and the $\eta - \phi$ cells for the central barrel region	21

226	3.7	Diagram of the ATLAS muon system	23
227	3.8	$R - \phi$ schematic of the ATLAS detector and various particle signatures	26
228	5.1	Example feynman diagrams for the 3 $t\bar{t}H$ multi-lepton categories.	30
229	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.	33
230	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	33
235	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.	46
240	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	47
241	8.3	$W^\pm Z + b$ validation region: NJet, NElec, BJet MV1 and Mass Variables	48
243	8.4	Jet-inclusive 4-lepton ZZ validation region using the $t\bar{t}H$ lepton identification and momentum cuts	49
245	8.5	$ZZ + b$ validation region using the $t\bar{t}H$ lepton identification and momentum cuts	50
246	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.	54
253	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 re-weighted 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.	54
258	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.	55
260	8.9	Ratios of regions with tight and anti-tight leptons in 2-lepton and 3-lepton channels	57
261	8.10	2,3 Jet SS $2\ell ed$ (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 $t\bar{t}\text{MC}$ (red) is used for reference and not used in the actual calculation	58
264	8.11	4,5 Jet SS $2\ell ed$ (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 $t\bar{t}\text{MC}$ (red) is used for reference and not used in the actual calculation	59

268	8.12 Distributions of the properties of the anti-tight muons in data (dots), compared with 269 the total simulation (red line), rescaled to the integral of the data for a shape compari- 270 son. The uncertainty on the data distribution is statistical. The number of events for 271 each of them is also presented in the legend. The variables probed are, top: p_T and 272 $\Delta R(\mu, \text{closest selected jet})$; bottom: ptcone20/ p_T and Etcone20/ p_T . The selection is the 273 signal region event selection with one anti-tight muon (failing at least one of the isolation, 274 muon-jet overlap, or p_T selection criteria) A ratio plot is containing the 20% area around 275 1, is displayed, demonstrating that a 20% data-MC comparison systematic is sufficient.	61
276	8.13 Distributions of anti-tight electron variables. The variables presented are, from top left 277 to bottom right, p_T , η , Very Tight Likelihood word value, ptcone20/ p_T , Etcone20/ p_T . 278 The plotted regions have the same cuts as the signal region, except the anti-tight electron 279 must fail isolation for the plot of the verytight identification word or fail the verytight 280 identification word for the plots of the isolation. Data (dots) are compared with a stacked 281 histogram of the various simulated samples: top in red, V+jets (blue), VV (purple) and 282 ttV (yellow). The uncertainty on the data distribution is statistical.	62
283	8.14 Muon-xxp (above) and electron-xxd (below) anti-tight control regions: jet variables. Note: 284 the $t\bar{t}$ and top MC in the plots is used only as comparison, but is not included in the fake 285 measurement	63
286	8.15 3ℓ fake validation regions for nominal p_T selection (above) and relaxed p_T selection, > 10 287 GeV, (below). Plotted are the number of jets and the number of electrons in each event. 288 The data and MC ratio below each plot agree with 1 within the statistics of the region 289 and the overall systematic assigned for the fake component (red)	64
290	9.1 Effects on the jet multiplicities in 2 SS lepton $t\bar{t}H$ events from different choices of the 291 factorization and renormalization scales. “Static” refers to the variations by a factor of 2 292 of the nominal μ_0 , while “Dynamic” refers to the alternative choice of μ_0 which depends 293 on the event kinematic. The grey band in the lower panels represents the statistical 294 uncertainty of the nominal sample.	68
295	9.2 PDF systematic uncertainty on the jet multiplicities in $t\bar{t}H$ events with at least 2 leptons. 296 The dashed red lines in the lower panel indicate the systematic uncertainty on the $t\bar{t}H$ 297 production cross section.	69
298	9.3 Muon (left) and electron(right) identification efficiencies in Data and MC as a function of 299 η and p_T respectively. For electrons, the verytight likelihood operating point is used and 300 for muons the CB+ST (combined+segment tagged) operating point is used	70
301	9.4 Muon (left) and electron(right) isolation efficiency scale-factors from the Z control sample 302 as a function of the number of jets in the event. An additional systematic uncertainty of 303 1% is added to encompass the variation in the number of jets variable	71
304	9.5 JES systematic uncertainties as a function of jet η (for jets $p_T > 25$ GeV) and p_T (for jets 305 $ \eta < 0.4$). The combined systematic uncertainty is shown with contributions from the 306 largest sources	72
307	9.6 b-Tagging data-MC efficiency scale-factors versus jet p_T calculated in the $t\bar{t}$ sample from 308 2012 data. The uncertainties are combined statistical and systematic.	73

309

CHAPTER 1

310

Introduction

311

CHAPTER 2

312

Theoretical Background

313 The Standard Model of particle physics (SM) is an extraordinarily successful description of the fundamental constituents of matter and their interactions. Many experiments have verified the extremely 314 precise prediction of the SM. This success has culminated most recently in the discovery of the Higgs 315 Boson. This chapter provides a brief introduction to the structure of the SM and how scientists are 316 able to test it using hadron collider. It focuses primarily on the physics of the Higgs boson and its 317 decays to top quarks. I stress the importance of a measurement of the rate at which Higgs Bosons 318 are produced in association of top quarks, as a new, rigorous test of the SM. The experimental search 319 for this production mode in multi-lepton final states is the general subject of this thesis. 320

321 **2.1 The Standard Model**

322 **2.1.1 The Standard Model Structure**

323 The Standard Model (SM) [1, 2, 3, 4] is an example of a quantum field theory that describes the 324 interactions of all of the known fundamental particles. Particles are understood to be excitations of 325 the more fundamental object of the theory, the field. The dynamics and interactions of the fields are 326 derived from the Standard Model Lagrangian, which is constructed to be symmetric under transfor- 327 mations of the group $SU(3) \times SU(2)_L \times U(1)$. $SU(3)$ is the group for the color, $SU(2)_L$ is the group 328 for weak iso spin, and $U(1)$ is the group for weak hyper-charge.

329 Demanding these symmetries be local, gauge symmetries allows the theory to be re-normalizable 330 [5], meaning that unwanted infinities can be absorbed into observables from theory in a way that 331 allows the theory to be able to predict physics at multiple energy scales. Gauging the symmetries

332 results in the introduction of 8 massless gluons, or the boson¹ carriers of the strong force [6] from the
 333 generators $SU(3)$ symmetry, and the 4 massless bosons, carriers for the weak and electromagnetic
 334 forces from the 3 generators of the $SU(2)$ and 1 generator of the $U(1)$ group. The weak and the
 335 electromagnetic forces are considered part of a larger single unified electroweak group $SU(2) \times U(1)$
 336 and the associated generators mix.

337 Matter particles are fermion particles, defined as representations of the symmetry groups. Singlets
 338 of the $SU(3)$ are called leptons, do not have a color charge, and, therefore, do not interact with the
 339 strong force. Quarks, on the other hand, are triplets of the $SU(3)$ group do interact with the strong
 340 force. The SM is a chiral theory: the weak force violates parity, as it only couples to left-chiral
 341 particles or right-chiral antiparticles. This means that right-chiral and left-chiral fermions arise from
 342 different fields, which are different representations of the $SU(2)_L$ group.

343 The discovery of particles and new interactions in various experiments is intertwined with the
 344 development of the theory that spans many decades and is not discussed in detail here. But these
 345 experiments have proven the above model and symmetries to be an overwhelming success. So far, 3
 346 separate generations of both quarks and leptons have been discovered, differing only by mass. The
 347 gluons and the 4 electroweak bosons have also been discovered (W^+ , W^- , Z^0 , and γ). The reason for
 348 this 3-fold replication is not known. Figure 2.1 shows a table of the known SM particle content.

349 2.1.2 Electroweak Symmetry Breaking and the Higgs

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

356 To solve these problems, spontaneous electro-weak symmetry breaking (EWSB) is introduced via
 357 the Brout-Englert-Higgs mechanism [7, 8, 9]. A massive scalar field in an electro-weak doublet is
 358 added to the theory with 4 new degrees of freedom and a potential which includes a quartic self-
 359 interaction term. Each fermion field interacts with the scalar field via a different Yukawa coupling,
 360 which unites the left and right chiral fields of a single particle type. This field explicitly preserves
 361 all of the symmetries, but the minimum of the potential does not occur when the expectation of the

¹bosons are full integer spin particles that obey Bose-Einstein statistics, while fermions are half-integer spin particles that obey Fermi-Dirac statistics

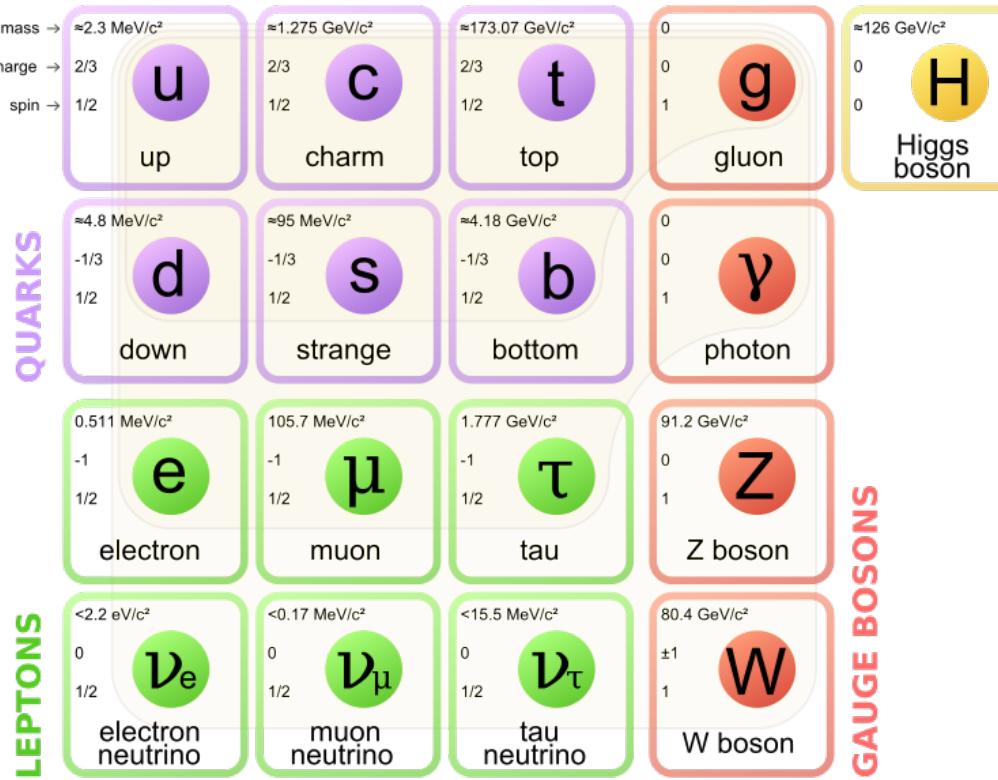


Figure 2.1: The Standard Model Particle Content

362 field is zero. The field eventually falls to a state, where it acquires a non-zero vacuum-expectation
 363 value. A non-vanishing field must point in a particular direction of weak-isospin space, breaking the
 364 symmetry.

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

372 2.1.3 The Standard Model Parameters

373 Confronting the SM with experiment requires the measurement of 17^2 free parameters, which are
 374 unconstrained from the theory. These free parameters include the fermion masses from the Yukawa
 375 couplings, the force coupling constants, the angles and phase of the mixing between quarks, and
 376 constants from the Higgs and electroweak sector³.

377 Experiments have provided a number of measurements of the parameters of the SM[10]. With
 378 the discovery of the Higgs boson and the measurement of the Higgs mass, all of the parameters of
 379 the SM can be estimated and statistical procedures can assess the relative agreement of overlapping
 380 measurements to test the self-consistency of the SM. The GFitter collaboration assembles all relevant
 381 electroweak observable measurements into a statistical model and then allows certain measurements
 382 to float within their uncertainty to allow for a fit among multiple correlated measurements[11]. These
 383 correlations arise for two reasons. First, measurements are made that often depend on multiple SM
 384 parameters. Second, radiative corrections often cause parameters to depend on each other. For
 385 instance, the Higgs mass is sensitive to both the W mass and top mass, through loop level corrections.

386 Figure 2.2 shows the fitted constraints on 4 key SM parameters (M_H , M_W , M_t , $\sin^2\theta_w$) with actual
 387 measurements overlaid. The plots show both the removal and inclusion in the fit of key measurements
 388 to assess their overall impact. The addition to the fit of the measured Higgs mass from the ATLAS and
 389 CMS collaborations creates a small tension, as the other observables prefer the mass to be much lower
 390 (~ 80 GeV). This tension in the combined electroweak fit as a result is not statistically significant
 391 with a p -value of 0.07. The SM seems to be self-consistent.

392 2.2 Collider Physics and the Higgs

393 To test the theory, physicists accelerate particles to extremely high energies and force them to interact
 394 through collisions. Typically, the particles accelerated are electrons or protons, since they are stable.
 395 Electron-positron collider machines have a rich history of discovery and measurement in particle
 396 physics. The advantage of electron accelerators is that the colliding element is itself a fundamental
 397 particle. However, due to synchrotron radiation, the curvature of the beam line becomes problematic
 398 for high energy beams. On the other hand, proton-proton and proton-anti-proton colliders can be
 399 accelerated in rings without large losses due to synchrotron radiation, but the actual colliding objects

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

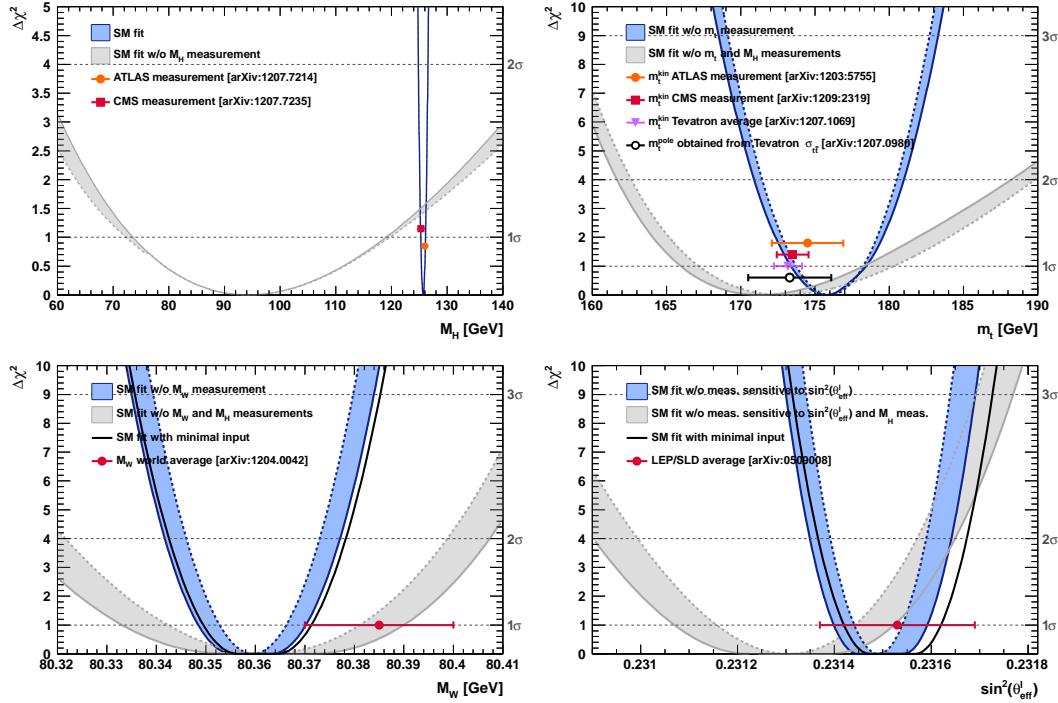


Figure 2.2: χ^2 as a function of the Higgs mass (top left), the top quark mass (top right), the W boson mass (bottom left) and the effective weak mixing angle (bottom right) for the combined SM fit from the GFitter group. The data points placed along $\chi^2 = 1$ represent direct measurements of the 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.

at high energies are the constituent quarks and gluons. This complicates analysis because the initial state of the hard-scattering system is not known on a per-collision basis and momentum of hard-scattering system is unknown along the beam direction.

For hadron colliders, physicists must rely on form-factor descriptions of the colliding hadrons that describe the fraction of momentum carried by the hadrons constituent ‘partons’. These are called parton-distribution functions, seen in Figure 2.3, and are factorized and integrated through the theoretical calculations of various collision processes [12].

At the Large Hadron Collider or LHC, the collider used in this thesis, protons are the colliders. The types of initial hard-scattering states at the LHC 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

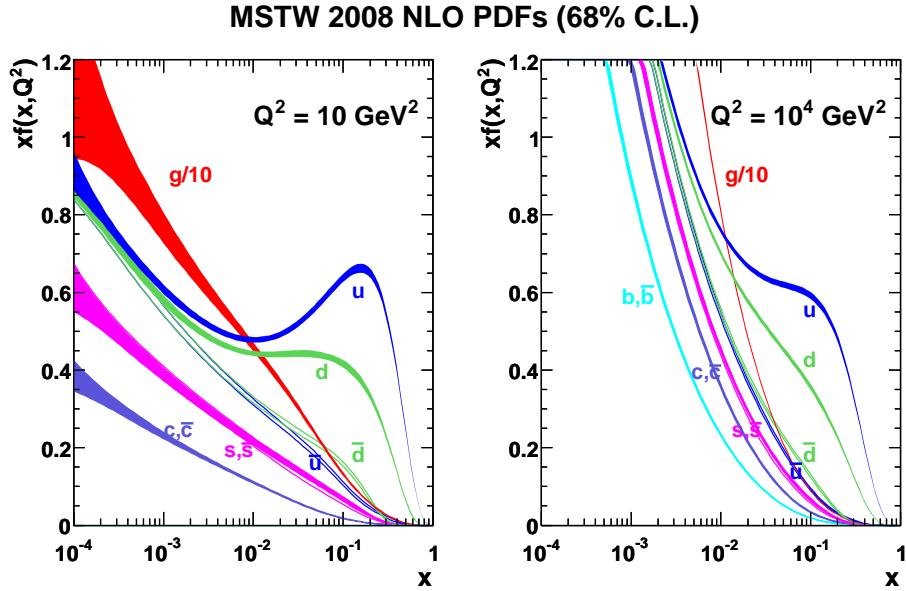


Figure 2.3: Proton Parton Distribution Functions (PDFs) form the MSTW Collaboration at $Q^2 = 10 \text{ GeV}^2$ and $Q^2 = 10^4 \text{ GeV}^2$

of the Higgs boson[13]. LEP and the Tevatron excluded large swaths of possible Higgs boson masses, especially below 114 GeV. The Higgs mass was also known to have a theoretically motivated upper bound. The unitarity of diagrams including the $WWWW$ vertex required the Higgs mass to be below about 1 TeV. This LHC was thus designed to be able to eventually find or exclude a Higgs particle in this range [10].

Reaching this discovery or exclusion required an enormous dataset with collisions at high energies. Despite the fact that the Higgs couples to nearly every particle, 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 2.4 and the production cross-sections as a function of Higgs mass for the 8 TeV LHC proton-proton running are shown in Figure 2.5 [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. Because the Higgs' couples to every massive particle, it has a rich set of decays also seen in Figure 2.5, especially for $m_H = 125$. Studies of Higgs properties at hadron colliders offers many tests of

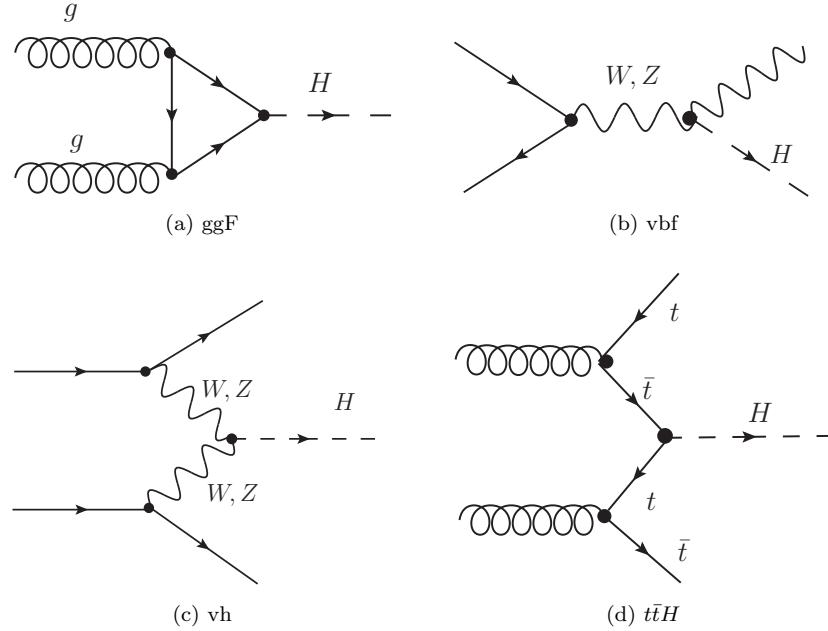


Figure 2.4: Dominant Higgs production modes at the LHC

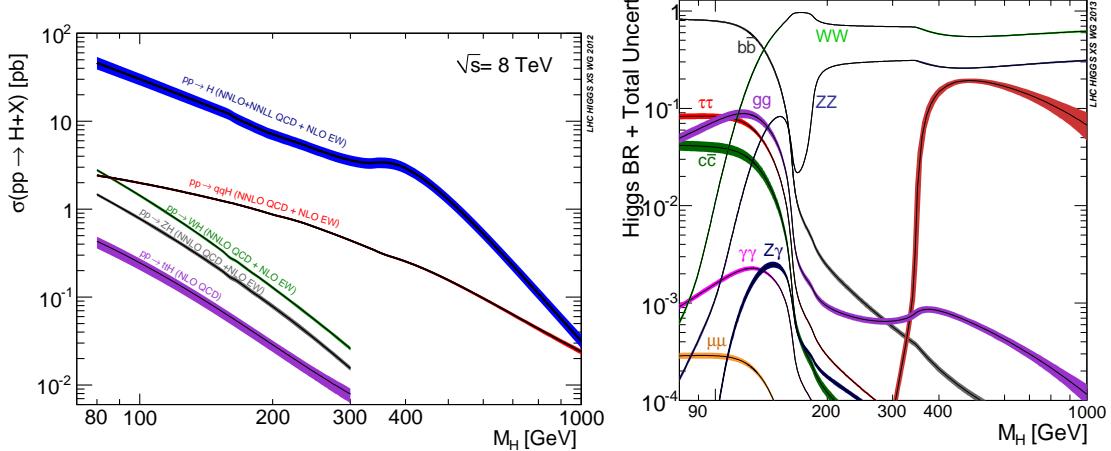


Figure 2.5: 8 TeV LHC Higgs production cross-sections (left) and decay branching fractions

430 the Standard Model and ample room for searches for new physics. These tests specifically can verify
 431 the link between Yukawa coupling and the particles mass and further constrain details of EWSB by
 432 examining Higgs coupling to the weak bosons.

433 **2.2.1 Higgs Discovery at the LHC**

434 In 2012 both ATLAS and CMS announced the discovery of a new boson consistent with the Higgs
 435 by examining the results of Higgs searches in a number of decay channels ($H \rightarrow W^+W^-$, $H \rightarrow Z^0Z^0$,
 436 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
 437 2013 and 2014, both experiments have updated and/or finalized their results for the full 2011 and
 438 2012 datasets [15, 16]. I will focus on the ATLAS results in the following. ATLAS measured both the
 439 Higgs mass[17] and spin[18], as well as provided initial constraints of the Higgs couplings to different
 440 particles.

441 Figure 2.6 show the results of the searches in all of the measurement channels as well as constraints
 442 on the SM Higgs coupling parameters in an example fit, where the couplings to the top-quark, bottom-
 443 quark, W,Z, and τ are allowed to fluctuate independently. These rely on measurements binned in
 444 different production and decay channels. They are dominated by higher statistics results in the gluon-
 445 fusion production modes, but measurements in the VH and VBF modes are close to SM sensitivity.

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

450 **2.2.2 The Importance $t\bar{t}H$ Production**

451 Notably absent thus far in the SM are searches for the Higgs in the $t\bar{t}H$ production channel, due to
 452 the low production rate and lack of statistics. Searches are underway and initial results are close to
 453 SM sensitivity for ATLAS and CMS.

454 Measuring the $t\bar{t}H$ production rate is important, because $t\bar{t}H$ production depends on the top
 455 Yukawa coupling at tree level. Comparing the predicted Yukawa coupling from top mass measurements
 456 to the coupling from the wholly independent Higgs production measurements is a very direct test of
 457 Higgs' involvement in providing mass the fermions in the SM.

