

¹ A SEARCH FOR THE HIGGS BOSON PRODUCED IN
² ASSOCIATION WITH TOP QUARKS IN MULTILEPTON
³ FINAL STATES AT ATLAS

⁴ Chris Lester

⁵ A DISSERTATION
⁶ in
⁷ Physics and Astronomy

⁸ Presented to the Faculties of The University of Pennsylvania
⁹ in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy
¹⁰ 2014

¹¹
¹² I. Joseph Kroll, Professor, Physics
¹³ Supervisor of Dissertation

¹⁴
¹⁵ Elliot Lipeles, Associate Professor, Physics
¹⁶ Graduate Group Chairperson

¹⁷ Dissertation Committee
¹⁸ Justin Khouri, Associate Professor, Physics
¹⁹ I. Joseph Kroll, Professor, Physics
²⁰ Burt Ovrut, Professor, Physics
²¹ Evelyn Thompson, Associate Professor, Physics
²² I. Joseph Kroll, Professor, Physics

²³ A SEARCH FOR THE HIGGS BOSON PRODUCED IN ASSOCIATION WITH TOP
²⁴ QUARKS IN MULTILEPTON FINAL STATES AT ATLAS

²⁵ COPYRIGHT
²⁶ 2014
²⁷ Chris Lester

²⁸ All rights reserved.

29

Acknowledgements

30 Acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements
31 acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements
32 acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements
33 acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements
34 acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements
35 acknowledgements acknowledgements acknowledgements.

36 Acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements
37 acknowledgements 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 acknowledgements acknowledgements.

42 Acknowledgements acknowledgements acknowledgements acknowledgements acknowledgements
43 acknowledgements 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 edgements acknowledgements acknowledgements acknowledgements.

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 ab-

55 stract abstract abstract abstract abstract abstract abstract abstract abstract abstract

56 abstract abstract abstract abstract abstract abstract abstract abstract abstract ab-

57 stract abstract abstract abstract abstract abstract abstract abstract abstract abstract

58 abstract abstract abstract abstract abstract abstract abstract abstract abstract ab-

59 stract abstract abstract abstract abstract abstract abstract abstract abstract abstract

60 abstract abstract abstract abstract abstract abstract abstract abstract abstract ab-

61 stract abstract abstract abstract abstract abstract abstract abstract abstract abstract

62 abstract abstract abstract abstract.

Contents

64	Acknowledgements	iii
65	Abstract	v
66	Contents	vi
67	List of Tables	xii
68	List of Figures	xiii
69	Preface	xvi
70	1 Introduction	1
71	2 Theoretical Background	2
72	2.1 The Standard Model	2
73	2.1.1 The Standard Model Structure	2
74	2.1.2 Electroweak Symmetry Breaking and the Higgs	4
75	2.1.3 The Standard Model Parameters	5
76	2.2 Collider Physics and the Higgs	6
77	2.2.1 Higgs Discovery at the LHC	9

78	2.2.2 $t\bar{t}H$ Production	11
79	2.3 Conclusion	12
80	3 The Large Hadron Collider and the ATLAS Experiment	14
81	3.1 The Large Hadron Collider	14
82	3.1.1 The Accelerator Complex	15
83	3.1.2 Beam Parameters and Collisions	16
84	3.2 The ATLAS Experiment	16
85	3.2.1 Detector Coordinate System	18
86	3.2.2 The Inner Detector	19
87	3.2.3 The Calorimeter	22
88	3.2.4 The Muon Spectrometer	25
89	3.2.5 The Trigger System	26
90	3.2.6 Reconstruction: Jets, Muons and Electrons	27
91	3.2.6.1 Tracks and Clusters	28
92	3.2.6.2 Electrons	29
93	3.2.6.3 Muons	29
94	3.2.6.4 Jets	30
95	3.2.6.5 B-Tagged Jets	30
96	4 Electrons	32
97	4.1 Electrons at Hadron Colliders	32
98	4.2 Identification of Electrons at ATLAS	33
99	4.2.1 Pile-up and Electron identification	33
100	4.2.2 2011 Menu and Trigger	34
101	4.2.3 2012 Menu and Trigger	34

102	4.2.4 Electron Likelihood	34
103	4.3 Measurement of Electron Efficiency at ATLAS	34
104	4.3.1 Techniques	34
105	4.3.2 Issues	34
106	5 $t\bar{t}H$ Analysis Summary	35
107	5.1 Signal Characteristics	36
108	5.2 Background Overview	37
109	5.3 Analysis Strategy	38
110	6 Dataset and Simulation	39
111	6.1 Data	39
112	6.1.1 The 2012 Dataset	39
113	6.2 Simulation	41
114	6.2.1 Signal Simulation	41
115	6.2.2 Background Simulation	42
116	7 Object and Event Selection	44
117	7.1 2ℓ Same-Charge Signal Region	44
118	7.2 3ℓ Signal Region	46
119	7.3 4ℓ Signal Region	46
120	7.4 Electron Selection	47
121	7.5 Muon Selection	48
122	7.6 Jet and b-Tagged Jet Selection	49
123	7.7 Tau Selection	49
124	7.8 Object Summary and Overlap	50

125	7.9 Optimization	50
126	8 Background Estimation	51
127	8.1 Vector Boson (W^\pm, Z) production in association with top quarks: $t\bar{t}V, tZ$	53
128	8.1.1 $t\bar{t}Z$ Validation Region	54
129	8.2 Di-boson Background Estimation: $W^\pm Z, ZZ$	55
130	8.2.1 $W^\pm Z$ Uncertainty	56
131	8.2.2 ZZ Uncertainty	59
132	8.3 Charge-Misidentification Background	61
133	8.3.1 Likelihood Method	62
134	8.3.2 Results	64
135	8.3.3 Systematic and Statistical Uncertainties	65
136	8.4 Fake Lepton Backgrounds	66
137	8.4.1 2ℓ SS Fakes	69
138	8.4.2 3ℓ Fakes	71
139	8.4.3 4ℓ Fakes	78
140	9 Summary of Systematic Uncertainties	79
141	9.1 Systematic Uncertainties on Signal Cross-section and Acceptance	80
142	9.2 Experimental and Detector Systematic Uncertainties	83
143	9.2.1 Lepton Identification, Energy Scale, and Trigger	84
144	9.2.2 Lepton Isolation and Impact Parameter	85
145	9.2.3 Jet Energy	86
146	9.2.4 B-Tagged Jet Efficiency	87
147	9.2.5 Summary	88
148	9.3 Summary of Background and Signal Normalization Uncertainties	89

¹⁴⁹	10 Results and Statistical Model	90
¹⁵⁰	10.1 Results in Signal Regions	90
¹⁵¹	10.2 Statistical Model	90
¹⁵²	10.2.1 The Likelihood	90
¹⁵³	10.2.2 Test Statistic and Profile Likelihood	91
¹⁵⁴	10.2.3 CL_s Method	92
¹⁵⁵	10.2.4 Exclusion Limits	93
¹⁵⁶	10.2.5 μ Measurements	93
¹⁵⁷	10.2.6 Nuisance Parameter Impact on the Signal Strength	94
¹⁵⁸	11 Conclusions	95
¹⁵⁹	11.1 Higgs Results in Review	95
¹⁶⁰	11.2 Prospects for Future	95
¹⁶¹	Bibliography	96

List of Tables

163	5.1	Contributions of the main Higgs decay modes to the 3 multi-lepton $t\bar{t}H$ signatures at generation level.	36
165	6.1	Monte Carlo samples used for signal description.	42
166	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.	43
169	7.1	Selections in the 2ℓ SS, 3ℓ and 4ℓ Signal Regions	45
170	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.	50
173	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 $t\bar{t}\text{fake}$).	52
177	8.2	NLO cross section and theoretical uncertainty calculations derived from MadGraph5_aMC@NLO.	54
178	8.3	Truth Origin of highest energy b-tagged jet in the $W^\pm Z + b$ VR and $3l$ SR	58
179	8.4	Truth Origin of highest energy b-tagged jet in the $ZZ + b$ VR and $4l$ SR	59
180	8.5	Expected and measured values of the θ factors.	71
181	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)	72
186	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	76
188	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.	82
189	9.2	Uncertainties on $t\bar{t}H$ acceptance in signal regions due to PDF variation.	82

191	9.3	Sum in quadrature of all the systematic uncertainties on the number of event yields per channel.	88
193	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.	89
196	10.1	95%CL limits on μ for all channels and combination with cumulative uncertainties.	93

List of Figures

198 2.1	χ^2 as a function of the Higgs mass (top left), the top quark mass (top right), the W boson mass (bottom left) and the effective weak mixing angle (bottom right) for the combined SM fit from the GFitter group. The data points placed along $\chi^2 = 1$ represent direct measurements of the respective observable and their $\pm 1\sigma$ uncertainties. The grey (blue) bands show the results when excluding (including) the new M_H measurements from (in) the fits.	6
204 2.2	Proton Parton Distribution Functions (PDFs) form the MSTW Collaboration at $Q^2 = 10 \text{ GeV}^2$ and $Q^2 = 10^4 \text{ GeV}^2$	7
206 2.3	Dominant Higgs production modes at the LHC	8
207 2.4	8 TeV LHC Higgs production cross-sections (left) and decay branching fractions	9
208 2.5	ATLAS Higgs combination results for all SM measurement channels as ratios of the measured to SM production cross-sections (left) and extracted Higgs coupling constraint scale-factors for a combined fit to the measurement channels, where the W,Z, top-quark, b-quark, and τ couplings are allowed to float. The p-value of this particular model is 0.13 and in agreement with SM expectations	10
213 2.6	RGE for the running of the SM parameter, λ for the Higgs self-coupling term with present values and uncertainty bands for M_H and M_t (left). The two-dimensional plot colored (right) shows regions for which the SM is stable, unstable, and metastable based on this RGE	12
217 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.	15
221 3.2	Diagram of the ATLAS detector and subsystems	18
222 3.3	Diagram of the ATLAS ID in the $R - \phi$ plane showing the barrel view of the Pixel, SCT, and TRT detectors.	20
224 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.	20
226 3.5	Diagram of the ATLAS calorimeters	23
227 3.6	Diagram of the ATLAS LAr EM calorimeter showing the longitudinal segmentation and the $\eta - \phi$ cells for the central barrel region	24

229	3.7	Diagram of the ATLAS muon system	26
230	3.8	$R - \phi$ schematic of the ATLAS detector and various particle signatures	31
231	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.	40
234	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	40
237	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.	55
243	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	57
245	8.3	$W^\pm Z + b$ validation region: NJet, NElec, BJet MV1 and Mass Variables	58
246	8.4	Jet-inclusive 4-lepton ZZ validation region using the $t\bar{t}H$ lepton identification and momentum cuts	60
248	8.5	$ZZ + b$ validation region using the $t\bar{t}H$ lepton identification and momentum cuts	60
249	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.	65
256	8.7	Closure test on simulated $Z \rightarrow e^+e^- + jets$ events (a) and data (b). The black circles show the distribution of same-sign events while the blue histograms show the distribution of the 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.	66
261	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.	67
264	8.9	Ratios of regions with tight and anti-tight leptons in 2-lepton and 3-lepton channels	69
265	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 $t\bar{t}MC$ (red) is used for reference and not used in the actual calculation	71
268	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 $t\bar{t}MC$ (red) is used for reference and not used in the actual calculation	72

272	8.12 Distributions of the properties of the anti-tight muons in data (dots), compared 273 with the total simulation (red line), rescaled to the integral of the data for a shape 274 comparison. The uncertainty on the data distribution is statistical. The number of 275 events for each of them is also presented in the legend. The variables probed are, 276 top: p_T and $\Delta R(\mu, \text{closest selected jet})$; bottom: $\text{ptcone20}/p_T$ and $\text{Etcone20}/p_T$. 277 The selection is the signal region event selection with one anti-tight muon (failing 278 at least one of the isolation, muon-jet overlap, or p_T selection criteria) A ratio 279 plot is containing the 20% area around 1, is displayed, demonstrating that a 20% 280 data-MC comparison systematic is sufficient.	74
281	8.13 Distributions of anti-tight electron variables. The variables presented are, from 282 top left to bottom right, p_T , η , Very Tight Likelihood word value, $\text{ptcone20}/p_T$, 283 $\text{Etcone20}/p_T$. The plotted regions have the same cuts as the signal region, except 284 the anti-tight electron must fail isolation for the plot of the verytight identification 285 word or fail the verytight identification word for the plots of the isolation. Data 286 (dots) are compared with a stacked histogram of the various simulated samples: 287 top in red, V+jets (blue), VV (purple) and ttV (yellow). The uncertainty on the 288 data distribution is statistical.	75
289	8.14 Muon-xxp (above) and electron-xxd (below) anti-tight control regions: jet vari- 290 ables. Note: the $t\bar{t}$ and top MC in the plots is used only as comparison, but is not 291 included in the fake measurement	76
292	8.15 3ℓ fake validation regions for nominal p_T selection (above) and relaxed p_T selection, 293 > 10 GeV, (below). Plotted are the number of jets and the number of electrons 294 in each event. The data and MC ratio below each plot agree with 1 within the 295 statistics of the region and the overall systematic assigned for the fake component 296 (red)	77
297	9.1 Effects on the jet multiplicities in 2 SS lepton $t\bar{t}H$ events from different choices of 298 the factorization and renormalization scales. “Static” refers to the variations by 299 a factor of 2 of the nominal μ_0 , while “Dynamic” refers to the alternative choice 300 of μ_0 which depends on the event kinematic. The grey band in the lower panels 301 represents the statistical uncertainty of the nominal sample.	81
302	9.2 PDF systematic uncertainty on the jet multiplicities in $t\bar{t}H$ events with at least 2 303 leptons. The dashed red lines in the lower panel indicate the systematic uncertainty 304 on the $t\bar{t}H$ production cross section.	83
305	9.3 Muon (left) and electron(right) identification efficiencies in Data and MC as a 306 function of η and p_T respectively. For electrons, the verytight likelihood operating 307 point is used and for muons the CB+ST (combined+segment tagged) operating 308 point is used	84
309	9.4 Muon (left) and electron(right) isolation efficiency scale-factors from the Z control 310 sample as a function of the number of jets in the event. An additional systematic 311 uncertainty of 1% is added to encompass the variation in the number of jets variable	86
312	9.5 JES systematic uncertainties as a function of jet η (for jets $p_T > 25$ GeV) and p_T (for 313 jets $ \eta < 0.4$). The combined systematic uncertainty is shown with contributions 314 from the largest sources	87
315	9.6 b-Tagging data-MC efficiency scale-factors versus jet p_T calculated in the $t\bar{t}$ sample 316 from 2012 data. The uncertainties are combined statistical and systematic.	88

317

Preface

318 This is the preface. Blah blah blah blah blah blah. Blah blah blah blah blah blah.

319 Blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah

320 blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah

321 blah blah blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah blah

322 blah. Blah blah blah blah blah. Blah blah blah blah blah. Blah blah blah blah. Blah

323 blah blah blah blah blah. Blah blah blah blah blah. Blah blah blah blah blah. Blah

324 blah.

325 Blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah blah blah

326 blah blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah blah blah.

327 Blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah

328 blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah blah blah. Blah

329 blah blah blah blah blah. Blah blah blah blah blah blah. Blah blah blah blah blah

330 blah. Blah blah blah blah blah blah.

Chris Lester

331

CERN, Fall 2014

332

CHAPTER 1

333

Introduction

334

CHAPTER 2

335

Theoretical Background

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

344 **2.1 The Standard Model**

345 **2.1.1 The Standard Model Structure**

346 The Standard Model (SM) [1, 2, 3, 4] is an example of a quantum field theory that describes
347 the interactions of all of the known fundamental particles. Particles are understood to be
348 excitations of the more fundamental object of the theory, the field. The dynamics and inter-
349 actions of the fields are derived from the Standard Model Lagrangian, which is constructed
350 to be symmetric under transformations of the group $SU(3) \times SU(2) \times U(1)$. $SU(3)$ is the

351 group for the color, $SU(2)$ is the group for weak iso spin, and $U(1)$ is the group for weak
352 hyper-charge.

353 Gauging the symmetries demands the introduction of 8 massless gluons, or the boson (full
354 integer spin) carriers of the strong force [5] from the generators $SU(3)$ symmetry, and the
355 4 massless bosons, carriers for the weak and electromagnetic forces from the 3 generators of
356 the $SU(2)$ and 1 generator of the $U(1)$ group. The weak and the electromagnetic forces are
357 considered part of a larger single unified electroweak group $SU(2) \times U(1)$ and the associated
358 generators mix.

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

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

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

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

372 2.1.2 Electroweak Symmetry Breaking and the Higgs

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

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

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

395 **2.1.3 The Standard Model Parameters**

396 Although the SM structure is set by specifying the gauge groups, the EWSB mechanism, and
397 acknowledging the 3-fold replication of the fermions, confronting the SM with experiment
398 requires the measurement of 17^2 free parameters, which are unconstrained from the theory.
399 These free parameters include the fermion masses (from the Yukawa coupling), the force
400 coupling constants, the angles and phase of the mixing between quarks, and constants from
401 the Higgs and electroweak sector³.

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

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

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

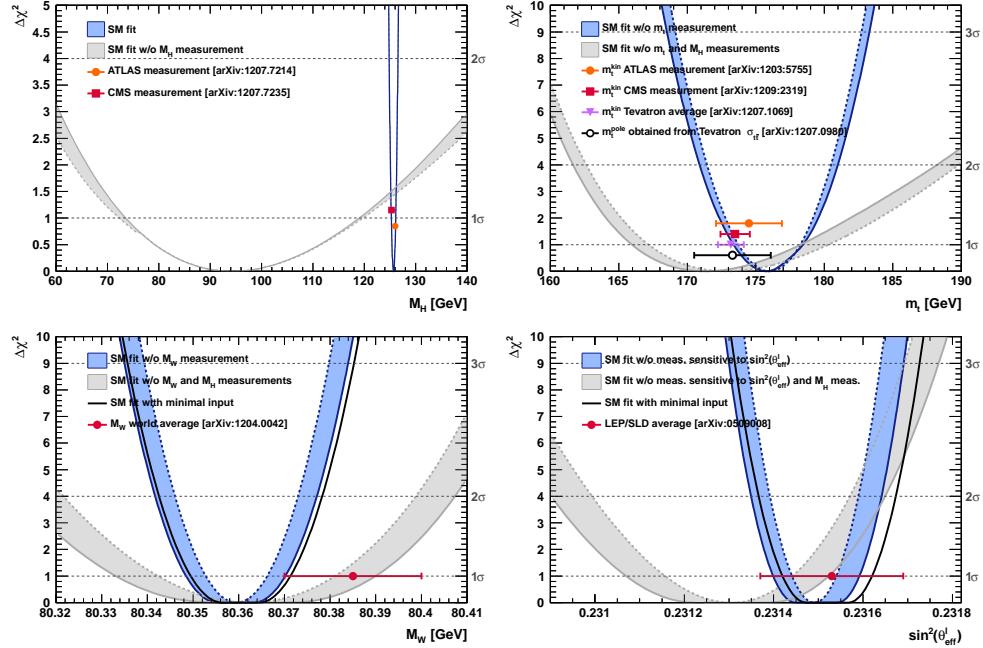


Figure 2.1: χ^2 as a function of the Higgs mass (top left), the top quark mass (top right), the W boson mass (bottom left) and the effective weak mixing angle (bottom right) for the combined SM fit from the GFitter group. The data points placed along $\chi^2 = 1$ represent direct measurements of the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

414 2.2 Collider Physics and the Higgs

415 To test the theory, physicists accelerate particles to extremely high energies and force them
 416 to interact through collisions. Typically, the particles accelerated are electrons or protons,
 417 since they are stable. Electron-positron collider machines have a rich history of discovery and
 418 measurement in particle physics. The advantage of electron accelerators is that the colliding
 419 element is itself a fundamental particle. However, due to synchrotron radiation, curvature of the
 420 beam line becomes problematic for high energy beams. Hadron colliders, specifically, proton-
 421 proton and proton-anti-proton colliders can be accelerated in rings without large losses due
 422 to synchrotron radiation, but the actual colliding objects at high energies are the constituent
 423 quarks and gluons. This complicates analysis because the initial state of the system is not

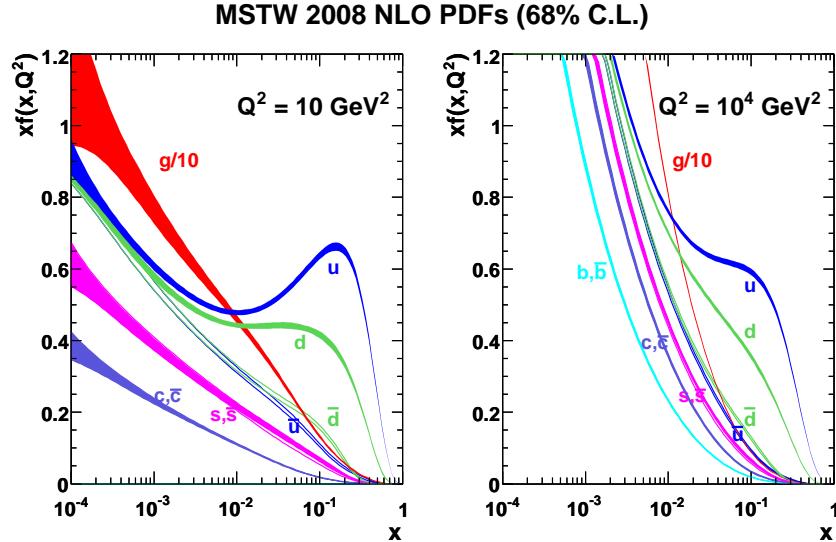


Figure 2.2: Proton Parton Distribution Functions (PDFs) form the MSTW Collaboration at $Q^2 = 10 \text{ GeV}^2$ and $Q^2 = 10^4 \text{ GeV}^2$

424 discernible on a per-collision basis and momentum of hard scatter system is unknown along
 425 the beam direction.

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

430 At the Large Hadron Collider or LHC (discussed in detail in Chapter 3 a proton-proton
 431 colluder, the types of initial states are quark-quark, quark-gluon, and gluon-gluon. Gluon
 432 collisions dominate overall, due to the large number of gluons inside the proton, though the
 433 relative importance of different initial states changes with the energy scale of the collision and
 434 the type of final state sought after.

435 A prime motivation for the construction of the Large Hadron Collider was the discovery
 436 or exclusion of the Higgs boson[13]. LEP and the Tevatron excluded large swaths of possible

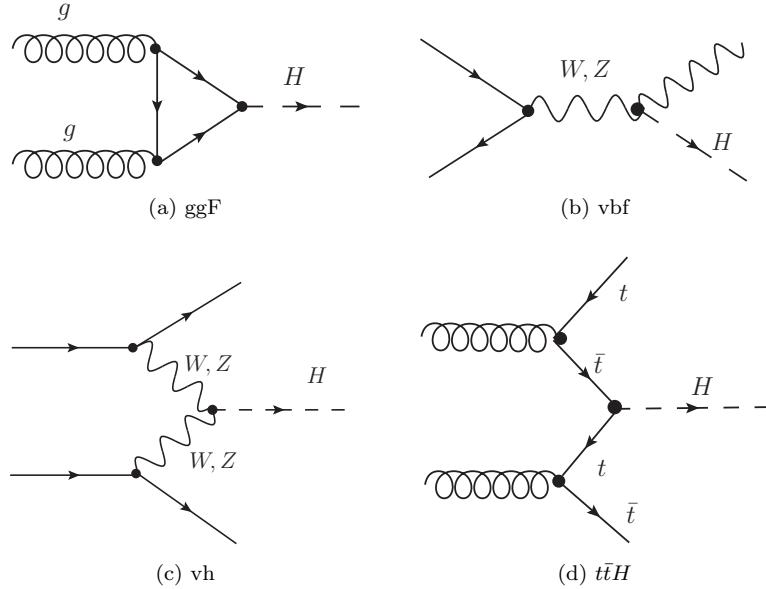


Figure 2.3: Dominant Higgs production modes at the LHC

437 Higgs boson masses, especially below 114 GeV and the unitarity of certain diagrams including
 438 the $WWWW$ vertex required the mass to be below about 1 TeV, a range that was achievable
 439 at the LHC with high luminosity running [10].

