A Search for Lepton-Flavor-Violating Decays of the 125 GeV Higgs Boson with Hadronically Decaying Tau Leptons in the 20.3 fb^{-1}, $\sqrt{s} = 8$ TeV Dataset Collected in 2012 by the ATLAS Detector at the Large Hadron Collider

by

Robert Najem Clarke

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Marjorie D. Shapiro, Chair
Professor Robert G. Jacobsen
Professor Karl A. van Bibber

Spring 2016

A Search for Lepton-Flavor-Violating Decays of the 125 GeV Higgs Boson with Hadronically Decaying Tau Leptons in the 20.3 fb^{-1}, $\sqrt{s} = 8$ TeV Dataset Collected in 2012 by the ATLAS Detector at the Large Hadron Collider

Abstract

A Search for Lepton-Flavor-Violating Decays of the 125 GeV Higgs Boson with Hadronically Decaying Tau Leptons in the 20.3 fb^{-1}, $\sqrt{s} = 8$ TeV Dataset Collected in 2012 by the ATLAS Detector at the Large Hadron Collider

by

Robert Najem Clarke

Doctor of Philosophy in Physics

University of California, Berkeley

Professor Marjorie D. Shapiro, Chair

This dissertation presents a search for lepton-flavor-violating (LFV) decays of the 125 GeV Higgs boson in final states of $\mu\tau_{\text{had}}$ and $e\tau_{\text{had}}$ in the full dataset collected by the ATLAS detector in 2012. The search is based on data samples of proton–proton collisions at $\sqrt{s} = 8$ TeV, corresponding to an integrated luminosity of 20.3 fb^{-1}. The topology of LFV events is exploited through an optimized selection of objects and events, with signal and control regions defined by event kinematics. A binned likelihood fit searching for the presence of LFV Higgs decays is performed using the reconstructed Higgs boson mass as the discriminating variable. No such decays are observed. The expected (observed) upper limits at 95% confidence on the branching ratios for the muon and electron final states are found to be $\text{Br}(H \to \mu\tau) = 1.24\%$ (1.85%) and $\text{Br}(H \to e\tau) = 2.07\%$ (1.81%).

For all my dearest family and friends—I wrote this for you.

Contents

Contents **ii**

List of Figures **vi**

List of Tables **xi**

1 Introduction **1**

2 Theory **4**
2.1 The Standard Model 4
2.1.1 Phenomenology 4
2.1.1.1 Electroweak Interactions 7
2.1.1.2 Quantum Chromodynamics 14
2.1.1.3 Higgs Mechanism 18
2.1.2 Higgs Boson Production at Proton–Proton Colliders 21
2.1.3 Decays of the Tau Lepton 24
2.2 Lepton Flavor Violation 25
2.2.1 Motivation 25
2.2.2 Prior Searches 27
2.2.2.1 Searches with the Discovered Higgs Boson 27
2.2.2.2 Other Searches 28
2.2.3 Phenomenology 28
2.2.3.1 Supersymmetric Models 30
2.2.3.2 Inverse Seesaw Model 32
2.2.3.3 Randall–Sundrum Models 33

3 The LHC and the ATLAS Experiment **35**
3.1 The Large Hadron Collider 35
3.1.1 Layout 36
3.1.2 Injection Chain 38
3.1.3 Magnets 40
3.1.4 Luminosity 42

3.2 The ATLAS Detector . . . 43
3.2.1 Coordinates . . . 45
3.2.2 Inner Detector . . . 46
3.2.2.1 Pixel Detector . . . 48
3.2.2.2 Semiconductor Tracker . . . 50
3.2.2.3 Transition Radiation Tracker . . . 51
3.2.3 Magnet Systems . . . 53
3.2.4 Calorimeters . . . 55
3.2.4.1 Electromagnetic Calorimeter . . . 56
3.2.4.2 Hadronic Calorimeters . . . 59
3.2.5 Muon Systems . . . 61
3.2.5.1 Monitored Drift Tubes . . . 63
3.2.5.2 Cathode Strip Chambers . . . 64
3.2.5.3 Resistive Plate Chambers . . . 64
3.2.5.4 Thin Gap Chambers . . . 65
3.2.6 Triggers and Data Acquisition . . . 65
3.2.7 Luminosity Measurement . . . 68

4 Object Reconstruction **71**
4.1 Track and Vertex Reconstruction . . . 71
4.1.1 Track Reconstruction . . . 72
4.1.2 Vertex Reconstruction . . . 73
4.2 Muon Reconstruction and Identification . . . 75
4.3 Electron Reconstruction and Identification . . . 76
4.4 Jet Reconstruction . . . 78
4.4.1 Reconstruction . . . 79
4.4.2 Identification of b-quark Initiated Jets . . . 81
4.5 Hadronic Tau Identification . . . 83
4.6 Missing Transverse Energy Reconstruction . . . 86
4.7 Higgs Boson Mass Reconstruction . . . 87

5 Analysis Strategy **90**

6 Modeling of Signal and Background Processes **93**
6.1 Signal . . . 96
6.2 Background . . . 96
6.2.1 Method of Estimation . . . 96
6.2.2 W + jets . . . 98
6.2.3 $Z \to \tau\tau$ + jets . . . 100
6.2.4 $Z \to \mu\mu/ee$ + jets . . . 103
6.2.5 Top quark . . . 104
6.2.6 Dibosons, VV . . . 105

6.2.7 Standard Model $H \to \tau\tau$ . . . 106
6.2.8 Same-Sign Data, Multijet Background, and r_{QCD} . . . 106

7 Event Selection **109**
7.1 Data . . . 109
7.2 Trigger . . . 109
7.3 Object Definitions . . . 110
7.3.1 Muons and Electrons . . . 110
7.3.2 Jets . . . 110
7.3.3 Hadronic Taus . . . 111
7.4 Object Overlap Removal . . . 111
7.5 Preselection . . . 111

8 Signal and Control Regions **113**
8.1 Event Categorization . . . 113
8.2 Signal Region 1 . . . 118
8.3 Signal Region 2 . . . 119
8.4 W + jets Control Region . . . 119
8.5 Top Control Region . . . 120
8.6 $Z \to \mu\mu/ee$ Control Region . . . 121
8.7 Multijet Control Regions . . . 122
8.7.1 Multijet Control Region 1 . . . 123
8.7.2 Multijet Control Region 2 . . . 123

9 Systematic Uncertainties **134**
9.1 Detector Uncertainties . . . 135
9.1.1 Muon Uncertainties . . . 135
9.1.2 Electron Uncertainties . . . 136
9.1.3 Jet Uncertainties . . . 137
9.1.3.1 Energy Scale and Resolution . . . 137
9.1.3.2 b-Tagging . . . 138
9.1.4 Hadronic Tau Uncertainties . . . 138
9.1.5 Missing Transverse Energy Uncertainties . . . 140
9.1.6 Luminosity . . . 141
9.2 Theory Uncertainties . . . 142
9.2.1 QCD Scale Uncertainties . . . 142
9.2.2 Parton Distribution Function Uncertainties . . . 143
9.2.3 Pile-up Reweighting Uncertainty . . . 144
9.2.4 $H \to \tau\tau$ Branching Ratio Uncertainty . . . 145
9.3 Methodological Uncertainties . . . 145
9.3.1 Uncertainties with the $Z \to \tau\tau$ Embedding Technique . . . 145
9.3.2 r_{QCD} Measurement Uncertainty . . . 145

9.3.3 Shape Uncertainties . . . 146
9.3.4 Extrapolation Uncertainties . . . 146
9.3.5 Uncertainties on k-Factors . . . 147

10 Signal Extraction **148**
10.1 Statistical Tools and Techniques . . . 149
10.1.1 Hypothesis Testing . . . 150
10.1.2 Parameter Estimation . . . 153
10.1.3 Nuisance Parameters . . . 155
10.1.4 Test Statistics and Exclusion Limits . . . 156
10.1.5 Expected Limits . . . 160
10.2 The `HistFactory` Tool . . . 160
10.3 The Structure of the Fit Model . . . 163
10.3.1 Parameters of Interest . . . 164
10.3.2 Free Parameters . . . 164
10.3.3 Nuisance Parameters . . . 165

11 Results **176**
11.1 Muon and Electron Measurements . . . 176
11.2 Combination with the Leptonic Tau Analysis . . . 177

12 Conclusions **185**

Bibliography **187**

A Shape Corrections **199**

B Tests of the Fit Model **208**

List of Figures

2.1 The Strong Coupling α_S as a function of momentum transfer [10]. 15
2.2 An example of a proton–proton collision. 16
2.3 Parton Distribution Functions for $Q^2 = 10$ GeV^2 and $Q^2 = 10^4$ GeV^2 [22]. . . . 17
2.4 The Higgs Potential [23]. 19
2.5 Branching ratios of the Higgs boson as a function of its mass [24]. 21
2.6 Diagrams of the dominant mechanisms for Higgs boson production at a proton–proton collider. 22
2.7 Higgs boson production cross sections as a function of its mass [24]. 23
2.8 Diagram showing a hadronically decaying tau lepton. 24
2.9 Bounds placed on the lepton-flavor-violating Yukawa couplings $|y_{e\tau}|$, $|y_{\tau e}|$, $|y_{\mu\tau}|$, and $|y_{\tau\mu}|$ [45]. The dashed line in the lower left corner of each plot indicates the largest theoretical bound that maintains the observed mass hierarchy $m_\tau > m_\mu > m_e$ without precise selection of theoretical model parameters. 30

3.1 Layout of the LHC Experiments [55]. 37
3.2 The LHC Injection Chain [58]. 39
3.3 Cross-sectional schematic of an LHC dipole magnet [59]. 41
3.4 Cut-away depicting the ATLAS Detector [64]. 45
3.5 Quarter Schematic of the ATLAS Inner Detector [64]. 47
3.6 Cut-away depicting the ATLAS Inner Detector [64]. 48
3.7 Cut-away depicting the ATLAS Pixel Detector [66]. 49
3.8 The ATLAS toroid and solenoid magnets [64]. 54
3.9 Magnetic field strength in the Inner Detector for fixed ϕ (left) and in the Monitored Drift Tubes in the Muon System (right) [64]. 55
3.10 Cut-away depicting the ATLAS calorimeters [64]. 56
3.11 Material budget of the electromagnetic and hadronic calorimeters in terms of interaction lengths (left) and radiation lengths (right) [64]. 57
3.12 Schematic illustrating different absorber layers in the barrel of the ATLAS electromagnetic calorimeter [64]. 58
3.13 Schematic of a tile module in the ATLAS hadronic tile calorimeter [64]. 60
3.14 Cut-away depicting the ATLAS Muon System [64]. 62
3.15 Schematic of the ATLAS Muon System, barrel (left) and end-cap (right) [64]. . 63

3.16 Flowchart illustrating the trigger and data acquisition systems [67]. The 2012 rates are given in the grey boxes with the design values indicated above. 66
3.17 Diagram illustrating trigger algorithms for electrons, photons, and hadronically decaying tau leptons [64]. 67
3.18 Total Integrated Luminosity in 2012 delivered by the LHC, recorded by ATLAS, and good for physics [67]. 70

4.1 Particle detection at ATLAS [70]. 72
4.2 Tracking Reconstruction efficiency as a function of η for muons, pions, and electrons with $p_{\mathrm{T}} = 5$ GeV. 74
4.3 Muon reconstruction efficiency as a function of η for different reconstruction techniques [78]. The drop in acceptance for $|\eta| < 0.1$ is from the placement of detector services, resulting in fewer MS stations. 76
4.4 Electron Reconstruction efficiency for different electron classifications as a function of E_{T} and η [79]. 78
4.5 b-tagging efficiency as a function of p_{T} and $|\eta|$ [87]. 82
4.6 Tau identification efficiency and rejection for one and three prong taus [88]. . . . 84

5.1 Diagram depicting the decays $H \to \mu\tau$ / $H \to e\tau$. The visible part of the tau decay, usually consisting of one or three charged pions and two or fewer neutral pions, is identified as a hadronic tau. The neutrino presence in the detector is characterized by missing transverse energy. 91

6.1 Measured cross sections for various SM processes at energies of 7 TeV and 8 TeV [67]. 94
6.2 Mean Number of Interactions per Bunch Crossing [67]. 95
6.3 Examples of leading order diagrams for W/Z + jets production. 99
6.4 A flowchart illustrating the steps in the embedding procedure for modeling the $Z \to \tau\tau$ background [108]. 101
6.5 Examples of leading order diagrams for $t\bar{t}$ (upper row) and single top production (lower row). 105
6.6 Examples of leading order diagrams for diboson production. 106
6.7 Examples of leading order diagrams for dijet production. 107

8.1 Comparisons of the relative p_{T} spectra of muons, electrons, and hadronic taus for signal LFV and background processes for the muon (top row) and electron (bottom row). These spectra are obtained after the preselection requirements outlined in Section 7.5 have been applied. All distributions are computed using MC-simulated data. 114
8.2 Two-dimensional transverse mass distributions for opposite-sign data, W + jets, $Z \to \tau\tau$, and signal events in the muon measurement [115]. All events pass the preselection criteria outlined in Section 7.5. The boxed areas in the data plot outline the W + jets Control Region and signal regions. 115

8.3 Two-dimensional transverse mass distributions for opposite-sign data, W + jets, $Z \to \tau\tau$, and signal events in the electron measurement [34]. All events pass the preselection criteria outlined in Section 7.5. The boxed areas in the data plot outline the W + jets Control Region and signal regions. 116
8.4 $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in SR1 for the muon measurement. 118
8.5 $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in SR1 for the electron measurement. . . . 119
8.6 $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in SR2 for the muon measurement. 120
8.7 $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in SR2 for the electron measurement. . . . 121
8.8 $m_{\mu\tau}^{\mathrm{MMC}}$, $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in WCR for the muon measurement. 122
8.9 $m_{e\tau}^{\mathrm{MMC}}$, $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in WCR for the electron measurement.123
8.10 $m_{\mu\tau}^{\mathrm{MMC}}$, $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in TCR for the muon measurement. 124
8.11 $m_{e\tau}^{\mathrm{MMC}}$, $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in TCR for the electron measurement.125
8.12 $m_{\mu\tau}^{\mathrm{MMC}}$, $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in ZmmCR for the muon measurement.126
8.13 $m_{e\tau}^{\mathrm{MMC}}$, $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in ZeeCR for the electron measurement. 127
8.14 $m_{\mu\tau}^{\mathrm{MMC}}$, $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in QCDCR for the muon measurement.128
8.15 $m_{e\tau}^{\mathrm{MMC}}$, $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in QCDCR for the electron measurement. 129
8.16 $m_{\mu\tau}^{\mathrm{MMC}}$, $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in QCDCR2 for the muon measurement. 130
8.17 $m_{e\tau}^{\mathrm{MMC}}$, $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in QCDCR2 for the electron measurement. 131

9.1 Distributions of N_{jet} for LFV Higgs boson signal events produced through gluon–gluon fusion in Signal Region 1, Signal Region 2, and after preselection (inclusive).144

10.1 The sampling distribution for z_0, the standard normal distribution. The red shaded areas are the critical region, bounded by $|z_0| > 1.96$ and chosen to comprise 5% of the total area of the distribution, indicating a significance level of 0.05. The line corresponds to the observed z_0 computed in the example. 153
10.2 The sampling distributions for the null and alternative hypotheses used in the example, centered respectively at zero and three. The red shaded area equals the probability of making a type II error since the null hypothesis would not be rejected in favor of the (true) alternative hypothesis. This area is chosen under the assumption that the null hypothesis is being tested at a significance level of 0.05. 154
10.3 An example of two sampling distributions for $\tilde{q}_\mu$ assuming signal + background ($\mu = 1$) and background-only ($\mu = 0$) hypotheses. The p-values for each hypothesis equal the area under the respective curves to the right of the observed value [128]. 159
10.4 Examples of systematic variations used in the fit model, both taken from the electron measurement. 169

11.1 Post-fit $m_{\mu\tau}^{\text{MMC}}$ distributions in SR1 and SR2 in the muon measurement. The red histogram indicates the predicted $H \to \mu\tau$ distribution, assuming $\text{Br}(H \to \mu\tau) = 25\%$ [115]. 177
11.2 Combined Post-fit $m_{\mu\tau}^{\text{MMC}}$ distribution for the muon measurement. The red histogram indicates the predicted $H \to \mu\tau$ distribution for the best-fit $\text{Br}(H \to \mu\tau) = 0.77\%$ [115]. 178
11.3 Post-fit $m_{e\tau}^{\text{MMC}}$ distributions in SR1 and SR2 for the electron measurement. The red histogram indicates the predicted $H \to e\tau$ distribution, assuming $\text{Br}(H \to e\tau) = 25\%$ [34]. 179
11.4 Combined Post-fit $m_{e\tau}^{\text{MMC}}$ distribution for the electron measurement. The red histogram indicates the predicted $H \to e\tau$ distribution, assuming $\text{Br}(H \to e\tau) = 1\%$ [34]. 180
11.5 Expected and Observed Limits in the muon measurement for both hadronic and leptonic tau analyses [34]. 182
11.6 Expected and Observed Limits in the electron measurement for both hadronic and leptonic tau analyses [34]. 183

A.1 Shape corrections applied to the W + jets background in SR1 and QCDCR in the muon measurement. The correction is applied to both OS and SS events. . . 200
A.2 Shape corrections applied to the W + jets background in SR1 and QCDCR in the electron measurement. The correction is applied to both OS and SS events. 201
A.3 Shape corrections applied to the W + jets background for OS events in WCR, TCR, and ZmmCR in the muon measurement. 202
A.4 Shape corrections applied to the W + jets background for SS events in WCR, TCR, and ZmmCR in the muon measurement. 203
A.5 Shape corrections applied to the W + jets background for OS events in WCR, TCR, and ZeeCR in the electron measurement. 204
A.6 Shape corrections applied to the W + jets background for SS events in WCR, TCR, and ZeeCR in the electron measurement. 205
A.7 Shape corrections applied to $Z \to \mu\mu$ + jets events where a jet is misidentified as a hadronic tau. The correction is applied in all signal and control regions in the muon measurement. 206
A.8 Shape corrections applied to $Z \to \mu\mu$ + jets events where a jet is misidentified as a hadronic tau. The correction is applied in all signal and control regions in the electron measurement. 207

B.1 Top ranked nuisance parameters in the muon measurement. 209
B.2 Top ranked nuisance parameters in the electron measurement. 210
B.3 Profiles of the top six ranked nuisance parameters in the muon measurement. . 211
B.4 Profiles of the top six ranked nuisance parameters in the electron measurement. 212

B.5 Fitted values of the `NormFactor` parameters used in the muon and electron measurements. All values are consistent with unity, indicating no underlying issues with the background modeling. 213

List of Tables

2.1 Standard Model particles and their properties [10]. Antiparticles have the opposite charge, baryon number, and lepton number to their particle counterparts. . 7
2.2 The 125 GeV Higgs boson production cross sections in picobarns for proton–proton collisions at $\sqrt{s} = 8$ TeV [24]. 23
2.3 Measured branching ratios for different decay modes of the tau lepton. All decay modes include the tau (anti)neutrino [10]. 25
2.4 90% confidence limits on branching ratios for various lepton-flavor-violating processes, with the experimental collaboration and reference. The limit obtained for $\mu \to e$ conversion in titanium was for ground state (GS) conversions of the nuclei. 29

3.1 Values of the run parameters of the LHC. The abbreviation IP refers to the interaction point. The 2012 value of the geometric correction is computed based on a formula in [60]. 44

6.1 Summary of all scale factors used for the background estimation. The normalization factors are listed with no uncertainties because they are used as free parameters in the binned likelihood fit used to perform the measurement. The uncertainty on the r_{QCD} factor is statistical + systematic. All other uncertainties are statistical only. 108

8.1 Event selection criteria defining the signal and control regions in both muon and electron measurements. The symbol ℓ indicates a muon or electron. Asterisks (*) indicate that the transverse mass requirements for QCDCR are the union of those for SR1 and SR2. 117
8.2 Measured values of r_{QCD} in QCDCR and QCDCR2 for both muon and electron measurements. The listed uncertainties are statistical only. The measured numbers are in excellent agreement with the values of r_{QCD} obtained in the SM $H \to \tau\tau$ analysis: $r_{\mathrm{QCD}} = 1.10 \pm 0.14$ and 1.00 ± 0.13, respectively, for the muon and electron measurements. 124

8.3 Event yields for signal and background predictions in SR1, SR2, WCR, and TCR after applying the optimized selection cuts. The signal predictions are obtained assuming a branching ratio of $\mathrm{Br}(H \to \mu\tau) = 1.0\%$. Statistical and systematic uncertainties are respectively quoted for each process. The systematic uncertainties on the total background predictions account for correlations between the uncertainties for different background components. 132
8.4 Event yields for signal and background predictions in SR1, SR2, WCR, and TCR after applying the optimized selection cuts. The signal predictions are obtained assuming a branching ratio of $\mathrm{Br}(H \to e\tau) = 1.0\%$. Statistical and systematic uncertainties are respectively quoted for each process. The systematic uncertainties on the total background predictions account for correlations between the uncertainties for different background components. 133

10.1 The binning scheme used in the fit model. The second column indicates the number of bins in each channel histogram. The third column lists the range of $m^{\mathrm{MMC}}_{\mu/e\tau}$ values corresponding to a given bin. 163
10.2 The floating parameters (`NormFactor` elements) used in the fit model. 165
10.3 Nuisance parameters used in the fit model. 165
10.3 Nuisance parameters used in the fit model. (continued) 166
10.4 Nuisance parameters used in the fit model for SR1 in the muon measurement following the described pruning procedure. The labels "O Sys", "HSys", and "OHSys" indicate that an `OverallSys` element, `HistoSys` element, or both, respectively, are implemented for the given nuisance parameter. The label "ATLAS_" is omitted for brevity. 170
10.4 Nuisance parameters used in the fit model for SR1 in the muon measurement following the described pruning procedure. The labels "O Sys", "HSys", and "OHSys" indicate that an `OverallSys` element, `HistoSys` element, or both, respectively, are implemented for the given nuisance parameter. The label "ATLAS_" is omitted for brevity. (continued) 171
10.5 Nuisance parameters used in the fit model for SR2 in the muon measurement following the described pruning procedure. The labels "O Sys", "HSys", and "OHSys" indicate that an `OverallSys` element, `HistoSys` element, or both, respectively, are implemented for the given nuisance parameter. The label "ATLAS_" is omitted for brevity. 171
10.5 Nuisance parameters used in the fit model for SR2 in the muon measurement following the described pruning procedure. The labels "O Sys", "HSys", and "OHSys" indicate that an `OverallSys` element, `HistoSys` element, or both, respectively, are implemented for the given nuisance parameter. The label "ATLAS_" is omitted for brevity. (continued) 172

10.6 Nuisance parameters used in the fit model for SR1 in the electron measurement following the described pruning procedure. The labels "O Sys", "HSys", and "OHSys" indicate that an `OverallSys` element, `HistoSys` element, or both, respectively, are implemented for the given nuisance parameter. The label "ATLAS_" is omitted for brevity. 173
10.6 Nuisance parameters used in the fit model for SR1 in the electron measurement following the described pruning procedure. The labels "O Sys", "HSys", and "OHSys" indicate that an `OverallSys` element, `HistoSys` element, or both, respectively, are implemented for the given nuisance parameter. The label "ATLAS_" is omitted for brevity. (continued) 174
10.7 Nuisance parameters used in the fit model for SR2 in the electron measurement following the described pruning procedure. The labels "O Sys", "HSys", and "OHSys" indicate that an `OverallSys` element, `HistoSys` element, or both, respectively, are implemented for the given nuisance parameter. The label "ATLAS_" is omitted for brevity. 174
10.7 Nuisance parameters used in the fit model for SR2 in the electron measurement following the described pruning procedure. The labels "O Sys", "HSys", and "OHSys" indicate that an `OverallSys` element, `HistoSys` element, or both, respectively, are implemented for the given nuisance parameter. The label "ATLAS_" is omitted for brevity. (continued) 175

11.1 Events yields in the 110 GeV $< m_{\mu\tau}^{\text{MMC}} <$ 150 GeV region in the muon measurement [115]. The signal yields assume $\text{Br}(H \to \mu\tau) = 0.77\%$. Event yields for the background processes are obtained after a combined fit to SR1, SR2, WCR, and TCR. The uncertainties on each process are respectively statistical and systematic. The systematic uncertainty on the total background prediction accounts for all correlations between nuisance parameters. 181
11.2 Events yields in the 110 GeV $< m_{e\tau}^{\text{MMC}} <$ 150 GeV region in the electron measurement. The signal yields assume $\text{Br}(H \to e\tau) = 1.0\%$. Event yields for the background processes are obtained after a combined fit to SR1, SR2, WCR, and TCR. The uncertainties on each process are respectively statistical and systematic. The systematic uncertainty on the total background prediction accounts for all correlations between nuisance parameters. 181
11.3 Results of the searches for LFV $H \to e\tau$ and $H \to \mu\tau$ decays. The limits are computed assuming that either $\text{Br}(H \to \mu\tau) = 0$ or $\text{Br}(H \to e\tau) = 0$. The expected and observed upper limits at 95% confidence level and the best-fit values of the branching ratios for the individual categories and their combination are listed. [34] 184

A.1 Binning scheme used to obtain shape corrections for the W + jets and $Z \to \mu\mu/ee$ + jets backgrounds. The symbol ℓ indicates a muon or electron. 199

Acknowledgments

This dissertation marks the end of a long journey that began eight years ago, carrying me across this country and abroad, in Europe. Though the path to a Ph.D. was often lonely, I could not have done it on my own. I do not know of adequate words to truly express my gratitude to the people who helped me along the way, but I want to acknowledge key contributions.

I want to first thank my adviser, Professor Marjorie D. Shapiro, for accepting me as one of her students and allowing me to pursue this journey. Her generous support and patience were critical to my success. She kept a watchful eye on my progress, while allowing me time to reflect and learn. I am also grateful to her and Ian Hinchliffe for graciously offering concert tickets, which provided a welcome respite from my studies.

It is difficult to me to imagine how this dissertation would have proceeded without Aliaksandr "Sasha" Pranko, who took great pains and responsibility for my education. Sasha's experience with the $H \to \tau\tau$ analysis was invaluable to my success. From him, I learned how to be a physicist—how to present information so that it is best comprehended, how to navigate the politics of working with a group of collaborators, and how to argue convincingly. I can't count the number of 5:00 AM meetings that he attended to ensure that my presentations went smoothly. I will miss working with him.

I want to thank the following people of the ATLAS LBNL group (past and present) for their insight and support: Ian Hinchliffe, Juerg Beringer, Lina Galtieri, Mike Hance, Mark Tibbetts, Mark Cooke, Simon Viel, Maurice Garcia-Sciveres, Zach Marshall, Martina Hurwitz, Seth Zenz, Lauren Tompkins, Max Scherzer, Andre Bach, Peter Loscutoff, Louise Skinnari, Andreas Korn, Shih-Chieh Hsu, David Yu, Ana Ovcharova, Alex Sood, Jackie Brosamer, Steven Farrell, Rebecca Carney, Timon Heim, Michel Trottier-McDonald, Brian Amadio, and Jennet Dickinson. I particularly want to thank Dimitris Varouchas, who was a wonderful mentor during my first years in the group. I also want to acknowledge Peter, Alex, and Michel for their well-written, pedagogical dissertations, which served as examples to this one.

I met many people at CERN who offered advice and insight. I would like to thank Stan Lai, Stefania Xella, Alex Tuna, Ryan Reece, Martin Flechl, Michel Janus, Koji Nakamura, Elias Coniavitis, and Sinead Farrington for their feedback on my tau studies and dissertation analysis. I want to thank Carl Jeske and Nils Ruthmann for their help with the common ntuples, and Nils also for his help with the statistics software. I want to especially thank Michel Trottier-McDonald and Emma Ideal for their warmth and friendship.

I was also lucky to work with an exceptional editorial board and collaborators for the LFV Higgs/Z boson paper. I want to thank Kevin Black, Swagato Banerjee, Stefania Xella, Lorenzo Bellagamba, and Erez Etzion for their useful feedback and Swagato for teaching me how to make workspaces, as well. I want to thank Olga Igonkina, Hartger Weits, and Shikma Bressler for their collaboration. I especially want to thank Davide Gerbaudo for helping me with plot issues, and for his role in editing the paper.

I want to thank all those who agreed to read my dissertation, but especially my grandmother, Ellen Clarke, who provided detailed editorial commentary over its entirety. I also want to thank Katherine Gray for her editorial feedback, as well. All errors, of course, remain my own.

My success could not have been possible without the love and support of my friends and family. I want to thank my paternal grandparents, James and Ellen Clarke, and my maternal grandfather, G. Reza Najem, for their wonderful example, encouragement, and support along every step of the way. I am especially grateful to Grandma for her calls, emails, and care packages, and I consider myself lucky to have spent time with my paternal grandparents in England.

I want to thank my parents, David and Shima, without whom I cannot imagine how I would have achieved this degree. I want to thank Mom for her calls and emails, for always asking if I needed anything, and for traveling to visit me, no matter where I was in the world. I want to thank Dad for his love and encouragement, and for instilling a love and appreciation for music. I'm grateful for his visits.

I want to thank my brother, Christopher Clarke, and my sister, Kathleen, two of my greatest inspirations. Chris is an outstanding physician, and Kathleen excels at whatever she applies herself. I want to thank Chris for visiting me at CERN and for coming to my graduation.

I want to thank Auntie, Uncle Mike, Catherine, Jamie, Sean, Mamah Jawad, Aunt Frozan, Jaffar, Alex, Zohal, Sophie, Julie, Aunt Leslie, Melissa, Claire, Aunt Cathy, Alexis, Carlton, Tessa, Elizabeth, Mamah Ali, Aunt Linda Lee, and Grandmother Leila for their love and support. I especially want to thank Mamah Jahed, Khala Ruhina, Sabrina, and Tawab for treating me like their own son and brother, helping me feel at home. I can't thank you enough for your generosity and kindness. I also want to thank Kathryn and Richard Warren for their great hospitality and kind encouragement. I hope to see you two soon.

I want to thank Mae Blackwell and Nicki Langway for their love and friendship. I want to thank Mae for visiting me in Berkeley and Europe, for sending me the most thoughtful care packages, and for always being there. I want to thank Nicki for visiting me in Berkeley and for checking-in on me frequently, offering her compassion and kindness. To both of you, I apologize for sometimes letting months pass between conversations. I'll try to minimize this in the future.

I am grateful to Sandy Miarecki for her friendship and support. It's difficult for me to imagine how I would have made it through the first years of graduate school without her.

I also want to thank Rick Alfaro for his friendship, treating me like family, having me over at his house often, inviting me out to the movies, showing me uncommon generosity and kindness, and teaching me how to drive a manual-transmission vehicle. You're a brave man, Rick.