458 The top Yukawa coupling is already constrained from current measurements of the ggF production
 459 process, since ggF loop is dominated by the top quark. However, new, colored particled could be
 460 present in the loop. Comparison of the gluon-fusion and the $t\bar{t}H$ modes would allow for disentangling
 461 the effects of these possible new particles[19]. The simplest of new phyiscs models, allowing for the
 462 modification of the ggF loop, introduce a new generation of quarks. However, fourth generation
 463 quarks, which obtain mass from a Higgs Yukawa coupling, are already largely excluded due to their
 464 enormous effects on the Higgs production cross-section[20]. Other exotic scenarios allow for new

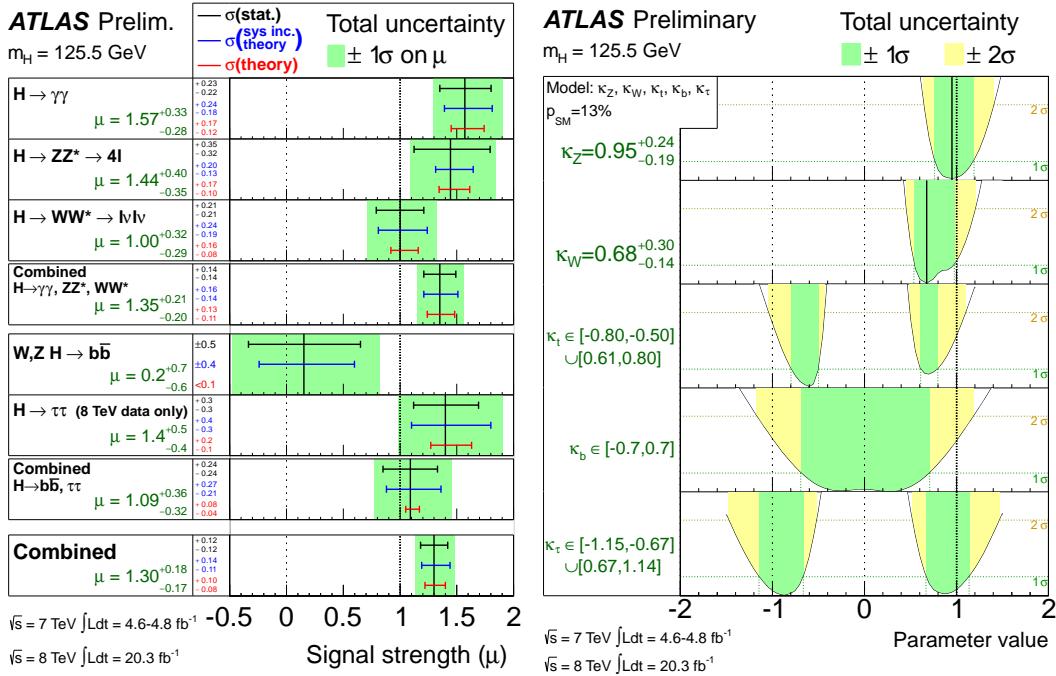


Figure 2.6: ATLAS Higgs combination results for all SM measurement channels as ratios of the measured to SM production cross-sections (left) and extracted Higgs coupling constraint scale-factors for a combined fit to the measurement channels, where the W, Z , top-quark, b-quark, and τ couplings are allowed to float. The p-value of this particular model is 0.13 and in agreement with SM expectations

465 colored particles, which are not entirely constrained by present measurements[21, 22, 23]. These
466 include, for instance, supersymmetric models involving the stop quark.

467 With the level of statistics available in Run I dataset, very strict constraints on the top Yukawa
468 coupling are simply not possible and the measurement presented in this thesis is a first step. Future,
469 high-statistics datasets will have the ability to provide better measurements and $t\bar{t}H$ production will
470 become very important. Despite similar uncertainties on the overall production cross-sections for $t\bar{t}H$
471 and the ggF, $t\bar{t}H$ has the advantage that most of these uncertainties would cancel for $t\bar{t}H$ if normalized
472 to the topologically similar $t\bar{t}Z$. Finally, the uniqueness of the experimental signature means that
473 searches for $t\bar{t}$ signatures can be performed for a variety of Higgs decays ($\gamma\gamma, b\bar{b}, WW, ZZ$, and $\tau\bar{\tau}$)
474 with roughly similar degrees of sensitivity (within a factor of 10)[19].

475 It is important to note the importance of the top Yukawa coupling due to its enormous size
476 compared to other couplings. For instance, the top Yukawa is 350000x as large as the electron
477 Yukawa coupling. The top Yukawa coupling, along with the Higgs mass, is one of the most important
478 pieces of the renormalization group equations (RGE) responsible for the running of the parameter that
479 determines the Higgs self-coupling λ . If this parameter runs negative, then the potential responsible

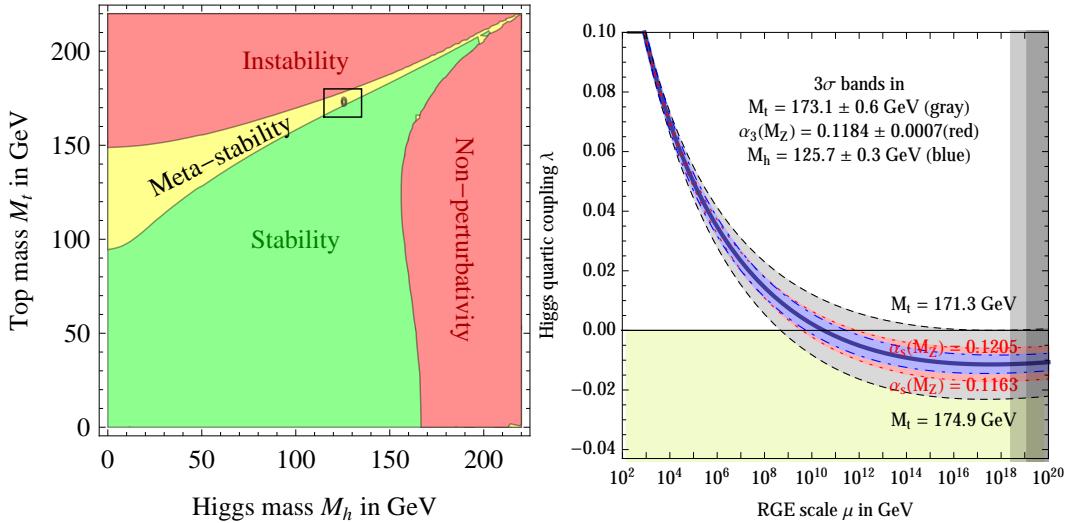


Figure 2.7: RGE for the running of the SM parameter, λ for the Higgs self-coupling term with present values and uncertainty bands for M_H and M_t (left). The two-dimensional plot colored (right) shows regions for which the SM is stable, unstable and metastable based on this RGE

for the entire mechanism of EWSB no longer has a minimum and becomes unbounded, resulting in instability in the universe [24]. Metastability occurs when the shape of the potential allows for a false local minimum. Figure 2.7 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. This is a sort of fanciful aside, intended only to highlight the importance of the top Yukawa coupling and to suggest that new discoveries in the top-Higgs sector have far reaching consequences.

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

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

496 topic of this thesis, provide tree-level access to a central parameter of the theory, the top Yukawa
497 coupling, as well as access a variety of Higgs decays, which will eventually provide a rigorous new test
498 of the SM.

CHAPTER 3

The Large Hadron Collider and the ATLAS Experiment

502

3.1 The Large Hadron Collider

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

506 The LHC is the world's highest energy particle accelerator and is located 100m underneath the
 507 Franco-Swiss border at the European Organization for Nuclear Research (CERN) in a 26.7 km tunnel.
 508 The technology involved in the development of the LHC is an enormous achievement in its own right
 509 and is documented in detail here [25, 26, 27]. The LHC is a circular machine capable of accelerating
 510 beams of protons and colliding them at center of mass energies up to $\sqrt{s} = 14$ TeV at 4 collision sites
 511 around the ring, where 4 experiments are housed (ATLAS[28], CMS[29], LHCb[30], and ALICE[31]).
 512 Figure 3.1 is a diagram of the layout of the LHC and its experiments[32]. The LHC also operates in a
 513 modes with beams of heavy ions. The LHC is composed of thousands of super-conducting Niobium-
 514 Titanium magnets, cooled to 2.7° C with liquid Helium, which steer and focus the particle beams,
 515 and a superconducting resonant-frequency (RF) cavity, which boosts the beam to higher energies.

516

3.1.1 The Accelerator Complex

517 The accelerator complex is a progressive series of machines with the LHC as the final stage. Protons
 518 are obtained from hydrogen atoms and are accelerated to 50 MeV using the Linac2, a linear acceler-
 519 ator, before being injected into the Proton-Synchrotron Booster (PSB). In the PSB the protons are
 520 accelerated to energies of 1.4 GeV for injection into the Proton-Synchrotron (PS). The PS accelerates

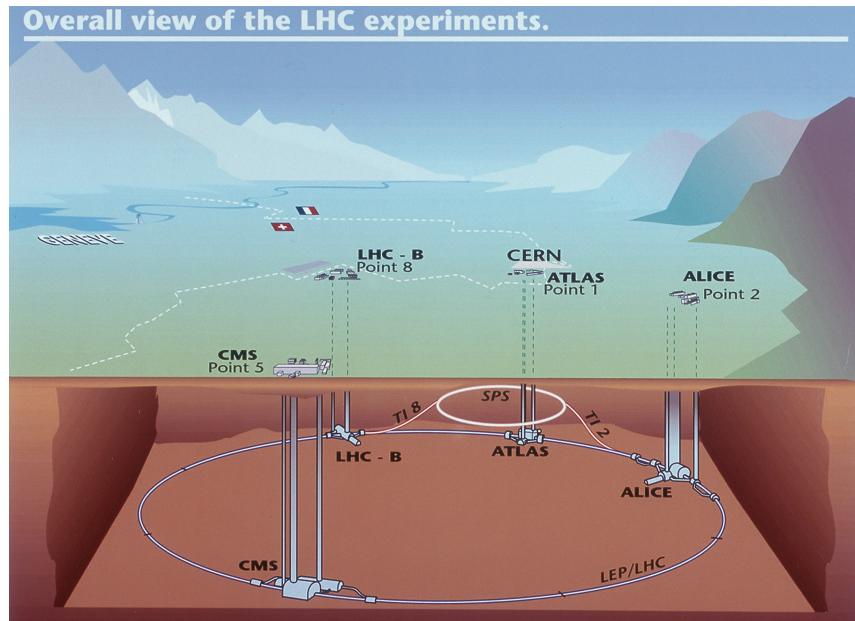


Figure 3.1: Diagram of the Large Hadron collider and location of the 4 main experiments (ATLAS, CMS, LHCb, and ALICE). Around the ring. The diagram also shows the location of the SPS, the final booster ring the accelerator complex that accelerates the protons to 450 GeV before injection into the LHC.

521 the protons to 25 GeV and dumps bunches into the Super Proton Synchrotron (SPS), where they are
 522 accelerated to 450 GeV and finally dumped into the LHC for full acceleration. The PS and SPS are
 523 circular accelerators that were important in past physics discoveries and have been re-purposed for
 524 use in the LHC complex.

525 **3.1.2 Beam Parameters and Collisions**

526 For the physics studied at the ATLAS experiment, the two most important parameters of the collisions
527 are the center of mass energy and instantaneous luminosity. High center of mass energies are necessary
528 for the production of new high mass particles, and, because the constituents of the actual collisions
529 are the partons of the proton, the CME of the collisions must in general be much higher than the
530 mass of the particles needed to be produced.

531 The instantaneous luminosity of the collisions, \mathcal{L} , is a measure of the collision rate. The integrated
532 luminosity over time is a measure of the size of the dataset and when multiplied by the cross-section of a
533 particular process gives the total number of expected events produced for that process. Instantaneous
534 luminosity depends on the number of colliding bunches of protons, the intensity of those bunches, the
535 revolution frequency, and the normalized transverse spread of the beam in momentum and position
536 phase space, called the emittance, and the transverse beam size. The LHC has the option for colliding
537 beams with 2808 bunches of protons, each with around 10^{11} protons, at a rate of one bunch collision
538 every 25 ns, or 40 MHz. These parameters correspond to a design luminosity of around $10^{34} \text{ cm}^2 \text{ s}^{-1}$
539 or $10 \text{ nb}^{-1} \text{ s}^{-1}$, or 1 Higgs every 5 seconds.

540 **3.2 The ATLAS Experiment**

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

546 ATLAS, shown in Figure 3.2, possesses cylindrical symmetry around the beam pipe. It weights
547 7000 tons, has a diameter of roughly 25m and length of 46m. ATLAS was designed to be a multi-
548 purpose hermetic, particle detector, able to identify many types of particles, and designed to provide
549 a snapshot of the entire collision event. The detector sub-systems form concentric rings around the
550 beam-line at increasing distance. From closest to the beam outward, they are:

- 551 • **Inner Detector:** The inner detector (ID)[33, 34] is immersed in a solenoidal magnetic field[35]
552 and provides measurements of charge particle tracks, through three subsystems: the Pixel
553 Detector[36, 37], the Semi-Conductor Tracker (SCT)[38, 39], and Transition Radiation Tracker
554 [40, 41, 42].

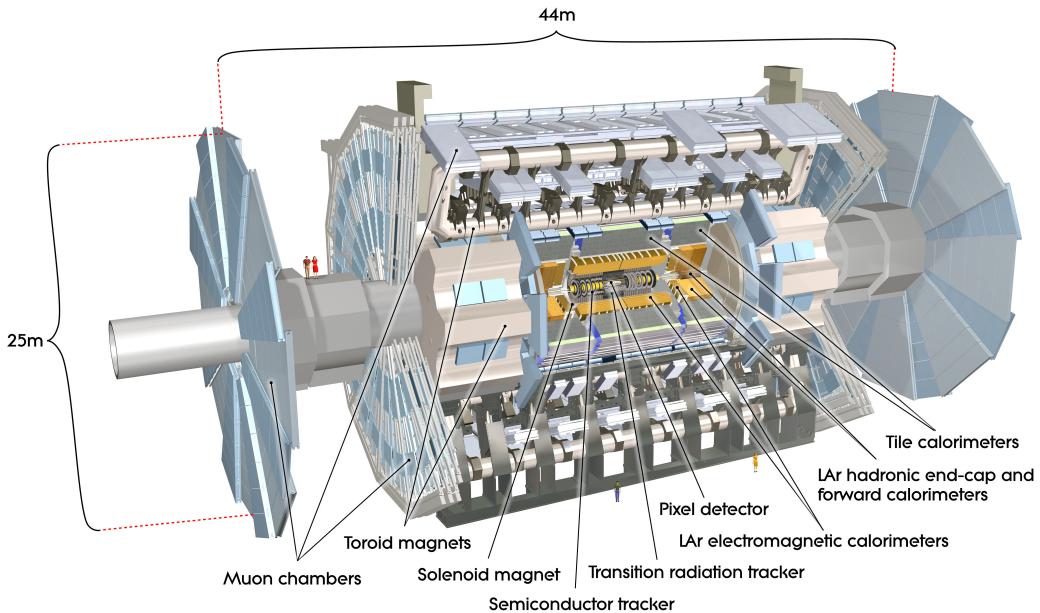


Figure 3.2: Diagram of the ATLAS detector and subsystems

- **Calorimeter:** The calorimeters measure the energy of particles that participate in the electromagnetic (photons, electrons) and hadronic interactions (pions, protons, neutrons, etc.), by forcing them to shower in dense material. The hermeticity of the calorimeters allows for missing transverse energy measurements. The calorimeter is composed of the liquid argon electromagnetic calorimeter (LAr)[43], the hadronic tile calorimeter[44], the liquid argon hadronic endcap calorimeter, and the forward calorimeters.
- **Muon Spectrometer:** The muon spectrometer (MS) sub-systems[45] 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 [46, 47, 48].
- **Triggering Systems:** The trigger and data acquisition systems[49, 50] 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.

570 **3.2.1 Detector Coordinate System**

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

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

585 The 'barrel' and 'endcap' are classifications that are used to label the position of sub-detectors.
586 Barrel sub-detectors occupy positions more central to the detector at $|\eta|$ values roughly less than
587 1-2, while the endcap calorimeters extend farther in $|\eta|$. The barrel-endcap transition region contains
588 detector services. Also, the orientation of the detector elements are often different in the barrel and
589 endcap to have optimal particle flux.

590 **3.2.2 The Inner Detector**

591 The ID makes measurements of the position of charged particles as they move through the detectors
592 3 sub-systems (Pixel, TRT, SCT). The individual position measurements can be strung together
593 to form a particle track. The entire ID is immersed in a 2T solenoidal magnetic field allowing for
594 measurements of particle momenta through the curvature of the tracks. The ID is contained with
595 a radius of 1.15 m and has a total length of 7m, allowing for particle tracking out to $|\eta| < 2.5$.
596 Figures 3.3 and 3.4 show the placement of the ID sub-systems in the $R - \phi$ and $R - z$ planes.

597 The Pixel detector has 80 million silicon read out channels (pixels) and is closest to the interaction
598 point with the finest granularity. As charged particles traverse the silicon, they create electron-hole
599 pairs, which subsequently drift in an electric field and can be captured and registered as a current
600 pulse. The detector has three concentric layers of pixels in the barrel (to $|\eta| < 1.9$) and three endcap
601 disks on each side of the barrel (to $|\eta| < 2.5$). The closest barrel layer to the beam pipe is called the

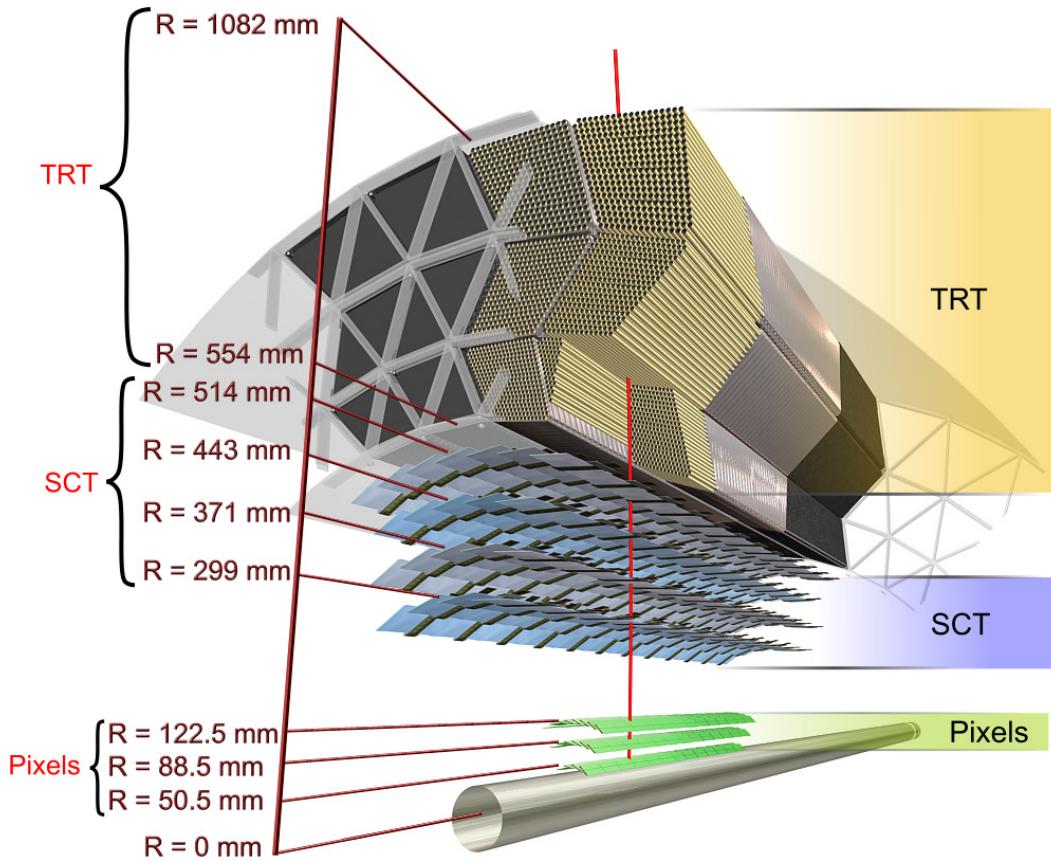


Figure 3.3: Diagram of the ATLAS ID in the $R - \phi$ plane showing the barrel view of the Pixel, SCT, and TRT detectors.

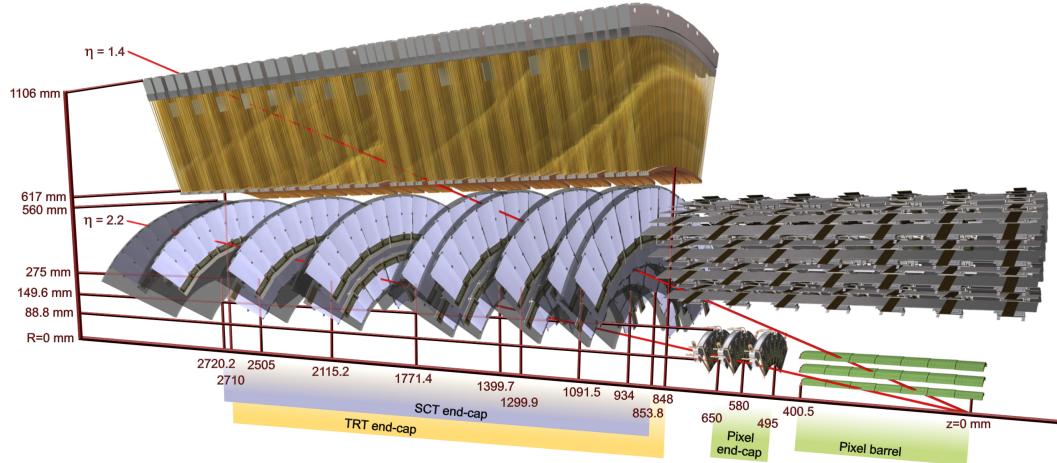


Figure 3.4: Diagram of the ATLAS ID in the $R - z$ plane showing the endcap view of the Pixel, SCT, and TRT detectors. Only one side of the endcap is shown.

602 b-layer. The pixels provide excellent hit resolution ($R - \phi$ accuracy of $10\ \mu\text{m}$ and $z(R)$ accuracy of
603 $115\ \mu\text{m}$ in the barrel (endcap)).

604 The SCT uses similar silicon technology to the pixels, but the each SCT layer contains a double
605 layer of silicon strips, which are much longer in length than width. The SCT has 4 million read out
606 channels and is comprises 4 barrel layers and 9 endcap layers with coverage to $|\eta| < 2.5$. The double
607 layers are inclined slightly with respect to each other so that these 1D sensors have 2D resolution for
608 coincident hits. The resolutions are $580\ \mu\text{m}$ in $z(R)$ for the barrel(endcap) and $17\ \mu\text{m}$ in $R - \phi$.

609 The TRT is comprised of straw drift tubes filled with a gas mixture, containing mostly Xenon
610 gas. Particles traversing the straws ionize the gas, and the liberated electrons drift to a wire at the
611 center of the straw, which has an applied voltage, and induce an signal on the wire. The TRT has
612 $\sim 300,000$ straws . The barrel straws are arranged cylindrically along the z direction out to $\sim \eta < 1$
613 and the endcap straws point radially outward in the R direction. For this reason, the barrel(endcap)
614 straws provide no measurement in the $R(z)$ directions. The drift tubes provide individual position
615 measurements with resolutions of $\sim 130\ \mu$. Each particle track has on average a large number, 35,
616 hits.

617 The TRT is unique in that it also provides particle identification measurements via transition
618 radiation. Charged particles emit transition radiation when traversing a boundary between materials
619 of different dielectric constants. The volume between the straws is filled with a radiator material, a
620 polymer foil or foam, to provide this boundary condition. Transition radiation photons are emitted
621 in the direction of the particle trajectory in the keV range and cause a much larger signal amplitude
622 within the straw. Hits that cause a signal at a higher threshold are thus indicative of transition
623 radiation. The probability for emission transition radiation depends on the relativistic γ of the
624 traversing particle. Because electrons are much lighter than any other charged particle, their γ -
625 factors tend to be high enough to induce transition radiations, as opposed to pions, muons and other
626 particles.

627 Combined tracking of particles through the 3 sub-detectors results track momentum measurements
628 from 500 MeV, the minimum energy need to leave the ID due to the magnetic field, and a few TeV.
629 The track p_{T} resolution is roughly $0.05\% \cdot p_{\text{T}} \oplus 1\%$.