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

444 The Higgs boson can be produced through collision at the LHC via 4 mechanisms: gluon-
 445 fusion (ggF), vector-boson fusion (VBF), Higgsstrahlung (VH), and production in association
 446 with top quarks ($t\bar{t}H$). The diagrams are shown in Figure 2.3 and the production cross-
 447 sections as a function of Higgs mass for the 8 TeV LHC proton-proton running are shown in
 448 Figure 2.4 [14]. The largest production cross-section is via the gluon fusion channel at 20
 449 pb, which proceeds through a fermion loop that is dominated by the top quark, because of

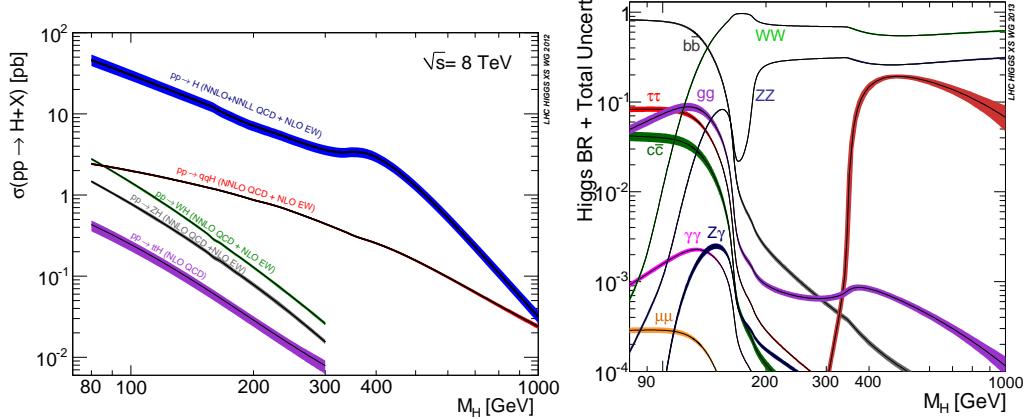


Figure 2.4: 8 TeV LHC Higgs production cross-sections (left) and decay branching fractions

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

456 2.2.1 Higgs Discovery at the LHC

457 In 2012 both ATLAS and CMS announced the discovery of a new boson consistent with
 458 the Higgs by examining the results of Higgs searches in a number of decay channels ($H \rightarrow$
 459 W^+W^- , $H \rightarrow Z^0Z^0$, and $H \rightarrow \gamma\gamma$) in the 2011 dataset at $\sqrt{s} = 7$ TeV and part of the 2012
 460 dataset at $\sqrt{s} = 8$ TeV. By 2013 and 2014, both experiments have updated and/or finalized
 461 their results for the full 2011 and 2012 datasets [15, 16]. I will focus on the ATLAS results in
 462 the following. The ATLAS measurements constrain both the Higgs mass[17] and spin[18], as
 463 well as provide constraints to the Higgs couplings to different particles.

464 Figure 2.5 show the results of the searches in all of the measurement channels as well as

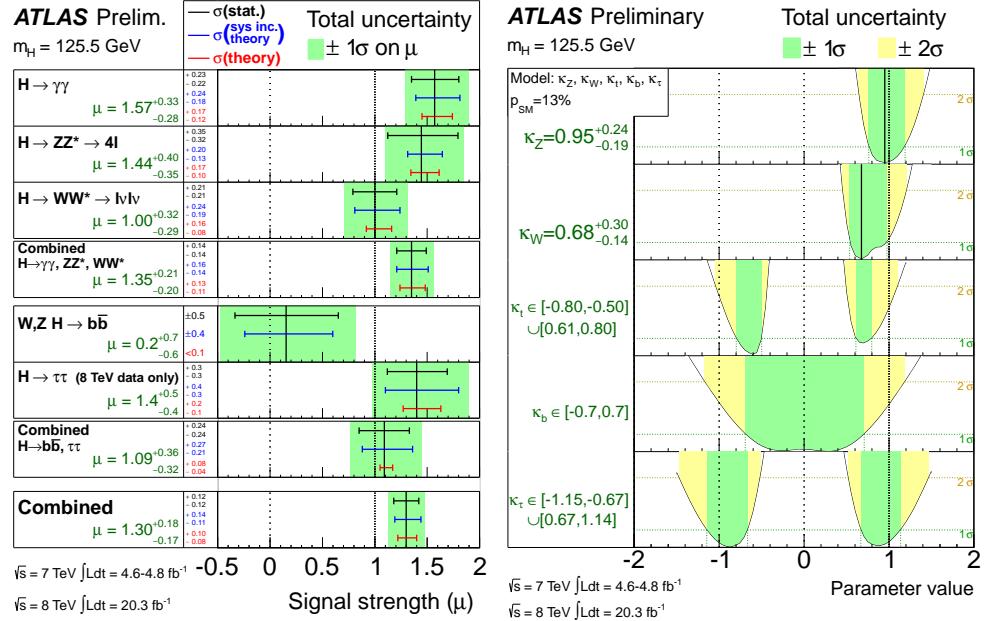


Figure 2.5: ATLAS Higgs combination results for all SM measurement channels as ratios of the measured to SM production cross-sections (left) and extracted Higgs coupling constraint scale-factors for a combined fit to the measurement channels, where the W,Z, top-quark, b-quark, and τ couplings are allowed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

465 constraints on the SM Higgs coupling parameters in an example fit, where the couplings to
 466 the top-quark, bottom-quark, W,Z, and τ are allowed to fluctuate independently. These rely
 467 on measurements binned in different production and decay channels. They are dominated by
 468 higher statistics results in the gluon-fusion production modes, but measurements in the VH
 469 and VBF modes are close to SM sensitivity.

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

474 2.2.2 $t\bar{t}H$ Production

475 Notably absent thus far in the SM are searches for the Higgs in the $t\bar{t}H$ production channel, due
476 to the lack of statistics. Searches are underway and initial results are close to SM sensitivity
477 for ATLAS and CMS. More will be discussed about particular searches later. $t\bar{t}H$ production
478 depends on the top Yukawa coupling at tree level. Comparison of the gluon-fusion and the $t\bar{t}H$
479 modes would allow for disentangling the effects of new particles in the gluon-fusion loop[19].
480 For instance, generic models would allow for the introduction of new colored particles into
481 the loop. The simplest of these models would be the addition of a new generation of quarks.
482 Fourth generation quarks, which obtain mass from a Higgs Yukawa coupling are already
483 largely excluded due to their enormous effects on the Higgs production cross-section[20]. Other
484 exotic scenarios allow for new colored particles in the gluon-fusion loop, which are not entirely
485 constrained by present measurements[21, 22, 23]. These include, for instance, Supersymmetric
486 models involving the stop quark.

487 Aside from the loop effects, measurement of the $t\bar{t}H$ production cross-section would provide
488 a precise measurement of the top Yukawa coupling. When compared with the measured top
489 quark mass, this tests the Higgs mass generation properties for up-type quarks. Despite
490 similar uncertainties on the overall production cross-sections for $t\bar{t}H$ and the gluon-fusion
491 modes, both of which depend on the top Yukawa, most of these uncertainties would cancel for
492 $t\bar{t}H$ if normalized to the topologically similar $t\bar{t}Z$. Finally, the uniqueness of the experimental
493 signature means that searches for $t\bar{t}$ signatures can be performed for a variety of Higgs decays
494 ($\gamma\gamma$, $b\bar{b}$, WW, ZZ , and $\tau\bar{\tau}$ with roughly similar degrees of sensitivity (within a factor of 10)[19].

495 It is important to note the importance of the top Yukawa coupling to the overall structure
496 of the SM, due to its enormous size compared to other couplings. The top Yukawa is for
497 instance 350000x as large as the electron Yukawa coupling. Because of this the top Yukawa

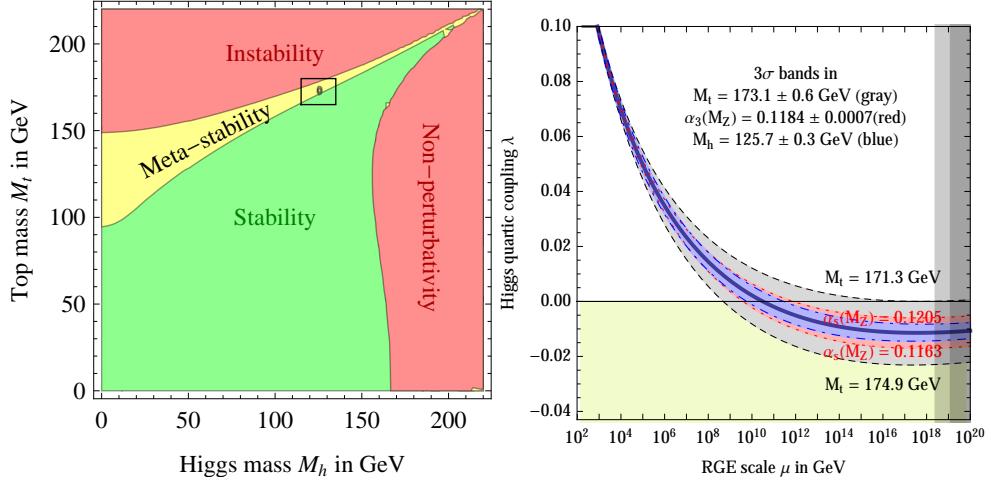


Figure 2.6: RGE for the running of the SM parameter, λ for the Higgs self-coupling term with present values and uncertainty bands for M_H and M_t (left). The two-dimensional plot colored (right) shows regions for which the SM is stable, unstable, and metastable based on this RGE

498 coupling, along with the Higgs mass, is one of the most important pieces of the renormalization
 499 group equations (RGE) responsible for the running of the parameter that determines the
 500 Higgs self-coupling λ . If this parameter runs negative, then the potential responsible for the
 501 entire mechanism of EWSB no longer has a minimum and becomes unbounded, resulting in
 502 instability in the universe [24]. Metastability occurs when the shape of the potential allows
 503 for a false local minimum. Figure 2.6 shows the running of this parameter, the regions for
 504 which the universe is stable, unstable and metastable. Current measurements suggest that
 505 universe lies in a metastable island⁴.

506 2.3 Conclusion

507 The Standard Model, despite its success in providing a unified description of fundamental
 508 particles and interactions into single theory, has its flaws. These have been discussed in depth
 509 elsewhere, but include issues like the description of massive neutrinos, the failure to include

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

510 gravity, and the unnaturalness of large quantum corrections to Higgs parameters. For these
511 reasons, it seems the SM might be a lower energy approximation to a more fundamental
512 theory. The discovery of the Higgs boson at the LHC provided a stunning verification of one
513 of the fundamental aspects of the theory but at the same time offers new area to search for
514 glimpses of something grander. The production of samples of Higgs bosons allows for a rich
515 array of new tests of the Standard Model, which is now finally over-constrained by experiment.
516 Searches for the $t\bar{t}H$ production, one category of which is the topic of this thesis, provide tree-
517 level access to a central parameter of the theory, the top Yukawa coupling, as well as access
518 a variety of Higgs decays, which will eventually provide a rigorous new test of the SM.

519

CHAPTER 3

520

521

The Large Hadron Collider and the ATLAS Experiment

522

3.1 The Large Hadron Collider

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

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

529 The technology involved in the development of the LHC and very briefly touched upon
530 in this chapter is an enormous achievement in its own right and is documented in detail here
531 [25, 26, 27].

532 The LHC is a circular machine capable of accelerating beams of protons and colliding
533 them at center of mass energies up to $\sqrt{s} = 14\text{TeV}$ at 4 collision sites around the ring, where
534 4 experiments are housed (ATLAS[28], CMS[29], LHCb[30], and ALICE[31]. Figure 3.1 is
535 a diagram of the layout of the LHC and its experiments[32]. The LHC also operates in a
536 modes with beams of heavy ions. The LHC is composed of thousands of super-conducting

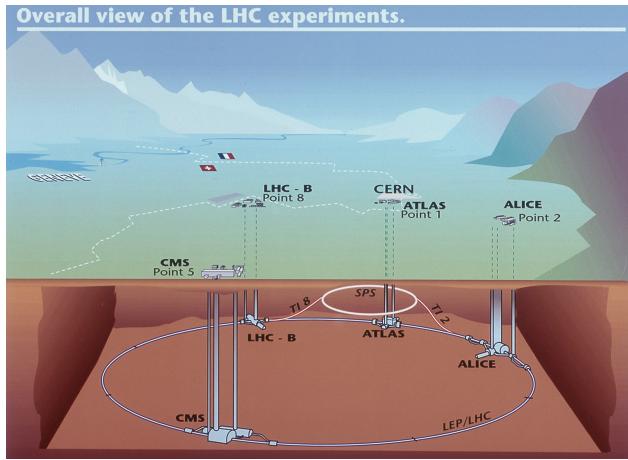


Figure 3.1: Diagram of the Large Hadron collider and location of the 4 main experiments (ATLAS, CMS, LHCb, and ALICE). Around the ring. The diagram also shows the location of the SPS, the final booster ring the accelerator complex that accelerates the protons to 450 GeV before injection into the LHC.

537 Niobium-Titanium magnets, cooled to 2.7° C with liquid Helium, which steer and focus the
 538 particle beams, and a superconducting resonant-frequency (RF) cavity, which boosts the beam
 539 to higher energies.

540 3.1.1 The Accelerator Complex

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

544 accelerated to energies of 1.4 GeV for injection in to the Proton-Synchrotron (PS). The PS
545 accelerates the protons to 25 GeV and dumps bunches into the Super Proton Synchrotron
546 (SPS), where they are accelerated to 450 GeV and finally dumped into the LHC.

547 **3.1.2 Beam Parameters and Collisions**

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

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

563 **3.2 The ATLAS Experiment**

564 This section provides a brief overview of the ATLAS experiment. The ATLAS detector is
565 centered on one of the LHC collisions points, located 100m underground. Through the com-

566 bination of a number of subsystems, it designed to identify the particle content arising from
567 these collisions, measure energy and momentum of of these particles, and make fast decisions
568 about the content of each collision, in order to save a small fraction of measured collision
569 events for offline study.

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

- 576 • **Inner Detector:** The inner detector (ID)[33, 34] is immersed in a solenoidal magnetic
577 field[35] and provides measurements of charge particle tracks, through three subsystems:
578 the Pixel Detector[36, 37], the Semi-Conductor Tracker (SCT)[38, 39], and Transition
579 Radiation Tracker [40, 41, 42]. and is capable of making particle
- 580 • **Calorimeter:** The calorimeters measures the energy of particles that participate in the
581 electromagnetic (photons, electrons) and hadronic interactions (pions, protons, neutrons,
582 etc.). The hermeticity of the calorimeters allows for missing transverse energy mea-
583 surements. The calorimeter is composed of the liquid argon electromagnetic calorime-
584 ter (LAr)[43] and the hadronic tile calorimeter[44], the liquid argon hadronic endcap
585 calorimeter, and the forward calorimeters
- 586 • **Muon Systems** The muon systems(MS)[45] form the outermost detector systems and
587 measure the momentum of minimum ionizing muon tracks, as all other particles are
588 stopped by the calorimeters. The muon systems are immersed in a toroidal magnetic

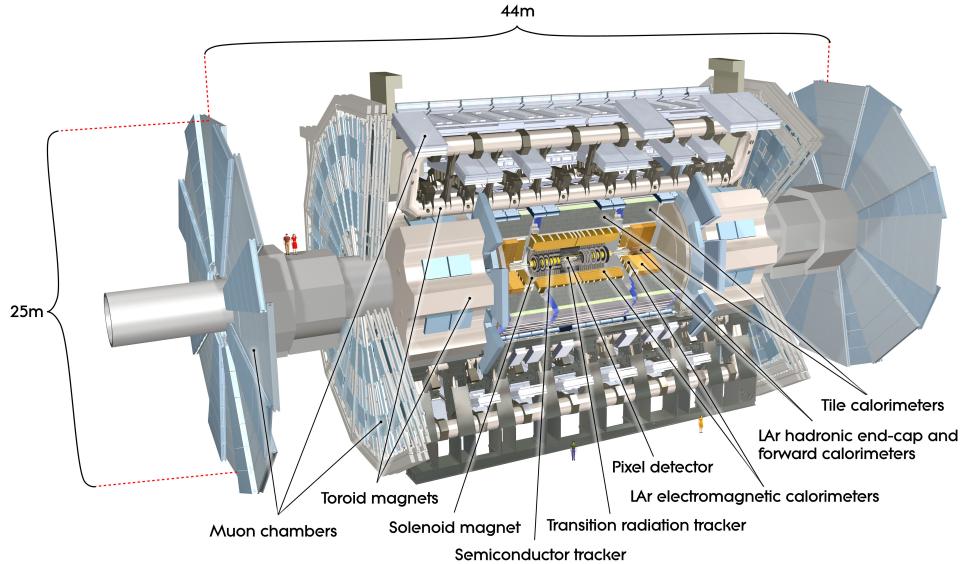


Figure 3.2: Diagram of the ATLAS detector and subsystems

589 field [35] and are composed of 4 different sub-systems for triggering and tracking mea-
 590 surements [46, 47, 48].

- 591 • **Triggering Systems:** The trigger and data acquisition systems[49, 50] read out data from
 592 the detector through a three-tiered hardware and software decision making framework
 593 to record the most interesting physical processes for a broad physics analysis program.

594 These systems are discussed in depth in the following sections.

595 3.2.1 Detector Coordinate System

596 ATLAS uses a right-handed coordinate system centered at the nominal proton interaction
 597 point. The beam line defines the z -axis, and the $x - y$ plane is perpendicular to the beam line.
 598 The $x - y$ plane is referred to as the transverse plane. The transverse plane holds special
 599 significance in reporting measurements, because the initial momentum of the hard collision
 600 system is 0 along the transverse plane in the laboratory rest frame. The momentum of the

601 colliding proton-proton system is also 0 along the z -axis but the colliding partons may have
602 vastly different momenta. Thus, momentum of the hard colliding system along the z -axis
603 differs collision to collision. Particle momenta measured along the transverse plane is called
604 transverse momenta, and labeled p_T .

605 Because ATLAS possesses nominal cylindrical symmetry, cylindrical and polar coordinates
606 are used to describe particle trajectories and detector positions. The radial coordinate, r ,
607 describes transverse distances from the beam line. An azimuthal angle, ϕ , describes angles
608 around the z -axis, and a polar coordinate θ describes angles away from the z -axis. The
609 polar angle is often expressed in terms of pseudo-rapidity, defined as $\eta = -\ln(\tan(\theta/2))$.
610 Distances in $\eta - \phi$ space are often used to describe the proximity of objects in the detector,
611 $\Delta R = \sqrt{\eta^2 + \phi^2}$.

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

617 3.2.2 The Inner Detector

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

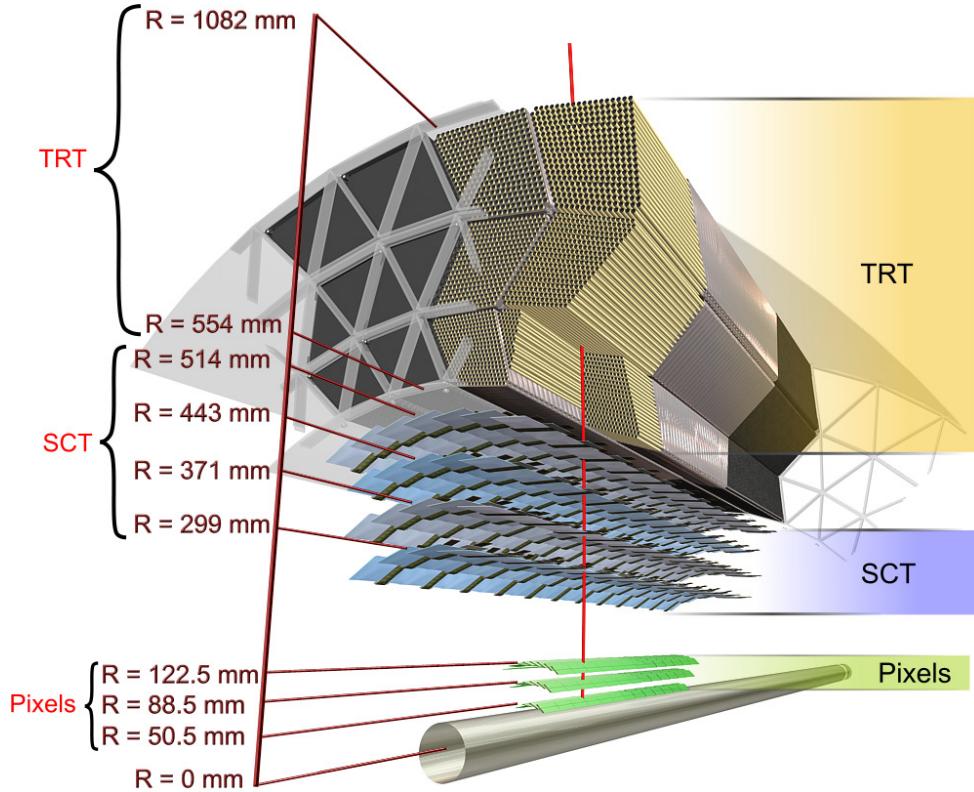


Figure 3.3: Diagram of the ATLAS ID in the $R - \phi$ plane showing the barrel view of the Pixel, SCT, and TRT detectors.

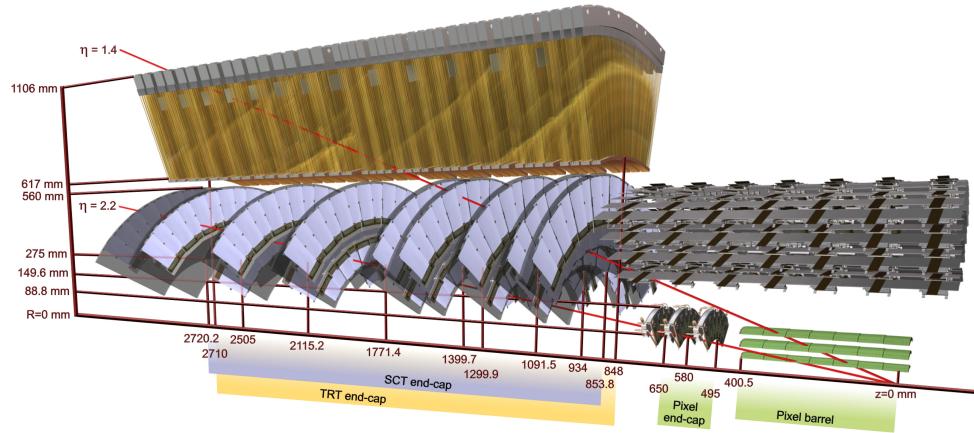


Figure 3.4: Diagram of the ATLAS ID in the $R - z$ plane showing the endcap view of the Pixel, SCT, and TRT detectors. Only one side of the endcap is shown.

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

632 The SCT uses similar silicon technology to the pixels, but the each SCT layer contains a
633 double layer of silicon strips, which are much longer in length than width. The double layers
634 are inclined slightly with respect to each other so that these 1D sensors have 2D resolution for
635 coincident hits of $580 \mu\text{m}$ in $z(R)$ for the barrel(endcap) and $17 \mu\text{m}$ in $R - \phi$. The SCT has
636 4 million read out channels and is comprises 4 barrel layers and 9 endcap layers with coverage
637 to $|\eta| < 2.5$.

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

646 The TRT is unique in that it also provides particle identification measurements via trans-
647 sition radiation. Charged particles emit transition radiation, when traversing a boundary
648 between materials of different dielectric constants. The volume between the straws is, there-
649 fore filled with a radiator material, a polymer foil or foam, to provide this boundary condition.

650 Transition radiation photons are emitted in the direction of the particle trajectory in the keV
651 range and cause a much larger signal amplitude within the straw. Hits that cause a signal
652 at a higher threshold are thus indicative of transition radiation. The probability for emission
653 transition radiation depends on the relativistic γ of the traversing particle. Because electrons
654 are much lighter than any other charged particle, their γ -factors tend to be high enough in the
655 GeV range to induce transition radiations, as opposed to pions, muons and other particles.

656 Combined tracking of particles through the 3 sub-detectors results track momentum mea-
657 surements from 500 MeV, the minimum energy need to leave the ID due to the magnetic field,
658 and a few TeV. The track p_T resolution is roughly $0.05\% \cdot p_T + 1\%$.

659 **3.2.3 The Calorimeter**

660 The ATLAS calorimeters measure the energy of electron, photons and hadrons with $|\eta| < 4.5$.
661 They induce a particle shower via electro-magnetic and nuclear interactions with the detec-
662 tor material and contain enough radiation lengths to ensure that all or most of the shower
663 energy remains contained. Muons are minimum ionizing particle that do not participate in
664 the strong interaction and therefore pass through the ATLAS calorimeters leaving relatively
665 little energy behind. ATLAS calorimeters are sampling calorimeters meaning that the active
666 material of the detector only measures a small fraction of the energy produced by the shower
667 and the overall shower energy is inferred from this fractional measurement. The rest of the
668 material is inactive, heavy dense material, designed to induce showers. The calorimetry system
669 is grossly divided longitudinally (radially) into electro-magnetic(EM) and then hadronic seg-
670 ments, operated with different technologies. Figure 3.5 diagrams the layout of the calorimeter
671 system.