Music has also been crucial to the successful completion of my degree. I want to thank Professors Davitt Moroney and Marika Kuzma for graciously allowing me to participate in the University Baroque Ensemble and University Chorus, respectively. I cannot express how much those experiences meant to me. I am grateful to Johann Sebastian Bach, George

Frideric Handel, Sylvius Leopold Weiss, Antonio Vivaldi, John Dowland, and the other masters whose music kept me company during the long hours writing.

Lastly, I want to thank Katherine for her endless love, encouragement, and support these past years. You will be an outstanding singer and scholar.

Chapter 1

Introduction

Physics is the science of natural phenomena and like all sciences, may be divided into disciplines. *Particle physics* is the study of the fundamental constituents of matter and their interactions. Its philosophical beginnings can be traced back millennia, but the empirical foundations were established in the latter part of the nineteenth century through advances in chemistry and the discovery of the electron by J.J. Thomson in 1897. Advances in quantum field theory and experimental developments in high energy colliders culminated in the formulation of the *Standard Model* (SM) in the 1960s and early 1970s.

It is difficult to understate the success of the SM. Though it has faced experimental scrutiny over many orders of magnitude in energy, every test has confirmed its predictions. Notable cases include the discovery of weak neutral currents in Gargamelle at CERN in 1973 [1], followed by the discovery of the W and Z bosons and the verification of their predicted masses in 1983 [2]–[5]. The discovery of the top quark in 1995 at the Tevatron [6], [7] and the most recent discovery of a neutral, scalar Higgs boson at the Large Hadron Collider (LHC) in 2012 [8], [9] have further underscored the predictive power of the theory. With the discovery of the Higgs boson, all of the pieces of the Standard Model have been experimentally verified.

Despite these great successes, the SM is known to be an incomplete theory. Notably, it offers no explanation for gravity. In addition, though many of the predicted particles are massive, the SM cannot account for why the particles have their respective masses. This issue is known as the *hierarchy problem*. There are many phenomenological mysteries that are not predicted in the SM. Astrophysical observations indicate that galaxies are suffused with a form of matter called *dark matter* that is invisible but interacts gravitationally and possibly weakly. Experimental measurements of the energy content of the known universe estimate that dark matter comprises $\sim$27%, which is staggering compared to the $\sim$5% due to ordinary matter [10]! Despite its large energy content, the nature of dark matter remains a mystery, though there are many research efforts under way to discern its character.

The nature of the remaining $\sim$68% [11] of the energy content of the universe poses another mystery. The current cosmological model of the universe is that it originated roughly 14 billion years ago in a singular event called *the Big Bang*. According to this theory, the universe is not static, but rather expanding. It might be expected that this expansion should

wane over time with the universe eventually contracting, but measurements indicate that the expansion rate is *increasing*.[1] That is, it appears the universe is growing larger at an accelerated rate! The mechanism for this expansion lies outside of the SM and is called *dark energy*.

Another mystery concerns the nature of neutrinos. While they will be discussed in more detail later, one of the predictions of the SM is that neutrinos are massless particles. Yet, measurements indicate that they possess a mass (albeit small). The mechanism by which neutrinos obtain mass is absent in the SM.

The recently discovered Higgs boson offers several experimental avenues for probing physics *beyond the Standard Model* (BSM). This dissertation presents an analysis of one such avenue: a search for decays of the Higgs boson to two leptons of different flavors, or *lepton-flavor-violating* (LFV) decays. Chapter 2 presents a theoretical overview of the Standard Model and the phenomenology of LFV decays. Chapters 3 and 4 discuss the Large Hadron Collider, ATLAS detector, and how measurements are combined to create the physics objects used for this analysis. Chapter 5 provides a brief overview of the analysis strategy. Chapter 6 describes the modeling of the LFV Higgs boson decay and physics processes that can mimic it. Chapter 7 discusses the selection criteria applied to choose data most sensitive to the LFV Higgs decay. Chapter 8 explains how events are categorized into different regions of interest to improve the analysis sensitivity. Chapter 9 discusses sources of uncertainty in the analysis. The statistical machinery used to make the measurement are described in Chapter 10, with the obtained results presented in Chapter 11. The interpretation of the results and future improvements to the analysis are discussed in Chapter 12.

Conventions and Mathematical Notation

In this dissertation, natural units are used such that $\hbar = c = 1$ and the elementary charge $e = 1$. Thus, all masses and momenta are given in units of GeV, unless otherwise stated. The Minkowski metric is

$$\eta^{\mu\nu} = \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Four-vectors are denoted as $C^\mu = (C^0, \mathbf{C})$, where C^0 is the time component of the four-vector and $\mathbf{C}$ is the spatial (vector) component. Lowering the index, in keeping with the metric signature, yields $C_\mu = (C^0, -\mathbf{C})$. The Einstein summation convention is used, where repeated indices imply summation over all (usually four) components. For example, $C^\mu C_\mu = C \cdot C = (C^0)^2 - \mathbf{C} \cdot \mathbf{C}$. Commonly used four-vectors are the position four-vector $x^\mu = (t, \mathbf{r})$, where t and $\mathbf{r}$ are the time coordinate and position vector; the energy–momentum four-vector $p^\mu = (E, \mathbf{p})$, where E and $\mathbf{p}$ are the energy and momentum; and the derivative four-vector

[1] This measurement culminated in a Nobel Prize awarded in 2011.

$\partial_\mu = (\frac{\partial}{\partial t}, \nabla)$, where ∇ is the del operator from vector calculus. The invariant mass m of an energy–momentum four-vector p^μ is computed as $m^2 = p^\mu p_\mu = E^2 - \mathbf{p} \cdot \mathbf{p}$.

From electrodynamics, $F^{\mu\nu}$ refers to the *electromagnetic field tensor*, and $A^\mu = (\Phi, \mathbf{A})$ is the electromagnetic four-potential, where Φ is the scalar potential (voltage or potential difference) and $\mathbf{A}$ is the magnetic vector potential, related to the electromagnetic field tensor by the relation $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$.

The *Dirac Matrices* γ^μ are 4×4 matrices that satisfy the anti-commutation relation $\{\gamma^\mu, \gamma^\nu\} = \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu} \cdot \mathbf{I}_{4\times4}$, where $\mathbf{I}_{4\times4}$ is the 4×4 identity matrix. For a *Dirac spinor* ψ, its conjugate, or *Dirac adjoint*, is defined as $\overline{\psi} = \psi^\dagger\gamma^0$, where $\psi^\dagger$ is the Hermitian conjugate of ψ.

Chapter 2

Theory

What is the motivation for conducting the experiment, and how does it fit into what is already known? This chapter describes the fundamental particles organized into the Standard Model, particularly the role of the Higgs boson. The means of Higgs boson production at a proton–proton collider, the phenomenon of *lepton flavor violation* (LFV), and theoretical and experimental studies of LFV are then discussed.

2.1 The Standard Model

The *Standard Model* (SM) is a framework describing the known fundamental building blocks, or *particles*, of the Universe and their interactions.[1] The SM formalism was developed throughout the 1960s and has been subject to numerous experimental tests since its inception. It is a phenomenally successful theory, but yet, it is known not to account for all observed phenomena. Furthermore, the SM has a complex structure that is presently not understood.[2]

2.1.1 Phenomenology

The *phenomenology* of the SM refers to the entire set of particle constituents and their parent *fields*. The concept of a field comes from *quantum field theory* (QFT), which is the theoretical framework describing the particles of the SM. According to QFT, every particle is an excitation of a parent field, which itself is an object that has different values at every location in spacetime.[3] Fields carry energy and momentum (linear and angular) that is transferred by their associated particles. Every particle excitation of a particular field is

[1] The key property of particles is a lack of internal structure. That is, they are point-like objects.

[2] Contrast this with the periodic table of the elements. Its structure comes from the number of valence electrons in each of the constituent atoms, determined by quantum mechanics.

[3] The basic example of a field is the Earth's gravitational field. No matter where someone is on the Earth's surface, that person experiences a force directed toward its center. The fact that one experiences such a force everywhere is evidence of the existence of the gravitational field.

identical. For example, every electron in the universe is an excitation of the electron field, and all electrons are fundamentally identical and indistinguishable.

Particles are broadly classified by their *spin*.[4] *Fermions* are particles with half-integral spin,[5] and *bosons* are those with integral spin.[6] These associations are a consequence of the *spin–statistics theorem* that describes the behavior of systems of a number of indistinguishable bosons or fermions. The theorem states that the *wavefunctions* for systems of bosons are symmetric under interchange of any pair of bosons, while those for fermions are antisymmetric under exchange.[7] Thus, fermions obey the *Pauli Exclusion Principle*, while bosons have no such constraints.[8]

Fermions come in two varieties, *quarks* (q) and *leptons* (ℓ). There are six known species, or *flavors*, of quarks: *up* (u), *down* (d), *charm* (c), *strange* (s), *top* (t), and *bottom* (b). The quarks are typically represented in the following matrix:

$$\begin{pmatrix} u & c & t \\ d & s & b \end{pmatrix}$$

Each column of this matrix is a grouping called a *generation*.[9] The generations are ordered by the mass of their constituent quarks. The quarks in the top row of the matrix are called *up-type* quarks, and each has electric charge $Q = +\frac{2}{3}$. Those in the bottom row are called *down-type* quarks and have charge $Q = -\frac{1}{3}$.

Like quarks, there are six known lepton flavors arranged in three generations, but a crucial difference is that lepton flavors are either electrically charged or neutral. The charged leptons are the *electron* (e), *muon* (μ), and *tau lepton* (τ),[10] and each has electric charge $Q = -1$. Each charged lepton has an associated *neutrino* (ν) that is electrically neutral ($Q = 0$). The matrix for leptons may be written as follows:

$$\begin{pmatrix} \nu_e & \nu_\mu & \nu_\tau \\ e & \mu & \tau \end{pmatrix}$$

[4] *Spin* is an intrinsic property of fundamental particles that reflects the amount of angular momentum contained within each particle, although it is incorrect to state that particles are actually spinning since they are point-like and do not occupy any volume. Unlike *classical*, or macroscopic, spin, quantum spin occurs only in discrete quantities. For a fascinating account of the history of spin, see [12].

[5] e.g. $\frac{1}{2}$, $\frac{3}{2}$, $\frac{5}{2}$,...

[6] 0, 1, 2,...

[7] A wavefunction is a type of probability amplitude and provides information about the most probable location of a particle or particles in a system. A wavefunction $\psi(x_1, x_2)$ is symmetric if $\psi(x_1, x_2) = \psi(x_2, x_1)$ and antisymmetric if $\psi(x_1, x_2) = -\psi(x_2, x_1)$. An important consequence of antisymmetry is that the probability of finding two fermions with identical spin occupying the same location in space is zero. This is called the *Pauli Exclusion Principle*.

[8] Multiple bosons can be gathered at the same spatial location to form a state of matter called a *Bose-Einstein condensate*.

[9] The individual 2×1 columns are called *doublets*.

[10] The name "tau lepton" is a little awkward, and it is usually simply called the *tau* (and will be in this dissertation) but is also called the *tauon*.

The structure of the lepton matrix is similar to that of the quark matrix. The bottom row consists of the charged leptons, while the top row consists of the associated neutrinos.[11] Each column represents a generation, and the second and third generation lepton masses are increasingly larger.

There are two categories of known bosons, *vector* and *scalar* bosons.[12] The vector bosons are spin-1 particles that mediate interactions among the different fermions and consist of the *photon* (γ), *gluon* (g), and W and Z bosons. Photons are the particles responsible for the *electromagnetic* force and interact with all particles that have electric charge.[13] Gluons bind quarks together to make *hadrons* and mediate the *strong* force.[14] The W and Z bosons mediate a variety of interactions between and among quarks and leptons and are carriers of the *weak* force.[15] The photon, gluon, and Z boson are electrically neutral, while the W boson has unit electric charge. The photon and gluon are also massless particles, while the W and Z bosons are each quite massive, roughly 80 and 90 times heavier than the proton. Consequently, the W and Z bosons have relatively short lifetimes and mediate interactions only over short distances.[16]

At present, only one scalar boson has been found in Nature: the *Higgs boson*, discovered in 2012. It is electrically neutral and is about 125 times heavier than the proton. This boson is the observable excitation of the *Higgs field*, responsible for providing mass to the W and Z bosons and all fermions (except neutrinos[17]) and is discussed in detail in Section 2.1.1.3.

A consequence of special relativity is that all particles have associated *antiparticles* that have the opposite electric charge, but same mass, as their counterparts. Neutral particles, except neutrinos, are their own antiparticles.[18] Antiparticles are usually named by placing "anti" before the name of their counterpart and are represented with an overbar over the corresponding symbol (e.g. antineutrino, $\overline{\nu}$). An exception is that antielectrons are called *positrons*. The lepton and W boson symbols are typically written with the negative charge in superscript (e^-, μ^-, τ^-, and W^-), while their corresponding antiparticles are indicated with a positive charge (e^+, μ^+, τ^+, and W^+).

The SM particles and their properties are listed in Table 2.1.

[11] The row ordering will be explained in the discussion of electroweak interactions.

[12] The names vector and scalar refer to the types of mathematical objects used to represent each boson. Vector bosons are represented by four-vectors, while scalar bosons are represented by a single number.

[13] Virtually every force experienced on a macroscopic scale, with the exception of gravity, is an electromagnetic force.

[14] The most abundant hadrons are the protons and neutrons in atoms.

[15] This force is known mostly for its role in nuclear decay.

[16] This is a consequence of the *Heisenberg Energy–Time Uncertainty Principle*, $\Delta E \Delta t \geq \hbar/2$.

[17] Neutrinos are massless in the SM, but numerous experiments have shown that they have very tiny masses whose precise values are an active research area. For reasons to be discussed, the SM Higgs field is not believed to be responsible for these neutrino masses.

[18] There is research into whether or not neutrinos are their own antiparticles, but so far, the answer appears to be no.

Particle	Spin	Electric Charge	Mass [GeV]	Color Charge?	Baryon Number	Lepton Number
Leptons						
Electron (e)	$\frac{1}{2}$	-1	5.11×10^{-4}	No	0	$+1$
Muon (μ)	$\frac{1}{2}$	-1	0.106	No	0	$+1$
Tau (τ)	$\frac{1}{2}$	-1	1.78	No	0	$+1$
Electron neutrino (ν_e)	$\frac{1}{2}$	0	$< 2 \times 10^{-9}$	No	0	$+1$
Muon neutrino (ν_μ)	$\frac{1}{2}$	0	$< 2 \times 10^{-9}$	No	0	$+1$
Tau neutrino (ν_τ)	$\frac{1}{2}$	0	$< 2 \times 10^{-9}$	No	0	$+1$
Quarks						
Up (u)	$\frac{1}{2}$	$+\frac{2}{3}$	2.3×10^{-3}	Yes	$+\frac{1}{3}$	0
Charm (c)	$\frac{1}{2}$	$+\frac{2}{3}$	1.28	Yes	$+\frac{1}{3}$	0
Top (t)	$\frac{1}{2}$	$+\frac{2}{3}$	173	Yes	$+\frac{1}{3}$	0
Down (d)	$\frac{1}{2}$	$-\frac{1}{3}$	4.8×10^{-3}	Yes	$+\frac{1}{3}$	0
Strange (s)	$\frac{1}{2}$	$-\frac{1}{3}$	95×10^{-3}	Yes	$+\frac{1}{3}$	0
Bottom (b)	$\frac{1}{2}$	$-\frac{1}{3}$	4.18	Yes	$+\frac{1}{3}$	0
Bosons						
Photon (γ)	1	0	0	No	0	0
Gluon (g)	1	0	0	Yes	0	0
Z boson (Z)	1	0	91.2	No	0	0
W boson (W)	1	-1	80.4	No	0	0
Higgs boson (H)	0	0	125	No	0	0

Table 2.1: Standard Model particles and their properties [10]. Antiparticles have the opposite charge, baryon number, and lepton number to their particle counterparts.

2.1.1.1 Electroweak Interactions

Electroweak (EWK) *interactions* are mediated by photons, W, and Z bosons. The photon was the first of these particles to be described theoretically in the framework of *quantum electrodynamics* (QED), which is the quantum formulation of *classical electromagnetism* as encapsulated in *Maxwell's Equations*, which provides the theoretical framework for macroscopic electromagnetic phenomena. According to the classical theory, the electromagnetic force manifests through the interaction of electric and magnetic fields with electric charges and currents. These fields can have energy of arbitrary value. In QED, the electromagnetic field energy is not a continuous, but discrete, spectrum. The differences in the energy levels are carried by the *field quanta*, or photons.

Why does the field even have to be quantized at all? In fact, the existence of the

photon is essential for electrically charged fermions to interact. This is best demonstrated by constructing the QED *Lagrangian density*.[19] To begin, first consider the equation of motion for free (i.e. non-interacting) spin-$\frac{1}{2}$ particles, given by the Dirac equation. Its corresponding Lagrangian is given by Equation 2.1:

$$\mathcal{L}_{\text{Dirac}} = \overline{\psi}(i\gamma^\mu \partial_\mu - m)\psi \tag{2.1}$$

where ψ is the fermion *spinor*,[20] and m is the fermion mass. The first term encodes information about the kinetic energy of the system, while the second term accounts for the mass–energy. $\overline{\psi}$ is the antifermion spinor.[21]

One may notice that there is symmetry present in Equation 2.1. If ψ is scaled by an arbitrary complex number, $\psi \to e^{i\alpha}\psi$ implying that $\overline{\psi} \to e^{-i\alpha}\overline{\psi}$, then the overall Lagrangian is unchanged. This type of symmetry is called *U(1) symmetry*.[22]

So far, all of this is nice, but it reflects a fairly boring universe of non-interacting particles. To make things more interesting, suppose that the parameter α varies over spacetime. That is, $\alpha \to \alpha(x^\mu)$. In this case, making the transformations to ψ and $\overline{\psi}$ does *not* leave the Lagrangian invariant. It can be shown that Equation 2.1 becomes

$$\mathcal{L}_{\text{Dirac}} \to \mathcal{L}_{\text{Dirac}} - \overline{\psi}\gamma^\mu(\partial_\mu \alpha(x^\mu))\psi \tag{2.2}$$

The additional term in Equation 2.2 is utterly undesirable. Its dependence on $\alpha(x^\mu)$ indicates that the particle dynamics depends on their spacetime location. While one solution is positing that α is simply a constant, its spacetime dependence is essential for electric charge conservation, which will be discussed later.

The other option is to modify the Dirac Lagrangian to account for this extra term. Specifically, the derivative is modified to become the *covariant derivative*, $\partial_\mu \to D_\mu = \partial_\mu + ieA_\mu$, where e is the electric charge, and A_μ is the electromagnetic four-potential.

An interesting property of A_μ is that it may be modified arbitrarily such that $A_\mu \to A_\mu + \partial_\mu \alpha(x^\mu)$, where $\alpha(x^\mu)$ is a scalar quantity that is a function of spacetime. This modification

[19] The Lagrangian is a function containing the dynamical information regarding a system. It is used to calculate the *action*, $S = \int L dt$. *Hamilton's Principle*, or *the Principle of Least Action*, states that dynamical systems behave such that the action is minimized. Consequently, the Lagrangian is the function satisfying $\delta S = 0$ and can be used to calculate the equations of motion of a system using the *Euler–Lagrange equations*, which are derivable from Hamilton's Principle. Also, the Lagrangian density $\mathcal{L}$ is related to the Lagrangian by $L = \int \mathcal{L} d^3x$. However, in most QFT literature and in this dissertation, the terms Lagrangian and Lagrangian density will be used interchangeably.

[20] A spinor is a 4×1 mathematical object containing information regarding the spin, energy, and momentum of the fermion.

[21] The Dirac equation predicts both fermions and antifermions, so both come along for the ride in the Lagrangian. One cannot have a relativistic theory of spin-$\frac{1}{2}$ particles with just one particle—its antipartner is also required! As a historical note, this equation was the first to predict antimatter, which was discovered some years later.

[22] The *U(1) group* is the set of all complex numbers of unit norm. A *representation* of $U(1)$ is $e^{i\theta}$, where θ is any real number. Since all group elements $e^{i\theta}$ *commute*, the group is *Abelian*.

preserves the *electromagnetic field tensor* $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, which contains the physical fields. The ability to arbitrarily choose A_μ in this manner is called *gauge freedom*.

By rewriting the Dirac Lagrangian with the covariant derivative and choosing the gauge such that $A_\mu \rightarrow A_\mu - \frac{1}{e}\partial_\mu \alpha(x^\mu)$, the offending term in Equation 2.2 is canceled. Incorporating the energy stored in the electromagnetic fields into the Lagrangian yields the QED Lagrangian:

$$\begin{aligned} \mathcal{L}_{\text{QED}} &= \overline{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} \\ &= \overline{\psi}(i\gamma^\mu \partial_\mu - m)\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \overline{\psi}e\gamma^\mu A_\mu \psi \\ &= \mathcal{L}_{\text{Dirac}} + \mathcal{L}_{\text{EM}} + \mathcal{L}_{\text{Interaction}} \end{aligned} \tag{2.3}$$

By substituting the QED Lagrangian into the Euler–Lagrange equations for ψ, $\overline{\psi}$, and A_μ, it can be shown that modified forms of the Dirac equation, including interactions with an electromagnetic field, and Maxwell's equations for a current source having total charge e are obtained. It can further be shown [13] that by choosing the *Lorenz gauge*, $\partial_\mu A^\mu = 0$, the field A^μ corresponds to a massless plane wave with two modes of polarization transverse with respect to the direction of propagation—this is the photon.[23]

By examining the interaction Lagrangian in Equation 2.3, one may notice that the photon, antifermion, and fermion terms are coupled together with strength equal to the electronic charge. The presence of the γ^μ indicates that the photon transforms like a *vector* under a *Lorentz transformation*[24] and thus, is a spin-1 vector boson. In addition, fermions and antifermions are created and destroyed in pairs.

Earlier, it was asserted that the parameter α had to be a function of spacetime to preserve electric charge conservation. Electric charge is a quantity that is locally conserved—it cannot be destroyed in one location and instantly recreated someplace far away. By requiring that α be a function of spacetime, it can be shown that there is a continuity equation for a current with total charge e.[25]

Like the photon, the gluon, W, and Z bosons are *gauge bosons* whose existence and interactions come from additional gauge symmetries present in the SM Lagrangian.[26] The W and Z bosons mediate interactions far beyond those of the photon. The basic theory of weak interactions was developed by E. Fermi to explain nuclear beta decay. Fermi's theory of weak interactions consisted of a neutron spontaneously decaying into a proton, electron, and a neutrino, which was unknown experimentally at the time, but crucial on theoretical grounds to conserve momentum in the interaction. This theory worked well for low energy

[23] The fact that no mass terms of the form $\frac{1}{2}m_\gamma^2 A^\mu A_\mu$ appear in the Lagrangian also implies that gauge invariance requires the photon to be massless.

[24] This is a change in inertial frame of reference.

[25] *Noether's Theorem* states that conserved quantities in a system correspond to symmetries in the Lagrangian. Thus, local conservation of charge is a direct consequence of the $U(1)$ symmetry in QED.

[26] The QCD Lagrangian will be discussed in Section 2.1.1.3.

interactions, but was *nonrenormalizable*,[27] so it became evident that there must be some massive boson mediating the weak interaction to have a renormalizable theory.

This massive boson (now known as the W boson) is responsible for mediating charged weak interactions. It can transform charged leptons into the neutrinos in their generation and vice versa, and it can decay into charged leptons and antineutrinos (or charged antileptons and neutrinos). The W boson also transforms up-type quarks into down-types (and vice versa) and can likewise decay into up-type and antidown-type quarks (or vice versa). However, there is a complication with the quarks. When a W boson decays into an up-type and down-type quark, the down-type quark is not necessarily of the same generation as the up-type quark. Each down-type quark has an observational probability given by the *Cabbibo–Kobayashi–Maskawa* (CKM) matrix [14], [15].[28]

Many symmetries in electromagnetic and strong interactions are not present in weak interactions, particularly *charge* (C) symmetry, *parity* (P) symmetry, and *charge–parity* (CP) symmetry. The charge operation replaces particles with their antiparticles, and the parity operation inverts all spatial coordinates.[29] Parity and charge symmetry are violated in weak interactions based on the empirical observation of only left-handed neutrinos and right-handed antineutrinos,[30] while observations of neutral kaon decays by J.W. Cronin and V.L. Fitch in 1964 [16] conclusively established CP violation. Though the Dirac equation posits that antiparticles and particles behave identically, CP violation suggests that weak processes involving only particles do not behave in the same manner as those involving only antiparticles. In the SM formalism, CP violation in the quark sector is contained in the CKM matrix.

The theory of electroweak interactions formulated by S. Glashow, S. Weinberg, and A. Salam [17]–[19] accounts for these experimental observations by positing that the left- and right-handed components of fermion fields transform differently.[31] Specifically, the left-handed field components can be arranged in a doublet that transforms under *SU(2)*,[32] while the right-handed field components are written as singlets (that is, they don't transform under

[27] One of the problems that plagued theorists for decades was that NLO and higher calculations led to infinities in the predictions, which are nonsensical. These infinities are the result of ignorance of physical structure at very short length scales and are treated by *renormalizing* the quantum fields so that the infinities are "absorbed" into the calculation, with measurable physical parameters used in place. Not every theory can be treated in this way, and those that can are renormalizable theories. Far from being merely a mathematical trick, renormalization is a consequence of the variation in the laws of physics over different length and energy scales.

[28] In the parlance of linear algebra, the weak interactions participate through weak eigenstates, which are not the same as the experimentally observable mass eigenstates. The mass and weak eigenstates are rotated with respect to one another, and the unitary CKM matrix describes the angles between the different bases.

[29] Mathematically, if ψ and $\overline{\psi}$ are particle and antiparticle states, $\hat{C}\psi = \overline{\psi}$, and $\hat{P}\psi(x, y, z) = \psi(-x, -y, -z)$.

[30] The handedness, or *helicity*, of the neutrino refers to the direction of its spin angular momentum.

[31] The handedness of a field is known as its *chirality*.

[32] $SU(2)$ is the group of 2×2 unitary matrices with determinant = 1. The classic representation of $SU(2)$ consists of the Pauli spin matrices and the 2×2 identity matrix.

$SU(2)$). For example:

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, e_R, \begin{pmatrix} u \\ d \end{pmatrix}_L, u_R, d_R$$

In other words, since charged weak currents transform (left-handed) electrons into neutrinos, such an interaction is akin to an $SU(2)$ rotation in "weak-space". Parity violation occurs because the right-handed fields do not transform in this manner.

To illustrate, consider the following Lagrangian consisting of only electrons and neutrinos (but can be extended to include the other fermions):

$$\mathcal{L}_{\text{Weak}} = \overline{e}(i\gamma^\mu\partial_\mu - m_e)e + \overline{\nu_e}(i\gamma^\mu\partial_\mu - m_{\nu_e})\nu_e \tag{2.4}$$

where e, $\overline{e}$, ν_e, and $\overline{\nu_e}$ refer to the electron, position, neutrino, and antineutrino Dirac spinors. Equation 2.4 can be written into a more compact form using the doublet notation previously mentioned and separating the right- and left-handed field components:[33]

$$\mathcal{L}_{\text{Weak}} = \overline{\psi}_L(i\gamma^\mu\partial_\mu - \mathbf{M})\psi_L + \overline{e}_R(i\gamma^\mu\partial_\mu - m_e)e_R \tag{2.5}$$

where

$$\psi_L = \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \overline{\psi}_L = \begin{pmatrix} \overline{\nu_e} & \overline{e} \end{pmatrix}_L,$$

and

$$\mathbf{M} = \begin{pmatrix} m_\nu & 0 \\ 0 & m_e \end{pmatrix}$$

is the mass matrix.

Similarly to QED, the doublet ψ_L can be transformed according to $\psi_L \rightarrow e^{\frac{ig}{2}\boldsymbol{\alpha}(x^\mu)\cdot\boldsymbol{\sigma}}\psi_L$, where g is a "weak charge" analogous to electric charge,[34] $\boldsymbol{\alpha}(x^\mu)$ is a "vector" of phase angles in $SU(2)$ space with each "dimension" covered by the appropriate Pauli spin matrix in the vector $\boldsymbol{\sigma}$.[35] To accommodate the additional terms in the Lagrangian, the covariant derivative for left-handed spinors is formed from three gauge fields W_μ^a ($a = 1, 2, 3$), one for each phase angle in $\boldsymbol{\alpha}(x^\mu)$. That is, $\partial_\mu \rightarrow D_\mu = \partial_\mu - \frac{ig}{2}\mathbf{W}_\mu \cdot \boldsymbol{\sigma}$, where $\mathbf{W}_\mu = (W_\mu^1, W_\mu^2, W_\mu^3)$. The gauge transformation for $\mathbf{W}_\mu$ is similar to A_μ: $\mathbf{W}_\mu \rightarrow \mathbf{W}_\mu + \partial_\mu\boldsymbol{\alpha}(x^\mu) + g\mathbf{W}_\mu \times \boldsymbol{\alpha}(x^\mu)$. The additional cross product term occurs because the Pauli matrices do not commute[36] and indicates that the different components of $\mathbf{W}_\mu$ interact with each other.

[33] $e = e_L + e_R$. There are no right-handed neutrinos nor left-handed antineutrinos in the SM.

[34] The phase factor in QED also could have been written as $e^{ie\alpha(x^\mu)}$, with the gauge transformation of A^μ adjusted accordingly.

[35] $\boldsymbol{\alpha}(x^\mu) \cdot \boldsymbol{\sigma} = \alpha_1(x^\mu)\sigma_1 + \alpha_2(x^\mu)\sigma_2 + \alpha_3(x^\mu)\sigma_3$. Since any unitary 2×2 matrix can be specified by a linear combination of the identity matrix and the Pauli matrices, the phase shift indicated is a completely general 2×2 unitary matrix (a power series expansion of the exponential reveals the identity matrix). The basis matrices (Pauli matrices in this case) are called the *generators* of a transformation group.

[36] The commutation relation for the Pauli matrices is $[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k$, where ϵ_{ijk} is the Levi–Civita symbol. Incidentally, the factor of 2 in the right-hand side of the commutator is the reason for including the factor of $\frac{1}{2}$ with the weak charge.