630 3.2.3 The Calorimeter

631 The ATLAS calorimeters measure the energy of electron, photons and hadrons with $|\eta| < 4.5$. They
632 induce a particle shower via electro-magnetic and nuclear interactions with the detector material and
633 are deep enough to ensure that all or most of the shower energy remains contained. Muons are min-

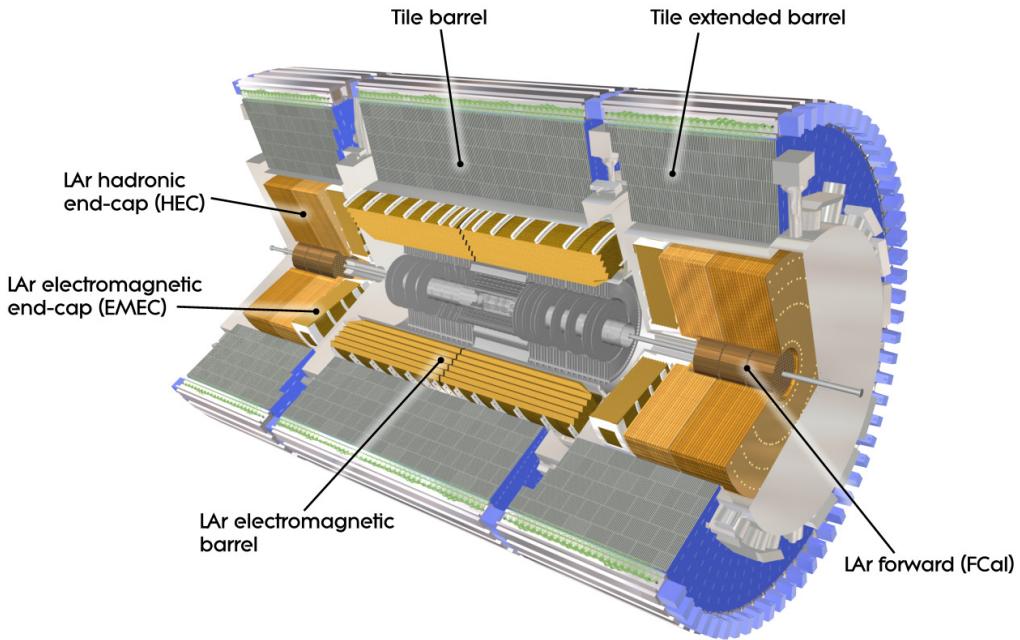


Figure 3.5: Diagram of the ATLAS calorimeters

634 imum ionizing particle that do not participate in the strong interaction and therefore pass through
 635 the ATLAS calorimeters leaving relatively little energy behind. ATLAS calorimeters are sampling
 636 calorimeters meaning that the active material of the detector only measures a small fraction of the
 637 energy produced by the shower. The overall shower energy is inferred from this fractional measure-
 638 ment. The rest of the material is inactive, dense material, designed to induce showers. The calorimetry
 639 system is grossly divided longitudinally (radially) into electro-magnetic(EM) and then hadronic seg-
 640 ments, operated with different technologies. Figure 3.5 diagrams the layout of the calorimeter system.

641 The EM calorimeter (LAr), which is located directly outside of the solenoid magnet but within
 642 the same cryostat, has a accordion design with lead absorber and liquid argon active material. The
 643 accordion design ensures uniform coverage in ϕ . The barrel and endcap LAr extend to $|\eta| < 2.47$. The
 644 LAr provides highly granular measurements in $\eta - \phi$ with 4 longitudinal segments, totaling $\sim 25\text{-}35$
 645 radiation lengths with the exception of the barrel/endcap transition region ($1.37 < |\eta| < 1.52$). The
 646 geometry of the barrel LAr calorimeter can be seen in Figure ???. The first longitudinal segment
 647 is called the pre-sampler, composed of a thin layer of active liquid argon, designed to detect early
 648 particle showers. The second segment is the most highly granular segment called the 'strips', as it
 649 is composed of thin liquid argon cells. The strips have a size of $0.025/8 \times 0.1$ in $\eta - \phi$ in the barrel

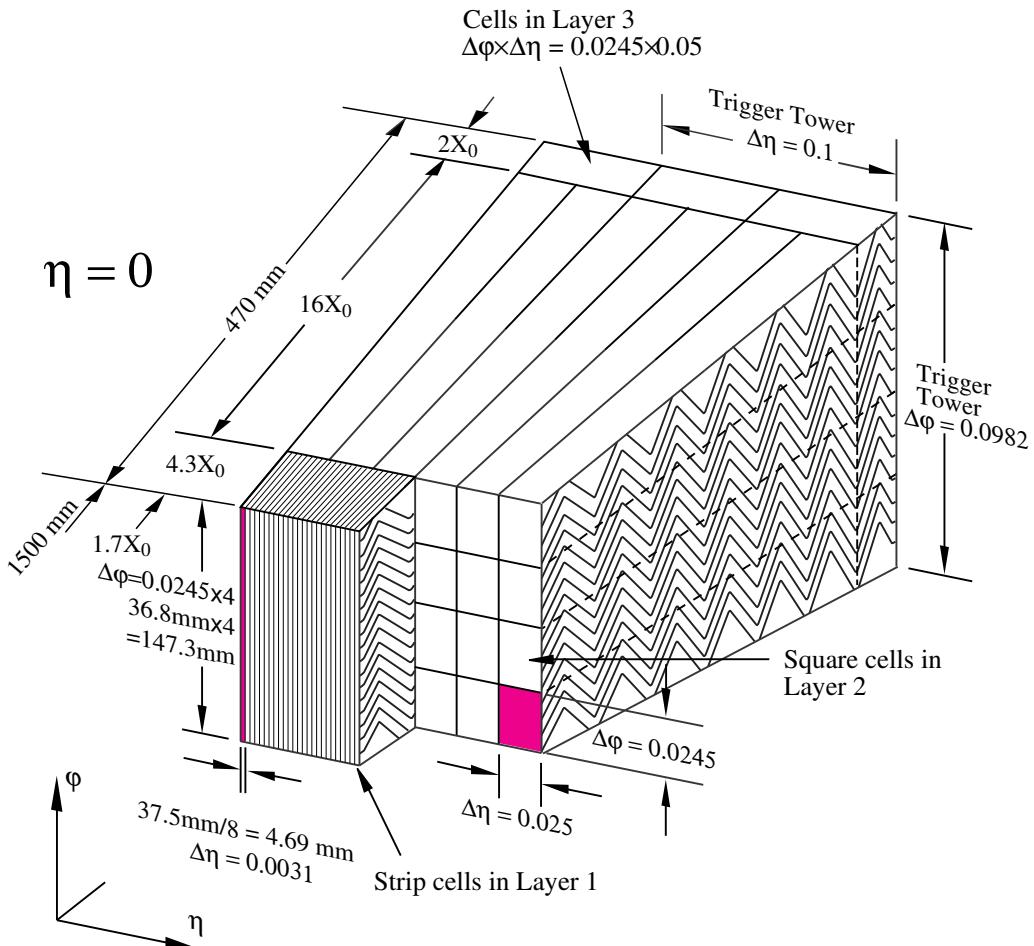


Figure 3.6: Diagram of the ATLAS LAr EM calorimeter showing the longitudinal segmentation and the $\eta - \phi$ cells for the central barrel region

with similar sizes in the endcap and are designed to be able to resolve single and double particle showers. This resolution is particularly useful in distinguishing $\pi^0 \rightarrow \gamma\gamma$ signatures from electron and photon signatures. The bulk of the radiation lengths and therefore the primary energy measurement come from the the third layer⁵. Each cell in this layer is 0.025×0.025 in $\eta - \phi$. The final layer is coarser in thinner and designed to estimate energy leaking out of the EM calorimeter. The forward EM calorimeters extend the η range and use the same technology, but are not used in this analysis. The energy resolution of the EM calorimeters is $\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$, measured in test beam data and confirmed in collision data.

The hadronic calorimeter is located directly behind the EM calorimeter and composed of tiles

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

of iron absorber and plastic scintillator in the barrel ($|\eta| < 1.6$), called the TileCal, and copper-liquid argon in the endcap ($1.5 < |\eta| < 3.2$), called the HEC. The calorimeters contain $\sim 10\text{-}19$ hadronic interactions lengths with multiple longitudinal segments to contain showers induced by the nuclear interaction of hadronic particles. The energy resolution of the hadronic calorimeters is $\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$. The intrinsic resolution of hadron calorimeters is much worse than electromagnetic calorimeters, because much of the energy is lost to the inelasticity of the nuclear interactions.

3.2.4 The Muon Spectrometer

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

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

The chambers in the barrel and most of the endcap are constructed from Monitored Drift Tubes (MDTs) with an Argon gas mixture. Each chamber contains 3-7 drift tubes and provide hit resolutions of $80\text{ }\mu\text{m}$ per tube and $35\text{ }\mu\text{m}$ per chamber in the bending plane. For $|\eta| > 2.0$, Cathode Strip Chambers (CSCs) are used, primarily to handle the higher incident particle flux. They are composed of cathode strips crossed with anode wires in the gas mixture, but use similar drift technology as the MDTs and have resolutions in the bending plane $40\text{ }\mu\text{m}$ per chamber.

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

3.2.5 The Trigger System

The ATLAS trigger system is designed to make quick decisions about individual particle collisions to reduce the enormous collision rate of 20 MHz to a much more manageable 400 Hz to be stored for offline analysis. Saving the full ATLAS data-stream would require space for 40 TB of raw data per second, but, more importantly, most of these collisions result in the uninteresting inelastic break-up of the colliding protons. To select out collisions to allow for a diverse physics program, ATLAS devotes a large portion of the bandwidth to general purpose single lepton triggers (~ 250 Hz). The presence of

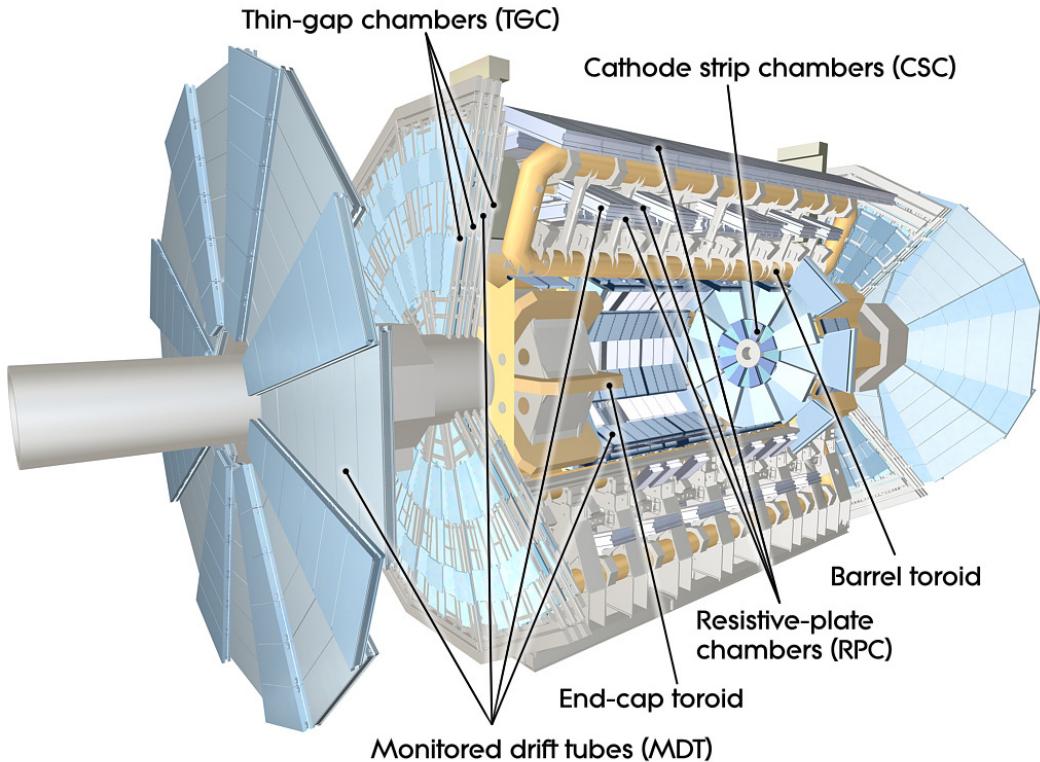


Figure 3.7: Diagram of the ATLAS muon system

leptons in the event indicates the presence of the weak or electro-magnetic interaction and therefore occurs at many order of magnitude less frequently than interactions involving the strong interaction. Moreover, many interesting physics signatures that are analyzable by ATLAS involve leptonic final states. The remaining bandwidth is allocated to jet, missing energy, tau, and unbiased supporting triggers.

The ATLAS trigger system is composed of 3 levels: level-1 (L1), level (L2), and the event filter (EF). The first level is hardware only trigger that reduces the input 20 MHz rate to ~ 75 kHz, selecting 1 out of every 250 collisions. The available buffering on the FPGA chips means that the decisions need to be made within $2.5 \mu\text{s}$. The L1 selection is based on calorimeter clustering and tracking finding in the MS for small areas of the detector called regions-of-interest (ROIs). It selects ROIs with significant energy.

The second and third stages L2 and EF are software based. The L2 algorithms perform more detailed object reconstruction for leptons, jets and photons inside of the ROIs provided by L1, by performing tracking and in depth calorimeter clustering algorithms. The decisions are made within 50

704 ms per event and pass 1 out of every 15 events to the EF. At the EF, events undergo full reconstruction
705 using similar but faster versions of the algorithms used offline. The EF makes decisions on the presence
706 of fully id-objects in the event and event topological quantities within 4s to reduce the L2 output by
707 a factor of 10. The events that pass this stage are then written to tape for offline study.

708 **3.2.6 Reconstruction: Jets, Muons and Electrons**

709 Physicists analyze the collision event as a collection of identified objects, expressed as momentum
710 4-vectors. The process of converting the disparate detector signatures and signals into a unified 4-
711 momentum description of individual objects is called reconstruction. These objects arise from the
712 final state particles in the event, which can be combined and counted to infer properties of the hard
713 scatter. The particles that make detectable signatures are those that are stable enough to pass through
714 the detector: muons, electrons, photons, and quarks and gluons. Jets and b-tagged jets, muons, and
715 electron, are used in the $t\bar{t}H$ analysis to define our search regions and to separate the Higgs signal
716 from backgrounds. Other analyses may used photons, taus and missing energy⁶, but these are not
717 discussed in depth here. Figure 3.8 shows an $R - \phi$ schematic of the interaction of various particle
718 signatures in the ATLAS detector.

719 **3.2.6.1 Tracks and Clusters**

720 The basic components of reconstruction are sensor measurements, or hits, in trackers (ID, MS) and
721 energy measurements in the calorimeter. Hits in the ID and MS undergo pattern recognition, which
722 identifies hits that belong to a single track, and fitting, which fits a curve to the track to assess
723 the particle trajectory. Charged particle trajectories are generally helical in a magnetic field, but
724 the fitting algorithm takes into more detailed information about energy loss to material along the
725 tracks length. The result of the fitting is an estimation of particle momentum 3-vector. Electron,
726 photon, and hadronic particles leave clustered deposits of energy in the EM and hadronic calorimeters
727 from their showers. Electron and photon showers are primarily contained with the EM calorimeter,
728 while hadronic showers are deposit most of their energy in the hadronic calorimeters. The process
729 of associating individual read-out cells of energy in the calorimeter to clusters of energy from the
730 showers of individual particles is called clustering. From the basic pieces of tracks and clusters, more
731 complex objects can be created.

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

732 **3.2.6.2 Electrons**

733 Electrons leave both a track in the ID and a narrow and isolated cluster of energy in the EM calorime-
734 ter, $\Delta R < 0.1$. Electron reconstruction proceeds using a sliding window algorithm, which scans a
735 fixed size rectangle in $\eta - \phi$ space over the EM calorimeter cells to find relative maxima of energy in
736 the window[51]. These maxima seed the clustering algorithms. Because electrons are light, they both
737 lose energy to the material gradually through scattering and more catastrophically through the emis-
738 sion of a high energy photon, through interaction with nuclei. This process is called bremsstrahlung.
739 Tracks for electrons are reconstructed differently because they must include the hypothesis that the
740 electron loses significant energy through bremsstrahlung. Generally, the emitted photon is contained
741 within the same energy cluster and therefore the sliding window algorithm is always wider in the di-
742 rection of bending, ϕ . A single track is then matched to the cluster within certain minimum matching
743 requirements in η, ϕ , and p_T . Electrons are distinguished from photon conversions, which also have a
744 track, by lack of association with conversion vertices, found with a dedicated algorithm.

745 Electron have many lever arms for further identification to suppress backgrounds from fake sources.
746 The narrowness of the shower shape, quality of track, and presence of transition radiation are used
747 by cut-based and multivariate algorithms are used by identification algorithms. This is discussed in
748 depth in Chapter 4. Electrons are reliably reconstructed and identified with energies above 7 GeV.

749 **3.2.6.3 Muons**

750 Muons are reconstructed from a combination of ID and MS tracks, when possible. The two tracks
751 must meet matching criteria to ensure they are from the same particle. The muon momentum 3-vector
752 comes from the combined ID/MS fit. Muons leave little energy in the calorimeters and are generally
753 isolation from other particle, when produced from electro-weak bosons. Identification algorithms
754 make requirements on the number of tracking hits in the ID and MS and the quality of the matching
755 of the two tracks. Muons are reliably reconstructed and identified with energies above 5 GeV. More
756 about muon reconstruction and identification can be found here [52].

757 **3.2.6.4 Jets**

758 Quarks and gluons are colored objects that cannot exist alone. When emitted, they undergo a process
759 called hadronization, in which the convert into 'jets' of colorless hadrons that emerge collimated from
760 the interaction point. The majority of these hadrons are charged and neutral pions, though other
761 hadrons are often present. Jets are reconstructed using conglomerations of calorimeter energy clusters
762 chosen via an anti- k_t algorithm, with a radius of $\Delta R < 0.4$ [53]. The algorithm has been shown to

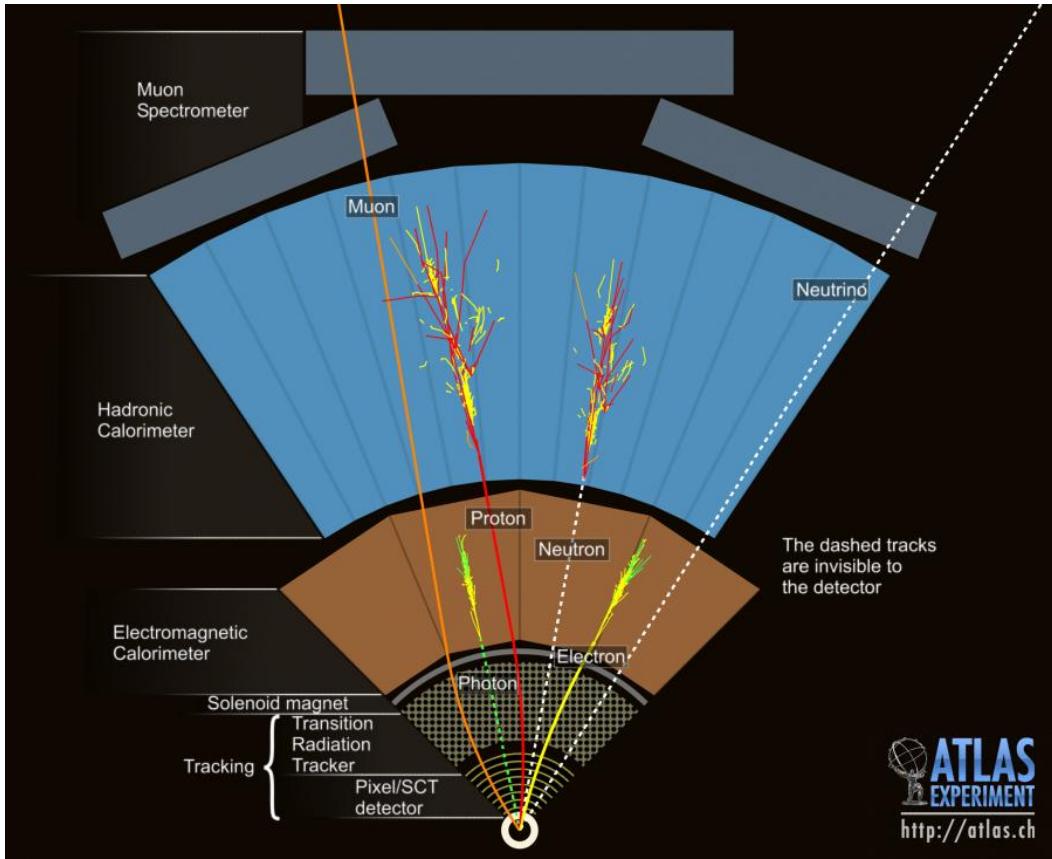


Figure 3.8: $R - \phi$ schematic of the ATLAS detector and various particle signatures

763 be infrared safe, meaning the jet quantities are not sensitive to low energy, small angle radiative
 764 divergences. Jets at ATLAS are reconstructed from 10 GeV, calibration of the energy scale and
 765 resolution are only available for energies greater than 20-25 GeV.

766 3.2.6.5 B-Tagged Jets

767 Generally, the flavor of the initiating quark is not known from the reconstructed jet, although gluon
 768 initiated jets and quark initiated jets have slightly different properties. B-quark jets, however, are
 769 unique in that the long life-time of the produced b-mesons allow for measurable decays in flight. This
 770 property is used to tag b-quark initiated jets. This analysis uses the MV1 tagging algorithm [54],
 771 which is a neural network based algorithm that looks for secondary displaced decay vertices inside
 772 the event and takes into account jet track parameters and energy flow with respect to these vertices.
 773 B-quark jets often involve b-meson decays to leptons, especially muons, which can be used to tag an
 774 orthogonal b-jet sample for studying tagging efficiencies.

775

CHAPTER 4

776

Electrons

777 This chapter details the contributions I made to electron identification and efficiency measurements.
778 It is not essential to continuity of the thesis in general but provides in depth documentation of the
779 work I completed for the experiment. I focus on the electron identification

780 **4.1 Electrons at Hadron Colliders**

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

793 At ATLAS, the primary datasets for most analyses are collected with electron and muon triggers.
794 Electron triggers are particularly important, because the muon triggers system has a 20% smaller
795 acceptance than the electrons. The challenge in identifying electrons is distinguishing the production of
796 electrons from direct production of W and Z decays from electrons produced in the more copiously
797 produced b-meson decays, fake-electron signatures from rare jet fragmentations into charged and

798 neutral pions, and photon conversions in the inner detector. The identification of electrons, the
799 precise measurement of the identification efficiency, and the measurement of the rate of fake electron
800 signatures lead are often the most important and challenging pieces of an analysis. The following sections
801 discuss the identification of electrons for the primary electron trigger and offline physics analyses as
802 well as the measurement of the electron identification efficiency in 2012. Because I had a major role in
803 these projects, I will at times discuss their historical evolution and not simply focus on the particular
804 measurement relevant to the $t\bar{t}H$ analysis.

805 **4.2 Identification of Electrons at ATLAS**

806 Electron reconstruction

807 **4.2.1 Pile-up and Electron identification**

808 Plots of pile up differences in distributions

809 **4.2.2 2011 Menu and Trigger**

810 **4.2.3 2012 Menu and Trigger**

811 **4.2.4 Electron Likelihood**

812 **4.3 Measurement of Electron Efficiency at ATLAS**

813 **4.3.1 Techniques**

814 **4.3.2 Issues**

815

CHAPTER 5

816

Analysis Summary

817 This chapter provides an overview of analysis searching for SM production of the Higgs boson
818 in association with top quarks in multi-lepton final states. The analysis searches in signal regions
819 with 2 same-sign, 3 and 4 light leptons (e, μ), which are sensitive to Higgs decays to vector bosons,
820 $H \rightarrow W^\pm W^\pm$ and $H \rightarrow Z^\pm Z^\pm$. We refer to these channels as 2ℓ SS, 3ℓ , and 4ℓ through the rest of
821 this document.

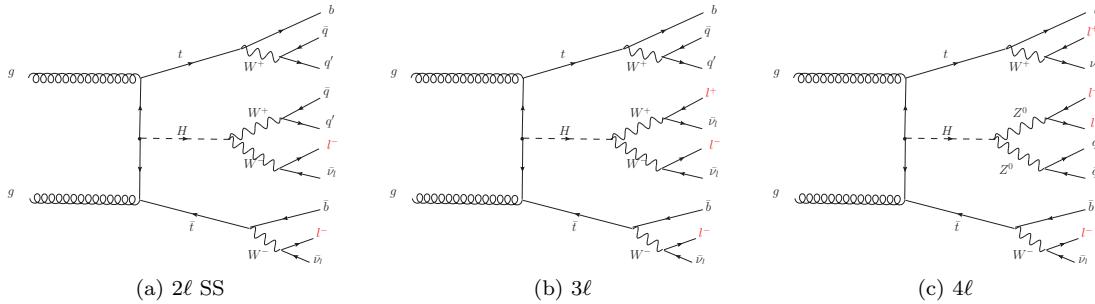
822 The multi-lepton channels form a complement to already completed $t\bar{t}H$ searches in final states
823 targeting the $H \rightarrow b\bar{b}$ [55], $H \rightarrow \gamma\gamma$ [56]. The $t\bar{t}H$ searches in the $H \rightarrow \tau\tau$ decay modes were
824 developed concurrently with the multi-lepton searches, but we do not discuss these here. Of this set
825 of complementary searches, the multi-lepton and $b\bar{b}$ are the most sensitve.

