

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

Chris Lester

A DISSERTATION

in

Physics and Astronomy

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

I. Joseph Kroll, Professor, Physics
Supervisor of Dissertation

Marija Drndic, Professor, Physics
Graduate Group Chairperson

Dissertation Committee

Justin Khoury, Associate Professor, Physics

Elliot Lipeles, Associate Professor, Physics

Burt Ovrut, Professor, Physics

Evelyn Thompson, Associate Professor, Physics

I. Joseph Kroll, Professor, Physics

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

COPYRIGHT
2014
Chris Lester

All rights reserved.

Acknowledgements

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

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

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

ABSTRACT

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

Chris Lester

Joseph Kroll

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

Contents

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

82	3.1.2 Beam Parameters and Collisions	15
83	3.2 The ATLAS Experiment	15
84	3.2.1 Detector Coordinate System	17
85	3.2.2 The Inner Detector	17
86	3.2.3 The Calorimeter	19
87	3.2.4 The Muon Spectrometer	22
88	3.2.5 The Trigger System	22
89	3.2.6 Reconstruction: Jets, Muons and Electrons	24
90	3.2.6.1 Tracks and Clusters	24
91	3.2.6.2 Electrons	25
92	3.2.6.3 Muons	25
93	3.2.6.4 Jets	25
94	3.2.6.5 B-Tagged Jets	26
95	4 Electrons	28
96	4.1 Identification of Electrons at ATLAS	28
97	4.1.1 2011 Menu	31
98	4.1.2 2012 Menu and Pile-up	32
99	4.1.3 Electron Likelihood	33
100	4.2 Measurement of Electron Identification Efficiency at ATLAS	35
101	5 Analysis Summary	39
102	5.1 Signal Characteristics	39
103	5.2 Background Overview	40
104	5.3 Analysis Strategy	41
105	6 Dataset and Simulation	42
106	6.1 Data	42
107	6.1.1 The 2012 Dataset	42
108	6.2 Simulation	43
109	6.2.1 Signal Simulation	44
110	6.2.2 Background Simulation	44
111	7 Object and Event Selection	46
112	7.1 2ℓ Same-Charge Signal Region	46

113	7.2	3 ℓ Signal Region	48
114	7.3	4 ℓ Signal Region	48
115	7.4	Electron Selection	49
116	7.5	Muon Selection	49
117	7.6	Jet and b-Tagged Jet Selection	50
118	7.7	Object Summary and Overlap	51
119	8	Background Estimation	52
120	8.1	Vector Boson (W^\pm, Z) production in association with top quarks: $t\bar{t}V, tZ$	52
121	8.1.1	$t\bar{t}Z$ Validation Region	54
122	8.2	Di-boson Background Estimation: $W^\pm Z, ZZ$	56
123	8.2.1	$W^\pm Z$ Normalization Uncertainty	57
124	8.2.2	ZZ Normalization Uncertainty	58
125	8.3	Charge-Misidentification Background	61
126	8.3.1	Likelihood Method	62
127	8.3.2	Results	63
128	8.3.3	Systematic and Statistical Uncertainties	64
129	8.4	Fake Lepton Backgrounds	66
130	8.4.1	2 ℓ SS Fakes	67
131	8.4.2	3 ℓ Fakes	69
132	8.4.3	4 ℓ Fakes	74
133	9	Summary of Systematic Uncertainties	76
134	9.1	Systematic Uncertainties on Signal Cross-section and Acceptance	76
135	9.2	Experimental and Detector Systematic Uncertainties	78
136	9.2.1	Lepton Identification, Energy Scale, and Trigger	79
137	9.2.2	Lepton Isolation and Impact Parameter	80
138	9.2.3	Jet Energy	80
139	9.2.4	B-Tagged Jet Efficiency	82
140	9.2.5	Summary	82
141	9.3	Summary of Background and Signal Normalization Uncertainties	82
142	10	Results and Statistical Model	85
143	10.1	Results in Signal Regions	85

144	10.1.1 2ℓ SS	85
145	10.1.2 3ℓ	85
146	10.1.3 4ℓ	85
147	10.2 Statistical Model	85
148	10.2.1 The Likelihood	85
149	10.2.2 Test Statistic and Profile Likelihood	86
150	10.2.3 CL_s Method	86
151	10.2.4 Exclusion Limits	87
152	10.2.5 μ Measurements	87
153	10.2.6 Nuisance Parameter Impact on the Signal Strength	88
154	11 Conclusions	89
155	11.1 Higgs Results in Review	89
156	11.2 Prospects for Future	89
157	Bibliography	90

List of Tables

159	5.1	Contributions of the main Higgs decay modes to the 3 multi-lepton $t\bar{t}H$ signatures at generation level.	40
161	6.1	Monte Carlo samples used for signal description.	44
162	6.2	Monte Carlo samples used for background description. Unless otherwise specified MadGraph samples use Pythia 6 for showering and Alpgen samples use Herwig+Jimmy. $t\bar{t}$, single top, and $Z+jets$ samples are replaced with data-driven estimates for the final result	45
165	7.1	Selections in the 2ℓ SS, 3ℓ and 4ℓ Signal Regions	47
166	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.	51
168	8.1	Expected number of signal and background events in 2ℓ SS, 3ℓ and 4ℓ signal regions. For data-driven backgrounds, Monte Carlo only numbers are given for reference. The expected sensitivity $\frac{s}{\sqrt{s+b}}$ is provided for data-driven, MC only and for the inclusion of ad-hoc systematic uncertainties (20% for $t\bar{V}$ and 30% for $t\bar{t}$ fake).	53
172	8.2	NLO cross section and theoretical uncertainty calculations derived from MadGraph5+aMC@NLO.	54
173	8.3	Truth origin of highest energy b-tagged jet in the $W^\pm Z + b$ VR and $3l$ SR	58
174	8.4	Truth origin of highest energy b-tagged jet in the $ZZ + b$ VR and $4l$ SR	61
175	8.5	Number of events of the main simulated background processes and of the data in the $e^\pm e^\pm$ and $\mu^\pm \mu^\pm$ channels used for the measurement of θ_e and θ_μ . VV , $V\gamma$, $t\bar{t}V$, tV and $t\bar{t}$ prompts (or charge misID) are the backgrounds which lead to prompt same-sign dileptons and are subtracted from the data to get a measured number of fakes. Uncertainties are statistical. The numbers labeled Data fakes are used to measure θ	68
180	8.6	Expected and measured values of the θ factors. NEEDS TO BE UPDATED FOR 2,3	69
181	8.7	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)	70
186	8.8	Summary of regions and inputs to the extraction of the number of $t\bar{t}$ events with a fake muon in the SR	71

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.	77
189	9.2	Uncertainties on $t\bar{t}H$ acceptance in signal regions due to PDF variation.	78
190	9.3	Sum in quadrature of all the systematic uncertainties on the number of event yields per channel.	83
191	9.4	Summary of systematic uncertainties for processes present in the signal regions in the anal- ysis, with their type, description, name, values and uncertainties, and status of inclusion in the final results.	84
192	10.1	95%CL limits on μ for all channels and combination with cumulative uncertainties.	87
193			
194			
195			
196			

List of Figures

198	2.1	The Standard Model Particle Content	4
199	2.2	χ^2 as a function of the Higgs mass (top left), the top quark mass (top right), the W boson mass (bottom left) and the effective weak mixing angle (bottom right) for the combined SM fit from the GFitter group. The data points placed along $\chi^2 = 1$ represent direct measurements of the respective observable and their $\pm 1\sigma$ uncertainties. The grey (blue) bands show the results when excluding (including) the new M_H measurements from (in) the fits.	6
200			
201			
202			
203			
204			
205	2.3	Proton Parton Distribution Functions (PDFs) from the MSTW Collaboration at $Q^2 = 10$ GeV 2 and $Q^2 = 10^4$ GeV 2	7
206			
207	2.4	Dominant Higgs production modes at the LHC	8
208	2.5	8 TeV LHC Higgs production cross-sections (left) and decay branching fractions	8
209	2.6	ATLAS Higgs combination results for all SM measurement channels as ratios of the measured to SM production cross-sections (left) and extracted Higgs coupling constraint scale-factors for a combined fit to the measurement channels, where the W, Z , top-quark, b-quark, and τ couplings are allowed to float. The p-value of this particular model is 0.13 and in agreement with SM expectations	10
210			
211			
212			
213			
214	2.7	RGE for the running of the SM parameter, λ for the Higgs self-coupling term with present values and uncertainty bands for M_H and M_t (left). The two-dimensional plot colored (right) shows regions for which the SM is stable, unstable and metastable based on this RGE.	11
215			
216			
217			
218	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 in the accelerator complex that accelerates the protons to 450 GeV before injection into the LHC.	14
219			
220			
221			
222	3.2	Diagram of the ATLAS detector and subsystems	16
223	3.3	Diagram of the ATLAS ID in the $R - \phi$ plane showing the barrel view of the Pixel Detector, SCT, and TRT.	18
224			
225	3.4	Diagram of the ATLAS ID in the $R - z$ plane showing the endcap view of the Pixel Detector, SCT, and TRT. Only one side of the endcap is shown.	18
226			
227	3.5	Diagram of the ATLAS calorimeters	20
228	3.6	Diagram of the ATLAS LAr EM calorimeter showing the longitudinal segmentation and the $\eta - \phi$ cells for the central barrel region	21
229			

230	3.7	Diagram of the ATLAS muon system	23
231	3.8	$R - \phi$ schematic of the ATLAS detector and various particle signatures	27
232	4.1	Electron reconstruction efficiency for an example p_T bin versus $ \eta $. The drop in efficiency at higher values of $ \eta $ is directly attributable to the increase in the amount of material in front of the EM calorimeter (left). The material causes bremsstrahlung, which makes track-cluster matching more difficult for electrons	29
236	4.2	Example single photon (left) and $\pi^0 \rightarrow \gamma\gamma$ (right) signatures in the ATLAS EM calorimeter. The fine segmentation of the cells in the strips allows for the distinguishing of two nearby showers from one shower and is used in electron identification	30
239	4.3	Electron and pion track high threshold probabilities as a function of their transverse momenta. The two scales are united through the γ -factor, on which the TR probability depends directly. Electron tracks are much more likely to have high threshold hits at electron energies typical of electro-weak decays.	31
243	4.4	Comparison of the ‘IsEM’ (left) and ‘IsEM++’ electron identification operating points for Loose, Medium, and Tight. The efficiency as a function of η for an example p_T bin is shown on top and the background rejection is shown on the bottom	32
246	4.5	Electron hadronic leakage fraction (R_{Had}) and transverse shower profile (R_η) in layer 2 for high and low pile-up conditions. Pile-up is measured here as the number of primary vertices in the event.	33
249	4.6	Schematic of the strategy to reduce the pile-up dependence of electron identification. A EM calorimeter wedge is shown overlaid with example electron (green) and pile-up particle (red) signatures. The strategy is to loosen the dependence of the identification on layer 2 and the hadronic calorimeter, which sample large volumes, and tighten selection on variables in the the 3rd layer and strips, which sample smaller volumes.	34
254	4.7	Comparison of the efficiency of the 2011 and 2012 electron identification menus versus the number of primary vertices in the event.	34
256	4.8	Example electron likelihood score output for electrons and fake electrons (left). A cut can be made at any point in this distribution to define a selection operating point (right). The cut-based operating point lies within the curve of possible likelihood operating points, showing that the likelihood indeed outperforms the cuts.	35
260	4.9	Comparison of the performance of the cut-based and likelihood operating points in the tight regime. Efficiency (top) and rejection (bottom) plots are shown versus $ \eta $ and E_T .	36
262	4.10	Example bins where the electron probe distribution from Z tag-and-probe is fit by an isolation template (left) invariant mass template (right) to subtract backgrounds.	37
264	4.11	Electron identification efficiency calculated in data and MC versus electron p_T	37
265	5.1	Example Feynman diagrams for the 3 $t\bar{t}H$ multi-lepton categories.	40
266	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.	43
269	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	43

272	7.1	Number of jets vs. number of b-tagged jet plot for the fully selected multi-lepton channels. Signal regions are outlined with a dashed line. Sub-channels are defined later in the 2ℓ SS and 4ℓ SRs. The fractional background contribution to each jet and b-tagged jet bin are shown for non-prompt (red), $t\bar{t}V + tZ$ (yellow), and VV (green). The expected signal fraction is shown in white. The expected non-prompt fraction contains charge misidentifications and fakes. It is shown for MC only, although data-based methods are used for the final result.	47
279	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
284	8.2	NJet spectrum for 2 tight-isolation leptons with 1 b-tagged jet (MV1_70)	57
285	8.3	3 lepton $W^\pm Z$ validation using the $t\bar{t}H$ lepton identification and momentum cuts: mass, number of jet and flavor variables	58
287	8.4	$W^\pm Z + b$ validation region: NJet, NElec, and Mass Variables	59
288	8.5	Jet-inclusive 4-lepton ZZ validation region using the $t\bar{t}H$ lepton identification and momentum cuts	60
290	8.6	$ZZ + b$ validation region using the $t\bar{t}H$ lepton identification and momentum cuts	60
291	8.7	Electron charge misidentification 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 misidentification 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.	64
298	8.8	Closure test on simulated $Z \rightarrow e^+e^-$ 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.	64
303	8.9	Relative systematic uncertainty contributions on the charge misidentification rate, for different bins in p_T and $ \eta $	65
305	8.10	Ratios of regions with tight and anti-tight leptons in 2-lepton and 3-lepton channels from $t\bar{t}$ MC. These ratios are the MC calculated transfer factors for each region, i.e. $\theta_e = eee/eed$, ee/ed and $\theta_\mu = mmp/mmm$, mm/mp , where 'd' refers to anti-tight electrons and 'p' refers to anti-tight muons. The transfer factors are seen to be similar in the 2ℓ and 3ℓ cases and stable as a function of the number of jets	67
310	8.11	4,5 Jet SS 2ℓ ed (above) and μp control regions with at least one b-tagged jet. After subtraction of prompt and charge misidentification 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 top MC (red) is used for reference but not in the actual calculation.	69

314	8.12 Distributions of the properties of the anti-tight muons in data (dots), compared with 315 the total simulation (red line), rescaled to the integral of the data for a shape compari- 316 son. The uncertainty on the data distribution is statistical. The number of events for 317 each of them is also presented in the legend. The variables probed are, top: p_T and 318 $\Delta R(\mu, \text{closest selected jet})$; bottom: $\text{ptcone20}/p_T$ and $\text{Etcone20}/p_T$. The selection is the 319 signal region event selection with one anti-tight muon (failing at least one of the isolation, 320 muon-jet overlap, or p_T selection criteria). A ratio plot is containing the 20% area around 321 1, is displayed, demonstrating that a 20% data-MC comparison systematic is sufficient.	72
322	8.13 Distributions of anti-tight electron variables. The variables presented are, from top left 323 to bottom right, p_T , η , VERYTIGHT Likelihood value, $\text{ptcone20}/p_T$, $\text{Etcone20}/p_T$. The 324 plotted regions have the same cuts as the signal region, except the anti-tight electron must 325 fail isolation for the plot of the VERYTIGHT identification word or fail the VERYTIGHT 326 identification word for the plots of the isolation. Data (dots) are compared with a stacked 327 histogram of the various simulated samples: top in red, $V+jets$ (blue), VV (purple) and 328 $t\bar{t}V$ (yellow). The uncertainty on the data distribution is statistical.	73
329	8.14 Muon-xxp (above) and electron-xxd (below) anti-tight control regions: jet variables. Note: 330 the $t\bar{t}$ and single top MC in the plots is used only as comparison, but is not included in 331 the fake measurement	74
332	8.15 3ℓ fake validation regions for nominal p_T selection (above) and relaxed p_T selection, > 10 333 GeV, (below). Plotted are the number of jets and the number of electrons in each event. 334 The data and MC ratio below each plot agree with 1 within the statistics of the region 335 and the overall systematic assigned for the fake component (red)	75
336	9.1 PDF systematic uncertainty on the jet multiplicities in $t\bar{t}H$ events with at least 2 leptons. 337 The dashed red lines in the lower panel indicate the systematic uncertainty on the $t\bar{t}H$ 338 production cross section.	78
339	9.2 Muon identification efficiency in Data and MC as a function of η . The CB+ST (com- 340 bined+segment tagged) operating point is used	79
341	9.3 Muon (left) and electron(right) isolation efficiency scale-factors from the Z control sample 342 as a function of the number of jets in the event. An additional systematic uncertainty of 343 1% is added to encompass the variation in the number of jets variable	81
344	9.4 JES systematic uncertainties as a function of jet η (for jets $p_T > 25$ GeV) and p_T (for jets 345 $ \eta < 0.4$). The combined systematic uncertainty is shown with contributions from the 346 largest sources	81
347	9.5 b-Tagging data-MC efficiency scale-factors versus jet p_T calculated in the $t\bar{t}$ sample from 348 2012 data. The uncertainties are combined statistical and systematic.	82

Preface

350 This analysis represents only the last year of work of my graduate career and was accomplished as
351 part of a ~ 10 person team. I was involved in some way at every level of the analysis, but for some
352 parts I had a more important role. I led the team that managed the datasets and analysis coding
353 infrastructure, developed the techniques for estimating the fake and vector-boson backgrounds and
354 optimized the selection for the 3 lepton signal region. I also directly supervised the work of multiple
355 students, each of which performed a critical task for the analysis: measurement of the isolation
356 selection efficiency for leptons, assessment of the impact of the experimental systematic uncertainties
357 on the analysis and development of the 4 lepton signal region.

358 Earlier in my graduate career, I focused primarily on performance work for the ATLAS experiment.
359 In 2010 I helped create an in-situ mapping of the inner detector material using photon conversions. As
360 a part of small Penn-led team, I developed the electron identification algorithms that formed the basis
361 of the primary electron trigger for the 2011 and 2012 runs and are used by many ATLAS analyses.
362 Electrons are critical pieces of analyses at hadron colliders but also are challenging to understand due
363 to large and diverse backgrounds. Our team oversaw the development of electron identification from
364 very basic cut-based roots to the extremely robust and high performing multi-variate algorithms used
365 in this thesis. We managed this at a time when the running conditions were changed rapidly. Later,
366 I pushed for the proper measurement of electron identification efficiency at lower energies that were
367 important for finding Higgs signals in a number of channels.

368 Before the $t\bar{t}H$ analysis, I worked on two additional analysis projects. The first was a search for
369 the Higgs in the W^+W^- decay channel. For this analysis I played a small role as the coordinator for
370 non- W^+W^- di-boson background measurements. The second was a search I explored for the Higgs in
371 the $W^\pm H$ production mode with like-sign dileptons. As many things in science, this ended up being
372 a dead-end: a difficult analysis without compelling sensitivity.

373 I had the privilege and luck of being part of some truly amazing research at a very important time

374 in the field. While experimental particle physics is a collaborative effort, I can look at my small set
375 of accomplishments with pride in the fact that they helped contribute to the overall success of the
376 experiment and to push our fundamental understanding of the universe the tiniest bit forward.

377

Chris Lester
CERN, Fall 2014

CHAPTER 1

Introduction

380 The discovery of the Higgs boson at the Large Hadron Collider (LHC) experiments has opened up a
381 new paradigm of research into the Standard Model of particle physics. This thesis primarily documents
382 a search for the production of Higgs boson in association with top quarks ($t\bar{t}H$) in multi-lepton
383 final states. Searching for this production mode of the Higgs is an important step toward a precise
384 measurement of the top Yukawa coupling, because it accesses this coupling via diagrams that do not
385 contain loops. Comparison of this coupling with the already well-measured top quark mass provides
386 a direct test of a fundamental provision of the Higgs mechanism: that it gives mass to the fermions.

387 The analysis uses the 2012 ATLAS experiment's dataset of proton-proton collisions at a center-of-
388 mass energy of 8 TeV provided by the LHC. The statistics available do not allow for an observation
389 of the $t\bar{t}H$ process at the Standard Model production cross-section, and the results of the search are
390 interpreted as a 95% exclusion on the production rate. The results will provide some of strictest
391 constraints on the rate to date and establish a program for future analyses on larger datasets that
392 will eventually observe this production mode.

393 Chapter 2 provides theoretical background and motivation for the study of this particular Higgs
394 production mode and Chapter 3 provides a basic review of the experimental apparatus, the LHC and
395 ATLAS. Chapter 4 is a brief diversion from the main text to elaborate on the techniques used to
396 identify electrons and measure their identification efficiency.

397 Chapters ?? are the main text, which discuss the full analysis procedure for the search and the
398 final measurement. The results of the analysis are currently undergoing approval in the ATLAS
399 collaboration and eventually will be documented for publication. Once approved the results will be
400 combined with other Higgs searches to set limits on Higgs couplings to other SM particles, particularly
401 the top quark.

CHAPTER 2

Theoretical Background

404 The Standard Model of particle physics (SM) is an extraordinarily successful description of the funda-
 405 mental constituents of matter and their interactions. Many experiments have verified the extremely
 406 precise prediction of the SM. This success has culminated most recently in the discovery of the Higgs
 407 Boson. This chapter provides a brief introduction to the structure of the SM and how scientists are
 408 able to test it using hadron colliders. It focuses primarily on the physics of the Higgs boson and its
 409 decay to top quarks. I stress the importance of a measurement of the rate at which Higgs Bosons are
 410 produced in association of top quarks, as a new, rigorous test of the SM. The experimental search for
 411 this production mode in multi-lepton final states is the general subject of this thesis.

412 **2.1 The Standard Model**

413 **2.1.1 The Standard Model Structure**

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

420 Demanding these symmetries be local, gauge symmetries allows the theory to be re-normalizable
 421 [5], meaning that unwanted infinities can be absorbed into observables from theory in a way that
 422 allows the theory to be able to predict physics at multiple energy scales. Gauging the symmetries

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

428 Matter particles are half-integer spin fermions and are representations of the symmetry groups.
 429 Singlets of the $SU(3)$ are called leptons, do not have a color charge, and, therefore, do not interact
 430 with the strong force. Quarks, as triplets of the $SU(3)$ group, do interact with the strong force.
 431 The SM is a chiral theory: the weak force violates parity, as it only couples to left-chiral particles
 432 or right-chiral antiparticles. This means that right-chiral and left-chiral fermions arise from different
 433 fields, which are different representations of the $SU(2)_L$ group.

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

440 2.1.2 Electroweak Symmetry Breaking and the Higgs

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

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

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

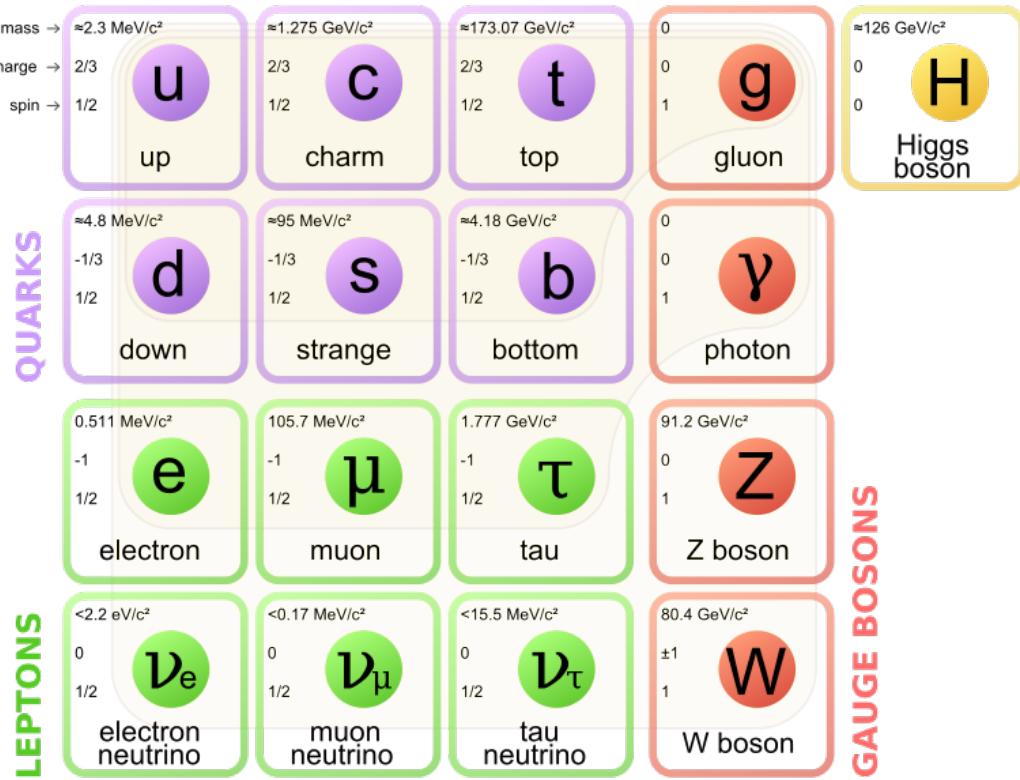


Figure 2.1: The Standard Model Particle Content

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

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

⁴⁶³ **2.1.3 The Standard Model Parameters**

⁴⁶⁴ Confronting the SM with experiment requires the measurement of 17^2 free parameters, which are
⁴⁶⁵ unconstrained from the theory. These free parameters include the fermion masses from the Yukawa
⁴⁶⁶ couplings, the force coupling constants, the angles and phase of the mixing between quarks, and
⁴⁶⁷ constants from the Higgs and electroweak sector³.

⁴⁶⁸ Experiments have provided a number of measurements of the parameters of the SM[10]. With
⁴⁶⁹ the discovery of the Higgs boson and the measurement of the Higgs mass, all of the parameters of
⁴⁷⁰ the SM can be estimated and statistical procedures can assess the relative agreement of overlapping
⁴⁷¹ measurements to test the self-consistency of the SM. The GFitter collaboration assembles all relevant
⁴⁷² electroweak observable measurements into a statistical model and then allows certain measurements
⁴⁷³ to float within their uncertainty to allow for a fit among multiple correlated measurements[11]. These
⁴⁷⁴ correlations arise for two reasons. First, measurements are made that often depend on multiple SM
⁴⁷⁵ parameters. Second, radiative corrections often cause parameters to depend on each other. For
⁴⁷⁶ instance, the Higgs mass is sensitive to both the W mass and top mass, through loop level corrections.

⁴⁷⁷ Figure 2.2 shows the fitted constraints on 4 key SM parameters (M_H , M_W , M_t , $\sin^2\theta_w$) with actual
⁴⁷⁸ measurements overlaid. The plots show both the removal and inclusion in the fit of key measurements
⁴⁷⁹ to assess their overall impact. The addition to the fit of the measured Higgs mass from the ATLAS and
⁴⁸⁰ CMS collaborations creates a small tension, as the other observables prefer the mass to be much lower
⁴⁸¹ (~ 80 GeV). This tension in the combined electroweak fit as a result is not statistically significant
⁴⁸² with a p -value of 0.07. The SM seems to be self-consistent.

⁴⁸³ **2.2 Collider Physics and the Higgs**

⁴⁸⁴ To test the theory, physicists accelerate particles to extremely high energies and force them to interact
⁴⁸⁵ through collisions. Typically, the particles accelerated are electrons or protons, since they are stable.
⁴⁸⁶ Electron-positron collider machines have a rich history of discovery and measurement in particle
⁴⁸⁷ physics. The advantage of electron accelerators is that the colliding element is itself a fundamental
⁴⁸⁸ particle. However, due to synchrotron radiation, the curvature of the beam line becomes problematic
⁴⁸⁹ for high energy beams. On the other hand, proton-proton and proton-anti-proton colliders can be
⁴⁹⁰ accelerated in rings without large losses due to synchrotron radiation, but the actual colliding objects

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

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

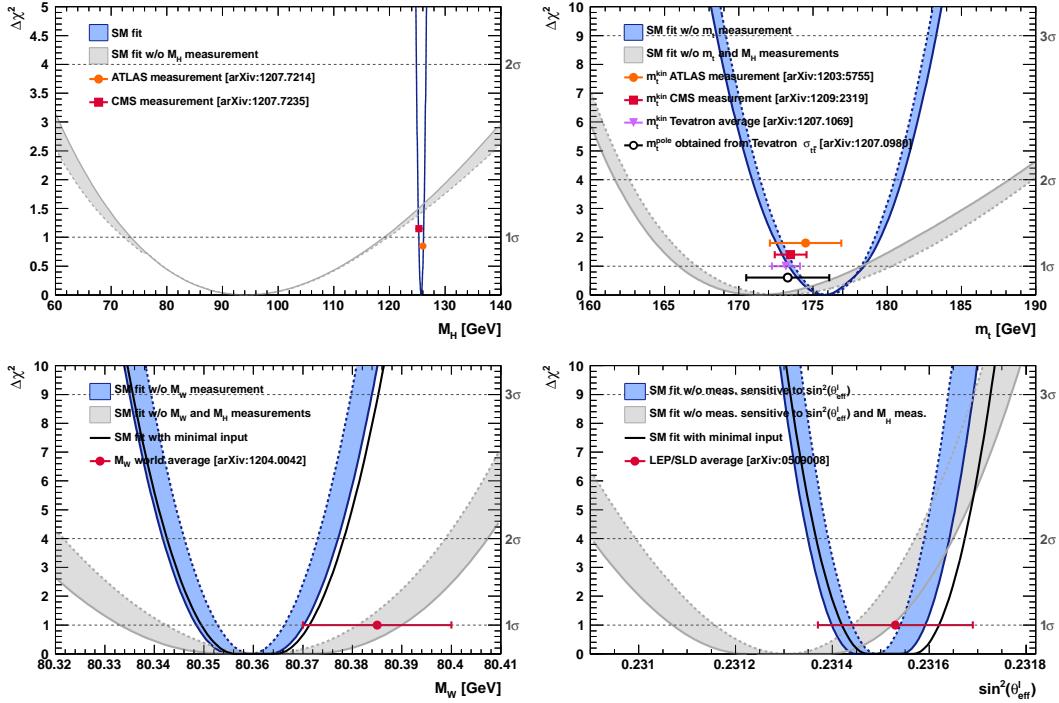


Figure 2.2: χ^2 as a function of the Higgs mass (top left), the top quark mass (top right), the W boson mass (bottom left) and the effective weak mixing angle (bottom right) for the combined SM fit from the GFitter group. The data points placed along $\chi^2 = 1$ represent direct measurements of the respective observable and their $\pm 1\sigma$ uncertainties. The grey (blue) bands show the results when excluding (including) the new M_H measurements from (in) the fits.

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

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

At the Large Hadron Collider (LHC) protons are collided. The types of initial hard-scattering states at the LHC are quark-quark, quark-gluon, and gluon-gluon. Gluon collisions dominate overall, due to the large number of gluons inside the proton, though the relative importance of different initial states changes with the energy scale of the collision and the type of final state selected.