Rewriting the Lagrangian in Equation 2.5 to include the interaction term yields

$$\begin{aligned}\mathcal{L}_{\text{Weak}} &\to \overline{\psi}_L(i\gamma^\mu\partial_\mu - \mathbf{M})\psi_L + \overline{e}_R(i\gamma^\mu\partial_\mu - m_e)e_R + \frac{g}{2}\overline{\psi}_L\mathbf{W}_\mu\cdot\boldsymbol{\sigma}\psi_L \\ &= \mathcal{L}_{\text{Weak}} + \mathcal{L}_{\text{Interaction}}\end{aligned} \tag{2.6}$$

By defining $W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2)$ and performing the matrix multiplication in the interaction Lagrangian, it can be shown [20] that the interaction Lagrangian becomes

$$\mathcal{L}_{\text{Interaction}} = \frac{g}{\sqrt{2}}\Big(\overline{\nu_e}\gamma^\mu W_\mu^- e_L + \overline{e}_L\gamma^\mu W_\mu^+\nu_e\Big) + \frac{g}{2}\Big(\overline{\nu_e}\gamma^\mu W_\mu^3\nu_e - \overline{e}_L\gamma^\mu W_\mu^3 e_L\Big) \tag{2.7}$$

There are two groups of terms in Equation 2.7, each corresponding to a different type of interaction. The first group contains *charged-current* interactions (that is, conversion of electrons into neutrinos and vice versa), while the second group *predicts* that there should be *neutral-current* interactions, whereby an electron remains an electron, and a neutrino remains a neutrino. The electrically charged spin-1 electroweak fields $W^\pm$ have the W bosons for field quanta, while the neutral current interactions suggest that there should be an electrically neutral spin-1 weak boson.

At this point, it is reasonable to wonder whether or not this weak theory is compatible with electromagnetic interactions in its current form. To accommodate both theories, the Lagrangian in Equation 2.7 would have to be invariant under simultaneous $SU(2)_L$ and $U(1)_{\text{EM}} = e^{iQ\alpha(x^\mu)}$ transformations, where Q is the electric charge of the particle.[37] However, the Lagrangian is *not* invariant under these transformations. Therefore, $U(1)_{\text{EM}}$ is not a symmetry that is compatible with both electromagnetic and weak interactions.

There is a $U(1)$ transformation that is compatible with both electromagnetic and weak interactions. The Lagrangian in Equation 2.7 is invariant under the additional transformation $f \to e^{\frac{ig'Y}{2}\chi(x^\mu)}f$, where f is the Dirac spinor for any fermion, g' is another weak charge, $\chi(x^\mu)$ is another phase angle (to distinguish it from $\alpha(x^\mu)$), and Y is the *hypercharge* associated with each particle, calculated according to the *Gell-Mann–Nishijima formula*, $Y = 2(Q - I^3)$, where Q is the particle electric charge and I^3 is the *weak isospin quantum number*. I^3 is 0 for right-handed fields, $+\frac{1}{2}$ for left-handed neutrino and up-type quark fields, and $-\frac{1}{2}$ for left-handed charged lepton and down-type quark fields. This symmetry is called $U(1)_Y$ symmetry, and the whole electroweak Lagrangian is invariant under $SU(2)_L \times U(1)_Y$ transformations.

The covariant derivative is modified to account for $\chi(x^\mu)$ by introducing a spin-1 gauge field B_μ, becoming $\partial_\mu \to D_\mu = \partial_\mu - \frac{ig}{2}\mathbf{W}_\mu\cdot\boldsymbol{\sigma} - \frac{ig'}{2}YB_\mu$. The Lagrangian in Equation 2.6 becomes

$$\begin{aligned}\mathcal{L}_{\text{EWK}} &= \overline{\psi}_L(i\gamma^\mu\partial_\mu - \mathbf{M})\psi_L + \overline{e}_R(i\gamma^\mu\partial_\mu - m_e)e_R + \overline{\psi}_L(\frac{g}{2}\mathbf{W}_\mu\cdot\boldsymbol{\sigma} + \frac{g'}{2}YB_\mu)\psi_L \\ &\quad + \overline{e}_R\frac{g'}{2}YB_\mu e_R \\ &= \mathcal{L}_{\text{Weak}} + \mathcal{L}_{\text{Interaction}}\end{aligned} \tag{2.8}$$

[37] More specifically, the Lagrangian would have to be invariant under $SU(2)_L \times U(1)_{\text{EM}}$.

The interaction term in Equation 2.8 may be written similarly to Equation 2.7 as

$$\mathcal{L}_{\text{Interaction}} = \frac{1}{2}\Big\{\overline{\nu_e}\gamma^\mu(g'YB_\mu + gW^3_\mu)\nu_e + \overline{e}_L\gamma^\mu(g'YB^\mu - gW^3_\mu)e_L\Big\} + \overline{e}_R\frac{g'}{2}YB_\mu e_R \qquad (2.9)$$
$$+ \frac{g}{\sqrt{2}}\Big(\overline{\nu_e}\gamma^\mu W^-_\mu e_L + \overline{e}_L\gamma^\mu W^+_\mu \nu_e\Big)$$

Examination of the first term of Equation 2.9 indicates that there is a gauge field interacting with neutrinos that cannot be the electromagnetic field since neutrinos are electrically neutral. This neutral electroweak gauge field is defined as $Z_\mu = \frac{g'YB_\mu + gW^3_\mu}{\sqrt{g^2+g'^2Y^2}}$ and has the Z boson for its quanta. The electromagnetic field is orthogonal by necessity to the neutral electroweak field and is defined as $A_\mu = \frac{gB_\mu - g'YW^3_\mu}{\sqrt{g^2+g'^2Y^2}}$. Writing Equation 2.9 in terms of Z_μ and A_μ and replacing the hypercharge by its numerical value for the given fermion spinor yields

$$\mathcal{L}_{\text{Interaction}} = \frac{1}{2\sqrt{g^2+g'^2}}\Big\{(g^2+g'^2)\overline{\nu_e}\gamma^\mu Z_\mu\nu_e - 2gg'\overline{e}\gamma^\mu A_\mu e + (g'^2-g^2)\overline{e}_L\gamma^\mu Z_\mu e_L \qquad (2.10)$$
$$+ 2g'^2\overline{e}_R\gamma^\mu Z_\mu e_R\Big\} + \frac{g}{\sqrt{2}}\Big(\overline{\nu_e}\gamma^\mu W^-_\mu e_L + \overline{e}_L\gamma^\mu W^+_\mu\nu_e\Big)$$

Each term in Equation 2.10 describes the coupling between a particular gauge field and fermions. For example, the first term indicates that the Z boson couples to neutrinos with strength $\frac{1}{2}\sqrt{g^2+g'^2}$. Looking at the term with A_μ, the relationship between the electronic charge and the weak charges can be written as $e = \frac{gg'}{\sqrt{g^2+g'^2}}$. Since Z_μ and A_μ are a rotation of B_μ and W^3_μ, a *weak mixing angle* (θ_W) may be defined such that $\cos\theta_W = \frac{g}{\sqrt{g^2+g'^2}}$. Notably, Equation 2.10 suggests that the W and Z bosons couple equally to the corresponding fermions of different generations (e.g. all charged leptons). This property is called *universality*.

While the theory accounts for the electroweak gauge bosons, there are two issues to be resolved. The first concerns the mass matrix in Equation 2.8. The term $\overline{\psi}_L\mathbf{M}\psi_L$ cannot remain invariant under the $SU(2)_L$ transformation unless $\mathbf{M}$ is proportional to the identity matrix,[38] which requires the electron and neutrino to have equal mass. This is far from experimental fact. Compounding this is that only fields with well-defined chirality are massless fields. Thus, all fermion fields in this theory have to be massless (killing the mass matrix term). The second issue is that all gauge fields have to be massless to maintain gauge-invariance. Yet, the W and Z bosons are very massive, while the photon remains massless. How can this be?

The answer is that the symmetry of $SU(2)_L \times U(1)_Y$ has to be broken to eliminate the requirement for gauge invariance, leaving only the symmetry $U(1)_{\text{EM}}$. While many ideas were put forward to explain this *electroweak symmetry breaking*, the idea with the strongest experimental support is the *Higgs mechanism*, described in Section 2.1.1.3.

[38] This is because the only matrix that commutes with all three Pauli matrices is the identity matrix.

2.1.1.2 Quantum Chromodynamics

Quantum Chromodynamics (QCD) is the framework that describes the strong interactions between quarks and gluons. An understanding of QCD is critical to modeling physics processes at proton–proton colliders since all physics starts from the collision of two protons. The basic premise is that the nuclei of matter are composed of quarks, bound together by gluons. Furthermore, each quark possesses a quantity called *color charge*, of which there are three species: *red*, *green*, and *blue*. Antiquarks possess anticolor charge: *antired*, *antigreen*, and *antiblue*. Gluons possess both a color and anticolor charge, and there are eight different types. Like electric charge, color charge is a conserved quantity.

QCD has unique properties that give rise to interesting behavior. One property is called *confinement*, under which all observable states are *color singlets*.[39] Examples of color singlets include *baryons*, which consist of three quarks, each with a different color, and *mesons*, which are quark–antiquark states, each having corresponding color and anticolor.[40] A consequence of confinement is that bare quarks and gluons are unobservable. If one tries to pull apart the quarks and antiquarks in a color singlet, the energy in the gluon fields increases until there is enough to produce a quark–antiquark pair out of the vacuum.[41] This breaking occurs at distances of $\sim$1 fm, corresponding to an energy of $\sim$200 MeV.

The second property is called *asymptotic freedom*, under which quarks and gluons behave more like free particles as the interaction energy increases. That is, they are capable of traveling farther distances apart before the gluon fields generate quark–antiquark pairs.

To understand asymptotic freedom and confinement, it is important to understand the *strong coupling constant* (α_S), which reflects the strength of interactions between sets of particles and plays a critical role in determining the rates of physical processes (e.g. particle decay rates, production rates, or scattering rates). The larger the coupling constant, the stronger the interaction. Figure 2.1 shows the value of α_S as a function of momentum (or energy) transferred through the binding gluon fields. What is unusual about α_S is that it is large at low energy interactions and gets smaller with increased energy transfer.[42] If one pulls apart two quarks gently, they cannot be pulled apart very far without generating quark–antiquark pairs. However, if one violently yanks at the quarks, they can be pulled apart further before pair production occurs. The decrease of the coupling constant with energy directly leads to asymptotic freedom. Another consequence of low α_S is that quark–antiquark pairs are constantly blinking in and out of existence all throughout space. These types of quarks are called *sea quarks*.

The running of α_S also results in complications for any sort of calculation involving quarks and gluons. These calculations rely on the numerical values of *Feynman diagrams*

[39] This is a rather technical name meaning all observable states have to be *colorless*.

[40] The LHCb Collaboration has recently observed *pentaquark* states consisting of four quarks and one antiquark [21].

[41] This is analogous to breaking a bar magnet in two. One does not have a north pole in one hand and a south pole in the other, but rather two separate magnets, each with its own north and south poles.

[42] Readers may wonder why α_S is called a constant since its value changes. The name is due to historical reasons, and this dependence on energy is called the *running of the coupling*.

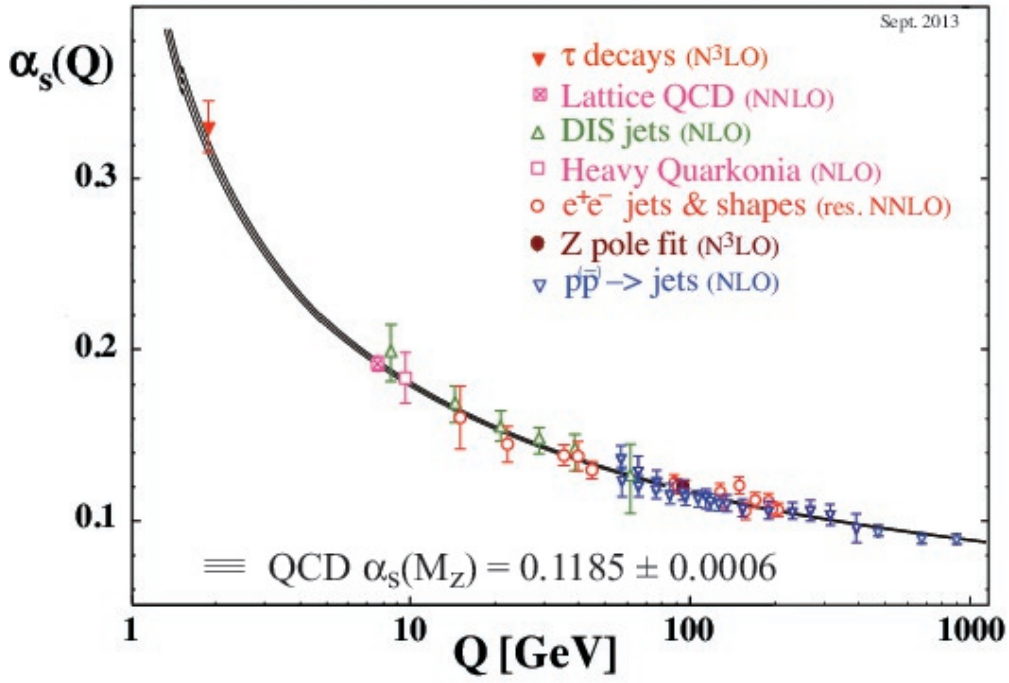


Figure 2.1: The Strong Coupling α_S as a function of momentum transfer [10].

that illustrate processes on the particle level. An infinite number of Feynman diagrams are needed to calculate the rates for physical processes to exact precision. Fortunately, not all Feynman diagrams are equal—they each depend on various powers of the coupling constant. The diagrams in which the coupling constant has the lowest power are called *Leading Order* (LO) diagrams. Those with the second and third lowest powers are respectively called *Next-to-Leading Order* (NLO) and *Next-to-Next-to-Leading Order* (NNLO). If the coupling constant is small, then the LO diagrams are sufficient for an accurate determination of a physical process since the contributions by higher order diagrams contribute marginally. The precision can be increased by performing NLO and NNLO computations. Calculations in which higher order computations contribute marginally are called *perturbative* calculations.[43] The increase of α_S with decreasing energy means that perturbative calculations are not possible for low energy QCD processes. Understanding this regime quantitatively is an active research area.

Protons are composed of three *valence quarks*:[44] two u-quarks and a d-quark. In a proton–proton collision, two of the quarks or gluons collide with high energy transfer, represented in Figure 2.2. These collisions can be calculated perturbatively. The colliding quarks

[43]Think of summing an infinite series, say $\sum_{n=0}^{\infty} \frac{1}{2^n}$. The answer is 2, but summing the first ten terms yields 1.998.

[44]The valence quarks are responsible for the properties of a hadron, but each hadron also contains some number of sea quarks and antiquarks.

separate, radiating gluons that form quark–antiquark pairs. As the quarks and antiquarks lose energy through gluon radiation, they bind together to form hadrons in a process called *hadronization*.[45] Hadronization is a non-perturbative process, but it may be simulated using models such as the *Lund String Model*. The resulting hadrons form in a column around the original quark, manifesting in the detector as an object called a *jet*.

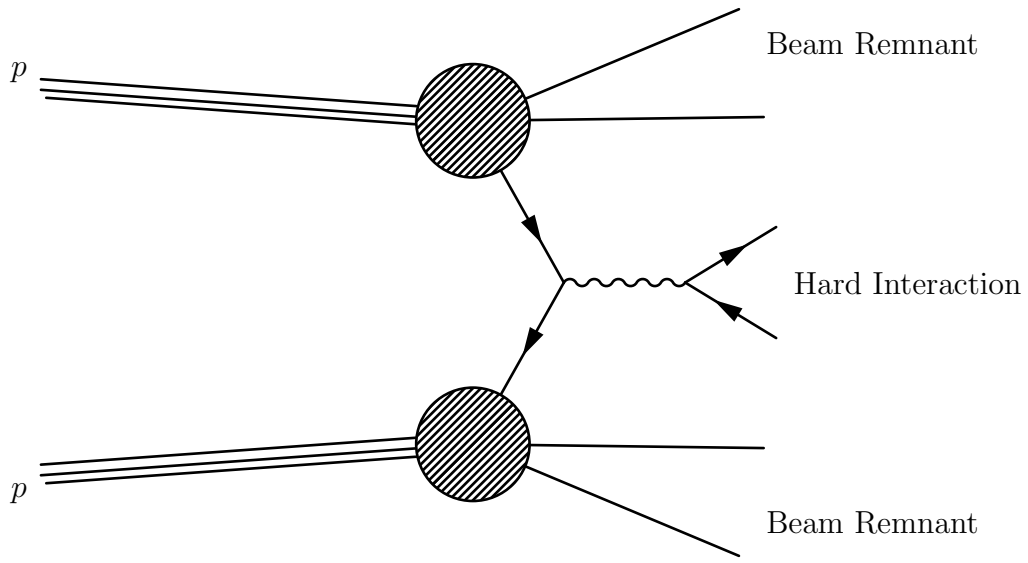


Figure 2.2: An example of a proton–proton collision.

Understanding a proton–proton collision also requires having knowledge of the *kinematics*[46] of the proton constituents before the collision. Although the interaction of these constituents is non-perturbative and cannot be calculated, the proton structure has been thoroughly studied in atomic and nuclear experiments. This empirical knowledge is used to model the proton structure with a *parton distribution function* (PDF) that describes the probability that a quark or gluon within the proton carries a particular fraction of the proton momentum. PDFs are calculated at different energy scales to account for the running of α_S. Two examples of PDFs are shown in Figure 2.3. The non-perturbative calculations based on PDFs are combined with the perturbative calculations of the colliding quarks in a process called *factorization*, which models the fragmentation of the proton during the collision.

[45]Because the top quark is so massive, it decays before it can hadronize. Thus, top quarks actually exist as "bare" quarks because they are ephemeral.

[46]The kinematics of a particle or system refer to the description of motion and are usually characterized by energy, momentum, and mass.

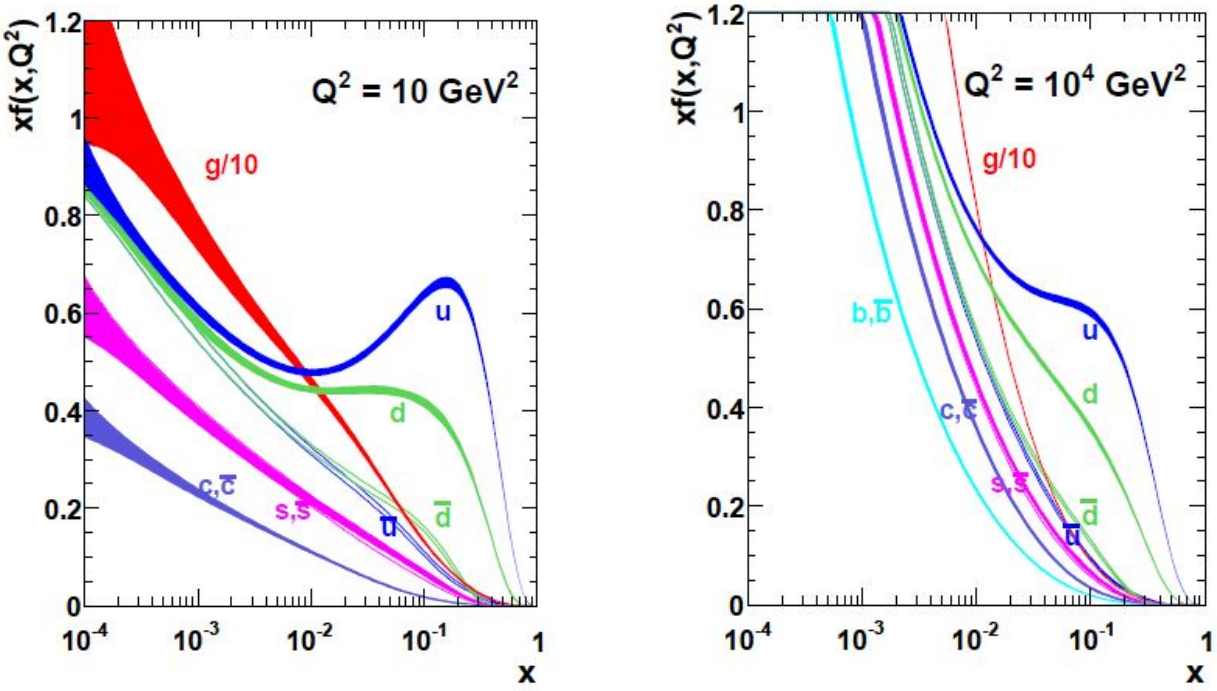


Figure 2.3: Parton Distribution Functions for $Q^2 = 10$ GeV^2 and $Q^2 = 10^4$ GeV^2 [22].

As the dynamics of electroweak processes emerge from gauge considerations of $SU(2)_L \times U(1)_Y$, the dynamics of QCD emerge from gauge considerations of $SU(3)_C$.[47] The subscript C refers to color charge. Quarks of a particular species can be arranged in a color triplet:

$$\psi_C = \begin{pmatrix} \psi_R \\ \psi_G \\ \psi_B \end{pmatrix}$$

where R, G, and B refer to the three colors red, green, and blue. In the Dirac Lagrangian (Equation 2.1), the triplet ψ_C transforms according to $\psi_C \rightarrow e^{\frac{ig_S}{2}\boldsymbol{\kappa}(x^\mu)\cdot\boldsymbol{\lambda}}\psi_C$, where $g_S = \sqrt{4\pi\alpha_S}$ and $\boldsymbol{\kappa}(x^\mu)\cdot\boldsymbol{\lambda} = \kappa_1(x^\mu)\lambda_1 + \kappa_2(x^\mu)\lambda_2 + \cdots + \kappa_8(x^\mu)\lambda_8$. Each λ_i corresponds to one of the eight generators of $SU(3)$[48] and has an associated phase angle $\kappa_i(x^\mu)$. The covariant derivative is written as $D_\mu = \partial_\mu - \frac{ig_S}{2}\mathbf{G}_\mu\cdot\boldsymbol{\lambda}$, where the vector $\mathbf{G}_\mu$ contains eight massless spin-1 gauge fields, each representing a different gluon. Six of the gluons mediate the transfer of color from one quark to another, while the remaining two gluons mediate interactions between quarks and/or antiquarks of the same color. To maintain symmetry under $SU(3)_C$ transformations, the gluon fields transform as $G^i_\mu \rightarrow G^i_\mu + \partial_\mu\kappa_i(x^\mu) + g_S f_{ijk}\kappa_j(x^\mu)G^k_\mu$, where f_{ijk} are the $SU(3)_C$ *structure constants* and have different values for combinations of $i, j, k =$

[47] The $SU(3)$ group is the set of all 3×3 unitary matrices with determinant $= +1$.

[48] The generators are typically represented with the *Gell-Mann matrices*.

1 . . . 8. The second term in the transformation indicates that the eight different gluon fields interact with each other, with strength dependent on the color combinations.

2.1.1.3 Higgs Mechanism

By considering $SU(3)_C \times SU(2)_L \times U(1)_Y$ symmetry, it was shown in Sections 2.1.1.1 and 2.1.1.2 that there are twelve massless bosons mediating strong and electroweak interactions among massless fermions. In reality, the weak bosons and fermions are massive, but the photon and gluons are massless. The $SU(3)_C \times SU(2)_L \times U(1)_Y$ symmetry must therefore be broken, resulting in the observed $SU(3)_C \times U(1)_{\mathrm{EM}}$. A simple mechanism for breaking the symmetry consists of another field (or fields) permeating spacetime that interacts with the $\mathbf{W}_\mu$ and B_μ fields, but does not interact with the resulting photon nor gluons. Since spacetime is invariant under rotations, such a field cannot contribute any angular momentum and is consequently spin-0.

The *Higgs mechanism* posits that there is a field $\mathbf{\Phi}$ consisting of a doublet that transforms under $SU(2)_L \times U(1)_Y$:

$$\mathbf{\Phi} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1^+(x^\mu) + i\phi_2^+(x^\mu) \\ \phi_3^0(x^\mu) + i\phi_4^0(x^\mu) \end{pmatrix}$$

The field has four degrees of freedom, ϕ_1^+, ϕ_2^+, ϕ_3^0, and ϕ_4^0, and is constructed such that the first row has weak isospin $I^3 = +\frac{1}{2}$, while the second row has $I^3 = -\frac{1}{2}$. Requiring that the hypercharge of $\mathbf{\Phi}$ be $+1$ to avoid coupling to photons results in the top row having an electric charge of $+1$, while the bottom row is electrically neutral.

The Lagrangian of this field is assumed to have the form shown in Equation 2.11:

$$\mathcal{L}_{\mathrm{Higgs}} = (D_\mu \mathbf{\Phi})^\dagger (D^\mu \mathbf{\Phi}) + \mu^2 \mathbf{\Phi}^\dagger \mathbf{\Phi} - \lambda (\mathbf{\Phi}^\dagger \mathbf{\Phi})^2 \tag{2.11}$$

where the first term refers to the kinetic behavior of the field (with covariant derivative $D_\mu = \partial_\mu - \frac{ig}{2}\mathbf{W}_\mu \cdot \boldsymbol{\sigma} - \frac{ig'}{2} Y B_\mu$) and the term $V(\mathbf{\Phi}^\dagger \mathbf{\Phi}) = -\mu^2 \mathbf{\Phi}^\dagger \mathbf{\Phi} + \lambda (\mathbf{\Phi}^\dagger \mathbf{\Phi})^2$ is the *potential* of the field (with $\mu^2 > 0$ and $\lambda > 0$), illustrated in Figure 2.4:

The most significant feature of the potential is that its minimum value occurs at $\mathbf{\Phi}^\dagger \mathbf{\Phi} = \frac{\mu^2}{2\lambda}$, rather than $\mathbf{\Phi}^\dagger \mathbf{\Phi} = 0$. The expected value of the field $\mathbf{\Phi}$ occurs at the minimum of the potential and may be written

$$\langle \mathbf{\Phi} \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}$$

where $v = \frac{\mu}{\sqrt{\lambda}}$ is the *vacuum expectation value* (vev). The Higgs field vev is computed using measurements of the Fermi constant to be roughly 246 GeV.[49]

Inspection of Figure 2.4 shows that the minimum value of the potential is not unique, but rather a continuous circle of values. Since the field can only pick one specific minimum

[49] Some authors absorb the factor of $\frac{1}{\sqrt{2}}$ into the calculation of the vev, reporting it as 174 GeV.

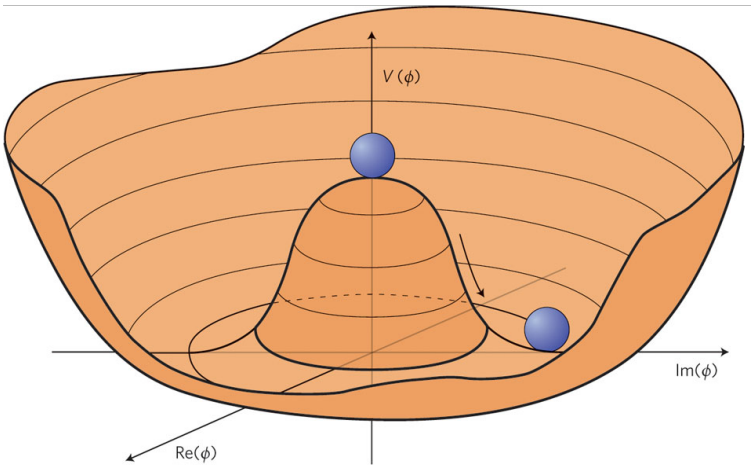


Figure 2.4: The Higgs Potential [23].

in the circle, any choice breaks the $SU(2)_L \times U(1)_Y$ symmetry.[50] There is gauge freedom to choose $\langle \mathbf{\Phi} \rangle$. For convenience, $\phi_1^{+,\min} = \phi_2^{+,\min} = \phi_4^{0,\min} = 0$ and $\phi_3^{0,\min} = v$ are chosen.

There are two ways the field can move about $\langle \mathbf{\Phi} \rangle$. It can oscillate radially, or it can move along the circle of minima. When written about the minimum, $\mathbf{\Phi}$ becomes

$$\mathbf{\Phi}(x^\mu) = e^{i\boldsymbol{\zeta}(x^\mu)\cdot\boldsymbol{\sigma}} \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x^\mu) \end{pmatrix} \tag{2.12}$$

where $H(x^\mu)$ is an excitation along the radial direction (or along ϕ_3^0) and $\boldsymbol{\zeta}(x^\mu)$ are excitations along the minima of the potential.[51] As predicted by the *Nambu–Goldstone theorem*, gauge invariance of the excitations $\boldsymbol{\zeta}(x^\mu)$ manifests in the Lagrangian as three massless scalar bosons called *Goldstone bosons*, each associated to the three non-radial degrees of freedom of the field $\mathbf{\Phi}$. The gauge may be chosen such that the Goldstone bosons are absorbed into the W and Z bosons, providing them with their longitudinal polarizations.[52]

Expanding the potential radially using the doublet in Equation 2.12 yields

$$V(H) = -\frac{1}{4}\mu^2 v^2 + \frac{1}{2}(2\mu^2)H^2 + \lambda v H^3 + \frac{1}{4}\lambda H^4 \tag{2.13}$$

The second term in Equation 2.13 suggests that the excitation H has a mass of $\sqrt{2\mu^2}$, or $\sqrt{2\lambda}v$. In fact, this excitation is the *Higgs boson*. The third and fourth terms in Equation

[50] This may seem suspect, but such symmetry breaking occurs frequently in Nature. A simple home experiment is to take a flexible rod and compress it on both ends. Eventually it will buckle and bend in a random direction. This spontaneously breaks its rotational symmetry.

[51] The doublet transforms under $SU(2)_L \times U(1)_Y$, so there are three such excitations, corresponding to the other fields whose minima were chosen to be zero.

[52] Massless particles can only have two polarization modes, each transverse to its momentum. Massive particles have an additional polarization mode parallel (or longitudinal) to their momentum.

2.13 suggest cubic or quartic Higgs boson couplings. That is, a single Higgs boson could decay into a pair or trio of Higgs bosons.

The masses of the W and Z bosons and their couplings to the Higgs boson appear when expanding the Lagrangian in Equation 2.11 about the Higgs field minimum. It can be shown that the masses of the W and Z bosons are $m_W = \frac{gv}{2}$ and $m_Z = \frac{g}{2}\sqrt{g^2 + g'^2}$.[53] The weak mixing angle is related to the masses of the W and Z by $m_W = m_Z \cos\theta_W$. Also indicated are the coupling strengths of the Higgs boson to pairs of weak bosons. These strengths are given by $g_{HVV} = \frac{2m_V^2}{v}$, where $V = W$, Z [10].