672 The EM calorimeter, which is located directly outside of the solenoid magnet but within
673 the same cryostat, has an accordion design with lead absorber and liquid argon (LAr) active

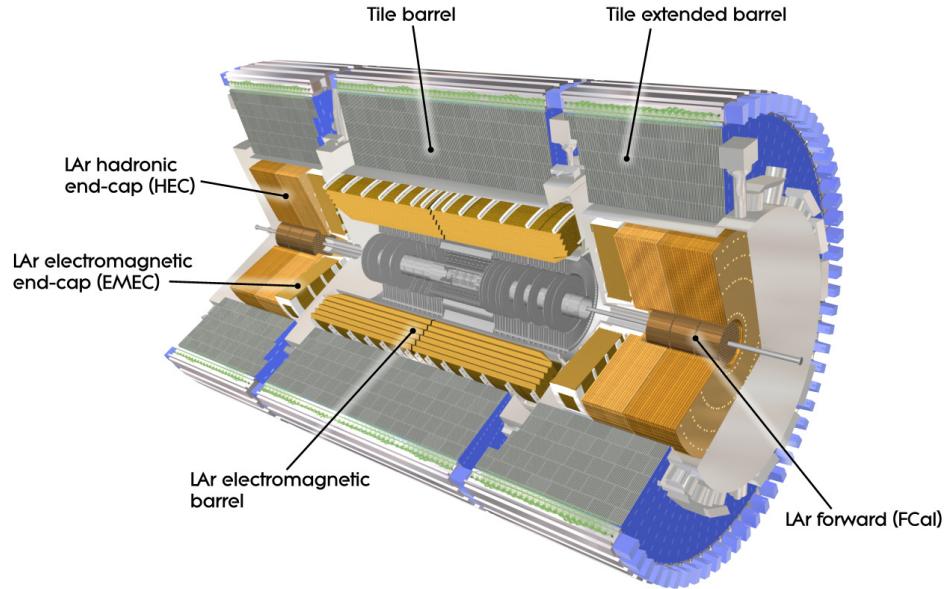


Figure 3.5: Diagram of the ATLAS calorimeters

material. The accordion design ensures uniform coverage in ϕ . The barrel and endcap LAr extend to $|\eta| < 2.47$. The LAr provides highly granular measurements in $\eta - \phi$ with 4 longitudinal segments, totaling $\sim 25\text{-}35$ radiation lengths with the exception of the barrel/endcap transition region ($1.37 < |\eta| < 1.52$). The geometry of the barrel LAr calorimeter can be seen in Figure ???. The first longitudinal segment is called the pre-sampler, composed of a thin layer of active liquid argon, designed to detect early particle showers. The second segment is the most highly granular segment called the 'strips', as it is composed of thin strips. The strips have a size of $0.025/8 \times 0.1$ in $\eta - \phi$ in the barrel with similar sizes in the endcap 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 signature. 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

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

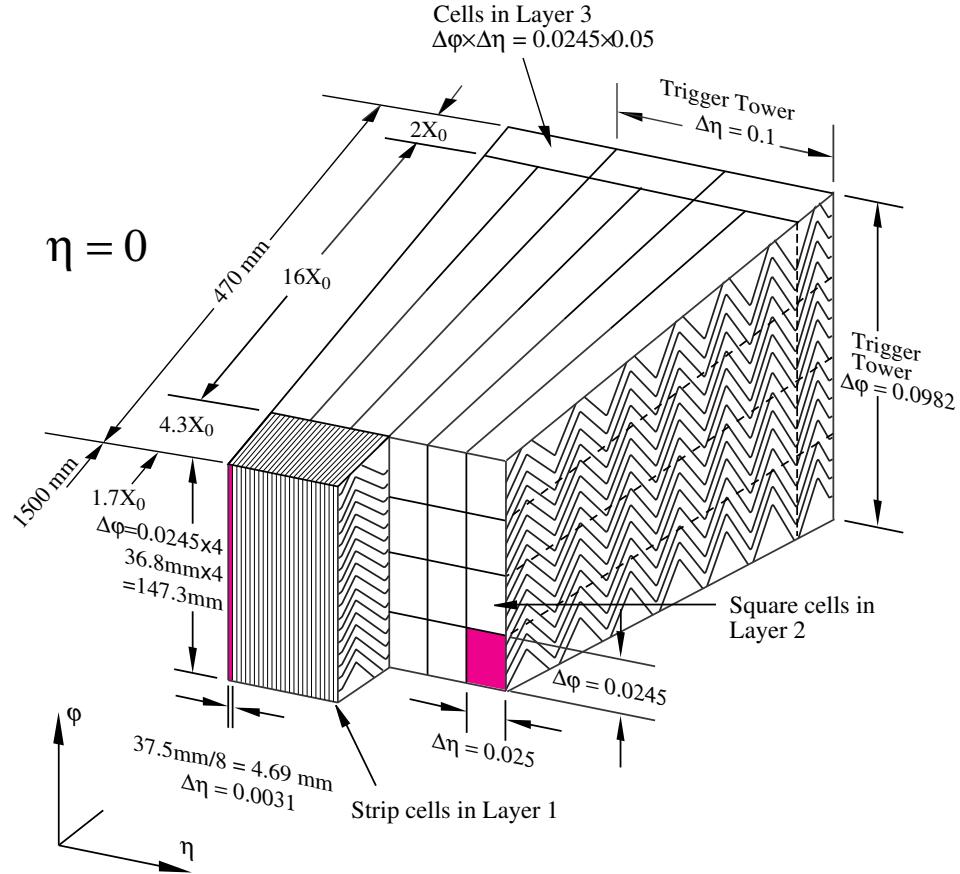


Figure 3.6: Diagram of the ATLAS LAr EM calorimeter showing the longitudinal segmentation and the $\eta - \phi$ cells for the central barrel region

686 designed to estimate energy leaking out of the EM calorimeter. The forward EM calorimeters
 687 extend the η range and use the same technology, but are not used in this analysis. The energy
 688 resolution of the EM calorimeters is $\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$, measured in test beam data
 689 and confirmed in collision data.

690 The hadronic calorimeter is located directly behind the EM calorimeter and composed of
 691 tiles of iron absorber and plastic scintillator in the barrel ($|\eta| < 1.6$), called the TileCal, and
 692 copper-liquid argon in the endcap ($1.5 < |\eta| < 3.2$), called the HEC. The calorimeters contain
 693 $\sim 10\text{-}19$ hadronic interaction lengths with multiple longitudinal segments to contain showers

694 induced by the nuclear interaction of hadronic particles. The energy resolution of the hadronic
695 calorimeters is $\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$. The intrinsic resolution of hadron calorimeters is
696 much worse than electro-magnetic calorimeters, because much of the inelasticity of the nuclear
697 interactions.

698 **3.2.4 The Muon Spectrometer**

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

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

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

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

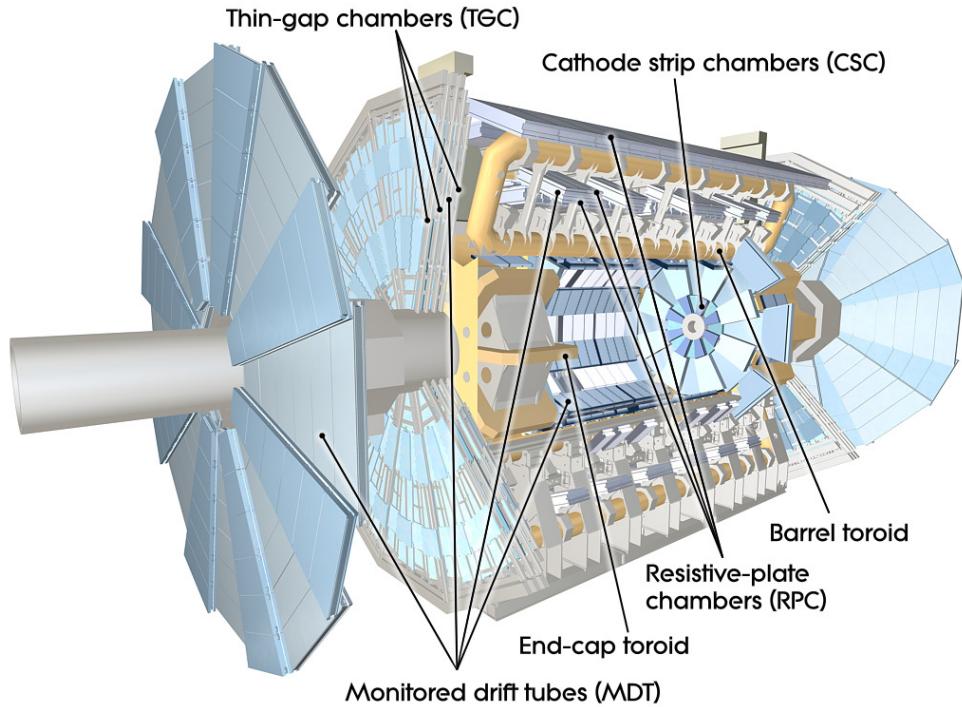


Figure 3.7: Diagram of the ATLAS muon system

⁷¹⁸ too slow for the level-1 trigger.

⁷¹⁹ 3.2.5 The Trigger System

⁷²⁰ The ATLAS trigger system is designed to make quick decisions about individual particle
⁷²¹ collisions to reduce the enormous collision rate of 20 MHz to a much more manageable 400
⁷²² Hz to be stored for offline analysis. Saving the full ATLAS data-stream would require space
⁷²³ for 40 TB of raw data per second, but, more importantly, most of these collisions result in
⁷²⁴ the uninteresting inelastic break-up of the colliding protons. To select out collisions to allow
⁷²⁵ for a diverse physics program, ATLAS devotes a large portion of the bandwidth to general
⁷²⁶ purpose single lepton triggers. The presence of leptons in the event indicates the presence
⁷²⁷ of the weak or electro-magnetic interaction and therefore occurs at many order of magnitude

728 less frequently than interactions involving the strong interaction. Moreover, many interesting
729 physics signatures that are analyzable by ATLAS involve leptonic final states. The remaining
730 bandwidth is allocated to jet, missing energy, tau, and unbiased supporting triggers.

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

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

745 **3.2.6 Reconstruction: Jets, Muons and Electrons**

746 Physicists analyze the collision event as a collection of identified objects, expressed as mo-
747 mentum 4-vectors. These objects arise from the final state particles in the event, which can
748 be combined and counted to infer properties of the hard scatter. The particles that make
749 detectable signatures are those that are stable enough to pass through the detector: muons,
750 electrons, photons, and quarks and gluons. Jets and b-tagged jets, muons and electrons, are
751 used in the $t\bar{t}H$ analysis to define our search regions and to separate the Higgs signal from

752 backgrounds. Other analyses may used photons, taus and missing energy⁶, but these are not
753 discussed in depth here. Figure 3.8 shows an $R - \phi$ schematic of the interaction of various
754 particle signatures in the ATLAS detector.

755 **3.2.6.1 Tracks and Clusters**

756 The process of converting the disparate detector signatures and signals into a unified 4-
757 momentum description of individual objects is called Reconstruction. The basic components
758 of reconstruction are sensor measurements, or hits, in trackers (ID, MS) and energy measure-
759 ments in the calorimeter. Hits in the ID and MS undergo pattern recognition, which identifies
760 hits that belong to a single track, and fitting, which fits a curve to the track to assess the par-
761 ticle trajectory. Charged particle trajectories are generally helical in a magnetic field, but the
762 fitting algorithm takes into more detailed information about energy loss to material along the
763 tracks length. The result of the fitting is an estimation of the particles momentum 3-vector.
764 ATLAS track reconstruction is discussed in depth here []. Electron, photon, and hadronic
765 particles leave clustered deposits of energy in the EM and hadronic calorimeters from their
766 showers. Electron and photon showers are primarily contained with the EM calorimeter, while
767 hadronic showers are deposit most of their energy in the hadronic calorimeters. The process of
768 associating individual read-out cells of energy in the calorimeter to clusters of energy from the
769 showers of individual particles is called clustering. ATLAS clustering algorithms are discussed
770 in depth here ???. From these basic, pieces, tracks and clusters, more complex objects can be
771 created.

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

772 3.2.6.2 Electrons

773 Electrons leave both a track in the ID and a narrow and isolated cluster of energy in the
774 EM calorimeter, $\Delta R < 0.1$. Electron reconstruction proceeds using a sliding window algo-
775 rithm, which scans a fixed size rectangle in $\eta - \phi$ space over the EM calorimeter cells to find
776 relative maxima of energy in the window[51]. These maxima seed the clustering algorithms.
777 Because electrons are light, they both lose energy to the material gradually via scattering and
778 more catastrophically through the emission of a high energy photon, through interaction with
779 the nuclear. This process is called bremsstrahlung. Tracks for electrons are reconstructed
780 uniquely because they must include the hypothesis that the electron loss energy through
781 bremsstrahlung. Generally, the emitted photon is contained within the same energy cluster
782 and therefore the sliding window algorithm is always wider in the direction of bending, ϕ . A
783 single track is then matched to the cluster within certain minimum matching requirements in
784 η, ϕ , and p_T . Electrons are distinguished from photon conversions, which also have a track, by
785 association with conversion vertices, found with a dedicated algorithm.

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

791 3.2.6.3 Muons

792 Muons are reconstructed from a combination of ID and MS tracks when possible. The two
793 tracks must meet matching criteria to ensure they are from the same particle and the muon
794 momentum 3-vector comes from the combined ID,MS fit. Muons leave little energy in the

795 calorimeters and are generally isolation from other particle, when produced from electro-weak
796 bosons. Identification algorithms make requirements on the number of tracking hits in the ID
797 and MS and the quality of the matching of the two tracks. Muons are reliably reconstructed
798 and identified with energies above 5 GeV. More about muon reconstruction and identification
799 can be found here [52].

800 **3.2.6.4 Jets**

801 Quarks and gluons are colored objects that cannot exist alone. When emitted, they undergo a
802 process called hadronization, in which the convert into 'jets' of colorless hadrons that emerge
803 collimated from the interaction point. The majority of these hadrons are charged and neutral
804 pions, though other hadrons are often present. Jets are reconstructed using conglomerations
805 of calorimeter energy clusters chosen via anti- k_t algorithm, with a radius of $\Delta R < 0.4$ [53].
806 The algorithm has been shown to be infrared safe, meaning the jet quantities are not sensitive
807 to low energy, small angle radiative divergences. Jets at ATLAS are reconstructed from 10
808 GeV, calibration of the energy scale and resolution are only available for energies greater than
809 20-25 GeV.

810 **3.2.6.5 B-Tagged Jets**

811 Generally, the flavor of the initiating quark is not known from the reconstructed jet, although
812 gluon initiated jets and quark initiated jets have slightly different properties. B-quark jets,
813 however, are unique in that the long life-time of the produced b-mesons allow for measurable
814 decays in flight. This property is used to tag b-quark initiated jets. This analysis used the MV1
815 tagging algorithm [54], which is a neural network based algorithm that looks for secondary
816 displaced decay vertices inside the event and takes into account jet track parameters and
817 energy flow with respect to these vertices. B-quark jets often involve b-meson decays to

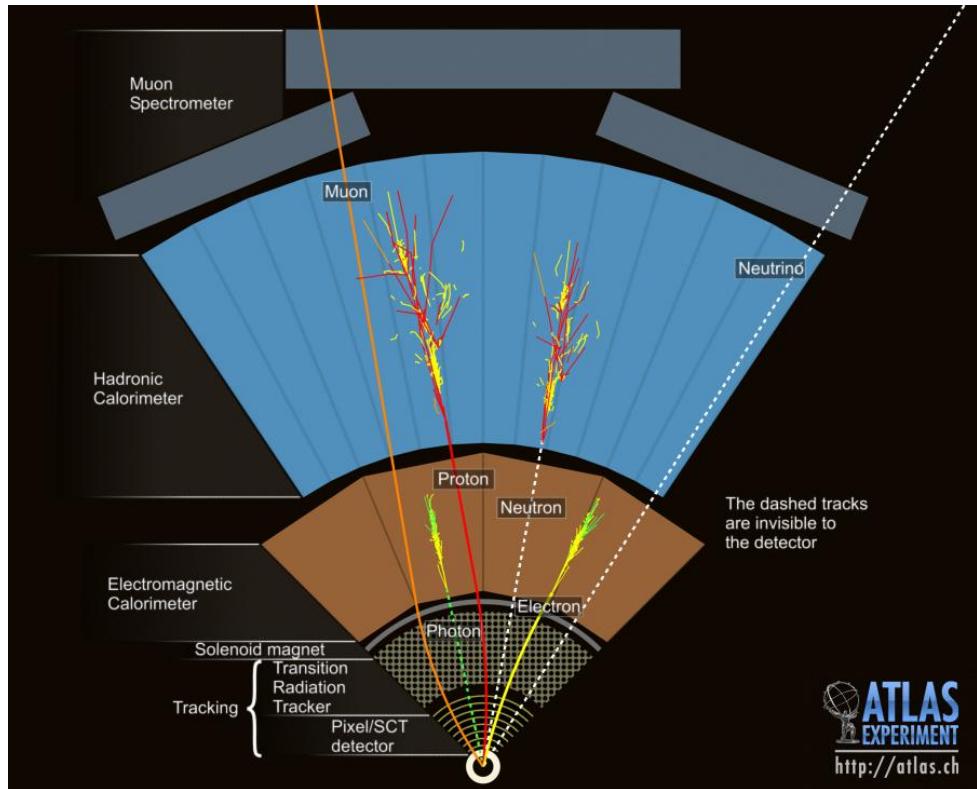


Figure 3.8: $R - \phi$ schematic of the ATLAS detector and various particle signatures

818 leptons, especially muons, which can be used to tag an orthogonal b-jet sample for studying

819 tagging efficiencies.

820

CHAPTER 4

821

Electrons

822 This chapter details the contributions I made to electron identification and efficiency mea-
823 surments. It is not essential to continuity of the thesis in general but provides in depth
824 documentation of the work I completed for the experiment. I focus on the electorn identifica-
825 tion

826 **4.1 Electrons at Hadron Colliders**

827 High energy electron signatures are important elements of seaches and measurements at
828 hadron colliders. The overwhemling majority of collisions that deposit energy in the de-
829 tectors are the result of strong-force mediate interactions of the constituent partons. These
830 collisions result in the production of high energy jets in the detector. Figure XX shows the
831 cross-sections of various processes as a function of the center of mass energy of the collision.

832 Physics invovling the electroweak interaction or even strong production of massive states occur
833 many orders of magnitdue less frequently than the total inelastic cross-section. Interesting
834 physics signatures, both stardard model and beyond, often involve the production of light
835 leptons as a result of the decay of massive particles. Choosing events that have high energy
836 electrons or muons targets events that contain electroweak vertices and dramatically reduce

837 the background from the more copiously produced strong physics. Electron and muon energy
838 and momenta are also relatively well-measured compared to jets. This allows for the use
839 of well-resolved kinematic shapes used to discriminate the signatures of different processes in
840 analyses.

841 At ATLAS, the primary datasets for most analyses are collected with electron and muon
842 triggers. Electron triggers are particularly important, because the muon triggers system has a
843 20% smaller acceptance than the electrons. The challenge in identifying electrons is distinguishing
844 the production of electrons from direct production of W and Z decays from electrons pro-
845 duced in the more copiously produced b-meson decays, fake-electron signatures from rare jet
846 fragmentations into charged and neutral pions, and photon conversions in the inner detector.
847 The identification of electrons, the precise measurement of the identification efficiency, and the
848 measurement of the rate of fake electron signatures are often the most important and
849 challenging pieces of an analysis. The following sections discuss the identification of electrons
850 for the primary electron trigger and offline physics analyses as well as the measurement of the
851 electron identification efficiency in 2012. Because I had a major role in these projects, I will
852 at times discuss their historical evolution and not simply focus on the particular measurement
853 relevant to the $t\bar{t}H$ analysis.

854 **4.2 Identification of Electrons at ATLAS**

855 Electron reconstruction

856 **4.2.1 Pile-up and Electron identification**

857 Plots of pile up differences in distributions

858 **4.2.2 2011 Menu and Trigger**

859 **4.2.3 2012 Menu and Trigger**

860 **4.2.4 Electron Likelihood**

861 **4.3 Measurement of Electron Efficiency at ATLAS**

862 **4.3.1 Techniques**

863 **4.3.2 Issues**

864

CHAPTER 5

865

$t\bar{t}H$ Analysis Summary

866 This chapter provides an overview of the set of analyses searching for the Standard Model
867 (SM) production of the Higgs boson in association with top quarks in multi-lepton final states
868 with multiple jets (including b-quark tagged jets). Searches in $t\bar{t}H$ final states with 2 same-
869 charge, 3 and 4 light leptons (e, μ) are discussed in depth. These final states target specifically
870 Higgs decays to vector bosons, $H \rightarrow W^\pm W^\pm$ and $H \rightarrow Z^\pm Z^\pm$ and form a complement to
871 searches for $t\bar{t}H$ production in final states targeting the $H \rightarrow b\bar{b}$ [55], $H \rightarrow \gamma\gamma$ [56], and
872 $H \rightarrow \tau\tau$ decay modes.

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

879 Detailed description of the event and objection section are provided in Chapter 7, back-
880 ground modeling in Chapter 8, the effect of systematic errors and the statistical analysis in
881 Chapter 9 and final results in Chapter 10.

882 **5.1 Signal Characteristics**

883 $t\bar{t}H$ can be observed in a number of different final states related to the Higgs boson and
884 the top quark decay modes.

885 Three Higgs boson decays are relevant for this analysis: W^+W^- , $\tau^+\tau^-$ and ZZ . The
886 top and anti-top quarks decay in $W^\pm b$. Each W^\pm boson decays either leptonically ($l=e^\pm$,
887 μ^\pm, τ^\pm) with missing energy or hadronically. Table 5.1 provides the fractional contribution of
888 the main Higgs decay modes at the generator level to $t\bar{t}H$ search channels. These numbers
889 will be modified by lepton acceptances.

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%

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

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

- 897 • In the same-sign channel, the $t\bar{t}H$ final state contains 6 quarks. These events are then
898 characterized by a large jet multiplicity.
- 899 • In the 3 lepton channel, the $t\bar{t}H$ final state contains 4 quarks from the hard scatter.

- 900 • In the 4 lepton channel, the $t\bar{t}H$ final state contains a small number of light quarks, 0
 901 ($H \rightarrow W^+W^-$ case), 2 or 4 ($H \rightarrow ZZ$ case).

902 **5.2 Background Overview**

903 Background processes can be sorted into two categories:

- 904 • Events with a non prompt or a fake lepton selected as prompt lepton. These pro-
 905 cesses cannot lead to a final state compatible with the signal signature without a mis-
 906 reconstructed object. This category includes events with a prompt lepton but with
 907 mis-reconstructed charge⁷ and events with jets that "fake" leptons. These processes
 908 are rejected with tight object isolation and identification criteria, requiring a large jet
 909 multiplicity, and veto-ing events consistent with a leptonically decaying Z boson.

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

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

⁷Charge mis-identification is almost exclusively a phenomenon of electrons at our energy scales. While it is possible for both electrons and muons to have an extremely straight track, whose direction of curvature is difficult to measure, the happens only at extremely high momentum. Electrons on the other hand may interact with the detector material, resulting in bremsstrahlung. The bremsstrahlung photon may subsequently convert resulting in 2 additional tracks near the original electron. This process, called a trident process, may cause a mismatching of the electron cluster with the original electron track and thus a charge-mis identification and happens also at low energies

917 **5.3 Analysis Strategy**

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

919 The analysis search is conducted in 3 channels, based on counting of fully identified leptons:

920 2 SS leptons, 3 leptons, and 4 leptons. The lepton counting occurs before additional object

921 cuts are made in each individual channel to ensure orthogonality . The division into lepton

922 channels rather than channels targeting specific decay modes allows channels with different

923 sensitivities to be considered separately. We further divide the 2ℓ SS into sub channels based on

924 the number of jets and flavor of the leptons and the 4ℓ channel into sub-channels enriched

925 and depleted in OS leptons arising from Z decays.

926 The channels are fed into a Poisson model

927

CHAPTER 6

928

Dataset and Simulation

929 6.1 Data

930 6.1.1 The 2012 Dataset

931 The LHC successfully produced datasets for physics studies in 2010, 2011 and 2012. The 2012
932 proton-proton dataset was delivered with collisions with a CME of 8 TeV with bunch collisions
933 every 50 ns and reached a total integrated luminosity of around 20 fb^{-1} [57]. Figure 6.1 shows
934 the accumulation of this dataset over time. Despite doubling the bunch spacing (thereby
935 halving the bunch collision frequency), the luminosity neared the design luminosity due to
936 unexpected improvements in the transverse beam profile[58]. This increased the amount of
937 pile-up, or number of collisions per bunch crossing and in general collision events were busier
938 due to these multiple interactions. Figure 6.2 shows the average number of interaction per
939 bunch crossing for the 2011 and 2012 datasets. The 2012 dataset shows an average of 20-25
940 interactions.

941 The $t\bar{t}H$ analysis uses the entire 2012 ATLAS dataset, collected from April to December.

942 The size of the dataset corresponds to 20.3 fb^{-1} , after passing data quality requirements,
943 ensuring the proper operation of the tracking, calorimeter and muon subsystems.

944 The datasets used in the analysis were collected with the primary electron (EF_e24vhi_medium1

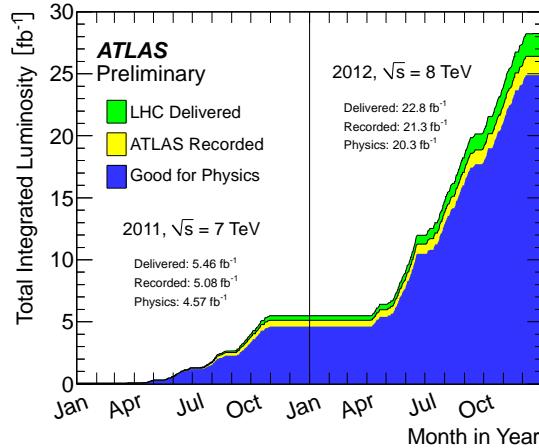


Figure 6.1: Plot showing the accumulation of the integrated luminosity delivered to the ATLAS experiment over 2011 and 2012. The rough size of the usable, physics ready dataset for 2012 is 20 fb^{-1} and is the dataset used for the following analysis.