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

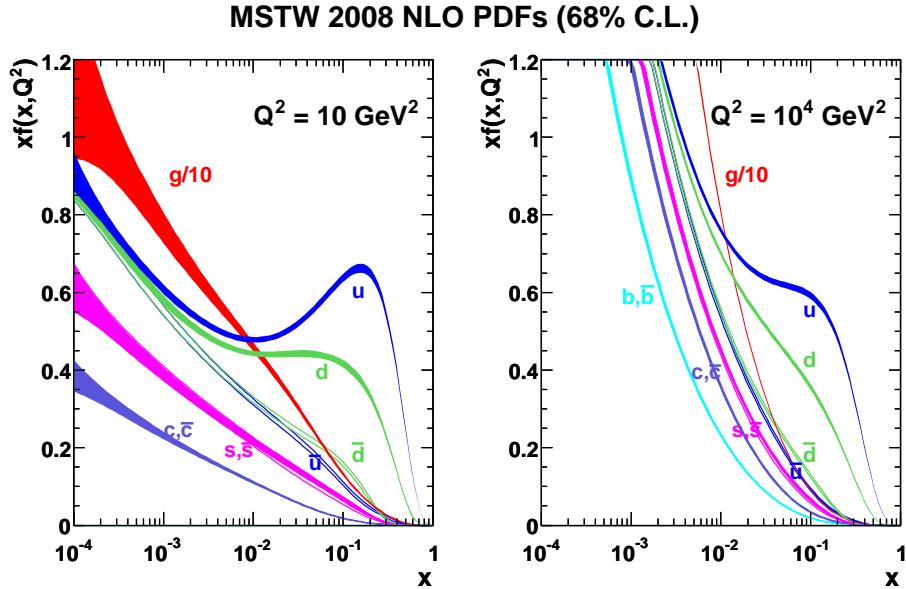


Figure 2.3: Proton Parton Distribution Functions (PDFs) from the MSTW Collaboration at $Q^2 = 10 \text{ GeV}^2$ and $Q^2 = 10^4 \text{ GeV}^2$

especially below 114 GeV. The Higgs mass was also known to have a theoretically motivated upper bound. The unitarity of diagrams including the $WWWW$ vertex required the Higgs mass to be below about 1 TeV. The LHC was designed to be able to eventually find or exclude a Higgs particle in this range [10].

Reaching this discovery or exclusion required an enormous dataset with collisions at high energies. Despite the fact that the Higgs couples to nearly every particle, Higgs boson production at the LHC is a low rate process. Because it couples to fermions proportional to mass and because the colliding particles must be stable and therefore light, production of the Higgs must occur through virtual states.

The Higgs boson can be produced through collision at the LHC via 4 mechanisms: gluon-fusion (ggF), vector-boson fusion (VBF), Higgsstrahlung (VH), and production in association with top quarks ($t\bar{t}H$). The diagrams are shown in Figure 2.4 and the production cross-sections as a function of Higgs mass for the 8 TeV LHC proton-proton running are shown in Figure 2.5 [14]. The largest production cross-section is via the gluon fusion channel at 20 pb, which proceeds through a fermion loop that is dominated by the top quark, because of its large Yukawa coupling to the Higgs. Because the Higgs couples to every massive particle, it has a rich set of decays also seen in Figure 2.5, especially for $m_H = 125 \text{ GeV}$. Studies of Higgs properties at hadron colliders offers many tests of the Standard Model and ample room for new physics searches. These tests specifically can verify

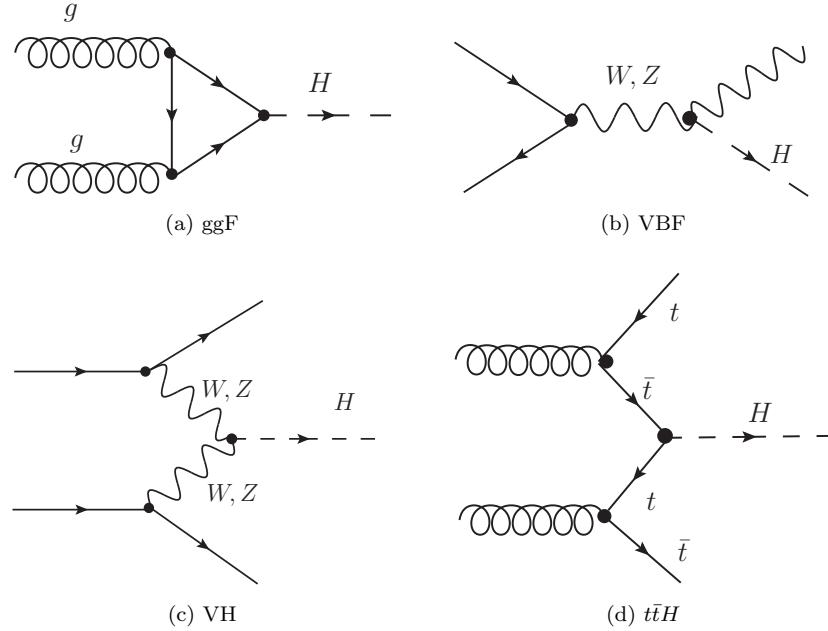


Figure 2.4: Dominant Higgs production modes at the LHC

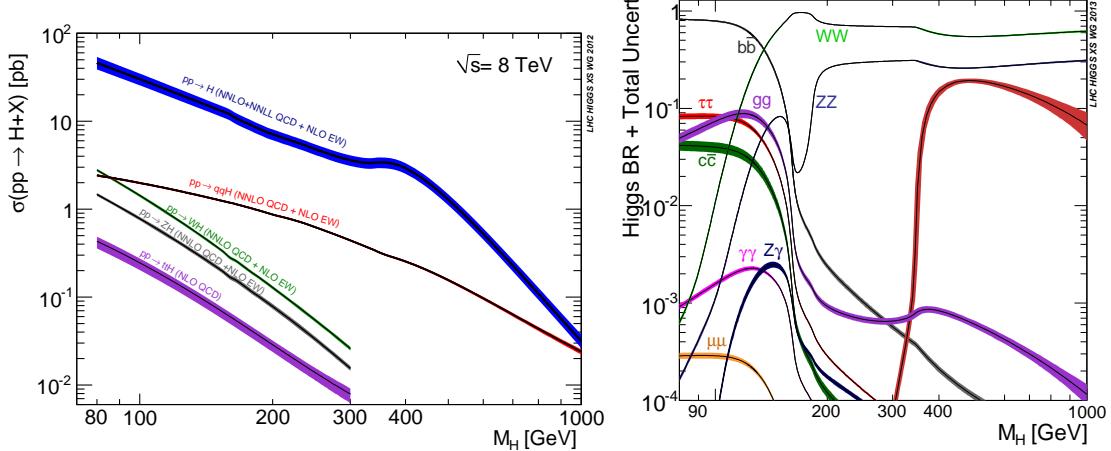


Figure 2.5: 8 TeV LHC Higgs production cross-sections (left) and decay branching fractions

521 the link between Yukawa coupling and the particles mass and further constrain details of EWSB by
 522 examining Higgs coupling to the weak bosons.

523 **2.2.1 Higgs Discovery at the LHC**

524 In 2012 both ATLAS and CMS announced the discovery of a new boson consistent with the Higgs
 525 by examining the results of Higgs searches in a number of decay channels ($H \rightarrow W^+W^-$, $H \rightarrow Z^0Z^0$,
 526 and $H \rightarrow \gamma\gamma$) in the 2011 dataset at $\sqrt{s} = 7$ TeV and part of the 2012 dataset at $\sqrt{s} = 8$ TeV. By
 527 2013 and 2014, both experiments have updated and/or finalized their results for the full 2011 and
 528 2012 datasets [15, 16]. I will focus on the ATLAS results, which measured both the Higgs mass[17]
 529 and spin[18], as well as provided initial constraints of the Higgs couplings to different particles.

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

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

539 **2.2.2 The Importance $t\bar{t}H$ Production**

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

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

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

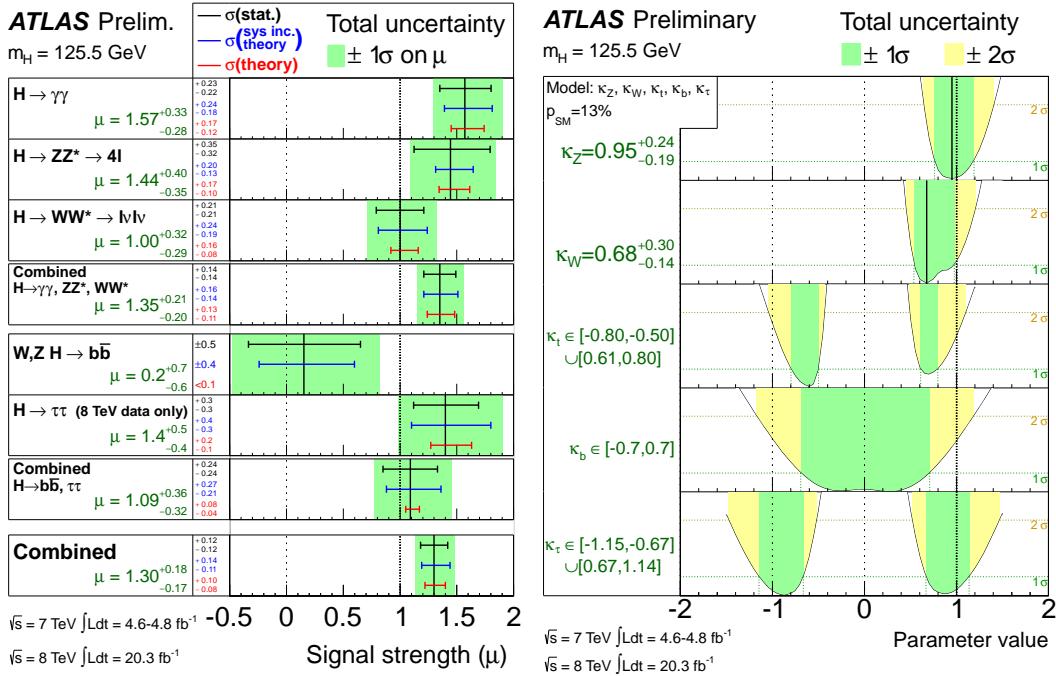


Figure 2.6: ATLAS Higgs combination results for all SM measurement channels as ratios of the measured to SM production cross-sections (left) and extracted Higgs coupling constraint scale-factors for a combined fit to the measurement channels, where the W, Z , top-quark, b -quark, and τ couplings are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

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

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

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

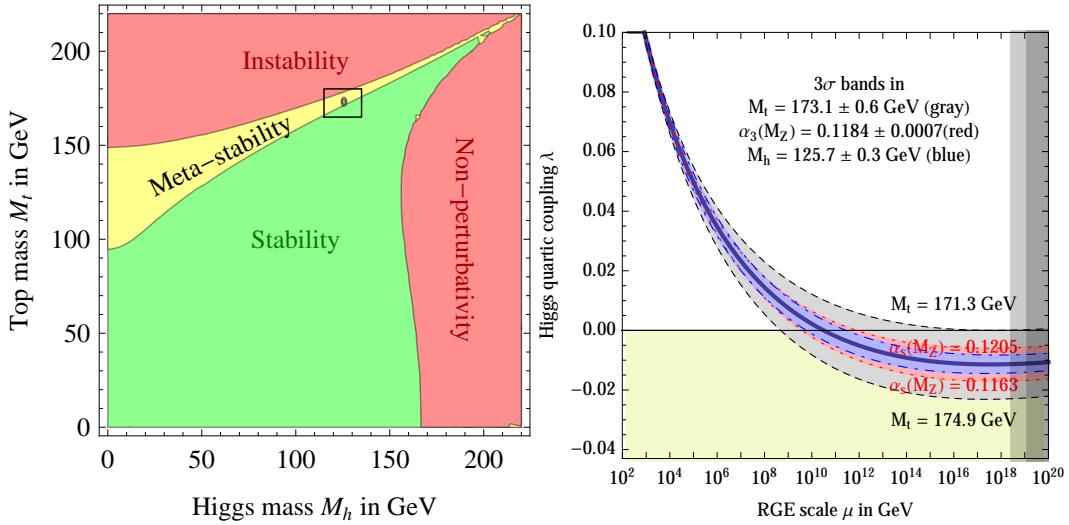


Figure 2.7: RGE for the running of the SM parameter, λ for the Higgs self-coupling term with present values and uncertainty bands for M_H and M_t (left). The two-dimensional plot colored (right) shows regions for which the SM is stable, unstable and metastable based on this RGE.

for the entire mechanism of EWSB no longer has a minimum and becomes unbounded, resulting in instability in the universe [24]. Metastability occurs when the shape of the potential allows for a false local minimum. Figure 2.7 shows the running of this parameter, the regions for which the universe is stable, unstable and metastable. Current measurements suggest that universe lies in a metastable island⁴. This is a sort of fanciful aside, intended only to highlight the importance of the top Yukawa coupling and to suggest that new discoveries in the top-Higgs sector have far reaching consequences.

2.3 Conclusion

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

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

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

CHAPTER 3

589
590

The Large Hadron Collider and the ATLAS Experiment

591

3.1 The Large Hadron Collider

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

595 The LHC is the world's highest energy particle accelerator and is located 100m underneath the
596 Franco-Swiss border at the European Organization for Nuclear Research (CERN) in a 26.7 km tunnel.
597 The technology involved in the development of the LHC is an enormous achievement in its own right
598 and is documented in detail here [25, 26, 27]. The LHC is a circular machine capable of accelerating
599 beams of protons and colliding them at center of mass energies up to $\sqrt{s} = 14$ TeV at 4 collision sites
600 around the ring, where 4 experiments are housed (ATLAS[28], CMS[29], LHCb[30], and ALICE[31]).
601 Figure 3.1 is a diagram of the layout of the LHC and its experiments[32]. The LHC also operates in
602 a mode with beams of heavy ions. The LHC is composed of thousands of super-conducting Niobium-
603 Titanium magnets, cooled to 1.9° K with liquid Helium, which steer and focus the particle beams,
604 and a superconducting resonant-frequency (RF) cavity, which boosts the beam to higher energies.

605

3.1.1 The Accelerator Complex

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

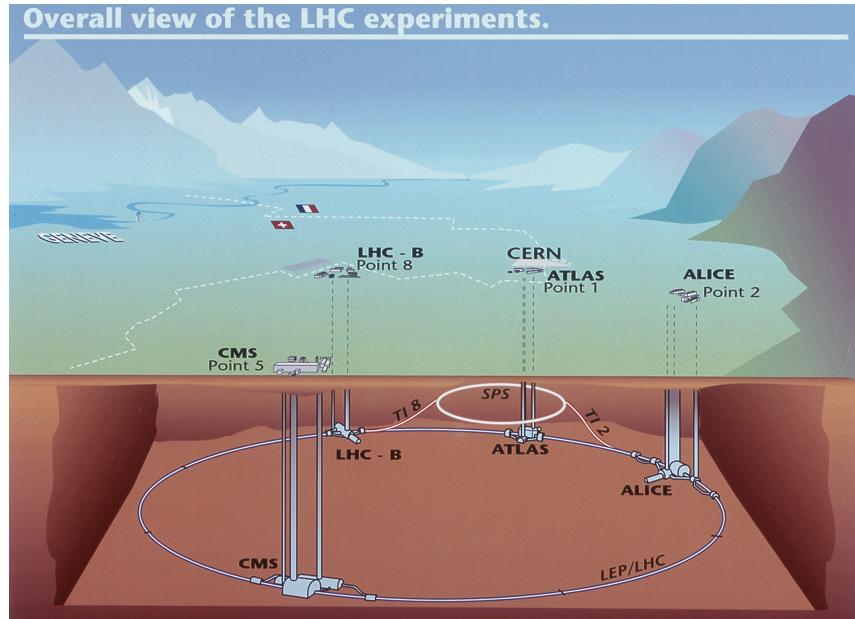


Figure 3.1: Diagram of the Large Hadron collider and location of the 4 main experiments (ATLAS, CMS, LHCb, and ALICE) around the ring. The diagram also shows the location of the SPS, the final booster ring in the accelerator complex that accelerates the protons to 450 GeV before injection into the LHC.

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

614 **3.1.2 Beam Parameters and Collisions**

615 For the physics studied at the ATLAS experiment, the two most important parameters of the collisions
616 are the center of mass energy (CME) and instantaneous luminosity (\mathcal{L}). High center of mass energies
617 are necessary for the production of new high mass particles, and, because the constituents of the
618 actual collisions are the partons of the proton, the CME of the collisions must in general be much
619 higher than the mass of the particles produced.

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

629 **3.2 The ATLAS Experiment**

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

635 ATLAS, shown in Figure 3.2, possesses cylindrical symmetry around the beam pipe. It weights
636 7000 tons, has a diameter of roughly 25m and length of 46m. ATLAS was designed to be a multi-
637 purpose hermetic, particle detector, able to identify many types of particles, and designed to provide
638 a snapshot of the entire collision event. The detector sub-systems form concentric rings around the
639 beam-line at increasing distance. From closest to the beam outward, they are:

- 640 • **Inner Detector:** The inner detector (ID)[33, 34] is immersed in a solenoidal magnetic field[35]
641 and provides measurements of charge particle tracks, through three subsystems: the Pixel
642 Detector[36, 37], the Semi-Conductor Tracker (SCT)[38, 39], and Transition Radiation Tracker(TRT)
643 [40, 41, 42].

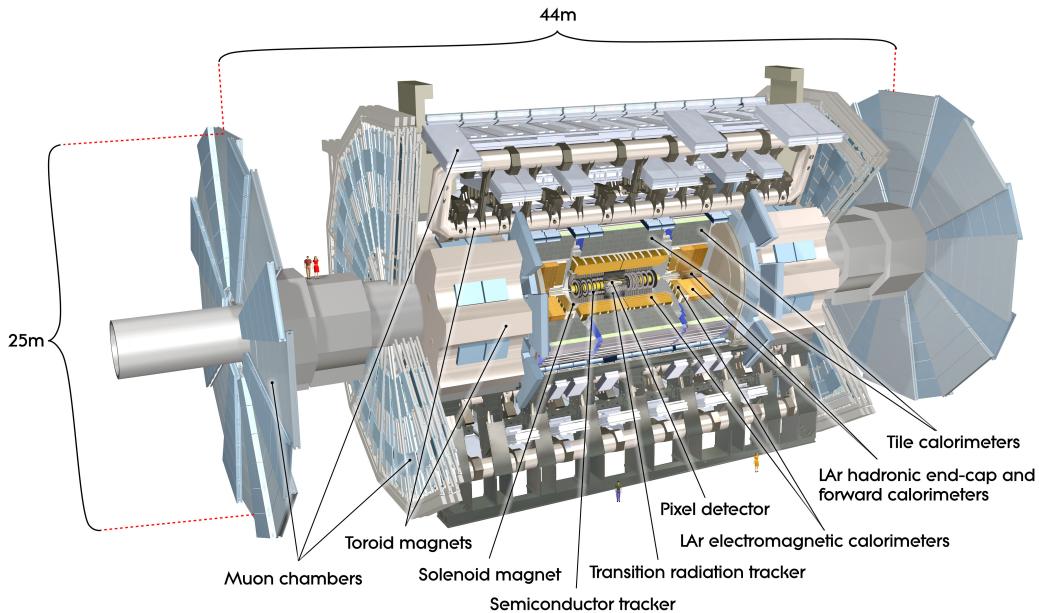


Figure 3.2: Diagram of the ATLAS detector and subsystems

- **Calorimeter:** The calorimeters measure the energy of particles that participate in the electromagnetic (photons, electrons) and hadronic interactions (pions, protons, neutrons, etc.), by forcing them to shower in dense material. The hermeticity of the calorimeters allows for missing transverse energy measurements. The calorimeter is composed of the liquid argon electromagnetic calorimeter (LAr)[43], the hadronic tile calorimeter[44], the liquid argon hadronic endcap calorimeter, and the forward calorimeters.
- **Muon Spectrometer:** The muon spectrometer (MS) sub-systems[45] form the outermost detector systems and measure the momentum of minimum ionizing muon tracks, as all other particles are stopped by the calorimeters. The muon systems are immersed in a toroidal magnetic field [35] and are composed of 4 different sub-systems for triggering and tracking measurements [46, 47, 48].
- **Triggering Systems:** The trigger and data acquisition systems[49, 50] read out data from the detector through a three-tiered hardware and software decision making framework to record the most interesting physical processes for a broad physics analysis program.

These systems are discussed in depth in the following sections.

659 **3.2.1 Detector Coordinate System**

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

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

674 The ‘barrel’ and ‘endcap’ are classifications that are used to label the position of sub-detectors.
675 Barrel sub-detectors occupy positions more central to the detector at $|\eta|$ values roughly less than
676 1-2, while the endcap calorimeters extend farther in $|\eta|$. The barrel-endcap transition region contains
677 detector services. Also, the orientation of the detector elements are often different in the barrel and
678 endcap to have optimal particle flux.

679 **3.2.2 The Inner Detector**

680 The ID makes measurements of the position of charged particles as they move through the detectors
681 3 sub-systems (Pixel Detector, SCT, TRT). The individual position measurements can be strung
682 together to form a particle track. The entire ID is immersed in a 2T solenoidal magnetic field allowing
683 for measurements of particle momenta through the curvature of the tracks. The ID is contained with
684 a radius of 1.15 m and has a total length of 7m, allowing for particle tracking out to $|\eta| < 2.5$.
685 Figures 3.3 and 3.4 show the placement of the ID sub-systems in the $R - \phi$ and $R - z$ planes.

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

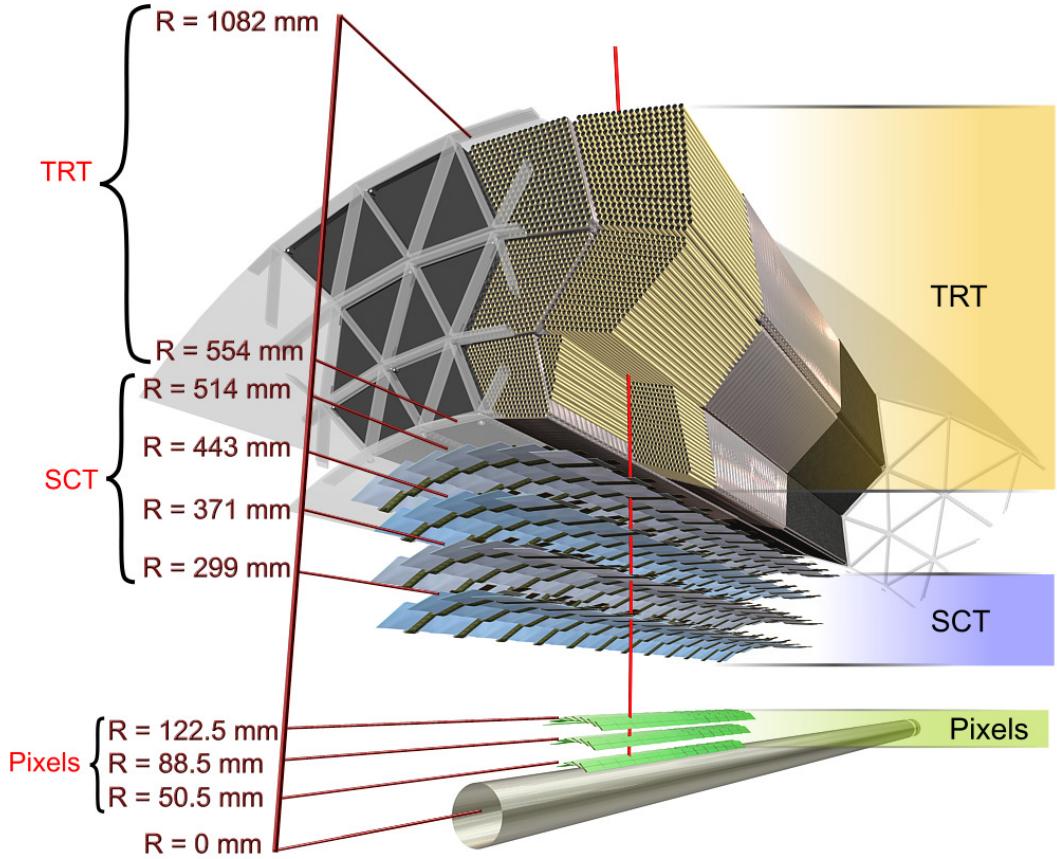


Figure 3.3: Diagram of the ATLAS ID in the $R - \phi$ plane showing the barrel view of the Pixel Detector, SCT, and TRT.

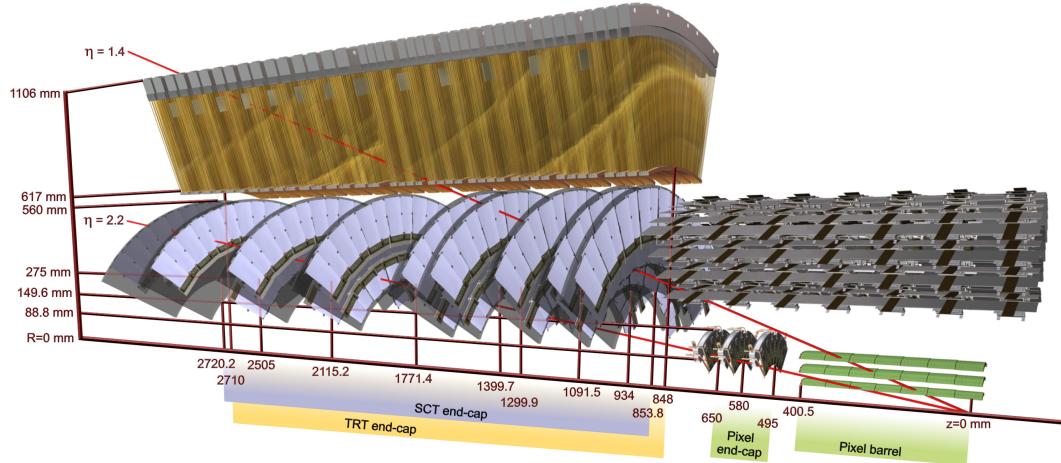


Figure 3.4: Diagram of the ATLAS ID in the $R - z$ plane showing the endcap view of the Pixel Detector, SCT, and TRT. Only one side of the endcap is shown.

691 b-layer. The pixels provide excellent hit resolution ($R - \phi$ accuracy of $10\ \mu\text{m}$ and $z(R)$ accuracy of
 692 $115\ \mu\text{m}$ in the barrel (endcap)).

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

698 The TRT is comprised of straw drift tubes filled with a gas mixture, containing mostly Xenon
 699 gas. Particles traversing the straws ionize the gas, and the liberated electrons drift to a wire at the
 700 center of the straw, which has an applied voltage, and induce an signal on the wire. The TRT has
 701 $\sim 300,000$ straws. The barrel straws are arranged cylindrically along the z direction out to $\sim \eta < 1$
 702 and the endcap straws point radially outward in the R direction. For this reason, the barrel(endcap)
 703 straws provide no measurement in the $R(z)$ directions. The drift tubes provide individual position
 704 measurements with resolutions of $\sim 130\ \mu\text{m}$. Each particle track has on average 35 hits, which is
 705 large compared to the Pixel and SCT tracks, which have on average 7 hits.

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

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

719 3.2.3 The Calorimeter

720 The ATLAS calorimeters measure the energy of electrons, photons and hadrons with $|\eta| < 4.5$. They
 721 induce a particle shower via electro-magnetic and nuclear interactions with the detector material and
 722 are deep enough to ensure that all or most of the shower energy remains contained. Exceptioanlly,

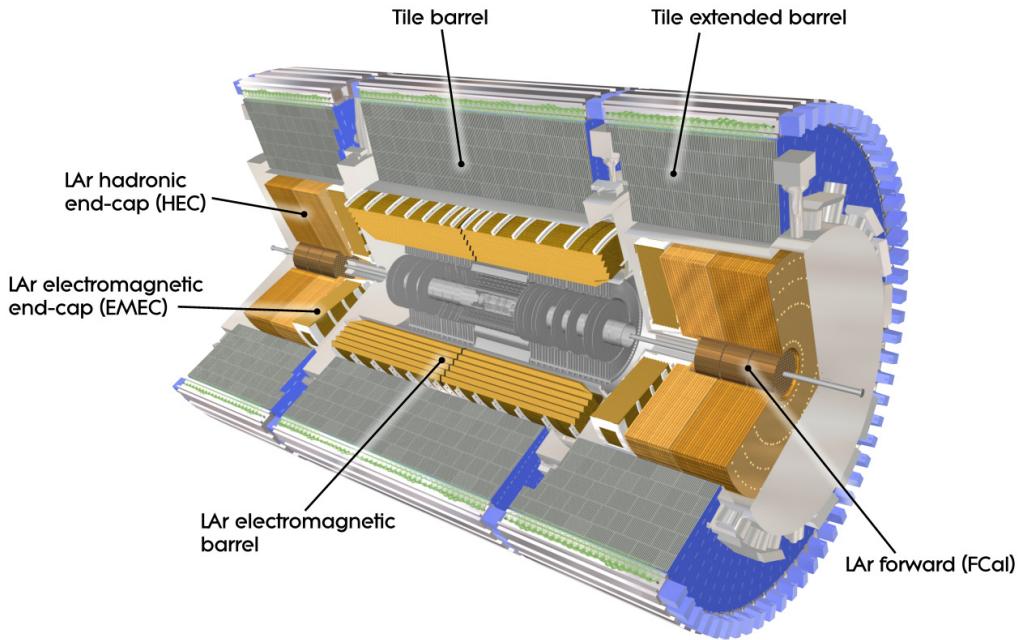


Figure 3.5: Diagram of the ATLAS calorimeters

723 muons pass through the ATLAS calorimeters leaving relatively little energy behind. ATLAS calorime-
 724 ters are sampling calorimeters meaning that the active material of the detector only measures a small
 725 fraction of the energy produced by the shower. The overall shower energy is inferred from this frac-
 726 tional measurement. The rest of the material is inactive, dense material, designed to induce showers.
 727 The calorimetry system is grossly divided longitudinally (radially) into electro-magnetic (EM) and
 728 then hadronic segments, operated with different technologies. Figure 3.5 diagrams the layout of the
 729 calorimeter system.