The Higgs mechanism also provides mass to the fermions through inclusion of the following terms to the Lagrangian:

$$\Delta\mathcal{L}_f = -y_f \overline{\boldsymbol{\psi}}_L \cdot \boldsymbol{\Phi} f_R + h.c. \tag{2.14}$$

where $\overline{\boldsymbol{\psi}}_L$ is the left-handed $SU(2)_L$ doublet associated with the fermion species f, f_R is the right-handed fermion SU(2) singlet, y_f is the *Yukawa coupling* of the fermion to the Higgs field, and *h.c.* is the Hermitian conjugate of the term. In the absence of electroweak symmetry breaking, $\boldsymbol{\Phi} = 0$ and the fermions are massless. Once the symmetry is broken, expansion about the minimum point of the Higgs field yields fermion mass terms, each having mass $m_f = \frac{y_f v}{\sqrt{2}}$.[54] The Yukawa couplings indicate the likelihood of the Higgs boson to decay to fermion–antifermion pairs (or vice versa: for fermion–antifermion pairs to produce a Higgs boson). They are directly proportional to the fermion mass, suggesting that the Higgs boson prefers to interact with heavier particles.

The Higgs boson has many decay modes due to its couplings to all charged fermions and the massive vector bosons. It can decay indirectly into pairs of photons or gluons through intermediate fermion loops. The fraction of time the Higgs boson decays to a particular set of particles (called the *branching ratio* (Br)) depends on the Higgs boson mass, a free parameter in the SM. The largest branching ratios for different Higgs boson decay modes are shown in Figure 2.5.

For the 125 GeV Higgs boson, the dominant decay mode is to a pair of b-quarks. This process is difficult to measure due to the large background of SM processes that also produce a pair of b-quarks. The Higgs boson discovery was made in the $H \to \gamma\gamma$ channel. It has also been experimentally confirmed in the $H \to WW \to \ell\nu\ell\nu$, $H \to ZZ \to \ell\ell\ell\ell$, and $H \to \tau\tau$ channels.

[53] These terms appear in the expanded Lagrangian as $\frac{1}{2}m_W^2 W_\mu^+ W^{+,\mu}$, $\frac{1}{2}m_W^2 W_\mu^- W^{-,\mu}$, and $\frac{1}{2}m_Z^2 Z_\mu Z^\mu$.

[54] The careful reader may notice that the choice of $\boldsymbol{\Phi}$ seems to suggest that the Higgs boson cannot couple to up-type quarks or neutrinos. For these fermions, the *charge conjugate Higgs field* $\boldsymbol{\Phi}^c = \hat{C}\boldsymbol{\Phi} = i\sigma^2\boldsymbol{\Phi}^* = \frac{1}{\sqrt{2}}\begin{pmatrix} v+h \\ 0 \end{pmatrix}$ is used to provide the mass. Note that the observational absence of right-handed neutrinos suggests that the Higgs field cannot be responsible for the neutrino mass since there is no singlet ν_R for the Higgs field to couple. More details to come.

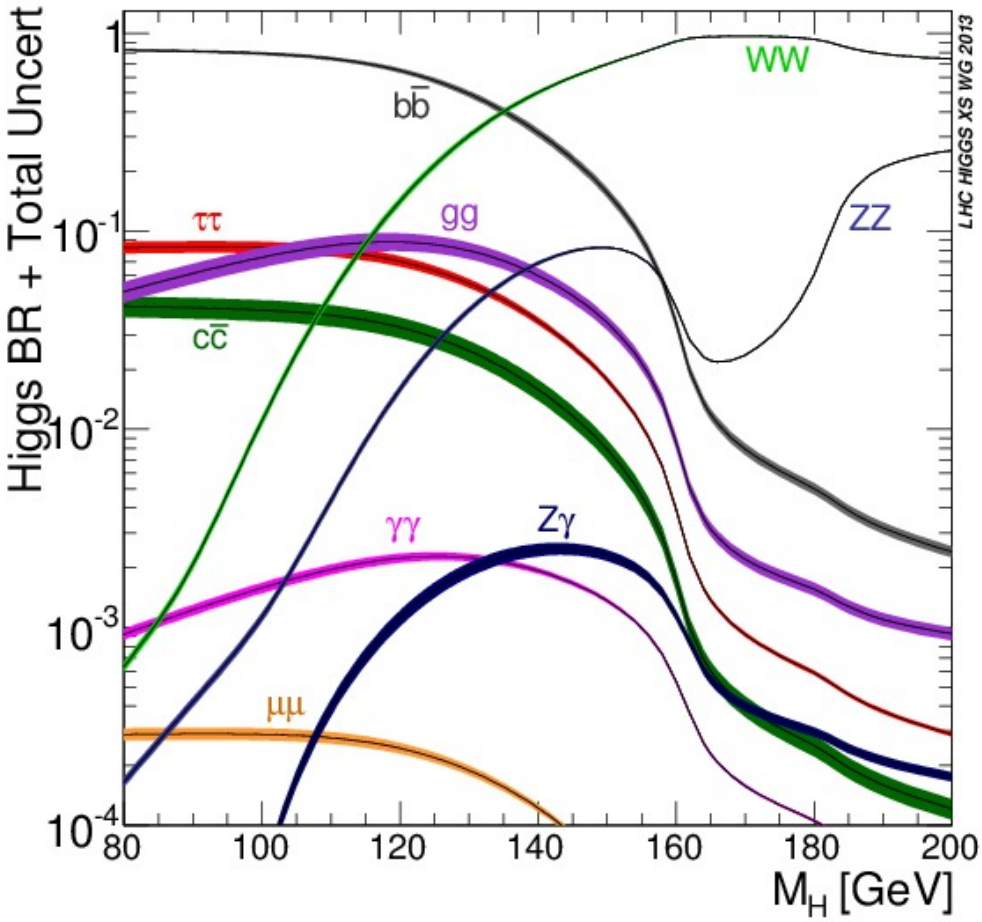


Figure 2.5: Branching ratios of the Higgs boson as a function of its mass [24].

2.1.2 Higgs Boson Production at Proton–Proton Colliders

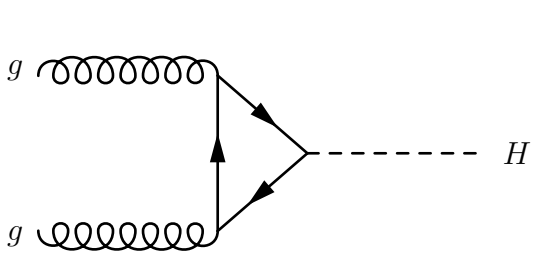
At a proton–proton collider, the Higgs boson is produced primarily through four mechanisms, each involving either the interaction of two gluons or two quarks. Feynman diagrams depicting these processes are shown in Figure 2.6.

The largest production mechanism is *gluon–gluon fusion* (ggF), shown in Figure 2.6a. This mode consists of two gluons interacting to form a triangular quark loop, producing the Higgs boson.[55] Each quark contributes to the loop, but the top quark contribution dominates due to its large mass. Accurate measurement of this production *cross section*[56] could offer indirect hints of the existence of higher generation quarks.

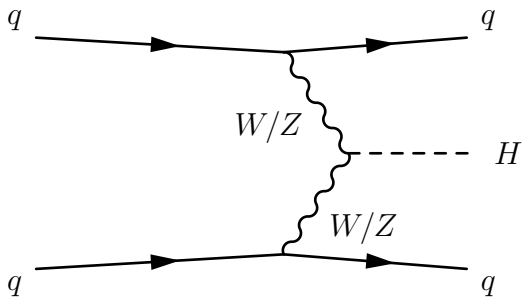
The next largest production mode is *vector boson fusion* (VBF), shown in Figure 2.6b. In this mode, two quarks emit vector bosons that fuse to form the Higgs boson. A distinctive

[55]The Higgs boson cannot couple directly to a pair of gluons because they are massless.

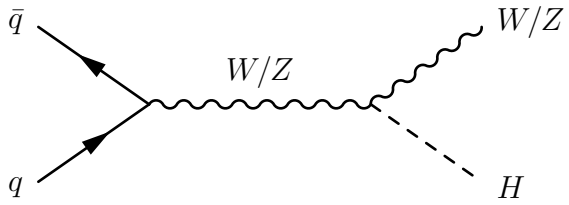
[56]The cross section refers to the effective area over which an interaction occurs.



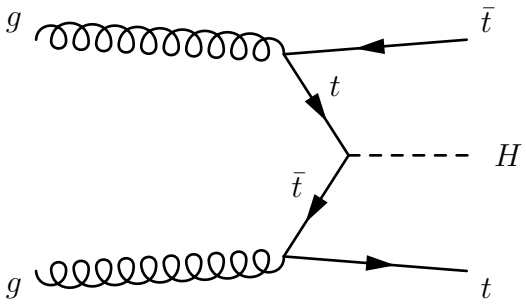
(a) Gluon–Gluon Fusion



(b) Vector Boson Fusion



(c) Associated Production



(d) $t\bar{t}$ Fusion

Figure 2.6: Diagrams of the dominant mechanisms for Higgs boson production at a proton–proton collider.

signature of VBF is two high-momentum jets, widely separated, with the Higgs boson located in-between. Since the color charge of the quarks is not changed by the emission of a vector boson, there is little QCD radiation between the quark jets.

The smallest production modes are *associated production* with a vector boson (VH)[57] or through fusion of two top quarks ($t\bar{t}$ fusion). Respective diagrams are shown in Figures 2.6c and 2.6d. Associated production consists of a vector boson radiating a Higgs boson and offers an excellent experimental signature when the vector boson decays leptonically. The $t\bar{t}$ fusion mode is the rarest production mechanism due to the large amount of energy required

[57]This mode is also known as *Higgs–strahlung*, etymologically similar to *bremsstrahlung*.

in each gluon to produce a pair of top quarks. Since the cross section for this process is smaller than the uncertainty of the ggF cross section, it is not considered further in this dissertation.

The values of the Higgs boson production cross sections depend on its mass and center-of-mass energy of the colliding protons, $\sqrt{s}$,[58] and are shown for the dominant production modes in Figure 2.7:

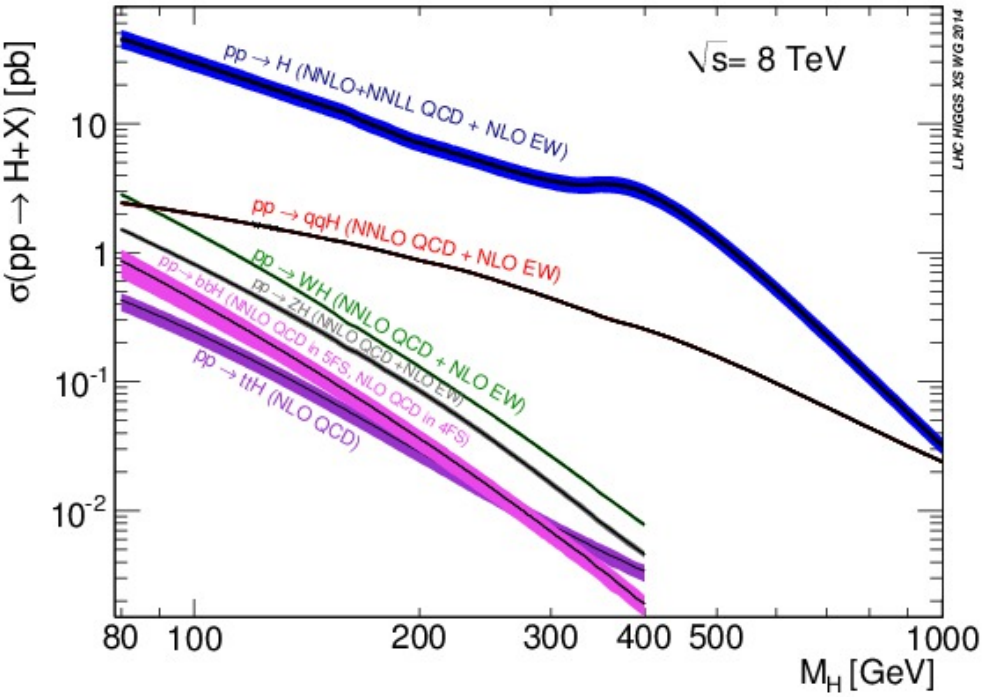


Figure 2.7: Higgs boson production cross sections as a function of its mass [24].

The production cross sections for the 125 GeV Higgs boson in $\sqrt{s} = 8$ TeV proton–proton collisions are shown in Table 2.2:

Production Mode	Cross-Section (pb)
ggF	19.27
VBF	1.58
WH	0.70
ZH	0.41

Table 2.2: The 125 GeV Higgs boson production cross sections in picobarns for proton–proton collisions at $\sqrt{s} = 8$ TeV [24].

[58] s is one of the *Mandelstam variables* and is equal to the square of the center-of-mass energy of a system.

2.1.3 Decays of the Tau Lepton

The tau lepton is the heaviest of the known charged leptons and has a mass of 1.78 GeV. The lifetime of the tau is $\tau_\tau = 2.9 \times 10^{-13}$ s, corresponding to a proper decay length of $c\tau_\tau = 87$ μm. It decays before leaving the beam pipe in which it is produced. The tau decays exclusively through the weak interaction, but is massive enough that the mediating W boson can decay into quarks as well as leptons. Taus that decay into quarks are called *hadronically decaying tau leptons* or simply *hadronic taus*, while those decaying into muons or electrons are *leptonically decaying*, called *leptonic taus*. A Feynman diagram of a hadronic tau decay is shown in Figure 2.8.

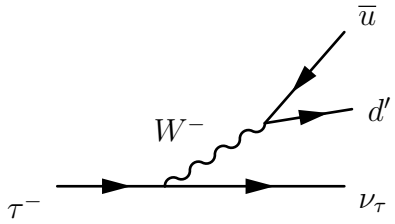


Figure 2.8: Diagram showing a hadronically decaying tau lepton.

As shown in the diagram, there is one neutrino among the decay products. A second neutrino and light lepton are produced in the W boson decay for leptonically decaying tau leptons. The produced neutrinos do not interact with the detector, but appear as *missing energy*. Hadronic taus predominantly decay into pions (ud), but rarely decay into kaons (us). The pions can be produced in an excited state, radiating gluons leading to neutral or charged pions. Most hadronic taus have either one or three charged pions[59] and two or fewer neutral pions. A listing of the branching ratios for the most common tau decay modes is provided in Table 2.3.

Hadronic taus may be confused for jets since they have many decay products in common.[60] However, the taus produced from the decays of weak bosons or Higgs bosons are heavily boosted due to the large mass difference between the tau and the boson, producing tightly collimated decay products. This boosted nature is exploited by algorithms designed to identify hadronic taus from jets. No such identification algorithms exist for leptonically decaying taus since their signature of a light lepton is difficult to separate from other processes producing light leptons.

[59]Hadronic taus with one and three charged decay products are respectively called *one–prong* and *three–prong* taus.

[60]Background processes having the same decay signature as the process of interest are called *irreducible backgrounds*.

Tau Decay Mode	Branching Ratio (%)
Leptonic	35.2
One-prong hadronic	50.1
$\pi^{\pm}$	10.8
$K^{\pm}$	0.7
$\pi^{\pm}\pi^{0}$	25.5
$\pi^{\pm}\pi^{0}\pi^{0}$	9.3
Other one-prong decays	3.8
Three-prong hadronic	14.6
$\pi^{\pm}\pi^{\mp}\pi^{\pm}$	9.3
$\pi^{\pm}\pi^{\mp}\pi^{\pm}\pi^{0}$	4.6
$\pi^{\pm}\pi^{\mp}\pi^{\pm}\pi^{0}\pi^{0}$	0.5
Other three-prong decays	0.2
Other decays	0.1

Table 2.3: Measured branching ratios for different decay modes of the tau lepton. All decay modes include the tau (anti)neutrino [10].

2.2 Lepton Flavor Violation

Lepton flavor refers to a particular lepton generation. The three known lepton flavors are electrons, muons, and taus. Standard Model processes conserve lepton flavor. For example, when a muon decays, its flavor is preserved in the resulting muon neutrino, while the created electron and electron antineutrino respectively carry electron flavor and antiflavor, resulting in no net electron flavor. This accounting is quantified by assigning a *lepton flavor number* of $+1$ to each member of a lepton doublet and -1 to the corresponding antileptons. Processes in which the lepton flavor numbers before and after disagree do not conserve lepton flavor. This non-conservation is called *lepton flavor violation* (LFV).

2.2.1 Motivation

While the Standard Model is an extraordinarily successful theory, its explanation of reality is incomplete. An observed phenomenon not predicted by the SM is *neutrino oscillation*, which is an LFV process. A 1968 experiment by R. Davis, designed to measure the solar neutrino flux, found that only one-third of the expected solar neutrinos were being detected [25]. It was then hypothesized by B. Pontecorvo that the neutrino flavors were changing in flight, thus becoming undetectable to Davis's experiment. This hypothesis was confirmed experimentally by the *Super-Kamiokande* [26] and *Sudbury Neutrino Observatory* [27], [28]

experiments by the early 2000s.

The theory of neutrino oscillations is a beautiful example of the predictive power of elementary quantum mechanics. For illustrative purposes, suppose that neutrinos can oscillate between two flavor states: $|\nu_e\rangle$ and $|\nu_\mu\rangle$. In quantum mechanics, states evolve according to a *Hamiltonian* containing the dynamics of the various states. If a state remains constant under action by a Hamiltonian, then it is an *eigenstate* of the system.[61] Oscillation between flavor states suggests that they are not eigenstates of their governing Hamiltonian and must be some linear combination of the *energy eigenstates*,[62] $|\nu_1\rangle$ and $|\nu_2\rangle$, each with respective energies E_1 and E_2. The energy eigenstates can be written in terms of the flavor states as

$$\begin{aligned} |\nu_1\rangle &= \cos\theta|\nu_\mu\rangle - \sin\theta|\nu_e\rangle \\ |\nu_2\rangle &= \sin\theta|\nu_\mu\rangle + \cos\theta|\nu_e\rangle \end{aligned} \tag{2.15}$$

where θ represents the degree of mixing between the flavor states and energy eigenstates. Equation 2.15 can be inverted to obtain

$$\begin{aligned} |\nu_\mu\rangle &= \cos\theta|\nu_1\rangle + \sin\theta|\nu_2\rangle \\ |\nu_e\rangle &= -\sin\theta|\nu_1\rangle + \cos\theta|\nu_2\rangle \end{aligned} \tag{2.16}$$

The time evolution of the states in Equation 2.16 can be written as

$$\begin{aligned} |\nu_\mu(t)\rangle &= \cos\theta e^{-iE_1t}|\nu_1(0)\rangle + \sin\theta e^{-iE_2t}|\nu_2(0)\rangle \\ |\nu_e(t)\rangle &= -\sin\theta e^{-iE_1t}|\nu_1(0)\rangle + \cos\theta e^{-iE_2t}|\nu_2(0)\rangle \end{aligned} \tag{2.17}$$

Assuming that the neutrino is initially an electron neutrino, $|\nu_1(0)\rangle = -\sin\theta|\nu_e\rangle$ and $|\nu_2(0)\rangle = \cos\theta|\nu_e\rangle$. Substituting these relations into Equation 2.17, it can be shown [13] that the probability of an electron neutrino oscillating into a muon neutrino is

$$P(\nu_e \to \nu_\mu) = |\langle\nu_\mu(t)|\nu_e(t)\rangle|^2 = \left[\sin 2\theta \sin(\frac{E_2 - E_1}{2})t\right]^2 \tag{2.18}$$

Equations 2.17 and 2.18 have interesting implications for neutrino properties. Equation 2.17 shows that the energy of a particular neutrino flavor is not well-defined. The time-dependent terms in the equation further show that the energy eigenstates are waves that oscillate from being in phase to being out of phase. Since these waves propagate through space, this happens if the waves for $|\nu_1\rangle$ and $|\nu_2\rangle$ travel at different speeds. Therefore, $|\nu_1\rangle$ and $|\nu_1\rangle$ necessarily have different masses, as massless waves always propagate at the speed of light.[63]

[61] The theory of quantum mechanics was developed from the linear algebra theory of eigenvalues and eigenvectors. In matrix notation, a state $|\psi\rangle$ is an eigenvector of a (matrix) operator $\hat{A}$ if $\hat{A}|\psi\rangle = \lambda|\psi\rangle$, where λ is a number called the eigenvalue.

[62] Eigenstates of the Hamiltonian are called energy eigenstates.

[63] This argument has been made for two neutrino flavors, but it generalizes in the same manner to more flavors. There are actually three neutrino mass eigenstates, with three mixing angles between the mass and flavor states. Each of the mixing angles has been measured, and the mixing matrix (similar to the CKM matrix) is called the *Pontecorvo–Maki–Nakagawa–Sakata* (PMNS) matrix [29].

Because only (right)left-handed (anti)neutrinos had been observed, it was thought that neutrinos were massless, as only massless particles have a well-defined helicity.[64] The nature of the neutrino mass mechanism is unclear since the SM Higgs boson field requires the presence of right-handed neutrinos to provide a neutrino mass.[65] If neutrinos are their own antiparticles, then they may possess a *Majorana mass*, which appears as a term in the Lagrangian of the form $m_\nu \chi\chi$, where χ is the neutrino *Majorana spinor* satisfying the Majorana equation. In any case, Equation 2.18 shows that neutrino oscillation is an inherently LFV process.

A Higgs boson with LFV decays could account for neutrino mass and oscillations. Many experimental searches have been made for charged LFV processes. While no search has found any evidence of these processes, experimental bounds have set constraints on the many theoretical models beyond the SM featuring charged LFV processes.

2.2.2 Prior Searches

Many experimental searches for charged LFV decays have been conducted, with no experiment offering conclusive evidence of such a process. The presentation of the prior experimental searches has been divided into those looking for LFV decays in the discovered Higgs boson and those in other modes.

2.2.2.1 Searches with the Discovered Higgs Boson

The Higgs boson was discovered in dedicated searches by the *ATLAS* [8] and *CMS Collaborations* [9] using data obtained from 2011–2012 at center-of-mass energies of 7 TeV and 8 TeV. Studies of the $H \to \tau\tau$ channel by both collaborations [30], [31] showed that the Higgs boson does couple to leptons, motivating the possibility that it could have LFV decays. Many studies continue to be performed to discern the exact nature of the recently discovered boson and see whether it offers any hints of physics beyond the SM.

The CMS Collaboration has performed a direct search to find whether or not the discovered Higgs boson also has LFV decays using their 19.7 fb^{-1}, 8 TeV dataset taken in 2012 [32]. They performed a full analysis of the $H \to \mu\tau$ channel, considering both leptonic and hadronic tau decays. In each tau decay channel, 0-, 1-, and 2-jet categories were considered to improve sensitivity. The discriminating variable used is the collinear mass, obtained from the kinematics of the visible decay products[66] and missing transverse energy in the direction of the visible products [33]. Intriguingly, CMS observed an excess of data events in their signal region 2.4 standard deviations from the predicted number of background events. This excess leads to a best fit on the branching ratio $\mathrm{Br}(H \to \mu\tau) = 0.84^{+0.39}_{-0.37}\%$ and an upper limit at 95% confidence of $\mathrm{Br}(H \to \mu\tau) < 1.51\%$.

[64] For massive particles, an observer can always boost to a frame where the helicity of the particle changes sign. This cannot be done for massless particles, which always travel at the speed of light.

[65] See Equation 2.14.

[66] All decay products except neutrinos.

A search for LFV Higgs boson decays with leptonic taus has also been performed at ATLAS [34].[67] This search used a data-driven technique [35] that exploits the fact that SM processes involving oppositely charged muons and electrons are symmetric under their interchange. For example, the number of events in which the muon has higher momentum than the electron is equal to the opposite case. An LFV process violates this symmetry since the leptons produced in the secondary decay of the tau will be softer relative to the other lepton produced in the Higgs boson decay. Other effects that could spoil the symmetry were considered in the search. The collinear mass of the system was used as the discriminating variable, and two signal regions consisting of events with and without jets were used to improve sensitivity. No evidence for any LFV Higgs boson decays was observed, with upper limits at 95% confidence on the branching ratios found to be $\mathrm{Br}(H \to e\tau) < 1.36$ and $\mathrm{Br}(H \to \mu\tau) < 1.79$ for the electron and muon measurements, respectively.

2.2.2.2 Other Searches

While the best limits on the branching ratio of LFV decays of the Higgs boson have been placed by the searches described in the previous section, LFV searches have been ongoing since the 1940s. A review of the history of those searches may be found in [36]. The experimental limits on the branching ratios of LFV processes have been used by theorists to build models of LFV decays under various mechanisms. There are many types of LFV decays involving hadrons, decays of muons and taus, and conversions of muons to electrons in the presence of atomic nuclei. The best limits on the branching ratio at 90% confidence level for some LFV processes are shown in Table 2.4. A full list of limits on branching ratios of LFV processes may be found in [10].

2.2.3 Phenomenology

There are three possible LFV Higgs boson decay modes: $H \to \mu\tau$, $H \to e\tau$, and $H \to e\mu$. Studies have been performed that predict the experimental signatures of these decay modes at ATLAS and CMS. A common feature of these studies is that the Higgs field Yukawa couplings are parametrized as a 3×3 matrix in the basis (e, μ, τ), with each element denoted as y_{ij}. This matrix is diagonal in the SM, with $y_{ee} = \frac{\sqrt{2}m_e}{v}$, $y_{\mu\mu} = \frac{\sqrt{2}m_\mu}{v}$, $y_{\tau\tau} = \frac{\sqrt{2}m_\tau}{v}$, and $y_{ij} = 0$ for $i \neq j$. One study [44] found that existing limits on the branching ratios of LFV decays, including many of those in Table 2.4, and other precision measurements lead to the following conclusions concerning LFV Higgs boson decays:

- The largest possible branching ratio for $H \to \mu\tau$ and $H \to e\tau$ is $\sim$10%.
- The tight constraints on $\mu \to e\gamma$ and $\mu \to e$ conversions suggest that $H \to e\mu$ decays are highly suppressed and consequently unobservable at ATLAS and CMS.

[67] As of the time of writing, the manuscript is considered for publication in *European Physics Journal C*.

LFV Process	90% CL on Br	Collaboration	Reference
$\mu \to e\gamma$	5.7×10^{-13}	MEG	[37]
$\mu \to eee$	1.0×10^{-12}	SINDRUM	[38]
$\tau \to \mu\gamma$	4.4×10^{-8}	BaBar	[39]
$\tau \to e\gamma$	3.3×10^{-8}	BaBar	[39]
$\tau^- \to \mu^-\mu^+\mu^-$	2.1×10^{-8}	Belle	[40]
$\tau^- \to e^-e^+e^-$	2.7×10^{-8}	Belle	[40]
$\tau^- \to e^-\mu^+\mu^-$	2.7×10^{-8}	Belle	[40]
$\tau^- \to e^+\mu^-\mu^-$	1.7×10^{-8}	Belle	[40]
$\tau^- \to \mu^-e^+e^-$	1.8×10^{-8}	Belle	[40]
$\tau^- \to \mu^+e^-e^-$	1.5×10^{-8}	Belle	[40]
$\mu(\mathrm{Au}) \to e(\mathrm{Au})$	7×10^{-13}	SINDRUM II	[41]
$\mu^-(\mathrm{Ti}) \to e^-(\mathrm{Ti})^{\mathrm{GS}}$	4.3×10^{-12}	SINDRUM II	[42]
$\mu^-(\mathrm{Ti}) \to e^+(\mathrm{Ca})^{\mathrm{GS}}$	1.7×10^{-12}	SINDRUM II	[43]

Table 2.4: 90% confidence limits on branching ratios for various lepton-flavor-violating processes, with the experimental collaboration and reference. The limit obtained for $\mu \to e$ conversion in titanium was for ground state (GS) conversions of the nuclei.

- The constraints on $\mu \to e\gamma$ also imply that either $H \to \mu\tau$ or $H \to e\tau$ may be observed, but not both.

Another study [45] calculated bounds on the Yukawa couplings using previous measurements and a reinterpretation of the ATLAS $H \to \tau\tau$ search using 7 TeV data. The findings are shown in Figure 2.9. The dashed line in the lower left corner of each plot indicates the largest theoretical bound that preserves the mass hierarchy $m_\tau > m_\mu > m_e$. The hierarchy can be maintained above this limit, but only at the cost of fine-tuning different parameters in the modeling.[68]

Many theoretical models have been developed showing how LFV Higgs boson decays arise in the framework of existing theories. The remaining sections discuss LFV Higgs boson decays in the context of supersymmetry, an inverse seesaw model, and Randall–Sundrum Models.

[68]Fine-tuning refers to the precise selection of model parameters such that divergent integrals in the calculation cancel each other out. The minimization of fine-tuning in a theoretical model is called *naturalness* and is generally considered a benchmark of a sound model. Theories that require fine-tuning usually require an explanation of why such tuning is necessary. A related concept is *naturalness*, which stipulates that parameters in a model should be on the order of 1. Otherwise, fine-tuning might be present in the model.

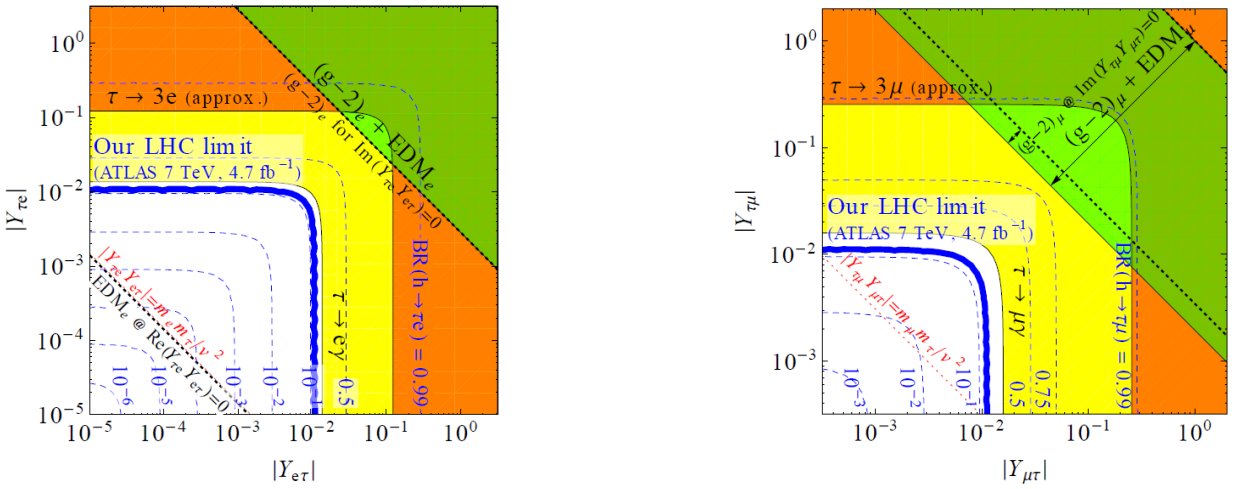


Figure 2.9: Bounds placed on the lepton-flavor-violating Yukawa couplings $|y_{e\tau}|$, $|y_{\tau e}|$, $|y_{\mu\tau}|$, and $|y_{\tau\mu}|$ [45]. The dashed line in the lower left corner of each plot indicates the largest theoretical bound that maintains the observed mass hierarchy $m_\tau > m_\mu > m_e$ without precise selection of theoretical model parameters.