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

831 Detailed description of the event and objection section are provided in Chapter 7, background
832 modeling in Chapter 8, the effect of systematic errors and the statistical analysis in Chapter 9 and
833 final results in Chapter 10.

834 5.1 Signal Characteristics

835 The signal is expected to be characterized by the presence of 2 b-quark jets from the top quark
836 decays, isolated leptons from vector boson and tau decays, a high jet multiplicity, and missing energy
837 from neutrinos. Three Higgs boson decays are relevant for this analysis: W^+W^- , $\tau^+\tau^-$ and ZZ .

Figure 5.1: Example Feynman diagrams for the 3 $t\bar{t}H$ multi-lepton categories.

838 All modes are generally dominated by the WW signature, though the 3ℓ and 4ℓ channels possess
 839 some contribution from the $\tau\tau$ and ZZ decays. Table 5.1 provides the fractional contribution of the
 840 main Higgs decay modes at the generator level to $t\bar{t}H$ search channels and Figure 5.1 shows example
 841 diagrams for each channel. In general, the number of leptons is anti-correlated with the number of
 842 jets, since a vector boson can either decay leptonically or hadronically, such that:

- 843 • in the 2ℓ SS channel, the $t\bar{t}H$ final state contains 6 quarks⁷. These events are then characterized
 844 by the largest jet multiplicity.
- 845 • In the 3ℓ , the $t\bar{t}H$ final state contains 4 quarks
- 846 • In the 4ℓ channel, the $t\bar{t}H$ final state contains a small number of light quarks, 0 ($H \rightarrow$
 847 W^+W^- case), 2 or 4 ($H \rightarrow ZZ$ case).

Table 5.1: Contributions of the main Higgs decay modes to the 3 multi-lepton $t\bar{t}H$ 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%

848 5.2 Background Overview

849 For all channels after selection, the size of the signal is of similar order to the expected size of
 850 background. Background processes can be sorted into two categories:

⁷this does not include additional quarks from radiation

- 851 • **Reducible:** These processes cannot lead to a final state compatible with the signal signature
852 without a mis-reconstructed object. This category includes events with a prompt lepton but
853 with mis-reconstructed charge and events with jets that "fake" leptons. The main backgrounds
854 of this sort are $t\bar{t}$ and $Z+jets$. Data-driven techniques are used to measure the rate of these
855 processes and strict object selection and used to reduce their rate.
- 856 • **Irreducible:** Events which can lead to the same final state as the signal. The main background
857 of this category are: $t\bar{t}V$, $W^\pm Z$, and ZZ . They are modeled using the Monte Carlo simulations.
858 In general, these backgrounds are combatted with jet and b-tagged jet requirements. Although
859 the jet multiplicity of $t\bar{t}V$ is high, the multiplicity of $t\bar{t}H$ events is still higher.

860

5.3 Analysis Strategy

861 The analysis search is conducted in 3 channels, based on counting of fully identified leptons: 2ℓ SS,
862 3ℓ , and 4ℓ , with cuts optimized separately for each. We further divide the 2ℓ SS into sub channels
863 based on the number of jets and flavor of the leptons and the 4ℓ channel into sub-channels enriched
864 and depleted in OS leptons arising from Z decays.

865 This analysis is a counting experiment, meaning that the only quantities significant to measured
866 result are the event counts in the signal regions and not the event shapes. The measured background
867 rates, expected signal rates and systematic uncertainties are fed into a Poisson model and fit to the
868 observed data. The parameter of interest in the fit and the result of this measurement is, μ , the ratio
869 of the fitted number of $t\bar{t}H$ events in the signal regions to expected number of $t\bar{t}H$ events in the
870 signal regions. Since we assume SM branching ratios, μ can be considered the ratio of the measured
871 $t\bar{t}H$ cross-section to the observed $t\bar{t}H$ cross-section, and we the fitted μ to be close to 1 with large
872 statistical errors.

873 We express the final result as a measurement of μ with uncertainties and 95% upper limit on the
874 value of μ : μ -values higher than this value will be considered excluded. We provide these results for
875 each channel individually and combined.

CHAPTER 6

Dataset and Simulation

878 6.1 Data

879 6.1.1 The 2012 Dataset

880 The $t\bar{t}H$ analysis uses the entire 2012 ATLAS dataset only, collected from April to December. The
 881 size of the dataset corresponds to 20.3 fb^{-1} , after passing data quality requirements, ensuring the
 882 proper operation of the tracking, calorimeter and muon subsystems. The LHC successfully produced
 883 datasets for physics studies in 2010, 2011 and 2012. The 2012 proton-proton dataset was delivered
 884 with collisions with a CME of 8 TeV with bunch collisions every 50 ns[57].

885 Figure 6.1 shows the accumulation of the 2012 dataset over time. Despite doubling the bunch
 886 spacing above the design of 25 ns, the luminosity neared the design luminosity due to unexpected
 887 improvements in the transverse beam profile[58]. This increased the amount of pile-up, or number
 888 of collisions per bunch crossing and in general collision events were busier due to these multiple
 889 interactions. Figure 6.2 shows the average number of interaction per bunch crossing for the 2011 and
 890 2012 datasets. The 2012 dataset shows an average of 20-25 interactions.

891 The dataset must contain either a primary muon or primary electron trigger (`EF_e24vhi_medium1`
 892 OR `EF_e60_medium1` OR `EF_24i_tight` OR `EF_36_tight`). The electron triggers require a electron with
 893 at least 25 GeV of calorimeter energy, passing the medium identification requirement and loose tracking
 894 isolation. Above 60GeV, the isolation requirement is dropped and the identification is loosened slightly.
 895 The muon trigger requires a good inner detector track and matching hits in the muon spectrometer,
 896 as well as loose tracking isolation, which also is dropped about 36 GeV.

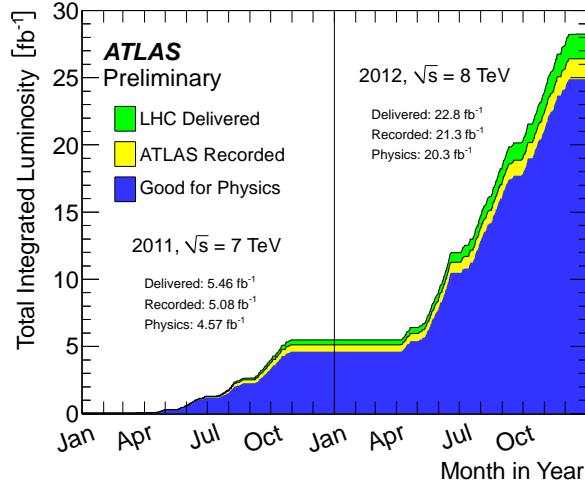


Figure 6.1: Plot showing the accumulation of the integrated luminosity delivered to the ATLAS experiment over 2011 and 2012. The rough size of the usable, physics ready dataset for 2012 is 20 fb^{-1} and is the dataset used for the following analysis.

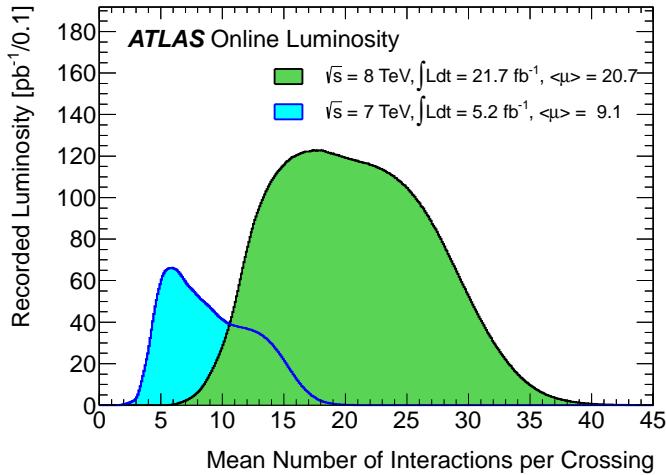


Figure 6.2: The average number of interactions per bunch-crossing for the 2012 and 2011 LHC proton-proton dataset. Most of these interactions are uninteresting but leave energetic signatures in particle detectors called pile-up which interfere with measurements

897 6.2 Simulation

898 Simulation samples are used to determine the overall event selection acceptance and efficiency and
 899 model the number of events in the signal regions for prompt backgrounds and signal. The simulated
 900 samples are created using parton distribution function (PDF) and use Monte Carlo (MC) techniques
 901 to model the hard parton scatter, underlying event activity and parton showering and hadronization.

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

902 The samples are then passed through a full ATLAS detector simulation[59] based on GEANT4 [60].
 903 Small corrections are then applied to re-scale object identification efficiencies, energy scales, and the
 904 pile-up based on control regions from data. These corrections are discussed in Chapter 9.

905 6.2.1 Signal Simulation

906 The $t\bar{t}H$ production is modeled using matrix elements obtained from the HELAC-Oneloop package [61]
 907 that corresponds to the next-to-leading order (NLO) QCD accuracy. POWHEG BOX [62, 63, 64]
 908 serves as an interface to the parton shower Monte Carlo programs. The samples created using this
 909 approach are referred to as PowHel samples. CT10NLO PDF sets are used and the factorization (μ_F)
 910 and renormalization (μ_R) scales are set to $\mu_0 = \mu_F = \mu_R = m_t + m_H/2$. Pile-up and the underlying
 911 events are simulated by Pythia 8.1 [65] with the CTEQ61L set of parton distribution functions and
 912 AU2 underlying event tune. The Higgs boson mass is set to 125 GeV and the top quark mass is set
 913 to 172.5 GeV.

914 The signal Monte Carlo samples are summarized in Table 6.1. These large samples are generated with
 915 inclusive Higgs boson decays with branching fractions set to the LHC Higgs Cross Section Working
 916 Group (Yellow Report) recommendation for $m_H = 125$ GeV [66]. The inclusive cross section (129.3
 917 fb at $m_H = 125$ GeV) is also obtained from the Yellow Report [66].

918 6.2.2 Background Simulation

919 The background simulations used for this analysis are listed in Table 6.2. In general, the Alpgen[67],
 920 MadGraph[68], and AcerMC[69] samples use the CTEQ6L1[70] parton distribution function, while
 921 the Powheg[71], Sherpa[72], are generated with the CT10 PDF. The exception is the MadGraph $t\bar{t}t\bar{t}$
 922 sample, which is generated with the MSTW2008 PDF[73]. The highest order calculations available
 923 are used for the 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. $t\bar{t}$, single top, and $Z+jets$ samples are replaced with data-driven estimates for the final result

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
$W\gamma+4p$	Alpgen	Full
W^+W^-	Sherpa	Full
$W^\pm Z$	Sherpa	Full
Same-sign WW	Madgraph+Pythia8	AF2
ZZ	Powheg+Pythia8,gg2ZZ+Herwig	Full
$Z\gamma^*$	Sherpa	Full
$Z+jets$	Sherpa	Full
ggF Higgs	Powheg+Pythia8	Full

924

CHAPTER 7

925

Object and Event Selection

926 OVERLAP REMOVAL The analysis is divided into 3 signal regions based on lepton counting: 2
 927 same-sign leptons, 3 leptons and 4 leptons. The lepton counting occurs for fully identified leptons
 928 with full overlap removal with transverse momenta over 10 GeV to ensure orthogonality. Lepton
 929 selections are tightened afterward within each region.

930 The cuts for each signal region are provided in Table 7.1 and the object selections are detailed in the
 931 following selections. The selections are based on optimizations of the region sensitivity performed using
 932 MC (event for data driven backgrounds) and ad-hoc values for normalization systematic uncertainties⁸
 933 All signal regions are comprised of three basic requirements: the presence of b-tagged jets, the presence
 934 of additional light jets, and a veto of same flavor opposite sign leptons with an invariant mass within
 935 the Z window. Additional requirements on the invariant mass of the leptons, the missing transverse
 936 energy in the event, and the total object energy (H_T) proved to have negligible additional benefit at
 937 our level of statistics.

938 **7.1 2ℓ Same-Charge Signal Region**

939 The 2 lepton same-sign signal region (2ℓ SS) requires two leptons of similar charge. The signal is
 940 symmetric in charge but the background from opposite-sign $t\bar{t}$ di-lepton production would be over-
 941 whelming. Requiring only two leptons allows the extra 2 W bosons in the event to decay hadronically,
 942 resulting in on average 4 additional light jets plus 2 additional b-quark jets from the top decays.

943 We require a leading lepton with transverse momentum of at least 25 GeV that matches to a
 944 trigger and a sub-leading lepton of at least 20 GeV, a b-tagged jet, and at least 4 jets in total.

⁸the sensitivity was approximated using the $\frac{s}{\sqrt{b+\Delta b}}$ formula. The systematic errors considered were 20% for $t\bar{t}V$ and VV and 30% for fakes. These ended up being close the final systematic errors assessed in Chapter 9. The objects of optimization were the lepton momenta, identification operating points, isolation and event kinematic variables

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^9	=2	=3	=4
Lepton Charge Sum	+2 or -2	+1 or -1	0
Lepton Momentum (GeV) ¹⁰	$p_{T0} > 25$ $p_{T1} > 20$	$p_{T0} > 10$ $p_{T1,2} > 25, 20$	$p_{T0} > 25$ $p_{T1} > 20$ $p_{T2,3} > 10$
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)	$ M_{ee} - M_Z < 10$	$ M_{SFOS} - M_Z < 10$	$M_{SFOS} > 10$ $150 < M_{4\ell} < 500$ $ M_{SFOS} - M_Z < 10$
Sub-channels	$2 (N_{Jet} = 4, N_{Jet} \geq 5) \times 3 (ee, e\mu, \mu\mu)$	none	2 (No SFOS leps, SFOS leps)

945 In order to suppress non-prompt backgrounds, the lepton isolation criteria for tracking and
 946 calorimeter are tightened from less than 10% of the lepton momentum to 5%. To suppress charge
 947 mis-identification, the electron is required to be extremely central ($|\eta| < 1.37$) to avoid the material-
 948 rich regions of the detector. Additionally, ee events with a lepton pair invariant mass within 10 GeV
 949 of the Z pole are removed. To maintain orthogonality with the τ analyses, events with fully identified
 950 taus are vetoed.

951 For the statistical combination the channel is divided into 6 sub-channels: 2 jets counting bins
 952 ($N_{Jet} = 4, N_{Jet} \geq 5$) \times 3 lepton flavor bins ($ee, \mu\mu, e\mu$).

953 7.2 3ℓ Signal Region

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

- 956 • **lep0:** the lepton that is opposite in charge to the other two leptons
- 957 • **lep1:** the lepton that is closer in ΔR to lep0
- 958 • **lep2:** the lepton that is farther in ΔR from lep1

959 Since events with a fake lepton arise exclusively from opposite sign di-lepton processes, $t\bar{t}$ and
 960 Z+jets, where additional jets are mis-identified as the third lepton, lep0 is never the fake lepton. As
 961 a result, the transverse momentum requirement of lep0 (> 10 GeV) is lower than the other two, > 20
 962 GeV. One lepton must match a trigger and have $p_T > 25$ GeV.

963 The 3ℓ channel further requires at least one b-tagged jets and at least 4 jets in total, or two
 964 b-tagged jets and exactly 3 jets in total. Additionally, to suppress $W^\pm Z$ and $Z+jets$ events, events
 965 with same-flavor opposite sign pairs within 10 GeV of the Z pole are vetoed.

966 Additional cuts, including a di-lepton mass cut, and splittings were investigated but low statistics
 967 proved to wash out any advantages. The di-lepton mass cut will be a useful discriminant in future
 968 analyses since the spin statistics of Higgs decay in W bosons often causes the two emitted opposite-sign
 969 leptons to point in the same direction, resulting in a small measured invariant mass.

970 7.3 4ℓ Signal Region

971 In the four lepton signal region, selected events must have exactly four leptons with a total charge
 972 of zero. At least one lepton must be matched to one of the applied single lepton trigger and have
 973 a transverse momentum above 25 GeV. The leading and sub-leading leptons are required to have
 974 transverse momentum of 25 and 15 GeV respectively. In order to suppress background contributions
 975 from low-mass resonances and Drell-Yan radiation, all opposite-sign-same-flavor (OS-SF) lepton pairs
 976 are required to have a dilepton invariant mass of at least 10 GeV.

977 The four-lepton invariant mass is required to be between 100 and 500 GeV. This choice of mass
 978 window suppresses background from the on-shell $Z \rightarrow 4\ell$ peak and exploits the high-mass differences
 979 between the signal and the dominant $t\bar{t}Z$ background. Events containing an OS-SF lepton pair
 980 within 10 GeV of the Z boson mass are discarded. This Z -veto procedure greatly reduces background
 981 contributions from ZZ and $t\bar{t}Z$. Finally, selected events are required to have at least two jets, at least
 982 one of which must be tagged as a b-quark jet.

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

987 7.4 Electron Selection

988 The electrons are reconstructed by a standard algorithm of the experiment [51] and the electron
 989 cluster is required to be fiducial to the barrel or endcap calorimeters: $|\eta_{cluster}| < 2.47$. Electrons in
 990 the transition region, $1.37 < |\eta_{cluster}| < 1.52$, are vetoed. Electrons must have $p_T > 10$ GeV and pass
 991 the the **VERYTIGHT** likelihood identification criteria.

992 In order to reject jets misidentified as electrons, electron candidates must also be well isolated
 993 from additional tracks and calorimeter energy around the electron cluster. Both the tracking and
 994 calorimeter energy within $\Delta R = 0.2$ of the electron cluster must be less than 5% of the electron trans-
 995 verse momentum: $\text{ptcone20}/P_T < 0.05$ and $\text{Etcone20}/E_T < 0.05$. All quality tracks with momentum
 996 greater than 400 MeV contribute to the isolation energy. Calorimeter isolation energy is calculated
 997 using topological clusters with corrections for energy leaked from the electron cluster. Pile-up and
 998 underlying event corrections are applied using a median ambient energy density correction.

999 The electron track must also match the primary vertex. The longitudinal projection of the track
 1000 along the beam line, $z_0 \sin \theta$, must be less than 1 cm) and the transverse projection divided by the
 1001 parameter error, $d0$ significance, must be less than 4. These cuts are used in particular to suppress
 1002 backgrounds from conversions, heavy-flavor jets and electron charge-misidentifications.

1003 The electron selection is summarized in Table 7.2.

1004 7.5 Muon Selection

1005 Muons used in the analysis are formed by matching reconstructed inner detector tracks with either
 1006 a complete track or a track-segment reconstructed in the muon spectrometer (MS), called Chain 2
 1007 muons. The muons have $p_T > 10$ GeV and satisfy $|\eta| < 2.5$. The muon track are required to be a good
 1008 quality combined fit of inner detector hits and muon spectrometer segments, unless the muon is not
 1009 fiducial to the inner detector, $|\eta| > 2.47$. Muons with inner detector tracks are further required to
 1010 pass standard inner detector track hit requirements [52].

1011 As with electrons, muons are required to be isolated from additional tracking or calorimeter energy:
 1012 $\text{ptcone20}/P_T < 0.1$, $\text{Etcone20}/E_T < 0.1$. A cell-based $\text{Etcone20}/P_T$ relative isolation variable is used.
 1013 A pile-up energy subtraction based on the number of reconstructed vertices in the event is applied.
 1014 The subtraction is derived from a Z boson control sample.

1015 The muons must also originate from the primary vertex and have impact parameter requirements,
 1016 $d0$ significance < 3 , and $z_0 \sin \theta < 0.1$ cm, similar to the electrons.

1017 The muon selection is summarized in Table 7.2.

1018 7.6 Jet and b-Tagged Jet Selection

1019 Jets are reconstructed in the calorimeter using the anti- k_t [53] algorithm with a distance parameter of
 1020 0.4 using locally calibrated topologically clusters as input (LC Jets). Since the jets in the $t\bar{t}H$ signal

1021 mostly arise from the decay massive resonances and not radiation, they are expected to be central
1022 and high energy. Jets must have $p_T > 25 \text{ GeV}$ and $|\eta| < 2.5$.

1023 Jets must also pass loose quality requirement, ensuring the proper functioning of the calorimeter
1024 at the time of data taking. Jets near a hot Tile cell in data periods B1/B2 are rejected. The local
1025 hadronic calibration is used for the jet energy scale, and ambient energy corrections are applied to
1026 account for energy due to pileup.

1027 Jets within $|\eta| < 2.4$ and $p_T < 50 \text{ GeV}$ are further required to be associated with the primary
1028 vertex. The fraction of track p_T associated with the jet that comes from the primary vertex, must
1029 exceed 0.5 (or there must be no track associated to the jet). This requirement rejects jets that arise
1030 from pile-up vertices.

1031 B-jets are tagged using a Multi-Variate Analysis (MVA) method called MV1 and relying on infor-
1032 mation of the impact parameter and the reconstruction of the displaced vertex of the hadron decay
1033 inside the jet[54]. The output of the tagger is required to be above 0.8119 which corresponds to a
1034 70% efficient Working Point (WP).

1035 **7.7 Object Summary and Overlap**

1036 Since many fully identified objects maybe reconstructed as two different objects, an overlap removal
1037 procedure is applied. Electrons within $\Delta R < 0.1$ of muons are rejected in favor of the moun. Jets
1038 within $\Delta R < 0.3$ of electrons are then removed. Finally, muons within $\Delta R < 0.04 + 10\text{GeV}/p_T$ of
1039 jets are rejected, as these muons are thought to arise from jet fragmentation.

1040 **7.8 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^{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^{sig} $	$< 3\sigma$	
z_0	$< 1 \text{ cm}$	
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 multi-lepton $t\bar{t}H$ analysis. Some channels use a sub-sample of objects as explained in the Remarks column.

1041

CHAPTER 8

1042

Background Estimation

1043 The $t\bar{t}H$ multi-lepton signal regions discussed in Chapter 5 are contaminated by background contribu-
1044 tions at a similar order of magnitude to the signal. The dominant background for each region is
1045 vector boson production in association with top quarks ($t\bar{t}V$). Sub-dominant but important back-
1046 grounds include the production of vector boson pairs in association with jets and b-quark jets (VV)
1047 and $t\bar{t}$ production with a jet misidentified as a lepton (fakes). The 2ℓ SS regions possesses a unique
1048 background of charge misidentification from Z and top events. The methods for estimating these
1049 backgrounds are discussed in this chapter. Monte Carlo simulation is used for the prompt $t\bar{t}V$ and
1050 VV contributions. Systematic uncertainties on the overall normalization of these backgrounds in
1051 the signal region are provided from theoretical studies and past ATLAS analyses and are verified in
1052 data-based validation regions. The non-prompt backgrounds from $t\bar{t}$ jet-misidentification and charge-
1053 misidentification are estimated using data-driven methods.

1054 For reference, Table 8.1 provides a summary of the $t\bar{t}H$ signal and background expectation for
1055 each of the signal regions, including the data-driven estimates discussed in this section. For each
1056 region, the background contribution exceeds the size of the signal.

1057 **8.1 Vector Boson (W^\pm, Z) production in association with top quarks:**

1058 $t\bar{t}V, tZ$

1059 This section describes the estimation and $t\bar{t}V$ productions. Production of top quarks plus vector boson
1060 is an important background in all multi-lepton channels. A large part of the $t\bar{t}V$ component, arising
1061 from on-shell $Z \rightarrow \ell\ell$, can be removed via a Z mass veto on like-flavor, opposite sign leptons. However
1062 the $Z \rightarrow \tau\tau$ and γ^* components remain. The $t\bar{t}W^\pm$ and tZ processes generally require extra jets to
1063 reach the multiplicity of our signal regions. Uncertainties from the choice of the factorization (μ_F)

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 depleted		
	$e^\pm e^\pm$	$e^\pm \mu^\pm$	$\mu^\pm \mu^\pm$	$e^\pm e^\pm$	$e^\pm \mu^\pm$	$\mu^\pm \mu^\pm$	Z enriched	Z depleted	
tH	0.73 \pm 0.03	2.13 \pm 0.05	1.41 \pm 0.04	0.44 \pm 0.02	1.16 \pm 0.03	0.74 \pm 0.03	0.19 \pm 0.01	0.03 \pm 0.00	
tV	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, 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	
$s/\sqrt{b} \oplus 0.3\text{fake}(MC) \oplus 0.2t\bar{t}V$	0.33	0.58	0.47	0.12	0.22	0.21	0.63	0.207	
$s/\sqrt{b} \oplus 0.3\text{fake}(DD) \oplus 0.2t\bar{t}V$	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 symmetrized 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

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. $t\bar{t}W^\pm$ 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 re-weighting that is available with MadGraph5_aMC@NLO. The prescription for the scale envelope is taken from [74]: 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 [75]: calculate the PDF uncertainty using the MSTW2008nlo [73] 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. [76]. 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.