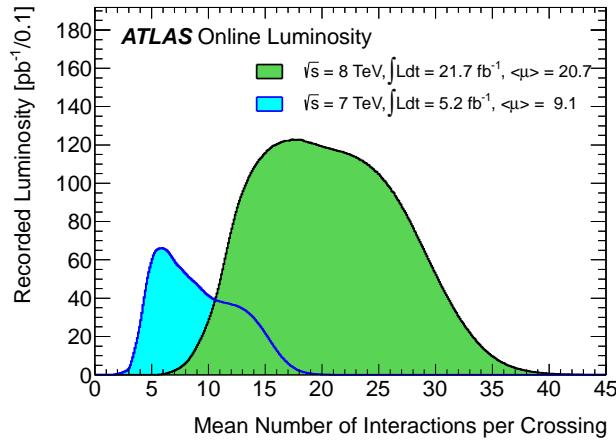


Figure 6.2: The average number of interactions per bunch-crossing for the 2012 and 2011 LHC proton-proton dataset. Most of these interactions are uninteresting but leave energetic signatures in particle detectors called pile-up which interfere with measurements

945 — EF_e60_medium1) and muon triggers (EF_24i_tight — EF_36_tight). The electron triggers require a electron with at least 25 GeV of calorimeter energy, passing the medium identification requirement and loose tracking isolation. Above 60GeV, the isolation requirement is 948 dropped and the identification is loosened slightly. Similarly, the muon trigger requires a good 949 inner detector track and matching hits in the muon spectrometer, as well as loose tracking

950 isolation, which also is dropped about 36 GeV. The data sample must contain either a primary
951 muon or primary electron trigger.

952 **6.2 Simulation**

953 Simulation samples based on are used to determine the overall event selection acceptance
954 and efficiency and for investigations not directly involved in the final result. The simulated
955 samples are created using parton distribution function (PDF) and model using Monte Carlo
956 (MC) techniques the hard parton scatter, underlying event activity and parton showering and
957 hadronization. The samples are then passed through a full ATLAS detector simulation[59]
958 based on GEANT4 [60]. Small corrections are then applied to the overall efficiencies to re-scale
959 object identification efficiencies, energy scales, and the pile-up, discussed in depth later.

960 **6.2.1 Signal Simulation**

961 The $t\bar{t}H$ production is modeled using matrix elements obtained from the HELAC-Oneloop
962 package [61] that corresponds to the next-to-leading order (NLO) QCD accuracy. POWHEG
963 BOX [62, 63, 64] serves as an interface to the parton shower Monte Carlo programs. The
964 samples created using this approach are referred to as PowHel samples. CT10NLO PDF sets
965 are used and the factorization (μ_F) and renormalization (μ_R) scales are set to $\mu_0 = \mu_F =$
966 $\mu_R = m_t + m_H/2$. Pile-up and the underlying events are simulated by Pythia 8.1 [65] with
967 the CTEQ61L set of parton distribution functions and AU2 underlying event tune. The Higgs
968 boson mass is set to 125 GeV and the Top quark mass is set to 172.5 GeV.

969 The signal Monte Carlo samples are summarized in Table 6.1. These large samples are gener-
970 ated with inclusive Higgs boson decays, with branching fractions set to the LHC Higgs Cross
971 Section Working Group (Yellow Report) recommendation for $m_H = 125$ GeV [66]. The in-

Table 6.1: Monte Carlo samples used for signal description.

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

972 clusive cross section (129.3 fb at $m_H = 125$ GeV) is also obtained from the Yellow Report
 973 [66].

974 **6.2.2 Background Simulation**

975 The background simulations used for this analysis are listed in Table 6.2. In general, the
 976 Alpgen[67], MadGraph[68], and AcerMC[69] samples use the CTEQ6L1[70] parton distribu-
 977 tion function, while the Powheg[71], Sherpa[72], are generated with the CT10 PDF. The
 978 exception is the MadGraph $t\bar{t}t\bar{t}$ sample, which is generated with the MSTW2008 PDF[73].
 979 The highest order calculations available are used for cross sections.

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

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

980

CHAPTER 7

981

Object and Event Selection

982 As stated in Chapter 5, the analysis is divided into 3 signal regions based on lepton count-
 983 ing: 2 same-charge leptons, 3 leptons and 4 leptons. The lepton counting occurs for fully
 984 identified leptons with full overlap removal with transverse momentum over 10 GeVto ensure
 985 orthogonality. Lepton selections are tightened afterward within each region.

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

994 **7.1 2ℓ Same-Charge Signal Region**

995 The 2ℓ same-charge signal region (2ℓ SS) requires two leptons similar charge. The
 996 signal is symmetric in charge but the background from $t\bar{t}$ di-lepton production is overwhelming,

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

Signal Region	2ℓ SS	3ℓ	4ℓ
Trigger Matched Lepton	Yes	Yes	Yes
N_l^8	=2	=3	=4
Lepton Charge Sum	+2 or -2	+1 or -1	0
Lepton Momentum (GeV) ⁹	$p_{T0} > 25$ $p_{T1} > 20$	$p_{T0} > 10$ $p_{T1,2} > 25, 20$	$p_{T0} > 25$ $p_{T1} > 20$ $p_{T2,3} > 10$
Jet Counting	$N_b \geq 1, N_{jet} = 4$	$N_b \geq 1, N_{jet} \geq 4$ or $N_b \geq 2, N_{jet} = 3$	$N_b \geq 1, N_{jet} \geq 2$
Mass Variables (GeV)	$ M_{ee} - M_Z < 10$	$ MSFOS - M_Z < 10$	$MSFOS > 10$ $150 < M_{4\ell} < 500$ $ MSFOS - 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)

necessitating the same-sign cut. Requiring only two leptons allows the extra 2 W bosons in the event to decay hadronically, resulting in on average 4 additional light jets plus 2 additional b-quark jets from the top decays.

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

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

In order to maintain orthogonality with the τ analyses, events with fully identified taus are vetoed.

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

1012 7.2 3 ℓ Signal Region

1013 The 3 lepton channel requires 3 leptons, whose summed charge is either -1 or $+1$. The leptons
1014 are ordered in this way:

- 1015* • lep0: the lepton that is opposite in charge to the other two leptons
- 1016* • lep1: the lepton that is closer in ΔR to lep0
- 1017* • lep2: the lepton that is farther in ΔR from lep1

1018 Since events with a "fake" lepton arise from di-lepton processes, $t\bar{t}$ and Z+jets, where
1019 additional jets are mis-identified as the third lepton, lep0 is never the fake lepton. As a result,
1020 the transverse momentum requirement of lep0 is lower than the other two, > 25 GeV. For the
1021 additional two leptons, one must match a trigger and have $p_T > 25$ GeV and the other
1022 must have $p_T > 10$ GeV.

1023 The 3 ℓ channel further requires at least one b-tagged jets and at least 4 jets in total, or two
1024 b-tagged jets and exactly 3 jets in total. Additionally, to suppress $W^\pm Z$ and Z+jet events,
1025 events with same-flavor opposite sign pairs within 10 GeV of the Z pole are vetoed.

1026 Additional cuts, including an M_{ll} cut, and splittings were investigated but low statistics
1027 proved to wash out any advantages.

1028 7.3 4 ℓ Signal Region

1029 In the four lepton signal region, selected events must have exactly four leptons with a total
1030 charge of zero. At least one leptons with must be matched to one of the applied single lepton
1031 trigger. The leading and sub-leading leptons are required to have a p_T of 25 and 15 GeV
1032 respectively. In order to suppress background contributions from low-mass resonances and

1033 Drell-Yan radiation, all opposite-sign-same-flavor (OS-SF) lepton pairs are required to have a
 1034 dilepton invariant mass of at least 10 GeV.

1035 The four-lepton invariant mass is required to be between 100 and 500 GeV. This choice
 1036 of mass window suppresses background from the on-shell $Z \rightarrow 4\ell$ peak and exploits the high-
 1037 mass differences between the signal and the dominant $t\bar{t}Z$ background. Events containing an
 1038 OS-SF lepton pair within 10 GeV of the Z boson mass are discarded. This Z-veto procedure
 1039 greatly reduces background contributions from ZZ production as well as $t\bar{t}Z$ and while it also
 1040 affects the signal by vetoing $H \rightarrow ZZ^*$, $Z \rightarrow \ell^+\ell^-$, these events constitute a small amount
 1041 of the total expected signal. Finally, selected events are required to have at least two jets, at
 1042 least one of which must be tagged as a b-quark initiated jet.

1043 The contribution from $t\bar{t}Z$ comprises approximately 75% of the total background in the
 1044 inclusive signal region. A signal region categorization which factorizes $t\bar{t}Z$ from the remaining
 1045 backgrounds is thus beneficial. The signal region is accordingly divided into two categories
 1046 based on the presence of OS-SF lepton pairs in the final state.

1047 7.4 Electron Selection

1048 The electrons are reconstructed by a standard algorithm of the experiment [51] and the elec-
 1049 tron cluster is required to be fiducial to the barrel or endcap calorimeters: $|\eta_{cluster}| < 2.47$.
 1050 Electrons in the transition region, $1.37 < |\eta_{cluster}| < 1.52$, are vetoed. Electron reconstruction
 1051 and identification is discussed in depth in Chapter 4. Electrons must pass the the VERYTIGHT
 1052 likelihood identification criteria.

1053 In order to reject jets (b-quark jets in particular) misidentified as electrons, electron can-
 1054 didates must also be well isolated from additional tracks and calorimeter energy around the
 1055 electron cluster. Both the tracking and calorimeter energy within $\Delta R = 0.2$ of the electron

1056 cluster must be less than 5% of the electron transverse momentum: $ptcone20/P_t < 0.05$ and
 1057 $Etcone20/E_T < 0.05$. All quality tracks with momentum greater than 400 MeV contribute
 1058 to the isolation energy. Calorimeter isolation energy is calculated using topological clusters
 1059 with corrections for energy leaked from the electron cluster. Pile-up and underlying event
 1060 corrections are applied using a median ambient energy density correction.

1061 The electron track must also match the primary vertex. The longitudinal projection of
 1062 the track along the beam line, $z0 \sin \theta$, must be less than 1 cm) and the transverse projection
 1063 divided by the parameter error, $d0$ significance,must be less than 4. These cuts are used in
 1064 particular to suppress backgrounds from conversions, heavy-flavor jets and electron charge-
 1065 misidentifications.

1066 The electron selection is provided in Table 7.2.

1067 7.5 Muon Selection

1068 Muons used in the analysis are formed by matching reconstructed inner detector tracks with
 1069 either a complete track or a track-segment reconstructed in the muon spectrometer (MS). The
 1070 muons must satisfy $|\eta| < 2.5$. The muon track are required to be a good quality combined
 1071 fit of inner detector hits and muon spectrometer segments, unless the muon is not fiducial to
 1072 the inner detector, $|\eta| > 2.4$. Muons with inner detector tracks are further required to pass
 1073 standard inner detector track hit requirements [52].

1074 As with electrons, muons are required to be isolated from additional tracking or calorimeter
 1075 energy: $ptcone20/P_t < 0.1$, $Etcone20/E_T < 0.1$) A cell-based $Etcone20/P_T$ relative isolation
 1076 variable is used. A pile-up energy subtraction based on the number of reconstructed vertices
 1077 in the event is applied. The subtraction is derived from a Z boson control sample.

1078 The muons must also originate from the primary vertex and have impact parameter re-

1079 requirements, $d0$ significance < 3 , and $z0 \sin \theta < 0.1$ cm, similar to the electrons.

1080 The muon selection is provided in Table 7.2.

1081 **7.6 Jet and b-Tagged Jet Selection**

1082 Jets are reconstructed in the calorimeter using the anti- k_t [53] algorithm with a distance
1083 parameter of 0.4 using locally calibrated topologically clusters as input (LC Jets).

1084 Jets must pass loose quality requirement, ensuring the proper functioning of the calorimeter
1085 at the time of data taking. Jets near a hot Tile cell in data periods B1/B2 are rejected. The
1086 local hadronic calibration is used for the jet energy scale, and ambient energy corrections are
1087 applied to account for energy due to pileup.

1088 p_T and η cuts are tuned based on the sensitivity to $t\bar{t}H$.

1089 For jets within $|\eta| < 2.4$ and $p_T < 50$ GeV, are required to be associated with the primary
1090 vertex, the “jet vertex fraction” (or JV), which is the fraction of track p_T associated with the
1091 jet that comes from the primary vertex, must exceed 0.5 (or there must be no track associated
1092 to the jet).

1093 B-jets are tagged using a Multi-Variate Analysis (MVA) method called MV1 and relying
1094 on information of the impact parameter and the reconstruction of the displaced vertex of the
1095 hadron decay inside the jet[54]. The output of the tagger is required to be above 0.8119 which
1096 corresponds to a 70% efficient Working Point (WP).

1097 **7.7 Tau Selection**

1098 The tau selection is important only in the 2ℓ SS channel, due the tau veto to ensure orthogo-
1099 nality with analyses searching for tau final states for a future combination.

¹¹⁰⁰ **7.8 Object Summary and Overlap**

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

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

¹¹⁰¹ **7.9 Optimization**

1102

CHAPTER 8

1103

Background Estimation

1104 The $t\bar{t}H$ multi-lepton signal regions discussed in Chapter 5 are contaminated by background
1105 contributions at a similar order of magnitude to the signal. The dominant background for
1106 each region is vector boson production in association with top quarks ($t\bar{t}V$). Sub-dominant
1107 but important backgrounds include the production of vector boson pairs in association with
1108 jets and b-quark jets (VV) and $t\bar{t}$ production with a jet misidentified as a lepton (fakes). The
1109 2ℓ SS regions possesses a unique background of charge misidentification from Z and top events.
1110 The methods for estimating these backgrounds are discussed in this chapter. Monte Carlo
1111 simulation is used for the prompt $t\bar{t}V$ and VV contributions. Systematic uncertainties on the
1112 overall normalization of these backgrounds in the signal region are provided from theoretical
1113 studies and past ATLAS analyses and are verified in data-based validation regions. The non-
1114 prompt backgrounds from $t\bar{t}$ jet-misidentification and charge-misidentification are estimated
1115 using data-driven methods.

1116 For reference, Table 8.1 provides a summary of the $t\bar{t}H$ signal and background expectation
1117 for each of the signal regions, including the data-driven estimates discussed in this section.
1118 For each region, the background contribution exceeds the size of the signal.

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				Z depleted	
	≥ 5 jets		4 jets		3 leptons		Z enriched		4 leptons	
	$e^\pm e^\mp$	$e^\pm \mu^\mp$	$\mu^\pm e^\mp$	$e^\pm \mu^\pm$	$\mu^\pm \mu^\pm$					
$t\bar{t}H$	0.73 ± 0.03	2.13 ± 0.05	1.41 ± 0.04	0.44 ± 0.02	1.16 ± 0.03	0.74 ± 0.03	2.34 ± 0.04	0.19 ± 0.01	0.03 ± 0.00	0.00 ± 0.00
$t\bar{t}V$	2.60 ± 0.13	7.42 ± 0.17	5.01 ± 0.16	3.05 ± 0.13	8.39 ± 0.24	5.79 ± 0.20	7.21 ± 0.24	0.74 ± 0.05	0.00 ± 0.00	0.00 ± 0.00
$t\bar{t}Z$										
VV	0.48 ± 0.25	0.37 ± 0.23	0.68 ± 0.30	0.77 ± 0.27	1.93 ± 0.80	0.54 ± 0.30	0.89 ± 0.25	0.71 ± 0.03	0.00 ± 0.00	0.00 ± 0.00
$t\bar{t}, tX$ (MC)	1.31 ± 0.67	2.55 ± 0.84	1.76 ± 0.67	4.99 ± 1.19	8.19 ± 1.41	3.70 ± 1.03	2.46 ± 0.19	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
$Z+\text{jets}$ (MC)	0.16 ± 0.16	0.28 ± 0.20	0.12 ± 0.12	1.37 ± 0.78	0	0.23 ± 0.23	0	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
fake leptons (DD)	2.31 ± 0.97	3.87 ± 1.01	1.24 ± 0.41	3.43 ± 1.38	6.82 ± 1.63	2.38 ± 0.78	2.62 ± 0.51	$(1.1 \pm 0.6) \cdot 10^{-3}$	$(0.09 \pm 0.03) \cdot 10^{-3}$	0.00 ± 0.00
Q misid (DD)	1.10 ± 0.09	0.85 ± 0.08	—	1.82 ± 0.11	1.39 ± 0.08	—	—	—	—	—
Tot Background (fake MC)	4.56 ± 1.17	10.62 ± 1.54	7.57 ± 1.31	10.18 ± 2.43	18.51 ± 2.54	10.26 ± 1.82	11.27 ± 0.40	0.83 ± 0.07	0.01 ± 0.00	0.0110 ± 0.0003
Tot Background (fake DD)	6.49 ± 1.04	12.51 ± 1.04	6.93 ± 0.52	9.07 ± 1.42	18.53 ± 1.83	8.71 ± 0.88	11.43 ± 0.62	0.831 ± 0.075	0.0110 ± 0.0003	
s/\sqrt{b} (fake MC)	0.34	0.65	0.51	0.14	0.27	0.23	0.70	0.21	0.30	
s/\sqrt{b} (fake DD)	0.29	0.60	0.54	0.15	0.27	0.25	0.69	0.21	0.29	
$s/\sqrt{b} \oplus 0.3\text{fake}(MC) \oplus 0.2t\bar{t}V$	0.33	0.58	0.47	0.12	0.22	0.21	0.63	0.207	0.30	
$s/\sqrt{b} \oplus 0.3\text{fake}(DD) \oplus 0.2t\bar{t}V$	0.27	0.53	0.50	0.14	0.23	0.23	0.62	0.207	0.286	

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

¹¹²⁰ **quarks:** $t\bar{t}V$, tZ

¹¹²¹ This section describes the estimation and $t\bar{t}V$ productions. Production of top quarks plus
¹¹²² vector boson is an important background in all multi-lepton channels. A large part of the
¹¹²³ $t\bar{t}V$ component, arising from on-shell $Z \rightarrow \ell\ell$, can be removed via a Z mass veto on like-flavor,
¹¹²⁴ opposite sign leptons. However the $Z \rightarrow \tau\tau$ and γ^* components remain. The $t\bar{t}W^\pm$ and tZ pro-
¹¹²⁵ cesses generally require extra jets to reach the multiplicity of our signal regions. Uncertainties
¹¹²⁶ from the choice of the factorization (μ_F) and renormalisation μ_R scales as well as from the
¹¹²⁷ PDF sets are considered evaluating their impacts on both the production cross sections and
¹¹²⁸ on the event selection efficiencies (particularly resulting from effects on the shape of number
¹¹²⁹ of jets spectrum).

¹¹³⁰ Monte Carlo events for these processes are generated with MadGraph 5 and showered with
¹¹³¹ Pythia 6. $t\bar{t}W^\pm$ events are generated with up to two extra partons at matrix element level,
¹¹³² while for $t\bar{t}Z$ up to one extra parton at matrix-element level is produced. The tZ process is
¹¹³³ simulated without extra partons. The next-to-leading-order (NLO) cross sections are imple-
¹¹³⁴ mented by applying a uniform k -factor to the leading-order (LO) events for each process. For
¹¹³⁵ $t\bar{t}Z$, there is a large component of off-shell production, and for the 3 and 4 ℓ channels low
¹¹³⁶ mass $\gamma^*/Z \rightarrow \ell\ell$ is an important background after on-shell production is removed with a Z
¹¹³⁷ veto. In this case the k -factor is determined by comparing LO and NLO cross sections for
¹¹³⁸ on-shell Z production only.

¹¹³⁹ The $t\bar{t}V$ uncertainties are calculated using the internal QDC scale and PDF re-weighting
¹¹⁴⁰ that is available with MadGraph5_aMC@NLO. The prescription for the scale envelope is taken
¹¹⁴¹ from [74]: the central value $\mu = \mu_R = \mu_F = m_t + m_V/2$ and the uncertainty envelope is
¹¹⁴² $[\mu_0/2, 2\mu_0]$. The PDF uncertainty prescription used is the recipe from [75]: calculate the PDF

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
$\bar{t}Z$	76.0	+5	-4	+7	-7	8.6

1143 uncertainty using the MSTW2008nlo [73] PDF for the central value and then the final PDF
 1144 uncertainty envelope is derived from three PDF error sets each with different α_S values (the
 1145 central value and the upper and lower 90% CL values). The final NLO cross section central
 1146 values and uncertainties are given in Table 8.2.

1147 The tZ process is normalized to NLO based on the calculation in Ref. [76]. Here the scales
 1148 are set to $\mu_0 = m_t$ and the scale variations are by a factor of four; the scale dependence is
 1149 found to be quite small.

1150 8.1.1 $t\bar{t}Z$ Validation Region

1151 Unlike $t\bar{t}W^\pm$, a $t\bar{t}Z$ validation region can be obtained by simply inverting the veto on same-
 1152 flavor opposite sign lepton pairs near the Z pole in the 3 lepton signal region. This region
 1153 thus requires 3 leptons (with momentum and identification cuts discussed in Chapter 7, at
 1154 least one opposite sign, same-flavor pair of leptons within 10 GeV of the Z mass, and either
 1155 4 jets and at least 1 b-tagged jet or exactly 3 jet and 2 or more b-tagged jets. The resulting
 1156 region has low statistics and is not used as a control region but is instead used as a validation
 1157 to demonstrate that the normalization uncertainty, discussed above, is properly evaluated.

1158 The region defined by this is predicted to be 67% $t\bar{t}Z$, 17% WZ , and 13% tZ . We predict
 1159 19.3 ± 0.5 events and observe 28, giving a observed-to-predicted ratio of $1.45 \pm 0.27 \pm 0.03$

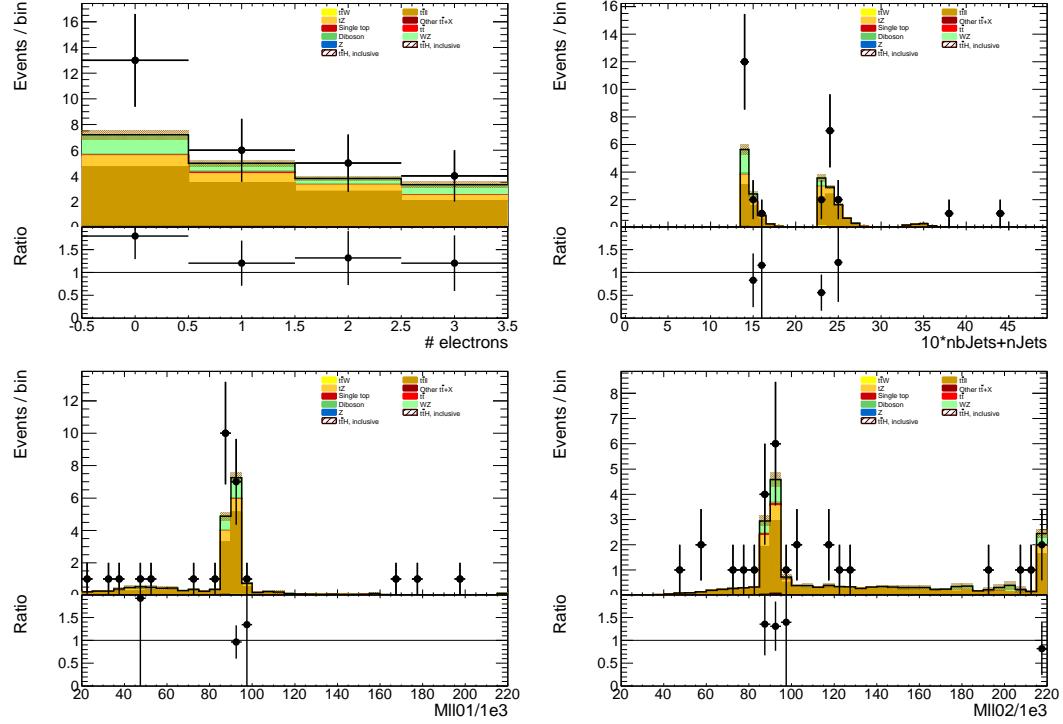


Figure 8.1: Data/MC comparison plots for $t\bar{t}Z$ control region A (≥ 4 jets, ≥ 1 b -tag and 3 jets, ≥ 2 b -tag). In all plots, the rightmost bin contains any overflows. Top left: number of electrons. Top right: 10^* the number of b -tags + the total number of jets. Middle left: the invariant mass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

1160 (where the errors are from data and simulation statistics, respectively). Given the large errors,
 1161 the region is still in agreement with the predictions to within $1-1.5 \sigma$. Distributions of various
 1162 variables are shown in Fig. 8.1.

1163 8.2 Di-boson Background Estimation: $W^\pm Z$, ZZ

1164 $W^\pm Z$ and ZZ di-boson production with additional and b -tagged jets constitute small contribu-
 1165 tions to the 3- and 4-lepton channels respectively. In the 3-lepton case $W^\pm Z$ comprises ~ 1
 1166 event of ~ 10 total background events while the ZZ contribution accounts for approximately
 1167 10% of the total background in the 4-lepton channel. Because of the small size of these con-

tributions, each of the above processes can be assigned a non-aggressive uncertainty based on similar previous analyses with ATLAS and cross-checked with data validation regions and MC truth studies. We assign an overall 50% error on both the $W^\pm Z$ 3-lepton signal region contribution and the ZZ 4-lepton signal region contribution. The details of this error assignment are discussed below.