2.2.3.1 Supersymmetric Models

Supersymmetry (SUSY) is a class of theories asserting that there is an internal symmetry between bosons and fermions [46]. For example [13], if ϕ and ψ are scalar and spinor fields, and each transforms such that $\delta\phi = 2\overline{\epsilon}\psi$ and $\delta\psi = -i\gamma^\mu\epsilon(\partial_\mu\phi)$, where ϵ is an infinitesimal spinor field, then the following Lagrangian is invariant under such transformations:

$$\mathcal{L}_{\text{SUSY}} = \frac{1}{2}\left[(\partial^\mu\phi^*)(\partial_\mu\phi) - m^2\phi^*\phi\right] + i\overline{\psi}\gamma^\mu\partial_\mu\psi - m^2\overline{\psi}\psi \tag{2.19}$$

The invariance of Equation 2.19 is contingent upon both fields having the same mass m. Thus, under a SUSY transformation, spin-$\frac{1}{2}$ particles are rotated into spin-0 particles, and vice versa. Similar transformations also rotate the spin-1 SM gauge bosons into spin-$\frac{1}{2}$ particles.

SUSY is capable of solving several difficult theoretical issues. It offers a candidate particle for dark matter. Under SUSY, the weak, strong, and EM couplings unify at a single point, while they do not under the SM. Finally, it can solve the *hierarchy problem*, the question of why the force of gravity is so feeble compared to the electroweak forces. This problem has an important impact on particle phenomenology. The Higgs boson mass is determined not only by parameters of the Higgs potential, but also by its self-interactions through other particles. Gravitational interactions are neglected in the SM due to the weak strength of their interactions. However, at energies on the order of 10^{19} GeV, gravitational effects manifest noticeably at the quantum level. While a theory of quantum gravity has yet to be developed, gravitational contributions to the Higgs boson mass are expected to make it

several orders of magnitude larger than the observed value. In the SM, significant fine-tuning of model parameters is required to cancel out the diagrams that drive up the Higgs boson mass. SUSY offers a complementary set of diagrams canceling out the large contributions naturally, theoretically motivating the relatively low 125 GeV Higgs boson mass. No evidence for SUSY has been observed, suggesting that (if it exists) SUSY is badly broken with the superpartner particles having masses much larger than the observed SM particles.

The simplest SUSY model is the *Minimal Supersymmetric Standard Model* (MSSM), which offers the fewest number of superpartners to the SM particles. Each fermion has two associated spin-0 superpartners, each for the left- and right-handed fields.[69] These superpartners are called *sfermions* and consist of *squarks* and *sleptons*.[70] The charged sleptons are the *selectron*, *smuon*, and *stau*, while the neutrino sleptons are called *sneutrinos*. Like the SM fermions, the left-handed sfermions transform as a doublet under $SU(2)$, and the right-handed sfermions transform as a singlet.

The Higgs field also has an associated spin-$\frac{1}{2}$ superpartner doublet called the *Higgsino*. To maintain $SU(2)_L \times U(1)_Y$ gauge invariance, there must be *two* Higgsinos, each with opposite hypercharge $= \frac{1}{2}$. Furthermore, the structure of SUSY theories requires that there be two Higgs doublets, one to provide mass to the up-type quarks and the other to provide mass to the down-type quarks and charged leptons.[71] Both Higgs fields have eight total degrees of freedom. Since three degrees of freedom provide mass to the weak bosons, the remaining five degrees of freedom result in *five* Higgs bosons: a light neutral, a heavy neutral, two charged, and a neutral that is antisymmetric under the CP operation. The ratio of the two Higgs field vevs is a parameter called $\tan\beta$.

The EWK $\mathbf{W}$ and B bosons have SUSY partners called the *wino* and *bino*, and each gluon has an associated *gluino*. The neutral wino and bino mix together to make the *zino* and *photino*, the superpartners of the photon and Z boson. Under electroweak symmetry breaking, the neutral components of the Higgsinos mix with the bino and neutral wino to form mass eigenstates called *neutralinos*. The charged components of the Higgsinos mix with those of the *wino* to form mass eigenstates called *charginos*.

Under many SUSY models, antisquarks can decay directly into pairs of quarks, which can lead to proton decay. To prevent this from happening, many models impose a conservation law called *R–parity*, defined as $R = (-1)^{3(B-L)+2s}$, where B and L are the baryon and lepton numbers for a given process. In such models, there is a neutral *lightest supersymmetric particle* (LSP) that cannot decay into SM particles since doing so would violate R–parity. These LSPs are good dark matter candidates since they are minimally interacting particles.

Some SUSY models have been developed examining LFV decays of the three neutral MSSM Higgs bosons given known LFV constraints and for different settings of various SUSY parameters. One study found that the MSSM neutral Higgs boson branching ratios for LFV decays get smaller with increasing SUSY mass scale [47]. Another study argued that any

[69] Note that since the superpartners are spin-0, they cannot be left- or right-handed.

[70] The top and bottom squarks are generally known as *stop* and *sbottom* while the remaining squarks do not use this naming convention.

[71] In the SM, the conjugate of the Higgs field $\mathbf{\Phi}^c$ provides the mass to the up-type quarks.

observation of LFV Higgs boson processes could be evidence for SUSY, depending on the permitted degree of slepton mixing and value of $\tan\beta$ [48]. While not explicitly depending on SUSY, one study examined LFV Higgs boson decays in the context of the two Higgs doublets previously discussed [49].

2.2.3.2 Inverse Seesaw Model

One of the mysteries in modern physics concerns the nature of the neutrino mass. In the SM, a right-handed neutrino is required for the Higgs mechanism to provide mass to the neutrino, yet none are observed. Even if such neutrinos were observed, the Yukawa coupling of the Higgs field to the neutrino would be on the order of 10^{-11} [20], which screams fine-tuning. A more natural mechanism for neutrino mass is the *seesaw mechanism* [20], which presumes that the neutrino has a right-handed component that transforms as a singlet under $SU(2)_L$ and does not interact with any other particles except the Higgs boson. If the left- and right-handed neutrino spinors are denoted as ν_L and ν_R, with respective charge conjugate spinors ν_L^c and ν_R^c,[72] then mass terms may be written as $\overline{\nu_L}\nu_L^c$, $\overline{\nu_R}^c\nu_R$, $\overline{\nu_L}\nu_R$, $\overline{\nu_R}^c\nu_L^c$, and their Hermitian conjugates.[73] The first two terms are Majorana mass terms, while the second pair of terms are *Dirac* mass terms, acquiring their mass through coupling to the Higgs field. These terms may be written in a Lagrangian as

$$\mathcal{L}_\nu = -\frac{1}{2}\begin{pmatrix} \overline{\nu_L} & \overline{\nu_R}^c \end{pmatrix}\begin{pmatrix} m_L & m_D \\ m_D & m_R \end{pmatrix}\begin{pmatrix} \nu_L^c \\ \nu_R \end{pmatrix} + h.c. \tag{2.20}$$

where m_L and m_R are the Majorana masses for ν_L and ν_R, and m_D is the Dirac mass obtained from coupling to the Higgs field. m_D is taken to be on the order of the electron mass.

The observable neutrino mass eigenstates are a mixture of the left- and right-handed neutrinos such that the 2×2 matrix in Equation 2.20 is diagonal in the mass-eigenstate basis. The diagonal entries are the observable masses and may be written as

$$m_\pm = \frac{1}{2}\left(m_L + m_R \pm \sqrt{(m_L - m_R)^2 + 4m_D^2}\right) \tag{2.21}$$

If $m_R \gg m_D$ and $m_R \gg m_L$, then approximate solutions to Equation 2.21 may be written as $m_+ \approx m_R$ and $m_- \approx \frac{-m_D^2}{m_R}$.[74] The observed neutrino mass is very tiny due to the suppression by m_R, while the neutrino Yukawa coupling is more natural since it is based on the Dirac mass. There is a right-handed neutrino state with mass m_R that is presently unobserved, according to the model.

[72] $\nu_L^c = i\nu_L^*$, $\nu_R^c = i\nu_R^*$.

[73] Unlike charged fermions, these types of mass terms are permitted for neutrinos because they are electrically neutral and consequently maintain gauge invariance under $SU(2)_L \times U(1)_Y$ transformations.

[74] In the Lagrangian shown in Equation 2.20, the negative sign in m_- can be eliminated by redefining the fields ν_L^c and ν_R to absorb the sign. The term "seesaw" comes from the observation that if m_R increases, then m_+ increases and m_- decreases, and vice versa if m_R decreases.

The seesaw mechanism provides an explanation for neutrino mass but fails to account for neutrino oscillations. This is accomplished in the *inverse seesaw model* (ISS) [50], which adds a second right-handed neutrino having opposite lepton number with respect to the first, mediating the neutrino LFV process.

LFV Higgs boson decays can be easily built into an ISS model. One study [51], [52] examined LFV Higgs boson decays in an ISS model with the following Lagrangian:

$$\mathcal{L}_{ISS} = -Y_\nu^{ij}\overline{\psi}_L \mathbf{\Phi}^c \nu_{R,j} - M_R^{ij}\overline{\nu_{R,i}}^c X_j - \frac{1}{2}\mu_X^{ij}\overline{X_i}^c X_j + h.c. \tag{2.22}$$

where $i, j = e, \mu, \tau$, Y_ν^{ij} are the Yukawa couplings for the (LFV) Higgs boson to different neutrino flavors,[75] $\overline{\psi}$ is the left-handed $SU(2)$ lepton doublet (see Section 2.1.1.1), $\mathbf{\Phi}^c$ is the charge conjugate Higgs field (see Section 2.1.1.3), X is the right-handed neutrino having the opposite lepton number as ν_R, and M_R and μ_X are mass matrices. Like the seesaw mechanism, a 9×9 mass matrix can be constructed using the basis (ν_L^c, ν_R, X):

$$M_{\text{ISS}} = \begin{pmatrix} 0 & m_D & 0 \\ m_D^T & 0 & M_R \\ 0 & M_R^T & \mu_X \end{pmatrix} \tag{2.23}$$

Each zero represents a 3×3 array of zeros. In this model, ν_L is taken to be massless, hence the zeros in the matrix. The matrix m_D corresponds to the Dirac mass in the seesaw mechanism, taken to be $m_D = vY_\nu$ where v is the Higgs vev. Diagonalizing this matrix with the assumption that the entries in μ_X are much smaller than others leads to three light neutrino mass eigenstates and six heavy mass eigenstates. The Higgs field and right-handed neutrino X are responsible for LFV in the model.

The study in [51] found that LFV Higgs boson decays would be highly suppressed under this model if the heavy neutrinos were degenerate in terms of mass and the matrices M_R, μ_X are diagonal, with the largest branching ratio approximately 10^{-11} for M_R on the order of 20 TeV. The suppression was shown to become worse as M_R decreased. By contrast, if the restriction on diagonal M_R and μ_X is relaxed, then the largest branching ratio increases to about 10^{-5}.

2.2.3.3 Randall–Sundrum Models

Many studies have been conducted predicting or explaining observed particle phenomena based on the structure of spacetime. The geometry of spacetime is encoded in an object called the *metric*, which establishes the relationship between various spacetime dimensions. In a flat spacetime, all dimensions are independent, but in a curved spacetime, the distance traveled in one dimension can depend on the location in another dimension.[76] *Randall–Sundrum* (RS) models [53] posit that the structure of spacetime is five-dimensional, with

[75] The Yukawa couplings can be represented in a 3×3 matrix denoted Y_ν.

[76] If this sounds strange, consider moving along lines of constant latitude on the Earth. The longest distance traveled is around the Equator, while it becomes considerably shorter as one approaches the poles.

the following metric:

$$ds^2 = e^{-2kr_c\phi}\eta_{\mu\nu}dx^\mu dx^\nu + r_c^2 d\phi^2 \tag{2.24}$$

where $\eta_{\mu\nu}dx^\mu dx^\nu$ is the conventional flat spacetime, warped by an extra dimension ϕ that is *compactified* in a circle of radius r_c.[77] The coordinate ϕ ranges from 0 to π and results in an exponential warping of the 4-D spacetime. Under some RS models, fermions and gauge bosons propagate in the full 5-D spacetime, but the Higgs field can only propagate in 4-D [11]. This results naturally in the observed hierarchy of fermion and boson masses since the couplings of the Higgs field to other fermion and boson fields are affected by the spacetime structure. It is possible for LFV Higgs boson decays to be built into these models using an approach similar to the ISS model discussed in the last section [54].

[77] A compactified dimension is one that does not have infinite extent.

Chapter 3

The Large Hadron Collider and the ATLAS Experiment

> *[W]hen you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind...* —William Thomson, Lord Kelvin, *Electrical Units of Measurement*, 1883

One cannot conduct a physical experiment without a measuring apparatus. This chapter describes the elaborate measurement apparatus used to investigate the subject of this dissertation. It begins with discussion of the Large Hadron Collider and how it creates two high-energy, counter-circulating beams of protons. The ATLAS detector and how its systems record the particles produced in the proton–proton collisions are then described.

3.1 The Large Hadron Collider

The *Large Hadron Collider* (LHC) is the world's most powerful particle collider. It straddles the Franco-Swiss border less than 10 km from the Geneva city center. Administration, construction, and maintenance of the LHC are performed by the *European Organisation for Nuclear Research*,[1] known internationally by the acronym *CERN*.[2] The LHC control center is located at the main CERN campus in Meyrin, Switzerland.

The LHC is arguably the largest and among the most complex machines ever constructed. It consists of a series of linear and circular accelerators that progressively accelerate particles, ultimately reaching the LHC ring, which has a circumference of 26.7 km and is buried roughly

[1] In French, *Organisation Européenne pour la Recherche Nucléaire*.

[2] CERN originally stood for *Conseil Européen pour la Recherche Nucléaire*, but was renamed *Organisation Européenne pour la Recherche Nucléaire* in 1954. It was decided that the acronym OERN was awkward, so the acronym CERN was kept.

170 m underground, with a small tilt of 1.4%. This final stage is designed to accelerate particles to center-of-mass energies of 14 TeV (7 TeV per beam).[3]

The history of the LHC begins with its immediate predecessor, the *Large Electron–Positron Collider* (LEP). The tunnel containing the LHC ring was originally constructed for LEP from 1984 to 1989, with LEP commencing operation after completion. In the early 1990s, discussions for converting LEP into a hadron collider began, motivated in part by calculations indicating that the Higgs boson, if it existed, had to have a mass less than 1 TeV, which is within reach of a hadron collider. In December 1994, plans were made to close LEP in 2000 and retrofit the existing infrastructure to house the LHC. Construction of the LHC and its experiments was completed in 2008, with first beams occurring on September 18 of that year. Unfortunately, the very next day, a faulty bus splice led to an electrical short that caused about a hundred of the superconducting[4] magnets to quench, or lose their superconducting properties. Consequently, there was a violent release of about six metric tons of liquid helium, severely damaging about fifty of the magnets in the vicinity and contaminating the beam pipe. It was about a year before the LHC was able to resume its operations, with first beams occurring on November 20, 2009. After a brief commissioning period, data collection for analysis commenced on March 30, 2010, with beams circulating at center-of-mass energy of 7 TeV. At the end of 2011, the LHC center-of-mass energy was upgraded to 8 TeV, and operations continued until the beginning of a long shutdown on February 13, 2013. After a roughly two year period, proton collisions at 13 TeV began on June 3, 2015.

A detailed overview of the LHC (from which much of the material in the following sections is obtained) may be found in [55].

3.1.1 Layout

The LHC is designed to collide two counter-circulating beams of protons moving at 99.9999991 percent of the speed of light. Its overall structure is shown in Figure 3.1. There are eight straight sections labeled *octants*, in which beams may be inserted, collided, or dumped, and four sections in which the beams intersect. Superconducting magnets are used to bend the beams around the ring and focus them to prepare for collisions. There are four major experiments that are situated at these intersection points: *A Large Toroidal ApparatuS* (ATLAS), *Compact Muon Solenoid* (CMS), *Large Hadron Collider beauty* (LHCb), and *A Large Ion Collider Experiment* (ALICE).[5] ATLAS and CMS are general-purpose detectors, while ALICE studies the physics of high energy ion collisions, and LHCb studies the physics associated with b-quarks.

The LHC is designed to deliver a very high rate of collisions, or *luminosity*, of 10^{34} cm^{-2} s^{-1} to ATLAS and CMS. The design luminosity of LHCb is 10^{32} cm^{-2} s^{-1}, or 1% that of ATLAS

[3] 7 TeV = 1.1 μJ. This might not seem like much, but it is an incredible amount of energy concentrated in a single proton!

[4] Anything superconductive has zero electrical resistance and can sustain an electric current indefinitely.

[5] For a musical introduction to the LHC, see [56]

and CMS. The two different luminosities are made possible by focusing the beams differently at each of the respective interaction points.

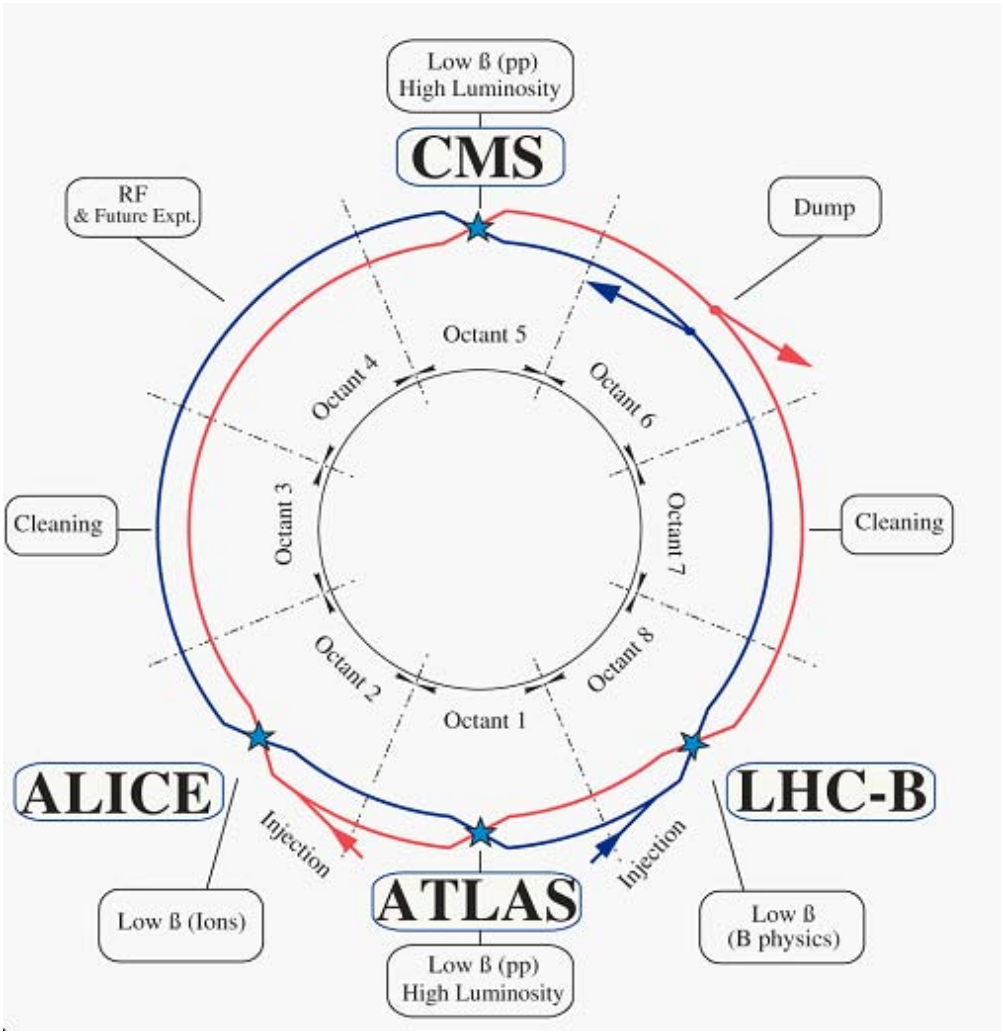


Figure 3.1: Layout of the LHC Experiments [55].

The beams consist of groups of protons (called *bunches*) surfing in the troughs of an electromagnetic wave, rotating at 8400 MHz in the ring. The troughs of this EM wave are called *radio-frequency* (RF) *cavities* since the cavity frequency is 400 MHz and are created using devices called *cryomodules*, with two of them per section to accommodate two counter-rotating beams. Power to the cryomodules is supplied by 300 kW *klystrons*, a type of power amplifier, providing a 2 MV voltage drop across each cavity and accelerating the protons. It is critical to successful operation that the protons be accelerated uniformly. This is accomplished by keeping the protons tightly bunched within each cavity and by setting the revolution frequency to a multiple of the RF frequency. At design energy and luminosity,

there are 2808 circulating bunches, each containing 1.15×10^{11} protons and spaced 25 ns apart.[6]

The beams are bent around the ring and focused using superconducting magnets kept at 1.9 K. This temperature is attained through immersion of the magnets in superfluid helium, which has its largest thermal conductivity at that temperature. Superfluid helium offers increased stability against thermal disturbances, reducing the risk of magnet quenching. Since the superconducting current depends on the temperature, maintaining the magnets at 1.9 K is critical to produce the magnetic fields with the required strength to bend the proton beams.

It is undesirable for protons to collide with anything outside of the dedicated interaction points. To prevent this from occurring, the beam pipe is evacuated to an internal pressure is 10^{-10}–10^{-11} mbar at room temperature. Furthermore, the beam pipe is cooled to 1.9 K to reduce beam energy loss from collisions with residual gases.

Successful operation of the LHC is contingent on controlling the high energy beams at all times. In the event that loss of beam stability seems likely, a safe and reliable method of dispersing the beam must be utilized to avoid damage to the LHC or experiments. The event of beam dispersal is called a *beam dump* and is facilitated by using fifteen *kicker* and *septum* magnets to rapidly spread the bunches in the transverse direction. Kicker magnets are dipole magnets capable of switching on and off very quickly, while septum magnets split the beam down the middle, allowing for faster dispersion. The beam is further dispersed by ten dilution magnets and absorbed in several layers of steel and concrete. Reliability studies of the beam dump system find that it will fail only once for every 10^6 hours of operation (once every 114 years).

The LHC may also be configured to collide lead nuclei ($^{208}\mathrm{Pb}^{82+}$) at 2.76 TeV/nucleon, or total center-of-mass energy of 1.15 PeV, with a luminosity of 10^{27} cm^{-2} s^{-1}. No ion collisions at these energies have been studied before. While primarily for the benefit of ALICE, ATLAS and CMS collect data from these collisions for study.

3.1.2 Injection Chain

The protons in the LHC ring collide at very high energy, but they first go through several stages of acceleration. Figure 3.2 shows the *injection chain*, or connected series of accelerators leading to the LHC ring. A detailed technical report of the injection chain may be found in [57].

The acceleration process begins with ordinary hydrogen gas. Since only the protons are accelerated, the electrons must be stripped away. This is done by passing the hydrogen gas through a *duoplasmatron*, a device that produces ion beams. Strong electric and magnetic fields inside the duoplasmatron ionize the hydrogen gas and expel the protons. RF-frequency quadrupole magnets located right outside the duoplasmatron focus the expelled protons into a beam with an energy of 750 keV.

[6] This seems like a large number of protons, but consider that the total mass is about half a nanogram.

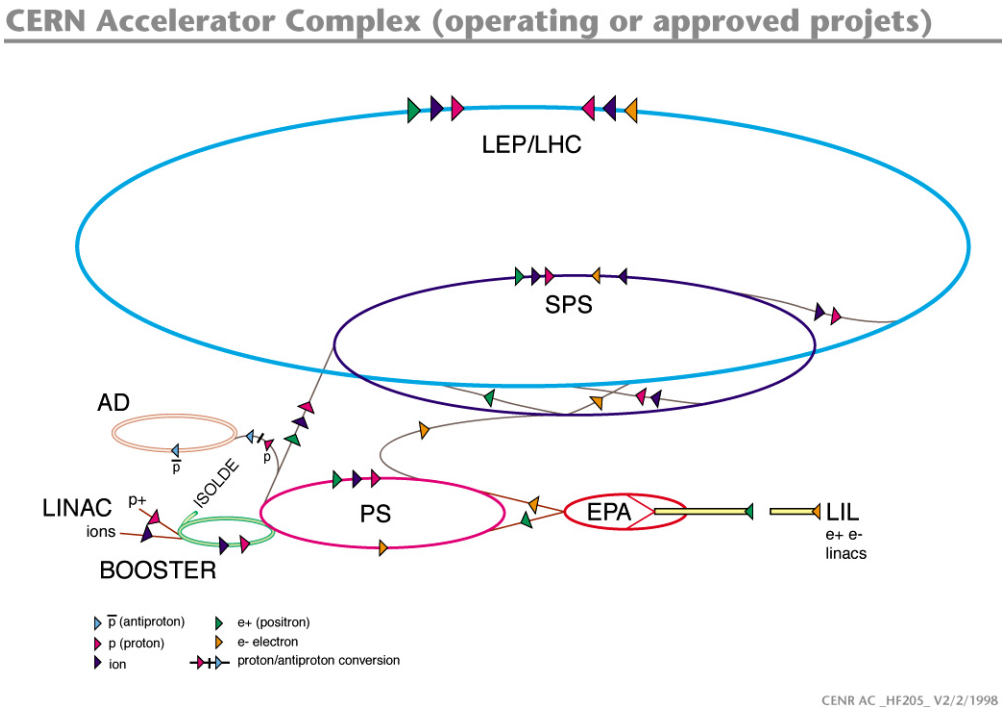


Figure 3.2: The LHC Injection Chain [58].

The beam is then accelerated to 50 MeV in *Linear Accelerator 2* (LINAC2), a straight (as opposed to circular) accelerator. It is then injected into the *Proton Synchrotron Booster* (PSB), a circular accelerator consisting of four stacked rings, each 157 m in circumference. The PSB accelerates the protons to 1.4 GeV and injects them into the *Proton Synchrotron* (PS), a circular accelerator 628 m in circumference.[7] The PS further accelerates the beam to 25 GeV and injects it into the *Super Proton Synchrotron* (SPS), a large ring 7 km in circumference.[8] Here, the protons undergo the last intermediate acceleration to 450 GeV, then are injected into the LHC where they are accelerated to their maximum energy.

The principles of accelerating the particles in the PSB, PS, and SPS are similar to those of the LHC. A guided EM wave is created in each synchrotron through the use of RF cavities. The oscillation of these cavities is set up such that the troughs of the cavity rotate around the ring at a given frequency. By increasing the rotational frequency, energy is added to the protons, and they are consequently accelerated to higher energies. Dipole and quadrupole magnets are also used to bend and focus beams, but since the energy required to bend the beams is substantially lower than that of the LHC, the magnets are capable of operating at

[7]Coming online in 1959, the PS is the oldest synchrotron at CERN. It has accelerated many types of particles and was used to make the neutrino beams that led to the discovery of neutral currents at Gargamelle in 1973.

[8]Like the PS, the SPS is a famous collider. It produced the beams that led to the discovery of the W and Z bosons by the UA1 and UA2 collaborations in 1983.

room temperature.

In addition to accelerating the beams, the PSB and PS establish the proton bunch structure and spacing. Six bunches are initially injected into the PSB with an "empty" bunch inserted between the first and sixth bunches. These six bunches are accelerated and split into 18 bunches upon injection in the PS. As the bunches are accelerated to 25 GeV, they are divided again into 72 bunches. The "empty" bunch injected is similarly expanded into twelve empty bunches. Each bunch sits in an RF cavity, or *bucket*. The bunches are spaced apart and designed such that one passes by every 25 ns in the LHC. Thus, 72 filled buckets followed by 12 empty buckets are injected from the PS to the SPS. The twelve empty buckets, producing a time gap of 320 ns, are placed to ensure that the beam dump system has sufficient time to activate and deflect the beams before any damage might occur. This configuration of 72 filled buckets, spaced apart by 25 ns and followed by 12 empty buckets corresponding to a gap of 320 ns, is known as a *bunch train*. The LHC is designed to have 2808 filled buckets, corresponding to 39 bunch trains. The 2012 LHC run conditions varied from the design specification and will be discussed in Section 3.1.4.

3.1.3 Magnets

Like all electrically charged particles, protons are subject to the Lorentz force law, $\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$, where $\mathbf{F}$ is the force acting on the proton, q is the charge of the proton, $\mathbf{v}$ is the proton velocity, and $\mathbf{E}$ and $\mathbf{B}$ are the electric and magnetic fields acting on the proton. In the presence of a uniform magnetic field, electrically charged particles move in circular orbits. The operation of a circular accelerator is based on this mechanism. An electric field is applied to accelerate the particles. The energy of a charged particle moving uniformly in a magnetic field is directly proportional to the square of the magnetic field strength. Consequently, high energy beams require sophisticated magnets capable of providing the field strength necessary to bend the beam trajectories. In addition to providing this bending force, magnets are used to focus the beams, keeping the protons in each bunch from spreading too far apart.

The magnets at the LHC are constructed from niobium–titanium (NbTi) *Rutherford cables*, superconducting by design. The cables are cooled to 1.9 K with superfluid helium, allowing them to operate in a superconducting state. These magnets can produce a magnetic field of 8.33 T,[9] produced by an electric current of 11850 A. This is a substantial improvement over earlier accelerator magnets that operated at 4.2 K with supercritical helium. The magnetic field generated by these earlier magnets was 5 T [55].

The workhorse magnets of the LHC are *dipole magnets*,[10] separated such that the protons pass between the two poles. The magnetic field is very uniform in this configuration, ensuring that the bunch trajectories are evenly bent. Successful operation requires that the field strength and direction vary by less than one part in ten thousand. A schematic of the dipole magnets is shown in Figure 3.3. There are 1232 such dipole magnets at the LHC.

[9] This is 166600 times stronger than the Earth's magnetic field at the surface.