1085 **8.1.1 $t\bar{t}Z$ Validation Region**

1086 Unlike $t\bar{t}W^\pm$, a $t\bar{t}Z$ validation region can be obtained by simply inverting the veto on same-flavor
1087 opposite sign lepton pairs near the Z pole in the 3 lepton signal region. This region thus requires 3
1088 leptons (with momentum and identification cuts discussed in Chapter 7, at least one opposite sign,
1089 same-flavor pair of leptons within 10 GeV of the Z mass, and either 4 jets and at least 1 b-tagged jet
1090 or exactly 3 jet and 2 or more b-tagged jets. The resulting region has low statistics and is not used as
1091 a control region but is instead used as a validation to demonstrate that the normalization uncertainty,
1092 discussed above, is properly evaluated.

1093 The region defined by this is predicted to be 67% $t\bar{t}Z$, 17% WZ , and 13% tZ . We predict
1094 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
1095 errors are from data and simulation statistics, respectively). Given the large errors, the region is still
1096 in agreement with the predictions to within 1-1.5 σ . Distributions of various variables are shown in
1097 Fig. 8.1.

1098 **8.2 Di-boson Background Estimation: $W^\pm Z, ZZ$**

1099 $W^\pm Z$ and ZZ di-boson production with additional and b-tagged jets constitute small contributions
1100 to the 3- and 4-lepton channels respectively. In the 3-lepton case $W^\pm Z$ comprises ~ 1 event of \sim
1101 10 total background events while the ZZ contribution accounts for approximately 10% of the total
1102 background in the 4-lepton channel. Because of the small size of these contributions, each of the
1103 above processes can be assigned a non-aggressive uncertainty based on similar previous analyses with
1104 ATLAS and cross-checked with data validation regions and MC truth studies. We assign an overall
1105 50% error on both the $W^\pm Z$ 3-lepton signal region contribution and the ZZ 4-lepton signal region
1106 contribution. The details of this error assignment are discussed below.

1107 Both $W^\pm Z$ and ZZ production have been studied by ATLAS [77][78] but neither process has
1108 been investigated thoroughly in association with multiple jets and b-quark jets. However, both $W + b$
1109 [79] and $Z + b$ [80] production in 7 TeV data have been shown to agree with MC models to within
1110 20-30%. A single W produced in association with b-tagged jets possesses a similar topology to the
1111 $W^\pm Z + b$ process at a different energy scale and has been shown to be dominated by charm mis-tags
1112 and b-jets from gluon splitting and multiple parton interaction. The $W + b$ analysis unfortunately
1113 uses Alpgen MC with Herwig PS modeling and only provides results to 1 additional jet and therefore
1114 is not directly applicable to this $t\bar{t}H$ analysis (where $W^\pm Z$ is modeled using Sherpa with massive
1115 c and b quarks). $Z + b$ production originates from slightly different diagrams than $ZZ + b$ however

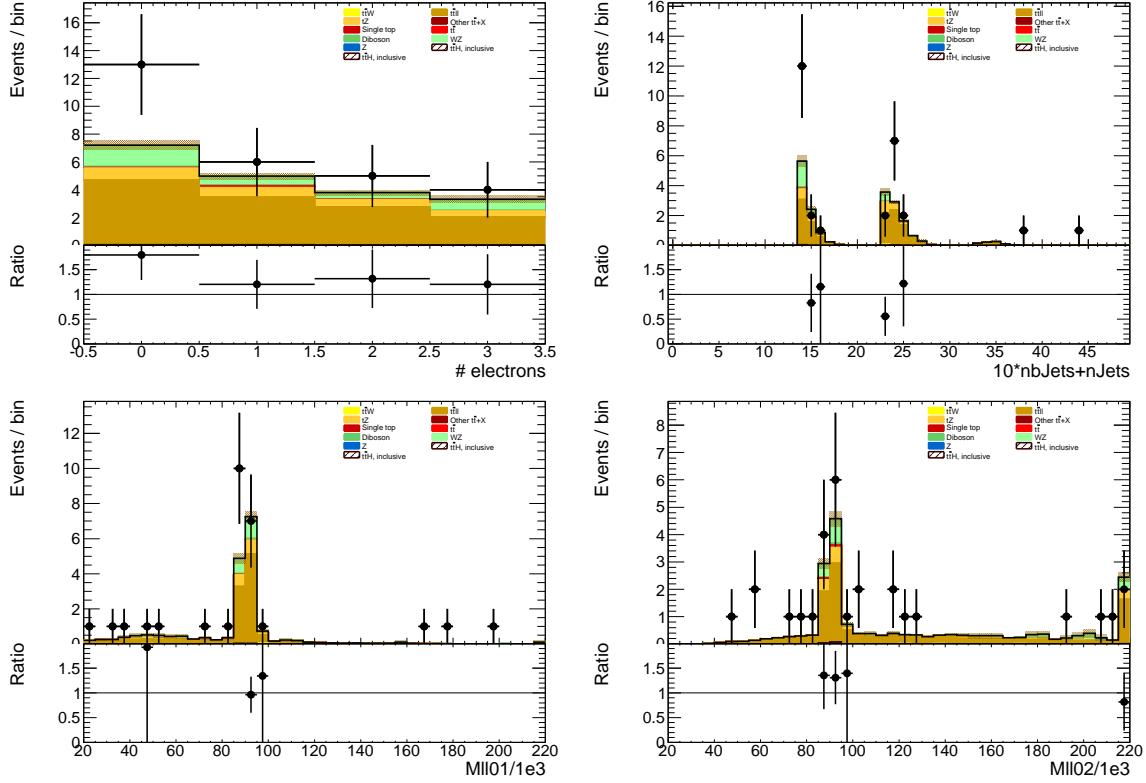


Figure 8.1: Data/MC comparison plots for $t\bar{t}Z$ control region A (≥ 4 jets, ≥ 1 b -tag and 3 jets, ≥ 2 b -tag). In all plots, the rightmost bin contains any overflows. Top left: number of electrons. Top right: 10^* the number of b -tags + the total number of jets. Middle left: the invariant mass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

1116 the sources of the b -tags are similar and the analysis above provides results with Sherpa MC with an
 1117 agreement of $\sim 30\%$.

1118 In the following two sections the uncertainty assignments for each of these two di-boson processes
 1119 will be reviewed in turn.

1120 8.2.1 $W^\pm Z$ Uncertainty

1121 The $t\bar{t}H$ analyses has two validation regions to test the Sherpa agreement with data for $W^\pm Z$: one
 1122 inclusive 3 lepton region, using the three-lepton channel object and p_T cuts; and a $W^\pm Z + b$ region
 1123 with 1 b -tagged jet, fewer than 4 jets (to remove $t\bar{t}V$), and a requirement that at least one same-flavor
 1124 opposite sign pair have an invariant mass within 10 GeV of the Z mass. Figure 8.2 shows kinematic

variables for the inclusive region ¹¹. 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.

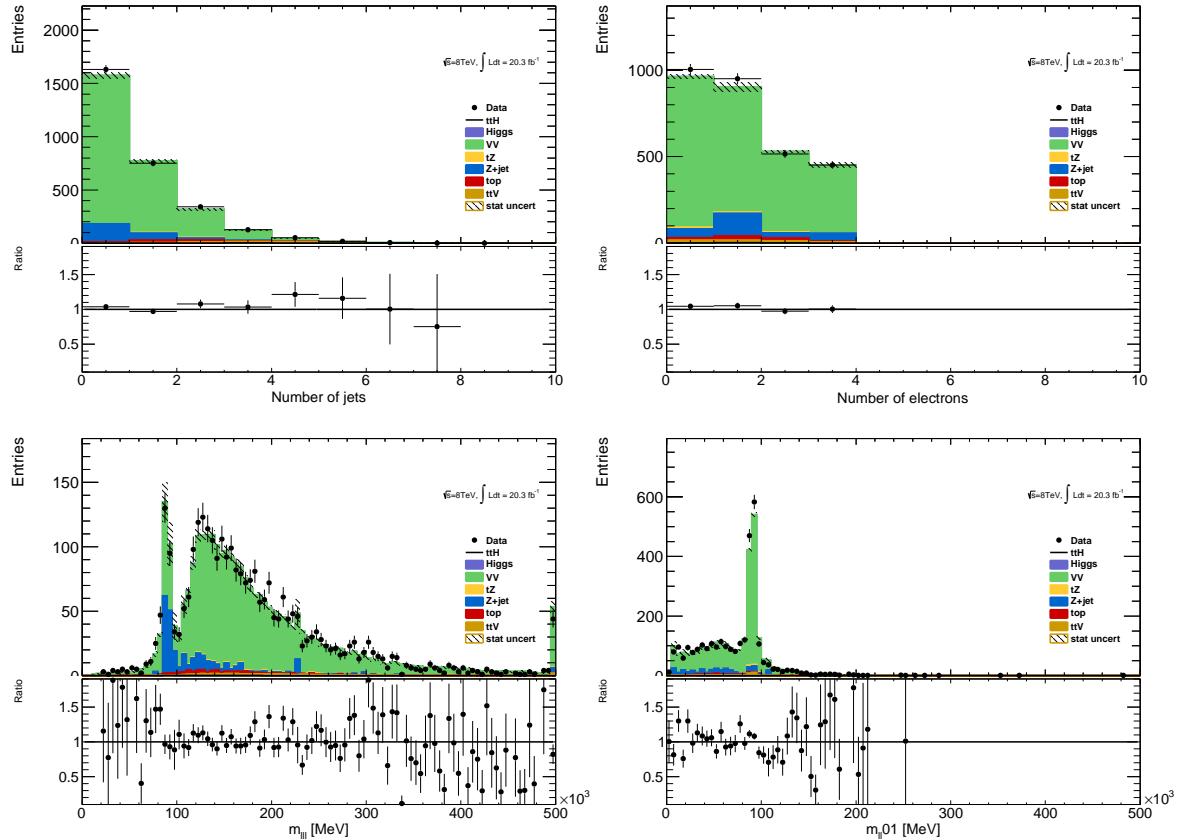


Figure 8.2: Inclusive 3 lepton $W^\pm Z$ validation region using the $t\bar{t}H$ lepton identification and momentum cuts: mass, number of jet and flavor variables

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

¹¹the fakes are taken directly from MC

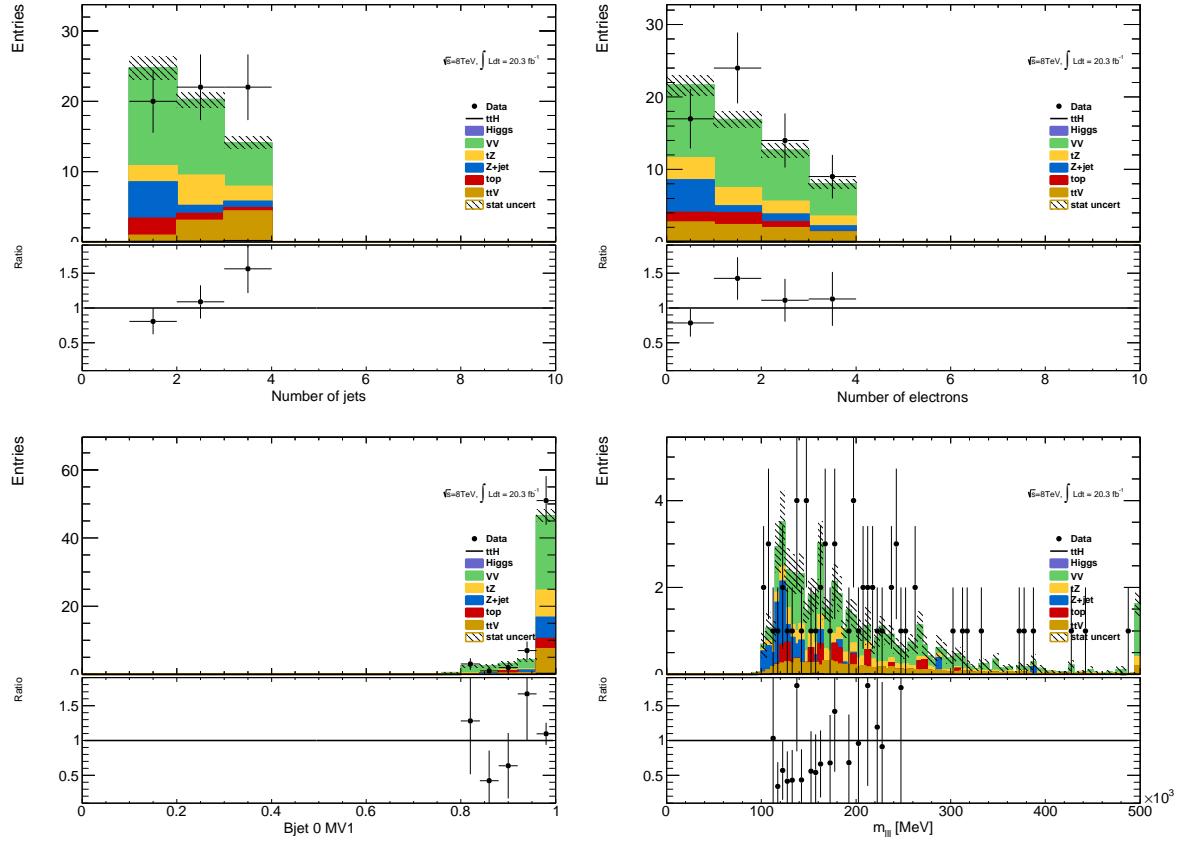


Figure 8.3: $W^\pm Z + b$ validation region: NJet, NElec, BJet MV1 and Mass Variables

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

	Bottom	Charm	Light
$W^\pm Z + b$ VR 1 Jet	0.25 ± 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

8.2.2 ZZ Uncertainty

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

di-lepton invariant mass within 10 GeV 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 modeling in this region which lies close to the 4-lepton signal region.

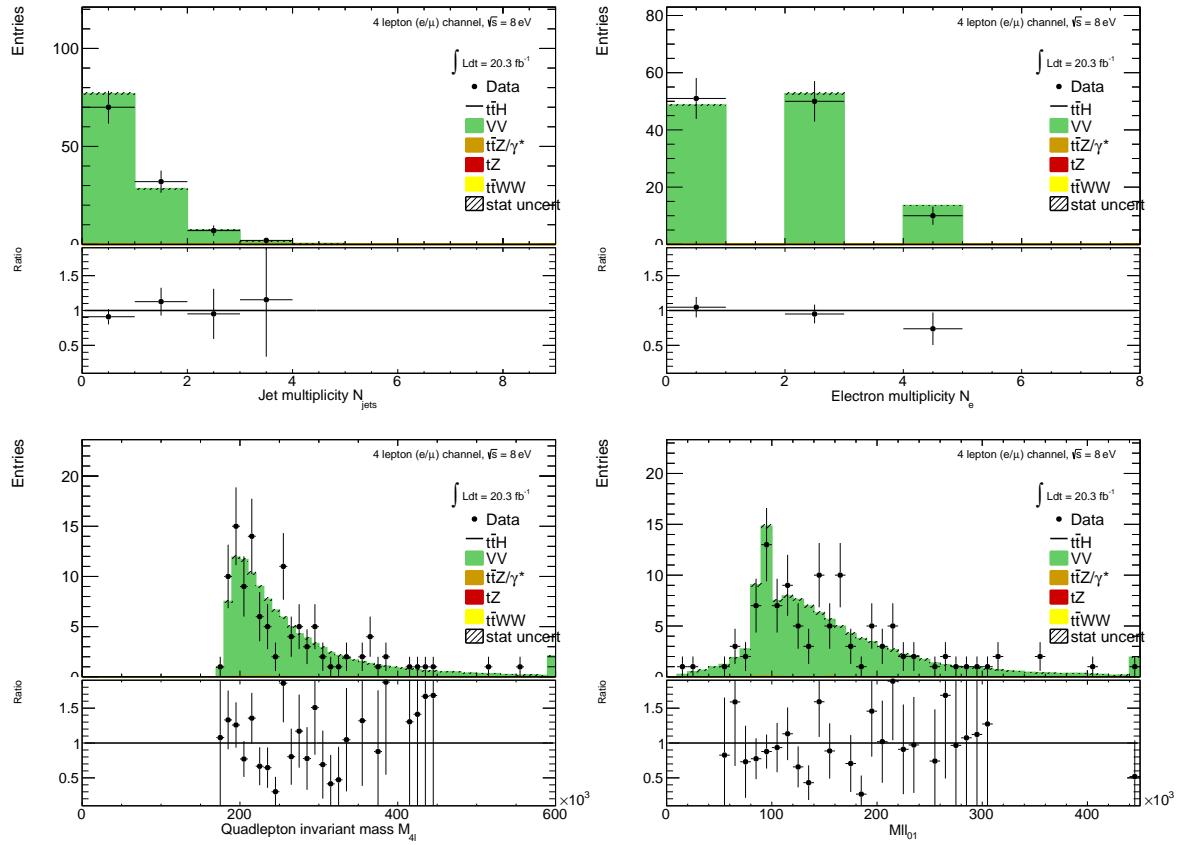


Figure 8.4: Jet-inclusive 4-lepton ZZ validation region using the $t\bar{t}H$ lepton identification and momentum cuts

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

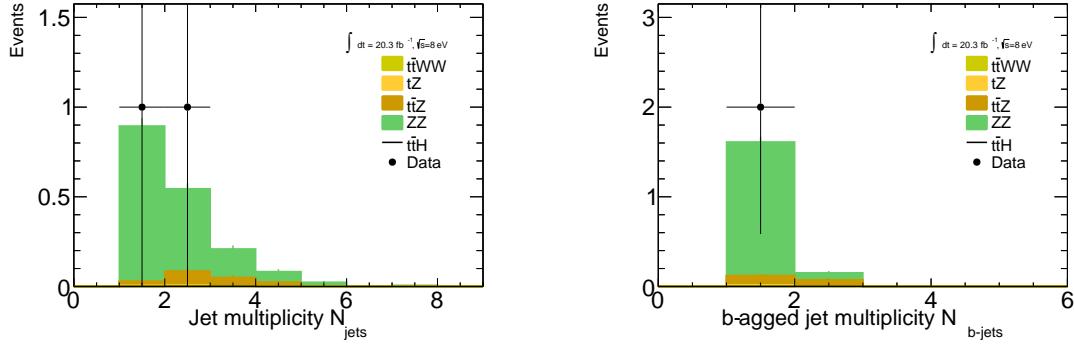


Figure 8.5: $ZZ + b$ validation region using the $t\bar{t}H$ lepton identification and momentum cuts

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 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. The second source of charge misidentification is from 'tridents', which only occurs for electrons, because their low mass allows for high rate bremsstrahlung in the detector material. In some cases, after an electron

releases a photon through bremsstrahlung, 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.

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

$$\bullet N^{ss} = \frac{\epsilon}{1 - \epsilon} N^{os} \text{ for the } e\mu \text{ channel,}$$

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

1199 rates of the electron charge are independent for different pseudo-rapidity regions. Therefore, the
 1200 probability to have a number of same-sign events (N_{ss}^{ij}) with electrons in $|\eta|$ region i and j can be
 1201 written as a function of the number of events N^{ij} as follows:

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

1202 If all the same-sign events in the Z peak are produced by charge mis-identification, then N_{ss}^{ij} is
 1203 described by a Poisson distribution:

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

1204 where k is the observed number of occurrences of the event, i.e. $k = N_{ss}^{ij}$, and λ is the expected number,
 1205 i.e. $\lambda = N^{ij}(\epsilon_i + \epsilon_j)$. Thus, the probability for both electrons to produce a charge mis-identification
 1206 is expressed by:

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

1207 The likelihood L for all the events is obtained by evaluating all the $|\eta|$ combinations:

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

1208 where the rates ϵ_i and ϵ_j can be obtained by minimizing the likelihood function. In this process, the
 1209 $-\ln L$ is used in order to simplify and make easier the minimization. Terms which do not depend on
 1210 the rates ϵ_i and ϵ_j are removed in this step. This way, the final function to minimize is given by the
 1211 following expression:

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

1212 The events are selected within the Z peak and stored –with the electron order by $|\eta|$ – in two
 1213 triangular matrices: one for the same-sign events N_{ss}^{ij} , and the other one for all events N^{ij} . The
 1214 likelihood method takes into account electron pairs with all $|\eta|$ combinations, which allows to use the
 1215 full available statistics getting therefore lower statistical uncertainties. Moreover, it does not bias the
 1216 kinematical properties of the electrons, compared to other methods like tag-and-probe.

1217 The likelihood method can be easily extended to measure the charge mis-identification rates as a
 1218 function of two parameters. In this study, the interest lies not only on the measurement of the rates
 1219 as a function of the pseudo-rapidity, but also transverse momentum. Thus, the probability to find a
 1220 same-sign event given the rates for each electron is $(\epsilon_{i,k} + \epsilon_{j,l})$, where the two indices represent binned
 1221 $|\eta|$ - and p_T -dependence. Thus, the Eq. 8.5 transforms into

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

1222 The likelihood method uses only Z *signal* events. Therefore, background coming from other
 1223 processes where the dilepton invariant mass corresponds to the one of the Z boson needs to be
 1224 subtracted. The background subtraction is done using a simple side-band method. This method
 1225 consists in dividing the Z invariant mass in three regions, i.e. A , B and C , where B is the central
 1226 region corresponding to the Z peak. The number of events is counted in the regions on the sides of
 1227 the peak, i.e. n_A and n_C , and removed from the total number of events in the peak region B , n_B .
 1228 This way, the number of signal events N_Z is given by

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

1229 Once the background has been subtracted, the likelihood method can be applied. MINUIT is used
 1230 for the minimization and MIGRAD to compute the uncertainty on these rates.

1231 8.3.2 Results

1232 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] by
 1233 4 p_T bins [15,60,90,130,1000]. For p_T bins above 130 GeV, the Z dataset becomes too small and the
 1234 rates are calculated using $t\bar{t}$ MC, scaled by the data-MC ratio of the rates in the lower p_T bins, [90-130]
 1235 GeV. Figure 8.6 shows the extracted rates in all bins.

1236 To validate the likelihood approach, we apply the full method to the Z MC samples (extracting
 1237 rates via a likelihood fit and applying them to opposite sign events) and compare to the MC predicted
 1238 number of same-sign events. The invariant mass of the Z from our charge mis-identification and
 1239 directly from the MC can be seen on Figure 8.7. In the simulated Z samples, the number of same-
 1240 sign Z events is 5 049 while the estimation is $5\,031^{+375}_{-365}$. The uncertainties combine both statistical
 1241 systematic uncertainties, which are discussed in depth below. The validation gives compatible results
 1242 within uncertainties.