730 The EM calorimeter (LAr), which is located directly outside of the solenoid magnet but within
 731 the same cryostat, has an accordion design with lead absorber and liquid argon active material. The
 732 accordion design ensures uniform coverage in ϕ . The barrel and endcap LAr extend to $|\eta| < 2.47$. The
 733 LAr provides highly granular measurements in $\eta - \phi$ with 4 longitudinal segments, totaling $\sim 25\text{-}35$
 734 radiation lengths with the exception of the barrel/endcap transition region ($1.37 < |\eta| < 1.52$). The
 735 geometry of the barrel LAr calorimeter can be seen in Figure 3.6. The first longitudinal segment
 736 is called the pre-sampler, composed of a thin layer of active liquid argon, designed to detect early
 737 particle showers. The second segment is the most highly granular segment called the ‘strips’, as it
 738 is composed of thin liquid argon cells. The strips have a size of $0.025/8 \times 0.1$ in $\eta - \phi$ in the barrel

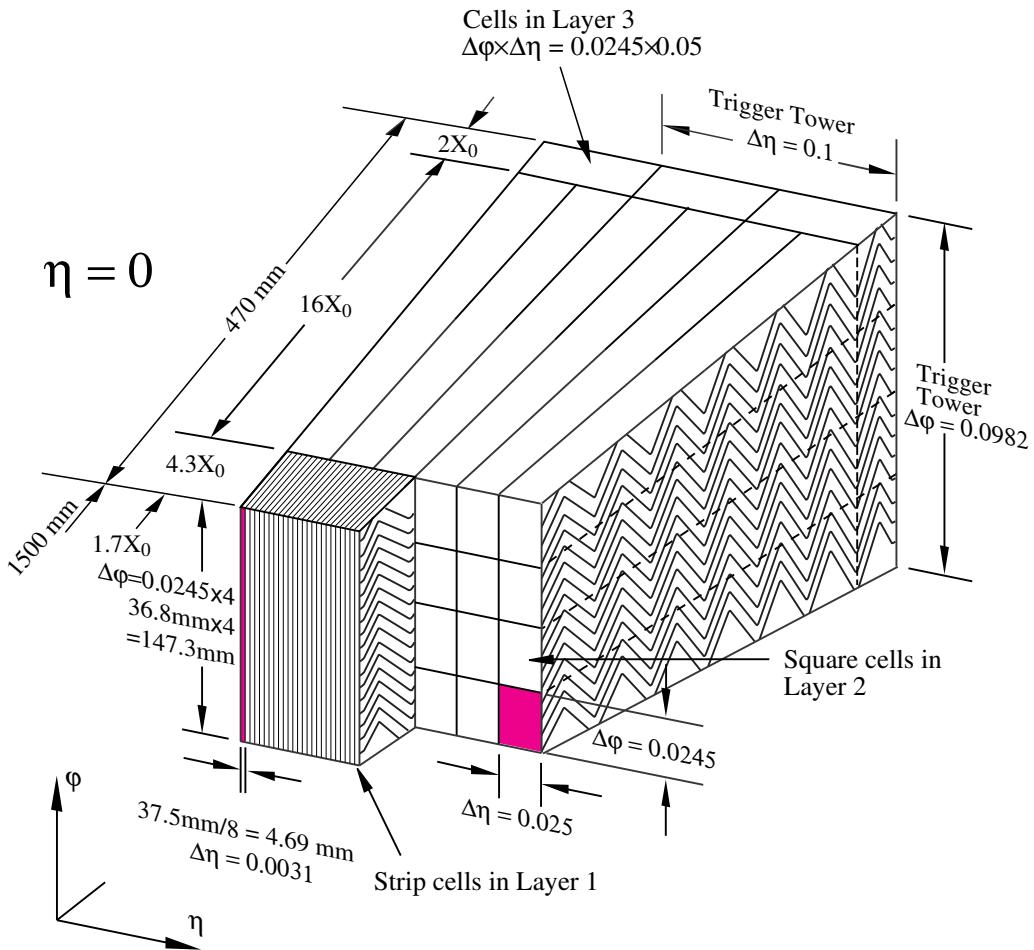


Figure 3.6: Diagram of the ATLAS LAr EM calorimeter showing the longitudinal segmentation and the $\eta - \phi$ cells for the central barrel region

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

747 The hadronic calorimeter is located directly behind the EM calorimeter. It is composed of tiles

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

748 of iron absorber and plastic scintillator in the barrel ($|\eta| < 1.6$), called the TileCal, and copper-
749 liquid argon in the endcap ($1.5 < |\eta| < 3.2$), called the HEC. The calorimeters contain $\sim 10\text{-}19$
750 hadronic interactions lengths with multiple longitudinal segments to contain showers induced by
751 the nuclear interaction of hadronic particles. The energy resolution of the hadronic calorimeters is
752 $\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$. The intrinsic resolution of hadron calorimeters is much worse than electro-
753 magnetic calorimeters, because much of the energy is lost to the inelasticity of nuclear break-up.

754 **3.2.4 The Muon Spectrometer**

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

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

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

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

773 **3.2.5 The Trigger System**

774 The ATLAS trigger system is designed to make quick decisions about individual particle collisions to
775 reduce the enormous collision rate of 20 MHz to a much more manageable 400 Hz to be stored for
776 offline analysis. Saving the full ATLAS data-stream would require space for 40 TB of raw data per
777 second, but, more importantly, most of these collisions result in the uninteresting inelastic break-up of
778 the colliding protons. To select out collisions to allow for a diverse physics program, ATLAS devotes a

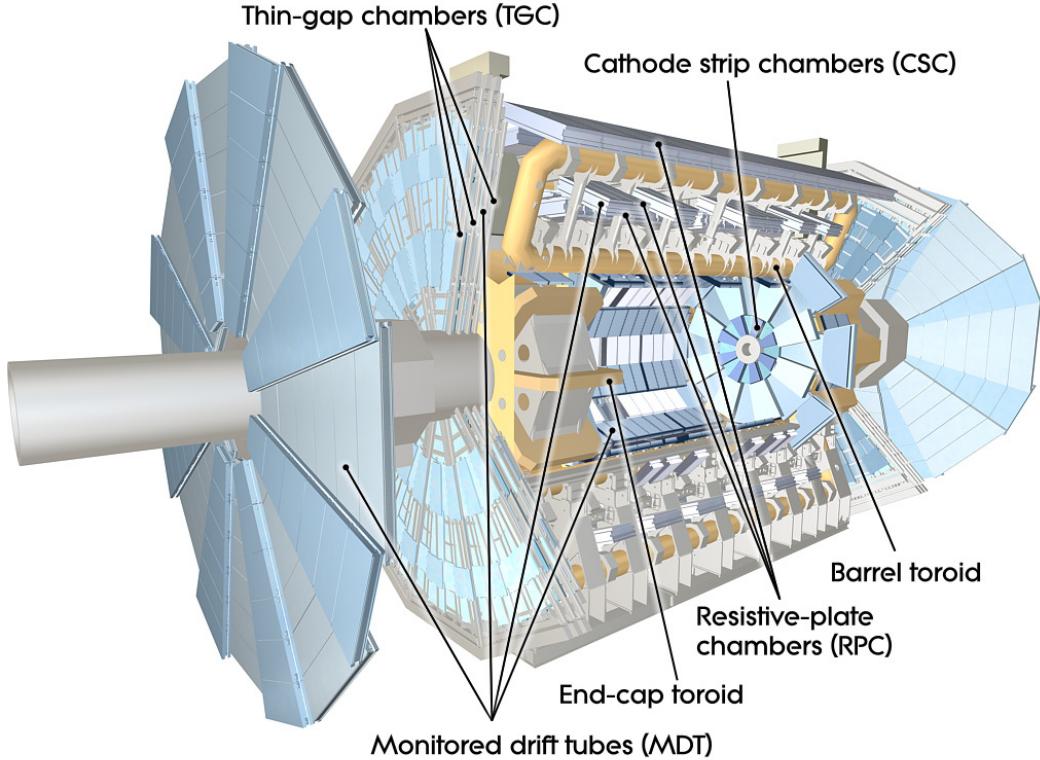


Figure 3.7: Diagram of the ATLAS muon system

779 large portion of the bandwidth to general purpose single lepton triggers (~ 250 Hz). The presence of
 780 leptons in the event indicates the presence of the weak or electro-magnetic interaction and therefore
 781 occurs at many order of magnitude less frequently than interactions involving the strong interaction.
 782 Moreover, many interesting physics signatures that are analyzable by ATLAS involve leptonic final
 783 states. The remaining bandwidth is allocated to jet, missing energy, tau, and unbiased supporting
 784 triggers.

785 The ATLAS trigger system is composed of 3 levels: level-1 (L1), level (L2), and the event filter
 786 (EF). The first level is hardware only trigger that reduces the input 20 MHz rate to ~ 75 kHz, selecting
 787 1 out of every 250 collisions. The available buffering on the FPGA chips means that the decisions
 788 need to be made within $2.5 \mu\text{s}$. The L1 selection is based on calorimeter clustering and tracking
 789 finding in the MS for small areas of the detector called regions-of-interest (ROIs). It selects ROIs
 790 with significant energy.

791 The second and third stages L2 and EF are software based. The L2 algorithms perform more
 792 detailed object reconstruction for leptons, jets and photons inside of the ROIs provided by L1, by

793 performing tracking and in depth calorimeter clustering algorithms. The decisions are made within 50
794 ms per event and pass 1 out of every 15 events to the EF. At the EF, events undergo full reconstruction
795 using similar but faster versions of the algorithms used offline. The EF makes decisions on the presence
796 of fully id-objects in the event and event topological quantities within 4s to reduce the L2 output by
797 a factor of 10. The events that pass this stage are then written to tape for offline study.

798 **3.2.6 Reconstruction: Jets, Muons and Electrons**

799 Physicists analyze the collision event as a collection of identified objects, expressed as momentum
800 4-vectors. The process of converting the disparate detector signatures and signals into a unified 4-
801 momentum description of individual objects is called reconstruction. These objects arise from the
802 final state particles in the event, which can be combined and counted to infer properties of the hard
803 scatter. The particles that make detectable signatures are those that are stable enough to pass through
804 the detector: muons, electrons, photons, and quarks and gluons. Jets and b-tagged jets, muons and
805 electrons are used in the $t\bar{t}H$ analysis to define our search regions and to separate the Higgs signal
806 from backgrounds. Other analyses may use photons, taus and missing energy⁶, but these are not
807 discussed in depth here. Figure 3.8 shows an $R - \phi$ schematic of the interaction of various particle
808 signatures in the ATLAS detector.

809 **3.2.6.1 Tracks and Clusters**

810 The basic components of reconstruction are sensor measurements, or hits, in trackers (ID, MS) and
811 energy measurements in the calorimeter. Hits in the ID and MS undergo pattern recognition, which
812 identifies hits that belong to a single track, and fitting, which fits a curve to the track to assess
813 the particle trajectory. Charged particle trajectories are generally helical in a magnetic field, but
814 the fitting algorithm takes into more detailed information about energy loss to material along the
815 tracks length. The result of the fitting is an estimation of particle momentum 3-vector. Electrons,
816 photons and hadronic particles leave clustered deposits of energy in the EM and hadronic calorimeters
817 from their showers. Electron and photon showers are primarily contained with the EM calorimeter,
818 while hadronic showers are deposit most of their energy in the hadronic calorimeters. The process
819 of associating individual read-out cells of energy in the calorimeter to clusters of energy from the
820 showers of individual particles is called clustering. From the basic pieces of tracks and clusters, more
821 complex objects can be created.

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

822 **3.2.6.2 Electrons**

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

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

839 **3.2.6.3 Muons**

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

847 **3.2.6.4 Jets**

848 Quarks and gluons are colored objects that cannot exist alone on the time scales of detector measure-
849 ments, due to confinement, a property of the strong force . When emitted, they undergo a process
850 called hadronization, in which the convert into 'jets' of colorless hadrons that emerge collimated from
851 the interaction point. The majority of these hadrons are charged and neutral pions, though other
852 hadrons are often present. Jets are reconstructed using conglomerations of calorimeter energy clusters

853 chosen via an anti- k_t algorithm, with a radius of $\Delta R < 0.4$ [53]. The algorithm has been shown to be
854 infrared safe, meaning the jet quantities are not sensitive to low energy, small angle radiative diver-
855 gences. Jets at ATLAS are reconstructed from 10 GeV, calibration of the energy scale and resolution
856 are only available for energies greater than 20-25 GeV.

857 **3.2.6.5 B-Tagged Jets**

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

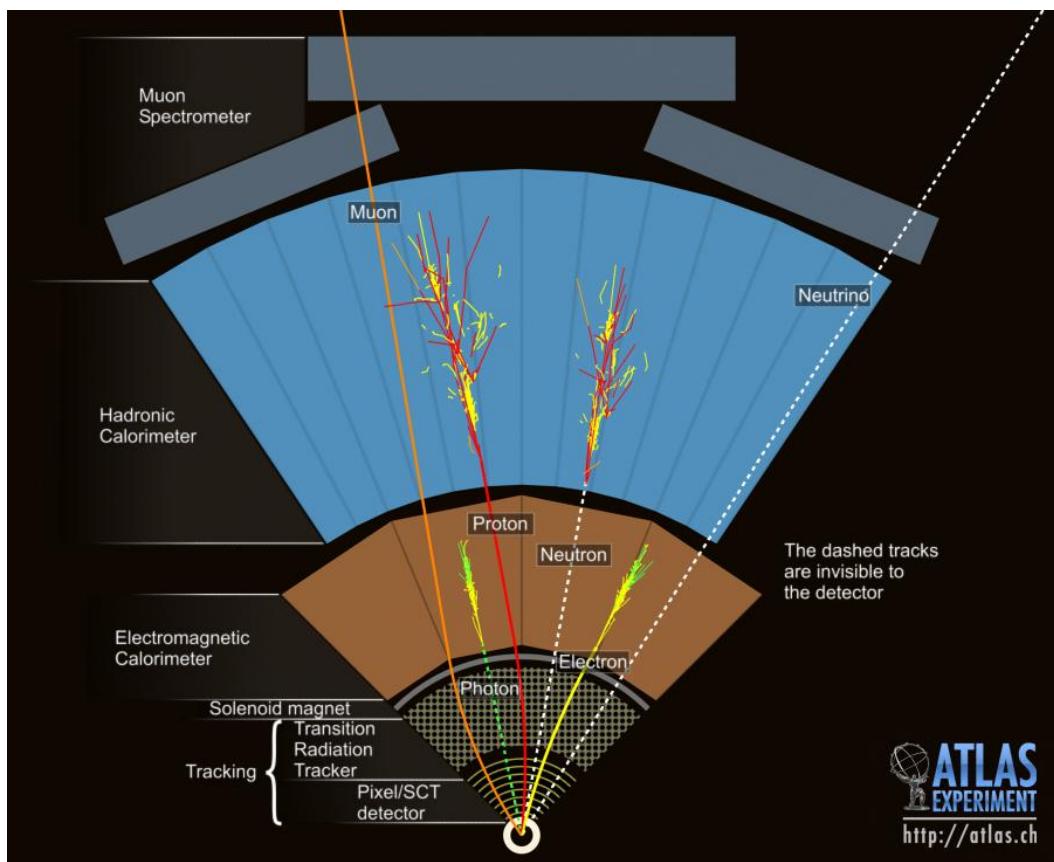


Figure 3.8: $R - \phi$ schematic of the ATLAS detector and various particle signatures

866

CHAPTER 4

867

Electrons

868 High energy electron signatures are important elements of searches and measurements at hadron
869 colliders, because they signal the presence of important electro-weak processes in the event. Requiring
870 well-identified electrons in collision events suppresses the overwhelming rate of strong-force mediated
871 scattering and allows for the collection of a manageable-sized dataset with interesting physics for
872 study. For this reason, electron signatures form one of the two pillars of the HLT trigger at ATLAS,
873 as discussed in Chapter 3, and rigorous electron identification is an important piece of many ATLAS
874 analyses. This section summarizes the development of ATLAS electron identification for the high
875 luminosity 2011 and 2012 datasets and discusses the techniques involved in measuring the electron
876 identification efficiency.

877 **4.1 Identification of Electrons at ATLAS**

878 Electron reconstruction is discussed briefly in Chapter 3 and, in depth here [51]. The result of
879 electron reconstruction is called an electron candidate, which is comprised of a narrow calorimeter
880 energy cluster with $|\eta| < 2.47$ and an ID track that matches loosely in η and ϕ . If the electron
881 has $|\eta| < 2.01$, the ID detector track is fiducial to the TRT and has the possibility of having high-
882 threshold hits, indicative of transition radiation (TR). Electron cluster reconstruction is extremely
883 efficient. The track-matching requirement is less efficient, because the presence of hard bremsstrahlung
884 may in certain cases cause the electron cluster and emitted photon cluster to have a wide separation
885 in the calorimeter [55]. Figure 4.1 shows the reconstruction efficiency as a function $|\eta|$ for an example
886 p_T bin as well as a plot of the amount of material in front of the EM calorimeter. The efficiency loss
887 at high $|\eta|$ is caused mostly by material-induced hard bremsstrahlung.

888 Objects that are not isolated electrons are often reconstructed as electrons, as the reconstruction

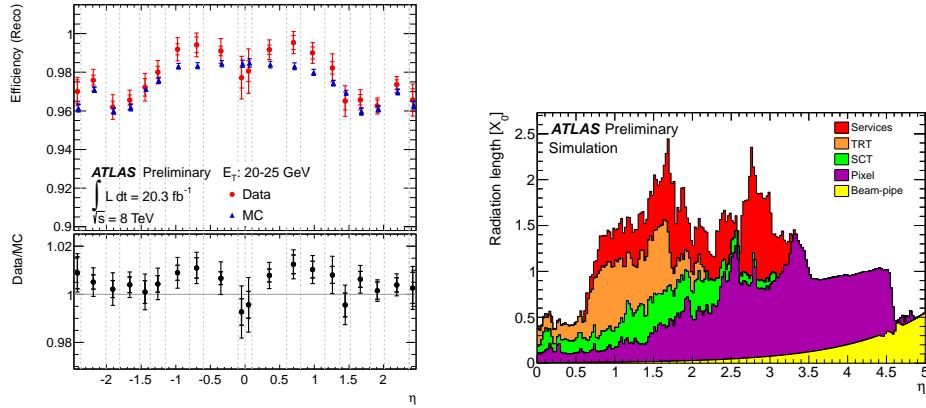


Figure 4.1: Electron reconstruction efficiency for an example p_T bin versus $|\eta|$. The drop in efficiency at higher values of $|\eta|$ is directly attributable to the increase in the amount of material in front of the EM calorimeter (left). The material causes bremsstrahlung, which makes track-cluster matching more difficult for electrons

889 requirements are quite loose. Objects that often ‘fake’ isolated electrons are light quark and gluon
 890 jets, heavy flavor jets that include real decays to electrons, and converted photons. Light quark
 891 and gluon jets fragment into a number of collimated hadronic particles. In rare cases, the jet may
 892 fragment most of its energy into a single charged pion, which showers early in the EM calorimeter
 893 and fakes an electron signature. In other cases, the jet may fragment mostly into a neutral pion,
 894 which subsequently decays into a pair of photons. If one of these photons converts, a track will point
 895 to the EM energy cluster and possibly fake an electron signature. These cases would result in a
 896 reconstructed electron candidate. Although the probability for these misidentifications to happen is
 897 small, the enormous jet production rate means that it is a significant background. In general, light
 898 quark and gluon jet ‘fakes’ have larger transverse shower profiles and more energy leakage into the
 899 hadronic calorimeter. For the neutral pion case, there are generally two separated showers for lower
 900 energy decays. For both cases, there are often other particle signatures nearby. Heavy-flavor jet
 901 decays and photon conversions contain real electrons. However, heavy flavor decays also involve the
 902 production of additional hadronic particles within the jet. Both photon conversion and heavy flavor
 903 decays involved secondary vertices displaced from the primary interaction point.

904 In order to distinguish these fake signatures from real, isolated electrons, electron identification

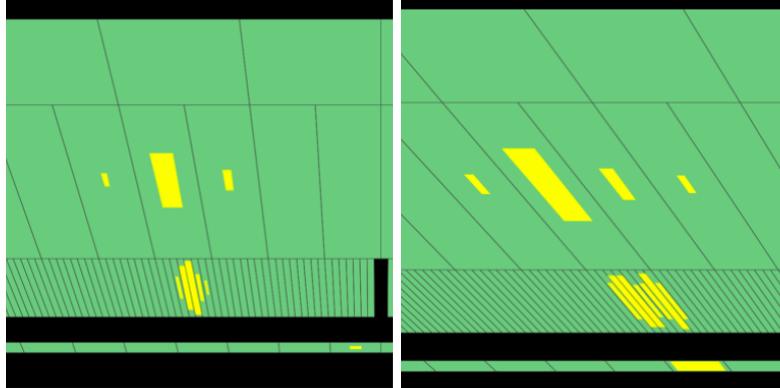


Figure 4.2: Example single photon (left) and $\pi^0 \rightarrow \gamma\gamma$ (right) signatures in the ATLAS EM calorimeter. The fine segmentation of the cells in the strips allows for the distinguishing of two nearby showers from one shower and is used in electron identification

algorithms use a number of reconstructed variables describing the electron shower in the detector and the electron track. The details of the calculated variables can be found here [56]. In general, electron identification takes advantage of the narrowness of isolated electron shower in the transverse plane and lack of energy deposition in the hadronic calorimeter. The transverse variables include measurements of the shower width in both layer 2 and the strips, where more refined measurements are possible. In fact, the strips were designed to separate single photon and electron showers from multiple showers from neutral pion decays, shown in Figure 4.2. The shower width variables are generally measured mostly in η as bremsstrahlung tends to smear the electron energy in ϕ . Electron tracks are required to have an adequate number of hits in the Pixel Detector, SCT and TRT. These hit requirements, especially the b-layer requirement suppress electron conversions which occur in the detector material. Track-cluster matching and geometric impact parameter variables require ID tracks to match the calorimeter energy well and to arise from the primary interaction point. Electrons with tracks explicitly associated with a conversion vertex can be rejected. Finally, the high threshold fraction of hits on the track, made by transition radiation, is an uncorrelated discriminator of pion and electron tracks. Figure 4.3 shows the average high threshold hit probability for pions and electrons as a function of their p_T .

Electron identification algorithms make selections in 9 bins of $|\eta|$, [0.10, 0.60, 0.80, 1.10, 1.37, 1.52, 1.81, 2.01, 2.37, 2.47] and bins of p_T , [7, 10, 15, 20, 30, 40, 50, 60, 70, 80+] GeV. The $|\eta|$ binning changes with the calorimeter geometry, which in turn affect the shower shape distributions. The shape of most of the identification variable distributions, tracking and calorimeter, are p_T dependent.

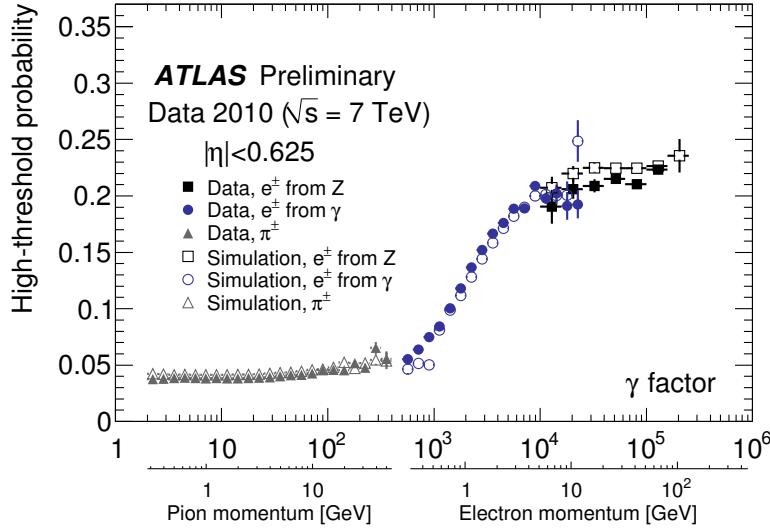


Figure 4.3: Electron and pion track high threshold probabilities as a function of their transverse momenta. The two scales are united through the γ -factor, on which the TR probability depends directly. Electron tracks are much more likely to have high threshold hits at electron energies typical of electro-weak decays.

925 4.1.1 2011 Menu

926 Electron identification in 2011 was accomplished through rectangular cuts on the identification vari-
 927 ables at 3 operating points: Loose, Medium and Tight. The medium operating point was used online
 928 as the primary electron trigger. At the beginning of the 2011 run, the 3 operating points possessed the
 929 same cut-values, but tighter operating points had cuts on more variables. The Loose operating point
 930 only cut on shower shape variables in layer 2 and hadronic leakage, the medium operating point added
 931 cuts on shower width variables in the strips, and tight added TR cuts, strict-track cluster matching,
 932 conversion rejection and a b-layer requirement. This menu, called the ‘IsEM’ menu, was the first fully
 933 data-optimized cut menu for electrons.

934 The demands of increasing luminosity demanded a tightening of the medium operating point
 935 midway through the data-taking, in order to maintain a EF trigger rate of around 20-25 Hz on the
 936 primary electron trigger. To accomplish this, variables cut on at the tight operating point were added
 937 to the medium operating point, and the entire set of cuts was optimized to provide the targeted fake
 938 rejection and reduction in the trigger rates at the highest possible efficiency. The same procedure was
 939 applied to the loose operating point, where the target was to provide an efficiency of 95% and the
 940 highest possible fake rejection. The re-inventing of the menu in this way allowed for not only better

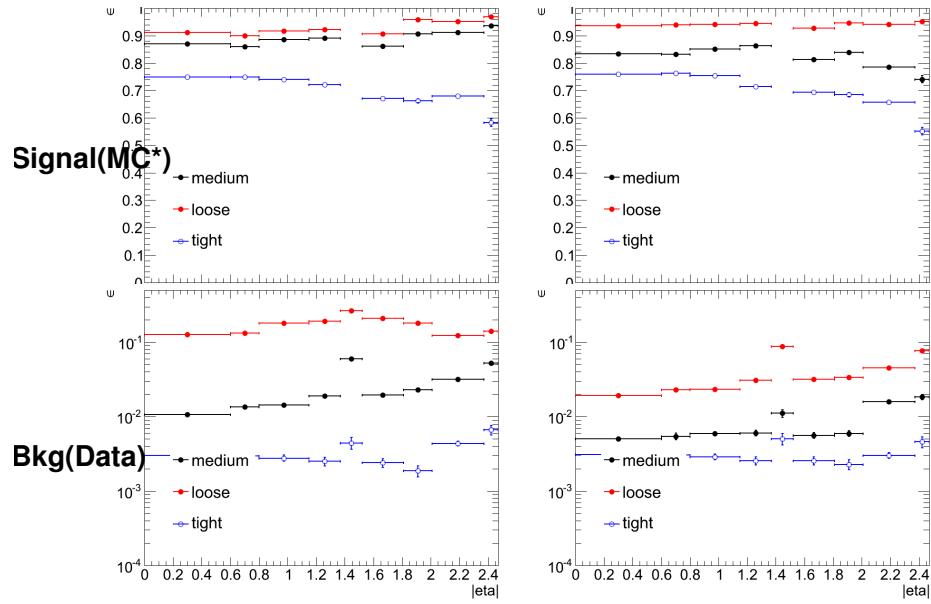


Figure 4.4: Comparison of the ‘IsEM’ (left) and ‘IsEM++’ electron identification operating points for Loose, Medium, and Tight. The efficiency as a function of η for an example p_T bin is shown on top and the background rejection is shown on the bottom

941 performance, due to the inclusion of more variables, but a more stable tightening of the backgrounds
 942 from loose to medium to tight, where the same background types were targeted at each level. The new
 943 menu was called the ‘IsEM++’ menu and the operating points were renamed ‘Loose++’, ‘Medium++’
 944 and ‘Tight++’. Figure 4.4 shows the comparison of the operating points for the new menu and old
 945 menu.

946 4.1.2 2012 Menu and Pile-up

947 Improvements in the running conditions for 2012, in particular narrowing the transverse beam emit-
 948 tance and size, resulted in large increase in number of proton-proton interactions during every 50
 949 ns bunch crossing. In 2011 the average number of reconstructed primary vertices in each event, an
 950 indicator of the number of interaction per bunch crossing, was around 7, while in 2012 the average
 951 grew to 25. Some events during 2012 running had 40 reconstructed primary vertices.

952 The increase in energy in the calorimeters from these additional collisions, called pile-up, caused
 953 a worsening of the resolution of electron identification variables, particularly the shower shapes and
 954 hadronic leakage. The presumed cause was the increase in the number of showers of low energy
 955 hadronic particles near electrons. Figure 4.5 shows two example distributions, the hadronic leakage

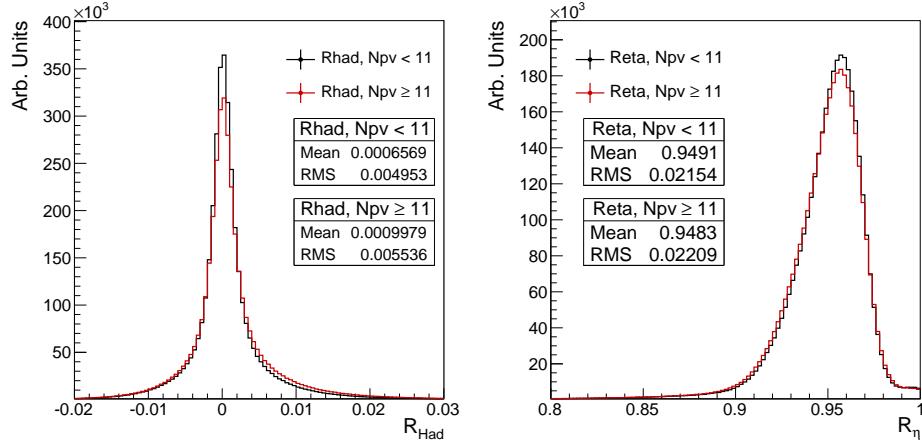


Figure 4.5: Electron hadronic leakage fraction (R_{Had}) and transverse shower profile (R_η) in layer 2 for high and low pile-up conditions. Pile-up is measured here as the number of primary vertices in the event.

(R_{Had}) and the transverse shower profile (R_η), for high and low pile-up conditions. The distributions shows a clear widening for higher pile-up which results in a loss of efficiency.

In order to combat this loss, the ‘IsEM++’ menu was once again optimized to have a flatter efficiency profile as a function of the amount of pile-up in the event with similar performance to the 2011 menu. The strategy for this menu was to loosen selections on variables sensitive to pile-up energy. It was expected and confirmed that relying more on the strip variables for the shower shape selection and the energy in layer 3 of the EM calorimeter for the hadronic leakage would sample a smaller volume of the calorimeter and thus be less sensitive to additional energy in the neighborhood of the electron. The strategy is outlined pictorially in Figure 4.6.

The efficiency of the 2011 operating points compared to the 2012 operating points is shown in Figure 4.7, demonstrating a clear improvement in efficiency of the selections for higher pile-up conditions.

4.1.3 Electron Likelihood

A natural step forward for the electron identification is the use of multi-variate algorithms. Multi-variate identification algorithms use many identification variables at once. Signals and backgrounds can be separated in a multi-dimensional variable space in ways that go beyond simple rectangular cuts. For the case of electron identification, it was found that using a likelihood function, trained with electron identification variables, provided clear performance gains with respect to rectangular cuts, while also providing stable and easily understandable results. The likelihood scores each electron based

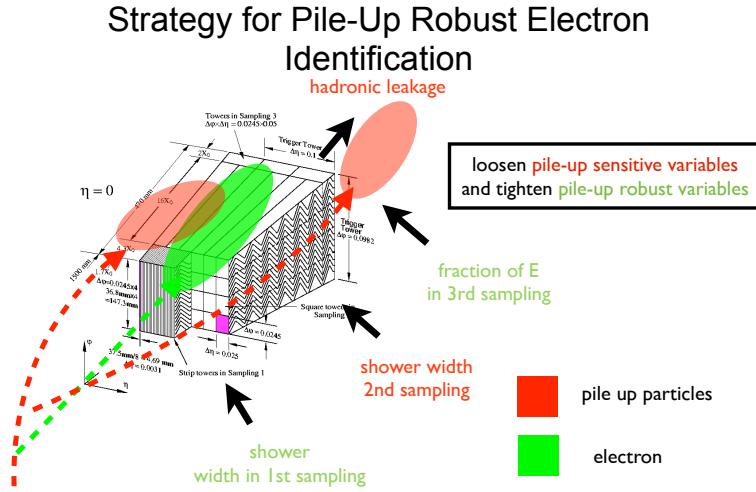


Figure 4.6: Schematic of the strategy to reduce the pile-up dependence of electron identification. A EM calorimeter wedge is shown overlaid with example electron (green) and pile-up particle (red) signatures. The strategy is to loosen the dependence of the identification on layer 2 and the hadronic calorimeter, which sample large volumes, and tighten selection on variables in the the 3rd layer and strips, which sample smaller volumes.