[10] So named because they consist of a north pole and a south pole

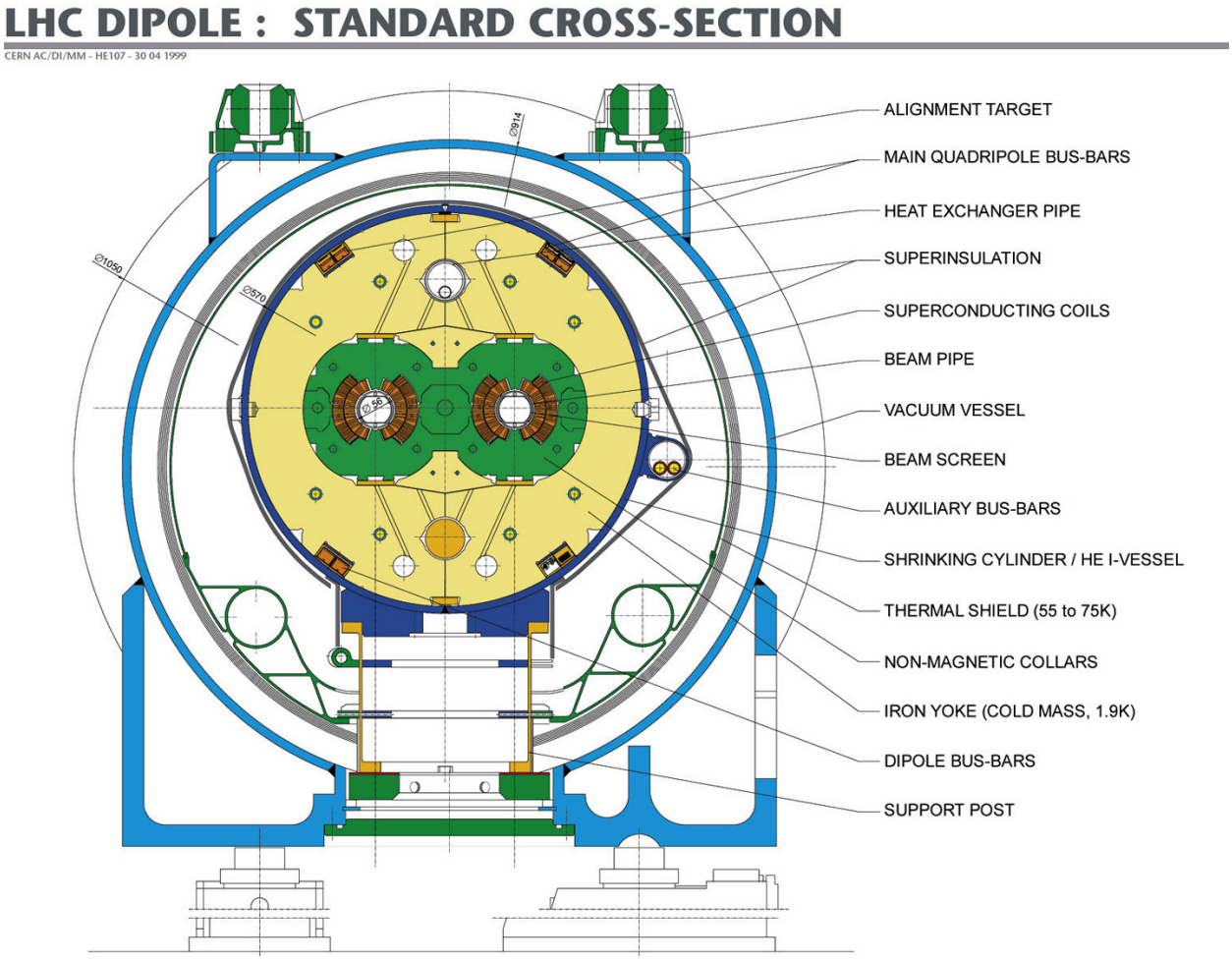


Figure 3.3: Cross-sectional schematic of an LHC dipole magnet [59].

The operational requirements of these magnets posed some engineering challenges. There are two counter-circulating beams of particles with the same charge, so dipole magnets for one beam require opposite polarity to those used for the other beam. However, very limited space in the tunnel necessitated compact magnets. As shown in Figure 3.3, a twin bore design was used to solve this challenge. The two beam pipes share the same iron yoke[11] and *cryostat*, the device that maintains the cold temperature. The superconducting coils are wound such that the orientation of the magnetic flux circling one pipe is opposite to that circling the other. As a result, both beam pipes are mechanically and magnetically bound together.

The magnets are also subjected to incredibly strong forces, on the order of a few meganewtons, during operation. Considerable structural support is provided to resist these forces. Since the magnets contract when cooled from room temperature, a rigid collar is applied to

[11] Also known as the *cold mass*.

pre–stress the magnet and furnish mechanical support.

The other main class of magnets at the LHC are *quadrupole magnets*. While dipole magnets are composed of two magnetic poles, generating a uniform magnetic field, quadrupole magnets consist of four magnetic poles (two north and two south) that "squeeze" the proton bunches, focusing them more tightly. There are 392 such magnets at the LHC. Since quadrupole magnets can only squeeze in one direction, several of them with different orientations are used for focusing. Larger order multipole magnets are used to correct additional deviations in the beam structure. Very high precision quadrupole magnets are placed in the vicinity of the interaction points for ATLAS and CMS to tightly focus the beams from 0.2 mm to nominally 16.7 μm. In 2012 operations, the focused beam size was slightly larger at 19 μm [60].

3.1.4 Luminosity

Upon creation, the proton beams are collided to produce physics processes for measurement. Many such measurements are dependent upon the collision rate. This quantity is called the *instantaneous luminosity* and is of such importance that each LHC experiment makes a dedicated effort to measure it. The basic equation for the instantaneous luminosity is given by Equation 3.1:

$$\mathcal{L} = \frac{1}{\sigma}\frac{dN}{dt} \tag{3.1}$$

where $\mathcal{L}$ is the instantaneous luminosity, measured in $\text{cm}^{-2}\ \text{s}^{-1}$, σ is the *scattering cross section*[12] for the process in question (two colliding protons, in this case), and $\frac{dN}{dt}$ is the rate of change of events corresponding to the cross section (N is the number of collisions as a function of time).

While the instantaneous luminosity is an excellent metric for assessing collider performance, a quantity of interest to measurements is the *integrated luminosity*, which is proportional to the total number of collisions or events. The integrated luminosity can be calculated by integrating Equation 3.1 with respect to time:

$$L = \int \mathcal{L}dt = \frac{N}{\sigma} \implies N = L\sigma \tag{3.2}$$

where L is the integrated luminosity,[13] specified in units of *inverse femtobarns*, or fb^{-1}. The total number of events is the product of the integrated luminosity and the collision cross section. At the LHC, 1 fb^{-1} corresponds to approximately 100 trillion (10^{14}) proton–proton

[12]The scattering cross section is essentially the interaction area. For example, if two identical balls with the same radius were launched toward each other, the scattering cross section would be the area πr^2. Since charged particles exert electromagnetic forces on each other over all distances, their cross section is a bit more complicated.

[13]Confusingly, the word "luminosity" is frequently used without specifying whether it refers to the instantaneous or integrated variety.

events. Equation 3.2 is important for the measurement of cross sections for various processes. If the integrated luminosity is known with high precision and the number of events has been accurately counted, then the cross section of a given process may be computed with high precision.

The instantaneous luminosity can be computed based on the accelerator configuration, as shown in Equation 3.3:

$$\mathcal{L} = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi \epsilon_n \beta^*} F \tag{3.3}$$

where N_b is the number of protons per bunch, n_b is the number of bunches per beam, f_{rev} is the frequency of revolution, γ_r is the Lorentz factor from relativity theory ($\gamma = 1/\sqrt{1 - v^2/c^2}$). The beam emittance, ϵ_n, is a parameter specifying the spread of a particle beam in both position and momentum. In a beam with low emittance, the particles occupy a small volume and have a low spread in momentum. The parameter β^* indicates the transverse size of the beam at the interaction point. The parameter F is a geometric correction factor that accounts for the nonzero crossing angle at the interaction point between the two beams, necessary to prevent "parasitic collisions" between beams outside of the interaction point. Further information about these parameters may be found in [61]. Table 3.1 shows the design values [61] and 2012 values [60], [62] of the LHC run parameters.

There is an alternate method of calculating the luminosity based on the beam geometry. It is given by Equation 3.4:

$$\mathcal{L} = \frac{N_b^2 n_b f_{rev}}{2\pi \Sigma_x \Sigma_y} \tag{3.4}$$

where N_b, n_b, and f_{rev} are as before, and Σ_x and Σ_y are the *root-mean-square* (rms) widths of the beam. The advantage of this method is that the rms widths of the beams may be measured accurately using the technique of beam-separation scans, or *van der Meer* (vdM) scans. The technique consists of fixing one beam and sweeping the other vertically and horizontally, in a manner akin to a chessboard. By measuring the rate of collisions at each separation, the size of the beam may be determined. vdM scans are performed routinely to ensure an accurate measurement of the instantaneous luminosity. More information may be found in [63].

3.2 The ATLAS Detector

ATLAS is one of the four main LHC experiments, located next to the main CERN campus at Meyrin. It is a general purpose detector, composed of an array of subdetectors enabling a variety of measurements of physical processes. The machine is cylindrical in shape, with length and diameter respectively 44 m and 25 m, and is designed to be symmetric with respect to the two counter-circulating beams. It weighs about 7000 metric tons. The detector is

LHC Run Parameter	Design Value	2012 Value
Beam energy [TeV]	7	4
Bunch spacing [ns]	25	50
Revolution Frequency [kHz]	11.245	11.245
Number of protons per bunch, N_b [$\times 10^{11}$]	1.15	1.6–1.7
Number of bunches per beam, n_b	2808	$\sim$1380
Lorentz factor, γ_r	7461	4263
Beam emittance, ϵ_n [μm]	3.75	2.6
β^* [m]	0.55	0.6
Crossing Angle [μrad]	285	290
Geometric Correction, F	0.836	$\sim$0.82
Beam size at IP, $\Sigma_{x,y}$ [μm]	16.7	19
Peak Luminosity, $\mathcal{L}$ [$\times 10^{34}$ cm^{-2} s^{-1}]	1	0.77

Table 3.1: Values of the run parameters of the LHC. The abbreviation IP refers to the interaction point. The 2012 value of the geometric correction is computed based on a formula in [60].

capable of measuring events very nearly to a solid angle of 4π by arraying smaller detectors in the shape of a barrel with two end-caps. A cut-away of the detector is shown in Figure 3.4.

The detector is constructed such that the beam pipe passes through the center. The beam pipe is constructed from beryllium with a thickness of 0.8 mm. Beryllium is chosen as the material because of its low ratio of atomic number to atomic mass, reducing the probability of *multiple scattering* by particles produced in the collision.[14] The detector system closest to the interaction point is the *Inner Detector* (ID), which is responsible for making measurements enabling accurate tracking of charged particles as well as measuring their momenta and charge. It consists of three subdetectors: the *Pixel Detector* (PIX), *Semiconductor Tracker* (SCT), and *Transition Radiation Tracker* (TRT). Located outside of the ID is a *solenoid* magnet that provides the magnetic field necessary for the ID to perform its momentum and charge measurements. Following the solenoid is an elaborate *calorimetry system* that stops electromagnetically and strongly interacting particles and measures their energy. Three *toroid* magnets are then placed to provide the magnetic field for the outermost system, the *muon system* (MS), responsible for measuring the kinematics of muons. The detector is structured such that the systems closer to the interaction point have higher precision than those farther away.

[14]Multiple scattering occurs when a particle "bounces off," or otherwise interacts many times with nuclei and/or electrons in the material, altering its trajectory. It is undesirable because it degrades information about the original kinematics of the particle.

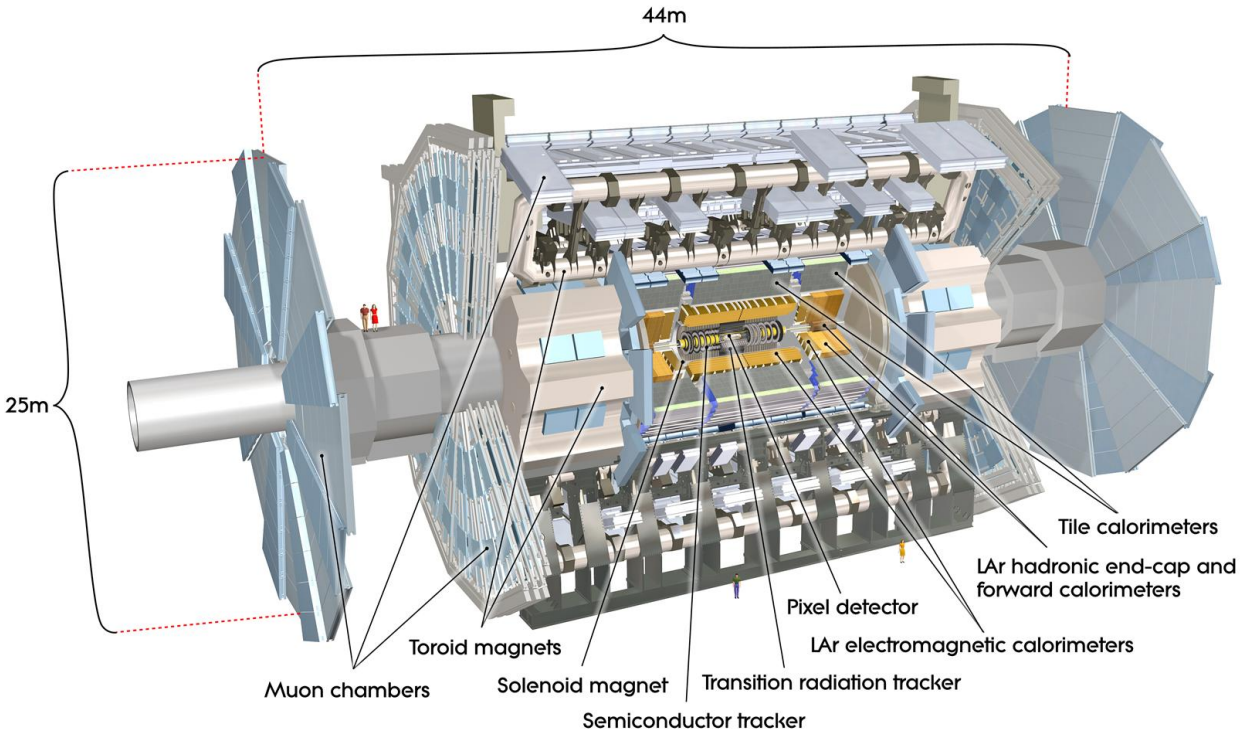


Figure 3.4: Cut-away depicting the ATLAS Detector [64].

During 2012 operations, the ATLAS experiment measured 20 million events per second. However, only 1000 of those events were able to be recorded each second for analysis. A multi-tier *trigger system* is implemented in both hardware and software to decide which of those events is worthy of being kept.

As mentioned in Section 3.1.4, accurate measurements of the luminosity are critical for determining the cross sections of various processes to high precision. Smaller forward detectors are installed for the purpose of measuring the luminosity.

Unless otherwise mentioned, more details about the following information presented may be found in [64].

3.2.1 Coordinates

ATLAS employs a right-handed coordinate system, defined with the origin located at the center of the detector, the x-axis pointing toward the center of the LHC ring, the y-axis pointing upward toward the Earth's surface, and the z-axis pointing along the direction of the counterclockwise beam. The xy-plane is called the *transverse plane*.

A polar coordinate system is also used. The polar angle θ is defined with respect to the positive z-axis, and the azimuthal angle ϕ is defined in the transverse plane with respect to

the positive x-axis.

Because many of the particles of interest are heavily boosted, it is useful to consider relativistic quantities such as the *rapidity*, well-suited for massive objects:

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z} \tag{3.5}$$

where y, E, and p_z are the rapidity, particle energy, and z-component of the particle momentum, respectively. For massless particles, Equation 3.5 reduces to a quantity called the *pseudorapidity*:

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right) \tag{3.6}$$

where η is the pseudorapidity.

Both rapidity and pseudorapidity share the property that differences in each quantity are *boost invariant*[15] along the z-axis. Pseudorapidity offers a geometric interpretation, so it is generally used to characterize the particle position instead of θ, which is not boost invariant. Another boost invariant quantity is ΔR, defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

Many quantities are specified with respect to the transverse plane (e.g. momentum). Such quantities will have a subscript "T". For example, a particle's momentum in the transverse plane will be denoted as p_{T}.

3.2.2 Inner Detector

The Inner Detector (ID) is located at the heart of ATLAS. Its primary objective is to record the information necessary to reconstruct the trajectory of charged particles exiting the interaction point. The detector is bathed in a 2 T magnetic field, enabling measurements of momentum based on the curvature of the reconstructed trajectories, or *tracks*. About fifty measurements are made of each charged particle traversing the ID. Pattern recognition software is utilized to reconstruct the tracks. Because there are about 1000 charged particles emerging from the proton–proton collisions in each bunch crossing, the ID is designed to operate in a densely populated environment. Its components have ultra-high resolution, or *granularity*, enabling track reconstruction with very high accuracy and precision. In addition to measuring momentum, tracks are used to identify *vertices*, locations from two or more charged particles originate. Vertices are found by extrapolating reconstructed tracks back to the beam pipe and looking for intersections of tracks. The ID tracking volume extends out to $|\eta| < 2.5$, and tracks are reconstructed for particles with $p_{\mathrm{T}} < 100$ MeV.

The ID consists of three subdetectors: the Pixel Detector (PIX), Semiconductor Tracker (SCT), and Transition Radiation Tracker (TRT). The PIX and SCT consist of silicon sensors doped to form a *pn*-junction and held at reverse bias. Passing charged particles induce charge in the sensor, resulting in a current registered as a *hit*. The TRT consists of *drift*

[15] Boost invariant quantities are independent of the frame of reference of the observer.

tubes interleaved with material stimulating production of transition photons. Since the presence of transition radiation depends on the mass of the traversing particle, the TRT can distinguish electrons from other heavier particles. Figure 3.5 shows a schematic drawing of the ID with the locations of the subdetectors with respect to the interaction point and their coverage in $|\eta|$. A cut-away depicting the ID subsystems is shown in Figure 3.6.

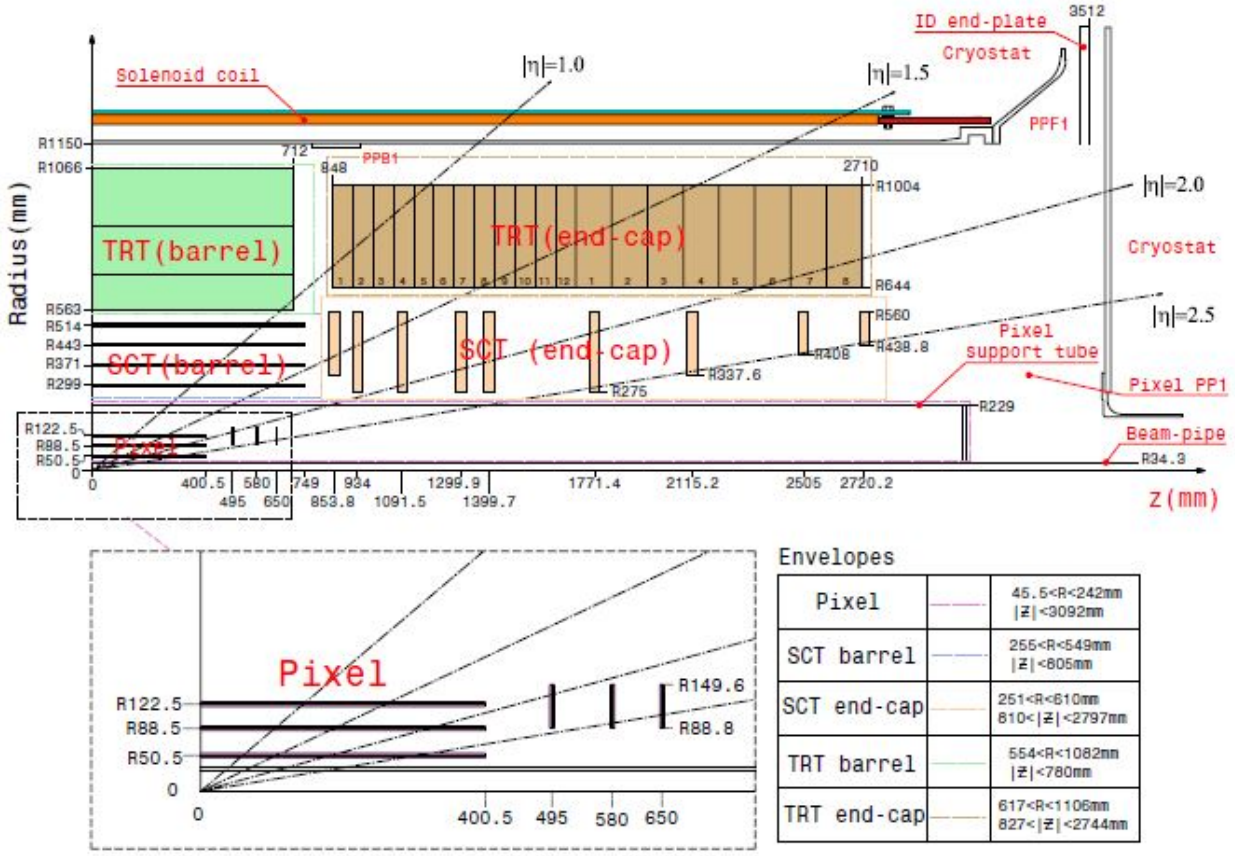


Figure 3.5: Quarter Schematic of the ATLAS Inner Detector [64].

Close proximity to the interaction point subjects the ID to the highest intensity of radiation, necessitating high requirements for performance. All sensors are *radiation-hard*, providing stable measurements despite being subjected to large doses of radiation. To reduce sensor electronic noise, even after radiation damage, the PIX and SCT are kept at a temperature between -5°C and -10°C. High structural stability is necessary to provide accurate tracking information. However, this stability has to be maintained with a modicum of material to reduce the likelihood of multiple scattering, which can compromise the integrity of tracking information due to the energy lost by the particle. A low material budget is also essential to reduce electron–positron pair production from photons (adding to the particle background) and the probability of an interaction between the ID and a neutral particle.

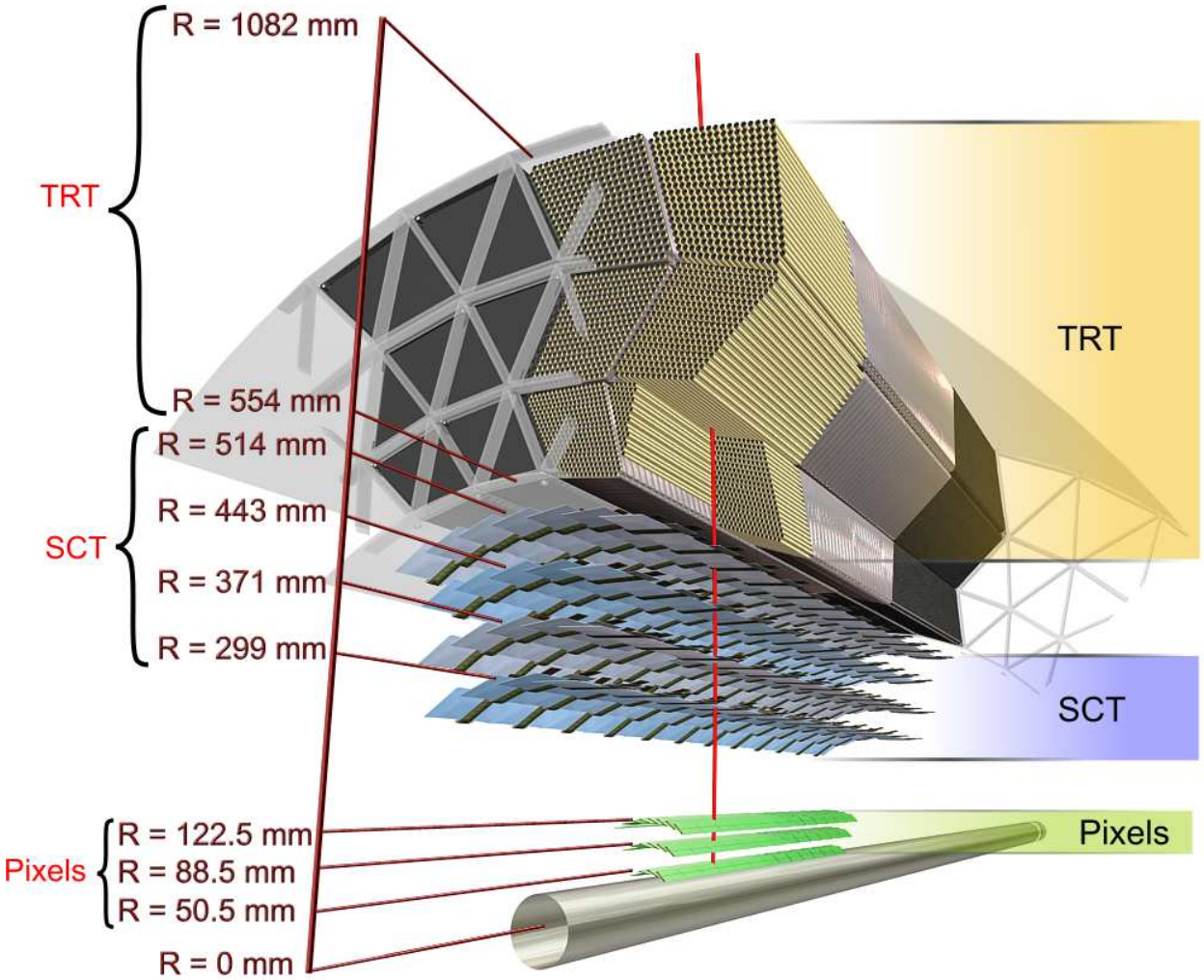


Figure 3.6: Cut-away depicting the ATLAS Inner Detector [64].

3.2.2.1 Pixel Detector

The Pixel Detector (PIX) is the closest subsystem of the ID to the interaction point. It consists of semiconductor sensors, *pixels*, connected to readout electronics to form *modules*, which are arranged in three cylindrical *barrel layers* and three annular *disk layers* (or *end-caps*) located on both sides of the barrel layers.[16] The innermost barrel layer is called the *B-layer*. The barrel and end-cap layers are illustrated in Figures 3.5 and 3.6. The barrel layers have radii of 50.5 mm, 88.5 mm, and 122.5 mm and extend out to 400.5 mm in the z-direction, while the disk layers are located along the z-direction 495 mm, 580 mm, and 650 mm on either side of the interaction point. The disk layers have inner and outer radii of

[16] A fourth barrel layer with radius 25.7 mm called the *Insertible B-Layer* (IBL) was installed over the long shutdown from 2013 and 2015 to accommodate the higher luminosity expected in higher energy runs. Details of its operation may be found in [65].

88.8 mm and 149.6 mm, respectively. This configuration allows for tracking up to $|\eta| < 2.5$. The detector occupies a volume equivalent to a cylinder with a length of 1442 mm and a diameter of 430 mm, as shown in Figure 3.7.

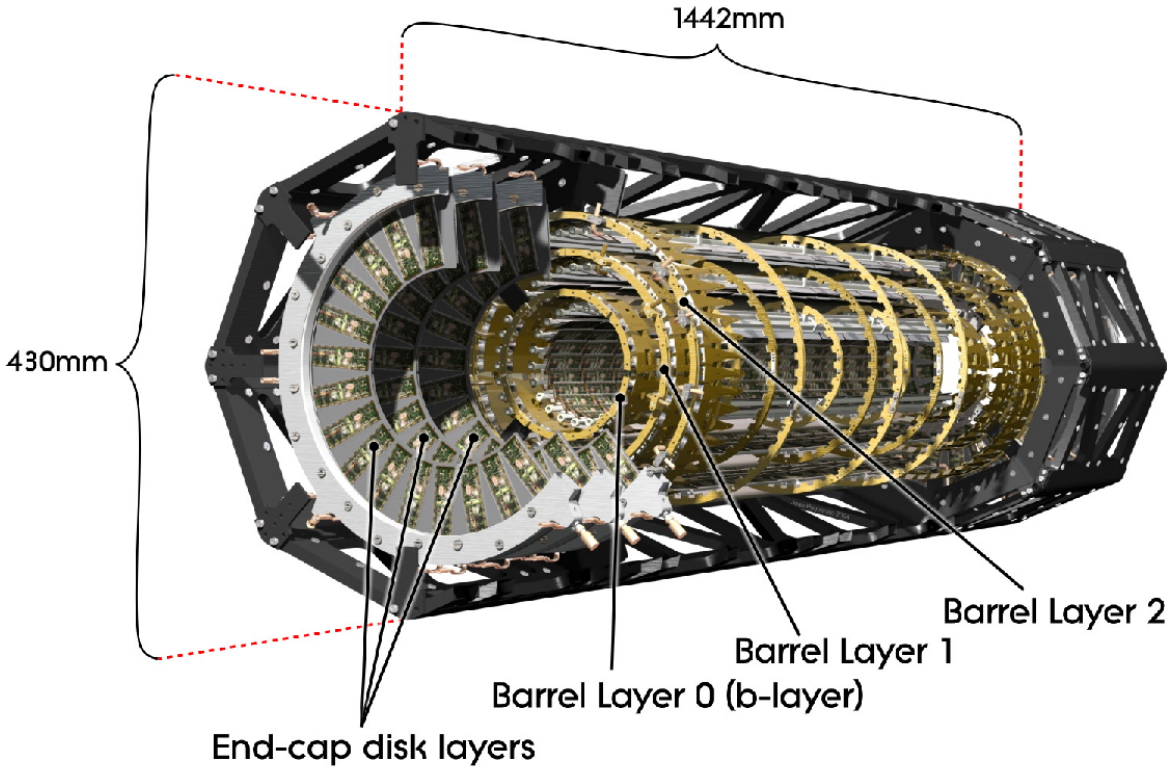


Figure 3.7: Cut-away depicting the ATLAS Pixel Detector [66].

The fundamental unit of the PIX is the pixel module, which consists of a silicon sensor containing 47232 pixels connected below by bump bonds to sixteen *front-end chips* that provide the electronic readout. These chips are affixed to a flexible printed circuit board. A *Master Control Chip* is bonded to the top of the sensor to enable communication between the front-end chips and other electronics. The dimensions of a module are 19 x 63 mm^2. There are 80.4 million pixels on 1744 such modules in the detector. Each pixel has a thickness of 250 μm, and $\sim$90% of the pixels have dimensions of 50 x 400 μm^2. The remaining $\sim$10% have dimensions of 50 x 600 μm^2 and are mostly located near the front-end chips. The modules are arranged in units called *staves* in the barrel layers, with thirteen modules per stave. In the disk layers, the modules are arranged in *sectors*, with six modules per sector.

The individual pixels consist of *n-type* silicon[17] doped on the front-end side with n-type

[17]This is silicon with impurities added to increase the concentration of electrons relative to holes. (Silicon

impurities and with *p-type* impurities[18] on the sensor side to form a *pn junction*. The readout side is doped with n-type material to maintain conductivity as the sensor is exposed to ionizing radiation. Over time, this radiation causes the n-type bulk to become p-type, but oxygen is added to the bulk to maintain the charge collection capability of the sensor even after type inversion.