1243 8.3.3 Systematic and Statistical Uncertainties

1244 Statistical uncertainties dominate the combined uncertainty on the charge mis-identification estimate.
 1245 The statistical uncertainties come primarily from the size of the Z same-sign sample in data and are
 1246 especially large for central, material-poor regions where the charge mis-identification rate is extremely
 1247 low. Additionally systematic uncertainties are included for a comparison between the positron and
 1248 electron rate, the per-bin MC closure test discussed above in the Results section, and for the effect of
 1249 varying the invariant mass window used for the background subtraction for three different cases. The
 1250 high p_T extrapolation induces a statistical error only in the last p_T bin. This bins is essentially irrelevant

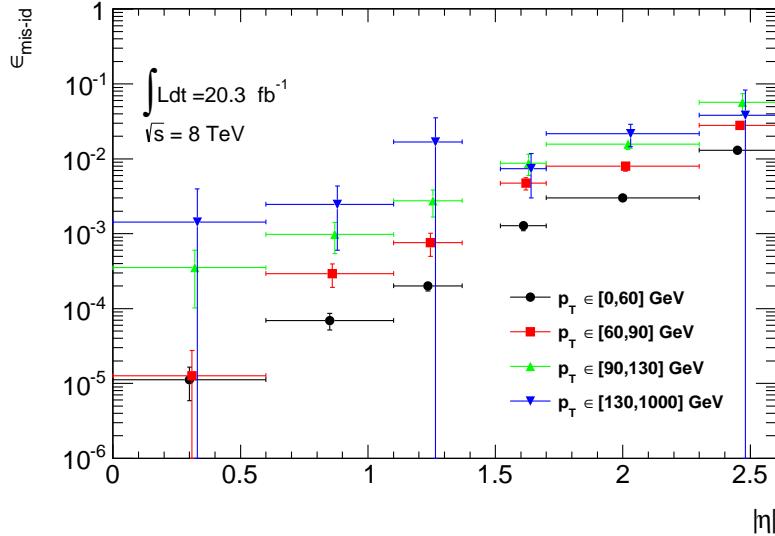


Figure 8.6: Electron charge mis-identification rates measured in data with the likelihood method on Z events (black points, red squares and blue triangles) as a function of $|\eta|$ and parametrized in p_T . The full 2012 dataset has been used to estimate the rates below 130 GeV. Above this value, the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

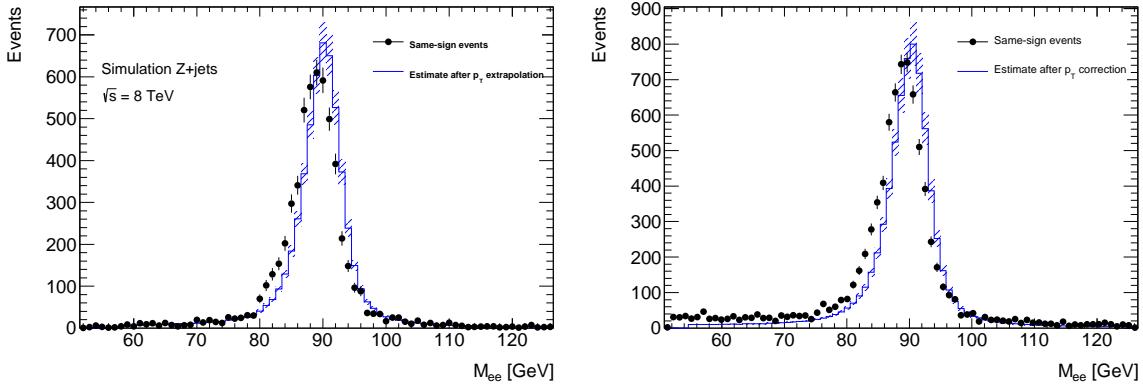


Figure 8.7: Closure test on simulated $Z \rightarrow e^+e^- + \text{jets}$ events (a) and data (b). The black circles show the distribution of same-sign events while the blue histograms show the distribution of the re-weighted opposite-sign events together with the statistical and systematic uncertainties.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

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

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

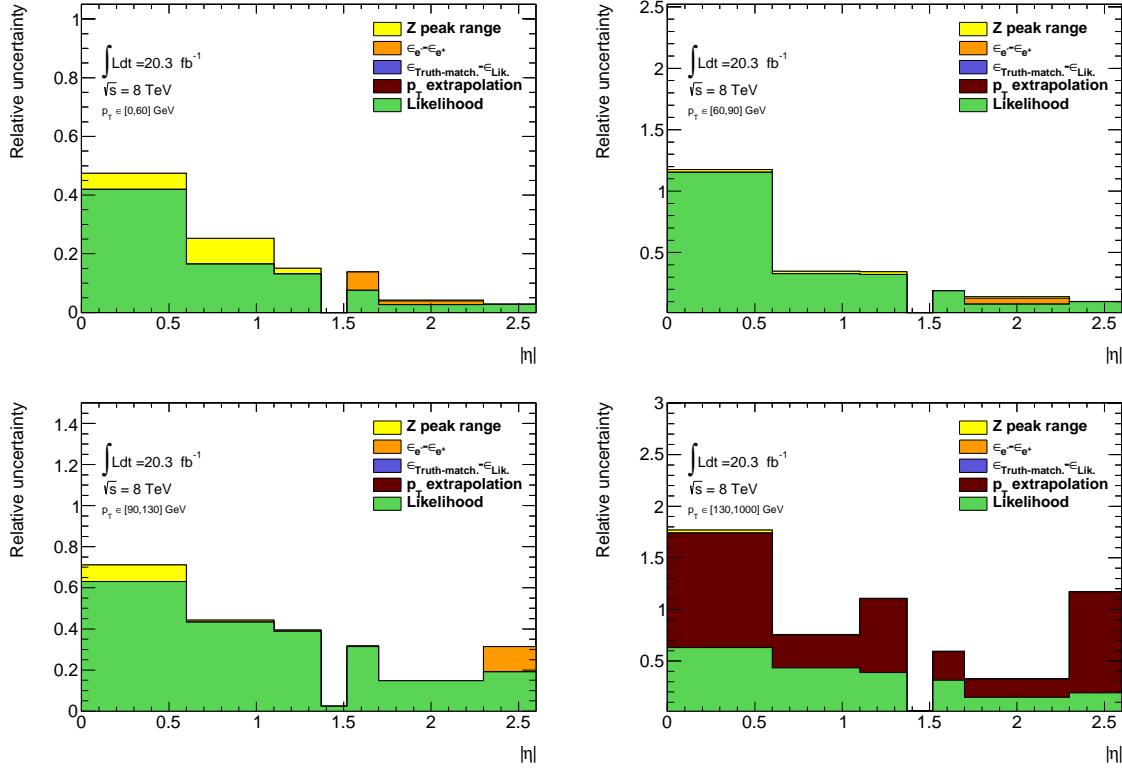


Figure 8.8: Relative systematic uncertainty contributions on the charge mis-identification rate, for different bins in p_T and $|\eta|$. Tight++ electrons have been used to produce this plot.

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

8.4 Fake Lepton Backgrounds

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

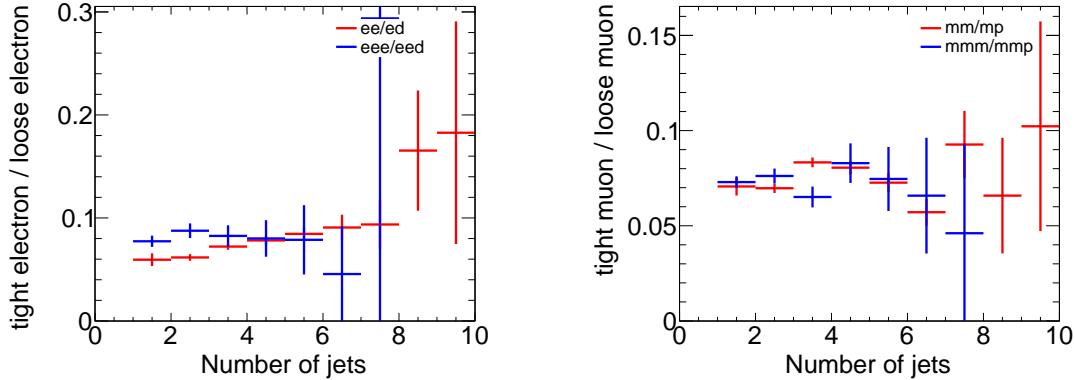


Figure 8.9: Ratios of regions with tight and anti-tight leptons in 2-lepton and 3-lepton channels

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

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

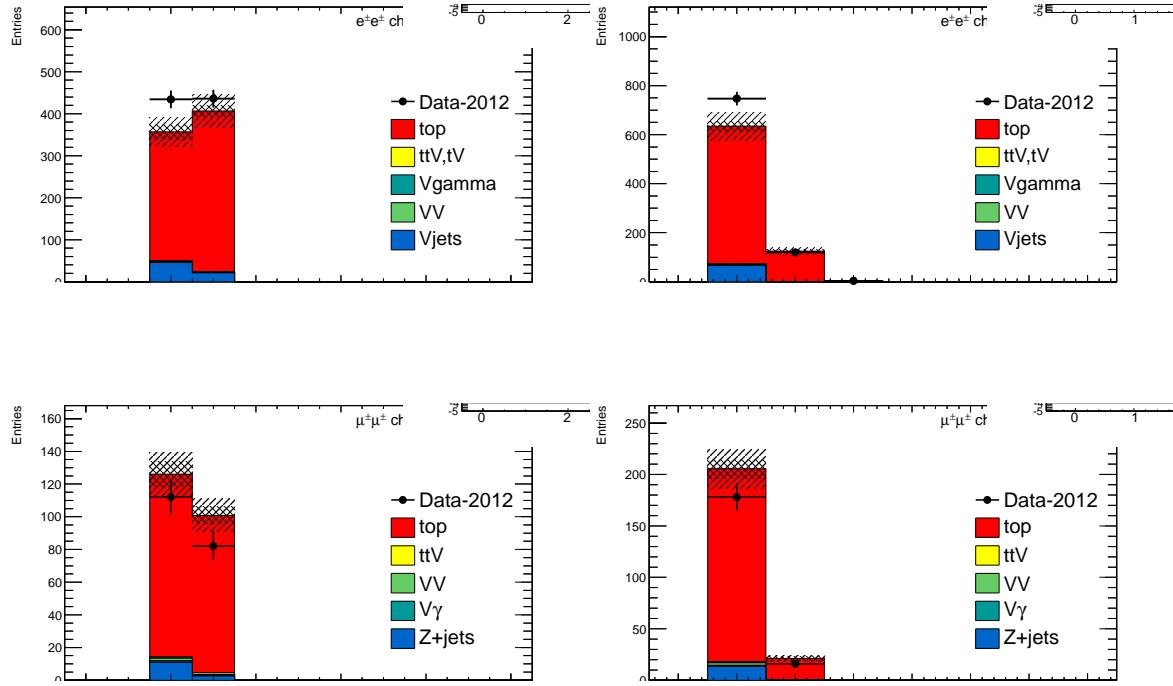


Figure 8.10: 2,3 Jet SS 2ℓ ed (above) and μp control regions with at least one b-tagged jet. This region and the associated tight-id regions are used to determine θ_e and θ_μ . The $t\bar{t}$ MC (red) is used for reference and not used in the actual calculation

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.

1311 The transfer factors obtained from the 2 and 3 jet regions are shown in 8.5 with statistical errors
 1312 and propagated systematic errors from the subtraction of relevant backgrounds ($t\bar{t}V$ and charge mis-
 1313 identification). The MC values are just for comparison. An additional systematic error is added by
 1314 comparing the transfer factors, obtained from the low jet control region, to those obtained from the
 1315 higher jet signal regions, using $t\bar{t}$ MC. The value of this systematic is about 20 % and can be seen in
 1316 the above Figure 8.9. The overall systematic uncertainties and contribution from each source in all of
 1317 the sub-channels of the signal region are shown in Table 8.5 and the final contribution of fake events
 1318 to the signal region are show in Table 8.1 found at the beginning of the chapter.

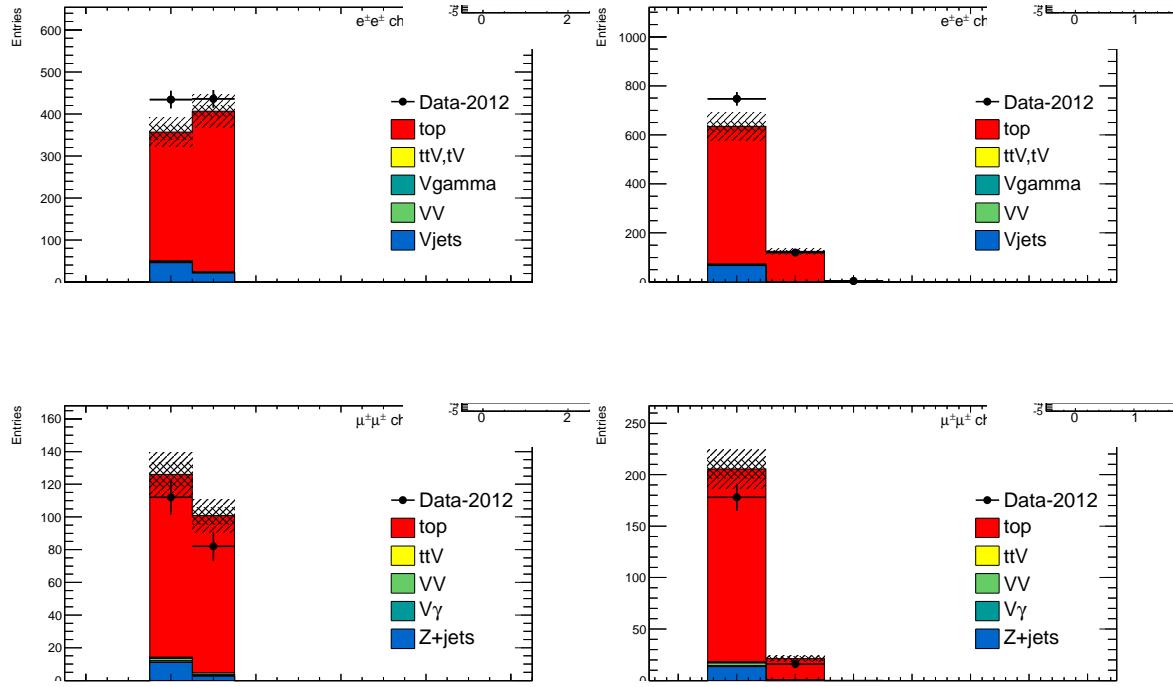


Figure 8.11: 4,5 Jet SS 2ℓ ed (above) and μp control regions with at least one b-tagged jet. After subtraction of prompt and charge mis-identification backgrounds, these regions are scaled by the transfer factors, θ_e and θ_μ to obtain the final number of fake events in the CR. The $t\bar{t}$ MC (red) 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)

1319 8.4.2 3ℓ Fakes

1320 The 3ℓ fake method follows the same general strategy as the 2ℓ SS case. Transfer factors are used
 1321 extrapolate from an anti-tight, fake-rich control region in data into the signal region. However, the
 1322 equivalent low jet control regions are too low in statistics to provide the transfer factors from data
 1323 directly, as above. Instead, the transfer factors are obtained directly from the $t\bar{t}$ simulation and data
 1324 control regions are used to determine the modeling of the identification and isolation variables, used in
 1325 the transfer factor extrapolation. The low jet regions are still employed in a low statistics validation
 1326 of the entire fake procedure.

1327 Anti-tight electrons and muons are defined in slightly different ways, compared to the 2ℓ SS case:

- 1328 • **anti-tight electron (d):** fails to verify the verytight likelihood operating point, but still verifies
 1329 the veryloose operating point. the isolation selection is released $E_T^{rel} > 0.05$, $p_T^{rel} > 0.05$.
- 1330 • **anti-tight muon(p):** muons must pass identification but the p_T cuts is lowered to 6 GeV, the
 1331 overlap removal with jets and isolation cuts are released.

1332 The transfer factors, θ_e and θ_μ , extracted from MC, is defined as the ratio of the number of top ($t\bar{t}$ +
 1333 single-top) events in the signal region, to the number of $t\bar{t}$ events in the anti-tight regions. The factors
 1334 are calculated in separate flavor regions to ensure that the electron jet fakes and muon jet fakes are
 1335 calculated separately. The calculation follows the same for as in Equation 8.8, but now lep0, which
 1336 by construction is almost never a fake is allowed to be either electron or muon in both cases, denoted
 1337 below in Equation 8.12.

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

1338 The MC modeling of the variables involved in the transfer factor can be verified when another
 1339 variable fails. For instance, the MC modeling of the electron isolation variable can be compared to
 1340 data when the particle identification variable fails and vice-versa. The modeling of muon-jet ΔR ,
 1341 involved in the overlap removal, can be compared when either the isolation variable or the p_T fails.
 1342 The comparison of the electron variables in this manner can be seen in Figure 8.13 and the muon
 1343 variables in Figure 8.12. The regions used have the same selection as the signal region with an added
 1344 missing transverse energy requirement, > 60 GeV, to ensure only top fakes. 20% and 21% systematic
 1345 uncertainties are assigned to the muon and electron transfer factors, respectively, to account for data-
 1346 MC discrepancies. This method for evaluating data-MC agreement for individual electron and muon
 1347 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.

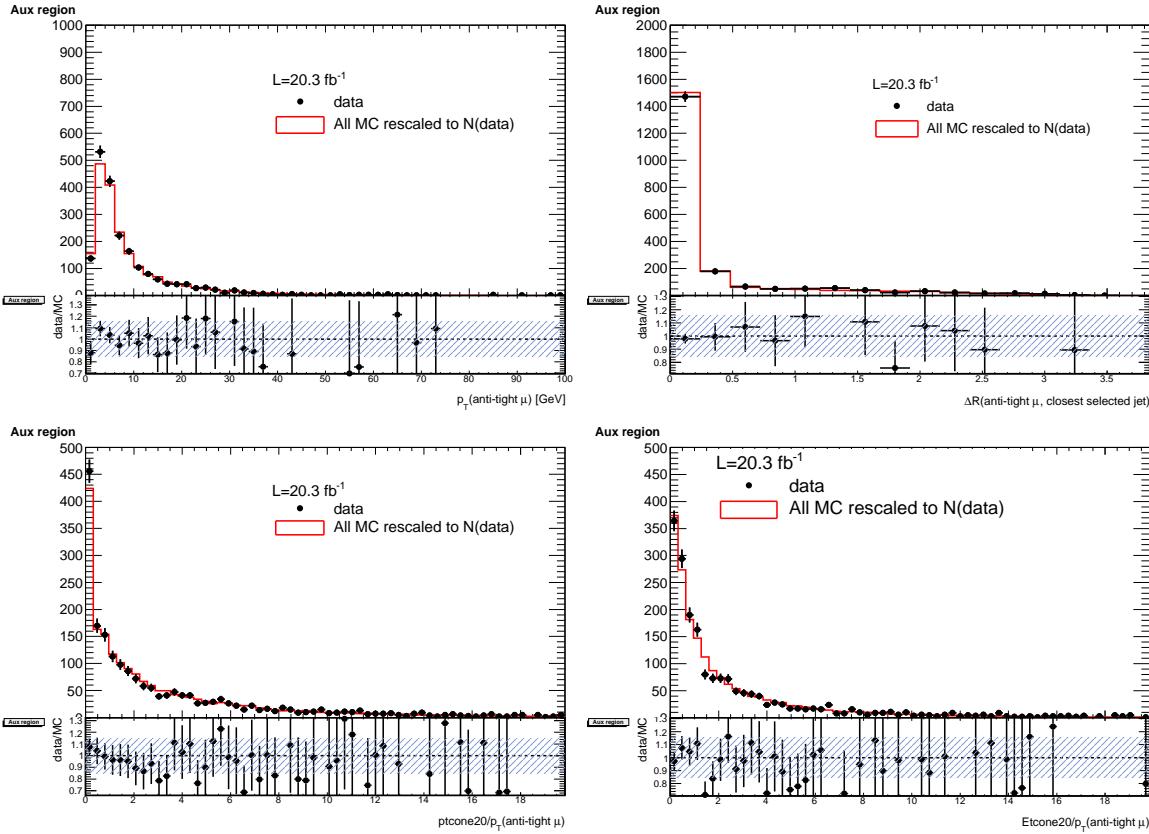


Figure 8.12: Distributions of the properties of the anti-tight muons in data (dots), compared with the total simulation (red line), rescaled to the integral of the data for a shape comparison. The uncertainty on the data distribution is statistical. The number of events for each of them is also presented in the legend. The variables probed are, top: p_T and $\Delta R(\mu, \text{closest selected jet})$; bottom: $ptcone20/p_T$ and $Etcone20/p_T$. 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 in

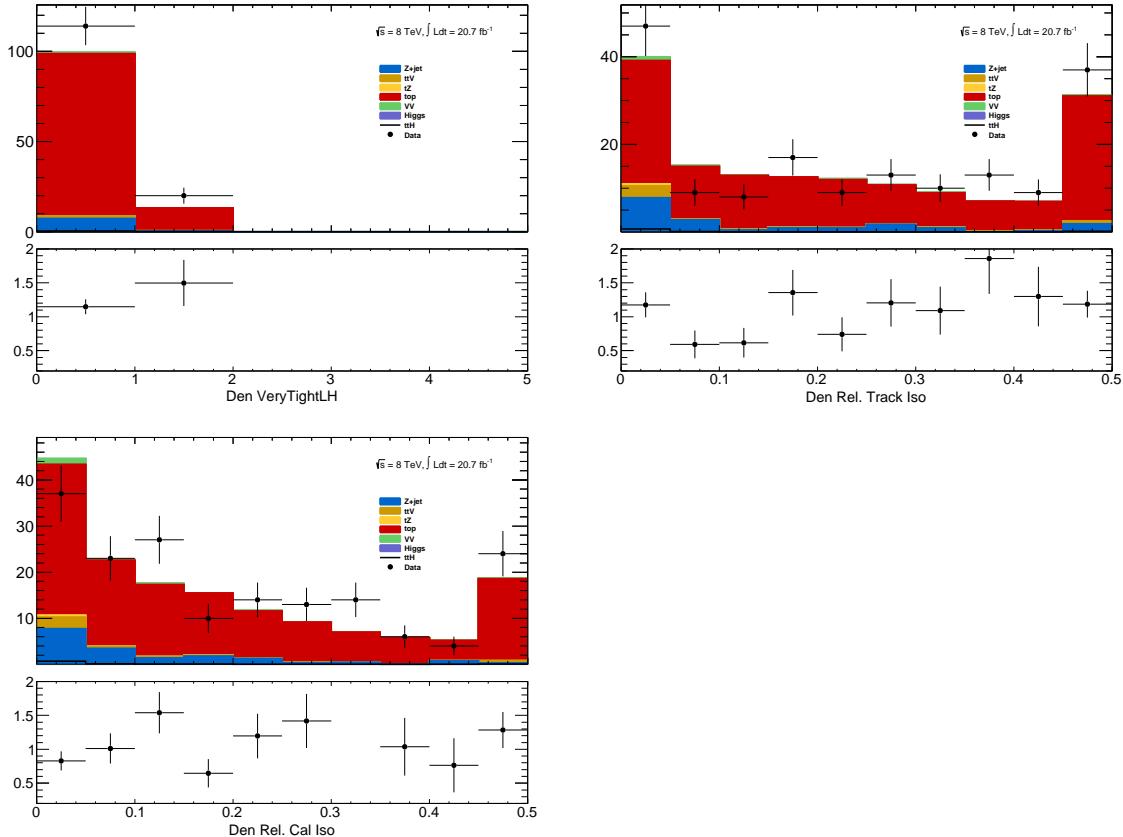


Figure 8.13: Distributions of anti-tight electron variables. The variables presented are, from top left to bottom right, p_T , η , Very Tight Likelihood word value, $\text{ptcone20}/p_T$, $\text{Etcone20}/p_T$. The plotted regions have the same cuts as the signal region, except the anti-tight electron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

the calculation are shown in Table 8.7.