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

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

8.2.1 $W^\pm Z$ Uncertainty

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

Figure 8.2 shows kinematic variables for the inclusive region ¹⁰. The NJet spectrum shows good agreement within statistics across the full spectrum, giving confidence about the Sherpa high NJet SR extrapolation. Figure 8.3 shows NJet spectrum for the $W^\pm Z + b$ validation region with good agreement in the 1 and 2 jet bins, but a slight data-MC discrepancy in the 3 jet bin. The region has low stats and around $\sim 60\%$ purity.

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

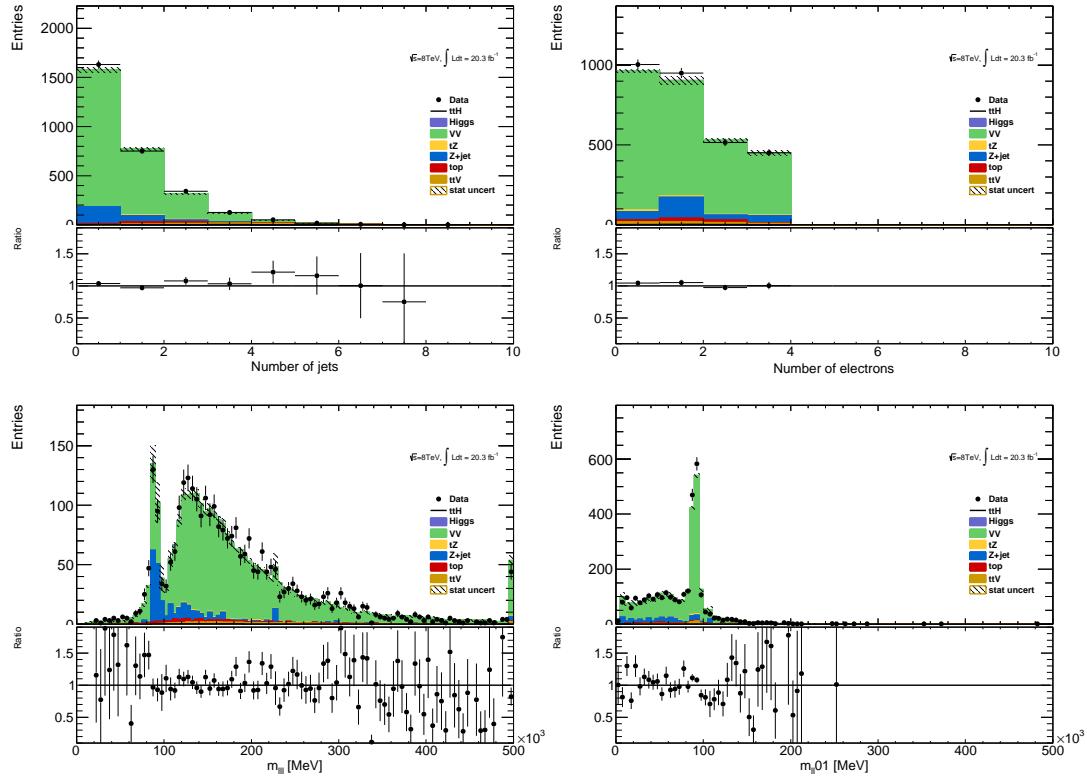
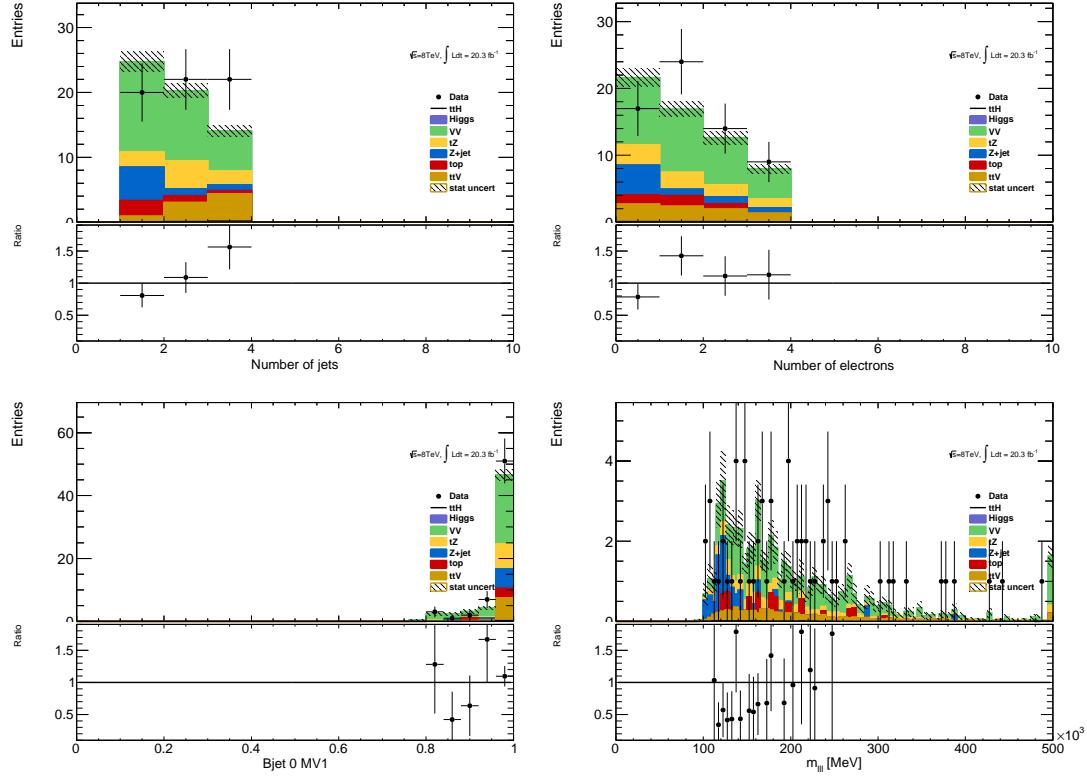


Figure 8.2: Inclusive 3 lepton $W^\pm Z$ validation region using the $t\bar{t}H$ lepton identification and momentum cuts: mass, number of jet and flavor variables

A cross-check is undertaken by examining the $W^\pm Z$ truth origins of the b-jet in the

¹⁰the fakes are taken directly from MC

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

1200 $W^\pm Z + b$ validation region (VR) and the signal region using the sherpa sample available.
 1201 Table 8.3 shows these fractions. As expected the charm and b contributions dominate, though
 1202 there is a small dependence on the number of jets. The composition of the VR is fairly similar
 1203 to that of the signal region, especially in the 3-jet bin. Importantly, also the tagged jet
 1204 composition is also similar to the composition in the $V + b$ analysis, already measured by
 1205 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

1206 **8.2.2 ZZ Uncertainty**

1207 In order to investigate the MC agreement with data in the ZZ case, two validation regions
 1208 similar to the $W^\pm Z$ case are defined. Firstly, a 4 lepton ZZ region is constructed using the
 1209 object selections for the 4-lepton channel and requiring exactly two pairs of opposite sign-same
 1210 flavor leptons with a di-lepton invariant mass within 10 GeV of the Z mass. Additionally, the
 1211 $ZZ + b$ process is investigated directly using a similar validation region which again requires
 1212 exactly two Z -candidate lepton pairs as well as at least 1 b-tagged jet. Some kinematic
 1213 distributions are shown in Figures 8.4 and 8.5, and particular attention should be paid to
 1214 the NJet spectrum which shows good data-MC agreement in the high-jet bins, with a slight
 1215 discrepancy in the 1-jet bin. The agreement for the region with at least 2 jets yields confidence
 1216 in the NJet MC modeling in this region which lies close to the 4-lepton signal region.

1217 Recall that in the $W^\pm Z$ case an overall systematic uncertainty of 50% was assigned to
 1218 cover the MC modeling. Based on the study of the ZZ and $ZZ + b$ validation regions and the
 1219 overall agreement noted with the $Z + b$ analysis, we expect a similar error to be appropriate
 1220 in the ZZ case. A similar truth origin study is undertaken in MC to demonstrate a similar
 1221 b-jet origin to the $W^\pm Z$ case. The true origin of the leading (highest energy) b-tagged jet
 1222 is shown in Table 8.4 for the 4-lepton signal region as well as the $ZZ + b$ validation region
 1223 described above divided into jet bins. It can be seen that in case, as it was in the $W^\pm Z$ case
 1224 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

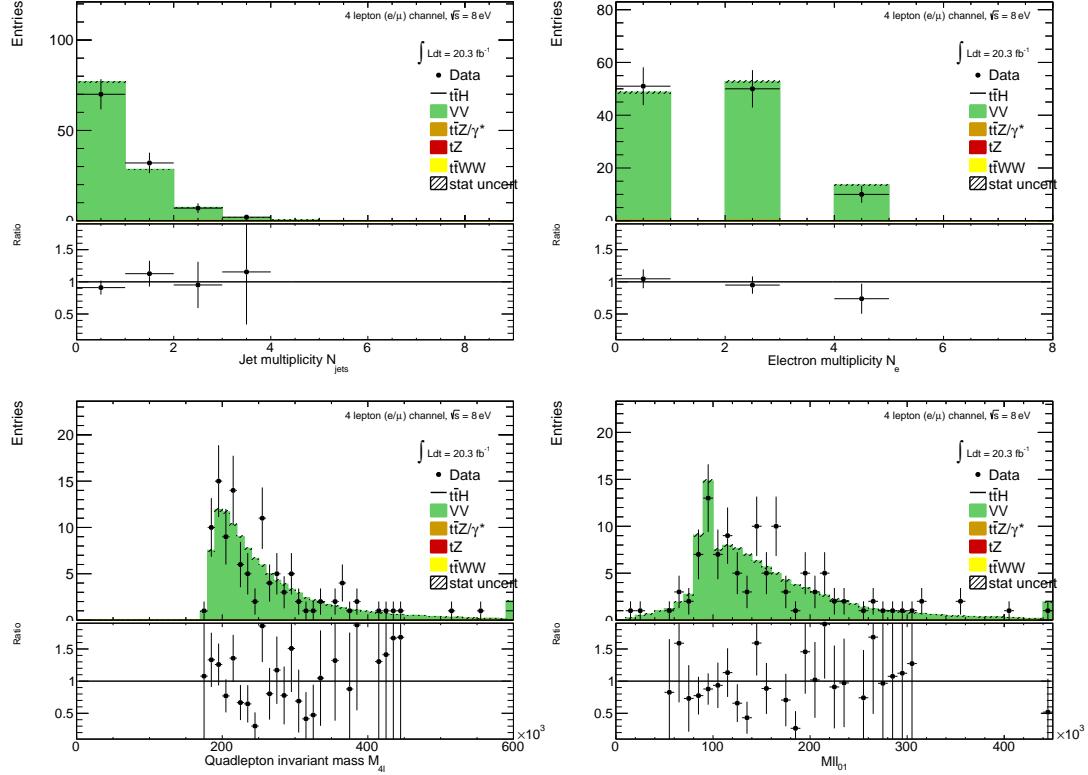


Figure 8.4: Jet-inclusive 4-lepton ZZ validation region using the $t\bar{t}H$ lepton identification and momentum cuts

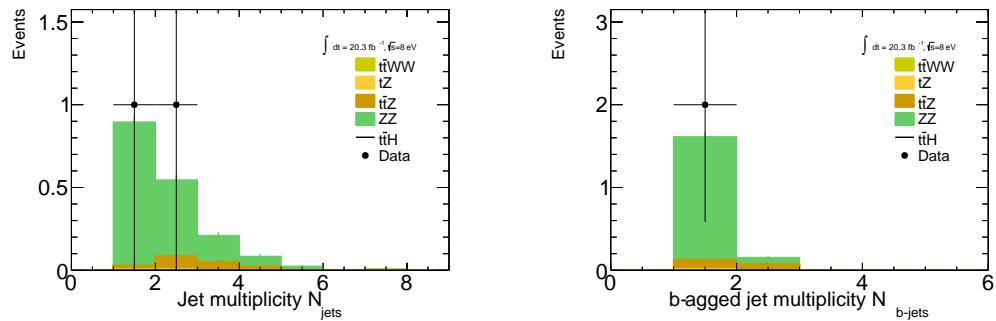


Figure 8.5: $ZZ + b$ validation region using the $t\bar{t}H$ lepton identification and momentum cuts

1225 8.3 Charge-Misidentification Background

1226 Charge-misidentification contributes to the background for 2ℓ SS case and only for flavor
 1227 channels, which include electrons. The same-sign require is useful in removing large SM
 1228 opposite sign-backgrounds, but because of their size even small charge mis-identification rates
 1229 result in contamination in same-sign regions. For the 2ℓ SS signal regions and low NJet
 1230 control regions, charge-misidentification background arise primarily from $t\bar{t}$ di-lepton events
 1231 with a smaller contribution from leptonic Z decays.

1232 In general, charge-misidentification can arise in two ways. The first occurs for ultra-high
 1233 energy electrons and muons, which leave tracks in the detector that are too straight for
 1234 the fit determine the direction of curvature with high confidence. This type of charge mis-
 1235 identification is not a concern to the $t\bar{t}H$ multi-lepton analysis, as most of the leptons have
 1236 momentum > 150 GeV. The second source of charge misidentification is from 'tridents', which
 1237 only occurs for electrons, because their low mass allows for high rate bremsstrahlung in the
 1238 detector material. In some cases, after an electron releases a photon through bremsstrahlung,
 1239 the photon may convert nearby resulting in three electron tracks. The reconstruction algo-
 1240 rithms may sometimes match the wrong track to the calorimeter energy deposit, resulting in
 1241 a possible charge mis-identification. As discussed in the selection, tight track-cluster geomet-
 1242 ric and energy matching requirements are applied on the electron candidates to reduce the
 1243 overall rate and the electron acceptance is narrowed to ($|\eta| < 1.37$), since most of the material
 1244 is concentrated more forward in the detector. For this reason, muon charge-misidentification
 1245 is considered negligible for this analysis.

1246 We estimate the contribution of charge-misidentification events in our 2ℓ SS signal regions
 1247 and relevant control regions by applying a weight per electron in the opposite-sign region
 1248 with otherwise same cuts. The weight is related to the charge-misidentification rates and

1249 is estimated using a likelihood method in the opposite-sign and same-sign $Z \rightarrow ee$ control
 1250 regions. The rate measured from these control regions is binned in electron p_T and η , to
 1251 account for dependencies in these variables. The method, validations and associated errors
 1252 are discussed in detail in the following sub-sections.

1253 **8.3.1 Likelihood Method**

1254 The number of reconstructed same-sign (N_{ss}) and opposite sign (N_{os}) $Z \rightarrow ee$ events are
 1255 related to number of produced $Z \rightarrow ee$ opposite sign events (N) via factors related to the
 1256 charge mis-identification rate. For a single per-electron charge mis-identification rate (ϵ , these
 1257 quantities are related as follows (with the assumption that ϵ is very small):

1258 • $N^{os} = (1 - 2\epsilon + 2\epsilon^2)N$ opposite-sign events,

1259 • $N^{ss} = 2\epsilon(1 - \epsilon)N \simeq 2\epsilon N$ same-sign events,

1260 Knowing ϵ , the charge-misidentification rate, and supposing we can have a different rate
 1261 per-electron, it is possible to estimate the number of same-sign events from the number of
 1262 opposite sign events.

1263 • $N^{ss} = \frac{\epsilon_i + \epsilon_j - 2\epsilon_i\epsilon_j}{1 - \epsilon_i - \epsilon_j + 2\epsilon_i\epsilon_j} N^{os}$ for the ee channel,

1264 • $N^{ss} = \frac{\epsilon}{1 - \epsilon} N^{os}$ for the $e\mu$ channel,

1265 where ϵ_i and ϵ_j are the charge mis-identification rates for the two different electrons.

1266 Although it is impossible from a typical same-sign $Z \rightarrow ee$ to know which electron's charge
 1267 was mis-identified, we can use a likelihood method over the whole sample to measure charge
 1268 mis-identification rate (ϵ) depends on the electron p_T and η . The likelihood method assumes
 1269 that the mis-identification rates of the electron charge are independent for different pseudo-
 1270 rapidity regions. Therefore, the probability to have a number of same-sign events (N_{ss}^{ij}) with

1271 electrons in $|\eta|$ region i and j can be written as a function of the number of events N^{ij} as
 1272 follows:

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

1273 If all the same-sign events in the Z peak are produced by charge mis-identification, then
 1274 N_{ss}^{ij} is described by a Poisson distribution:

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

1275 where k is the observed number of occurrences of the event, i.e. $k = N_{ss}^{ij}$, and λ is the expected
 1276 number, i.e. $\lambda = N^{ij}(\epsilon_i + \epsilon_j)$. Thus, the probability for both electrons to produce a charge
 1277 mis-identification is expressed by:

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

1278 The likelihood L for all the events is obtained by evaluating all the $|\eta|$ combinations:

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

1279 where the rates ϵ_i and ϵ_j can be obtained by minimizing the likelihood function. In this
 1280 process, the $-\ln L$ is used in order to simplify and make easier the minimization. Terms
 1281 which do not depend on the rates ϵ_i and ϵ_j are removed in this step. This way, the final
 1282 function to minimize is given by the following expression:

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

1283 The events are selected within the Z peak and stored –with the electron order by $|\eta|$ – in
 1284 two triangular matrices: one for the same-sign events N_{ss}^{ij} , and the other one for all events N^{ij} .
 1285 The likelihood method takes into account electron pairs with all $|\eta|$ combinations, which allows
 1286 to use the full available statistics getting therefore lower statistical uncertainties. Moreover,

1287 it does not bias the kinematical properties of the electrons, compared to other methods like
 1288 tag-and-probe.

1289 The likelihood method can be easily extended to measure the charge mis-identification rates
 1290 as a function of two parameters. In this study, the interest lies not only on the measurement
 1291 of the rates as a function of the pseudo-rapidity, but also transverse momentum. Thus, the
 1292 probability to find a same-sign event given the rates for each electron is $(\epsilon_{i,k} + \epsilon_{j,l})$, where the
 1293 two indices represent binned $|\eta|$ - and p_T -dependence. Thus, the Eq. 8.5 transforms into

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

1294 The likelihood method uses only Z *signal* events. Therefore, background coming from
 1295 other processes where the dilepton invariant mass corresponds to the one of the Z boson needs
 1296 to be subtracted. The background subtraction is done using a simple side-band method. This
 1297 method consists in dividing the Z invariant mass in three regions, i.e. A , B and C , where B
 1298 is the central region corresponding to the Z peak. The number of events is counted in the
 1299 regions on the sides of the peak, i.e. n_A and n_C , and removed from the total number of events
 1300 in the peak region B , n_B . This way, the number of signal events N_Z is given by

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

1301 Once the background has been subtracted, the likelihood method can be applied. MINUIT
 1302 is used for the minimization and MIGRAD to compute the uncertainty on these rates.

1303 8.3.2 Results

1304 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,
 1305 2.47] by 4 p_T bins [15,60,90,130,1000]. For p_T bins above 130 GeV, the Z dataset becomes too
 1306 small and the rates are calculated using $t\bar{t}$ MC, scaled by the data-MC ratio of the rates in the
 1307 lower p_T bins, [90-130] GeV. Figure 8.6 shows the extracted rates in all bins.

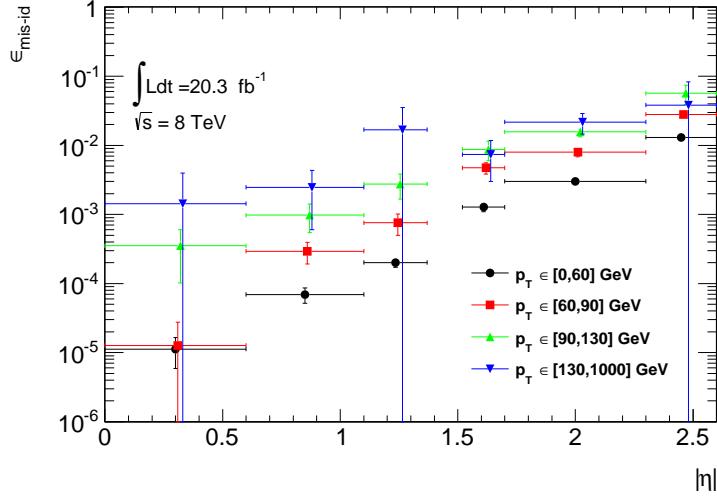


Figure 8.6: Electron charge mis-identification rates measured in data with the likelihood method on Z events (black points, red squares and blue triangles) as a function of $|\eta|$ and parametrized in p_T . The full 2012 dataset has been used to estimate the rates below 130 GeV. Above this value, the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

1308 To validate the likelihood approach, we apply the full method to the Z MC samples
 1309 (extracting rates via a likelihood fit and applying them to opposite sign events) and compare
 1310 to the MC predicted number of same-sign events. The invariant mass of the Z from our charge
 1311 mis-identification and directly from the MC can be seen on Figure 8.7. In the simulated Z
 1312 samples, the number of same-sign Z events is 5 049 while the estimation is $5\,031^{+375}_{-365}$. The
 1313 uncertainties combine both statistical systematic uncertainties, which are discussed in depth
 1314 below. The validation gives compatible results within uncertainties.

1315 8.3.3 Systematic and Statistical Uncertainties

1316 Statistical uncertainties dominate the combined uncertainty on the charge mis-identification
 1317 estimate. The statistical uncertainties come primarily from the size of the Z same-sign sample
 1318 in data and are especially large for central, material-poor regions where the charge mis-

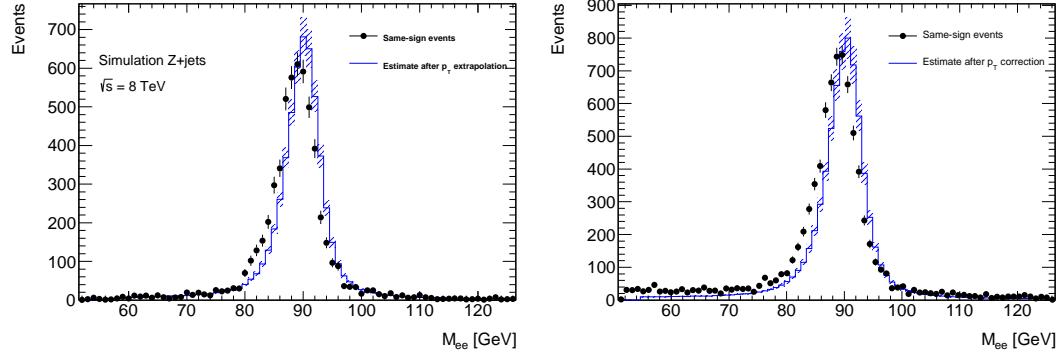


Figure 8.7: Closure test on simulated $Z \rightarrow e^+e^- + jets$ events (a) and data (b). The black circles show the distribution of same-sign events while the blue histograms show the distribution of the re-weighted opposite-sign events together with the statistical and systematic uncertainties.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

1319 identification rate is extremely low. Additionally systematic uncertainties are included for
 1320 a comparison between the positron and electron rate, the per-bin MC closure test discussed
 1321 above in the Results section, and for the effect of varying the invariant mass window used
 1322 for the background subtraction for three different cases. The high p_T extrapolation induces a
 1323 statistical error only in the last p_T bin. This bin is essentially irrelevant to the energy scales
 1324 considered in this analysis. Figure 8.8 shows the relative bin uncertainties for all rate bins.

1325 We apply the rates to estimate the charge mis-identification background in the 2ℓ SS signal
 1326 regions, and find $\sim 25\%$ contamination in the $e^\pm e^\pm$ regions and a $\sim 10\%$ contribution to the
 1327 $e^\pm \mu^\pm$ regions with a 10% systematic error overall. The low overall error can be attributed to
 1328 the fact that the statistical error is lowest where the bulk of charge misidentifications occur.

1329 8.4 Fake Lepton Backgrounds

1330 Fake Leptons, from the mis-identification of jets as either electrons or muons, primarily arise
 1331 from $t\bar{t}$ and single top processes in the 2ℓ SS, 3ℓ and 4ℓ channels. Smaller contributions come
 1332 from $Z + jet$ events. These backgrounds are sub-dominant but important in the 2ℓ SS and 3ℓ

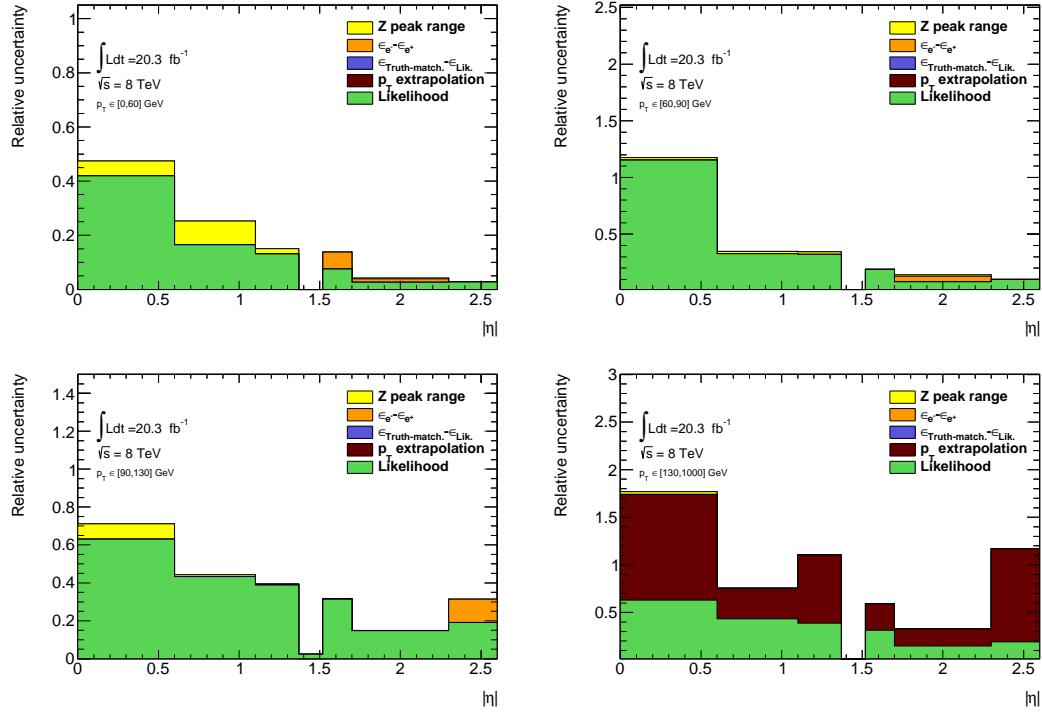


Figure 8.8: Relative systematic uncertainty contributions on the charge mis-identification rate, for different bins in p_T and $|\eta|$. Tight++ electrons have been used to produce this plot.