The pixel sensor is kept under reverse bias[19] at 150 V, though this voltage will increase to 600 V over time (∼10 years) to compensate for radiation damage. When a charged particle passes through, or near, the sensor, it ionizes the sensor's material, resulting in the creation of electron–hole pairs that respectively drift toward the positive and negative ends of the sensor. These electrons are collected and measured by readout electronics. If the number of ionization electrons exceeds a threshold value (nominally 3500), then the pixel registers a hit. Contiguous pixel hits are called *clusters*, and the average number of ionization electrons in a cluster is about 20000 [67].

Because the PIX is closest to the interaction point, it has the highest granularity of the detectors in ATLAS. The dimensions of each of the pixel sensors provide a resolution of 10 μm in the transverse plane and 115 μm in the z-direction. The high resolution in the transverse plane is necessary because the magnetic field in the ID bends particles in this plane. The measurement in the z-direction is important for reconstructing vertices, which tend to be spread out in this direction. Consequently, the longitudinal resolution is looser compared to that needed for momentum measurements in the transverse plane.

3.2.2.2 Semiconductor Tracker

The Semiconductor Tracker (SCT) is the next subsystem in the ID, located right outside of the PIX. It consists of large silicon microstrip sensors interfaced with readout electronics to form modules. Like the PIX, these modules are arranged in four barrel layers and nine disk layers in each end-cap. The barrel layers have radii of 299 mm, 371 mm, 443 mm, and 514 mm and extend out to 749 mm in the z-direction. The disks are located at distances of 853.8 mm, 934 mm, 1091.5 mm, 1299.9 mm, 1399.7 mm, 1771.4 mm, 2115.2 mm, 2505 mm, and 2720.2 mm in the z-direction, have varying inner radii of 275 mm, 337.6 mm, 408 mm, and 438.8 mm, and have an outer radius of 560 mm. Figure 3.5 indicates which disk has the given inner radius. This configuration allows the SCT to track charged particles for $|\eta| < 2.5$.

The SCT has 4088 modules, of which 2112 are located in the barrel layer, and roughly 6.3 million readout channels. SCT modules are double-sided, containing four silicon sensors

atoms can have up to eight valence electrons—any absence of these electrons is called a hole. Silicon atoms have four electrons and four holes.)

[18] n-type and p-type impurities respectively increase the concentration of electrons and holes.

[19] Reverse bias means keeping the p side of the pn-junction at negative voltage and the n side at positive voltage. This creates a large *depletion region* with no excess of electrons or holes, hence no current flows through.

(two per side). Like the PIX, these sensors are comprised of *p-in-n* silicon.[20] A sensor is composed of 768 strips with a thickness of 285 μm, 6 cm in length, and with a *pitch* of 80 μm.[21] The sensors on each side of the module are connected, resulting in a total length of 12 cm and are also affixed such that there is a 40 mrad *stereo angle* between the forward and rear pairs of sensors,[22] with one layer parallel to the beam line, to improve resolution in the z-direction. The larger sensor size is required due to the increased distance from the interaction point and to keep the channel count reasonable. The sensors are *AC–coupled* to the read-out electronics to reduce noise.[23] Mechanical support of the module is furnished by a graphite baseboard. Additionally, the modules in the barrel layer are tilted at an angle of $\sim$11° in the transverse plane to improve resolution in the ϕ direction. The disk sensors are arranged so that each strip is constant in ϕ and with the same pitch as the barrel modules. All sensors operate at an initial voltage of 150 V, but the voltage will be increased to 250–350 V in the future to maintain good charge collection capability in the face of radiation damage.

A charged particle passing through the SCT modules usually results in hits on two sensors due to the double-layer design. The average traversing particle registers eight hits, two for each barrel layer. The SCT has excellent resolution in both transverse and z-directions, 17 μm and 580 μm respectively.

3.2.2.3 Transition Radiation Tracker

The Transition Radiation Tracker is the outermost subdetector in the ID and is comprised mostly of drift tubes rather than silicon sensors. These drift tubes, known as *straws*, form the fundamental unit of the TRT. It has the fewest number of readout channels (351000) of the three detectors. The TRT is kept at room temperature during operation since it is relatively far from the interaction point.

Before describing the components of the TRT, the principles of drift tubes and transition radiation will be reviewed. A drift tube is a device consisting of an electrically conductive cylindrical shell with a wire at the center. The device is filled with a prepared gas mixture and sealed at both ends. A high voltage difference between the wire and shell is applied, with the wire usually kept at positive polarity.[24] Charged particles passing through the tube ionize the gas, resulting in free electrons that accelerate toward the wire, colliding with other gas atoms along the way. These collisions produce more ionized gas and free electrons, resulting in a cascade. The collected electrons signal that a charged particle passed through.

Relativistic charged particles passing through media with different indices of refraction emit transition radiation with energies proportional to the Lorentz factor of the original

[20] That is, the silicon is n-type with one side doped with n-type impurities, while the other is doped with p-type impurities.

[21] The pitch is the distance between the centers of adjacent strips.

[22] The stereo angle is a rotation angle between two parallel planes. Think of laying two pieces of paper flat on top of each other, then rotating one slightly.

[23] AC–coupling means that only AC (as opposed to DC) signals can be read out.

[24] Since the electrons migrate toward this wire, it is called the *cathode*. The shell is called the *anode*.

particle and are emitted at angles of $1/\gamma$ with respect to the particle trajectory. High energy transition photons entering the drift tube can convert into electron–positron pairs, resulting in a large cascade of electrons. Low energy photons are absorbed by the gas atoms, potentially resulting in further ionization. Since the Lorentz factor of a particle is inversely proportional to its mass, lighter particles are more likely to emit high-energy transition photons than heavier ones. Thus, the presence of transition radiation can provide a clue to the identity of the traversing particle.

The straw tube is at the heart of TRT operation. Each straw is 4 mm in diameter. The walls of the tube consist of two 35 μm films bonded together. The base layer of each film is a 25 μm thick polyimide layer[25] with a 0.2 μm aluminum coating on one side. A 5–6 μm graphite–polyimide layer is then applied to this coating for protection. On the other side of the base layer, a 5 μm polyurethane layer[26] is applied. The two films are put into contact with their polyurethane layers facing each other and bonded using heat, resulting in a total tube thickness of 70 μm. Carbon fibers are used to provide additional mechanical support. The electrical resistance of the tube is less than 300 Ω/m, and it is kept at a voltage of -1530 V during operation. At the center of each tube is a 31 μm diameter tungsten wire with a gold plating 0.5–0.7 μm thick. This wire has a resistance of 60 Ω/m, is kept electrically grounded, and is connected directly to the front-end electronics of the straw located at each end. Proper operation requires that the wire be offset from the center of the tube by no more than 300 μm. The sag of the wires is less than 15 μm. The tube is filled with a gas mixture of 70% xenon, 27% carbon dioxide, and 3% oxygen.

TRT straws are placed among a matrix of polypropylene fibers[27] 19 μm in diameter to form a module. These fibers serve as the transition radiation material. The straws in each module are placed $\sim$7 mm apart from one another and are cut to 144 cm in length.

The TRT is also configured into barrel and end-cap regions. There are three barrel layers, spanning radii from 563 mm to 1066 mm and extending out to 712 mm in the z-direction. Each layer of the barrel contains 32 modules, aligned to be parallel to the beam pipe. There are 73 straw layers in the barrel. The end-caps each contain two independent sets of wheels located between 848 mm and 2710 mm in the z-direction and with inner and outer radii of 644 mm and 1004 mm, respectively. The inner (outer) set of wheels consists of 12 (8) wheels, each consisting of eight straw layers spaced 8 mm (15 mm) apart. There are 768 straws 37 cm in length, oriented in the radial direction and spaced uniformly in ϕ, located in each layer. The straws are interleaved with polypropylene radiator foil layers 15 μm thick with a polypropylene net used to separate the layers, serving as the transition radiation material in the end-caps.

The TRT provides tracking coverage up to $|\eta| < 2.0$ in this configuration, with at least 36 straws traversed by a charged particle. However, particles passing through the barrel–end-cap transition region, located between $0.8 < |\eta| < 1.0$, will traverse at least 22 straws.

[25]Polyimide is a polymer that is used on insulation for spacecraft and water purification by reverse osmosis.

[26]Another polymer, polyurethane is a primary component of Spandex, tennis grips, and automobile seats.

[27]Polypropylene is another polymer with various uses, including packaging material, thermal underwear, automotive parts, and currency.

Because of the long lengths of the straws, the TRT only measures the transverse components (R, ϕ) of a particle's motion in the barrel (ϕ only in the end-caps). These measurements have a resolution of 130 μm, requiring that the position of every TRT tube wire be known to within 50 μm.

The TRT is also used for electron identification. Since electrons are the lightest charged particles, they emit transition photons with higher energy than other charged particles of equal energy. Two thresholds of hits are defined based on the amount of charge collected by the straw: low and high. Since transition photons can result in additional ionization to that produced by the charged particle, high threshold hits are typically produced by electrons. Electrons with energies of 2 GeV or more typically produce 7–10 high threshold hits.

3.2.3 Magnet Systems

The ATLAS magnet systems provide the magnetic fields that bend charged particles, allowing measurements of particle momentum. They consist of a central solenoid magnet that provides the magnetic field to the ID and three toroid magnets that provide the magnetic field to the muon system (MS).[28] One toroid magnet is dedicated to the barrel region of the MS, while the other two are dedicated to each end-cap. A diagram of the magnet system is shown in Figure 3.8. As discussed in Section 3.1.3, the trajectories of charged particles are curved in the presence of magnetic fields. The measured radius of curvature depends on the particle's momentum and the magnetic field strength. Consequently, the magnetic field must be known to high precision to have an accurate momentum measurement.

The central solenoid provides the magnetic field that bends particles in the ID. The field has a strength of 2 T and is aligned with the solenoid axis. A profile is shown in Figure 3.9a. The solenoid has dimensions of 5.8 m in length, and 2.46 m and 2.56 m for the inner and outer diameters, respectively. It is wound with 1154 turns of NbTi cables consisting of 12 strands, each having a diameter of 1.22 mm. Aluminum is added to the cables to provide further stability, with additional mechanical support furnished by an aluminum support structure 12 mm thick. The total assembly has a mass of 5.7 metric tons. The solenoid is kept at a temperature of 4.5 K using liquid helium, allowing it to operate in a superconducting state. The nominal current during operation is 7.73 kA. If a magnet quench occurs, the magnet's energy is discharged into the cold mass, raising its temperature to less than 120 K, cooled within a day to 4.5 K. The solenoid is encased in a *vacuum vessel*, similar to a Dewar flask or Thermos, to keep the temperature low.

Because the solenoid is located in front of the calorimeters, it is important that the amount of its material be as low as possible. Particles can interact with material in the solenoid, leading to production of additional particles as well as inaccurate calorimeter energy readings. The solenoid and liquid argon calorimeter share the same vacuum vessel to reduce

[28] A solenoid magnet consists of a cylinder with wire wrapped around it, providing reasonably uniform magnetic field inside. A conventional toroid magnet consists of a doughnut with wire wound around it in a manner similar to a solenoid. The ATLAS toroid magnets are a bit of a misnomer, as they are *not* of this kind.

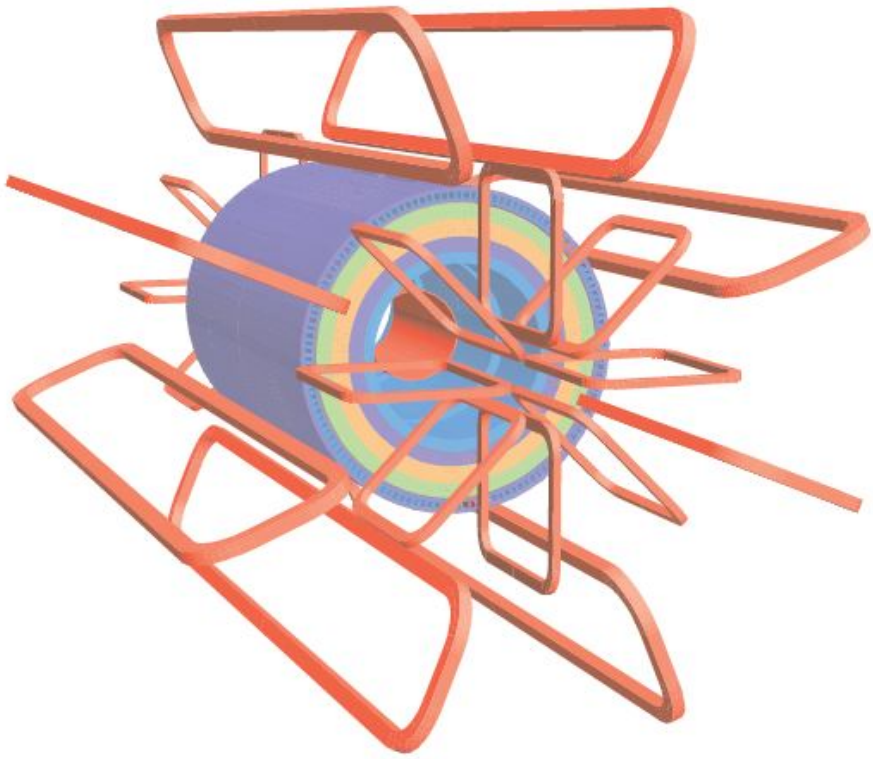


Figure 3.8: The ATLAS toroid and solenoid magnets [64].

the material budget. The magnetic flux generated is returned by the infrastructure of the hadronic calorimeter. The stored energy-to-mass ratio of 7.4 kJ/kg and the contribution of ∼0.66 *radiation lengths*[29] indicate successful compliance with the low material design requirements.

The toroid magnets provide magnetic fields of 1.0 T and 0.5 T to the MS barrel and end-cap detectors, respectively. A profile is shown in Figure 3.9b. The magnets each consist of eight individual coils with 120 (116) turns of NbTi wire in the barrel (end-caps), encased in stainless-steel vacuum vessels. The barrel toroid is 25.3 m in length, with inner and outer diameters of 9.4 m and 20.1 m, respectively, with a total mass of 830 metric tons. Each end-cap toroid is 5.0 m in length, with inner and outer diameters of 1.65 m and 10.7 m, and weighs 239 metric tons. The coils are cooled to 4.5 K using liquid helium, making them superconductive. The end-cap toroid coils are rotated 22.5° with respect to those of the barrel toroid to maintain the magnetic field strength at the barrel–end-cap boundary.

[29]The radiation length of a material is the average distance that a traversing charged particle loses all but $1/e$ of its energy.

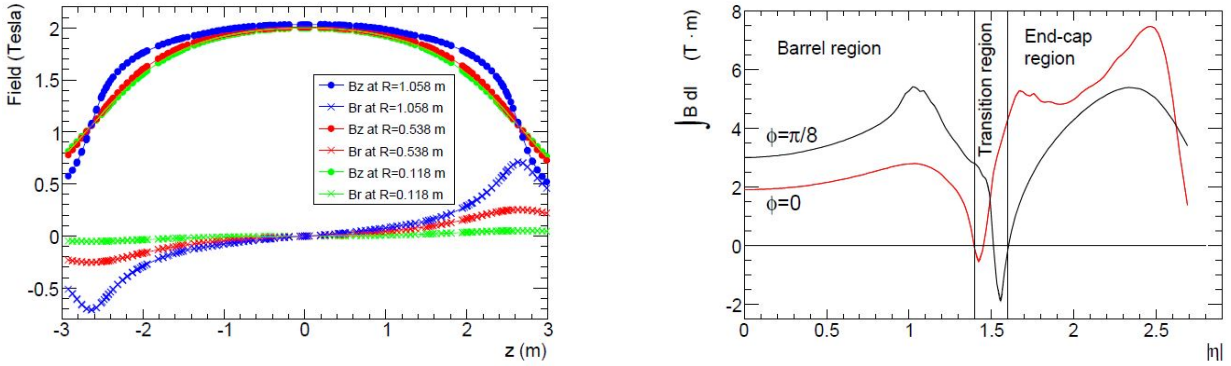


Figure 3.9: Magnetic field strength in the Inner Detector for fixed ϕ (left) and in the Monitored Drift Tubes in the Muon System (right) [64].

3.2.4 Calorimeters

The next major system of detectors outside of the ID and solenoid are the *calorimeters*. While the ID makes measurements of charged particles without disturbing their motion, the calorimeters are designed to stop particles, allowing for measurements of their energy. Consequently, the calorimeters have a large material budget.

The ATLAS calorimeters are *sampling calorimeters*. In such a calorimeter, a particle enters and interacts with the *active material*, stimulating production of additional particles that also interact to produce further particles, depending on the energy of the incident particle. This cascading chain is called a *shower*. The showering particles ionize and eventually come to rest inside *passive material* connected to instrumentation that measures the amount of ionization, which is proportional to the energy of the incident particle. The calorimeters are designed so that the showers are contained within their volume. Particles that escape the calorimeter are said to *punch-through* it.

There are two subsystems of calorimeters with different purposes. The inner subsystem is an *electromagnetic* (EM) *calorimeter*, while the outer subsystem is comprised of the *hadronic calorimeters*. The EM calorimeter is designed to stop particles that interact electromagnetically, mainly electrons and photons, using liquid argon (LAr) as the active material. The hadronic calorimeters stop and measure strongly interacting particles, particularly pions, neutrons, and protons. The hadronic calorimeters use LAr and scintillating tiles[30] as the active material. The EM calorimeter consists of barrel and end-cap (EMEC) LAr calorimeters and has coverage for $|\eta| < 3.2$. One of the prominent features of the EM calorimeter is fine granularity to resolve photons and electrons in high detail. The hadronic calorimeters consist of a tile calorimeter with extended barrel layers and a LAr end-cap calorimeter (HEC), which make measurements for $|\eta| < 3.2$. A LAr *forward calorimeter* (FCAL) with one EM and two hadronic layers extends the calorimeter coverage to $|\eta| < 4.9$, improving

[30]So-called because ionizing particles passing through stimulate production of photons.

the overall coverage of the ATLAS detector and enabling accurate measurements of jets and missing transverse energy. The hadronic calorimeters have coarser granularity compared to the EM calorimeters because of the larger volume occupied by jets, compared to electrons and photons. Figure 3.10 shows a profile of the calorimeter systems, and Figure 3.11 shows a distribution of their material budget.

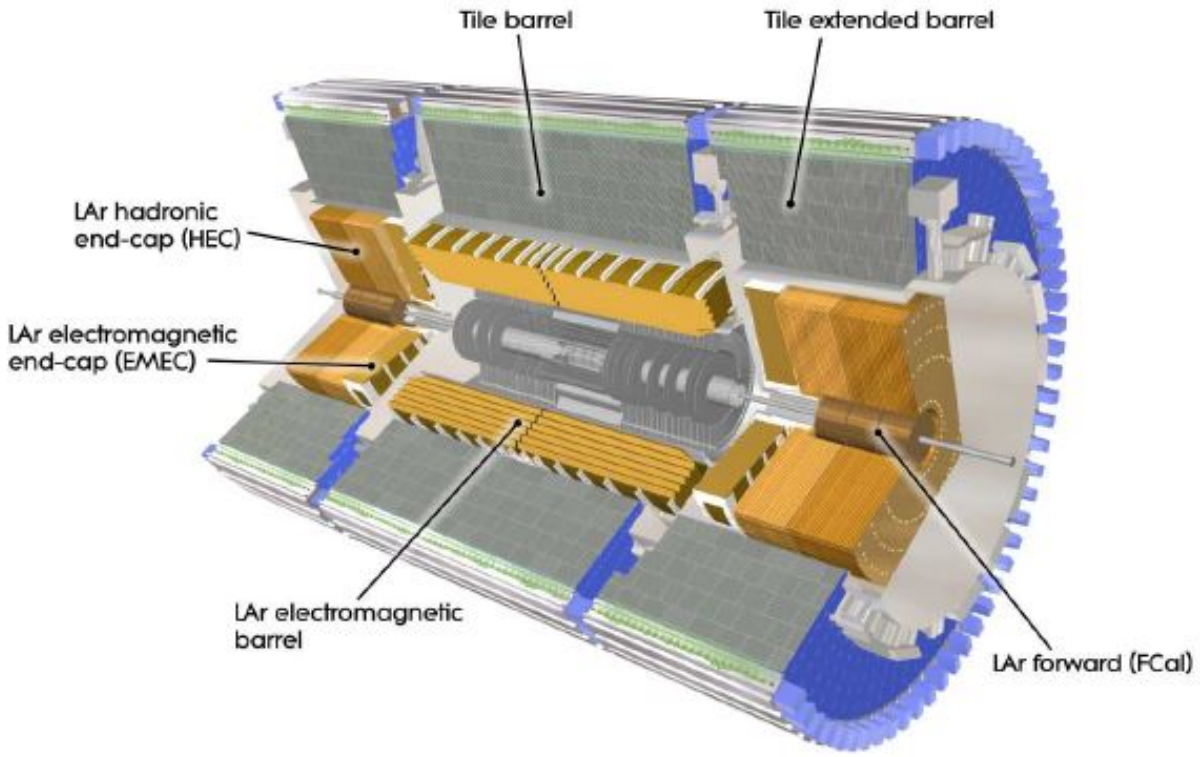


Figure 3.10: Cut-away depicting the ATLAS calorimeters [64].

3.2.4.1 Electromagnetic Calorimeter

The electromagnetic (EM) calorimeter is the innermost calorimeter and measures the energy of electromagnetically interacting particles, primarily photons and electrons. These particles interact with nuclei and electrons, converting to electron–position pairs or producing *bremsstrahlung*[31] photons. Thus, EM showers consist of photons and electrons. The basic unit of distance in the EM calorimeter is the *radiation length* X_0, defined as the mean distance through a material in which a charged particle loses energy through ionization such

[31]Bremsstrahlung is the process where an electron is deflected by another matter particle, emitting a photon in the process.

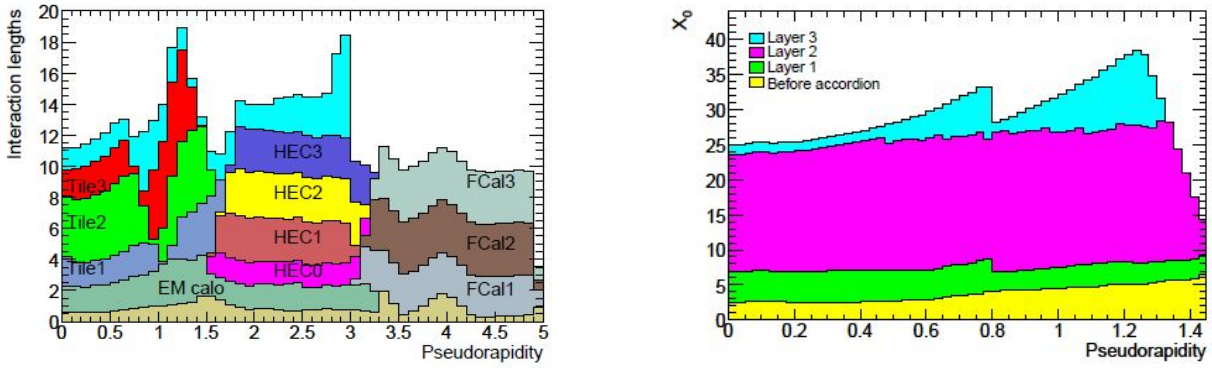


Figure 3.11: Material budget of the electromagnetic and hadronic calorimeters in terms of interaction lengths (left) and radiation lengths (right) [64].

that it only has $1/e$ times the original amount. The EM calorimeter is structured so that it has a thickness of at least 22 radiation lengths.

The EM calorimeter uses LAr as the active material in both the barrel and end-cap regions, with lead sheets arranged in an accordion structure serving as the passive material. The temperature of the LAr is between 88.5 K and 88.6 K, and must not vary by more than 100 mK due to the temperature-dependent performance of LAr [68]. A schematic picture of a section of the barrel is shown in Figure 3.12. The accordion geometry is chosen to ensure full coverage in ϕ and enable fast signal extraction in the front and rear of the electrodes. The absorber sheets are constructed from lead sheets that have thicknesses of 1.53 mm and 1.13 mm for $|\eta|$ less than and greater than 0.8, respectively, in the barrel. In the end-caps, the sheets are 1.7 mm and 2.2 mm thick for $|\eta|$ less than and greater than 2.5, respectively. The difference in thicknesses prevents a decrease in the *sampling fraction*[32] with the increase in $|\eta|$. Mechanical support is furnished by the addition of two 0.2 mm stainless-steel sheets affixed to each side. Electrodes consisting of three conductive copper layers insulated with polyimide sheets are interleaved between the absorber sheets to provide readout. The outer two electrodes in the barrel are kept at 2000 V, while the signal is read out of the middle electrode. In the end-caps, the voltages vary by $|\eta|$ between 1000 V and 2500 V.

In the EM calorimeter, the barrel is divided into two equal halves with a small 4 mm gap in-between at $z = 0$. The barrel region covers $|\eta| < 1.475$. Each half-barrel has a length of 3.2 m, inner and outer diameters of 2.8 m and 4 m, and a weight of 57 metric tons. 1024 accordion-shaped absorbers with readout electrodes are contained within each half-barrel and are divided into 16 modules, each consisting of three layers of varying granularity. The first layer, known as the *strip layer*, has a $\Delta\eta \times \Delta\phi$ granularity of 0.0031×0.098, allowing for precise photon resolution from processes such as neutral pion decays. The second layer $\Delta\eta \times \Delta\phi$ granularity is 0.025×0.0245, and the third is 0.0245×0.05. Figure 3.5 indicates

[32]The sampling fraction is a parameter inversely proportional to the calorimeter noise. Consequently, having high resolution requires keeping the sampling fraction high.

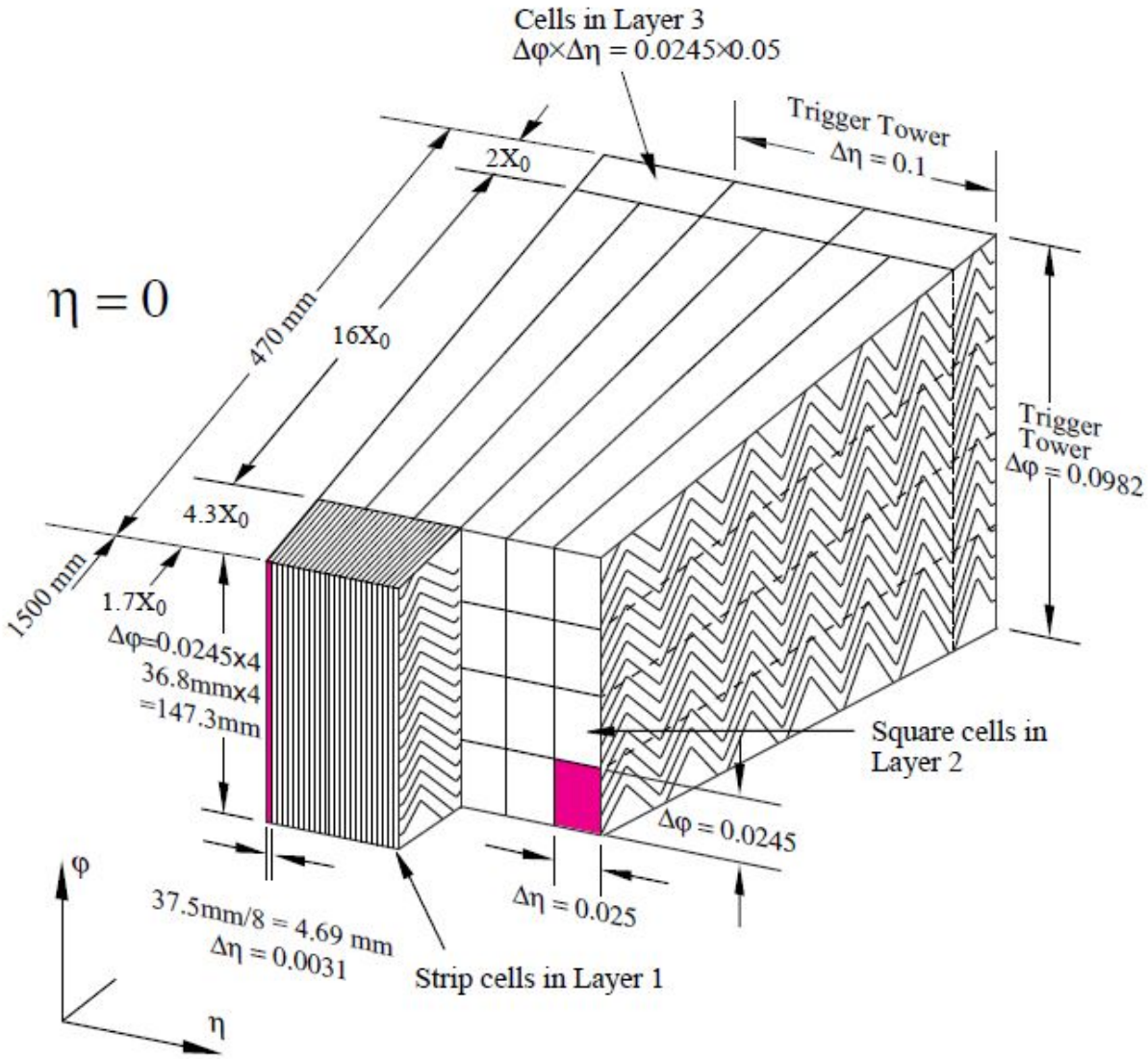


Figure 3.12: Schematic illustrating different absorber layers in the barrel of the ATLAS electromagnetic calorimeter [64].

specific dimensions of the layers. The depth of each module varies from 22 to 33 radiation lengths depending on $|\eta|$ position in the barrel. A *pre–sampler* layer of LAr, 11 mm in depth with instrumentation, is situated before the strip layer for $|\eta| < 1.8$ to measure the energy lost by particles before reaching the EM calorimeter.

The EMEC calorimeters are wheels 63 cm thick and weighing 27 metric tons, located on each side of the barrel LAr calorimeter. Each wheel has internal and external radii of 330 mm and 2098 mm, respectively, at ambient temperature. The wheels provide coverage

for $1.375 < |\eta| < 3.2$, overlapping with the barrel. In this overlap region, a presampler spanning $1.5 < |\eta| < 1.8$ accounts for energy lost in the barrel calorimeter. Each wheel is further subdivided into two coaxial wheels at $|\eta| = 2.5$ and eight wedge-shaped modules. The outer wheels each contains 768 absorbers with read-out electrodes, while the inner wheels each have 256.

The EM calorimeter has excellent performance in terms of resolution. Though required to have a resolution of $\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$,[33] dedicated test-beam studies of the barrel measured the constant term to be 0.17% for $\eta = 0.687$ [64]. Tests of the EMEC showed similar results. The detector response was also found to be linear for particles with energies larger than 15 GeV. Tests of the polar angle resolution of the first and second layers showed that it varied as (50–60 mrad)$/\sqrt{E}$ (GeV).