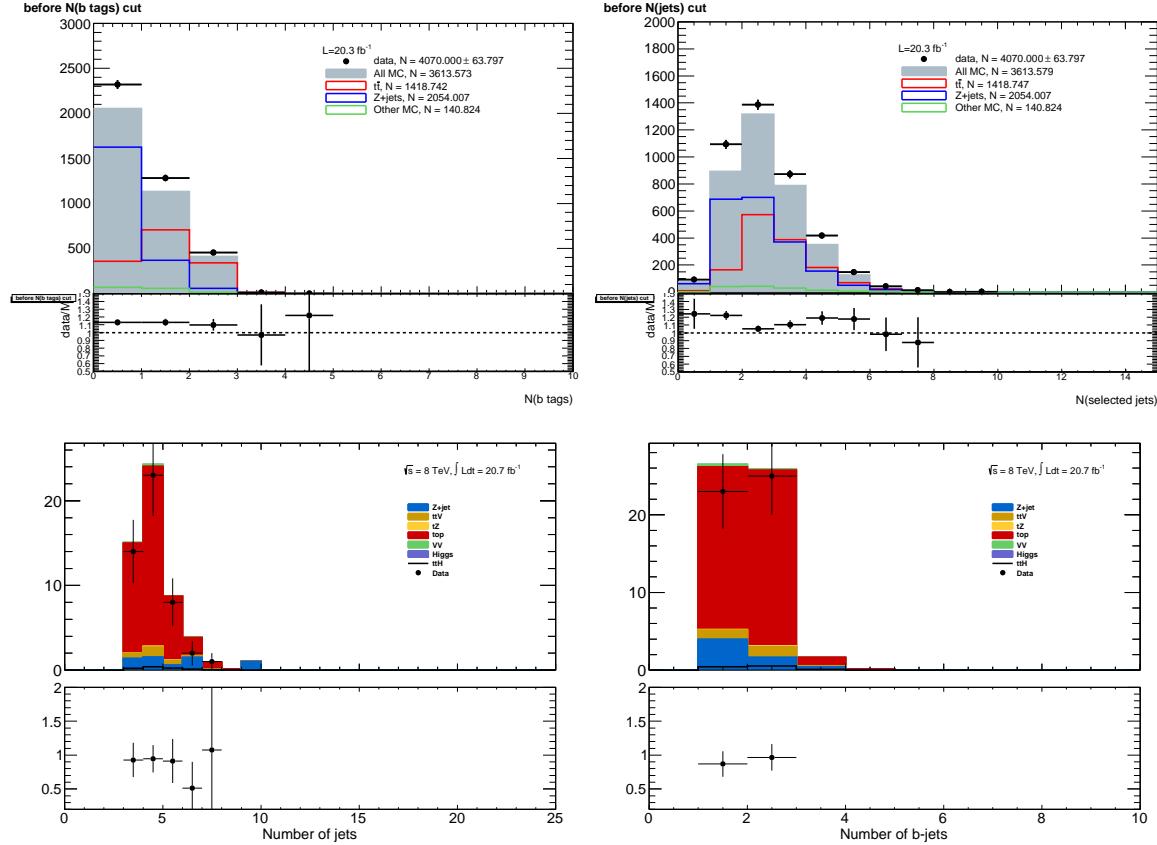


Figure 8.14: Muon-xxp (above) and electron-xxd (below) anti-tight control regions: jet variables. Note: the $t\bar{t}$ and top MC in the plots is used only as comparison, but is not included in the fake measurement

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

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

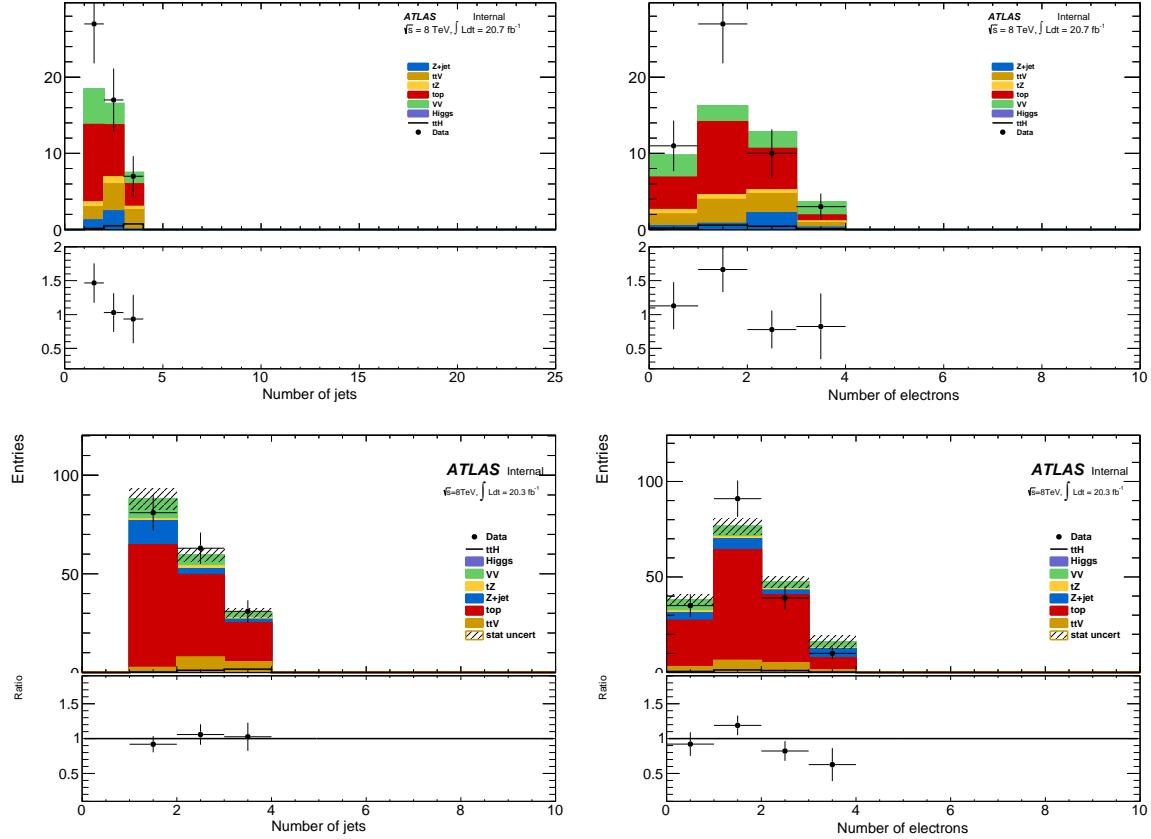


Figure 8.15: 3ℓ fake validation regions for nominal p_T selection (above) and relaxed p_T selection, > 10 GeV, (below). Plotted are the number of jets and the number of electrons in each event. The data and MC ratio below each plot agree with 1 within the statistics of the region and the overall systematic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

1372 fully identified and 2 anti-identified leptons are created. These control regions do not have a number
1373 of jets requirement in order to increase statistics. From these control regions, two extrapolations
1374 are made. First, a transfer factor is applied to extrapolate from the anti-tight to tight regions for
1375 electrons and muons. The regions are defined with identical object identification selection and reversal
1376 as the 3ℓ case, and the same transfer factors can be used. They must be used twice however, because
1377 there are two anti-identified leptons in each event. Second, the jet inclusive regions are extrapolated
1378 into the 2-jet signal region, using as a second extrapolation factor derived from $t\bar{t}$ events. Since, the
1379 majority of fake leptons arise from b-quark initiated jets, the jet spectrum $t\bar{t}$ events with the additional
1380 requirement of 2-b-tagged jets from data are used as a model for the jet extrapolation. The overall
1381 systematic uncertainty on this measurement arises from the statistics in the control regions and MC
1382 based assessments of non-closure and are 35%-50% depending on the sub-channel.

1383 CHAPTER 9

1384 Summary of Systematic Uncertainties

1385 This chapter summarizes the various systematic uncertainties that enter the measurement of the limit
1386 of $t\bar{t}H$ multi-lepton analysis, as well as structure of the statistical analysis model used to obtain
1387 the measurement. The systematic uncertainties arise from three main sources. The first are the
1388 normalization uncertainties on the background process estimation methods, which are discussed in
1389 depth in . The second source is the theoretical uncertainties on the $t\bar{t}H$ production cross-section and
1390 acceptance. The final source are the experimental and detector related systematic uncertainties related
1391 to event selection efficiencies and measurements and identification of the objects. They affect only
1392 the non-data driven backgrounds and the $t\bar{t}H$ signal, as simulation is used to model their acceptance
1393 and efficiency for the analysis selection.

1394 These systematic uncertainties, the estimated background and signal event counts in each of the
1395 signal regions, and the observed data in each signal region are combined in a statistical fit to an
1396 analysis model to extract the measurement of interest. We measure per-channel and combined ratios
1397 of the observed production rate to the theoretically predicted production rate of $t\bar{t}H$, a parameter
1398 called μ . In the absence of a statistically significant observation, this measurement is translated into
1399 a upper confidence limit on μ . The details of this procedure are discussed in the following sections
1400 and the results with the observed data are discussed in Chapter 10

1401 9.1 Systematic Uncertainties on Signal Cross-section and Acceptance

1402 The $t\bar{t}H$ signal is simulated with matrix elements at NLO (next-to-leading order) in QCD with Powhel
1403 and is discussed in Chapter 6.

1404 The production cross section and the Higgs boson decay branching fractions together with their
1405 theoretical uncertainties from the QCD scale and PDF choice are taken from the NLO theoretical

calculations reported in Ref. [66]. The uncertainty from the QCD scale estimated by varying μ_0 by a factor of 2 from the nominal value is $^{+3.8\%}_{-9.3\%}$, while the uncertainty from the PDF set and the value of α_S is $\pm 8.1\%$.

1409

1410 The impact of the choice of the QCD scale on the simulation of the $t\bar{t}H$ event selection efficiency
1411 is estimated in two independent ways.

1412 First, the factorization and renormalisation scales μ_0 are varied by a factor of 2, as $\mu = 2\mu_0$ and
1413 $\mu = \mu_0/2$. The effects of these new scales are estimated via the application of event re-weighting
1414 procedures on the nominal simulation using kinematic distributions at parton level. The weights used
1415 are dependent on the transverse momenta of both the $t\bar{t}H$ system and of the top quark, as described
1416 in Ref. [81].

1417 Second, the choice of the factorization and renormalisation scales, dependent on fixed (“static”)
1418 parameters in the nominal simulation, is tested comparing its prediction with an alternative (“dy-
1419 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.
1420 This comparison is performed via event re-weighting of the nominal static simulation based on weights
1421 derived as a function of the $t\bar{t}H$ transverse momentum [81]. In order to take the difference between
1422 the choices of scale as systematic uncertainties, a symmetric envelope around the nominal simulation
1423 is built applying the weights and also their inverses.

1424 Fig. 9.1 shows the impact of the different choices for the factorization and renormalization scales
1425 on the jet and b-jet multiplicities in events with 2 SS leptons. Similar variations are seen also in
1426 the other event categories. In order to not double-count the variations on the total cross section
1427 the predictions from the different QCD scales are normalized to the same total cross section. That
1428 means that the observed differences are only coming from the event selection. Significant variations
1429 on the jet multiplicities can be seen and these translate into different predictions on the signal event
1430 yields in the signal regions. Such differences, listed in Table 9.1, are taken as theoretical systematic
1431 uncertainties in addition to the ones affecting the total $t\bar{t}H$ production cross section. The “Static”
1432 uncertainties come from the variations by a factor of 2 from the nominal scale and they are correlated
1433 with the uncertainties on the total cross section, which are estimated with the same procedure. The
1434 “Dynamic” uncertainties come from the difference between the nominal and the alternative dynamic
1435 scale and are treated as an independent source of theoretical uncertainty.

1436 The uncertainty of the $t\bar{t}H$ event selection due to the PDF sets is estimated comparing the predic-
1437 tions with three different PDF sets, varying each set within errors and taking the width of the envelope
1438 as systematic uncertainty. The recommended sets are CT10, MSTW2008nlo68cl and NNPDF21_100. We

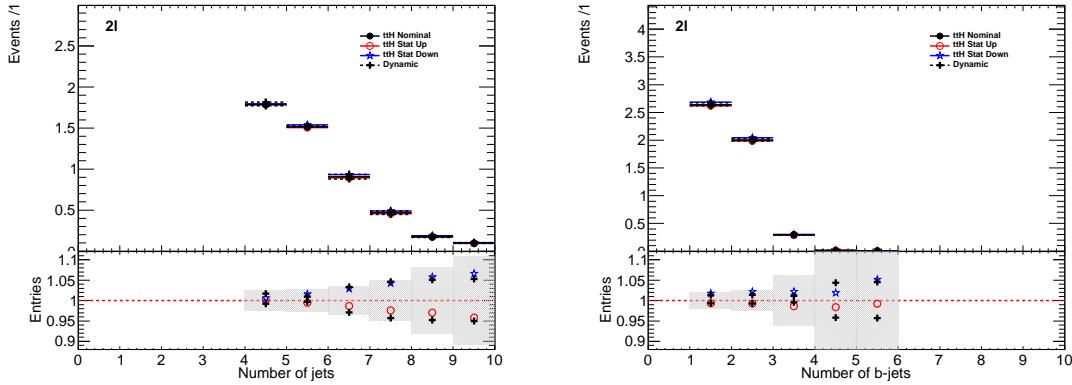


Figure 9.1: Effects on the jet multiplicities in 2 SS lepton $t\bar{H}$ events from different choices of the factorization and renormalization scales. “Static” refers to the variations by a factor of 2 of the nominal μ_0 , while “Dynamic” refers to the alternative choice of μ_0 which depends on the event 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 [%]	$2\ell 4\text{jets}$	$2\ell \geq 5\text{jets}$	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	$2\ell 4\text{j}$	$2\ell 5\text{j}$	3ℓ	4ℓ
$t\bar{H}$	0.3%	1.0%	0.5%	1.4%

1439 determine the change in the acceptance due to the PDF sets via the formula 9.1 to disentangle it from
 1440 the change in production cross section.

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

1441 Fig. 9.2 shows the estimated PDF systematic uncertainties as a function of the jet multiplicity in
 1442 $t\bar{H}$ events with at least two leptons. The uncertainties are compatible with the uncertainty on the
 1443 production cross section estimated in Ref. [66] and indicated by the dashed red lines in the lower panel.
 1444 Table 9.2 shows the half-width of the envelope of the acceptance under all eigenvector variations of
 1445 the three PDF sets. No significant dependence on the event topology is observed, so that the PDF
 1446 systematic uncertainty on the $t\bar{H}$ event selection is neglected.

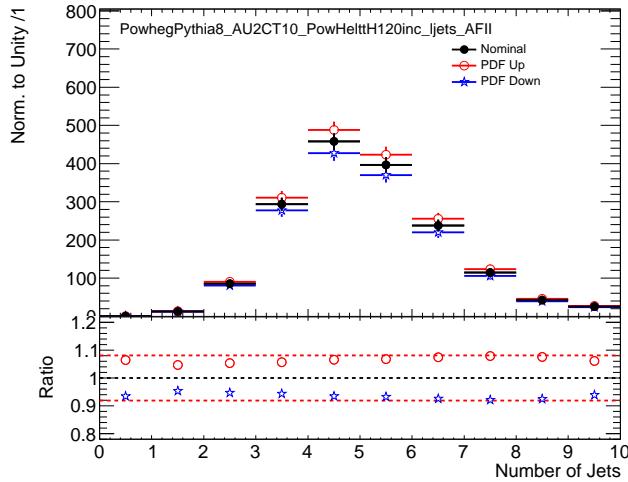


Figure 9.2: PDF systematic uncertainty on the jet multiplicities in $t\bar{t}H$ events with at least 2 leptons. The dashed red lines in the lower panel indicate the systematic uncertainty on the $t\bar{t}H$ production cross section.

1447 9.2 Experimental and Detector Systematic Uncertainties

1448 Experimental and detector systematic uncertainties arise from the efficiency of identifying objects
 1449 and the efficiency of the event selections. These affect only MC models of physics processes, $t\bar{t}V$,
 1450 $t\bar{t}H$, VV . Data-driven backgrounds take into account these effects by construction. We consider
 1451 systematic effects from a number of sources: the lepton and jet energy scale measurements, the lepton
 1452 identification and isolation selections, the efficiency and mis-identification rate associated with tagging
 1453 b-quark jets. Effects due to modeling the energy and objects from additional vertices were studied
 1454 and found to be negligible. The vast majority of the individual detector systematics effects are small.
 1455 The sum total of the systematic effects are comparable to some of the overall normalization and
 1456 cross-section uncertainties on some of the physics processes and is shown in Table ??.

1457 9.2.1 Lepton Identification, Energy Scale, and Trigger

1458 The electron[51] and muon identification efficiencies[82] are measured in data using Z boson and
 1459 J/Ψ control samples. They are shown in Figure 9.3. The uncertainty on the muon efficiencies are
 1460 measured as functions of η and p_T and are generally less 1%. The uncertainty on the electron and
 1461 muon efficiencies are also measured as functions of η and p_T and are at the 1 % level for p_T above 30
 1462 GeV, but become much larger 5-10% for the lower p_T regimes. These translate into sub-% level effects
 1463 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 the

1464 electrons. The effects become more important with increasing numbers of leptons.

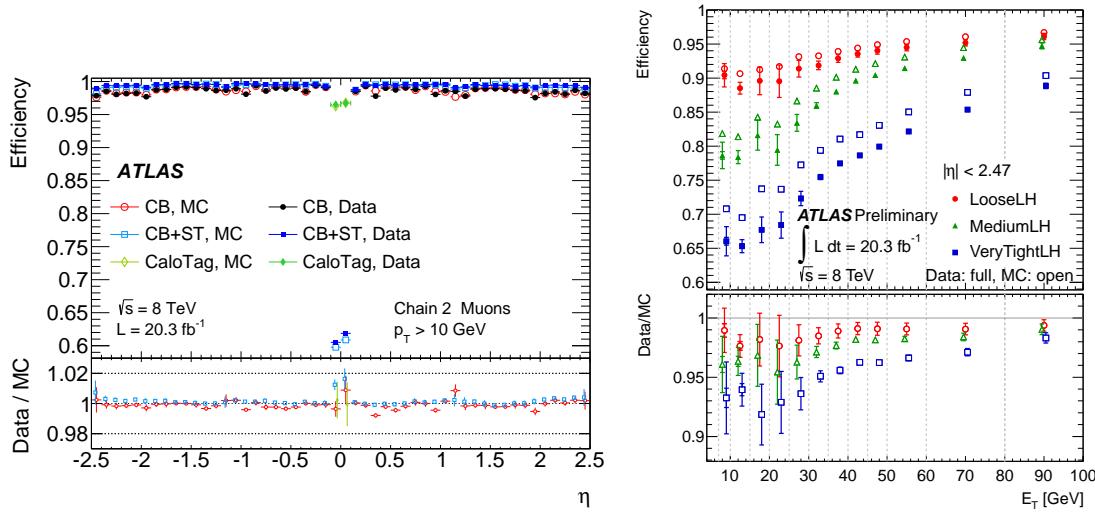


Figure 9.3: Muon (left) and electron(right) identification efficiencies in Data and MC as a function of η and p_T respectively. For electrons, the verytight likelihood operating point is used and for muons the CB+ST (combined+segment tagged) operating point is used

1465 The electron[83] and muon[82] energy scale and resolution are also measured using the Z -boson
 1466 control samples in data. The uncertainties related to the scale and resolution for the leptons affect
 1467 the overall event acceptance through the lepton momentum cuts primary and have negligible impact
 1468 on the event count uncertainties in the signal regions.

1469 The efficiencies for muons and electrons to pass muon[84] and electron triggers[85] have been
 1470 calculated with respect to the offline identification operating points using the Z boson control samples.
 1471 They are in the range of 90% for electron triggers and 70% for muon triggers, owing to gaps in
 1472 muon trigger coverage, and have % level errors. When statistically combined for 2,3,4 and lepton
 1473 signal regions, the overall trigger efficiency is high and the error on the number of expected events is
 1474 negligible.

1475 9.2.2 Lepton Isolation and Impact Parameter

1476 The isolation and impact parameter selections are specific to this analysis and are discussed in depth
 1477 in Chapter 7. We calculated their combined efficiency with respect to the full lepton identification

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.

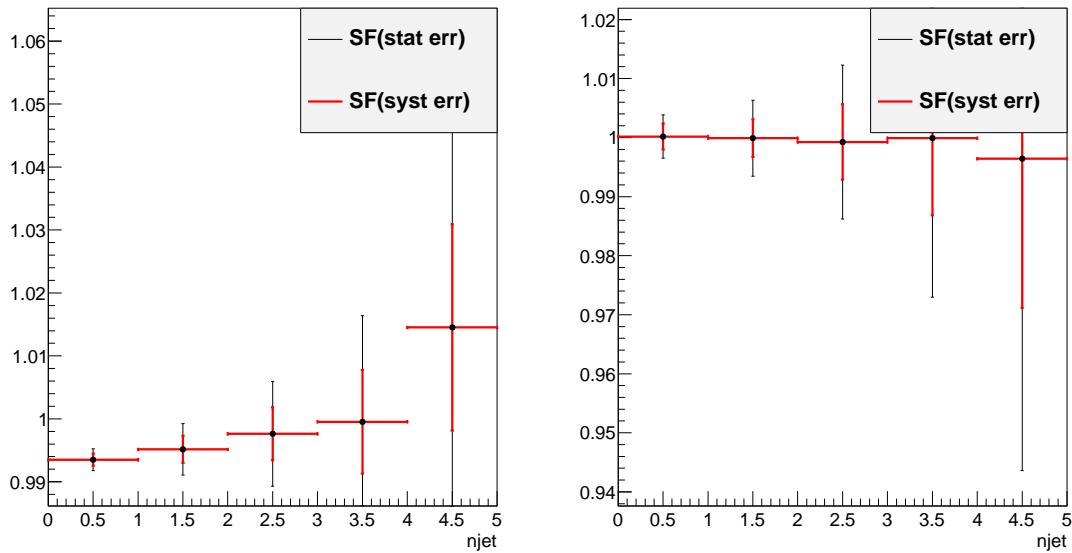


Figure 9.4: Muon (left) and electron (right) isolation efficiency scale-factors from the Z control sample as a function of the number of jets in the event. An additional systematic uncertainty of 1% is added to encompass the variation in the number of jets variable

9.2.3 Jet Energy

The jet energy scale (JES) is calculated using a combination of data-based in-situ techniques, where jet transverse momentum is balanced with respect to a reference photon or a Z boson, as well as single particle test-stand studies[86]. Additional smaller effects are taken into account including the b-quark jet specific response, near-by jets, the effects of pile-up and an inter-calibration of similar η regions using di-jet events. The JES systematic errors arises from numerous sources that are diagonalized into eigenvectors so that they can be combined in an uncorrelated way. The combined uncertainty 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 analysis. The jet energy resolution is calculated in a similar way with slightly larger errors, 10% [87]. Propagated to the event counts in the signal regions, the combined scale and resolution systematics are of non-negligible effects 6-7%.

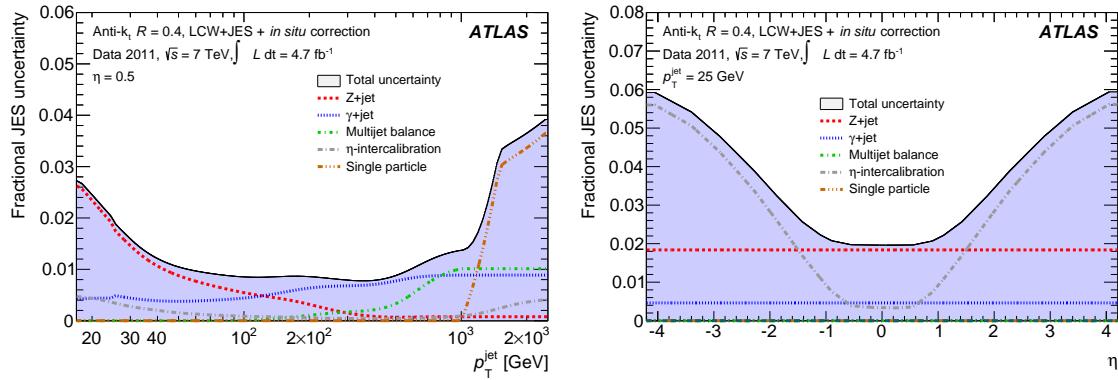


Figure 9.5: JES systematic uncertainties as a function of jet η (for jets $p_T > 25$ GeV) and p_T (for jets $|\eta| < 0.4$). The combined systematic uncertainty is shown with contributions from the largest sources

9.2.4 B-Tagged Jet Efficiency

The b -quark tagging efficiency must be calculated separately for charm, light and b -quarks. ATLAS uses three data based control regions: an inclusive jet sample for mis-tagged light quarks[88], the $t\bar{t}$ sample for b -quarks[89], and a sample of D^* mesons for charm quarks[90]. These efficiencies and rates are well-measured in MC and the data-based corrections are small. The data-MC efficiency scale-factor shown in Figure 9.6 is close to 1 and has an overall systematic uncertainty of around 5%. The uncertainties are applied to the analysis via a number of eigenvectors. Together these uncertainties have a 4 % effect in the event expectation in the signal regions.

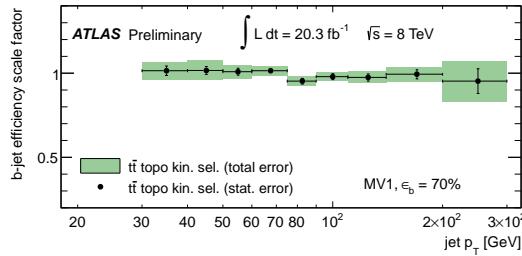


Figure 9.6: b-Tagging data-MC efficiency scale-factors versus jet p_T calculated in the $t\bar{t}$ sample from 2012 data. The uncertainties are combined statistical and systematic.

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 systematic uncertainties 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.

1523

CHAPTER 10

1524

Results and Statistical Model

10.1 Results in Signal Regions

10.2 Statistical Model

NEED TO DEFINE MU IN ANALYSIS SUMMARY CHAPTER

We use the above results to make two sets of measurements: an upper confidence limit on μ , the signal strength parameter, and a measurement of μ . These measurements are done for each channel individually and then combined. The interpretation of the results in the form of a statistical model follow the procedure, discussed here [91]. We interpret the results as counting experiments in each signal region. Therefore agreement in kinematic shapes do not affect the statistical procedure.

10.2.1 The Likelihood

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

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

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

1542 The signal and background expectations , s and b , depend on systematic errors. These are included
 1543 in the likelihood function in the form of a vector nuisance parameters, $\vec{\theta}$, which are constrained to
 1544 fluctuate within Gaussian distributions. These fluctuations affect the background and signal expecta-
 1545 tions by response functions, $\nu(\vec{\theta})$, set by uncertainties measured in the previous section. For instance,
 1546 the $W^\pm Z$ normalization uncertainty is 50% from Section 8.2 and is included in the fit as its own unit
 1547 gaussian, $G(\theta|0, 1)$. The fluctuations of the gaussian, θ_{WZ} scale the background contribution via the
 1548 form, $0.5 \cdot (1 + \theta_{WZ}) \cdot b_{WZ}$. For many of the detector systematics, the uncertainties are two sided
 1549 and are included as piecewise Gaussians. We add correlations to various uncertainties by hand, when
 1550 appropriate. With these nuisance parameters, the likelihood takes this form:

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

1551 10.2.2 Test Statistic and Profile Likelihood

1552 Values of μ are tested with the negative log quantity, $q_\mu = -2\ln(\lambda(\mu))$, where $\lambda(\mu)$ is the test statistic.
 1553 $\lambda(\mu)$ is defined as:

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

1554 where $\hat{\vec{\theta}}_\mu$ are values of the nuisance parameter vector that maximize the likelihood for a given value
 1555 of μ and $\hat{\mu}$ and $\hat{\vec{\theta}}$ are the fitted values of signal strength and nuisance parameters that maximize the
 1556 likelihood overall. μ is constrained to be positive.

1557 10.2.3 CL_s Method

1558 Exclusions limits on the signal strength are calculated with the test statistic using a modified fre-
 1559 quentist method, called the CL_s method[92]. CL_s is defined as a ratio of two frequentist quantities.
 1560 The numerator quantifies the probability of finding the observed data given the signal + background
 1561 hypothesis. The denominator quantifies the probability of the data given the background only hy-
 1562 pothesis.