1333 channels. They are extremely small in the 4ℓ channels. Truth studies suggest that these mis-
 1334 identified leptons arise overwhelmingly from b-quark initiated jets. The general method for
 1335 estimating fakes in all channels is to define a reversed object selection region (usually isolation)
 1336 for each lepton flavor with otherwise identical signal region selection (N_{CR}^e , N_{CR}^μ). This region
 1337 is fake-dominated with small contributions from prompt backgrounds, which are subtracted
 1338 from the data. The total number of fake events in this region is then scaled by a transfer
 1339 factor (θ) to estimate the number of fake events of the appropriate flavor in the signal region.
 1340 The transfer factor is defined in Equations 8.8 and the simple formula for determining fakes is
 1341 defined in Equations 8.9. 'd' refers to anti-identified electrons, and 'p' refers to anti-identified
 1342 muons.

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

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

1343 This approach factorizes the background model into two separate measurements. N_{CR} is
 1344 sensitive the overall $t\bar{t}$ production rate, especially in the presence of additional jets from QCD
 1345 ratio, as well as the object-level misidentification of a jet as a lepton. The transfer factor θ
 1346 is sensitive to only the object level properties of the mis-identified jet, and in particular only
 1347 the variables which are reversed in the anti-tight identification.

1348 The transfer factor is obtained in a different way for each channel, due to unique issues
 1349 with statistics and contamination, but each method relies heavily on the data-based control
 1350 regions with fewer jets. Figure 8.9 shows a truth study of the stability of the transfer factor
 1351 for the 2ℓ SS and 3ℓ cases as a function of the number of jets in the event for events with
 1352 one b-tagged jet. This suggest that the regions with fewer jets are a good model of the fakes
 1353 in the signal regions with more jets and is expected because of the homogeneity of origin of
 1354 the fakes across all jet bins.

1355 The details of the methods for each channel are discussed in depth in the following sections.
 1356 For all methods, the overall systematic uncertainty on the normalization of the fake estimate
 1357 is in the range 30%-50% and arise primarily from statistics and the closure on assumptions
 1358 used to obtain the transfer factor.

1359 Because these methods do provide a per-object transfer-factor that depends on the prop-
 1360 erties of the faking object, we must use the MC to model the shapes of the fake kinematic
 1361 distributions in the signal regions. This is not an essential issue, since the analysis only

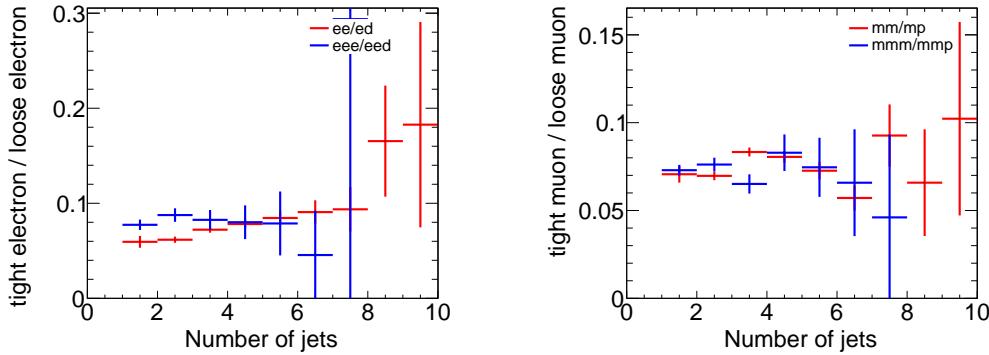


Figure 8.9: Ratios of regions with tight and anti-tight leptons in 2-lepton and 3-lepton channels

1362 considers only the total number of events in each signal region in the final measurement of
 1363 $t\bar{t}H$ production.

1364 8.4.1 2ℓ SS Fakes

1365 The 2ℓ SS fake method follows the procedure outlined in general above. We define anti-tight
 1366 electron and muon control regions with reversed particle identification criteria for each signal
 1367 region, including the 6 flavor and jet-counting sub regions. The anti-tight muon and electron
 1368 criteria are provided below:

1369 • anti-tight electron (d): fails to verify the verytight likelihood operating point, but still
 1370 verifies the veryloose operating point. fails relative tracking and calorimeter isolation,
 1371 $E_T^{rel} > 0.05$ and $p_T^{rel} > 0.05$.

1372 • anti-tight muon (p: $6 \text{ GeV} < p_T < 10 \text{ GeV}$

1373 The electron and muon transfer factors, θ_e and θ_μ , are calculated in the region with
 1374 signal region selection but fewer jets, $NJet == 2$ or $NJet == 3$ and are defined as the
 1375 ratio of the number of events for two fully identified leptons to the number of events with
 1376 one fully identified lepton and one anti-identified lepton, after the prompt and charge mis-

1377 identification backgrounds are subtracted. Only same-flavor channels are used to ensure that
 1378 muon and electron transfer factors maybe estimated separately: on every region, the prompt
 1379 and charge-misidentification backgrounds are subtracted from the data.

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

1380

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

1381

1382 The 2,2 jet anti-tight regions used in obtaining the transfer factors are shown in Figure
 1383 8.10 and the 4,5 jet anti-tight regions, scaled by the transfer factors to get the fake estimates
 1384 in the SR are shown in Figure 8.11. The $t\bar{t}MC$ is included in the plots for reference, although
 1385 it is not included in the measurements.

1386 The transfer factors obtained from the 2 and 3 jet regions are shown in 8.5 with statistical
 1387 errors and propagated systematic errors from the subtraction of relevant backgrounds ($t\bar{t}V$ and
 1388 charge mis-identification). The MC values are just for comparison. An additional systematic
 1389 error is added by comparing the transfer factors, obtained from the low jet control region, to
 1390 those obtained from the higher jet signal regions, using $t\bar{t}MC$. The value of this systematic
 1391 is about 20 % and can be seen in the above Figure 8.9. The overall systematic uncertainties
 1392 and contribution from each source in all of the sub-channels of the signal region are shown in
 1393 Table 8.5 and the final contribution of fake events to the signal region are show in Table 8.1
 1394 found at the beginning of the chapter.

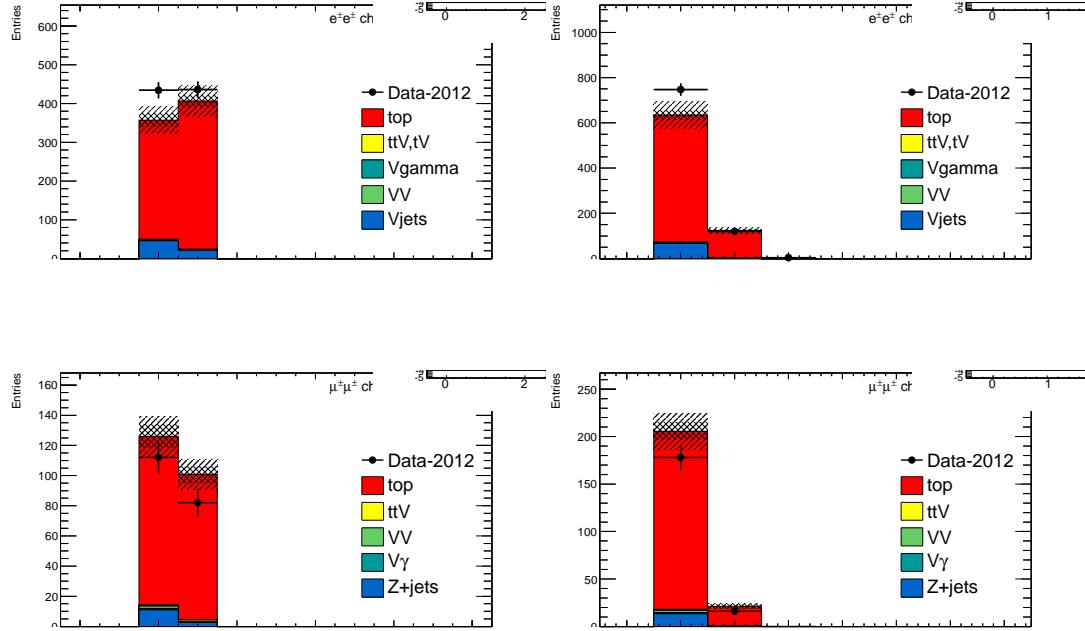


Figure 8.10: 2,3 Jet SS 2ℓ ed (above) and μp control regions with at least one b-tagged jet. This region and the associated tight-id regions are used to determine θ_e and θ_μ . The $t\bar{t}$ MC (red) is used for reference and not used in the actual calculation

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.

1395 8.4.2 3ℓ Fakes

1396 The 3ℓ fake method follows the same general strategy as the 2ℓ SS case. Transfer factors are
 1397 used extrapolate from an anti-tight, fake-rich control region in data into the signal region.
 1398 However, the equivalent low jet control regions are too low in statistics to provide the transfer
 1399 factors from data directly, as above. Instead, the transfer factors are obtained directly from the
 1400 $t\bar{t}$ simulation and data control regions are used to determine the modeling of the identification
 1401 and isolation variables, used in the transfer factor extrapolation. The low jet regions are still

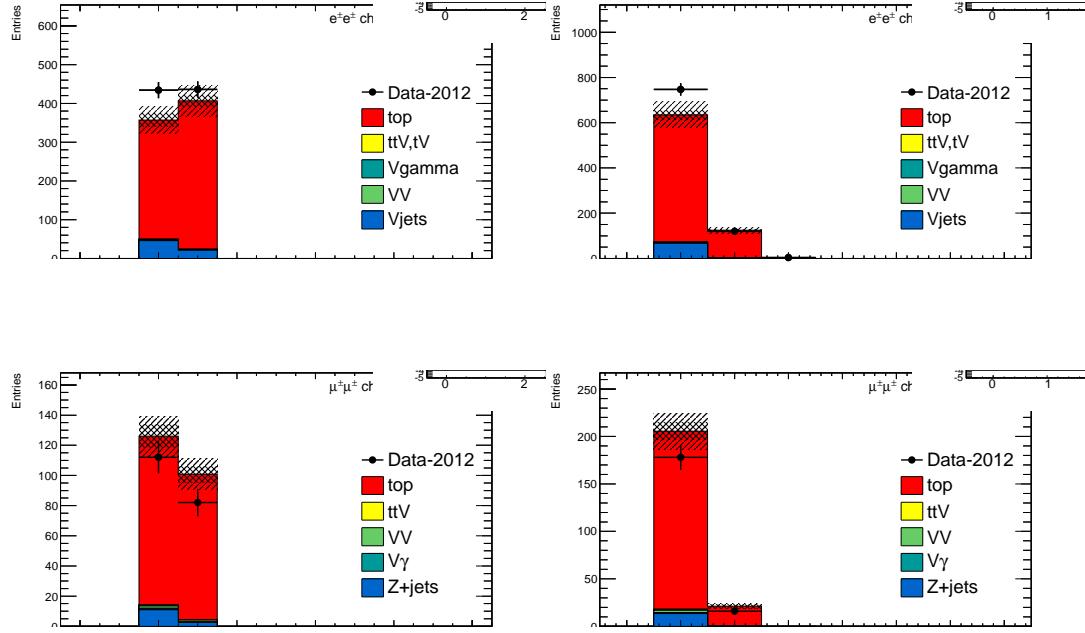


Figure 8.11: 4,5 Jet SS 2ℓ ed (above) and $\mu\mu$ control regions with at least one b-tagged jet. After subtraction of prompt and charge mis-identification backgrounds, these regions are scaled by the transfer factors, θ_e and θ_μ to obtain the final number of fake events in the CR. The $t\bar{t}$ MC (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
	$\Delta N_{\ell anti-\ell}(n jets)$ (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 ($\Delta N_{\ell anti-\ell}(n jets)$). For channels with muons, uncertainties between parenthesis are obtained when anti-tight muons are defined such as $p_T < 20$ GeV (includes blinded data)

1402 employed in a low statistics validation of the entire fake procedure.

1403 Anti-tight electrons and muons are defined in slightly different ways, compared to the 2ℓ

1404 SS case:

1405 • anti-tight electron (d): fails to verify the verytight likelihood operating point, but still

1406 verifies the veryloose operating point. the isolation selection is released $E_T^{rel} > 0.05$,

1407 $p_T^{rel} > 0.05$.

1408 • anti-tight muon(p): muons must pass identification but the p_T cuts is lowered to 6 GeV,

1409 the overlap removal with jets and isolation cuts are released.

1410 The transfer factors, θ_e and θ_μ , extracted from MC, is defined as the ratio of the number of

1411 top ($t\bar{t}$ + single-top) events in the signal region, to the number of $t\bar{t}$ events in the anti-tight

1412 regions. The factors are calculated in separate flavor regions to ensure that the electron jet

1413 fakes and muon jet fakes are calculated separately. The calculation follows the same for as

1414 in Equation 8.8, but now lep0, which by construction is almost never a fake is allowed to be

1415 either electron or muon in both cases, denoted below in Equation 8.12.

$$\theta_e = \frac{N_{xee}}{N_xed}, \theta_\mu = \frac{N_{x\mu\mu}}{N_{x\mu p}} \quad (8.12)$$

1416 The MC modeling of the variables involved in the transfer factor can be verified when

1417 another variable fails. For instance, the MC modeling of the electron isolation variable can be

1418 compared to data when the particle identification variable fails and vice-versa. The modeling

1419 of muon-jet ΔR , involved in the overlap removal, can be compared when either the isolation

1420 variable or the p_T fails. The comparison of the electron variables in this manner can be seen in

1421 Figure 8.13 and the muon variables in Figure 8.12. The regions used have the same selection

1422 as the signal region with an added missing transverse energy requirement, > 60 GeV, to

1423 ensure only top fakes. 20% and 21% systematic uncertainties are assigned to the muon and

1424 electron transfer factors, respectively, to account for data-MC discrepancies. This method for
 1425 evaluating data-MC agreement for individual electron and muon variables in turn relies on the
 1426 assumption that these variables are largely un-correlated and that the transfer factor itself is
 1427 factorizable into pieces for each variable. This factorized and fully correlated transfer factors
 1428 have therefore been compared using MC and shown to have differences than the systematic
 1429 quotes, suggesting that these assumptions are reasonable.

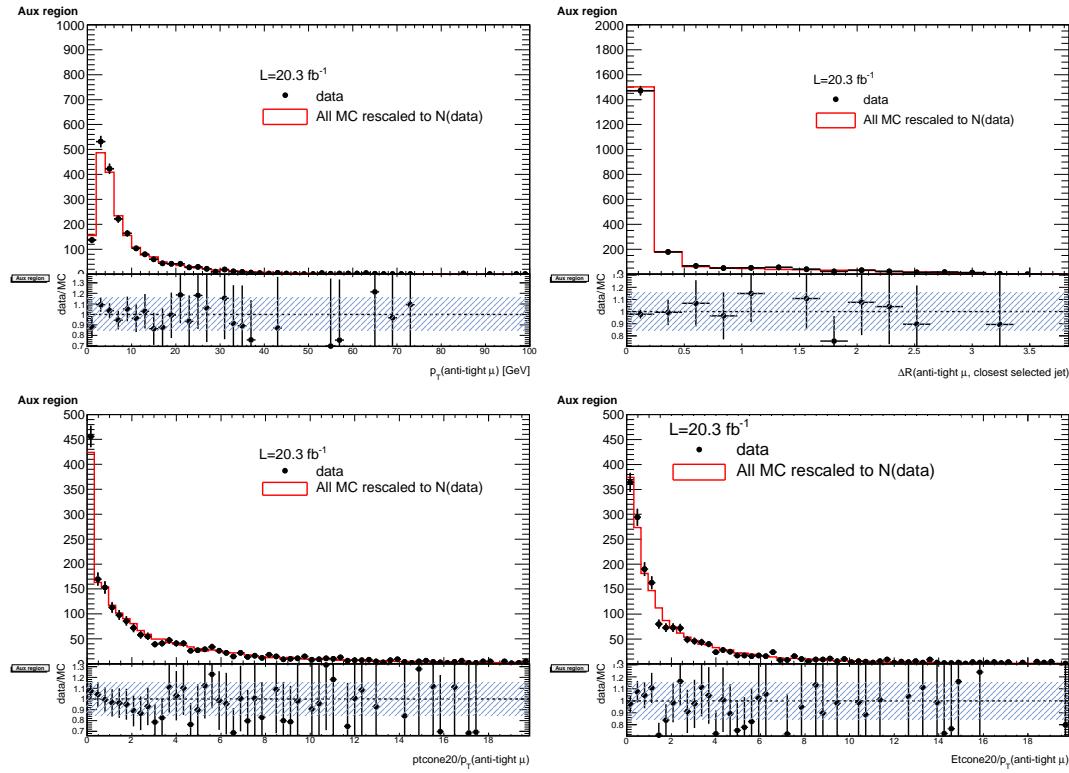


Figure 8.12: Distributions of the properties of the anti-tight muons in data (dots), compared with the total simulation (red line), rescaled to the integral of the data for a shape comparison. The uncertainty on the data distribution is statistical. The number of events for each of them is also presented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

1430 The electron and muon anti-tight control regions, which are scaled by the transfer factors

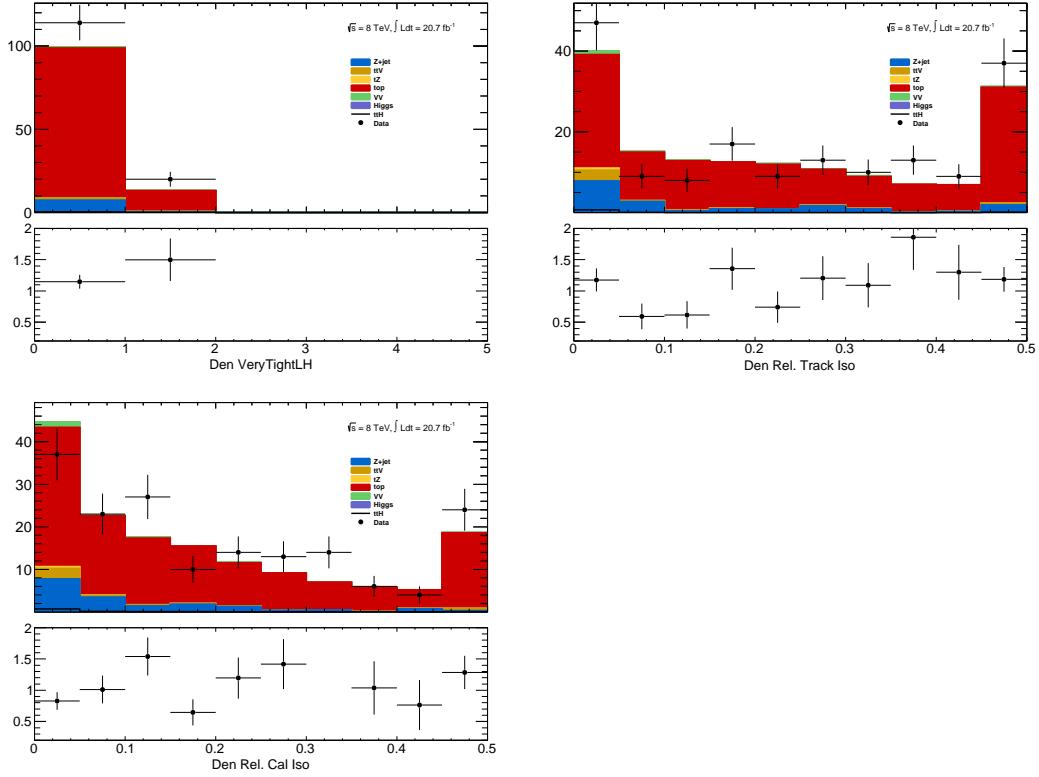


Figure 8.13: Distributions of anti-tight electron variables. The variables presented are, from top left to bottom right, p_T , η , Very Tight Likelihood word value, $ptcone20/p_T$, $Etcone20/p_T$. The plotted regions have the same cuts as the signal region, except the anti-tight electron must fail isolation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

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 the calculation are shown in Table 8.7.

Finally, the low jet region (1,2,3j) is used as a validation for the method, described above.

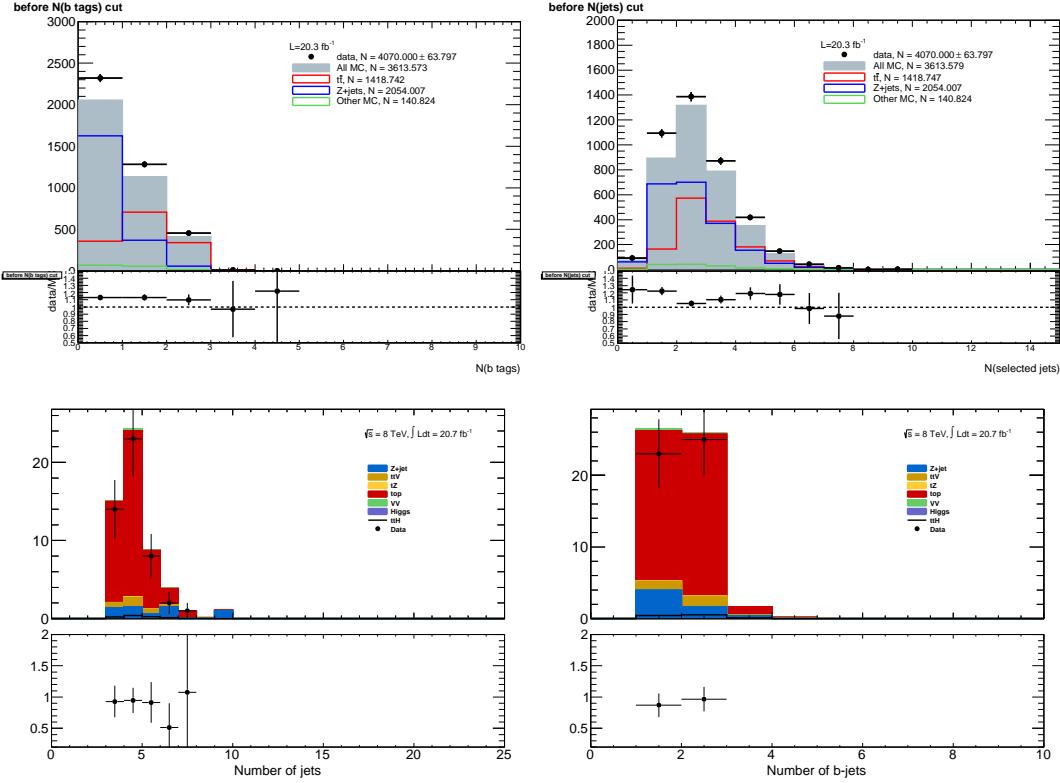


Figure 8.14: Muon-xxp (above) and electron-xxd (below) anti-tight control regions: jet variables. Note: the $t\bar{t}$ and top MC in the plots is used only as comparison, but is not included in the fake measurement

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

1438 The $t\bar{t}$ and single top fakes are estimated using the procedure above, but instead using the
 1439 lower jet region. Similar systematics are assessed. This region with the fake estimate is
 1440 plotted in Figure 8.15. The agreement of data and summed prediction for the fakes and
 1441 prompt backgrounds is well within the systematic and statistical uncertainties. The figure
 1442 also shows the same region with relaxed p_T cuts on all leptons to 10 GeV. This increases the
 1443 purity of fakes in the region as well as the statistics. The data and summed fake and prompt
 1444 predictions are also well within the statistical and systematic uncertainties.