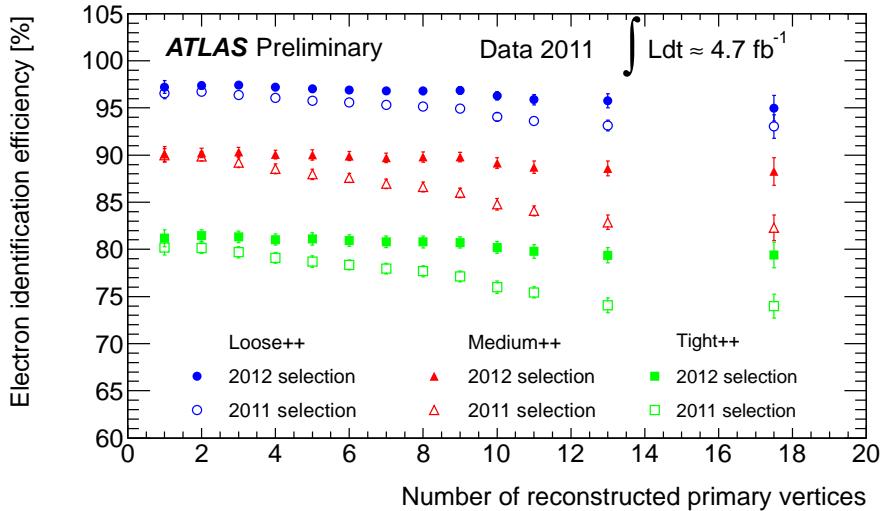


Figure 4.7: Comparison of the efficiency of the 2011 and 2012 electron identification menus versus the number of primary vertices in the event.

975 on how signal-like or background-like it is for each identification variable and then multiplies these
 976 individual scores together into a final score. Figure 4.8 shows the example output of the likelihood
 977 for real electrons and fake electrons. The output distributions can be cut on continuously to produce
 978 a curve of possible selections rather than a single selection point.

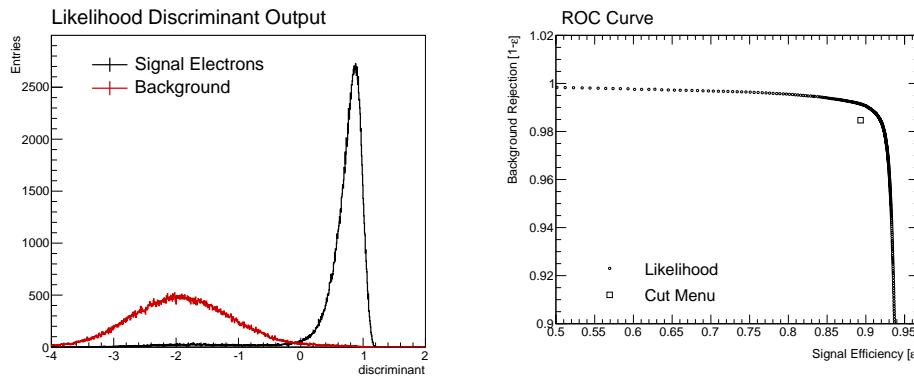


Figure 4.8: Example electron likelihood score output for electrons and fake electrons (left). A cut can be made at any point in this distribution to define a selection operating point (right). The cut-based operating point lies within the curve of possible likelihood operating points, showing that the likelihood indeed outperforms the cuts.

979 There are many advantages to a likelihood-based approach. First, variables that show significant
 980 shape differences between real and fake electrons but do not have a clear cut point can still be used
 981 in a likelihood. Second, the likelihood score takes into account the entire shape of the distribution
 982 and not simply an efficiency and fake rejection at a single cut point. Finally, the final cut on the
 983 likelihood output score can be tuned easily to achieve a desired efficiency and rejection in a way that
 984 does not overly bias selection on a single variable.

985 The likelihood menu for ATLAS was developed at the end of the 2012 run to be used on advanced
 986 2012 analyses. The menu uses similar variables to the cut-based menus but adds a few additional
 987 ones, including a measurement of the amount of energy the electron track lost as it traversed the ID.
 988 The likelihood menu makes cuts on the likelihood output score at 4 different operating points with
 989 the same binning as the cut menu but tunes the cuts based on the number of primary vertices in the
 990 event. This tuning ensures a stable response of the identification with varying degrees of pile-up.
 991 The likelihood menu greatly outperforms the rejection of the ‘IsEM++’ menu for similar efficiencies.
 992 Figure 4.9 shows a comparison of performance the cut-based and likelihood tight regime operating
 993 points. The likelihood menu, specifically the `VERYTIGHT` operating point, is used in the $t\bar{t}H$ analysis.

994 4.2 Measurement of Electron Identification Efficiency at ATLAS

995 Precise measurements of the electron identification efficiency are important pieces of many ATLAS
 996 analyses, including the $t\bar{t}H$ multi-leptons analysis. For analyses with low p_T leptons, systematic
 997 uncertainties on the electron identification efficiency can be some of the largest systematic effects.

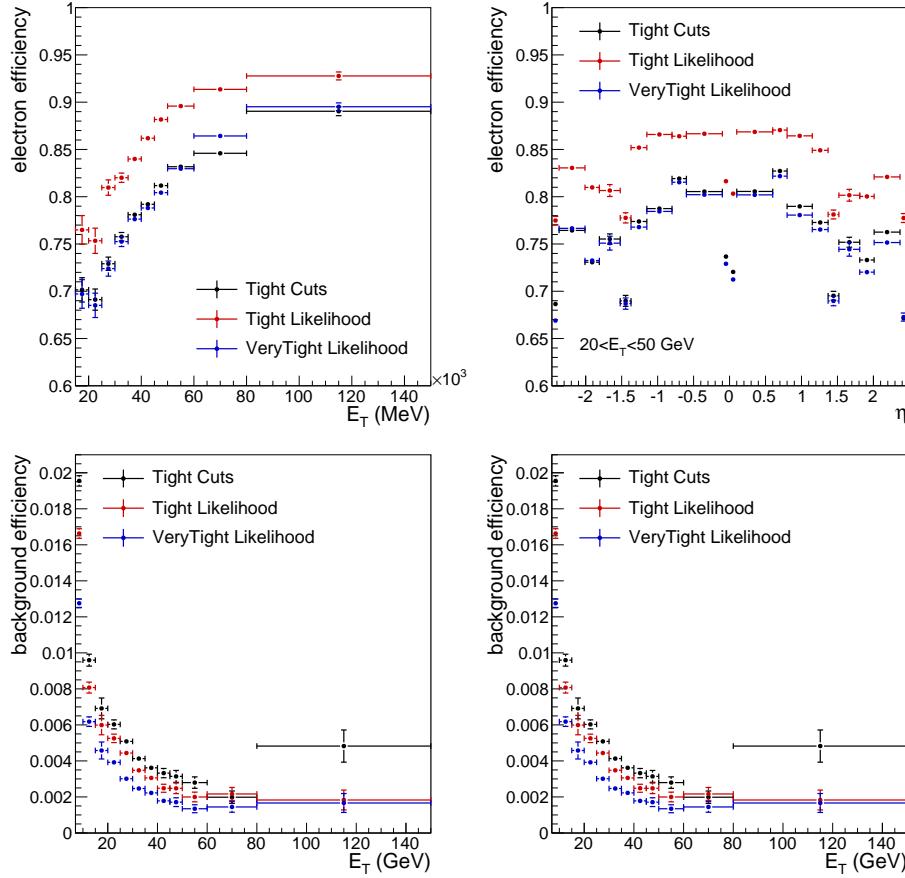


Figure 4.9: Comparison of the performance of the cut-based and likelihood operating points in the tight regime. Efficiency (top) and rejection (bottom) plots are shown versus $|\eta|$ and E_T

998 The methods used to measure the electron identification efficiency are described in depth here [51].
 999 Electron identification efficiencies are measured using a method called tag-and-probe for J/Ψ and
 1000 Z boson decays to electrons. One object from the decay is ‘tagged’, or fully identified, while the other
 1001 is left unidentified. There is reasonable confidence that the second object is an electron based on the
 1002 kinematic properties of the event, specifically the di-electron invariant mass is near the Z or J/Ψ pole.
 1003 The tag-and-probe method leaves a sample of unidentified and unbiased ‘probes’, where the efficiency
 1004 can be measured.

1005 As opposed to muons, contamination from fake electron make the tag-and-probe method difficult.
 1006 Backgrounds from fake electrons are subtracted using fits to the Z and J/Ψ invariant mass distribu-
 1007 tions. For Z electrons, fits to the electron isolation distribution are also used. The final efficiencies
 1008 reported are the result of statistical fit among all methods. The uncertainties are at low momenta are
 1009 around $\sim 5\%$ and are dominated by systematics effects from large background subtractions. They

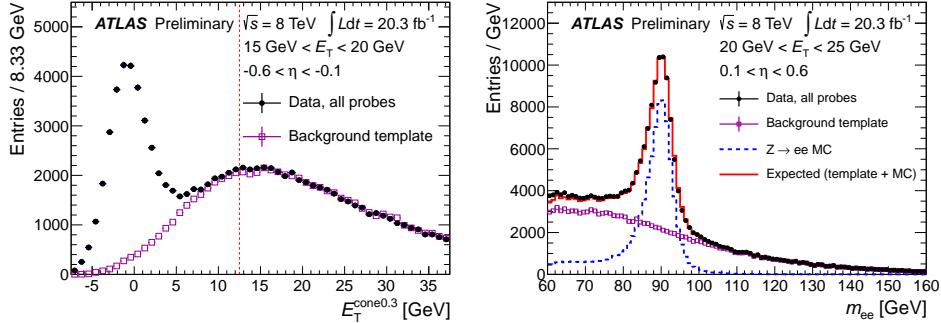


Figure 4.10: Example bins where the electron probe distribution from Z tag-and-probe is fit by an isolation template (left) invariant mass template (right) to subtract backgrounds.

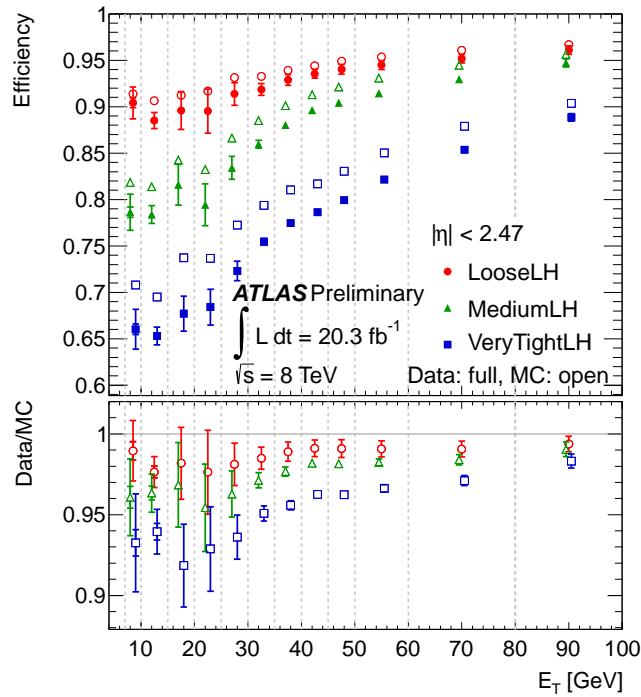


Figure 4.11: Electron identification efficiency calculated in data and MC versus electron p_T .

1010 are less than 1% at high momenta and dominated by tag-and-probe selection effects. The efficiency
1011 can be seen in Figure 4.11.

1012 The development of low p_T electron scale-factors was an integral piece in extending the sensitivity
1013 of Higgs searches in the WW and ZZ decays modes. Early efforts in 2012 to provide consistent,
1014 well-measured efficiencies in this region were complicated by large disagreements in the efficiencies
1015 obtained from the three different estimation methods: Z tag-and-probe with isolation background
1016 subtraction, Z tag-and-probe with invariant mass background subtraction, and J/Ψ tag-and-probe.
1017 This was especially true for electrons with p_T of 10-20 GeV. The energy scale of Z and J/Ψ decays
1018 disfavor electrons of this momentum, and backgrounds are high, revealing problems disguised in higher
1019 purity regions. The background subtraction methods were studied in depth to assess possible biases,
1020 and a new lower statistics but high purity tag-and-probe method using radiative $Z \rightarrow ee\gamma$ decays was
1021 developed. The result of these studies was the development of new background subtraction templates
1022 for the isolation and invariant mass subtraction and an appropriate uncertainty to cover the systematic
1023 effects from the biases of these subtractions.

1024

CHAPTER 5

1025

Analysis Summary

1026 This chapter provides an overview the of analysis searching for SM production of the Higgs boson in
1027 association with top quarks in multi-lepton final states. The analysis searches in signal regions (SRs)
1028 with 2 same-sign, 3 and 4 light leptons (e, μ), which are sensitive to Higgs decays to vector bosons,
1029 $H \rightarrow W^\pm W^\pm$ and $H \rightarrow Z^\pm Z^\pm$. We refer to these channels as 2ℓ SS, 3ℓ , and 4ℓ through the rest of
1030 this document.

1031 The multi-lepton channels form a complement to already completed $t\bar{t}H$ searches in final states
1032 targeting the $H \rightarrow b\bar{b}$ [57], $H \rightarrow \gamma\gamma$ [58]. The $t\bar{t}H$ searches in the $H \rightarrow \tau\tau$ decay modes were
1033 developed concurrently with the multi-lepton searches, but we do not discuss these here. Of this set
1034 of complementary searches, the multi-lepton and $b\bar{b}$ are the most sensitve.

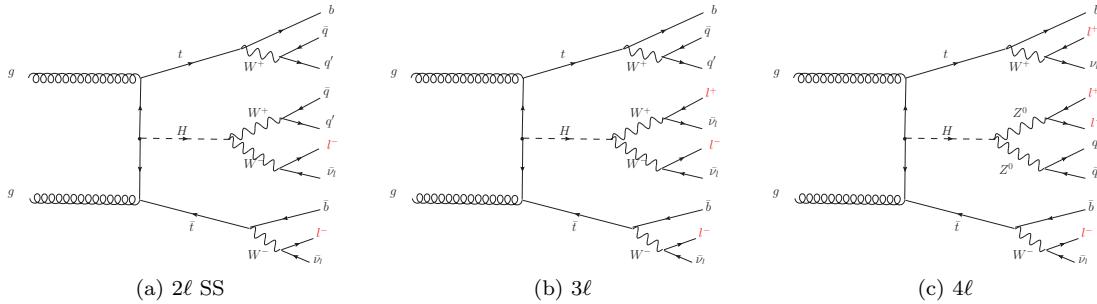
1035 Based on SM production cross-sections, observation lies just outside the sensitivity of the Run I
1036 dataset, even when combining all searches. Instead, the analyses provide an opportunity to constrain
1037 for the first time the $t\bar{t}H$ production mode with limits reasonably close to the actual production rate.
1038 The multi-lepton analysis is therefore optimized to overall sensitivity to the $t\bar{t}H$ production rather
1039 than individual decay modes, which would be more useful for constraining Higgs couplings.

1040 Detailed description of the event and objection section are provided in Chapter 7, background
1041 modeling in Chapter 8, the effect of systematic errors and the statistical analysis in Chapter 9 and
1042 final results in Chapter 10.

1043

5.1 Signal Characteristics

1044 The signal is expected to be characterized by the presence of 2 b-quark jets from the top quark
1045 decays, isolated leptons from vector boson and tau decays, a high jet multiplicity and missing energy
1046 from neutrinos. Three Higgs boson decays are relevant for this analysis: W^+W^- , $\tau^+\tau^-$ and ZZ .

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

1047 All modes are generally dominated by the WW signature, though the 3ℓ and 4ℓ channels possess
 1048 some contribution from the $\tau\tau$ and ZZ decays. Table 5.1 provides the fractional contribution of the
 1049 main Higgs decay modes at the generator level to $t\bar{t}H$ search channels and Figure 5.1 shows example
 1050 diagrams for each channel. In general, the number of leptons is anti-correlated with the number of
 1051 jets, since a vector boson can either decay leptonically or hadronically, such that:

- 1052 • in the 2ℓ SS channel, the $t\bar{t}H$ final state contains 6 quarks⁷. These events are then characterized
 1053 by the largest jet multiplicity.
- 1054 • In the 3ℓ , the $t\bar{t}H$ final state contains 4 quarks
- 1055 • In the 4ℓ channel, the $t\bar{t}H$ final state contains a small number of light quarks, 0 ($H \rightarrow$
 1056 W^+W^- case), 2 or 4 ($H \rightarrow ZZ$ case).

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

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

5.2 Background Overview

1058 For all channels after selection, the size of the signal is of similar order to the expected size of
 1059 background. Background processes can be sorted into two categories:

⁷this does not include additional quarks from radiation

- 1060 • **Reducible:** These processes cannot lead to a final state compatible with the signal signature
1061 without a mis-reconstructed object. This category includes events with a prompt lepton but
1062 with mis-reconstructed charge and events with jets that "fake" leptons. The main backgrounds
1063 of this sort are $t\bar{t}$ and $Z+jets$. Data-driven techniques are used to measure the rate of these
1064 processes and strict object selection and isolation requirements are used to reduce their rate.
- 1065 • **Irreducible:** Events which can lead to the same final state as the signal. The main background
1066 of this category are: vector boson production (V) associated with top quarks ($t\bar{t}V$), a Z boson
1067 produced in association with a top quark (tZ), $W^\pm Z$, and ZZ . They are modeled using the
1068 Monte Carlo simulations. In general, these backgrounds are combatted with jet and b-tagged
1069 jet requirements. Although the jet multiplicity of $t\bar{t}V$ is high, the multiplicity of $t\bar{t}H$ events is
1070 still higher.

1071 5.3 Analysis Strategy

1072 The analysis search is conducted in 3 channels, based on counting of fully identified leptons: 2ℓ SS,
1073 3ℓ , and 4ℓ , with cuts optimized separately for each. We further divide the 2ℓ SS into sub channels
1074 based on the number of jets and flavor of the leptons and the 4ℓ channel into sub-channels enriched
1075 and depleted in opposite-sign (OS) leptons arising from Z decays.

1076 This analysis is a counting experiment, meaning that the only quantities significant to the result
1077 are the event counts in the signal regions and not the event shapes. The measured background rates,
1078 expected signal rates and systematic uncertainties are fed into a Poisson model and fit to the observed
1079 data. The parameter of interest in the fit and the result of this measurement is, μ , the ratio of the
1080 fitted number of $t\bar{t}H$ events in the signal regions to expected number of $t\bar{t}H$ events in the signal
1081 regions. Since we assume SM branching ratios, μ can be considered the ratio of the measured $t\bar{t}H$
1082 cross-section to the observed $t\bar{t}H$ cross-section, and we expect the fitted μ to be close to 1 with large
1083 statistical errors.

1084 We express the final result as a measurement of μ with uncertainties and 95% upper limit on the
1085 value of μ : μ -values higher than this value will be considered excluded. We provide these results for
1086 each channel individually and combined.

1087 CHAPTER 6

1088 Dataset and Simulation

1089 6.1 Data

1090 6.1.1 The 2012 Dataset

1091 The $t\bar{t}H$ analysis uses the entire 2012 ATLAS dataset only, collected from April to December. The
1092 size of the dataset corresponds to 20.3 fb^{-1} , after passing data quality requirements, ensuring the
1093 proper operation of the tracking, calorimeter and muon subsystems. The LHC successfully produced
1094 datasets for physics studies in 2010, 2011 and 2012. The 2012 proton-proton dataset was delivered
1095 with collisions with a CME of 8 TeV with bunch collisions every 50 ns[59].

1096 Figure 6.1 shows the accumulation of the 2012 dataset over time. Despite doubling the bunch
1097 spacing above the design of 25 ns, the luminosity neared the design luminosity due to unexpected
1098 improvements in the transverse beam profile[60]. This increased the amount of pile-up, or number
1099 of collisions per bunch crossing and in general collision events were busier due to these multiple
1100 interactions. Figure 6.2 shows the average number of interaction per bunch crossing for the 2011 and
1101 2012 datasets. The 2012 dataset shows an average of 20-25 interactions.

1102 The dataset must contain either a primary muon or primary electron trigger (`EF_e24vhi_medium1`
1103 OR `EF_e60_medium1` OR `EF_24i_tight` OR `EF_36_tight`). The electron triggers require a electron with
1104 at least 25 GeV of calorimeter energy, passing the medium identification requirement and loose tracking
1105 isolation. Above 60GeV, the isolation requirement is dropped and the identification is loosened slightly.
1106 The muon trigger requires a good inner detector track and matching hits in the muon spectrometer,
1107 as well as loose tracking isolation, which is also dropped about 36 GeV.

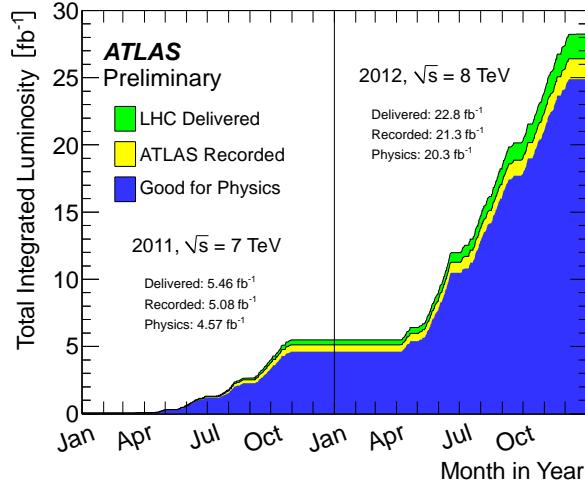


Figure 6.1: Plot showing the accumulation of the integrated luminosity delivered to the ATLAS experiment over 2011 and 2012. The rough size of the usable, physics ready dataset for 2012 is 20 fb^{-1} and is the dataset used.

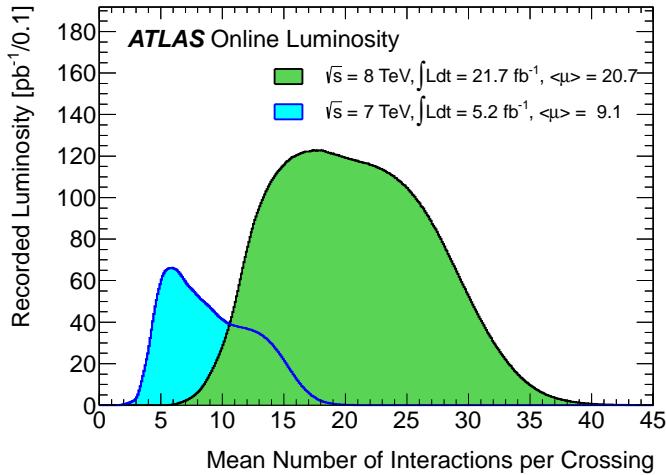


Figure 6.2: The average number of interactions per bunch-crossing for the 2012 and 2011 LHC proton-proton dataset. Most of these interactions are uninteresting but leave energetic signatures in particle detectors called pile-up which interfere with measurements

1108 6.2 Simulation

1109 Simulation samples are used to determine the overall event selection acceptance and efficiency and
 1110 model the number of events in the signal regions for prompt backgrounds and signal. The simulated
 1111 samples are created using parton distribution function (PDF) and use Monte Carlo (MC) techniques
 1112 to model the hard parton scatter, underlying event activity and parton showering and hadronization.

Table 6.1: Monte Carlo samples used for signal description.

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

1113 The samples are then passed through a full ATLAS detector simulation[61] based on GEANT4 [62].
 1114 Small corrections are then applied to re-scale object identification efficiencies, energy scales, and the
 1115 pile-up based on control regions from data. These corrections are discussed in Chapter 9.

1116 6.2.1 Signal Simulation

1117 The $t\bar{t}H$ production is modeled using matrix elements obtained from the HELAC-Oneloop package [63]
 1118 that corresponds to the next-to-leading order (NLO) QCD accuracy. Powheg BOX [64, 65, 66] serves
 1119 as an interface to the parton shower Monte Carlo programs. The samples created using this approach
 1120 are referred to as PowHel samples. CT10NLO PDF sets are used and the factorization (μ_F) and
 1121 renormalization (μ_R) scales are set to $\mu_0 = \mu_F = \mu_R = m_t + m_H/2$. Pile-up and the underlying
 1122 events are simulated by Pythia 8.1 [67] with the CTEQ61L set of parton distribution functions and
 1123 AU2 underlying event tune. The Higgs boson mass is set to 125 GeV and the top quark mass is set
 1124 to 172.5 GeV.

1125 The signal Monte Carlo samples are summarized in Table 6.1. These large samples are generated with
 1126 inclusive Higgs boson decays with branching fractions set to the LHC Higgs Cross Section Working
 1127 Group (Yellow Report) recommendation for $m_H = 125$ GeV [68]. The inclusive cross section (129.3
 1128 fb at $m_H = 125$ GeV) is also obtained from the Yellow Report [68].

1129 6.2.2 Background Simulation

1130 The background simulations used for this analysis are listed in Table 6.2. In general, the Alpgen[69],
 1131 MadGraph[70], and AcerMC[71] samples use the CTEQ6L1[72] parton distribution function, while
 1132 the Powheg[73], Sherpa[74], are generated with the CT10 PDF. The exception is the MadGraph $t\bar{t}t\bar{t}$
 1133 sample, which is generated with the MSTW2008 PDF[75]. The highest order calculations available
 1134 are used for the cross sections.

Table 6.2: Monte Carlo samples used for background description. Unless otherwise specified MadGraph samples use Pythia 6 for showering and Alpgen samples use Herwig+Jimmy. $t\bar{t}$, single top, and $Z+jets$ samples are replaced with data-driven estimates for the final result

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

1135

CHAPTER 7

1136

Object and Event Selection

1137 The analysis is divided into 3 signal regions based on lepton counting: 2 same-sign leptons, 3 leptons
 1138 and 4 leptons. The lepton counting occurs for fully identified leptons with full overlap removal with
 1139 transverse momenta over 10 GeV to ensure orthogonality. Lepton selections are tightened afterward
 1140 within each region.

1141 The cuts for each signal region are provided in Table 7.1 and the object selections are detailed in the
 1142 following selections. The selections are based on optimizations of the region sensitivity performed using
 1143 MC (event for data driven backgrounds) and ad-hoc values for normalization systematic uncertainties⁸
 1144 i All signal regions are comprised of three basic requirements: the presence of b-tagged jets, the
 1145 presence of additional light jets, and a veto of same flavor opposite sign leptons with an invariant
 1146 mass within the Z window. Additional requirements on the invariant mass of the leptons, the missing
 1147 transverse energy in the event, and the total object energy (H_T) proved to have negligible additional
 1148 benefit at our level of statistics. Figure 7.1 shows the background and signal fractions as a function
 1149 number of jets and number of b-tagged jets for otherwise fully selected events.

1150 **7.1 2ℓ Same-Charge Signal Region**

1151 The 2 lepton signal region requires two leptons of similar charge (2ℓ SS). The signal is symmetric
 1152 in charge but the background from opposite-sign $t\bar{t}$ di-lepton production would be overwhelming.
 1153 Requiring only two leptons allows the extra 2 W bosons in the event to decay hadronically, resulting
 1154 in on average 4 additional light jets plus 2 additional b-quark jets from the top decays.

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

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

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

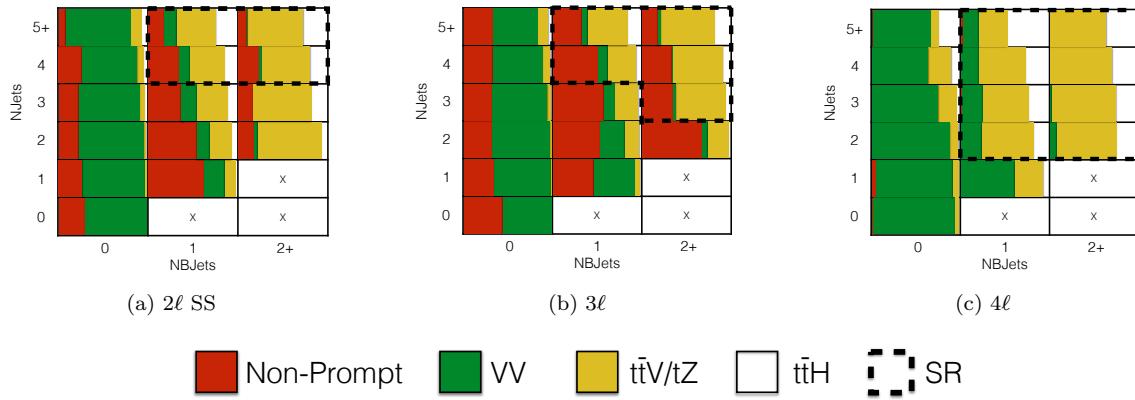


Figure 7.1: Number of jets vs. number of b-tagged jet plot for the fully selected multi-lepton channels. Signal regions are outlined with a dashed line. Sub-channels are defined later in the 2ℓ SS and 4ℓ SRs. The fractional background contribution to each jet and b-tagged jet bin are shown for non-prompt (red), $ttV + tZ$ (yellow), and VV (green). The expected signal fraction is shown in white. The expected non-prompt fraction contains charge misidentifications and fakes. It is shown for MC only, although data-based methods are used for the final result.

1155 We require a leading lepton with transverse momentum of at least 25 GeV that matches to a
 1156 trigger and a sub-leading lepton of at least 20 GeV, a b-tagged jet, and at least 4 jets in total.

1157 In order to suppress non-prompt backgrounds, the lepton isolation criteria for tracking and
 1158 calorimeter are tightened from less than 10% of the lepton momentum to 5%. To suppress charge
 1159 misidentification, the electron is required to be extremely central ($|\eta| < 1.37$) to avoid the material-
 1160 rich regions of the detector. Additionally, ee events with a lepton pair invariant mass within 10 GeV

1161 of the Z pole are removed. To maintain orthogonality with the τ analyses, events with fully identified
 1162 taus are vetoed.

1163 For the statistical combination the channel is divided into 6 sub-channels: 2 jets counting bins
 1164 ($N_{Jet} = 4, N_{Jet} \geq 5$) \times 3 lepton flavor bins (ee, $\mu\mu$,e μ).

1165 7.2 3 ℓ Signal Region

1166 The 3 lepton channel requires 3 leptons, whose summed charge is either -1 or $+1$. The leptons are
 1167 ordered in this way:

- 1168 • **lep0**: the lepton that is opposite in charge to the other two leptons
- 1169 • **lep1**: the lepton that is closer in ΔR to lep0
- 1170 • **lep2**: the lepton that is farther in ΔR from lep1

1171 Since events with a fake lepton arise exclusively from opposite sign di-lepton processes, $t\bar{t}$ and
 1172 Z+jets, where additional jets are misidentified as the third lepton, lep0 is never the fake lepton. As
 1173 a result, the transverse momentum requirement of lep0 (> 10 GeV) is lower than the other two, > 20
 1174 GeV. One lepton must match a trigger and have $p_T > 25$ GeV.

1175 The 3 ℓ channel further requires at least one b-tagged jet and at least 4 jets in total, or two b-
 1176 tagged jets and exactly 3 jets in total. Additionally, to suppress $W^\pm Z$ and Z+jets events, events with
 1177 same-flavor opposite-sign (SFOS) pairs within 10 GeV of the Z pole are vetoed.

1178 Additional cuts, including a di-lepton mass cut, and splittings were investigated but low statistics
 1179 proved to wash out any advantages. The di-lepton mass cut will be a useful discriminator in future
 1180 analyses since the spin statistics of Higgs decay in W bosons often causes the two emitted opposite-sign
 1181 leptons to point in the same direction, resulting in a small measured invariant mass.

1182 7.3 4 ℓ Signal Region

1183 In the four lepton signal region, selected events must have exactly four leptons with a total charge
 1184 of zero. At least one lepton must be matched to one of the applied single lepton trigger and have
 1185 a transverse momentum above 25 GeV. The leading and sub-leading leptons are required to have
 1186 transverse momentum of 25 and 15 GeV respectively. In order to suppress background contributions
 1187 from low-mass resonances and Drell-Yan radiation, all SFOS lepton pairs are required to have a
 1188 dilepton invariant mass of at least 10 GeV.