1563 Using the numerator alone has the undesirable property that, if the data fluctuates below the
 1564 expectation, an exclusion limit can be reached that is far beyond the sensitivity of the experiment.
 1565 Normalizing to the background only hypothesis penalizes these low sensitivity cases.

1566 The probability of obtaining an observation as extreme as the data given a particular signal +
 1567 background hypothesis is given by the p-value, p_μ defined as:

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

1568 and the probability of obtaining an observation as extreme as the data given the background hypothesis
 1569 p_b is:

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

1570 where $f(q_{\mu})$ is the distribution of q_{μ} for all possible observations for a given μ and q is defined above.
 1571 Therefore,

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

1572 . A μ value is considered excluded at 95% confidence when CL_s is less than 5%.

1573 10.2.4 Exclusion Limits

1574 Table ?? shows expected exclusion limits for all channels, including the analysis uncertainties cumula-
 1575 tively. It shows the relative importance of the statistical and systematic uncertainties to the analysis
 1576 sensitivity. The observed limits using observed data and predictions can be seen in Figures ??-?? for
 1577 splitting and combining the sub-channels and in Table XX by numbers. We expect a combined limit
 1578 of XXX (background only) and XXX (signal injected) and see a limit of XXX. The channel sensitivity
 1579 is dominated by the 2ℓ and 3ℓ channels.

Channels	Stat	+Fakes Unc.	+Theory	+ Experimental
2ℓ	2 ℓ ee	7.44	8.52	8.94
	2 ℓ em	3.46	3.81	4.18
	2 ℓ mm	4.03	4.14	4.57
	2 ℓ tau	8.08	8.92	10.03
	All	2.16	2.44	2.90
3ℓ		3.40	3.43	3.66
4ℓ		15.16	15.16	15.55
1l2tau		10.41	13.84	14.22
All		1.68	1.85	2.22

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

1580 10.2.5 μ Measurements

1581 In addition to setting a limit on the signal strength, we also fit the best value of the signal strength
 1582 for μ . We do this by minimizing the negative log likelihood value, q_{μ} or conversely maximizing the
 1583 likelihood. The $1-\sigma$ error band is set via a profile likelihood scan, where the value q_{μ} is scanned as
 1584 a function of μ . Values of μ that increase q_{μ} by 1 form the edges of the error band. The fitted
 1585 values of μ with errors are provided in Table XXX for each sub-channel fit as well as the combined
 1586 fit.

1587 10.2.6 Nuisance Parameter Impact on the Signal Strength

1588 Finally, we examine the post-fit impact of the various nuisance parameters on the final fit. We expect
1589 to have measured the various analysis uncertainties well and do not expect the fit to have much further
1590 constraint. For that reason, we expect the pulls of the nuisance parameters to be close to 0 and the
1591 measured uncertainties on those parameters to be consistent with the input uncertainties. Figures
1592 XXXX show.

1593

CHAPTER 11

1594

Conclusions

1595 **11.1 Higgs Results in Review**

1596 **11.2 Prospects for Future**

Bibliography

- 1598 [1] S. L. Glashow, *Partial-symmetries of weak interactions*, Nucl. Phys. **22** (1961) no. 4, 579. [2.1.1](#)
- 1599 [2] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264. [2.1.1](#)
- 1600 [3] A. Salam and J. C. Ward, *Gauge theory of elementary interactions*, Phys. Rev. **136** (1964) 763–768. [2.1.1](#)
- 1601 [4] S. Weinberg, *Non-abelian gauge theories of the strong interactions*, Phys. Rev. Lett. **31** (1973) 494–497. [2.1.1](#)
- 1602 [5] G. 't Hooft and M. Veltman, *Regularization and renormalization of gauge fields*, Nuclear Physics B **44** (1972) 189 – 213. [2.1.1](#)
- 1603 [6] D. J. Gross and F. Wilczek, *Ultraviolet behavior of non-abelian gauge theories*, Phys. Rev. Lett. **30** (1973) 1343–1346. [2.1.1](#)
- 1604 [7] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508. [2.1.2](#)
- 1605 [8] P. W. Higgs, *Spontaneous symmetry breakdown without massless bosons*, Phys. Rev. **145** (1966) 1156. [2.1.2](#)
- 1606 [9] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321–322. [2.1.2](#)
- 1607 [10] The ALEPH, CDF, DØ, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the Tevatron Electroweak Working Group, and the SLD electroweak and heavy flavour groups, *Precision Electroweak Measurements and Constraints on the Standard Model*, CERN-PH-EP-2010-095 (2010) , [arXiv:1012.2367 \[hep-ex\]](#). [2.1.3](#), [2.2](#)
- 1608 [11] M. Baak, M. Goebel, J. Haller, A. Hoecker, D. Kennedy, R. Kogler, K. Mnig, M. Schott, and J. Stelzer, *The electroweak fit of the standard model after the discovery of a new boson at the LHC*, The European Physical Journal C **72** (2012) no. 11, .
<http://dx.doi.org/10.1140/epjc/s10052-012-2205-9>. [2.1.3](#)
- 1609 [12] J. C. Collins, D. E. Soper, and G. Sterman, *Factorization for short distance hadron-hadron scattering*, Nuclear Physics B **261** (1985) 104 – 142. [2.2](#)
- 1610 [13] CERN, . CERN, Geneva, 1984. [2.2](#)

- 1625 [14] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and
 1626 R. Tanaka (Eds.), *Handbook of LHC Higgs Cross Sections: 2. Differential Distributions*,
 1627 CERN-2012-002 (CERN, Geneva, 2012) , [arXiv:1201.3084 \[hep-ph\]](https://arxiv.org/abs/1201.3084). **2.2**
- 1628 [15] *Updated coupling measurements of the Higgs boson with the ATLAS detector using up to 25*
 1629 *fb⁻¹ of proton-proton collision data*, Tech. Rep. ATLAS-CONF-2014-009, CERN, Geneva, Mar,
 1630 2014. **2.2.1**
- 1631 [16] CMS Collaboration Collaboration, *Precise determination of the mass of the Higgs boson and*
 1632 *studies of the compatibility of its couplings with the standard model*, Tech. Rep.
 1633 CMS-PAS-HIG-14-009, CERN, Geneva, 2014. **2.2.1**
- 1634 [17] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of the Higgs boson mass from*
 1635 *the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4\ell$ channels with the ATLAS detector using 25 fb⁻¹ of pp*
 1636 *collision data*, [arXiv:1406.3827 \[hep-ex\]](https://arxiv.org/abs/1406.3827). **2.2.1**
- 1637 [18] *Evidence for the spin-0 nature of the Higgs boson using {ATLAS} data*, Physics Letters B **726**
 1638 (2013) no. 13, 120 – 144.
<http://www.sciencedirect.com/science/article/pii/S0370269313006527>. **2.2.1**
- 1640 [19] S. Dawson, A. Gritsan, H. Logan, J. Qian, C. Tully, et al., *Working Group Report: Higgs*
 1641 *Boson*, [arXiv:1310.8361 \[hep-ex\]](https://arxiv.org/abs/1310.8361). **2.2.2**
- 1642 [20] O. Eberhardt, G. Herbert, H. Lacker, A. Lenz, A. Menzel, et al., *Impact of a Higgs boson at a*
 1643 *mass of 126 GeV on the standard model with three and four fermion generations*,
 1644 *Phys.Rev.Lett.* **109** (2012) 241802, [arXiv:1209.1101 \[hep-ph\]](https://arxiv.org/abs/1209.1101). **2.2.2**
- 1645 [21] M. Carena, S. Gori, N. R. Shah, C. E. Wagner, and L.-T. Wang, *Light Stops, Light Staus and*
 1646 *the 125 GeV Higgs*, *JHEP* **1308** (2013) 087, [arXiv:1303.4414](https://arxiv.org/abs/1303.4414). **2.2.2**
- 1647 [22] N. Arkani-Hamed, K. Blum, R. T. D’Agnolo, and J. Fan, *2:1 for Naturalness at the LHC?*,
 1648 *JHEP* **1301** (2013) 149, [arXiv:1207.4482 \[hep-ph\]](https://arxiv.org/abs/1207.4482). **2.2.2**
- 1649 [23] D. Carmi, A. Falkowski, E. Kuflik, and T. Volansky, *Interpreting LHC Higgs Results from*
 1650 *Natural New Physics Perspective*, *JHEP* **1207** (2012) 136, [arXiv:1202.3144 \[hep-ph\]](https://arxiv.org/abs/1202.3144). **2.2.2**
- 1651 [24] G. Degrassi, S. Di Vita, J. Elias-Miro, J. R. Espinosa, G. F. Giudice, et al., *Higgs mass and*
 1652 *vacuum stability in the Standard Model at NNLO*, *JHEP* **1208** (2012) 098, [arXiv:1205.6497](https://arxiv.org/abs/1205.6497)
 1653 [\[hep-ph\]](https://arxiv.org/abs/1205.6497 [hep-ph]). **2.2.2**
- 1654 [25] L. Evans and P. Bryant, *LHC Machine*, *JINST* **3** (2008) no. 08, S08001. **3.1**
- 1655 [26] T. S. Pettersson and P. Lefevre, *The Large Hadron Collider: conceptual design.*, Tech. Rep.
 1656 CERN-AC-95-05 LHC, CERN, Geneva, Oct, 1995. <https://cdsweb.cern.ch/record/291782>.
 1657 **3.1**
- 1658 [27] T. Linnecar et al., *Hardware and Initial Beam Commissioning of the LHC RF Systems.*
 1659 *oai:cds.cern.ch:1176380*, Tech. Rep. LHC-PROJECT-Report-1172.
 1660 CERN-LHC-PROJECT-Report-1172, CERN, Geneva, Oct, 2008.
 1661 <https://cdsweb.cern.ch/record/1176380>. **3.1**
- 1662 [28] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3**
 1663 (2008) S08003. **3.1**

- 1664 [29] The CMS Collaboration, *The CMS experiment at the CERN LHC*, *Journal of Instrumentation*
1665 **3** (2008) no. 08, S08004. [3.1](#)
- 1666 [30] The LHCb Collaboration, *The LHCb Detector at the LHC*, *Journal of Instrumentation* **3** (2008)
1667 no. 08, S08005. [3.1](#)
- 1668 [31] The ALICE Collaboration, *The ALICE experiment at the CERN LHC*, *Journal of*
1669 *Instrumentation* **3** (2008) no. 08, S08002.
1670 <http://stacks.iop.org/1748-0221/3/i=08/a=S08002>. [3.1](#)
- 1671 [32] A. Team, *The four main LHC experiments*, Jun, 1999. [3.1](#)
- 1672 [33] ATLAS Collaboration, *ATLAS inner detector: Technical Design Report 1*. Technical Design
1673 Report ATLAS. CERN, Geneva, 1997. <https://cdsweb.cern.ch/record/331063>. [3.2](#)
- 1674 [34] ATLAS Collaboration, *ATLAS inner detector: Technical Design Report, 2*. Technical Design
1675 Report ATLAS. CERN, Geneva, 1997. <https://cdsweb.cern.ch/record/331064>. [3.2](#)
- 1676 [35] ATLAS Collaboration, *ATLAS magnet system: Technical Design Report, 1*. Technical Design
1677 Report ATLAS. CERN, Geneva, 1997. <https://cdsweb.cern.ch/record/338080>. [3.2](#)
- 1678 [36] ATLAS Collaboration, *ATLAS pixel detector: Technical Design Report*. Technical Design
1679 Report ATLAS. CERN, Geneva, 1998. <https://cdsweb.cern.ch/record/381263>. [3.2](#)
- 1680 [37] ATLAS Collaboration, *ATLAS pixel detector electronics and sensors*, *JINST* **3** (2008) P07007.
1681 [3.2](#)
- 1682 [38] ATLAS Collaboration, *The barrel modules of the ATLAS semiconductor tracker*,
1683 *Nucl.Instrum.Meth.* **A568** (2006) 642–671. [3.2](#)
- 1684 [39] ATLAS Collaboration, *The ATLAS semiconductor tracker end-cap module*, *Nucl.Instrum.Meth.*
1685 **A575** (2007) 353–389. [3.2](#)
- 1686 [40] The ATLAS TRT Collaboration, *The ATLAS Transition Radiation Tracker (TRT) proportional
1687 drift tube: Design and performance*, *JINST* **3** (2008) P02013. [3.2](#)
- 1688 [41] The ATLAS TRT Collaboration, *The ATLAS TRT barrel detector*, *JINST* **3** (2008) P02014. [3.2](#)
- 1689 [42] E. Abat et al., *The ATLAS TRT electronics*, *J. Instrum.* **3** (2008) P06007. [3.2](#)
- 1690 [43] T. A. Collaboration, *ATLAS liquid argon calorimeter: Technical design report*, .
1691 CERN-LHCC-96-41. [3.2](#)
- 1692 [44] ATLAS Collaboration, *ATLAS tile calorimeter: Technical Design Report*. Technical Design
1693 Report ATLAS. CERN, Geneva, 1996. <https://cdsweb.cern.ch/record/331062>. [3.2](#)
- 1694 [45] ATLAS Collaboration, *ATLAS muon spectrometer: Technical Design Report*. Technical Design
1695 Report ATLAS. CERN, Geneva, 1997. <https://cdsweb.cern.ch/record/331068>. [3.2](#)
- 1696 [46] G. Aielli, A. Aloisio, M. Alviggi, V. Aprodu, V. Bocci, et al., *The RPC first level muon trigger
1697 in the barrel of the ATLAS experiment*, *Nucl.Phys.Proc.Suppl.* **158** (2006) 11–15. [3.2](#)
- 1698 [47] F. Bauer, U. Bratzler, H. Dietl, H. Kroha, T. Lagouri, et al., *Construction and test of MDT
1699 chambers for the ATLAS muon spectrometer*, *Nucl.Instrum.Meth.* **A461** (2001) 17–20. [3.2](#)

- 1700 [48] T. Argyropoulos, K. A. Assamagan, B. H. Benedict, V. Chernyatin, E. Cheu, et al., *Cathode*
 1701 *strip chambers in ATLAS: Installation, commissioning and in situ performance*, IEEE
 1702 *Trans.Nucl.Sci.* **56** (2009) 1568–1574. [3.2](#)
- 1703 [49] ATLAS Collaboration, *ATLAS level-1 trigger: Technical Design Report*. Technical Design
 1704 Report ATLAS. CERN, Geneva, 1998. <https://cdsweb.cern.ch/record/381429>. [3.2](#)
- 1705 [50] P. Jenni, M. Nessi, M. Nordberg, and K. Smith, *ATLAS high-level trigger, data-acquisition and*
 1706 *controls: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 2003.
 1707 <https://cdsweb.cern.ch/record/616089>. [3.2](#)
- 1708 [51] *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton*
 1709 *collision data*, Tech. Rep. ATLAS-CONF-2014-032, CERN, Geneva, Jun, 2014. [3.2.6.2](#), [7.4](#),
 1710 [9.2.1](#)
- 1711 [52] *Preliminary results on the muon reconstruction efficiency, momentum resolution, and*
 1712 *momentum scale in ATLAS 2012 pp collision data*, Tech. Rep. ATLAS-CONF-2013-088, CERN,
 1713 Geneva, February, 2013. [3.2.6.3](#), [7.5](#)
- 1714 [53] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti- $k(t)$ jet clustering algorithm*, *JHEP* **0804**
 1715 (2008) 063, [arXiv:0802.1189 \[hep-ph\]](https://arxiv.org/abs/0802.1189). [3.2.6.4](#), [7.6](#)
- 1716 [54] ATLAS Collaboration, *Commissioning of the ATLAS high-performance b-tagging algorithms in*
 1717 *the 7 TeV collision data*, Tech. Rep. ATLAS-CONF-2011-102, CERN, Geneva, Jul, 2011.
 1718 <https://cdsweb.cern.ch/record/1369219>. [3.2.6.5](#), [7.6](#)
- 1719 [55] ATLAS Collaboration Collaboration, G. Aad et al., *Search for $H \rightarrow \gamma\gamma$ produced in association*
 1720 *with top quarks and constraints on the Yukawa coupling between the top quark and the Higgs*
 1721 *boson using data taken at 7 TeV and 8 TeV with the ATLAS detector*, [arXiv:1409.3122](https://arxiv.org/abs/1409.3122)
 1722 [\[hep-ex\]](#). [5](#)
- 1723 [56] *Search for the Standard Model Higgs boson produced in association with top quarks and decaying*
 1724 *to $b\bar{b}$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector at the LHC*, Tech. Rep.
 1725 ATLAS-CONF-2014-011, CERN, Geneva, Mar, 2014. [5](#)
- 1726 [57] ATLAS Collaboration, G. Aad et al., *Improved luminosity determination in pp collisions at \sqrt{s}*
 1727 *= 7 TeV using the ATLAS detector at the LHC*, *Eur.Phys.J.* **C73** (2013) 2518, [arXiv:1302.4393](https://arxiv.org/abs/1302.4393)
 1728 [\[hep-ex\]](#). [6.1.1](#)
- 1729 [58] CERN, . CERN, Geneva, 2012. [6.1.1](#)
- 1730 [59] ATLAS Collaboration Collaboration, G. Aad et al., *The ATLAS Simulation Infrastructure*,
 1731 *Eur.Phys.J.* **C70** (2010) 823–874, [arXiv:1005.4568 \[physics.ins-det\]](https://arxiv.org/abs/1005.4568). [6.2](#)
- 1732 [60] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A Simulation toolkit*,
 1733 *Nucl.Instrum.Meth.* **A506** (2003) 250–303. [6.2](#)
- 1734 [61] G. Bevilacqua, M. Czakon, M. Garzelli, A. van Hameren, A. Kardos, C. Papadopoulos,
 1735 R. Pittau, and M. Worek. [6.2.1](#)
- 1736 [62] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, *11*
 1737 (2004) 040. [6.2.1](#)
- 1738 [63] S. Frixione, P. Nason, and C. Oleari **11** (2007) 070, [arXiv:0709.2092 \[hep-ph\]](https://arxiv.org/abs/0709.2092). [6.2.1](#)

- [64] S. Alioli, P. Nason, C. Oleari, and E. Re **06** (2010) 040, [arXiv:1002.2581 \[hep-ph\]](#). 6.2.1
- [65] T. Sjöstrand, S. Mrenna, and P. Skands, *A Brief Introduction to Pythia 8.1*, [arXiv:0710.3820 \[hep-ph\]](#). 6.2.1
- [66] LHC Higgs Cross Section Working Group Collaboration, S. Heinemeyer et al., *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, [arXiv:1307.1347 \[hep-ph\]](#). 6.2.1, 9.1
- [67] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, **JHEP 0307** (2003) 001, [arXiv:hep-ph/0206293 \[hep-ph\]](#). 6.2.2
- [68] F. Maltoni and T. Stelzer, *MadEvent: Automatic event generation with MadGraph*, **JHEP 0302** (2003) 027, [arXiv:hep-ph/0208156 \[hep-ph\]](#). 6.2.2
- [69] B. P. Kersevan and E. Richter-Was, *The Monte Carlo event generator AcerMC versions 2.0 to 3.8 with interfaces to PYTHIA 6.4, HERWIG 6.5 and ARIADNE 4.1*, **Comput.Phys.Commun.** **184** (2013) 919–985, [arXiv:hep-ph/0405247 \[hep-ph\]](#). 6.2.2
- [70] P. M. Nadolsky, H.-L. Lai, Q.-H. Cao, J. Huston, J. Pumplin, et al., *Implications of CTEQ global analysis for collider observables*, **Phys.Rev. D78** (2008) 013004, [arXiv:0802.0007 \[hep-ph\]](#). 6.2.2
- [71] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, **JHEP 0711** (2007) 070, [arXiv:0709.2092 \[hep-ph\]](#). 6.2.2
- [72] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, et al., *Event generation with SHERPA 1.1*, **JHEP 0902** (2009) 007, [arXiv:0811.4622 \[hep-ph\]](#). 6.2.2
- [73] A. Martin, W. Stirling, R. Thorne, and G. Watt, *Parton distributions for the LHC*, **Eur.Phys.J. C63** (2009) 189–285, [arXiv:0901.0002 \[hep-ph\]](#). 6.2.2, 8.1
- [74] 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
- [75] 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
- [76] 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
- [77] 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
- [78] 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
- [79] 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

- 1777 [80] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of differential production*
 1778 *cross-sections for a Z boson in association with b-jets in 7 TeV proton-proton collisions with*
 1779 *the ATLAS detector*, arXiv:1407.3643 [hep-ex]. 8.2
- 1780 [81] S. Guindon, E. Shabalina, J. Adelman, M. Alhroob, S. Amor dos Santos, A. Basye, J. Bouffard,
 1781 M. Casolino, I. Connelly, A. Cortes Gonzalez, V. Dao, S. D'Auria, A. Doyle, P. Ferrari,
 1782 F. Filthaut, R. Goncalo, N. de Groot, S. Henkelmann, V. Jain, A. Juste, G. Kirby, D. Kar,
 1783 A. Knue, K. Kroeninger, T. Liss, E. Le Menedeu, J. Montejo Berlingen, M. Moreno Llacer,
 1784 O. Nackenhorst, T. Neep, A. Onofre, M. Owen, M. Pinamonti, Y. Qin, A. Quadt, D. Quilty,
 1785 C. Schwanenberger, L. Serkin, R. St Denis, J. Thomas-Wilsker, and T. Vazquez-Schroeder,
 1786 *Search for the Standard Model Higgs boson produced in association with top quarks and decaying*
 1787 *to bb in pp collisions at sqrt(s) = 8 TeV with the ATLAS detector at the LHC*, Tech. Rep.
 1788 ATL-COM-PHYS-2013-1659, CERN, Geneva, Dec, 2013. The note contains internal
 1789 documentation of the ttH(bb) analysis approved as a preliminary result
 1790 (ATLAS-CONF-2014-011). 9.1
- 1791 [82] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of the muon reconstruction*
 1792 *performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data*,
 1793 arXiv:1407.3935 [hep-ex]. 9.2.1, 9.2.1
- 1794 [83] ATLAS Collaboration Collaboration, G. Aad et al., *Electron and photon energy calibration with*
 1795 *the ATLAS detector using LHC Run 1 data*, Eur.Phys.J. C74 (2014) no. 10, 3071,
 1796 arXiv:1407.5063 [hep-ex]. 9.2.1
- 1797 [84] ATLAS Collaboration Collaboration, *Performance of the ATLAS muon trigger in 2011*, Tech.
 1798 Rep. ATLAS-CONF-2012-099, CERN, Geneva, Jul, 2012. 9.2.1
- 1799 [85] ATLAS Collaboration Collaboration, *Performance of the ATLAS Electron and Photon Trigger*
 1800 *in p-p Collisions at sqrt(s) = 7 TeV in 2011*, Tech. Rep. ATLAS-CONF-2012-048, CERN,
 1801 Geneva, May, 2012. 9.2.1
- 1802 [86] ATLAS Collaboration Collaboration, G. Aad et al., *Jet energy measurement and its systematic*
 1803 *uncertainty in proton-proton collisions at sqrt(s) = 7 TeV with the ATLAS detector*,
 1804 arXiv:1406.0076 [hep-ex]. 9.2.3
- 1805 [87] ATLAS Collaboration Collaboration, G. Aad et al., *Jet energy resolution in proton-proton*
 1806 *collisions at sqrt(s) = 7 TeV recorded in 2010 with the ATLAS detector*, Eur.Phys.J. C73 (2013)
 1807 2306, arXiv:1210.6210 [hep-ex]. 9.2.3
- 1808 [88] ATLAS Collaboration Collaboration, *Measurement of the Mistag Rate with 5 fb^-1 of Data*
 1809 *Collected by the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2012-040, CERN, Geneva, 2012.
 1810 9.2.4
- 1811 [89] ATLAS Collaboration Collaboration, *Measuring the b-tag efficiency in a top-pair sample with*
 1812 *4.7 fb^-1 of data from the ATLAS detector*, Tech. Rep. ATLAS-CONF-2012-097, CERN,
 1813 Geneva, 2012. 9.2.4
- 1814 [90] ATLAS Collaboration Collaboration, *b-jet tagging calibration on c-jets containing D?+ mesons*,
 1815 Tech. Rep. ATLAS-CONF-2012-039, CERN, Geneva, 2012. 9.2.4
- 1816 [91] Glen Cowan, Kyle Cranmer, Eilam Gross, Ofer Vitells, *Asymptotic formulae for likelihood-based*
 1817 *tests of new physics*, Eur.Phys.J.C 71 (2011) 1554. 10.2

- 1818 [92] A. L. Read, *Presentation of search results: the CL s technique*, Journal of Physics G: Nuclear
1819 and Particle Physics **28** (2002) no. 10, 2693.
1820 <http://stacks.iop.org/0954-3899/28/i=10/a=313>. 10.2.3