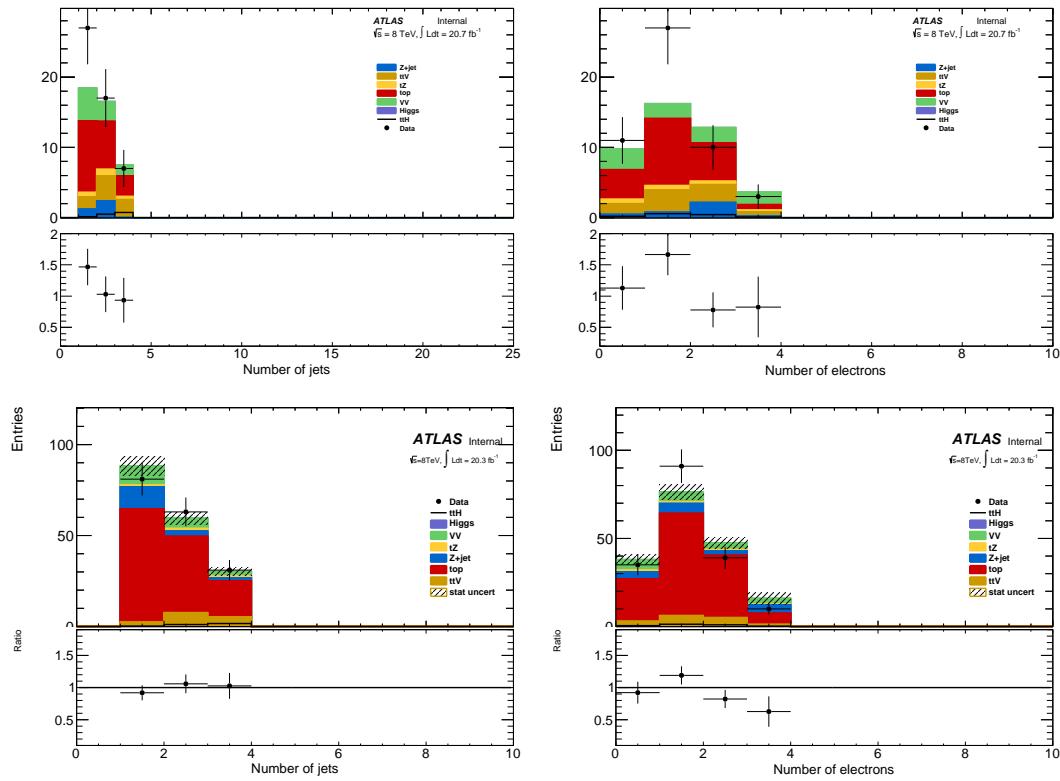


Figure 8.15: 3ℓ fake validation regions for nominal p_T selection (above) and relaxed p_T selection, > 10 GeV, (below). Plotted are the number of jets and the number of electrons in each event. The data and MC ratio below each plot agree with 1 within the statistics of the region and the overall systematic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

1445 8.4.3 4ℓ Fakes

1446 We will not discuss the 4ℓ fakes in depth, as it is a very small background - at the % level
1447 and will have almost no impact on the final result. It is important, however, to carry out
1448 the measurement using the data to ensure that this is indeed the case. The fake method
1449 used in the the 4ℓ case is similar to the 2ℓ and 3ℓ cases discussed above. All fakes arise
1450 from $t\bar{t}$ and single-top events, where two jets are mis-identified as leptons. To measure the
1451 contribution of this background, control regions with 2 fully identified and 2 anti-identified
1452 leptons are created. These control regions do not have a number of jets requirement in order
1453 to increase statistics. From these control regions, two extrapolations are made. First, a
1454 transfer factor is applied to extrapolate from the anti-tight to tight regions for electrons and
1455 muons. The regions are defined with identical object identification selection and reversal as
1456 the 3ℓ case, and the same transfer factors can be used. They must be used twice however,
1457 because there are two anti-identified leptons in each event. Second, the jet inclusive regions
1458 are extrapolated into the 2-jet signal region, using as a second extrapolation factor derived
1459 from $t\bar{t}$ events. Since, the majority of fake leptons arise from b-quark initiated jets, the jet
1460 spectrum $t\bar{t}$ events with the additional requirement of 2-b-tagged jets from data are used as
1461 a model for the jet extrapolation. The overall systematic uncertainty on this measurement
1462 arises from the statistics in the control regions and MC based assessments of non-closure and
1463 are 35%-50% depending on the sub-channel.

1464

CHAPTER 9

1465

Summary of Systematic Uncertainties

1466 This chapter summarizes the various systematic uncertainties that enter the measurement of
1467 the limit of $t\bar{t}H$ multi-lepton analysis, as well as structure of the statistical analysis model
1468 used to obtain the measurement. The systematic uncertainties arise from three main sources.

1469 The first are the normalization uncertainties on the background process estimation methods,
1470 which are discussed in depth in . The second source is the theoretical uncertainties on the
1471 $t\bar{t}H$ production cross-section and acceptance. The final source are the experimental and de-
1472 tector related systematic uncertainties related to event selection efficiencies and measurements
1473 and identification of the objects. They affect only the non-data driven backgrounds and the
1474 $t\bar{t}H$ signal, as simulation is used to model their acceptance and efficiency for the analysis
1475 selection.

1476 These systematic uncertainties, the estimated background and signal event counts in each
1477 of the signal regions, and the observed data in each signal region are combined in a statistical
1478 fit to an analysis model to extract the measurement of interest. We measure per-channel and
1479 combined ratios of the observed production rate to the theoretically predicted production rate
1480 of $t\bar{t}H$, a parameter called μ . In the absence of a statistically significant observation, this
1481 measurement is translated into a upper confidence limit on μ . The details of this procedure
1482 are discussed in the following sections and the results with the observed data are discussed in

1483 Chapter 10

1484 **9.1 Systematic Uncertainties on Signal Cross-section and**
1485 **Acceptance**

1486 The $t\bar{t}H$ signal is simulated with matrix elements at NLO (next-to-leading order) in QCD
1487 with Powheg and is discussed in Chapter 6.

1488 The production cross section and the Higgs boson decay branching fractions together with
1489 their theoretical uncertainties from the QCD scale and PDF choice are taken from the NLO
1490 theoretical calculations reported in Ref. [66]. The uncertainty from the QCD scale estimated
1491 by varying μ_0 by a factor of 2 from the nominal value is $^{+3.8\%}_{-9.3\%}$, while the uncertainty from the
1492 PDF set and the value of α_S is $\pm 8.1\%$.

1493

1494 The impact of the choice of the QCD scale on the simulation of the $t\bar{t}H$ event selection
1495 efficiency is estimated in two independent ways.

1496 First, the factorization and renormalisation scales μ_0 are varied by a factor of 2, as $\mu = 2\mu_0$
1497 and $\mu = \mu_0/2$. The effects of these new scales are estimated via the application of event re-
1498 weighting procedures on the nominal simulation using kinematic distributions at parton level.
1499 The weights used are dependent on the transverse momenta of both the $t\bar{t}H$ system and of
1500 the top quark, as described in Ref. [81].

1501 Second, the choice of the factorization and renormalisation scales, dependent on fixed
1502 (“static”) parameters in the nominal simulation, is tested comparing its prediction with an
1503 alternative (“dynamic”), but still physics motivated choice $\mu_0 = (m_T^t m_T^{\bar{t}} m_T^H)^{1/3}$, which depends
1504 on kinematic variables. This comparison is performed via event re-weighting of the nominal
1505 static simulation based on weights derived as a function of the $t\bar{t}H$ transverse momentum [81].

1506 In order to take the difference between the choices of scale as systematic uncertainties, a
 1507 symmetric envelope around the nominal simulation is built applying the weights and also
 1508 their inverses.

1509 Fig. 9.1 shows the impact of the different choices for the factorization and renormalization
 1510 scales on the jet and b-jet multiplicities in events with 2 SS leptons. Similar variations are
 1511 seen also in the other event categories. In order to not double-count the variations on the
 1512 total cross section the predictions from the different QCD scales are normalized to the same
 1513 total cross section. That means that the observed differences are only coming from the event
 selection. Significant variations on the jet multiplicities can be seen and these translate into

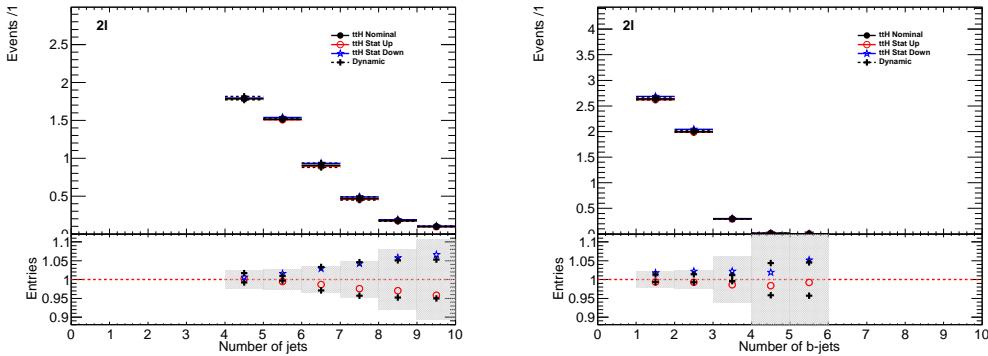


Figure 9.1: Effects on the jet multiplicities in 2 SS lepton $t\bar{t}H$ events from different choices of the factorization and renormalization scales. “Static” refers to the variations by a factor of 2 of the nominal μ_0 , while “Dynamic” refers to the alternative choice of μ_0 which depends on the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

1514
 1515 different predictions on the signal event yields in the signal regions. Such differences, listed in
 1516 Table 9.1, are taken as theoretical systematic uncertainties in addition to the ones affecting
 1517 the total $t\bar{t}H$ production cross section. The “Static” uncertainties come from the variations by
 1518 a factor of 2 from the nominal scale and they are correlated with the uncertainties on the total
 1519 cross section, which are estimated with the same procedure. The “Dynamic” uncertainties

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	$+2.7$ -1.3	$+2.3$ -0.8	$+0.9$ -0.2
Dynamic	$+1.7$ -0.8	$+2.0$ -2.6	$+1.7$ -1.1	$+0.5$ -0.0

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

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

1520 come from the difference between the nominal and the alternative dynamic scale and are
 1521 treated as an independent source of theoretical uncertainty.

1522 The uncertainty of the $t\bar{t}H$ event selection due to the PDF sets is estimated comparing the
 1523 predictions with three different PDF sets, varying each set within errors and taking the width
 1524 of the envelope as systematic uncertainty. The recommended sets are CT10, MSTW2008nlo68cl
 1525 and NNPDF21_100. We determine the change in the acceptance due to the PDF sets via the
 1526 formula 9.1 to disentangle it from the change in production cross section.

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

1527 Fig. 9.2 shows the estimated PDF systematic uncertainties as a function of the jet multiplicity
 1528 in $t\bar{t}H$ events with at least two leptons. The uncertainties are compatible with the uncertainty
 1529 on the production cross section estimated in Ref. [66] and indicated by the dashed red lines
 1530 in the lower panel. Table 9.2 shows the half-width of the envelope of the acceptance under all
 1531 eigenvector variations of the three PDF sets. No significant dependence on the event topology
 1532 is observed, so that the PDF systematic uncertainty on the $t\bar{t}H$ event selection is neglected.

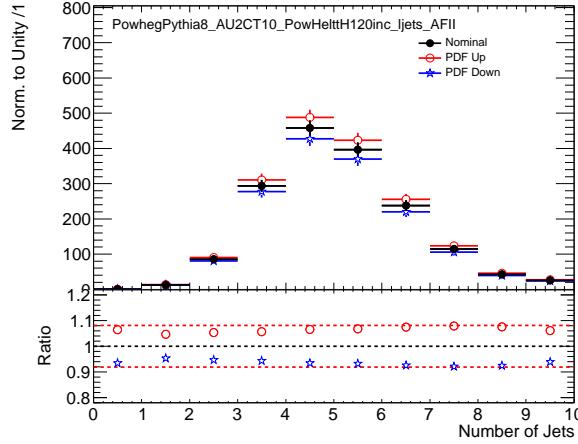


Figure 9.2: PDF systematic uncertainty on the jet multiplicities in $t\bar{t}H$ events with at least 2 leptons. The dashed red lines in the lower panel indicate the systematic uncertainty on the $t\bar{t}H$ production cross section.

1533 9.2 Experimental and Detector Systematic Uncertainties

1534 Experimental and detector systematic uncertainties arise from the efficiency of identifying
1535 objects and the efficiency of the event selections. These affect only MC models of physics
1536 processes, $t\bar{t}V$, $t\bar{t}H$, VV . Data-driven backgrounds take into account these effects by
1537 construction. We consider systematic effects from a number of sources: the lepton and jet
1538 energy scale measurements, the lepton identification and isolation selections, the efficiency
1539 and mis-identification rate associated with tagging b-quark jets. Effects due to modeling
1540 the energy and objects from additional vertices were studied and found to be negligible.
1541 The vast majority of the individual detector systematics effects are small. The sum total of
1542 the systematic effects are comparable to some of the overall normalization and cross-section
1543 uncertainties on some of the physics processes and is shown in Table ??.

1544 9.2.1 Lepton Identification, Energy Scale, and Trigger

1545 The electron[51] and muon identification efficiencies[82] are measured in data using Z boson
 1546 and J/Ψ control samples. They are shown in Figure 9.3. The uncertainty on the muon
 1547 efficiencies are measured as functions of η and p_T and are generally less 1%. The uncertainty
 1548 on the electron and muon efficiencies are also measured as functions of η and p_T and are at
 1549 the 1 % level for p_T above 30 GeV, but become much larger 5-10% for the lower p_T regimes.
 1550 These translate into sub-% level effects on the $t\bar{t}V$ and $t\bar{t}H$ event counts in the signal regions
 1551 for the muons and \sim % level effects for the electrons. The effects become more important
 1552 with increasing numbers of leptons.

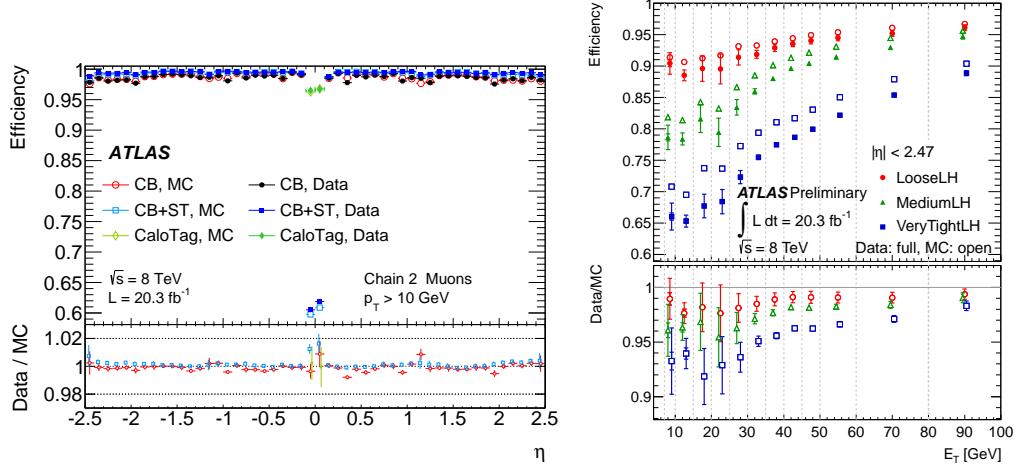


Figure 9.3: Muon (left) and electron(right) identification efficiencies in Data and MC as a function of η and p_T respectively. For electrons, the verytight likelihood operating point is used and for muons the CB+ST (combined+segment tagged) operating point is used

1553 The electron[83] and muon[82] energy scale and resolution are also measured using the
 1554 Z -boson control samples in data. The uncertainties related to the scale and resolution for the

1555 leptons affect the overall event acceptance through the lepton momentum cuts primary and
1556 have negligible impact on the event count uncertainties in the signal regions.

1557 The efficiencies for muons and electrons to pass muon[84] and electron triggers[85] have
1558 been calculated with respect to the offline identification operating points using the Z boson
1559 control samples. They are in the range of 90% for electron triggers and 70% for muon triggers,
1560 owing to gaps in muon trigger coverage, and have % level errors. When statistically combined
1561 for 2,3,4 and lepton signal regions, the overall trigger efficiency is high and the error on the
1562 number of expected events is negligible.

1563 **9.2.2 Lepton Isolation and Impact Parameter**

1564 The isolation and impact parameter selections are specific to this analysis and are discussed
1565 in depth in Chapter 7. We calculated their combined efficiency with respect to the full
1566 lepton identification selection using the Z boson control samples and define data-MC scale
1567 factors to correct the efficiency in the simulation. Background are subtracted using shape
1568 templates in the di-lepton invariant mass spectrum. The electron template is derived from
1569 MC, while the background template is derived from the same-sign control region, with certain
1570 object cuts reversed in the electron case. We measure the efficiency scale-factors in bins of
1571 lepton momentum. Uncertainties are assigned per-bin to account for the level of statistics and
1572 variations caused by the fit parameters. An additional 1% uncertainty envelope is added to
1573 both the electron and muon measurements to account for trends observed in the dependence
1574 of the data-MC efficiency scale-factor as a function of the number of jets. Stability of the
1575 efficiency scale-factor as a function of the number of jets is important for this analysis, because
1576 event activity in the low jet Z sample where the efficiency is measured is much different from
1577 the high jet signal regions where the efficiencies are applied. The dependence of the scale-
1578 factor on the number of jets can be seen Figure 9.4. The isolation scale-factor uncertainties

1579 are around 1-3% depending on the particle momentum, but these uncertainties propagate to
 1580 2-5 % (some of the largest) effects in the event counts in the signal regions. The uncertainties
 1581 are more important in the regions with more leptons.

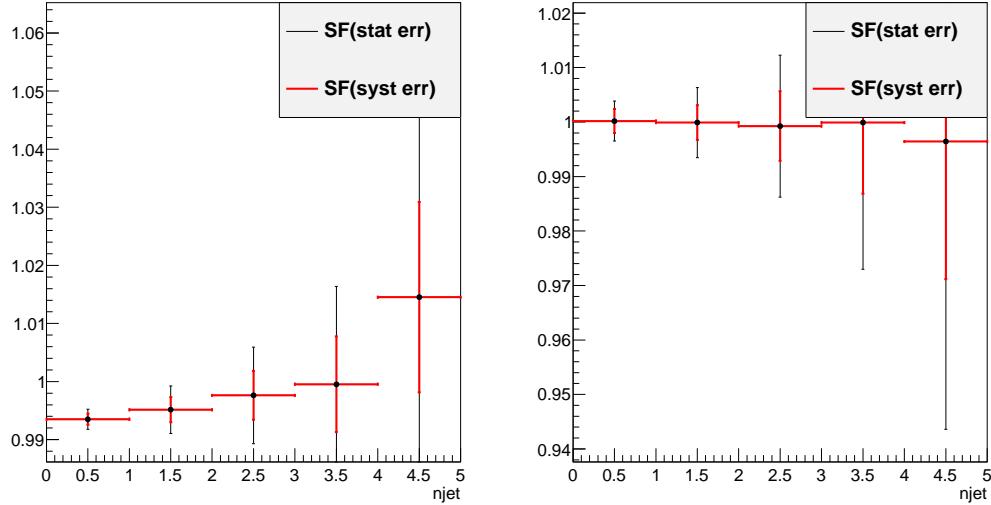


Figure 9.4: Muon (left) and electron(right) isolation efficiency scale-factors from the Z control sample as a function of the number of jets in the event. An additional systematic uncertainty of 1% is added to encompass the variation in the number of jets variable

1582 9.2.3 Jet Energy

1583 The jet energy scale (JES) is calculated using a combination of data-based in-situ techniques,
 1584 where jet transverse momentum is balanced with respect to a reference photon or a Z boson, as
 1585 well as single particle test-stand studies[86]. Additional smaller effects are taken into account
 1586 including the b-quark jet specific response, near-by jets, the effects of pile-up and an inter-
 1587 calibration of similar *eta* regions using di-jet events. The JES systematic errors arises from
 1588 numerous sources that are diagonalized into eigenvectors so that they can be combined in an
 1589 uncorrelated way. The combined uncertainty is plotted in Figure 9.5 as a function of jet η and

1590 p_T and is the range 2-4% for jets used in this analysis. The jet energy resolution is calculated
 1591 in a similar way with slightly larger errors, 10% [87]. Propagated to the event counts in
 1592 the signal regions, the combined scale and resolution systematics are of non-negligible effects
 1593 6-7%.

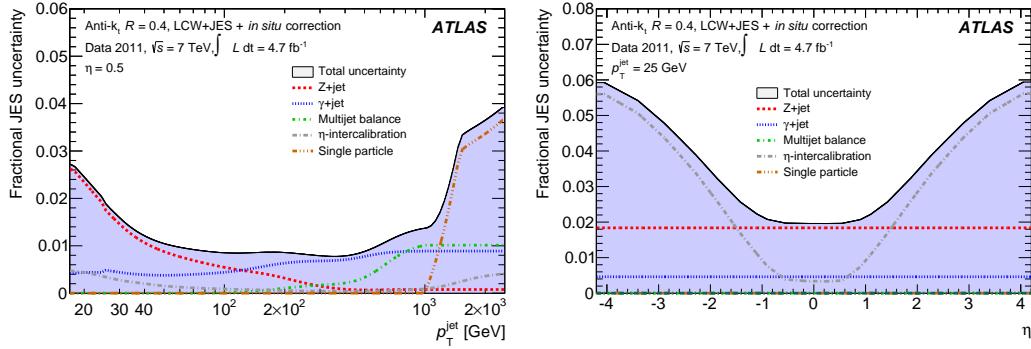


Figure 9.5: JES systematic uncertainties as a function of jet η (for jets $p_T > 25$ GeV) and p_T (for jets $|\eta| < 0.4$). The combined systematic uncertainty is shown with contributions from the largest sources

1594 9.2.4 B-Tagged Jet Efficiency

1595 The b -quark tagging efficiency must be calculated separately for charm, light and b -quarks.
 1596 ATLAS uses three data based control regions: an inclusive jet sample for mis-tagged light
 1597 quarks[88], the $t\bar{t}$ sample for b -quarks[89], and a sample of D^* mesons for charm quarks[90].
 1598 These efficiencies and rates are well-measured in MC and the data-based corrections are
 1599 small. The data-MC efficiency scale-factor shown in Figure 9.6 is close to 1 and has an
 1600 overall systematic uncertainty of around 5%. The uncertainties are applied to the analysis
 1601 via a number of eigenvectors. Together these uncertainties have a 4 % effect in the event
 1602 expectation in the signal regions.

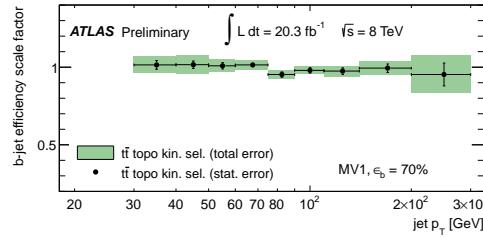


Figure 9.6: b-Tagging data-MC efficiency scale-factors versus jet p_T calculated in the $t\bar{t}$ sample from 2012 data. The uncertainties are combined statistical and systematic.

Total Systematic Uncertainty	2ee4j		2ee5jincl		2em4j		2em5jincl	
	Down	Up	Down	Up	Down	Up	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		2mm5jincl		3ℓ		4ℓ	
	Down	Up	Down	Up	Down	Up	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.

1603 9.2.5 Summary

1604 The combined effect of these detector and experimental systematics on the $t\bar{t}V$ and $t\bar{t}H$ is
 1605 provided in Table 9.3. The effects are smaller than the normalization uncertainties on some
 1606 of the backgrounds. However, since they effect all processes signal and background. They
 1607 are dominated by the lepton isolation scale-factor measurements and the electron identifi-
 1608 cation with smaller contributions from the JES and b-tagging efficiencies. These detector
 1609 systematic uncertainties enter the fit individually and their ranking of influence on the overall
 1610 measurement uncertainty can be seen in Figure??.

¹⁶¹¹ **9.3 Summary of Background and Signal Normalization**

¹⁶¹² **Uncertainties**

¹⁶¹³ Tab.9.4 gives the summary of the systematic uncertainties that are included in the analy-
¹⁶¹⁴ sis for the normalization and acceptance of each process. The relative importance of these
¹⁶¹⁵ 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.

1616

CHAPTER 10

1617

Results and Statistical Model

1618

10.1 Results in Signal Regions

1619

10.2 Statistical Model

1620

NEED TO DEFINE MU IN ANALYSIS SUMMARY CHAPTER

1621

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.

1627

10.2.1 The Likelihood

1628

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:

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

1631 where s_{exp}^i is the SM signal expectation in the signal region, b_{exp}^i are the background expecta-
 1632 tions, i counts over the signal regions, and P is the Poisson distribution. The signal strength
 1633 parameter is the parameter of interest in the model (POI) and acts as a simple scale-factor
 1634 to the SM $t\bar{t}H$ production rate and is common to all channels. Setting μ to 0 corresponds to
 1635 the background only scenario. The background parameter, b , is a sum over all background
 1636 processes.

1637 The signal and background expectations, s and b , depend on systematic errors. These are
 1638 included in the likelihood function in the form of a vector nuisance parameters, $\vec{\theta}$, which are
 1639 constrained to fluctuate within Gaussian distributions. These fluctuations affect the back-
 1640 ground and signal expectations by response functions, $\nu(\vec{\theta})$, set by uncertainties measured
 1641 in the previous section. For instance, the $W^\pm Z$ normalization uncertainty is 50% from Sec-
 1642 tion 8.2 and is included in the fit as its own unit gaussian, $G(\theta|0, 1)$. The fluctuations of the
 1643 gaussian, θ_{WZ} scale the background contribution via the form, $0.5 \cdot (1 + \theta_{WZ}) \cdot b_{WZ}$. For
 1644 many of the detector systematics, the uncertainties are two sided and are included as piece-
 1645 wise Gaussians. We add correlations to various uncertainties by hand, when appropriate.