1189 The four-lepton invariant mass is required to be between 100 and 500 GeV. This choice of mass
 1190 window suppresses background from the on-shell $Z \rightarrow 4\ell$ peak and exploits the high-mass differences
 1191 between the signal and the dominant $t\bar{t}Z$ background. Events containing an OS-SF lepton pair
 1192 within 10 GeV of the Z boson mass are discarded. This Z-veto procedure greatly reduces background
 1193 contributions from ZZ and $t\bar{t}Z$. Finally, selected events are required to have at least two jets, at
 1194 least one of which must be tagged as a b-quark jet.

1195 The contribution from $t\bar{t}Z$ comprises approximately 75% of the total background in the inclusive
 1196 signal region. A signal region categorization which factorizes $t\bar{t}Z$ from the remaining backgrounds is
 1197 thus beneficial. The signal region is accordingly divided into two categories based on the presence of
 1198 OS-SF lepton pairs in the final state.

1199 7.4 Electron Selection

1200 The electrons are reconstructed by a standard algorithm of the experiment [51] and the electron
 1201 cluster is required to be fiducial to the barrel or endcap calorimeters: $|\eta_{cluster}| < 2.47$. Electrons in
 1202 the transition region, $1.37 < |\eta_{cluster}| < 1.52$, are vetoed. Electrons must have $p_T > 10$ GeV and pass
 1203 the **VERYTIGHT** likelihood identification criteria.

1204 In order to reject jets misidentified as electrons, electron candidates must also be well isolated
 1205 from additional tracks and calorimeter energy around the electron cluster. Both the tracking and
 1206 calorimeter energy within $\Delta R = 0.2$ of the electron cluster must be less than 5% of the electron trans-
 1207 verse momentum: $\text{ptcone20}/P_T < 0.05$ and $\text{Etcone20}/E_T < 0.05$. All quality tracks with momentum
 1208 greater than 400 MeV contribute to the isolation energy. Calorimeter isolation energy is calculated
 1209 using topological clusters with corrections for energy leaked from the electron cluster. Pile-up and
 1210 underlying event corrections are applied using a median ambient energy density correction.

1211 The electron track must also match the primary vertex. The longitudinal projection of the track
 1212 along the beam line, $z_0 \sin \theta$, must be less than 1 cm) and the transverse projection divided by the
 1213 parameter error, $d0$ significance, must be less than 4. These cuts are used in particular to suppress
 1214 backgrounds from conversions, heavy-flavor jets and electron charge-misidentifications.

1215 The electron selection is summarized in Table 7.2.

1216 7.5 Muon Selection

1217 Muons used in the analysis are formed by matching reconstructed inner detector tracks with either
 1218 a complete track or a track-segment reconstructed in the muon spectrometer (MS), called Chain 2

1219 muons. The muons have $p_T > 10$ GeV and satisfy $|\eta| < 2.5$. The muon track are required to be a good
 1220 quality combined fit of inner detector hits and muon spectrometer segments, unless the muon is not
 1221 fiducial to the inner detector, $|\eta| > 2.47$. Muons with inner detector tracks are further required to
 1222 pass standard inner detector track hit requirements [52].

1223 As with electrons, muons are required to be isolated from additional tracking or calorimeter energy:
 1224 $\text{ptcone20}/P_T < 0.1$, $\text{Etcone20}/E_T < 0.1$. A cell-based Etcone20/ P_T relative isolation variable is used.
 1225 A pile-up energy subtraction based on the number of reconstructed vertices in the event is applied.
 1226 The subtraction is derived from a Z boson control sample.

1227 The muons must also originate from the primary vertex and have impact parameter requirements,
 1228 d_0 significance < 3 , and $z_0 \sin \theta < 0.1$ cm, similar to the electrons.

1229 The muon selection is summarized in Table 7.2.

1230 7.6 Jet and b-Tagged Jet Selection

1231 Jets are reconstructed in the calorimeter using the anti- k_t [53] algorithm with a distance parameter of
 1232 0.4 using locally calibrated topologically clusters as input (LC Jets). Since the jets in the $t\bar{t}H$ signal
 1233 mostly arise from the decay massive resonances and not radiation, they are expected to be central
 1234 and high energy. Jets must have $p_T > 25$ GeV and $|\eta| < 2.5$.

1235 Jets must also pass loose quality requirement, ensuring the proper functioning of the calorimeter
 1236 at the time of data taking. Jets near a hot calorimeter cell in data periods B1/B2 are rejected. The
 1237 local hadronic calibration is used for the jet energy scale, and ambient energy corrections are applied
 1238 to account for energy due to pileup.

1239 Jets within $|\eta| < 2.4$ and $p_T < 50$ GeV are further required to be associated with the primary
 1240 vertex. The the fraction of track p_T associated with the jet that comes from the primary vertex, must
 1241 exceed 0.5 (or there must be no track associated to the jet). This requirement rejects jets that arise
 1242 from pile-up vertices.

1243 B-jets are tagged using a Multi-Variate Analysis (MVA) method called MV1 and relying on infor-
 1244 mation of the impact parameter and the reconstruction of the displaced vertex of the hadron decay
 1245 inside the jet[54]. The output of the tagger is required to be above 0.8119 which corresponds to a
 1246 70% efficient Working Point (WP).

¹²⁴⁷ 7.7 Object Summary and Overlap

¹²⁴⁸ Since many fully identified objects may be reconstructed as two different objects, an overlap removal
¹²⁴⁹ procedure is applied. Electrons within $\Delta R < 0.1$ of muons are rejected in favor of the muon. Jets
¹²⁵⁰ within $\Delta R < 0.3$ of electrons are then removed. Finally, muons within $\Delta R < 0.04 + 10\text{GeV}/p_T$ of
¹²⁵¹ jets are rejected, as these muons are thought to arise from jet fragmentation.

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

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

1252

CHAPTER 8

1253

Background Estimation

1254 The $t\bar{t}H$ multi-lepton signal regions discussed in Chapter 5 are contaminated by background con-
 1255 tributions at a similar order of magnitude to the signal. The dominant background for each region
 1256 is $t\bar{t}V$. Sub-dominant but important backgrounds include the production of vector boson pairs in
 1257 associated with jets and b-quark jets (VV) and $t\bar{t}$ production with a jet misidentified as a lepton
 1258 (fakes). The 2ℓ SS regions possesses a unique background of charge misidentification from Z and top
 1259 events. The methods for estimating these backgrounds are discussed in this chapter. Monte Carlo
 1260 simulation is used for the prompt $t\bar{t}V$ and VV contributions. The non-prompt backgrounds from $t\bar{t}$
 1261 jet-misidentification and charge-misidentification are estimated using data-driven methods. Table 8.1
 1262 provides a summary of the $t\bar{t}H$ signal and background expectation for each of the signal regions,
 1263 including the data-driven estimates discussed in this section.

1264 **8.1 Vector Boson (W^\pm , Z) production in association with top quarks:**

1265 $t\bar{t}V$, tZ

1266 Production of top quarks plus vector boson is an important background in all multi-lepton channels.
 1267 A large part of the $t\bar{t}V$ component, arising from on-shell $Z \rightarrow \ell\ell$, can be removed via a Z mass veto
 1268 on like-flavor, opposite sign leptons. However the $Z \rightarrow \tau\tau$ and γ^* components remain. The $t\bar{t}W^\pm$ and
 1269 tZ processes generally require extra jets to reach the multiplicity of our signal regions, as such it is
 1270 important to ascertain uncertainties associated with QCD radiation. We consider uncertainties on both
 1271 the $t\bar{t}W^\pm$ and $t\bar{t}Z$ production cross-sections of these two processes and event selection efficiencies in
 1272 the signal regions. The latter is sensitive to the NJet modelling in the MC. We assess the size of these
 1273 uncertainties by investigating the effects of the choice of the factorization (μ_F) and renormalisation
 1274 μ_R scales and PDF sets.

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 $t\bar{t}$ fake).

Table 8.2: NLO cross section and theoretical uncertainty calculations derived from MadGraph5+aMC@NLO.

Process	$\sigma_{NLO} [fb]$	Scale Uncertainty [%]		PDF Uncertainty [%]		Total symmetrized uncertainty [%]
$t\bar{t}W^+$	144.9	+10	-11	+7.7	-8.7	13.3
$t\bar{t}W^-$	61.4	+11	-12	+6.3	-8.4	13.6
$t\bar{t}Z$	206.7	+9	-13	+8.0	-9.2	14.0
tZ	160.0	+4	-4	+7	-7	8.0
tZ	76.0	+5	-4	+7	-7	8.6

1275 Monte Carlo events for these processes are generated with MadGraph 5 and showered with Pythia
 1276 6. $t\bar{t}W^\pm$ events are generated with up to two extra partons at matrix element level, while for $t\bar{t}Z$ up
 1277 to one extra parton at matrix-element level is produced. The tZ process is simulated without extra
 1278 partons. The next-to-leading-order (NLO) cross sections are implemented by applying a uniform
 1279 k -factor to the leading-order (LO) events for each process. For $t\bar{t}Z$, the k -factor is determined by
 1280 comparing LO and NLO cross sections for on-shell Z production only and then applied to the off-shell
 1281 signal regions.

1282 The $t\bar{t}V$ uncertainties are calculated using the internal QCD scale and PDF re-weighting that is
 1283 available with MadGraph5+aMC@NLO. The prescription for the scale envelope is taken from [76]:
 1284 the central value $\mu = \mu_R = \mu_F = m_t + m_V/2$ and the uncertainty envelope is $[\mu_0/2, 2\mu_0]$. The
 1285 PDF uncertainty prescription used is the recipe from [77]: calculate the PDF uncertainty using the
 1286 MSTW2008nlo [75] PDF for the central value and then the final PDF uncertainty envelope is derived
 1287 from three PDF error sets each with different α_S values (the central value and the upper and lower
 1288 90% CL values). The final NLO cross section central values and uncertainties are given in Table 8.2.

1289 The tZ process is normalized to NLO based on the calculation in Ref. [78]. Here the scales are set
 1290 to $\mu_0 = m_t$ and the scale variations are by a factor of four; the scale dependence is found to be quite
 1291 small.

1292 8.1.1 $t\bar{t}Z$ Validation Region

1293 Unlike $t\bar{t}W^\pm$, a $t\bar{t}Z$ validation region can be obtained by simply inverting the veto on same-flavor
 1294 opposite-sign (SFOS) lepton pairs near the Z pole in the 3 lepton signal region. This region thus
 1295 requires 3 leptons (with momentum and identification cuts discussed in Chapter 7, at least one
 1296 opposite sign, same-flavor pair of leptons within 10 GeV of the Z mass, and either 4 jets and at least 1
 1297 b-tagged jet or exactly 3 jet and 2 or more b-tagged jets. The resulting region has low statistics and is

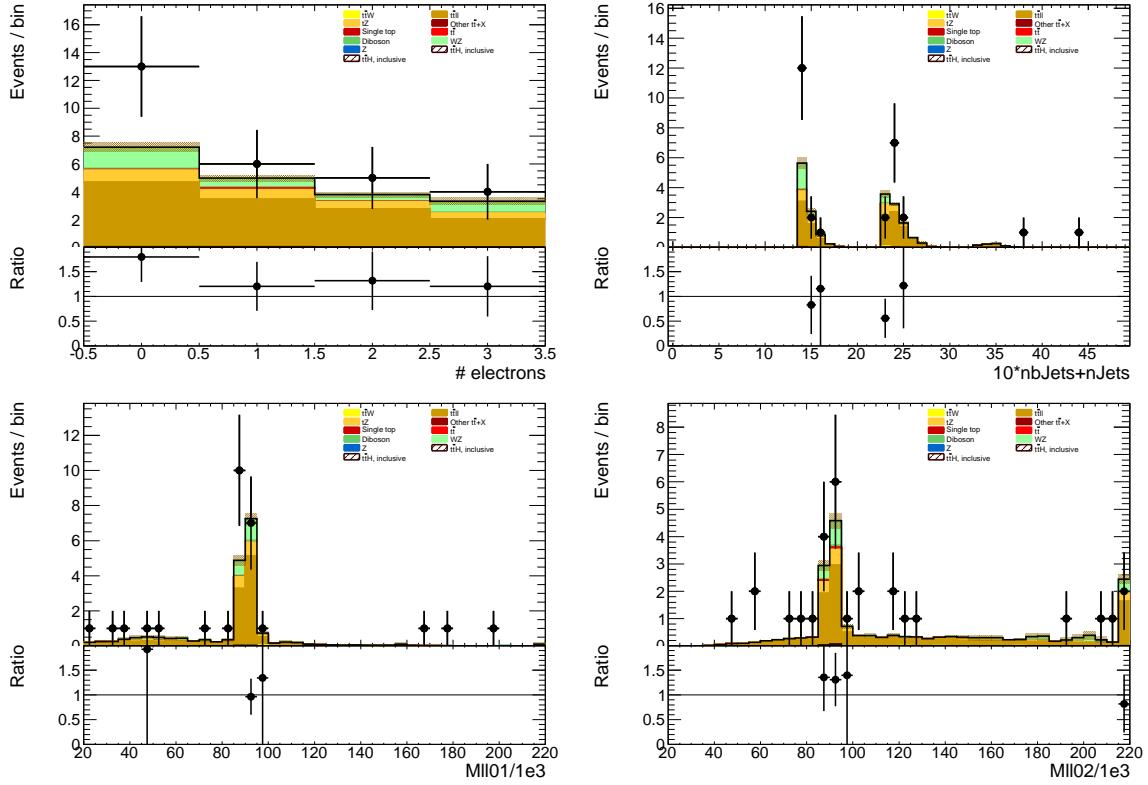


Figure 8.1: Data/MC comparison plots for $t\bar{t}Z$ control region A (≥ 4 jets, ≥ 1 b -tag and 3 jets, ≥ 2 b -tag). In all plots, the rightmost bin contains any overflows. Top left: number of electrons. Top right: 10^* the number of b -tags + the total number of jets. Middle left: the invariant mass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

1298 not used as a control region but is instead used as a validation to demonstrate that the normalization
 1299 uncertainty, discussed above, is properly evaluated.

1300 The region defined by this is predicted to be 67% $t\bar{t}Z$, 17% $W^\pm Z$, and 13% tZ . We predict
 1301 19.3 ± 0.5 events and observe 28, giving a observed-to-predicted ratio of $1.45 \pm 0.27 \pm 0.03$, where the
 1302 errors are from data and simulation statistics, respectively. Given the large errors, the region is still
 1303 in agreement with the predictions to within 1-1.5 σ . Distributions of various variables are shown in
 1304 Fig. 8.1.

1305 8.2 Di-boson Background Estimation: $W^\pm Z, ZZ$

1306 $W^\pm Z$ and ZZ di-boson production with additional and b-tagged jets constitute small contributions
 1307 to the 3ℓ and 4ℓ channels. For the 3ℓ case $W^\pm Z$ comprises $\sim 10\%$ of the total background, while for
 1308 the 4ℓ case ZZ contribution accounts comprises $\sim 10\%$ of the total background. Because of the small
 1309 size of these contributions, each of the above processes can be assigned a non-aggressive uncertainty
 1310 based on similar previous analyses with ATLAS and cross-checked with data validation regions and
 1311 MC truth studies.

1312 Both $W^\pm Z$ and ZZ production have been studied by ATLAS [79][80], but neither process has
 1313 been investigated thoroughly in association with multiple jets and b-quark jets. However, single
 1314 boson production with b-quark jets has been investiaged. Both $W + b$ [81] and $Z + b$ [82] production
 1315 in 7 TeV data have been shown to agree with MC models to within 20-30%.

1316 A single W produced in association with b-tagged jets possesses a similar topology to the $W^\pm Z + b$
 1317 process at a different energy scale and has been shown to be dominated by c mis-tags and b-jets from
 1318 gluon splitting and multiple parton interaction. The $W + b$ analysis, referenced above, uses Alpgen
 1319 MC with Herwig PS modeling, only provides results to 1 additional jet, and uses the CombNN tagger
 1320 (we use MV1). Its results are therefore not directly comparable to this $t\bar{t}H$ analysis (where $W^\pm Z$ is
 1321 modeled using Sherpa with massive c and b quarks). $Z + b$ production originates from slightly different
 1322 diagrams than $ZZ + b$, but the sources of the b-tags are similar. The 7 TeV analysis, referenced above,
 1323 provides results with Sherpa MC with an agreement of $\sim 30\%$. However, it also used the CombNN
 1324 tagger instead of MV1. Beause of the differences of the 2011 single boson analyses (type of tagger
 1325 used, type of MC and tunes used), we would like to verify the general 20-30% level of agreement in
 1326 2012 data with the simulation and tagger used in the $t\bar{t}H$ analysis: Sherpa MC, 2012 tunes, MV1.
 1327 With the data skims available to use we are able to do this in the $Z + b$ region but not the $W + b$.

1328 Figure 8.2 shows the spectrum of the number of reconstructed and selected jets (NJet) in a $Z + b$
 1329 validation region, defined by 2 tight-isolated leptons within 10 GeV of the Z mass and with at least
 1330 one b-tagged jet, using the $t\bar{t}H$ analysis definitions. The level of agreement in this region confirms at
 1331 the 30% level seen in the 7 TeV analysis, discussed above.

1332 In the following two sections, we assess the truth origin of jets in the $W^\pm Z + b$ and $ZZ + b$ regions
 1333 and leverage data/MC agreement where we can. We see that the data allows us to constrain the
 1334 $W^\pm Z$ to 50%. We claim this 50% as a systematic. The 20-30% agreement in the single boson regions
 1335 above bolsters our confidence in this number.

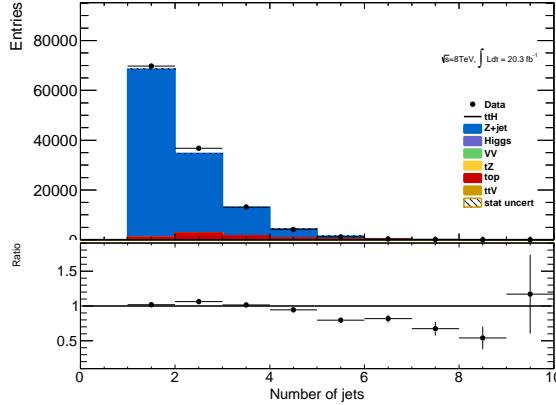


Figure 8.2: NJet spectrum for 2 tight-isolation leptons with 1 b-tagged jet (MV1_70)

8.2.1 $W^\pm Z$ Normalization 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 and a requirement that at least one SFOS pair have an invariant mass within 10 GeV of the Z mass. The region with fewer than 4 jets is $W^\pm Z$ dominated. Figure 8.3 shows kinematic variables for the inclusive region. The overall data normalization is $\sim 10\%$ higher than MC, but this will be well within our systematic uncertainty. The NJet shape shows good agreement across the full spectrum, giving confidence about the Sherpa high NJet SR extrapolation. Figure 8.4 shows NJet spectrum for the $W^\pm Z + b$ validation region with agreement with in statistical uncertainties. The region has low statistics and around $\sim 60\%$ purity and statistical analysis of the region suggests that a 50% normalization error on the $W^\pm Z$ component is enough to cover any possible mismodelings, especially in higher NJet bins, which are closer to the signal regions.

We also examine the $W^\pm Z$ truth origins of the b-jet in the $W^\pm Z + b$ validation region (VR) and the signal region using MC to assess the validity of the extrapolation from the VR to the SR and to confirm the similarity in jet origin to the single boson analyses, references above. The flavour of the closest matching truth particle ($p_T > 5$ GeV, after FSR) in ΔR determines the true-jet flavor. If there are no quarks, taus or gluons within ΔR of 0.3, the label defaults to light. Table 8.3 shows the origin fraction of b-tagged jets in the various $W^\pm Z + b$ VRs and the SR. If there are two b-tagged jets, the highest p_T is used, but this is a small fraction of the number of b-tags. As expected the c and b contributions dominate, as was the case with the 2011 single boson analyses referenced above. It is important also that the VR has similar composition to the SR. There is a small dependence on

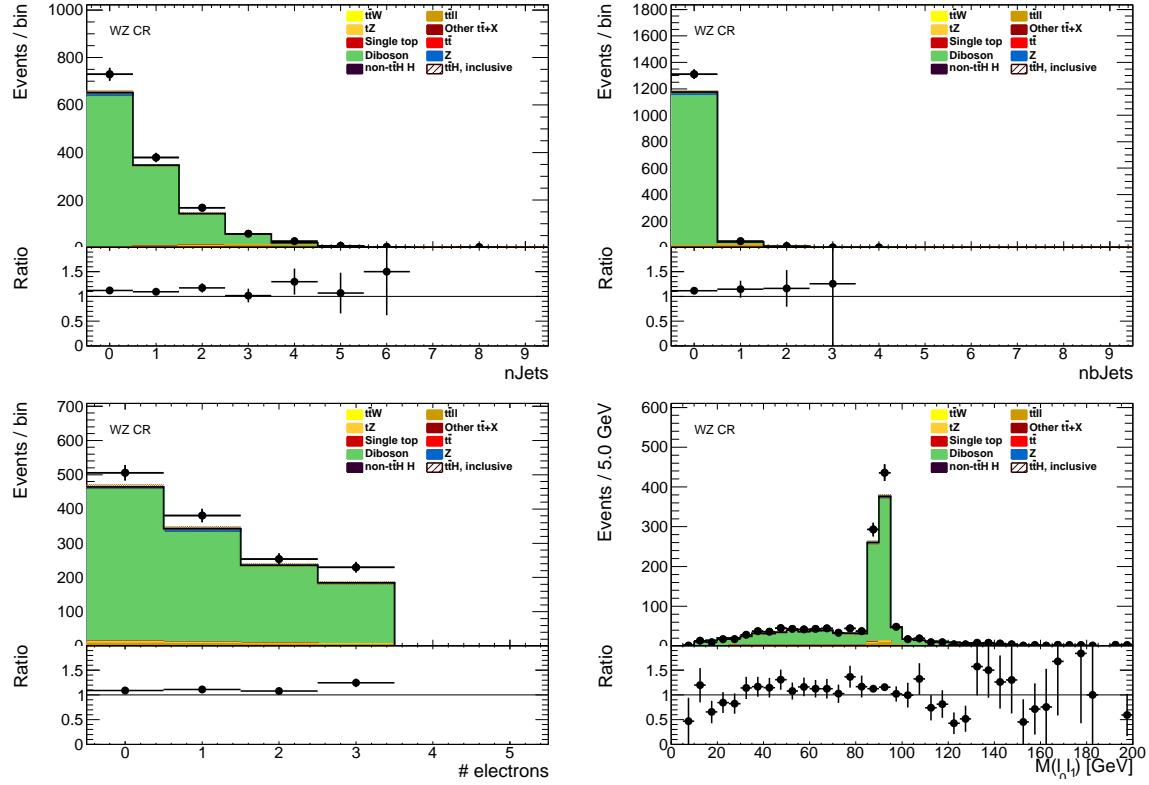


Figure 8.3: 3 lepton $W^\pm Z$ validation using the $t\bar{t}H$ lepton identification and momentum cuts: mass, number of jet and flavor variables

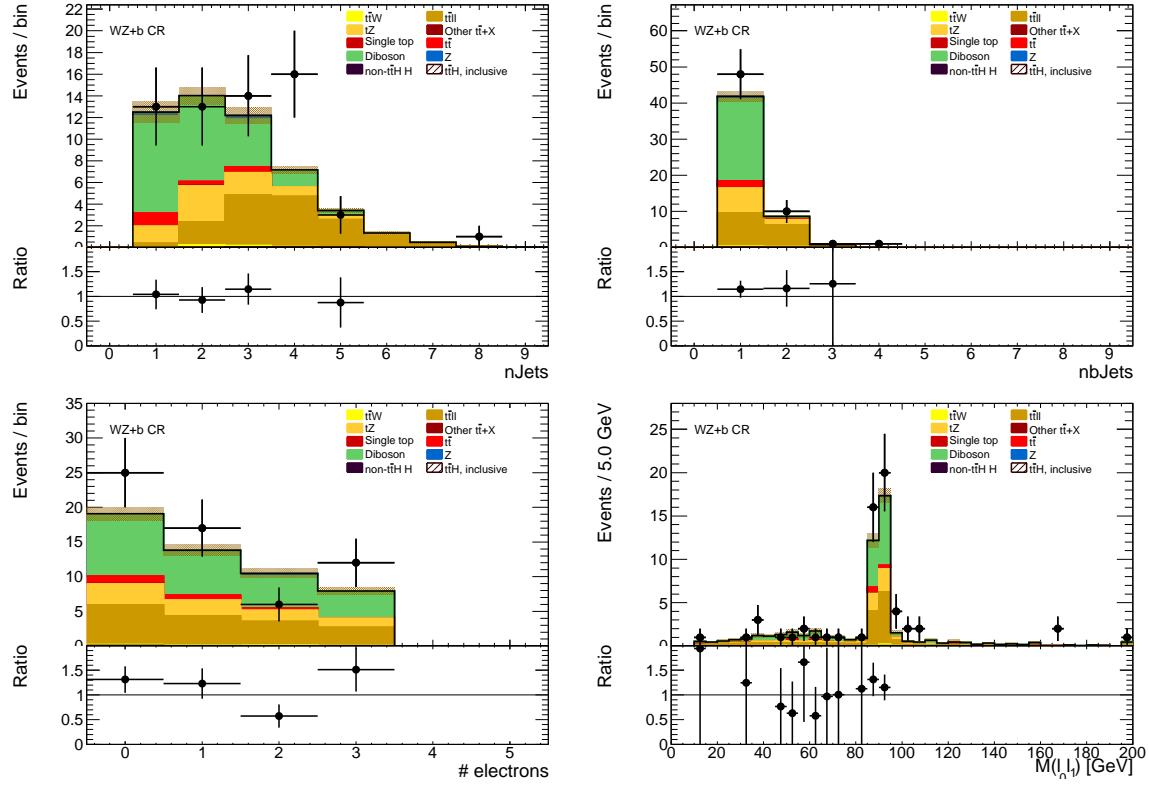
1357 the number of jets.

	Bottom	Charm	Light
$W^\pm Z + b$ VR 1 Jet	0.25 ± 0.03	0.54 ± 0.04	0.20 ± 0.03
$W^\pm Z + b$ VR 2 Jet	0.34 ± 0.04	0.52 ± 0.06	0.13 ± 0.03
$W^\pm Z + b$ VR 3 Jet	0.40 ± 0.07	0.41 ± 0.07	0.18 ± 0.04
$3l$ SR	0.43 ± 0.14	0.38 ± 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

1358 8.2.2 ZZ Normalization Uncertainty

In order to investigate the MC agreement with data in the ZZ case, two validation regions similar to the $W^\pm Z$ case are defined. First, a 4 lepton ZZ region is constructed using the object selections for the 4-lepton channel and requiring exactly two pairs of opposite sign-same flavor leptons with a di-lepton invariant mass within 10 GeV of the Z mass. Additionally, the $ZZ + b$ process is investigated directly using a similar validation region which again requires exactly two Z -candidate lepton pairs

Figure 8.4: $W^\pm Z + b$ validation region: NJet, NElec, and Mass Variables

as well as at least 1 b-tagged jet. Some kinematic distributions are shown in Figures 8.5 and 8.6, and particular attention should be paid to the NJet spectrum, which shows good data-MC agreement in the high-jet bins, with a slight discrepancy in the 1-jet bin. The agreement for the region with at least 2 jets yields confidence in the NJet MC modeling in this region which lies close to the 4-lepton signal region.

Based on the study of the ZZ and $ZZ + b$ validation regions and the overall agreement noted with the $Z + b$ analysis, we expect a similar error to $W^\pm Z$ to be appropriate in the ZZ case. A truth origin study is undertaken in MC to demonstrate a similar b-jet origin to the $W^\pm Z$ case. The true origin of the leading (highest energy) b-tagged jet is shown in Table 8.4 for the 4-lepton signal region as well as the $ZZ + b$ validation region described above divided into jet bins. As it was in the $W^\pm Z$ case above, the true origin of the b-jet in $ZZ + b$ is dominated by c and b. Taking this study in tandem with the results from the $W^\pm Z$ investigation, it is appropriate to take the central value of the $ZZ + b$ background contribution in the 4-lepton SR from MC and to assign an overall systematic of 50% in order to account for the MC modeling limitations.

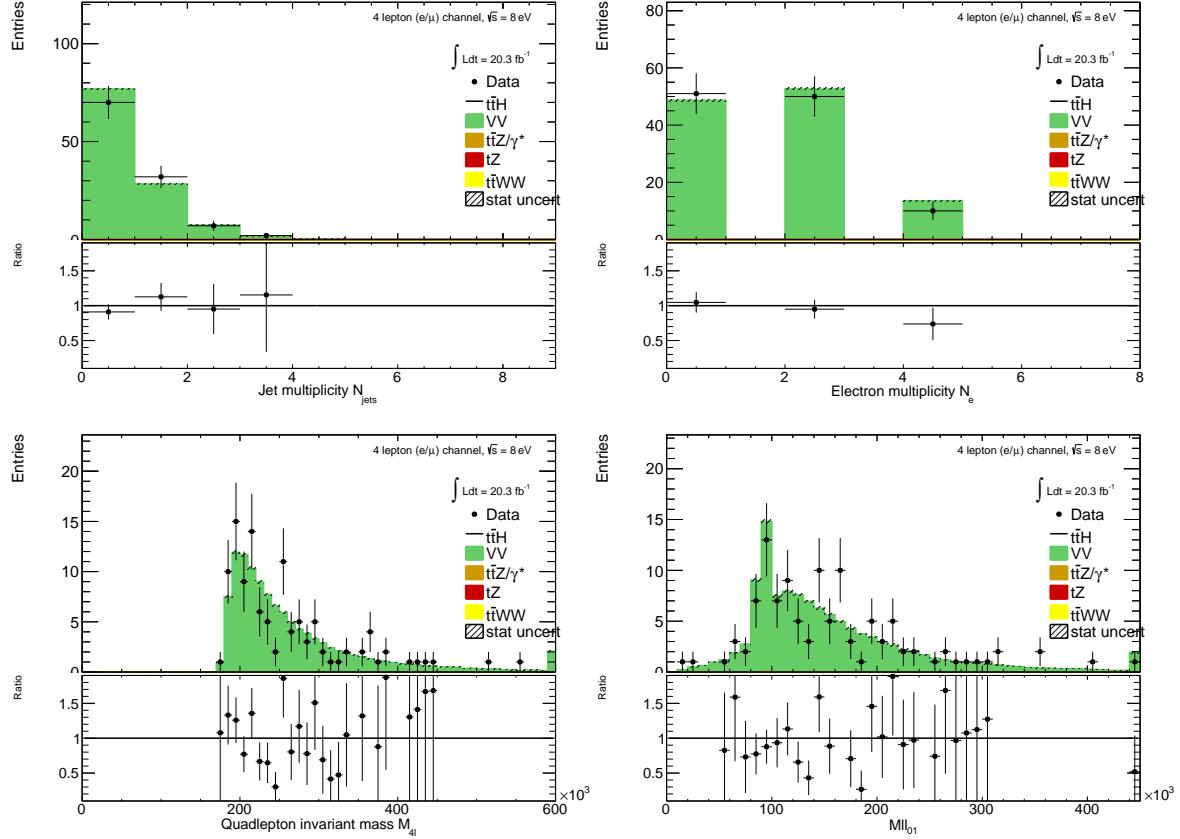


Figure 8.5: Jet-inclusive 4-lepton ZZ validation region using the $t\bar{t}H$ lepton identification and momentum cuts

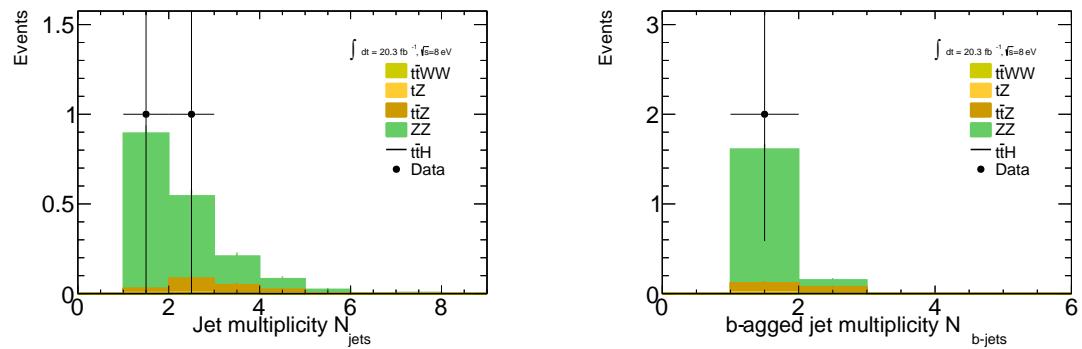


Figure 8.6: $ZZ + b$ validation region using the $t\bar{t}H$ lepton identification and momentum cuts

	Bottom	Charm	Light
$ZZ + b$ VR 1 Jet	0.56 ± 0.03	0.24 ± 0.01	0.20 ± 0.01
$ZZ + b$ VR 2 Jet	0.52 ± 0.05	0.25 ± 0.02	0.23 ± 0.02
$ZZ + b$ VR 3 Jet	0.53 ± 0.11	0.25 ± 0.08	0.22 ± 0.07
$4l$ SR	0.34 ± 0.15	0.42 ± 0.16	0.24 ± 0.10

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

1378 8.3 Charge-Misidentification Background

1379 Charge-misidentification contributes to the background for 2ℓ SS case and only for flavor channels,
 1380 which include electrons. The same-sign requirement is essential in removing large SM opposite sign-
 1381 backgrounds, but because of their size even small charge misidentification rates result in contamination
 1382 in same-sign regions. For the 2ℓ SS signal regions, charge-misidentification background arise primarily
 1383 from $t\bar{t}$ di-lepton events with a smaller contribution from leptonic Z decays.

1384 In general, charge-misidentification can arise in two ways. The first occurs for ultra-high energy
 1385 electrons and muons, which leave tracks in the detector that are too straight for the fit to determine
 1386 the direction of curvature with high confidence. This type of charge misidentification is not a concern
 1387 to the $t\bar{t}H$ multi-lepton analysis, as most of the leptons have transverse momentum < 150 GeV. The
 1388 second source of charge misidentification is from 'tridents', which only occurs for electrons, because
 1389 their low mass allows for high rate bremsstrahlung in the detector material. In some cases, after
 1390 an electron releases a photon through bremsstrahlung, the photon may convert nearby resulting
 1391 in three electron tracks. The reconstruction algorithms may sometimes match the wrong track to
 1392 the calorimeter energy deposit, resulting in a possible charge misidentification. As discussed in the
 1393 selection, tight track-cluster geometric and energy matching requirements are applied on the electron
 1394 candidates to reduce the overall rate and the electron acceptance is narrowed to ($|\eta| < 1.37$), since
 1395 most of the material is concentrated more forward in the detector.

1396 We estimate the contribution of charge-misidentification events in our 2ℓ SS signal regions and
 1397 relevant control regions by applying a weight per electron in the OS region with otherwise identical
 1398 cuts. The weight is related to the charge-misidentification rates. We measure these rates using a
 1399 likelihood method in the OS and SS $Z \rightarrow ee$ control region in data. The rate measured from these
 1400 control regions is binned in electron p_T and η , to account for dependencies in these variables. The
 1401 method, validations and associated errors are discussed in detail in the following sub-sections.

8.3.1 Likelihood Method

The number of reconstructed same-sign (N_{ss}) $Z \rightarrow ee$ events is related to total number of produced $Z \rightarrow ee$ (N) through factors related to the charge misidentification rate, ϵ :

$$N_{ss}^{ij} = N^{ij}(\epsilon_i + \epsilon_j - 2\epsilon_i\epsilon_j) \quad (8.1)$$

where ϵ_i and ϵ_j are the charge misidentification rates for each electron separately. If we drop terms quadratic in ϵ , we have:

$$N_{ss}^{ij} = N^{ij}(\epsilon_i + \epsilon_j). \quad (8.2)$$

Although it is impossible to know event-by-event which electron's charge was misidentified, we can use a likelihood method over the whole Z sample to measure how ϵ depends on the electron p_T and $|\eta|$. As illustration, we first consider the case, where ϵ depends on only one variable, $|\eta|$, and then generalize to the two-dimensional case of $|\eta|$ vs p_T .

N_{ss}^{ij} is described by a Poisson distribution:

$$f(k, \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad (8.3)$$

where k is the observed number of occurrences of the event, i.e. $k = N_{ss}^{ij}$, and λ is the expected number, i.e. $\lambda = N^{ij}(\epsilon_i + \epsilon_j)$. Thus, the probability for an observed number of same-sign Z events given the sample size and charge misidentification rates is expressed by:

$$P(N_{ss}^{ij}|N^{ij}, \epsilon_i, \epsilon_j) = \frac{[N^{ij}(\epsilon_i + \epsilon_j)]^{N_{ss}^{ij}} e^{-N^{ij}(\epsilon_i + \epsilon_j)}}{N_{ss}^{ij}!}. \quad (8.4)$$

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

$$L(\epsilon|N_{ss}, N) = \prod_{i,j} \frac{[N^{ij}(\epsilon_i + \epsilon_j)]^{N_{ss}^{ij}} e^{-N^{ij}(\epsilon_i + \epsilon_j)}}{N_{ss}^{ij}!}. \quad (8.5)$$

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

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

The likelihood can be easily extended to depend on the charge misidentification rates as a function of two parameters. The probability to find a same-sign event given the rates for each electron is $(\epsilon_{i,k} + \epsilon_{j,l})$, where the two indices represent binned $|\eta|$ - and p_T -dependence. Thus, the Eq. 8.6

1422 transforms into

$$-\ln L(\epsilon | N_{ss}, N) \approx \sum_{i,j,k,l} \ln[N^{ij,kl}(\epsilon_{i,k} + \epsilon_{j,l})]N_{ss}^{ij,kl} - N^{ij,kl}(\epsilon_{i,k} + \epsilon_{j,l}). \quad (8.7)$$

1423 We use events selected within the Z peak using the $t\bar{t}H$ electron object cuts. The events are stored
 1424 in two matrices: one for the same-sign events $N_{ss}^{ij,kl}$, and the other one for all events $N^{ij,kl}$. Small
 1425 backgrounds need to be subtracted. The background subtraction is done using a simple side-band
 1426 method. This method consists in dividing the Z invariant mass in three regions, i.e. A , B and C ,
 1427 where B is the central region corresponding to the Z peak. The number of events is counted in the
 1428 regions on the sides of the peak, i.e. n_A and n_C , and removed from the total number of events in the
 1429 peak region B , n_B . This way, the number of signal events N_Z is given by

$$N_Z = n_B - \frac{n_A + n_C}{2}. \quad (8.8)$$

1430 Once the background has been subtracted, the likelihood is minimized for the 2D matrix of ϵ bins.
 1431 Knowing ϵ as a function of $|\eta|$ and p_T for any single electron, it is now possible to estimate the number
 1432 of same-sign events from the number of opposite sign events in any sample:

1433 • $N^{ss} = \frac{\epsilon_i + \epsilon_j - 2\epsilon_i\epsilon_j}{1 - \epsilon_i - \epsilon_j + 2\epsilon_i\epsilon_j} N^{os}$ for ee channels

1434 • $N^{ss} = \frac{\epsilon}{1 - \epsilon} N^{os}$ for the $e\mu$ channels

1435 8.3.2 Results

1436 The charge misidentification rate is calculated in 7 $|\eta|$ bins [0.0, 0.6, 1.1, 1.37, 1.52, 1.7, 2.3, 2.47] by
 1437 4 p_T bins [15, 60, 90, 130, 1000] GeV. For p_T bins above 130 GeV, the Z dataset becomes too small
 1438 and the rates are calculated using $t\bar{t}$ MC, scaled by the data-MC ratio of the rates in the lower p_T
 1439 bins, [90-130] GeV. Figure 8.7 shows the extracted rates in all bins.

1440 As a cross-check, we apply the full method to the Z MC samples (extracting rates via a likelihood
 1441 fit and applying them to opposite sign events) and compare to the MC predicted number of same-sign
 1442 events. The invariant mass of the Z from our charge misidentification and directly from the MC
 1443 can be seen on Figure 8.8. In the simulated Z samples, the number of same-sign Z events is 5 049
 1444 while the estimation is $5 031^{+375}_{-365}$. The uncertainties combine both statistical systematic uncertainties,
 1445 which are discussed in depth below. The validation gives compatible results within uncertainties.

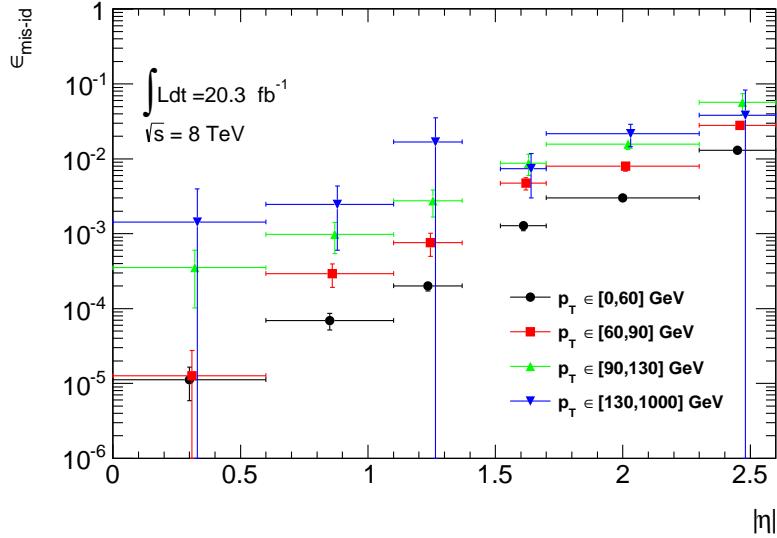


Figure 8.7: Electron charge misidentification rates measured in data with the likelihood method on Z events (black points, red squares and blue triangles) as a function of $|\eta|$ and parametrized in p_T . The full 2012 dataset has been used to estimate the rates below 130 GeV. Above this value, the charge misidentification rates have been estimated by extrapolating the rates in the region where the $p_T \in [90, 130]$ GeV with a p_T dependent factor extracted from simulated $t\bar{t}$ events (green triangles). Statistical and systematic uncertainties have been included in this plot.

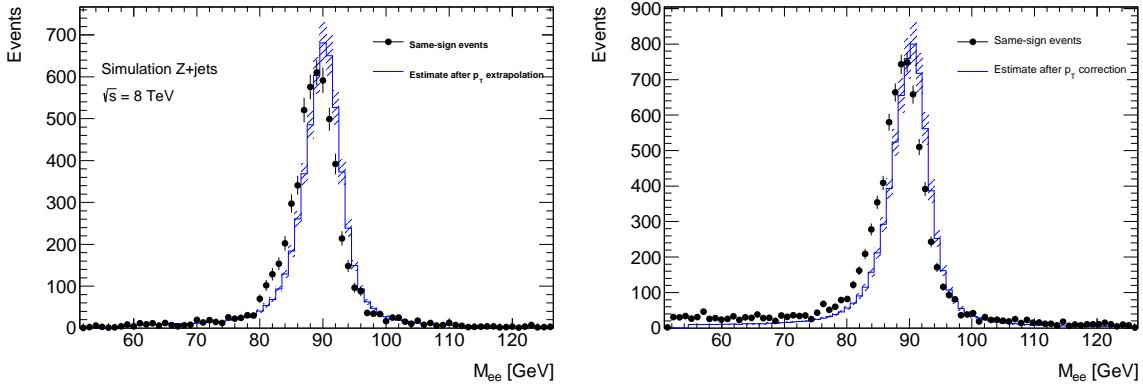


Figure 8.8: Closure test on simulated $Z \rightarrow e^+e^-$ events (a) and data (b). The black circles show the distribution of same-sign events while the blue histograms show the distribution of the re-weighted opposite-sign events together with the statistical and systematic uncertainties. The distributions are not expected to overlay exactly, due to the loss of energy of the trident electrons for the same-sign peak.

1446 8.3.3 Systematic and Statistical Uncertainties

1447 Statistical uncertainties dominate the combined uncertainty on the charge misidentification estimate.

1448 The statistical uncertainties come primarily from the size of the Z same-sign sample in data and are

especially large for central, material-poor regions where the charge misidentification rate is extremely low. Additional systematic uncertainties are included for a comparison between the positron and electron rate, the per-bin MC closure test discussed above, and for the effect of varying the invariant mass window used for the background subtraction for three different cases. Figure 8.9 shows the relative uncertainties for all rate bins.

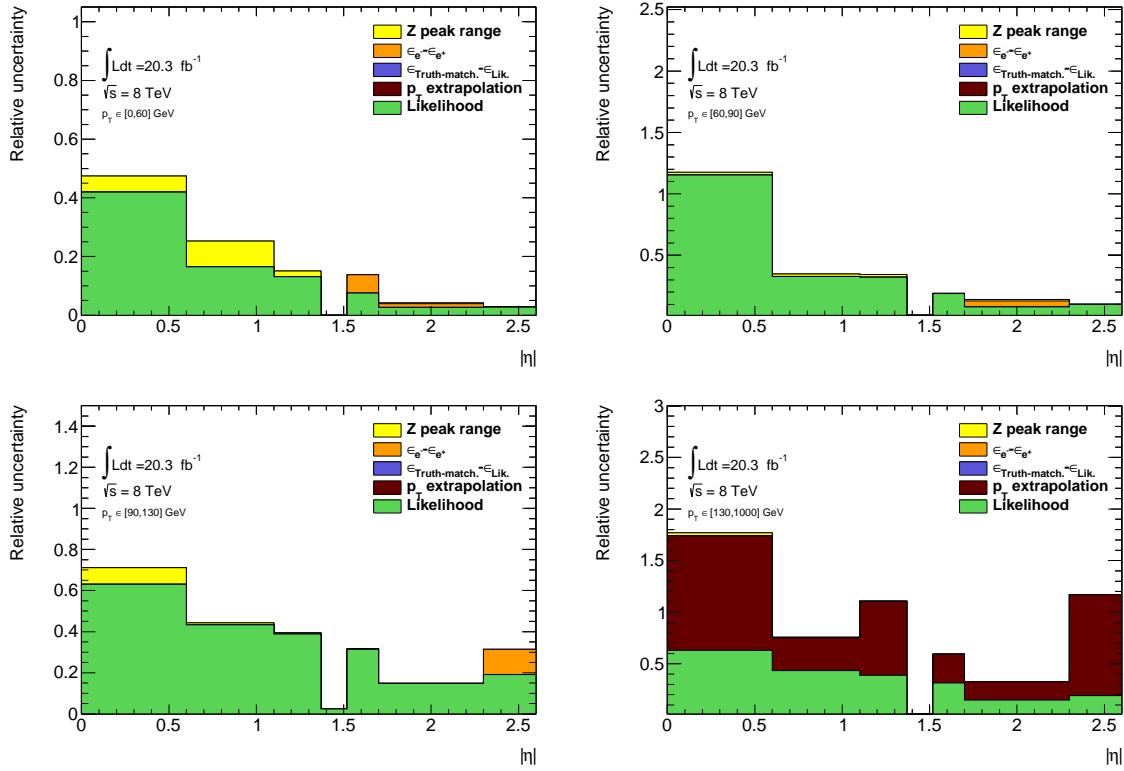


Figure 8.9: Relative systematic uncertainty contributions on the charge misidentification rate, for different bins in p_T and $|\eta|$.

We apply the rates to estimate the charge misidentification background in the 2ℓ SS signal regions, and find $\sim 25\%$ contamination in the $e^\pm e^\pm$ regions and a $\sim 10\%$ contribution to the $e^\pm \mu^\pm$ regions with a 10% systematic error overall. The low overall error can be attributed to the fact that the statistical error is lowest where the bulk of charge misidentifications occur. The charge flip contribution measured in the signal regions from this method is detailed in Table 8.1.

1459 8.4 Fake Lepton Backgrounds

1460 Fake Leptons, from the misidentification of jets as either electrons or muons, primarily arise from $t\bar{t}$
 1461 and single top processes in the 2ℓ SS, 3ℓ and 4ℓ channels. Smaller contributions come from $Z+jet$
 1462 events. Fake backgrounds are sub-dominant but important in the 2ℓ SS and 3ℓ channels. They
 1463 are extremely small in the 4ℓ channels. Truth studies suggest that these misidentified leptons arise
 1464 overwhelmingly from b-quark initiated jets. The general method for estimating fakes in all channels
 1465 is to define a reversed object selection control region (usually isolation) for each lepton flavor with
 1466 otherwise identical signal region selection (N_{CR}^e , N_{CR}^μ). This region is fake-dominated. The total
 1467 number of fake events in these regions are then scaled by transfer factors (θ) to estimate the number
 1468 of fake events of the appropriate flavor in the signal region. The simple formula for determining fakes
 1469 is defined in Equation 8.9.

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

1470 This approach factorizes the background model into two separate measurements. N_{CR} is sensitive
 1471 the overall $t\bar{t}$ production rate, especially in the presence of additional jets from QCD ratio, as well as
 1472 the object-level misidentification of a jet as a lepton. The transfer factors are sensitive to only the
 1473 object level properties of the misidentified jet, and in particular only the variables which are reversed
 1474 in the anti-tight identification.

1475 The transfer factors are obtained obtained in a different way for each channel, due to unique issues
 1476 with statistics and contamination, but each method relies heavily on the data-based control regions
 1477 with fewer jets. Figure 8.10 shows a truth study of the stability of the transfer factor for the 2ℓ SS
 1478 and 3ℓ cases as a function of the number of jets in the event for events with one-b-tagged jet. This
 1479 suggest that the regions with fewer jets are a good model of the fakes in the signal regions with more
 1480 jets and is expected because of the homogeneity of origin of the fakes across all jet bins.

1481 The details of the methods for each channel are discussed in depth in the following sections. For
 1482 all methods, the overall systematic uncertainty on the normalization of the fake estimate is in the
 1483 range 30%-50% and arise primarily from statistics and the closure on assumptions used to obtain the
 1484 transfer factor.

1485 Because these methods do not provide a per-object transfer-factor that depends on the properties
 1486 of the faking object, we must use the MC to model the shapes of the fake kinematic distributions
 1487 in the signal regions. This is not an essential issue, since the analysis only considers only the total
 1488 number of events in each signal region in the final measurement of $t\bar{t}H$ production.

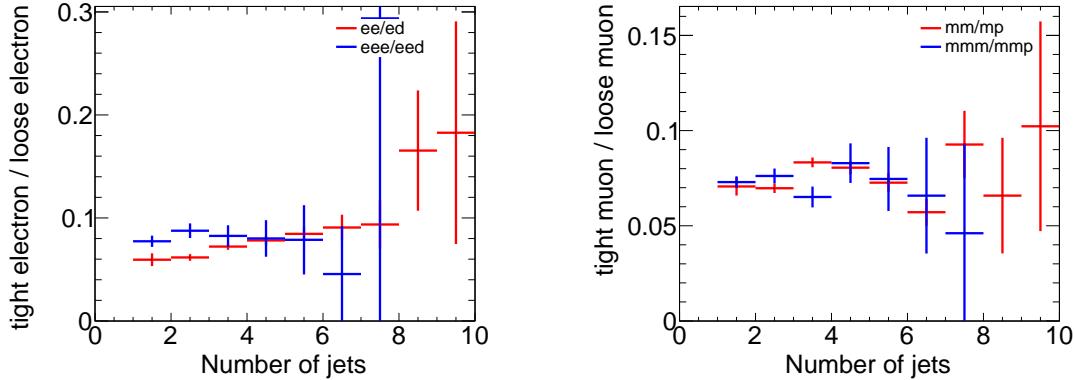


Figure 8.10: Ratios of regions with tight and anti-tight leptons in 2-lepton and 3-lepton channels from $t\bar{t}$ MC. These ratios are the MC calculated transfer factors for each region, i.e. $\theta_e = eee/eed$, ee/ed and $\theta_\mu = mmp/mmm$, mm/mp , where 'd' refers to anti-tight electrons and 'p' refers to anti-tight muons. The transfer factors are seen to be similar in the 2ℓ and 3ℓ cases and stable as a function of the number of jets

1489 8.4.1 2ℓ SS Fakes

1490 The 2ℓ SS fake method follows the procedure outlined in general above. We define anti-tight electron
 1491 and muon control regions with reversed particle identification criteria for each signal region, including
 1492 the 6 flavor and jet-counting sub regions. The anti-tight muon and electron criteria are provided
 1493 below:

- 1494 • **anti-tight electron (d):** fails the verytight likelihood operating point, but still passes the
 1495 veryloose operating point. fails relative tracking and calorimeter isolation, $E_T^{rel} > 0.05$ and
 1496 $p_T^{rel} > 0.05$.
- 1497 • **anti-tight muon (p):** $6 \text{ GeV} < p_T < 10 \text{ GeV}$

1498 The electron and muon transfer factors, θ_e and θ_μ , are calculated in the region with signal region
 1499 selection but fewer jets, $NJet == 2$ or $NJet == 3$ and are defined as the ratio of the number of
 1500 events for two fully identified leptons to the number of events with one fully identified lepton and
 1501 one anti-identified lepton, after the prompt and charge misidentification backgrounds are subtracted.
 1502 Only same-flavor channels are used to ensure that muon and electron transfer factors maybe estimated
 1503 separately: on every region, the prompt and charge-misidentification backgrounds are subtracted from
 1504 the data.

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

Process	N(events)
<i>ed</i> \leq 3 jets	
<i>VV</i>	7.13 ± 0.63
<i>V</i> γ	7.55 ± 1.27
<i>t</i> $\bar{t}V, tV$	6.68 ± 0.18
<i>V + jets</i>	59.4 ± 18.51
<i>t</i> $\bar{t}, t + X$	671.26 ± 12.76
<i>t</i> $\bar{t} prompts$	32.97 ± 2.83
Total MC	752.0 ± 22.5
Data	967
Data fakes (Data - prompts)	912.66
<i>ee</i> \leq 3 jets	
<i>VV</i>	3.30 ± 0.42
<i>V</i> γ	1.31 ± 0.65
<i>t</i> $\bar{t}V, tV$	3.96 ± 0.16
<i>V + jets</i>	8.3 ± 8.8
<i>t</i> $\bar{t}, t + X$	11.65 ± 1.67
Charge misID	8.54 ± 0.23
Total MC	28.52 ± 8.96
Data	32
Data fakes (Data - prompts)	14.26
<i>$\mu\mu$</i> \leq 3 jets	
<i>VV</i>	6.27 ± 0.56
<i>V</i> γ	0.06 ± 0.25
<i>t</i> $\bar{t}V, tV$	10.00 ± 0.27
<i>V + jets</i>	1.22 ± 11.78
<i>t</i> $\bar{t}, t + X$	15.4 ± 2.1
Total MC	32.95 ± 2.21
Data	44
Data fakes (Data - prompts)	27.65

Table 8.5: Number of events of the main simulated background processes and of the data in the $e^\pm e^\pm$ and $\mu^\pm \mu^\pm$ channels used for the measurement of θ_e and θ_μ . VV , $V\gamma$, $t\bar{t}V, tV$ and $t\bar{t}$ prompts (or charge misID) are the backgrounds which lead to prompt same-sign dileptons and are subtracted from the data to get a measured number of fakes. Uncertainties are statistical. The numbers labeled Data fakes are used to measure θ .

1505

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

1506

1507 The 2,3 jet anti-tight regions used in obtaining the transfer factors are shown in Table 8.5 and the
 1508 4,5 jet anti-tight regions, scaled by the transfer factors to get the fake estimates in the SR are shown
 1509 in Figure 8.11. The $t\bar{t}$ and single top MC are included in the plots and tables for reference, although
 1510 they are not used in the measurements.

1511 The transfer factors obtained from the 2 and 3 jet regions are shown in 8.6 with statistical errors
 1512 and propagated systematic errors from the subtraction of relevant backgrounds ($t\bar{t}V$ and charge
 1513 misidentification). The MC values are just for comparison. An additional systematic error is added
 1514 by comparing the transfer factors, obtained from the low jet control region, to those obtained from
 1515 the higher jet signal regions, using $t\bar{t}$ MC. The value of this systematic is about 20 % and can be seen

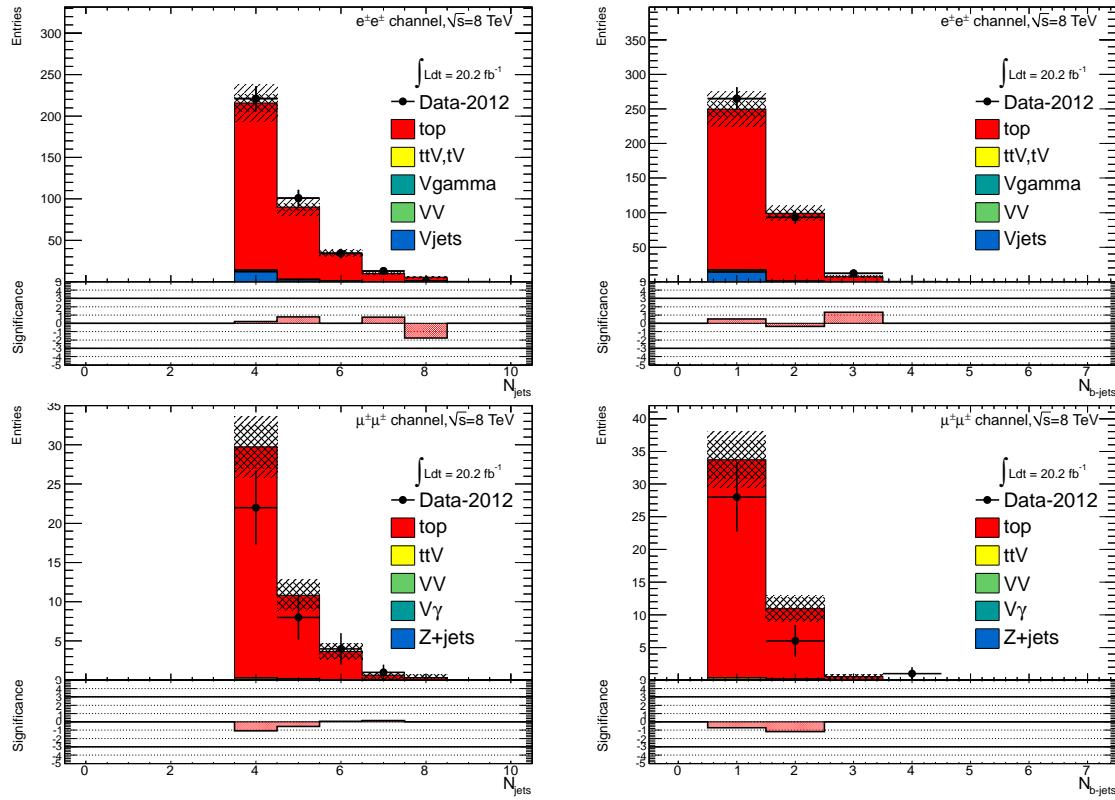


Figure 8.11: 4,5 Jet SS 2ℓ ed (above) and μp control regions with at least one b-tagged jet. After subtraction of prompt and charge misidentification backgrounds, these regions are scaled by the transfer factors, θ_e and θ_μ , to obtain the final number of fake events in the CR. The top MC (red) is used for reference but not 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 8.6: Expected and measured values of the θ factors. NEEDS TO BE UPDATED FOR 2,3

in the above Figure 8.10. The overall systematic uncertainties and contribution from each source in all of the sub-channels of the signal region are shown in Table 8.6 and the final contribution of fake events to the signal region are show in Table 8.1 found at the beginning of the chapter.

8.4.2 3ℓ Fakes

The 3ℓ fake method follows the same general strategy as the 2ℓ SS case. Transfer factors are used to extrapolate from anti-tight, fake-rich control regions in data into the signal region. However, the equivalent low jet control regions are too low in statistics to provide the transfer factors from data

Uncertainties		Channels			
		4 jets		≥ 5 jets	
		$e^\pm e^\pm$	$\mu^\pm \mu^\pm$	$e^\pm e^\pm$	$\mu^\pm \mu^\pm$
Statistical	$\Delta\theta_e^{\text{stat}}$	39.6	—	14.2	39.6
	$\Delta\theta_\mu^{\text{stat}}$	—	24.7	15.8	24.7
	$\Delta N_{\ell\text{anti}-\ell}(n \text{ jets})(\text{stat})$	6.8	18.1	21.3	25.9
Systematics	$\Delta\theta_e^{\text{syst}}$ (closure)	21.8	—	7.8	26.7
	$\Delta\theta_\mu^{\text{syst}}$ (closure)	—	23.3	18.4	31.2
	Q Mis Id ($\ell\text{anti}-\ell$)	2.2	—	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 8.7: 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)

1523 directly. Instead, the transfer factors are obtained directly from the $t\bar{t}$ simulation. Data control
 1524 regions, called auxiliary regions, are used to determine the modeling of the identification and isolation
 1525 variables, used in the transfer factor extrapolation. The low jet regions are still employed in a low
 1526 statistics validation of the entire fake procedure.

1527 Anti-tight electrons and muons are defined in slightly different ways, compared to the 2ℓ SS case:

- 1528 • **anti-tight electron (d):** fails to pass the verytight likelihood operating point, but still verifies
 1529 the veryloose operating point. The isolation selection is released $E_T^{\text{rel}} > 0.05$, $p_T^{\text{rel}} > 0.05$.
- 1530 • **anti-tight muon (p):** muons must pass identification but the p_T cuts is lowered to 6 GeV.
 1531 The overlap removal with jets and isolation cuts are released.

1532 The transfer factors, θ_e and θ_μ , extracted from MC, is defined as the ratio of the number of top ($t\bar{t}$
 1533 + single-top) events in the signal region, to the number of $t\bar{t}$ events in the anti-tight regions. The
 1534 factors are calculated in separate flavor regions to ensure that the electron jet fakes and muon jet
 1535 fakes are calculated separately. The calculation follows the same form as for the 2ℓ SS case, but now
 1536 lep0, which by construction is almost never a fake is allowed to be either electron or muon in both
 1537 cases, denoted below in Equations 8.12 and 8.13.

$$\theta_e = \frac{N_{xee}^{\text{top}}}{N_{xed}^{\text{top}}} \quad (8.12)$$

$$\theta_\mu = \frac{N_{x\mu\mu}^{\text{top}}}{N_{x\mu p}^{\text{top}}} \quad (8.13)$$

1538 The MC modeling of the variables involved in the transfer factor can be verified when another
 1539 variable fails. For instance, the MC modeling of the electron isolation variable can be compared to
 1540 data when the particle identification variable fails and vice-versa. The modeling of muon-jet ΔR ,
 1541 involved in the overlap removal, can be compared when either the isolation variable or the p_T fails.
 1542 The comparison of the electron variables in this manner can be seen in Figure 8.13 and the muon
 1543 variables in Figure 8.12. The regions used have the same selection as the signal region with an added
 1544 missing transverse energy requirement, > 60 GeV, to ensure only top fakes. 20% and 30% systematic
 1545 uncertainties are assigned to the muon and electron transfer factors, respectively, to account for data-
 1546 MC discrepancies. This method for evaluating data-MC agreement for individual electron and muon
 1547 variables in turn relies on the assumption that these variables are largely uncorrelated and that the
 1548 transfer factor itself is factorizable into pieces for each variable. Factorized and fully correlated transfer
 1549 factors have been compared using MC and shown to have differences smaller than the systematic
 1550 quoted, suggesting that the uncorrelated assumption is reasonable.