1646 With these nuisance parameters, the likelihood takes this form:

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

1647 10.2.2 Test Statistic and Profile Likelihood

1648 Values of μ are tested with the negative log quantity, $q_\mu = -2\ln(\lambda(\mu))$, where $\lambda(\mu)$ is the test
 1649 statistic. $\lambda(\mu)$ is defined as:

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

1650 where $\hat{\vec{\theta}}_\mu$ are values of the nuisance parameter vector that maximize the likelihood for a given
 1651 value of μ and $\hat{\mu}$ and $\hat{\vec{\theta}}$ are the fitted values of signal strength and nuisance parameters that

1652 maximize the likelihood overall. μ is constrained to be positive.

1653 **10.2.3 CL_s Method**

1654 Exclusions limits on the signal strength are calculated with the test statistic using a modified
 1655 frequentist method, called the CL_s method[92]. CL_s is defined as a ratio of two frequentist
 1656 quantities. The numerator quantifies the probability of finding the observed data given the
 1657 signal + background hypothesis. The denominator quantifies the probability of the data given
 1658 the background only hypothesis.

1659 Using the numerator alone has the undesirable property that, if the data fluctuates below
 1660 the expectation, an exclusion limit can be reached that is far beyond the sensitivity of the
 1661 experiment. Normalizing to the background only hypothesis penalizes these low sensitivity
 1662 cases.

1663 The probability of obtaining an observation as extreme as the data given a particular signal
 1664 + background hypothesis is given by the p-value, p_μ defined as:

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

1665 and the probability of obtaining an observation as extreme as the data given the background
 1666 hypothesis, p_b is:

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

1667 where $f(q_\mu)$ is the distribution of q_μ for all possible observations for a given μ and q is defined
 1668 above. Therefore,

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

1669 . A μ value is considered excluded at 95% confidence when CL_s is less than 5%.

1670 **10.2.4 Exclusion Limits**

1671 Table ?? shows expected exclusion limits for all channels, including the analysis uncertainties
 1672 cumulatively. It shows the relative importance of the statistical and systematic uncertainties
 1673 to the analysis sensitivity. The observed limits using observed data and predictions can be seen
 1674 in Figures ??-?? for splitting and combining the sub-channels and in Table XX by numbers.
 1675 We expect a combined limit of XXX (background only) and XXX (signal injected) and see a
 1676 limit of XXX. The channel sensitivity is dominated by the 2ℓ and 3ℓ channels.

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

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

1677 **10.2.5 μ Measurements**

1678 In addition to setting a limit on the signal strength, we also fit the best value of the signal
 1679 strength for μ . We do this by minimizing the negative log likelihood value, q_μ or conversely
 1680 maximizing the likelihood. The $1-\sigma$ error band is set via a profile likelihood scan, where the
 1681 value q_μ is scanned as a function of μ . Values of μ that increase q_{min}^m by 1 form the edges
 1682 of the error band. The fitted values of μ with errors are provided in Table XXX for each
 1683 sub-channel fit as well as the combined fit.

1684 10.2.6 Nuisance Parameter Impact on the Signal Strength

1685 Finally, we examine the post-fit impact of the various nuisance parameters on the final fit. We
1686 expect to have measured the various analysis uncertainties well and do not expect the fit to
1687 have much further constraint. For that reason, we expect the pulls of the nuisance parameters
1688 to be close to 0 and the measured uncertainties on those parameters to be consistent with the
1689 input uncertainties. Figures XXXX show.

1690

CHAPTER 11

1691

Conclusions

1692 **11.1 Higgs Results in Review**

1693 **11.2 Prospects for Future**

Bibliography

- 1695 [1] S. L. Glashow, *Partial-symmetries of weak interactions*, Nucl. Phys. **22** (1961) no. 4,
1696 579.
- 1697 [2] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.
- 1698 [3] A. Salam and J. C. Ward, *Gauge theory of elementary interactions*, Phys. Rev. **136**
1699 (1964) 763–768.
- 1700 [4] S. Weinberg, *Non-abelian gauge theories of the strong interactions*, Phys. Rev. Lett. **31**
1701 (1973) 494–497.
- 1702 [5] D. J. Gross and F. Wilczek, *Ultraviolet behavior of non-abelian gauge theories*, Phys.
1703 Rev. Lett. **30** (1973) 1343–1346.
- 1704 [6] G. 't Hooft and M. Veltman, *Regularization and renormalization of gauge fields*,
1705 Nuclear Physics B **44** (1972) 189 – 213.
- 1706 [7] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13**
1707 (1964) 508.
- 1708 [8] P. W. Higgs, *Spontaneous symmetry breakdown without massless bosons*, Phys. Rev.
1709 **145** (1966) 1156.
- 1710 [9] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*,
1711 Phys. Rev. Lett. **13** (1964) 321–322.
- 1712 [10] The ALEPH, CDF, DØ, DELPHI, L3, OPAL, SLD Collaborations, the LEP
1713 Electroweak Working Group, the Tevatron Electroweak Working Group, and the SLD
1714 electroweak and heavy flavour groups, *Precision Electroweak Measurements and*
1715 *Constraints on the Standard Model*, CERN-PH-EP-2010-095 (2010) , arXiv:1012.2367
1716 [hep-ex].
- 1717 [11] M. Baak, M. Goebel, J. Haller, A. Hoecker, D. Kennedy, R. Kogler, K. Mnig,
1718 M. Schott, and J. Stelzer, *The electroweak fit of the standard model after the discovery*
1719 *of a new boson at the LHC*, The European Physical Journal C **72** (2012) no. 11, .
1720 <http://dx.doi.org/10.1140/epjc/s10052-012-2205-9>.

- 1721 [12] J. C. Collins, D. E. Soper, and G. Sterman, *Factorization for short distance*
 1722 *hadron-hadron scattering*, Nuclear Physics B **261** (1985) 104 – 142.
- 1723 [13] CERN, . CERN, Geneva, 1984.
- 1724 [14] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and
 1725 R. Tanaka (Eds.), *Handbook of LHC Higgs Cross Sections: 2. Differential Distributions*,
 1726 CERN-2012-002 (CERN, Geneva, 2012) , arXiv:1201.3084 [hep-ph].
- 1727 [15] *Updated coupling measurements of the Higgs boson with the ATLAS detector using up to*
 1728 *25 fb⁻¹ of proton-proton collision data*, Tech. Rep. ATLAS-CONF-2014-009, CERN,
 1729 Geneva, Mar, 2014.
- 1730 [16] CMS Collaboration Collaboration, *Precise determination of the mass of the Higgs boson*
 1731 *and studies of the compatibility of its couplings with the standard model*, Tech. Rep.
 1732 CMS-PAS-HIG-14-009, CERN, Geneva, 2014.
- 1733 [17] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of the Higgs boson*
 1734 *mass from the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4\ell$ channels with the ATLAS detector using 25*
 1735 *fb⁻¹ of pp collision data*, arXiv:1406.3827 [hep-ex].
- 1736 [18] *Evidence for the spin-0 nature of the Higgs boson using {ATLAS} data*, Physics Letters
 1737 B **726** (2013) no. 13, 120 – 144.
 1738 <http://www.sciencedirect.com/science/article/pii/S0370269313006527>.
- 1739 [19] S. Dawson, A. Gritsan, H. Logan, J. Qian, C. Tully, et al., *Working Group Report:*
 1740 *Higgs Boson*, arXiv:1310.8361 [hep-ex].
- 1741 [20] O. Eberhardt, G. Herbert, H. Lacker, A. Lenz, A. Menzel, et al., *Impact of a Higgs*
 1742 *boson at a mass of 126 GeV on the standard model with three and four fermion*
 1743 *generations*, Phys.Rev.Lett. **109** (2012) 241802, arXiv:1209.1101 [hep-ph].
- 1744 [21] M. Carena, S. Gori, N. R. Shah, C. E. Wagner, and L.-T. Wang, *Light Stops, Light*
 1745 *Staus and the 125 GeV Higgs*, JHEP **1308** (2013) 087, arXiv:1303.4414.
- 1746 [22] N. Arkani-Hamed, K. Blum, R. T. D’Agnolo, and J. Fan, *2:1 for Naturalness at the*
 1747 *LHC?*, JHEP **1301** (2013) 149, arXiv:1207.4482 [hep-ph].
- 1748 [23] D. Carmi, A. Falkowski, E. Kuflik, and T. Volansky, *Interpreting LHC Higgs Results*
 1749 *from Natural New Physics Perspective*, JHEP **1207** (2012) 136, arXiv:1202.3144
 1750 [hep-ph].
- 1751 [24] G. Degrassi, S. Di Vita, J. Elias-Miro, J. R. Espinosa, G. F. Giudice, et al., *Higgs mass*
 1752 *and vacuum stability in the Standard Model at NNLO*, JHEP **1208** (2012) 098,
 1753 arXiv:1205.6497 [hep-ph].
- 1754 [25] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) no. 08, S08001.
- 1755 [26] T. S. Pettersson and P. Lefevre, *The Large Hadron Collider: conceptual design.*, Tech.
 1756 Rep. CERN-AC-95-05 LHC, CERN, Geneva, Oct, 1995.
 1757 <https://cdsweb.cern.ch/record/291782>.

- 1758 [27] T. Linnecar et al., *Hardware and Initial Beam Commissioning of the LHC RF Systems.*
1759 *oai:cds.cern.ch:1176380*, Tech. Rep. LHC-PROJECT-Report-1172.
1760 CERN-LHC-PROJECT-Report-1172, CERN, Geneva, Oct, 2008.
1761 <https://cdsweb.cern.ch/record/1176380>.
- 1762 [28] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*,
1763 JINST **3** (2008) S08003.
- 1764 [29] The CMS Collaboration, *The CMS experiment at the CERN LHC*, Journal of
1765 Instrumentation **3** (2008) no. 08, S08004.
- 1766 [30] The LHCb Collaboration, *The LHCb Detector at the LHC*, Journal of Instrumentation
1767 **3** (2008) no. 08, S08005.
- 1768 [31] The ALICE Collaboration, *The ALICE experiment at the CERN LHC*, Journal of
1769 Instrumentation **3** (2008) no. 08, S08002.
1770 <http://stacks.iop.org/1748-0221/3/i=08/a=S08002>.
- 1771 [32] A. Team, *The four main LHC experiments*, Jun, 1999.
- 1772 [33] ATLAS Collaboration, *ATLAS inner detector: Technical Design Report 1*. Technical
1773 Design Report ATLAS. CERN, Geneva, 1997.
1774 <https://cdsweb.cern.ch/record/331063>.
- 1775 [34] ATLAS Collaboration, *ATLAS inner detector: Technical Design Report, 2*. Technical
1776 Design Report ATLAS. CERN, Geneva, 1997.
1777 <https://cdsweb.cern.ch/record/331064>.
- 1778 [35] ATLAS Collaboration, *ATLAS magnet system: Technical Design Report, 1*. Technical
1779 Design Report ATLAS. CERN, Geneva, 1997.
1780 <https://cdsweb.cern.ch/record/338080>.
- 1781 [36] ATLAS Collaboration, *ATLAS pixel detector: Technical Design Report*. Technical
1782 Design Report ATLAS. CERN, Geneva, 1998.
1783 <https://cdsweb.cern.ch/record/381263>.
- 1784 [37] ATLAS Collaboration, *ATLAS pixel detector electronics and sensors*, JINST **3** (2008)
1785 P07007.
- 1786 [38] ATLAS Collaboration, *The barrel modules of the ATLAS semiconductor tracker*,
1787 Nucl.Instrum.Meth. **A568** (2006) 642–671.
- 1788 [39] ATLAS Collaboration, *The ATLAS semiconductor tracker end-cap module*,
1789 Nucl.Instrum.Meth. **A575** (2007) 353–389.
- 1790 [40] The ATLAS TRT Collaboration, *The ATLAS Transition Radiation Tracker (TRT)*
1791 *proportional drift tube: Design and performance*, JINST **3** (2008) P02013.
- 1792 [41] The ATLAS TRT Collaboration, *The ATLAS TRT barrel detector*, JINST **3** (2008)
1793 P02014.
- 1794 [42] E. Abat et al., *The ATLAS TRT electronics*, J. Instrum. **3** (2008) P06007.
- 1795 [43] T. A. Collaboration, *ATLAS liquid argon calorimeter: Technical design report*, .
1796 CERN-LHCC-96-41.

- 1797 [44] ATLAS Collaboration, *ATLAS tile calorimeter: Technical Design Report*. Technical
 1798 Design Report ATLAS. CERN, Geneva, 1996.
 1799 <https://cdsweb.cern.ch/record/331062>.
- 1800 [45] ATLAS Collaboration, *ATLAS muon spectrometer: Technical Design Report*. Technical
 1801 Design Report ATLAS. CERN, Geneva, 1997.
 1802 <https://cdsweb.cern.ch/record/331068>.
- 1803 [46] G. Aielli, A. Aloisio, M. Alviggi, V. Aprodu, V. Bocci, et al., *The RPC first level muon
 1804 trigger in the barrel of the ATLAS experiment*, Nucl.Phys.Proc.Suppl. **158** (2006) 11–15.
- 1805 [47] F. Bauer, U. Bratzler, H. Dietl, H. Kroha, T. Lagouri, et al., *Construction and test of
 1806 MDT chambers for the ATLAS muon spectrometer*, Nucl.Instrum.Meth. **A461** (2001)
 1807 17–20.
- 1808 [48] T. Argyropoulos, K. A. Assamagan, B. H. Benedict, V. Chernyatin, E. Cheu, et al.,
 1809 *Cathode strip chambers in ATLAS: Installation, commissioning and in situ
 1810 performance*, IEEE Trans.Nucl.Sci. **56** (2009) 1568–1574.
- 1811 [49] ATLAS Collaboration, *ATLAS level-1 trigger: Technical Design Report*. Technical
 1812 Design Report ATLAS. CERN, Geneva, 1998.
 1813 <https://cdsweb.cern.ch/record/381429>.
- 1814 [50] P. Jenni, M. Nessi, M. Nordberg, and K. Smith, *ATLAS high-level trigger,
 1815 data-acquisition and controls: Technical Design Report*. Technical Design Report
 1816 ATLAS. CERN, Geneva, 2003. <https://cdsweb.cern.ch/record/616089>.
- 1817 [51] *Electron efficiency measurements with the ATLAS detector using the 2012 LHC
 1818 proton-proton collision data*, Tech. Rep. ATLAS-CONF-2014-032, CERN, Geneva, Jun,
 1819 2014.
- 1820 [52] *Preliminary results on the muon reconstruction efficiency, momentum resolution, and
 1821 momentum scale in ATLAS 2012 pp collision data*, Tech. Rep. ATLAS-CONF-2013-088,
 1822 CERN, Geneva, February, 2013.
- 1823 [53] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti- $k(t)$ jet clustering algorithm*, JHEP
 1824 **0804** (2008) 063, arXiv:0802.1189 [hep-ph].
- 1825 [54] ATLAS Collaboration, *Commissioning of the ATLAS high-performance b-tagging
 1826 algorithms in the 7 TeV collision data*, Tech. Rep. ATLAS-CONF-2011-102, CERN,
 1827 Geneva, Jul, 2011. <https://cdsweb.cern.ch/record/1369219>.
- 1828 [55] ATLAS Collaboration Collaboration, G. Aad et al., *Search for $H \rightarrow \gamma\gamma$ produced in
 1829 association with top quarks and constraints on the Yukawa coupling between the top
 1830 quark and the Higgs boson using data taken at 7 TeV and 8 TeV with the ATLAS
 1831 detector*, arXiv:1409.3122 [hep-ex].
- 1832 [56] *Search for the Standard Model Higgs boson produced in association with top quarks and
 1833 decaying to $b\bar{b}$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector at the LHC*,
 1834 Tech. Rep. ATLAS-CONF-2014-011, CERN, Geneva, Mar, 2014.
- 1835 [57] ATLAS Collaboration, G. Aad et al., *Improved luminosity determination in pp
 1836 collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC*, Eur.Phys.J. **C73**
 1837 (2013) 2518, arXiv:1302.4393 [hep-ex].

- 1838 [58] CERN, . CERN, Geneva, 2012.
- 1839 [59] ATLAS Collaboration Collaboration, G. Aad et al., *The ATLAS Simulation*
1840 *Infrastructure*, Eur.Phys.J. **C70** (2010) 823–874, arXiv:1005.4568 [physics.ins-det].
- 1841 [60] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A Simulation toolkit*,
1842 Nucl.Instrum.Meth. **A506** (2003) 250–303.
- 1843 [61] G. Bevilacqua, M. Czakon, M. Garzelli, A. van Hameren, A. Kardos, C. Papadopoulos,
1844 R. Pittau, and M. Worek.
- 1845 [62] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo*
1846 *algorithms*, **11** (2004) 040.
- 1847 [63] S. Frixione, P. Nason, and C. Oleari **11** (2007) 070, arXiv:0709.2092 [hep-ph].
- 1848 [64] S. Alioli, P. Nason, C. Oleari, and E. Re **06** (2010) 040, arXiv:1002.2581 [hep-ph].
- 1849 [65] T. Sjöstrand, S. Mrenna, and P. Skands, *A Brief Introduction to Pythia 8.1*,
1850 arXiv:0710.3820 [hep-ph].
- 1851 [66] LHC Higgs Cross Section Working Group Collaboration, S. Heinemeyer et al.,
1852 *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, arXiv:1307.1347 [hep-ph].
- 1853 [67] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *ALPGEN, a*
1854 *generator for hard multiparton processes in hadronic collisions*, JHEP **0307** (2003) 001,
1855 arXiv:hep-ph/0206293 [hep-ph].
- 1856 [68] F. Maltoni and T. Stelzer, *MadEvent: Automatic event generation with MadGraph*,
1857 JHEP **0302** (2003) 027, arXiv:hep-ph/0208156 [hep-ph].
- 1858 [69] B. P. Kersevan and E. Richter-Was, *The Monte Carlo event generator AcerMC versions*
1859 *2.0 to 3.8 with interfaces to PYTHIA 6.4, HERWIG 6.5 and ARIADNE 4.1*,
1860 Comput.Phys.Commun. **184** (2013) 919–985, arXiv:hep-ph/0405247 [hep-ph].
- 1861 [70] P. M. Nadolsky, H.-L. Lai, Q.-H. Cao, J. Huston, J. Pumplin, et al., *Implications of*
1862 *CTEQ global analysis for collider observables*, Phys.Rev. **D78** (2008) 013004,
1863 arXiv:0802.0007 [hep-ph].
- 1864 [71] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton*
1865 *Shower simulations: the POWHEG method*, JHEP **0711** (2007) 070, arXiv:0709.2092
1866 [hep-ph].
- 1867 [72] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, et al., *Event*
1868 *generation with SHERPA 1.1*, JHEP **0902** (2009) 007, arXiv:0811.4622 [hep-ph].
- 1869 [73] A. Martin, W. Stirling, R. Thorne, and G. Watt, *Parton distributions for the LHC*,
1870 Eur.Phys.J. **C63** (2009) 189–285, arXiv:0901.0002 [hep-ph].
- 1871 [74] M. Garzelli, A. Kardos, C. Papadopoulos, and Z. Trocsanyi, *$t\bar{t}W^\pm$ and $t\bar{t}Z$*
1872 *Hadroproduction at NLO accuracy in QCD with Parton Shower and Hadronization*
1873 *effects*, JHEP **1211** (2012) 056, arXiv:1208.2665 [hep-ph].
- 1874 [75] J. M. Campbell and R. K. Ellis, *$t\bar{t}W^{+-}$ production and decay at NLO*, JHEP **1207**
1875 (2012) 052, arXiv:1204.5678 [hep-ph].

- 1876 [76] Campbell, John and Ellis, R. Keith and Röntsch, Raoul, *Single top production in*
 1877 *association with a Z boson at the LHC*, Phys.Rev. **D87** (2013) 114006, arXiv:1302.3856
 1878 [hep-ph].
- 1879 [77] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of WZ production in*
 1880 *proton-proton collisions at $\sqrt{s} = 7 \text{ TeV}$ with the ATLAS detector*, Eur.Phys.J. **C72**
 1881 (2012) 2173, arXiv:1208.1390 [hep-ex].
- 1882 [78] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of ZZ production in*
 1883 *pp collisions at $\sqrt{s} = 7 \text{ TeV}$ and limits on anomalous ZZZ and ZZ γ couplings with the*
 1884 *ATLAS detector*, JHEP **1303** (2013) 128, arXiv:1211.6096 [hep-ex].
- 1885 [79] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of the cross-section*
 1886 *for W boson production in association with b-jets in pp collisions at $\sqrt{s} = 7 \text{ TeV}$ with*
 1887 *the ATLAS detector*, JHEP **1306** (2013) 084, arXiv:1302.2929 [hep-ex].
- 1888 [80] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of differential*
 1889 *production cross-sections for a Z boson in association with b-jets in 7 TeV*
 1890 *proton-proton collisions with the ATLAS detector*, arXiv:1407.3643 [hep-ex].
- 1891 [81] S. Guindon, E. Shabalina, J. Adelman, M. Alhroob, S. Amor dos Santos, A. Basye,
 1892 J. Bouffard, M. Casolino, I. Connelly, A. Cortes Gonzalez, V. Dao, S. D'Auria,
 1893 A. Doyle, P. Ferrari, F. Filthaut, R. Goncalo, N. de Groot, S. Henkelmann, V. Jain,
 1894 A. Juste, G. Kirby, D. Kar, A. Knue, K. Kroeninger, T. Liss, E. Le Menedeu,
 1895 J. Montejo Berlingen, M. Moreno Llacer, O. Nackenhorst, T. Neep, A. Onofre,
 1896 M. Owen, M. Pinamonti, Y. Qin, A. Quadt, D. Quilty, C. Schwanenberger, L. Serkin,
 1897 R. St Denis, J. Thomas-Wilsker, and T. Vazquez-Schroeder, *Search for the Standard*
 1898 *Model Higgs boson produced in association with top quarks and decaying to $b\bar{b}$ in pp*
 1899 *collisions at $\sqrt{s} = 8 \text{ TeV}$ with the ATLAS detector at the LHC*, Tech. Rep.
 1900 ATL-COM-PHYS-2013-1659, CERN, Geneva, Dec, 2013. The note contains internal
 1901 documentation of the ttH(bb) analysis approved as a preliminary result
 1902 (ATLAS-CONF-2014-011).
- 1903 [82] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of the muon*
 1904 *reconstruction performance of the ATLAS detector using 2011 and 2012 LHC*
 1905 *proton-proton collision data*, arXiv:1407.3935 [hep-ex].
- 1906 [83] ATLAS Collaboration Collaboration, G. Aad et al., *Electron and photon energy*
 1907 *calibration with the ATLAS detector using LHC Run 1 data*, Eur.Phys.J. **C74** (2014)
 1908 no. 10, 3071, arXiv:1407.5063 [hep-ex].
- 1909 [84] ATLAS Collaboration Collaboration, *Performance of the ATLAS muon trigger in 2011*,
 1910 Tech. Rep. ATLAS-CONF-2012-099, CERN, Geneva, Jul, 2012.
- 1911 [85] ATLAS Collaboration Collaboration, *Performance of the ATLAS Electron and Photon*
 1912 *Trigger in p-p Collisions at $\sqrt{s} = 7 \text{ TeV}$ in 2011*, Tech. Rep.
 1913 ATLAS-CONF-2012-048, CERN, Geneva, May, 2012.
- 1914 [86] ATLAS Collaboration Collaboration, G. Aad et al., *Jet energy measurement and its*
 1915 *systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7 \text{ TeV}$ with the ATLAS*
 1916 *detector*, arXiv:1406.0076 [hep-ex].

- 1917 [87] ATLAS Collaboration Collaboration, G. Aad et al., *Jet energy resolution in*
1918 *proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector*,
1919 Eur.Phys.J. **C73** (2013) 2306, arXiv:1210.6210 [hep-ex].
- 1920 [88] ATLAS Collaboration Collaboration, *Measurement of the Mistag Rate with 5 fb^{-1} of*
1921 *Data Collected by the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2012-040, CERN,
1922 Geneva, 2012.
- 1923 [89] ATLAS Collaboration Collaboration, *Measuring the b-tag efficiency in a top-pair*
1924 *sample with 4.7 fb^{-1} of data from the ATLAS detector*, Tech. Rep.
1925 ATLAS-CONF-2012-097, CERN, Geneva, 2012.
- 1926 [90] ATLAS Collaboration Collaboration, *b-jet tagging calibration on c-jets containing $D^?+$*
1927 *mesons*, Tech. Rep. ATLAS-CONF-2012-039, CERN, Geneva, 2012.
- 1928 [91] Glen Cowan, Kyle Cranmer, Eilam Gross, Ofer Vitells, *Asymptotic formulae for*
1929 *likelihood-based tests of new physics*, Eur.Phys.J.C **71** (2011) 1554.
- 1930 [92] A. L. Read, *Presentation of search results: the CL s technique*, Journal of Physics G:
1931 Nuclear and Particle Physics **28** (2002) no. 10, 2693.
1932 <http://stacks.iop.org/0954-3899/28/i=10/a=313>.