1551 The electron and muon anti-tight control regions, which are scaled by the transfer factors are shown
 1552 in Figure 8.14. The prompt MC subtracted data in these regions are scaled by the transfer factors to
 1553 obtain the overall contribution of fake electron and muon events in the signal region. The systematic
 1554 uncertainties are split between the statistical error on the transfer factor and normalization of the
 1555 anti-tight control regions and the data-MC comparisons outlined above. For muon fakes the total
 1556 systematic uncertainty is 25% and for electrons it is 34%. The numbers and uncertainties involved in
 1557 the calculation are shown in Table 8.8.

Stage	Muon	Electron
Anti-Tight CR Normalization	364.62 ± 20.02 (5%)	38.2 ± 6.9 (17%)
Transfer factor	0.0047 ± 0.0011 (23%)	0.0240 ± 0.0064 (36%)
SR Contribution	1.71 ± 0.42 (25%)	0.91 ± 0.35 (39%)

Table 8.8: Summary of regions and inputs to the extraction of the number of $t\bar{t}$ events with a fake muon in the SR

1558 The low jet region (1, 2, 3) is used as a validation for the method. The $t\bar{t}$ and single top fakes in
 1559 this region are estimated using the procedure above. Similar systematics are assessed. This region
 1560 with the fake estimate is plotted in Figure 8.15. The agreement of data and summed prediction for
 1561 the fakes and prompt backgrounds is well within the systematic and statistical uncertainties. The
 1562 figure also shows the same region with relaxed p_T cuts on all leptons to 10 GeV, which enriches the
 1563 fake contributions greatly. The data and summed fake and prompt predictions are also well within
 1564 the statistical and systematic uncertainties.

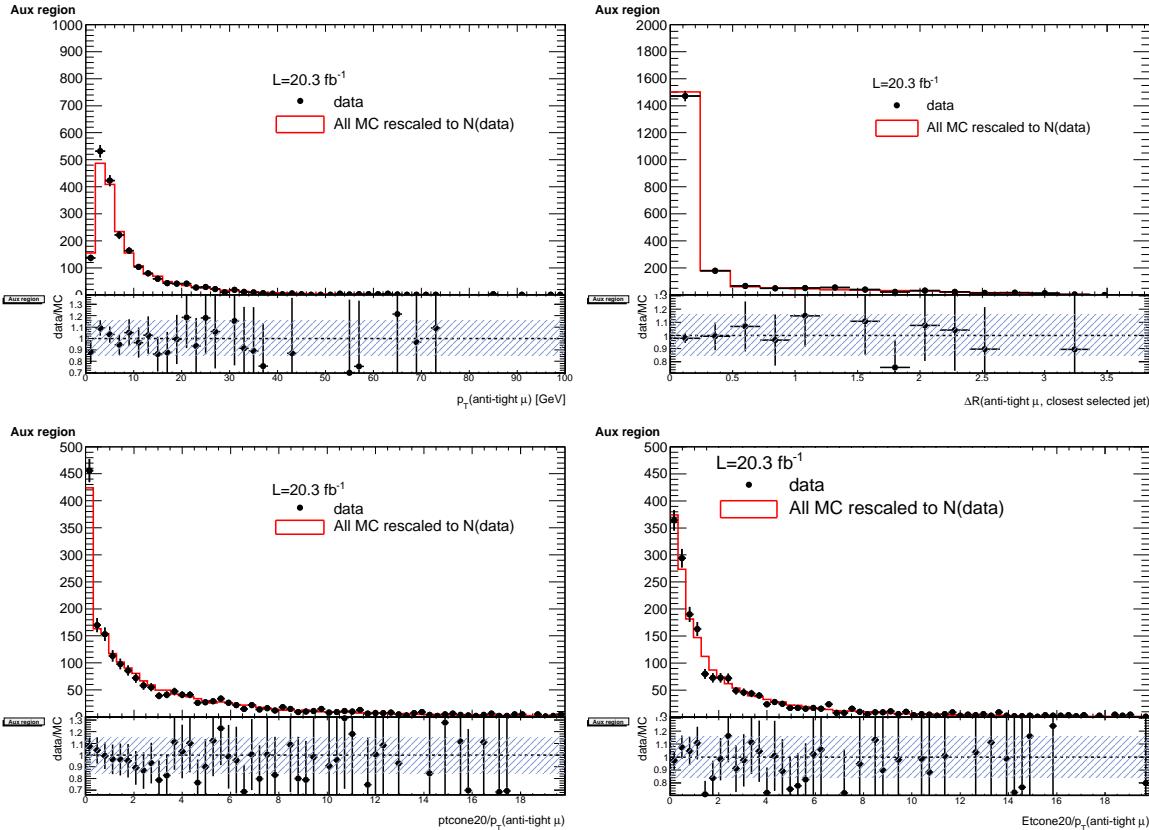


Figure 8.12: Distributions of the properties of the anti-tight muons in data (dots), compared with the total simulation (red line), rescaled to the integral of the data for a shape comparison. The uncertainty on the data distribution is statistical. The number of events for each of them is also presented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

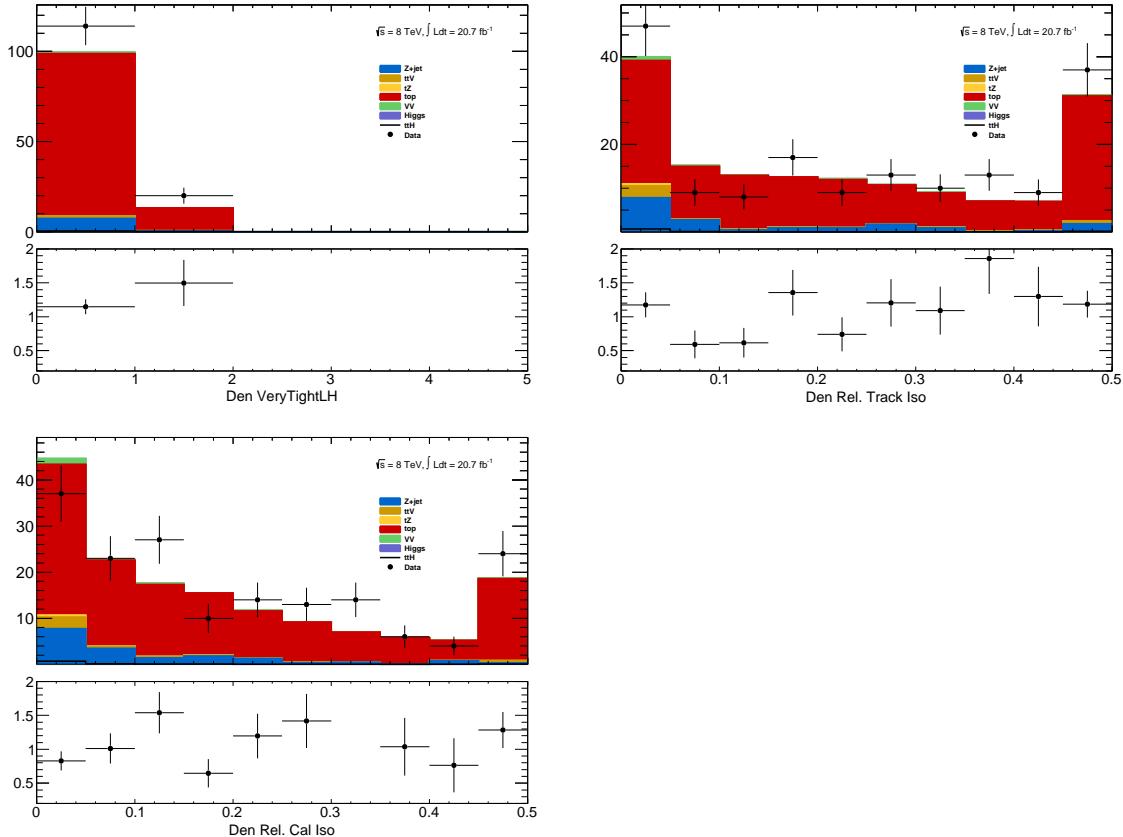


Figure 8.13: Distributions of anti-tight electron variables. The variables presented are, from top left to bottom right, p_T , η , VERYTIGHT Likelihood value, $ptcone20/p_T$, $Etcone20/p_T$. The plotted regions have the same cuts as the signal region, except the anti-tight electron must fail isolation 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 $t\bar{t}V$ (yellow). The uncertainty on the data distribution is statistical.

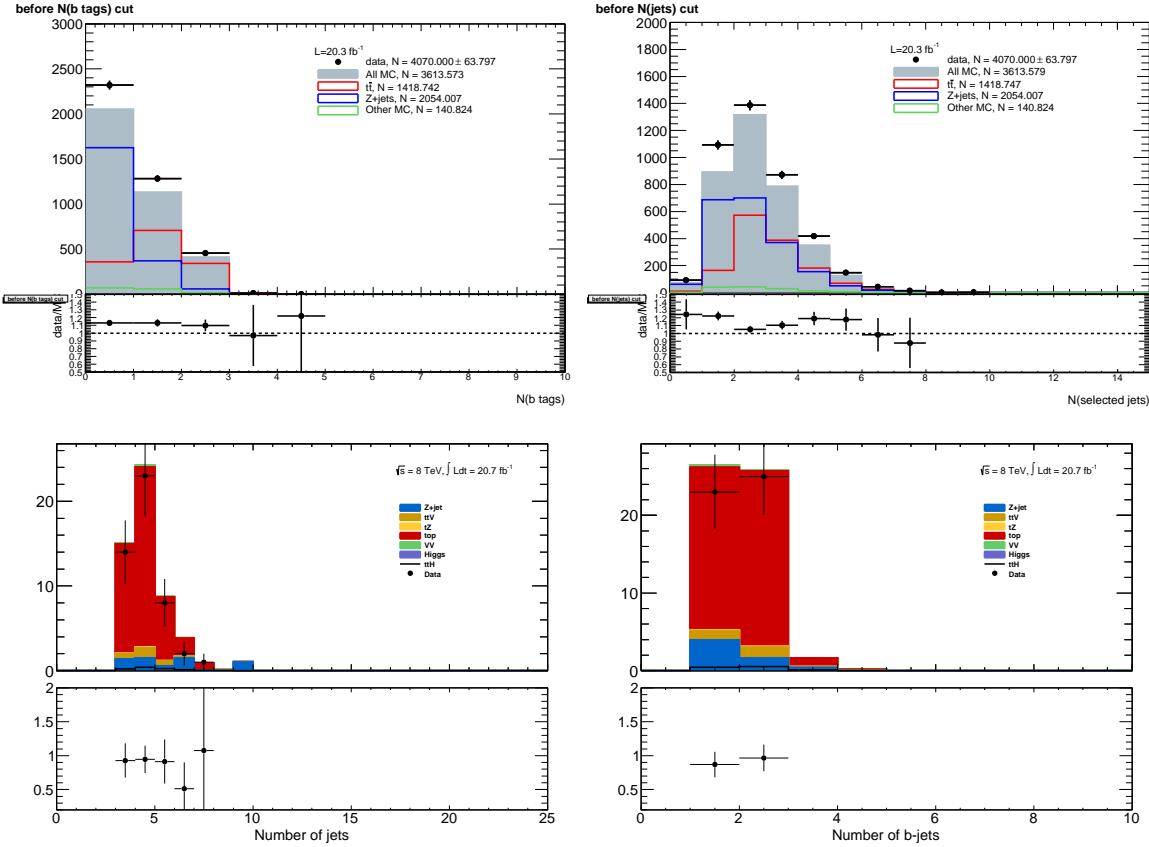


Figure 8.14: Muon-xxp (above) and electron-xxd (below) anti-tight control regions: jet variables. Note: the $t\bar{t}$ and single top MC in the plots is used only as comparison, but is not included in the fake measurement

1565 8.4.3 4ℓ Fakes

1566 We will not discuss the 4ℓ fakes in depth, as it is a very small background - at the % level and will have
 1567 almost no impact on the final result. The fake method used in the the 4ℓ case is similar to the 2ℓ and
 1568 3ℓ cases discussed above. All fakes arise from $t\bar{t}$ and single-top events, where *two* jets are misidentified
 1569 as leptons. To measure the contribution of this background, control regions with 2 fully identified and
 1570 2 anti-identified leptons are created. These control regions do not have a number of jets requirement
 1571 in order to increase statistics. From these control regions, two extrapolations are made. First, a
 1572 transfer factor is applied to extrapolate from the anti-tight to tight regions for electrons and muons.
 1573 The regions are defined with identical object identification selection and reversal as the 3ℓ case, and
 1574 the same transfer factors can be used. They must be used twice however, because there are two anti-
 1575 identified leptons in each event. Second, the jet inclusive regions are extrapolated into the 2-jet signal

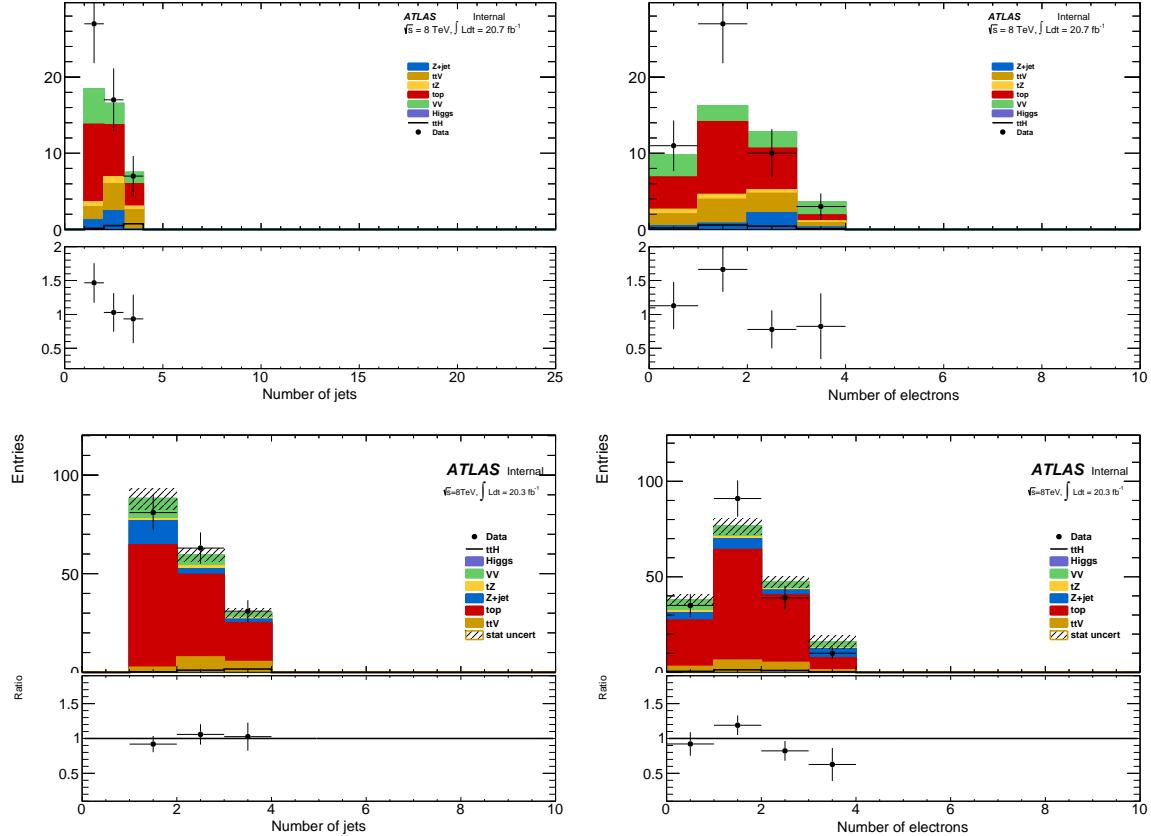


Figure 8.15: 3ℓ fake validation regions for nominal p_T selection (above) and relaxed p_T selection, > 10 GeV, (below). Plotted are the number of jets and the number of electrons in each event. The data and MC ratio below each plot agree with 1 within the statistics of the region and the overall systematic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

region, using a second extrapolation factor derived from $t\bar{t}$ events. Since, the majority of fake leptons arise from b-quark initiated jets, the jet spectrum $t\bar{t}$ events with the additional requirement of 2-b-tagged jets from data are used as a model for the jet extrapolation. The overall systematic uncertainty on this measurement arises from the statistics in the control regions and MC based assessments of non-closure and are 35%-50% depending on the sub-channel.

1581 CHAPTER 9

1582 Summary of Systematic Uncertainties

1583 This chapter summarizes the systematic uncertainties that enter the measurement of the limit of
1584 $t\bar{t}H$ multi-lepton analysis. The systematic uncertainties arise from three main sources. The first are
1585 the normalization uncertainties on the background process estimation methods, which are discussed
1586 in depth in Chapter 8. The second source is the theoretical uncertainties on the $t\bar{t}H$ production
1587 cross-section and acceptance. The final source are the experimental and detector related systematic
1588 uncertainties related to event selection efficiencies and measurements and identification of the objects.
1589 They affect only the non-data driven backgrounds and the $t\bar{t}H$ signal, as simulation is used to model
1590 their acceptance and efficiency for the analysis selection.

1591 These systematic uncertainties, the estimated background and signal event counts in each of the
1592 signal regions, and the observed data in each signal region are combined in a statistical fit to an
1593 analysis model to extract the measurement of interest. We measure per-channel and combined ratios
1594 of the observed production rate to the theoretically predicted production rate of $t\bar{t}H$, a parameter
1595 called μ . In the absence of a statistically significant observation, this measurement is translated into
1596 a upper confidence limit on μ . The details of this procedure are discussed in the following sections
1597 and the results with the observed data are discussed in Chapter 10

1598 9.1 Systematic Uncertainties on Signal Cross-section and Acceptance

1599 The $t\bar{t}H$ signal is simulated with matrix elements at NLO in QCD with Powheg. The simulation
1600 details are discussed in Chapter 6. The production cross section and the Higgs boson decay branching
1601 fractions together with their theoretical uncertainties from the QCD scale and PDF choice are taken
1602 from the NLO theoretical calculations reported in [68]. The uncertainty from the QCD scale estimated
1603 by varying μ_0 by a factor of 2 from the nominal value is $^{+3.8\%}_{-9.3\%}$, while the uncertainty from the PDF

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

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

1604 set and the value of α_S is $\pm 8.1\%$.

1605 The impact of the choice of the QCD scale on the simulation of the $t\bar{t}H$ event selection efficiency
1606 is estimated in two independent ways.

1607 First, the factorization and renormalisation scales μ_0 are varied by a factor of 2, as $\mu = 2\mu_0$ and
1608 $\mu = \mu_0/2$. The effects of these new scales are estimated via the application of event re-weighting
1609 procedures on the nominal simulation using kinematic distributions at parton level. The weights used
1610 are dependent on the transverse momenta of both the $t\bar{t}H$ system and of the top quark, as described
1611 in [83].

1612 Second, the choice of the factorization and renormalisation scales, dependent on fixed (static) pa-
1613 rameters in the nominal simulation, is tested comparing its prediction with an alternative (dynamic),
1614 but still physics motivated choice $\mu_0 = (m_T^t m_T^{\bar{t}} m_T^H)^{\frac{1}{3}}$, which depends on kinematic variables. This
1615 comparison is performed via event re-weighting of the nominal static simulation based on weights
1616 derived as a function of the $t\bar{t}H$ transverse momentum [83]. In order to take the difference between
1617 the choices of scale as systematic uncertainties, a symmetric envelope around the nominal simulation
1618 is built applying the weights and also their inverses.

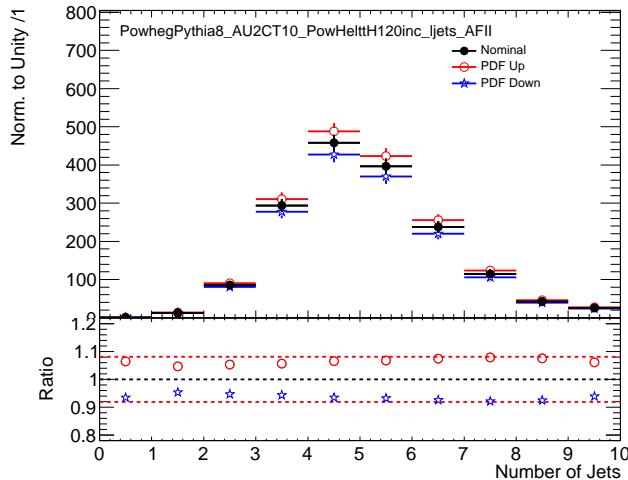
1619 In order to not double-count the variations on the total cross section the predictions from the
1620 different QCD scales are normalized to the same total cross section. That means that the observed
1621 differences are only coming from the event selection. Significant variations on the jet multiplicities can
1622 be seen and these translate into different predictions on the signal event yields in the signal regions.
1623 Such differences, listed in Table 9.1, are taken as theoretical systematic uncertainties in addition to
1624 the ones affecting the total $t\bar{t}H$ production cross section. The static uncertainties come from the
1625 variations by a factor of 2 from the nominal scale and they are correlated with the uncertainties on
1626 the total cross section, which are estimated with the same procedure. The dynamic uncertainties
1627 come from the difference between the nominal and the alternative dynamic scale and are treated as
1628 an independent source of theoretical uncertainty.

1629 The uncertainty of the $t\bar{t}H$ event selection due to the PDF sets is estimated comparing the predic-
1630 tions with three different PDF sets, varying each set within errors and taking the width of the envelope

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

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

as systematic uncertainty. The recommended sets are CT10, MSTW2008nlo68cl and NNPDF21_100. Figure 9.1 shows the estimated PDF systematic uncertainties as a function of the jet multiplicity in $t\bar{t}H$ events with at least two leptons. The uncertainties are compatible with the uncertainty on the production cross section estimated in [68] and indicated by the dashed red lines in the lower panel. Table 9.2 shows the half-width of the envelope of the acceptance under all eigenvector variations of the three PDF sets. No significant dependence on the event topology is observed, so that the PDF systematic uncertainty on the $t\bar{t}H$ event selection is neglected.

Figure 9.1: 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.

1637

9.2 Experimental and Detector Systematic Uncertainties

Experimental and detector systematic uncertainties affect the efficiencies of identifying objects and the efficiencies for events to pass our cuts. These uncertainties affect only MC models of physics processes, $t\bar{t}V$, $t\bar{t}H$, VV and thus alter the number of expected events from signal and background in our signal regions. Data-driven backgrounds take into account these uncertainties by construction. We consider systematic effects from a number of sources: the lepton and jet energy scale measurements,

the lepton identification and isolation selections, the efficiency and misidentification rate associated with tagging b-quark jets. Effects due to modeling the energy and objects from additional vertices were studied and found to be negligible. The vast majority of the individual detector systematics effects are small. The sum total of the systematic effects are comparable to the overall normalization and cross-section uncertainties on some of the physics processes and is shown in Table 9.3.

9.2.1 Lepton Identification, Energy Scale, and Trigger

The electron[51] and muon identification efficiencies[84] are measured in data using Z boson and J/Ψ control samples. They muon efficiencies are shown in Figure 9.2, while the lectron efficienceis are shown in Figure 4.11. The uncertainty on the muon efficiencies are measured as functions of η and p_T and are generally less 1%. The uncertainty on the electron and muon efficiencies are also measured as functions of η and p_T and are at the 1% level for p_T above 30 GeV, but become much larger 5-10% for the lower p_T regimes. These translate into sub-1% level effects on the $t\bar{t}V$ and $t\bar{t}H$ event counts in the signal regions for the muons and 1% level effects for the electrons. The effects become more important with increasing numbers of leptons.

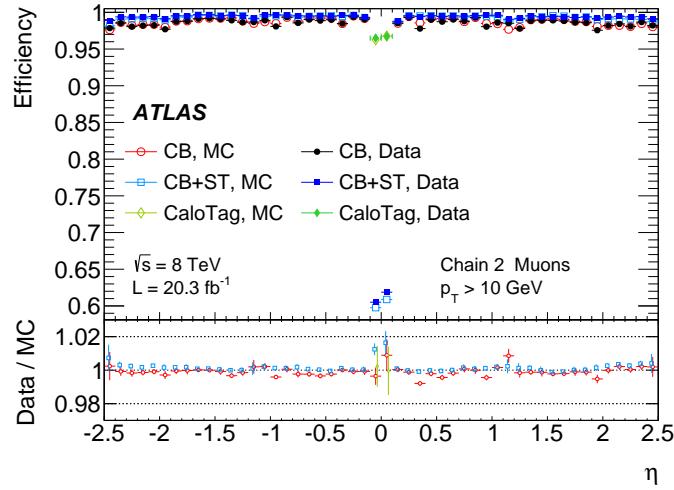


Figure 9.2: Muon identification efficiency in Data and MC as a function of η . The CB+ST (combined+segment tagged) operating point is used

The electron[85] and muon[84] energy scale and resolution are also measured using the Z -boson control samples in data. The uncertainties related to the scale and resolution for the leptons affect the overall event acceptance through the lepton momentum cuts primary and have negligible impact on the event count uncertainties in the signal regions.

1662 The efficiencies for muons and electrons to pass muon[86] and electron triggers[87] have been
1663 calculated with respect to the offline identification operating points using the Z boson control samples.
1664 They are in the range of 90% for electron triggers and 70% for muon triggers, owing to gaps in muon
1665 trigger coverage, and have 1% level uncertainties. When statistically combined for 2ℓ SS, 3ℓ and 4ℓ
1666 lepton signal regions, the overall trigger efficiency is high and the uncertainties on the number of
1667 expected events is negligible.

1668 **9.2.2 Lepton Isolation and Impact Parameter**

1669 The isolation and impact parameter selections are specific to this analysis and are discussed in depth
1670 in Chapter 7. We calculated their combined efficiency with respect to the full lepton identification
1671 selection using the Z boson control samples and define data-MC scale factors to correct the efficiency
1672 in the simulation. Background are subtracted using shape templates in the di-lepton invariant mass
1673 spectrum. The Z -event template is derived from MC, while the background template is derived from
1674 the same-sign control region. We measure the efficiency scale-factors in bins of lepton momentum.
1675 Uncertainties are assigned per-bin to account for the level of statistics and variations caused by the
1676 fit parameters. An additional 1% uncertainty envelope is added to both the electron and muon
1677 measurements to account for trends observed in the dependence of the data-MC efficiency scale-
1678 factor as a function of the number of jets. Stability of the efficiency scale-factor as a function of the
1679 number of jets is important for this analysis, because event activity in the low jet Z sample, where
1680 the efficiency is measured, is much different from the high jet signal regions, where the efficiency
1681 is applied. The dependence of the scale-factor on the number of jets can be seen Figure 9.3. The
1682 isolation scale-factor uncertainties are around 1-3% depending on the particle momentum, but these
1683 uncertainties propagate to 2-5 % (some of the largest) effects in the event counts in the signal regions.
1684 The uncertainties are more important in the regions with more leptons.

1685 **9.2.3 Jet Energy**

1686 The jet energy scale (JES) is calculated using a combination of data-based in-situ techniques, where
1687 jet transverse momentum is balanced with respect to a reference photon or a Z boson, as well as
1688 single particle test-stand studies[88]. Additional smaller effects are taken into account including the
1689 b-quark jet specific response, near-by jets, the effects of pile-up and an inter-calibration of similar
1690 η regions using di-jet events. These effects are measured in 2012 data. The JES systematic errors
1691 arises from numerous sources that are diagonalized into eigenvectors so that they can be combined
1692 in an uncorrelated way. The combined uncertainty is plotted in Figure 9.4 as a function of jet η and

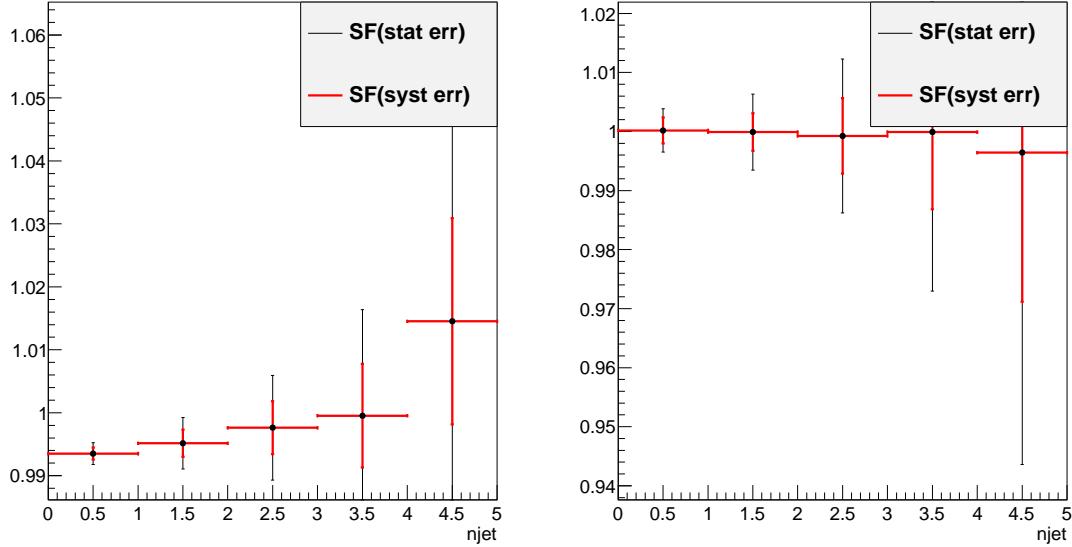


Figure 9.3: Muon (left) and electron (right) isolation efficiency scale-factors from the Z control sample as a function of the number of jets in the event. An additional systematic uncertainty of 1% is added to encompass the variation in the number of jets variable

1693 p_T and is the range 2-4% for jets used in this analysis. The jet energy resolution is calculated in a
 1694 similar way with slightly larger errors, 10% [89]. The combined scale and resolution systematics are
 1695 of non-negligible effects 6-7% on the signal and background event counts in the signal regions.

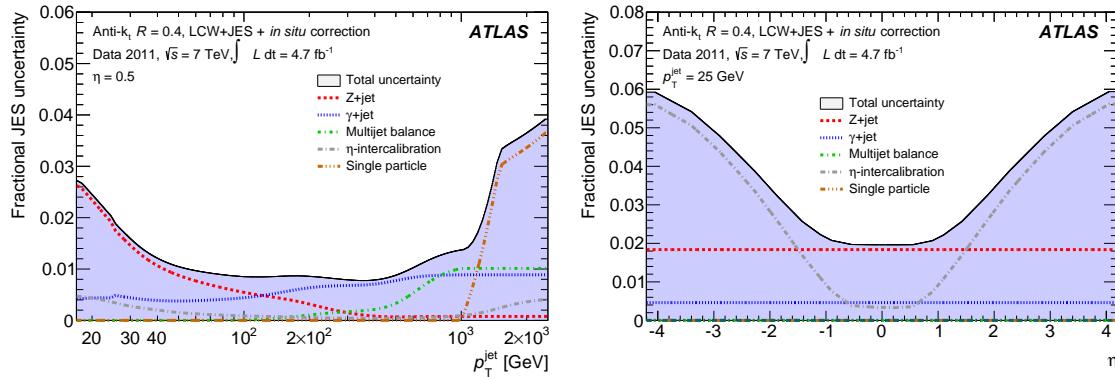


Figure 9.4: JES systematic uncertainties as a function of jet η (for jets $p_T > 25$ GeV) and p_T (for jets $|\eta| < 0.4$). The combined systematic uncertainty is shown with contributions from the largest sources

1696 9.2.4 B-Tagged Jet Efficiency

1697 The b -quark tagging efficiency must be calculated separately for charm, light and b -quarks. ATLAS
 1698 uses three data based control regions: an inclusive jet sample for mistagged light quarks[90], the $t\bar{t}$
 1699 sample for b -quarks[91], and a sample of D^* mesons for charm quarks[92]. These efficiencies and rates
 1700 are well-measured in MC and the data-based corrections are small. The data-MC efficiency scale-
 1701 factor shown in Figure 9.5 is close to 1 and has an overall systematic uncertainty of around 5%. The
 1702 uncertainties are applied to the analysis via a number of eigenvectors. Together these uncertainties
 1703 have a 4 % effect in the event expectation in the signal regions.

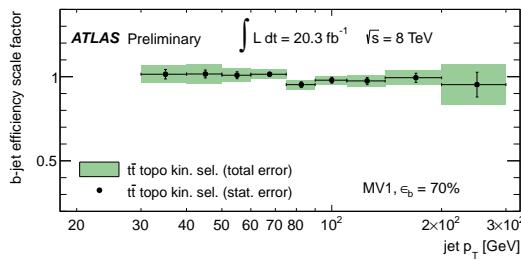


Figure 9.5: b-Tagging data-MC efficiency scale-factors versus jet p_T calculated in the $t\bar{t}$ sample from 2012 data. The uncertainties are combined statistical and systematic.

1704 9.2.5 Summary

1705 The combined effect of these detector and experimental systematics on the $t\bar{t}V$ and $t\bar{t}H$ is provided in
 1706 Table 9.3. The effects are smaller than the normalization uncertainties on some of the backgrounds.
 1707 They are dominated by the lepton isolation scale-factor measurements and the electron identification
 1708 with smaller contributions from the JES and b -tagging efficiencies. These detector systematic uncer-
 1709 tainties enter the fit individually and their ranking of influence on the overall measurement uncertainty
 1710 can be seen in Figure ??.

1711 9.3 Summary of Background and Signal Normalization Uncertainties

1712 Table 9.4 gives the summary of the systematic uncertainties that are included in the analysis for the
 1713 normalization and acceptance of each process. The relative importance of these uncertainties to the
 1714 final fit can be seen in Figure ??/

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.

Type	Description	Uncertainty
Signal (ttH)		
QCD Scale	Cross Section (Dynamic Scale)	+3.8% –9.3% (Section 9.1)
	Analyses Acceptance	0.-2.6%
PDF+ α_S	Cross Section	\pm 8.1%
	Analyses Acceptance	Negligible
ttW (Irreducible background)		
QCD Scale	Cross Section (Dynamic Scale)	\pm 15% (Section 8.1)
	Analyses Acceptance	0.4-3.5%
PDF+ α_S	Cross Section	\pm 13%
	Analyses Acceptance	1.1-4.8%
ttZ (Irreducible background)		
QCD Scale	Cross Section (Dynamic Scale)	\pm 12% (Section 8.1)
	Analyses Acceptance	0.1-3.1%
PDF+ α_S	Cross Section	\pm 9%
	Analyses Acceptance	0.9-2.7%
VV Backgrounds		
Normalization Uncertaintiy	WZ,ZZ Processes	\pm 50% (Section 8.2)
Data-Driven Backgrounds		
Normalization Uncertainty	Jet Fakes	\pm 30-50% (Section 8.4)
Normalization Uncertainty	Charge MisID	\pm 30-40% (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.

1715

CHAPTER 10

1716

Results and Statistical Model

1717 **10.1 Results in Signal Regions**

1718 **10.1.1 2ℓ SS**

1719 **10.1.2 3ℓ**

1720 **10.1.3 4ℓ**

1721 **10.2 Statistical Model**

1722 We use the above results to make two sets of measurements: an upper confidence limit on μ , the
1723 signal strength parameter, and a measurement of μ . These measurements are done for each channel
1724 individually and then combined. The interpretation of the results in the form of a statistical model
1725 follow the procedure, discussed here [93]. We interpret the results as counting experiments in each
1726 signal region. Therefore agreement in kinematic shapes do not affect the statistical procedure.

1727 **10.2.1 The Likelihood**

1728 The observed and expected event yields in the signal regions are analyzed using a binned likelihood
1729 function (\mathcal{L}), built from product of Poisson models of expected event counts for each bin, where the
1730 bins are the separate signal regions:

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

1731 where s_{exp}^i is the SM signal expectation in the signal region, b_{exp}^i are the background expectations, i
1732 counts over the signal regions, and P is the Poisson distribution. The signal strength parameter is the

1733 parameter of interest in the model (POI) and acts as a simple scale-factor to the SM $t\bar{t}H$ production
 1734 rate and is common to all channels. Setting μ to 0 corresponds to the background only scenario. The
 1735 background parameter, b , is a sum over all background processes.

1736 The signal and background expectations, s and b , depend on systematic errors. These are included
 1737 in the likelihood function in the form of a vector nuisance parameters, $\vec{\theta}$, which are constrained to
 1738 fluctuate within Gaussian distributions. These fluctuations affect the background and signal expec-
 1739 tations by response functions, $\nu(\vec{\theta})$, set by systematic uncertainties measured in the previous section.
 1740 For instance, the $W^\pm Z$ normalization uncertainty is 50% from Section 8.2 and is included in the fit
 1741 as its own unit Gaussian, $G(\theta|0, 1)$. The fluctuations of the Gaussian, θ_{WZ} , scale the background con-
 1742 tribution via the form, $0.5 \cdot (1 + \theta_{WZ}) \cdot b_{WZ}$. For many of the detector systematics, the uncertainties
 1743 are two sided and are included as piecewise Gaussians. We add correlations to various uncertainties
 1744 by hand, when appropriate. With these nuisance parameters, the likelihood takes this form:

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

1745 10.2.2 Test Statistic and Profile Likelihood

1746 Values of μ are tested with the negative log quantity, $q_\mu = -2\ln(\lambda(\mu))$, where $\lambda(\mu)$ is the test statistic.
 1747 $\lambda(\mu)$ is defined as:

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

1748 where $\hat{\vec{\theta}}_\mu$ are values of the nuisance parameter vector that maximize the likelihood for a given value
 1749 of μ and $\hat{\mu}$ and $\hat{\vec{\theta}}$ are the fitted values of signal strength and nuisance parameters that maximize the
 1750 likelihood overall. μ is constrained to be positive.

1751 10.2.3 CL_s Method

1752 Exclusions limits on the signal strength are calculated with the test statistic using a modified fre-
 1753 quentist method, called the CL_s method[94]. CL_s is defined as a ratio of two frequentist quantities.
 1754 The numerator quantifies the probability of finding the observed data given the signal + background
 1755 hypothesis. The denominator quantifies the probability of the data given the background only hy-
 1756 pothesis.

1757 Using the numerator alone has the undesirable property that, if the data fluctuates below the
 1758 expectation, an exclusion limit can be reached that is far beyond the sensitivity of the experiment.
 1759 Normalizing to the background only hypothesis penalizes these low sensitivity cases.

1760 The probability of obtaining an observation as extreme as the data given a particular signal +
 1761 background hypothesis is given by the p-value, p_μ defined as:

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

1762 , and the probability of obtaining an observation as extreme as the data given the background hy-
 1763 pothesis, p_b is:

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

1764 , where $f(q_\mu)$ is the distribution of q_μ for all possible observations for a given μ and q is defined above.

1765 Therefore,

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

1766 A μ value is considered excluded at 95% confidence when CL_s is less than 5%.

1767 10.2.4 Exclusion Limits

1768 Table ?? shows *expected* exclusion limits for all channels, including the analysis uncertainties cumula-
 1769 tively. It shows the relative importance of the statistical and systematic uncertainties to the analysis
 1770 sensitivity. The *observed* limits using observed data and predictions can be seen in Figures ??-?? for
 1771 splitting and combining the sub-channels and in Table XX by numbers. We expect a combined limit
 1772 of XXX (background only) and XXX (signal injected) and see a limit of XXX. The channel sensitivity
 1773 is dominated by the 2ℓ and 3ℓ channels.

Channels	Stat	+Fakes Unc.	+Theory	+ Experimental
2ℓ	2 ℓ ee	7.44	8.52	8.94
	2 ℓ em	3.46	3.81	4.18
	2 ℓ mm	4.03	4.14	4.57
	2 ℓ tau	8.08	8.92	10.00
	All	2.16	2.44	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.

1774 10.2.5 μ Measurements

1775 In addition to setting a limit on the signal strength, we also fit the best value of the signal strength
 1776 for μ . We do this by minimizing the negative log likelihood value, q_μ or conversely maximizing the

1777 likelihood. The $1-\sigma$ error band is set via a profile likelihood scan, where the value q_μ is scanned as a
1778 function of μ . Values of μ that increase q_μ^{min} by 1 form the edges of the error band. The fitted values
1779 of μ with errors are provided in Table XXX for each sub-channel fit as well as the combined fit.

1780 **10.2.6 Nuisance Parameter Impact on the Signal Strength**

1781 Finally, we examine the post-fit impact of the various nuisance parameters on the final fit. We expect
1782 to have measured the various analysis uncertainties well and do not expect the fit to have much further
1783 constraint. For that reason, we expect the pulls of the nuisance parameters to be close to 0 and the
1784 measured uncertainties on those parameters to be consistent with the input uncertainties. Figures
1785 XXXX show.

1786

CHAPTER 11

1787

Conclusions

1788 **11.1 Higgs Results in Review**

1789 **11.2 Prospects for Future**

Bibliography

- 1791 [1] S. L. Glashow, *Partial-symmetries of weak interactions*, Nucl. Phys. **22** (1961) no. 4, 579. [2.1.1](#)
- 1792 [2] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264. [2.1.1](#)
- 1793 [3] A. Salam and J. C. Ward, *Gauge theory of elementary interactions*, Phys. Rev. **136** (1964) 763–768. [2.1.1](#)
- 1794 [4] S. Weinberg, *Non-abelian gauge theories of the strong interactions*, Phys. Rev. Lett. **31** (1973) 494–497. [2.1.1](#)
- 1795 [5] G. 't Hooft and M. Veltman, *Regularization and renormalization of gauge fields*, Nuclear Physics B **44** (1972) 189 – 213. [2.1.1](#)
- 1796 [6] D. J. Gross and F. Wilczek, *Ultraviolet behavior of non-abelian gauge theories*, Phys. Rev. Lett. **30** (1973) 1343–1346. [2.1.1](#)
- 1801 [7] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508. [2.1.2](#)
- 1802 [8] P. W. Higgs, *Spontaneous symmetry breakdown without massless bosons*, Phys. Rev. **145** (1966) 1156. [2.1.2](#)
- 1803 [9] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321–322. [2.1.2](#)
- 1805 [10] The ALEPH, CDF, DØ, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the Tevatron Electroweak Working Group, and the SLD electroweak and heavy flavour groups, *Precision Electroweak Measurements and Constraints on the Standard Model*, CERN-PH-EP-2010-095 (2010) , [arXiv:1012.2367 \[hep-ex\]](#). [2.1.3](#), [2.2](#)
- 1811 [11] M. Baak, M. Goebel, J. Haller, A. Hoecker, D. Kennedy, R. Kogler, K. Mnig, M. Schott, and J. Stelzer, *The electroweak fit of the standard model after the discovery of a new boson at the LHC*, The European Physical Journal C **72** (2012) no. 11, .
<http://dx.doi.org/10.1140/epjc/s10052-012-2205-9>. [2.1.3](#)
- 1815 [12] J. C. Collins, D. E. Soper, and G. Sterman, *Factorization for short distance hadron-hadron scattering*, Nuclear Physics B **261** (1985) 104 – 142. [2.2](#)
- 1817 [13] CERN, . CERN, Geneva, 1984. [2.2](#)

- 1818 [14] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and
 1819 R. Tanaka (Eds.), *Handbook of LHC Higgs Cross Sections: 2. Differential Distributions*,
 1820 CERN-2012-002 (CERN, Geneva, 2012) , [arXiv:1201.3084 \[hep-ph\]](https://arxiv.org/abs/1201.3084). **2.2**
- 1821 [15] *Updated coupling measurements of the Higgs boson with the ATLAS detector using up to 25*
 1822 fb^{-1} *of proton-proton collision data*, Tech. Rep. ATLAS-CONF-2014-009, CERN, Geneva, Mar,
 1823 2014. **2.2.1**
- 1824 [16] CMS Collaboration Collaboration, *Precise determination of the mass of the Higgs boson and*
 1825 *studies of the compatibility of its couplings with the standard model*, Tech. Rep.
 1826 CMS-PAS-HIG-14-009, CERN, Geneva, 2014. **2.2.1**
- 1827 [17] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of the Higgs boson mass from*
 1828 *the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4\ell$ channels with the ATLAS detector using 25 fb^{-1} of pp*
 1829 *collision data*, [arXiv:1406.3827 \[hep-ex\]](https://arxiv.org/abs/1406.3827). **2.2.1**
- 1830 [18] *Evidence for the spin-0 nature of the Higgs boson using {ATLAS} data*, Physics Letters B **726**
 1831 (2013) no. 13, 120 – 144.
 1832 <http://www.sciencedirect.com/science/article/pii/S0370269313006527>. **2.2.1**
- 1833 [19] S. Dawson, A. Gritsan, H. Logan, J. Qian, C. Tully, et al., *Working Group Report: Higgs*
 1834 *Boson*, [arXiv:1310.8361 \[hep-ex\]](https://arxiv.org/abs/1310.8361). **2.2.2**
- 1835 [20] O. Eberhardt, G. Herbert, H. Lacker, A. Lenz, A. Menzel, et al., *Impact of a Higgs boson at a*
 1836 *mass of 126 GeV on the standard model with three and four fermion generations*,
 1837 Phys.Rev.Lett. **109** (2012) 241802, [arXiv:1209.1101 \[hep-ph\]](https://arxiv.org/abs/1209.1101). **2.2.2**
- 1838 [21] M. Carena, S. Gori, N. R. Shah, C. E. Wagner, and L.-T. Wang, *Light Stops, Light Staus and*
 1839 *the 125 GeV Higgs*, JHEP **1308** (2013) 087, [arXiv:1303.4414](https://arxiv.org/abs/1303.4414). **2.2.2**
- 1840 [22] N. Arkani-Hamed, K. Blum, R. T. D’Agnolo, and J. Fan, *2:1 for Naturalness at the LHC?*,
 1841 JHEP **1301** (2013) 149, [arXiv:1207.4482 \[hep-ph\]](https://arxiv.org/abs/1207.4482). **2.2.2**
- 1842 [23] D. Carmi, A. Falkowski, E. Kuflik, and T. Volansky, *Interpreting LHC Higgs Results from*
 1843 *Natural New Physics Perspective*, JHEP **1207** (2012) 136, [arXiv:1202.3144 \[hep-ph\]](https://arxiv.org/abs/1202.3144). **2.2.2**
- 1844 [24] G. Degrassi, S. Di Vita, J. Elias-Miro, J. R. Espinosa, G. F. Giudice, et al., *Higgs mass and*
 1845 *vacuum stability in the Standard Model at NNLO*, JHEP **1208** (2012) 098, [arXiv:1205.6497](https://arxiv.org/abs/1205.6497)
 1846 [\[hep-ph\]](https://arxiv.org/abs/1205.6497 [hep-ph]). **2.2.2**
- 1847 [25] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) no. 08, S08001. **3.1**
- 1848 [26] T. S. Pettersson and P. Lefevre, *The Large Hadron Collider: conceptual design.*, Tech. Rep.
 1849 CERN-AC-95-05 LHC, CERN, Geneva, Oct, 1995. <https://cdsweb.cern.ch/record/291782>.
 1850 **3.1**
- 1851 [27] T. Linnecar et al., *Hardware and Initial Beam Commissioning of the LHC RF Systems.*
 1852 [oai:cds.cern.ch:1176380](https://cds.cern.ch/record/1176380), Tech. Rep. LHC-PROJECT-Report-1172.
 1853 CERN-LHC-PROJECT-Report-1172, CERN, Geneva, Oct, 2008.
 1854 <https://cdsweb.cern.ch/record/1176380>. **3.1**
- 1855 [28] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3**
 1856 (2008) S08003. **3.1**

- 1857 [29] The CMS Collaboration, *The CMS experiment at the CERN LHC*, *Journal of Instrumentation*
1858 **3** (2008) no. 08, S08004. **3.1**
- 1859 [30] The LHCb Collaboration, *The LHCb Detector at the LHC*, *Journal of Instrumentation* **3** (2008)
1860 no. 08, S08005. **3.1**
- 1861 [31] The ALICE Collaboration, *The ALICE experiment at the CERN LHC*, *Journal of*
1862 *Instrumentation* **3** (2008) no. 08, S08002.
<http://stacks.iop.org/1748-0221/3/i=08/a=S08002>. **3.1**
- 1864 [32] A. Team, *The four main LHC experiments*, Jun, 1999. **3.1**
- 1865 [33] ATLAS Collaboration, *ATLAS inner detector: Technical Design Report 1*. Technical Design
1866 Report ATLAS. CERN, Geneva, 1997. <https://cdsweb.cern.ch/record/331063>. **3.2**
- 1867 [34] ATLAS Collaboration, *ATLAS inner detector: Technical Design Report, 2*. Technical Design
1868 Report ATLAS. CERN, Geneva, 1997. <https://cdsweb.cern.ch/record/331064>. **3.2**
- 1869 [35] ATLAS Collaboration, *ATLAS magnet system: Technical Design Report, 1*. Technical Design
1870 Report ATLAS. CERN, Geneva, 1997. <https://cdsweb.cern.ch/record/338080>. **3.2**
- 1871 [36] ATLAS Collaboration, *ATLAS pixel detector: Technical Design Report*. Technical Design
1872 Report ATLAS. CERN, Geneva, 1998. <https://cdsweb.cern.ch/record/381263>. **3.2**
- 1873 [37] ATLAS Collaboration, *ATLAS pixel detector electronics and sensors*, *JINST* **3** (2008) P07007.
1874 **3.2**
- 1875 [38] ATLAS Collaboration, *The barrel modules of the ATLAS semiconductor tracker*,
1876 *Nucl.Instrum.Meth.* **A568** (2006) 642–671. **3.2**
- 1877 [39] ATLAS Collaboration, *The ATLAS semiconductor tracker end-cap module*, *Nucl.Instrum.Meth.*
1878 **A575** (2007) 353–389. **3.2**
- 1879 [40] The ATLAS TRT Collaboration, *The ATLAS Transition Radiation Tracker (TRT) proportional
1880 drift tube: Design and performance*, *JINST* **3** (2008) P02013. **3.2**
- 1881 [41] The ATLAS TRT Collaboration, *The ATLAS TRT barrel detector*, *JINST* **3** (2008) P02014. **3.2**
- 1882 [42] E. Abat et al., *The ATLAS TRT electronics*, *J. Instrum.* **3** (2008) P06007. **3.2**
- 1883 [43] T. A. Collaboration, *ATLAS liquid argon calorimeter: Technical design report*, .
1884 CERN-LHCC-96-41. **3.2**
- 1885 [44] ATLAS Collaboration, *ATLAS tile calorimeter: Technical Design Report*. Technical Design
1886 Report ATLAS. CERN, Geneva, 1996. <https://cdsweb.cern.ch/record/331062>. **3.2**
- 1887 [45] ATLAS Collaboration, *ATLAS muon spectrometer: Technical Design Report*. Technical Design
1888 Report ATLAS. CERN, Geneva, 1997. <https://cdsweb.cern.ch/record/331068>. **3.2**
- 1889 [46] G. Aielli, A. Aloisio, M. Alviggi, V. Aprodu, V. Bocci, et al., *The RPC first level muon trigger
1890 in the barrel of the ATLAS experiment*, *Nucl.Phys.Proc.Suppl.* **158** (2006) 11–15. **3.2**
- 1891 [47] F. Bauer, U. Bratzler, H. Dietl, H. Kroha, T. Lagouri, et al., *Construction and test of MDT
1892 chambers for the ATLAS muon spectrometer*, *Nucl.Instrum.Meth.* **A461** (2001) 17–20. **3.2**

- [48] T. Argyropoulos, K. A. Assamagan, B. H. Benedict, V. Chernyatin, E. Cheu, et al., *Cathode strip chambers in ATLAS: Installation, commissioning and in situ performance*, IEEE Trans.Nucl.Sci. **56** (2009) 1568–1574. [3.2](#)
- [49] ATLAS Collaboration, *ATLAS level-1 trigger: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1998. <https://cdsweb.cern.ch/record/381429>. [3.2](#)
- [50] P. Jenni, M. Nessi, M. Nordberg, and K. Smith, *ATLAS high-level trigger, data-acquisition and controls: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 2003. <https://cdsweb.cern.ch/record/616089>. [3.2](#)
- [51] *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data*, Tech. Rep. ATLAS-CONF-2014-032, CERN, Geneva, Jun, 2014. [3.2.6.2](#), [4.1](#), [4.2](#), [7.4](#), [9.2.1](#)
- [52] *Preliminary results on the muon reconstruction efficiency, momentum resolution, and momentum scale in ATLAS 2012 pp collision data*, Tech. Rep. ATLAS-CONF-2013-088, CERN, Geneva, February, 2013. [3.2.6.3](#), [7.5](#)
- [53] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti- $k(t)$ jet clustering algorithm*, JHEP **0804** (2008) 063, [arXiv:0802.1189 \[hep-ph\]](https://arxiv.org/abs/0802.1189). [3.2.6.4](#), [7.6](#)
- [54] ATLAS Collaboration, *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*, Tech. Rep. ATLAS-CONF-2011-102, CERN, Geneva, Jul, 2011. <https://cdsweb.cern.ch/record/1369219>. [3.2.6.5](#), [7.6](#)
- [55] ATLAS Collaboration Collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, Tech. Rep. ATLAS-CONF-2012-047, CERN, Geneva, May, 2012. [4.1](#)
- [56] *Expected electron performance in the ATLAS experiment*, Tech. Rep. ATL-PHYS-PUB-2011-006, CERN, Geneva, Apr, 2011. [4.1](#)
- [57] ATLAS Collaboration Collaboration, G. Aad et al., *Search for $H \rightarrow \gamma\gamma$ produced in association with top quarks and constraints on the Yukawa coupling between the top quark and the Higgs boson using data taken at 7 TeV and 8 TeV with the ATLAS detector*, [arXiv:1409.3122 \[hep-ex\]](https://arxiv.org/abs/1409.3122). [5](#)
- [58] *Search for the Standard Model Higgs boson produced in association with top quarks and decaying to $b\bar{b}$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector at the LHC*, Tech. Rep. ATLAS-CONF-2014-011, CERN, Geneva, Mar, 2014. [5](#)
- [59] ATLAS Collaboration, G. Aad et al., *Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC*, Eur.Phys.J. **C73** (2013) 2518, [arXiv:1302.4393 \[hep-ex\]](https://arxiv.org/abs/1302.4393). [6.1.1](#)
- [60] CERN, . CERN, Geneva, 2012. [6.1.1](#)
- [61] ATLAS Collaboration Collaboration, G. Aad et al., *The ATLAS Simulation Infrastructure*, Eur.Phys.J. **C70** (2010) 823–874, [arXiv:1005.4568 \[physics.ins-det\]](https://arxiv.org/abs/1005.4568). [6.2](#)
- [62] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl.Instrum.Meth. **A506** (2003) 250–303. [6.2](#)

- [63] G. Bevilacqua, M. Czakon, M. Garzelli, A. van Hameren, A. Kardos, C. Papadopoulos, R. Pittau, and M. Worek. [6.2.1](#)
- [64] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, **11** (2004) 040. [6.2.1](#)
- [65] S. Frixione, P. Nason, and C. Oleari **11** (2007) 070, [arXiv:0709.2092 \[hep-ph\]](#). [6.2.1](#)
- [66] S. Alioli, P. Nason, C. Oleari, and E. Re **06** (2010) 040, [arXiv:1002.2581 \[hep-ph\]](#). [6.2.1](#)
- [67] T. Sjöstrand, S. Mrenna, and P. Skands, *A Brief Introduction to Pythia 8.1*, [arXiv:0710.3820 \[hep-ph\]](#). [6.2.1](#)
- [68] LHC Higgs Cross Section Working Group Collaboration, S. Heinemeyer et al., *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, [arXiv:1307.1347 \[hep-ph\]](#). [6.2.1, 9.1, 9.1](#)
- [69] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, **JHEP 0307** (2003) 001, [arXiv:hep-ph/0206293 \[hep-ph\]](#). [6.2.2](#)
- [70] F. Maltoni and T. Stelzer, *MadEvent: Automatic event generation with MadGraph*, **JHEP 0302** (2003) 027, [arXiv:hep-ph/0208156 \[hep-ph\]](#). [6.2.2](#)
- [71] B. P. Kersevan and E. Richter-Was, *The Monte Carlo event generator AcerMC versions 2.0 to 3.8 with interfaces to PYTHIA 6.4, HERWIG 6.5 and ARIADNE 4.1*, **Comput.Phys.Commun.** **184** (2013) 919–985, [arXiv:hep-ph/0405247 \[hep-ph\]](#). [6.2.2](#)
- [72] P. M. Nadolsky, H.-L. Lai, Q.-H. Cao, J. Huston, J. Pumplin, et al., *Implications of CTEQ global analysis for collider observables*, **Phys.Rev. D78** (2008) 013004, [arXiv:0802.0007 \[hep-ph\]](#). [6.2.2](#)
- [73] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, **JHEP 0711** (2007) 070, [arXiv:0709.2092 \[hep-ph\]](#). [6.2.2](#)
- [74] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, et al., *Event generation with SHERPA 1.1*, **JHEP 0902** (2009) 007, [arXiv:0811.4622 \[hep-ph\]](#). [6.2.2](#)
- [75] A. Martin, W. Stirling, R. Thorne, and G. Watt, *Parton distributions for the LHC*, **Eur.Phys.J. C63** (2009) 189–285, [arXiv:0901.0002 \[hep-ph\]](#). [6.2.2, 8.1](#)
- [76] M. Garzelli, A. Kardos, C. Papadopoulos, and Z. Trocsanyi, *$t\bar{t}W^\pm$ and $t\bar{t}Z$ Hadroproduction at NLO accuracy in QCD with Parton Shower and Hadronization effects*, **JHEP 1211** (2012) 056, [arXiv:1208.2665 \[hep-ph\]](#). [8.1](#)
- [77] J. M. Campbell and R. K. Ellis, *$t\bar{t}W^{+-}$ production and decay at NLO*, **JHEP 1207** (2012) 052, [arXiv:1204.5678 \[hep-ph\]](#). [8.1](#)
- [78] Campbell, John and Ellis, R. Keith and Röntsch, Raoul, *Single top production in association with a Z boson at the LHC*, **Phys.Rev. D87** (2013) 114006, [arXiv:1302.3856 \[hep-ph\]](#). [8.1](#)
- [79] ATLAS Collaboration, G. Aad et al., *Measurement of WZ production in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, **Eur.Phys.J. C72** (2012) 2173, [arXiv:1208.1390 \[hep-ex\]](#). [8.2](#)

- [80] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of ZZ production in pp collisions at $\sqrt{s} = 7$ TeV and limits on anomalous ZZZ and ZZ γ couplings with the ATLAS detector*, *JHEP* **1303** (2013) 128, arXiv:1211.6096 [hep-ex]. [8.2](#)
- [81] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of the cross-section for W boson production in association with b-jets in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, *JHEP* **1306** (2013) 084, arXiv:1302.2929 [hep-ex]. [8.2](#)
- [82] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of differential production cross-sections for a Z boson in association with b-jets in 7 TeV proton-proton collisions with the ATLAS detector*, arXiv:1407.3643 [hep-ex]. [8.2](#)
- [83] S. Guindon, E. Shabalina, J. Adelman, M. Alhroob, S. Amor dos Santos, A. Basye, J. Bouffard, M. Casolino, I. Connelly, A. Cortes Gonzalez, V. Dao, S. D'Auria, A. Doyle, P. Ferrari, F. Filthaut, R. Goncalo, N. de Groot, S. Henkelmann, V. Jain, A. Juste, G. Kirby, D. Kar, A. Knue, K. Kroeninger, T. Liss, E. Le Menedeu, J. Montejo Berlingen, M. Moreno Llacer, O. Nackenhorst, T. Neep, A. Onofre, M. Owen, M. Pinamonti, Y. Qin, A. Quadt, D. Quilty, C. Schwanenberger, L. Serkin, R. St Denis, J. Thomas-Wilsker, and T. Vazquez-Schroeder, *Search for the Standard Model Higgs boson produced in association with top quarks and decaying to bb in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector at the LHC*, Tech. Rep. ATL-COM-PHYS-2013-1659, CERN, Geneva, Dec, 2013. The note contains internal documentation of the ttH(bb) analysis approved as a preliminary result (ATLAS-CONF-2014-011). [9.1](#)
- [84] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data*, arXiv:1407.3935 [hep-ex]. [9.2.1](#), [9.2.1](#)
- [85] ATLAS Collaboration Collaboration, G. Aad et al., *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, *Eur.Phys.J.* **C74** (2014) no. 10, 3071, arXiv:1407.5063 [hep-ex]. [9.2.1](#)
- [86] ATLAS Collaboration Collaboration, *Performance of the ATLAS muon trigger in 2011*, Tech. Rep. ATLAS-CONF-2012-099, CERN, Geneva, Jul, 2012. [9.2.1](#)
- [87] ATLAS Collaboration Collaboration, *Performance of the ATLAS Electron and Photon Trigger in p-p Collisions at $\sqrt{s} = 7$ TeV in 2011*, Tech. Rep. ATLAS-CONF-2012-048, CERN, Geneva, May, 2012. [9.2.1](#)
- [88] ATLAS Collaboration Collaboration, G. Aad et al., *Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, arXiv:1406.0076 [hep-ex]. [9.2.3](#)
- [89] ATLAS Collaboration Collaboration, G. Aad et al., *Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector*, *Eur.Phys.J.* **C73** (2013) 2306, arXiv:1210.6210 [hep-ex]. [9.2.3](#)
- [90] ATLAS Collaboration Collaboration, *Measurement of the Mistag Rate with 5 fb^{-1} of Data Collected by the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2012-040, CERN, Geneva, 2012. [9.2.4](#)
- [91] ATLAS Collaboration Collaboration, *Measuring the b-tag efficiency in a top-pair sample with 4.7 fb^{-1} of data from the ATLAS detector*, Tech. Rep. ATLAS-CONF-2012-097, CERN, Geneva, 2012. [9.2.4](#)

- 2012 [92] ATLAS Collaboration Collaboration, *b-jet tagging calibration on c-jets containing D⁰ + mesons*,
2013 Tech. Rep. ATLAS-CONF-2012-039, CERN, Geneva, 2012. [9.2.4](#)
- 2014 [93] Glen Cowan, Kyle Cranmer, Eilam Gross, Ofer Vitells, *Asymptotic formulae for likelihood-based
2015 tests of new physics*, Eur.Phys.J.C **71** (2011) 1554. [10.2](#)
- 2016 [94] A. L. Read, *Presentation of search results: the CL_s technique*, Journal of Physics G: Nuclear
2017 and Particle Physics **28** (2002) no. 10, 2693.
- 2018 <http://stacks.iop.org/0954-3899/28/i=10/a=313>. [10.2.3](#)