3.2.4.2 Hadronic Calorimeters

The hadronic calorimeters are located outside of the EM calorimeters and are designed to stop and measure the energies of strongly interacting particles. These interactions occur between the particles and nuclei of the active and passive material in the calorimeter. Hadrons interacting with nuclei often produce other hadrons, resulting in a shower. LAr and scintillating tiles are used as the active material and are measured in *interaction lengths* λ, a quantity describing the energy loss of traversing hadrons, analogous to interaction lengths. Figure 3.11 shows the depth of the calorimeter in interaction lengths as a function of pseudorapidity.[34] The hadronic calorimeter is designed to fully contain hadronic showers.

The barrel region of the hadronic calorimeter is called the tile calorimeter, providing coverage for $|\eta| < 1.7$ and consisting of central ($|\eta| < 1.0$) and extended regions on each side ($0.8 < |\eta| < 1.7$). These have lengths of 5.8 m and 2.6 m, respectively, and inner and outer radii of 2.28 m and 4.25 m. This region has a thickness of about 7.4 interaction lengths in the radial direction. Each barrel contains 64 modules consisting of alternating steel plates and scintillating tiles connected to a supporting girder that are radially oriented in the calorimeter, each subtending about 0.1 rad in ϕ. A diagram is shown in Figure 3.13. The 4 mm steel plates are glued to a 5 mm master plate affixed to the girder. The scintillating tiles are 3 mm thick and made of polystyrene, doped with wavelength-shifting compounds. Each tile is encased in a protective plastic sleeve that also keeps scintillation photons from escaping the tile. The ratio of steel plates to scintillating tiles is 4.7:1 by volume. Hadrons passing through these tiles stimulate the production of ultraviolet photons that are then modulated into the visible spectrum by 1 mm thick wavelength shifting fibers connected to each scintillator. These fibers are attached to photomultiplier tubes (PMTs)[35] located in the girder that are wired to read-out instrumentation. The measured signals are indicative of the shower energy. The detector granularity is specified for three radial sampling depths, each

[33] The symbol $\oplus$ indicates that the quantities are added in quadrature.

[34] It should be noted that the interaction lengths for the EM calorimeter are also listed in this figure. This is because hadrons also strongly interact with material in the EM calorimeter.

[35] A photomultiplier tube is a device that generates an electric current when struck by photons.

roughly 1.5, 4.1, and 1.8 interaction lengths thick at the center of the detector. The first two layers have granularity $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$, and the third layer has $\Delta\eta \times \Delta\phi = 0.2 \times 0.1$.

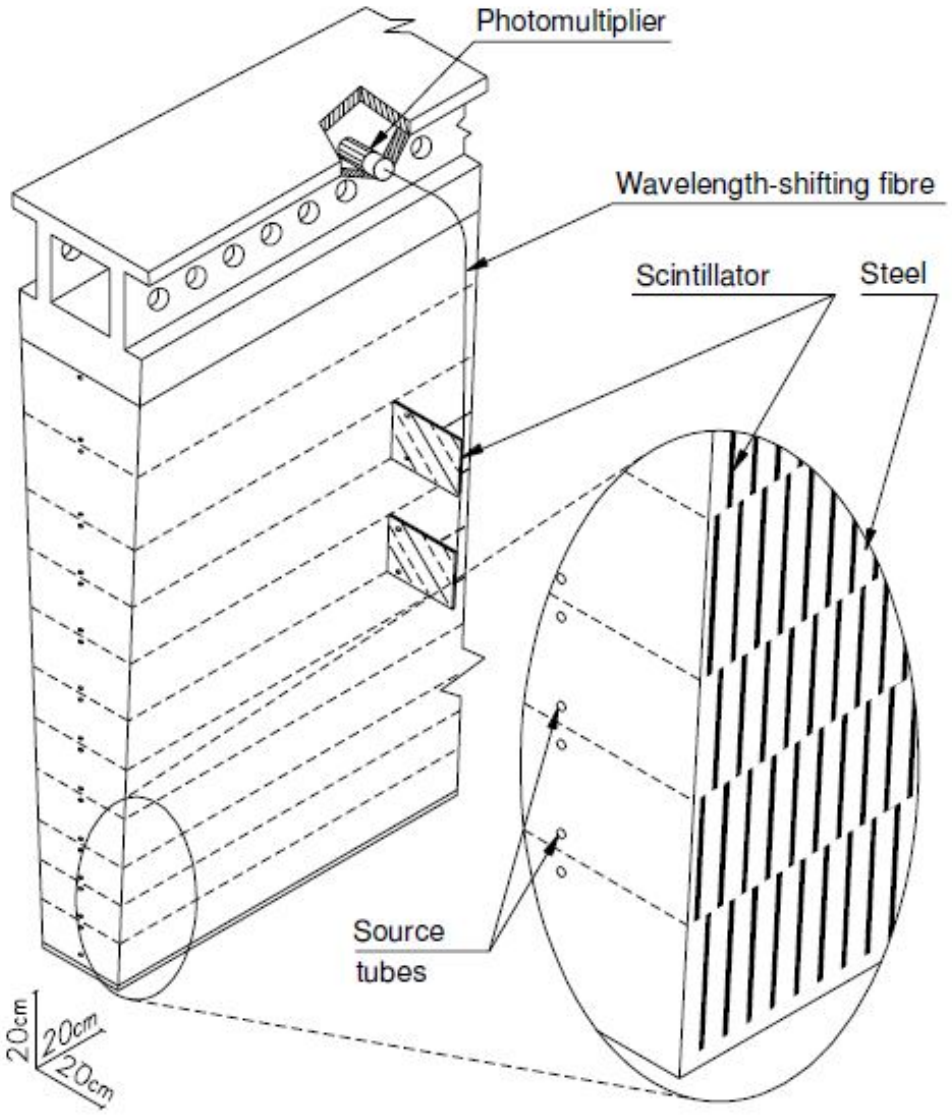


Figure 3.13: Schematic of a tile module in the ATLAS hadronic tile calorimeter [64].

The hadronic end-cap calorimeters (HEC) are located outside of the EMEC and cover the region $1.5 < |\eta| < 3.2$. The HEC uses LAr as the active material, has copper absorbers, and is divided into two wheels, a front and rear, each containing 32 modules. The modules in the front wheel consist of twenty-four 25 mm thick copper plates with a front plate 12.5 mm in thickness, while those in the rear wheel have sixteen 50 mm thick copper plates with a front plate 25 mm in thickness. The plates are assembled with gaps of 8.5 mm in-between.

Three electrodes are situated in each gap, providing LAr drift zones 1.8 mm in thickness. The middle electrode is used for read-out, while the other two are kept at 1800 V. The granularity of the HEC is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ for $|\eta| < 2.5$ and 0.2×0.2 otherwise.

The outermost calorimeter systems are the forward calorimeters (FCAL), designed to cover $3.1 < |\eta| < 4.9$. This extreme η range furthers the goal of making the detector as hermetic as possible, important especially for missing transverse energy measurements. Each FCAL is divided into three modules consisting of layers of perforated plates fitted with electrodes, each 45 cm in depth. The inner module is used for EM calorimetry, while the outer two modules are devoted to hadronic measurements. All use LAr as the active material. The plates in the first layer are made of copper and spaced apart 0.269 mm. The electrodes are formed from copper tubes with coaxial copper rods surrounded by plastic. The FCAL's relative proximity to the interaction point (4.7 m), subjects it to high particle fluxes, requiring that the plastic be radiation-hard. The small gap spacing also prevents ion build-up in the LAr. The two hadronic modules each use two copper end-plates 2.35 cm in thickness and are filled with electrodes similar to those in the EM module, but made from tungsten instead of copper. Tungsten slugs are placed in the spaces between the electrodes. Each module operates at a different voltage: 250 V (inner), 375 V (middle), and 500 V (outer). The amount of material in the FCAL is roughly $\sim$10 interaction lengths.

The hadronic calorimeter is designed to have energy resolutions of $\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$ for the tile and HEC calorimeters and $\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$ for the FCAL. Test-beam studies with pions [64] measured a resolution of $\sigma_E/E = 70.6\%/\sqrt{E} \oplus 5.8\%$ for the HEC and $\sigma_E/E = 70\%/\sqrt{E} \oplus 3.0\%$ for the FCAL after a reweighting procedure was applied. The energy resolution of the tile calorimeter was measured to be $\sigma_E/E = 56.4\%/\sqrt{E} \oplus 5.5\%$, but is expected to degrade by about 10% over the first ten years of LHC operations based on irradiation tests of the tile assembly.

3.2.5 Muon Systems

The muon system (MS), collectively also known as the *muon spectrometer*, comprises the outermost detector in ATLAS and is designed to measure the momenta of charged particles emerging from the calorimeter system. Such particles are likely muons for the following reasons: muons hardly lose energy by bremsstrahlung due to their large mass, are in a minimally ionizing regime when they have momenta on the GeV-scale (which applies to those produced at the LHC), and are not strongly interacting. Consequently, muons barely interact with the active material of the calorimeter system, passing through largely unimpeded. The calorimeter depth limits the punch-through of other particles.

The MS is bathed in a magnetic field generated by the toroid magnet system. The trajectories of entering muons are bent, enabling measurements of their momenta. The barrel toroid provides the field exclusively covering $|\eta| < 1.4$, while the end-cap toroids cover $1.6 < |\eta| < 2.7$. The barrel and end-cap toroid fields overlap for $1.4 < |\eta| < 1.6$, called the transition region. While the solenoid field is aligned with the beam pipe, the toroid field is oriented in the transverse plane, curling to make a roughly circular shape. Consequently,

entering muons bend in the η-direction. The MS is designed to measure momenta with high resolution, required to be $\sigma_{p_T}/p_T = 10\%$ for 1 TeV muons.

The MS is comprised of four subdetector systems: Monitored Drift Tubes (MDTs), Cathode Strip Chambers (CSCs), Resistive Plate Chambers (RPCs), and Thin Gap Chambers (TGCs). Each has a different purpose. A cut-away depicting the MS is shown in Figure 3.14, while schematics are shown in Figure 3.15. Each subdetector contributes to the MS coverage of $|\eta| < 2.7$ and allows for muon triggering for $|\eta| < 2.4$. Specifically, the MDTs and CSCs provide tracking information, while the RPCs and TGCs offer triggering and identification of the proton bunch-crossing associated with each muon. The muon chambers in the barrel region consist of concentric cylindrical shells aligned along the beam pipe, having approximate radii of 5 m, 7.5 m, and 10 m, while those in the end-cap region are large wheels aligned in the transverse plane at approximate distances in $|z|$ of 7.4 m, 10.8 m, 14 m, and 21.5 m. There is a small gap in muon coverage at $|\eta| = 0$, allowing for placement of calorimeter support infrastructure. The MS is also designed to tolerate the background radiation emanating from the walls of the ATLAS cavern.

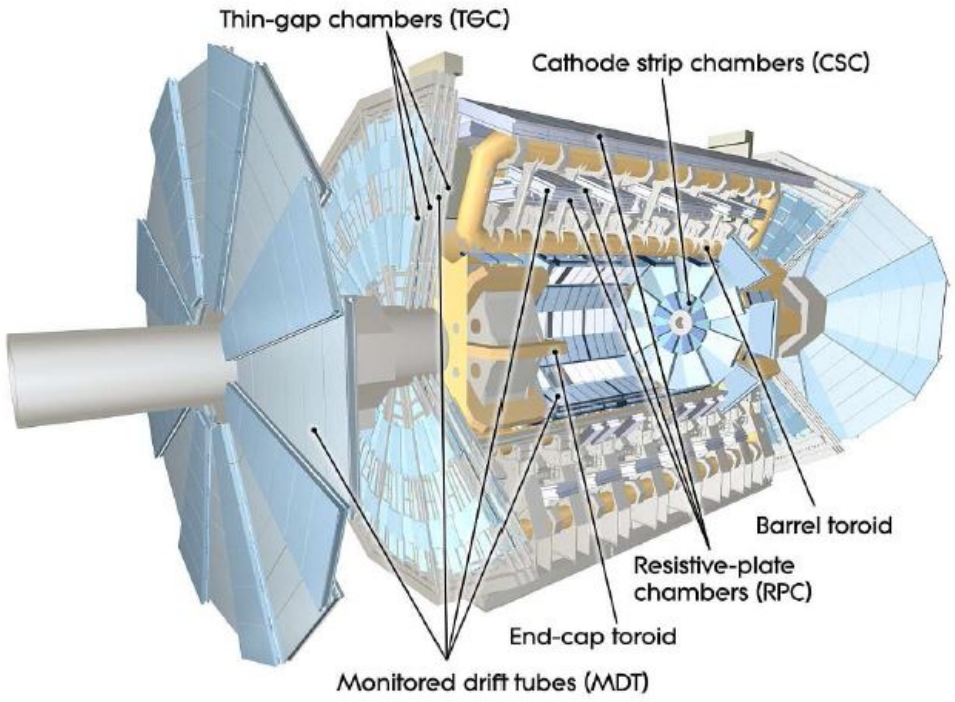


Figure 3.14: Cut-away depicting the ATLAS Muon System [64].

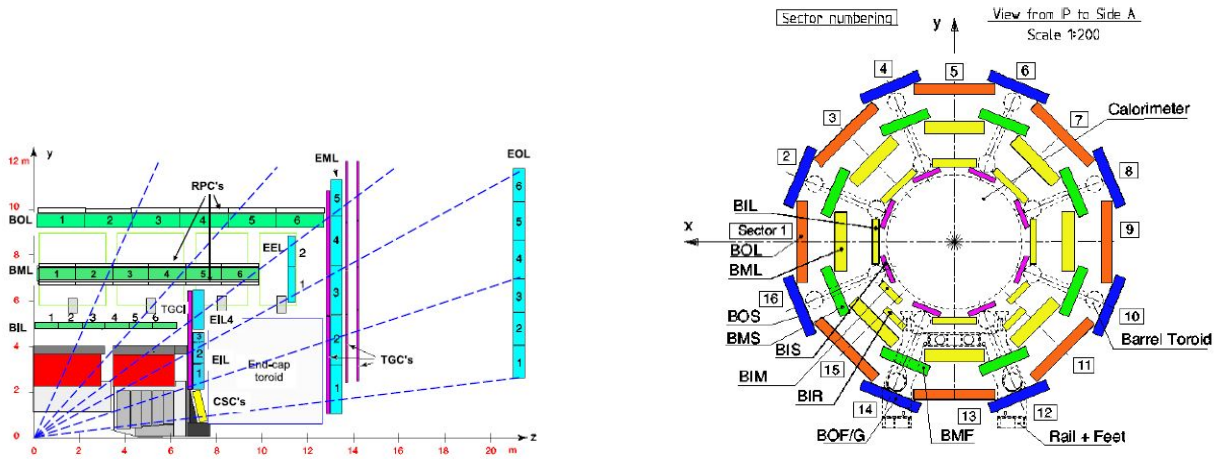


Figure 3.15: Schematic of the ATLAS Muon System, barrel (left) and end-cap (right) [64].

3.2.5.1 Monitored Drift Tubes

Monitored Drift Tubes (MDTs) consist of pressurized drift tubes and are designed to provide high-precision measurements of muon momentum. Drift tubes are chosen for their measurement accuracy and design simplicity. The MDTs provide coverage for $|\eta| < 2.7$, except in the first end-cap layer where coverage is limited to $|\eta| < 2.0$ due to an inability to accommodate the higher particle incidence rate that exceeds 150 Hz/cm^2. Each tube operates at a pressure of 3 bar and provides an average resolution of 80 μm, resulting in a total resolution of 35 μm per chamber. The tubes in both the barrel and end-caps are oriented along the ϕ direction with each tube center at normal incidence to the beam pipe.

Each drift tube is 29.970 mm in diameter and filled with a 93% argon, 7% carbon dioxide gas mixture to a pressure of 3 bar. The tube cathode consists of a tungsten–rhenium wire 50 μm in diameter and kept at 3080 V. A cylindrical plug keeps the wire in position and enables gas transfer into and out of the tube. The length of each tube varies from 1 m to 6 m, depending on placement within the barrel or end-caps. In this configuration, the maximum electron drift time is about 700 ns.

The tubes are arranged in 3–8 layers within each MDT chamber. Each barrel chamber has a rectangular shape, while those in the end-caps are trapezoidal. The shapes are chosen to ensure adequate solid angle coverage, and the end-cap chambers are further divided into large and small categories. The tube layers are joined to form multi-layers that are four layers deep in the innermost MS regions and three layers elsewhere. The four-layer choice in the innermost layer enhances the performance of tracking pattern recognition software. The multi-layers are held apart with mechanical spacers, and each chamber consists of two multi-layers. Over time, slight deformations of the tubes occur due to bending under their weight. Since the measurement precision depends on accurate knowledge of the tube wire position, the wire sag is corrected by adjusting its tension to bring the position uncertainty

within 10 μm.

3.2.5.2 Cathode Strip Chambers

Cathode Strip Chambers (CSCs) consist of *multi-wire proportional chambers* (MWPCs)[36] with the cathodes and anodes arranged in a grid pattern, allowing for measurements of the particle position based on the induced charge in the wires. Along with the MDTs, they are used for precision tracking and are located in the innermost end-cap layer, providing coverage of $2 < |\eta| < 2.7$. They can accommodate particle rates up to 1000 Hz/cm^2 in the forward direction and have superior time resolution, compared to the drift tubes. The CSCs are oriented along ϕ such that the cathodes point in the radial direction, with the center of each anode at normal incidence to the beam pipe. The CSC system is comprised of two disks, each with eight alternating pairs of large and small chambers, oriented along the ϕ direction.

The CSC MWPCs are oriented such that they provide measurements of both η and ϕ coordinates. Each chamber contains four CSC planes, making four η–ϕ measurements by looking at the induced charges on the cathode and anode. The measurements have resolutions of 60 μm in the η-plane and 5 mm in the ϕ-plane The η-resolution is determined from the read-out pitches of the anodes in the large and small chambers, respectively 5.31 mm and 5.56 mm. The cathode strips are 17 μm thick copper having widths of 12.52 mm and 21.00 mm in the small and large chambers and kept at 1900 V. The anodes are formed of gold-plated tungsten wires containing 3% rhenium and 30 μm in diameter. Each small and large chamber respectively has 250 and 402 anode wires. An 80% argon, 20% carbon dioxide gas mixture is used inside the CSCs.

3.2.5.3 Resistive Plate Chambers

The Resistive Plate Chambers (RPCs) offer fast triggering on muons, essentially providing track information between 15 ns and 25 ns. They are utilized in the barrel region, covering $|\eta| < 1.05$, and are comprised of parallel plates separated by 2 mm, made of electrically resistive phenolic–melaminic plastic. A gas mixture 94.7% $C_2H_2F_4$, 5% Iso-C_4H_{10}, and 0.3% SF_6 mixture fills the space between the plates. The plates are kept at a 9800 V potential difference, resulting in a 4.9 kV/mm electric field. Charged particles passing through the plates trigger an avalanche of charge in the gas medium. This avalanche charge is read out by metallic strips glued to the surfaces of each plate. The RPCs also provide measurements of η and ϕ and have a spatial resolution of 10 mm in both the η- and ϕ-planes.

Each RPC is composed of two contiguous rectangular detectors called *units*, each having a *gas volume*, which is composed of two layers of resistive plates and gas, and four read-out strip panels. The gas volumes are interleaved with a light paper honeycomb support

[36]Invented by Georges Charpak in 1968 at CERN, winning him a Nobel Prize in 1992, the multi-wire proportional chamber operates similarly to the drift tube, with large numbers of cathodes bathed in gas. MWPCs can handle higher rates of particle incidence than drift tubes, offering a significant advantage.

structure. External support panels are attached to furnish further mechanical support. The total thickness varies from 112 mm to 122 mm. Units are typically placed with the MDTs in their support framework, but some are used alone where there is no space for MDT units, typically around the ribs and feet of the magnet system. These special units maintain high trigger coverage.

3.2.5.4 Thin Gap Chambers

Thin Gap Chambers (TGCs) are fast triggering systems utilized in the end-cap, covering $1.05 < |\eta| < 2.4$. They are comprised of MWPCs and provide an additional measurement of ϕ, complementing the η-measurement performed by the MDTs. They can handle the high event rate in the forward region and have good time resolution. The TGCs are placed in seven layers along the middle MDT layer in the end-cap and two layers along the inner MDT layer. The layers are interlinked, with one triplet and two doublet groupings, to reject background events. Each layer consists of two concentric rings. The outer end-cap ring covers $1.05 < |\eta| < 1.92$, while the inner forward ring covers $1.92 < |\eta| < 2.4$.

The TGCs are MWPCs constructed with the wire–cathode distance smaller than the distance between individual wires, respectively 1.4 mm and 1.8 mm. The MWPC gas mixture is 55% carbon dioxide and 45% n-pentane (n-C_5H_{12}), chosen to be highly *quenching*, or resistant to ionization avalanches. The anode wires are 50 μm in diameter and are kept at 2900 V. Signals are capable of being read out within 25 ns 99% of the time in this configuration. The remaining 1% of signals pass at normal incidence between two wires, where the drift electric field vanishes. Copper strips are clad onto flame resistant material coated in graphite and placed such that the strips face *away* from the wires. Two of the copper strips are divided into read-out strips to provide ϕ information. A TGC chamber consists of a gas volume with two cathodes and two wire planes. A honeycomb paper stiffener 20 mm in thickness is placed between adjacent chambers in each module to provide mechanical support. The number of wires in each TGC varies from 6 to 31, depending on η, providing high granularity for the required momentum resolution. The wire and cathode strip alignment between successive layers is staggered to provide optimal position resolution and a ϕ-granularity of 2–3 mrad.

3.2.6 Triggers and Data Acquisition

The LHC produced 20 million collisions every second during 2012 operations. ATLAS was able to record only 1000 of these 20 million collisions, using a *trigger system* to decide which events to keep. This system consists of three levels, shown in a flowchart in Figure 3.16. The first level is the *L1 trigger*, implemented in hardware to make fast decisions and reducing the 20 million collisions down to 70000. It selects events with high-p_{T} muons, electrons, photons, jets, and hadronically decaying tau leptons. The second level is the *L2 trigger*, which reduces the 70000 events to 6500. It looks at Regions-of-Interest (RoIs) in the detector, identified by the L1 trigger as having interesting objects. The L2 trigger uses the RoI coordinates,

measured energy, and signature type to determine whether or not to keep the object. The third level is called the *Event Filter* (EF), and it reduces the 6500 events passed by the L2 trigger to 1000 for recording. The EF is implemented entirely in software, and makes event acceptance decisions using programmed algorithms. The L2 trigger and Event Filter are together called the *High Level Trigger* (HLT).

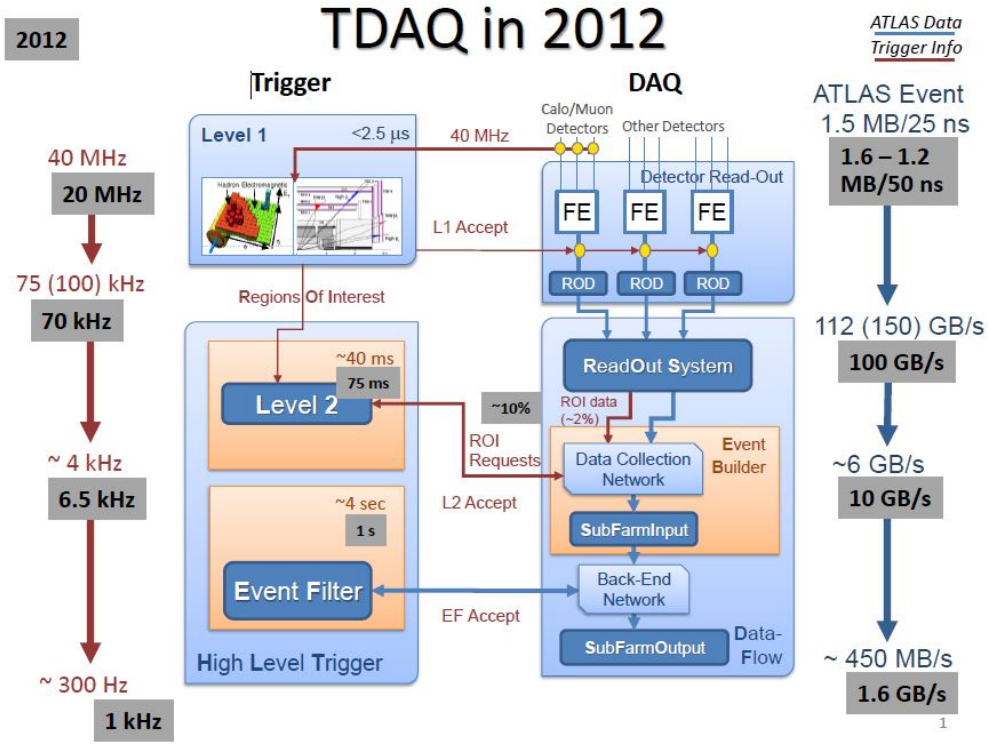


Figure 3.16: Flowchart illustrating the trigger and data acquisition systems [67]. The 2012 rates are given in the grey boxes with the design values indicated above.

The L1 trigger makes decisions based on information provided from the calorimeters and MS. The *L1 Calorimeter Trigger* (L1Calo) identifies objects and events with large transverse energy or missing transverse energy. It triggers on electrons, photons, jets, and hadronic taus. It can be configured to trigger on a large sum of jet transverse energy. The electron, photon, and tau triggers may also be configured to trigger on *isolated* objects of interest in which the amount of energy in a cone surrounding the object is less than a threshold amount. It is also possible to trigger on multiplicities of objects. The L1Calo is interfaced with calorimeter units called *trigger towers* consisting of groupings of calorimeter cells with granularity $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. Figures 3.12 and 3.17 illustrate these towers. Electron, photon, and tau triggers use 2×2 clusters within the trigger tower and have an isolation requirement that no objects be present in the ring of cells outside the 2×2 cluster. Jet triggers also use this 2×2 trigger tower arrangement. The L1 muon trigger uses information provided by the RPC and TGC to accept or reject muon candidates. Conceptually, a decision

is based on a coincidence of hits in three trigger stations, each located in the barrel and end-caps. The hits have to be located within a p_{T}-dependent width of the muon trajectory, called a *road*. All L1 trigger information is passed to the *Central Trigger Processor* (CTP) that ultimately decides on which events to trigger. The CTP has 256 trigger object definitions that can be set based on information from any of the detector systems. For example, a trigger could be defined for events with one high-p_{T} muon, one high-p_{T} hadronic tau, and large missing transverse energy. The L1 trigger decision is quickly made in 2.5 μs.

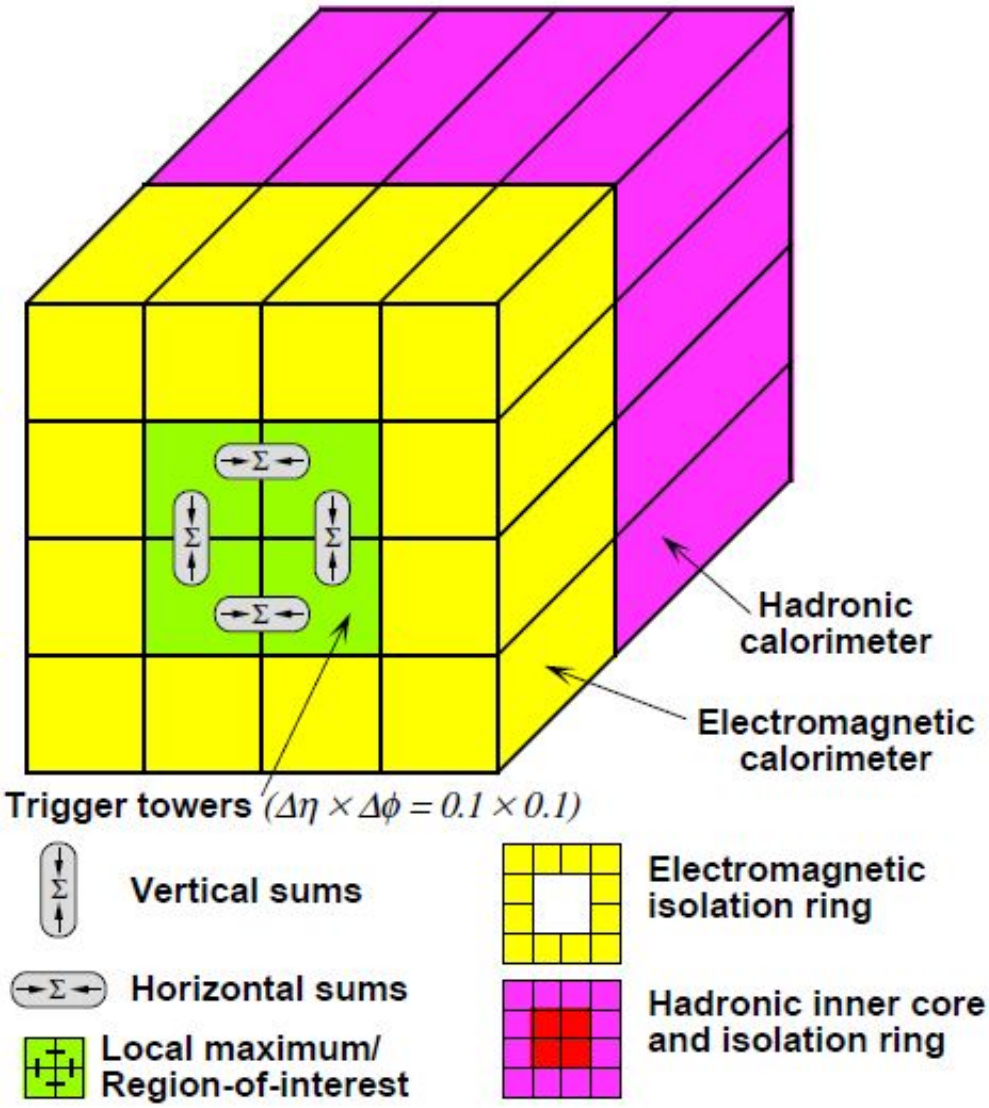


Figure 3.17: Diagram illustrating trigger algorithms for electrons, photons, and hadronically decaying tau leptons [64].

The HLT is comprised of the L2 trigger and Event Filter. It works by examining infor-

mation in RoIs provided by the L1 trigger system. All information in a given RoI and event data are combined into a single data structure. After analysis at a processing farm, the L2 trigger decides whether or not to accept the event. If accepted, the raw data from the event are passed to the Event Filter, which fully reconstructs the event using standard ATLAS reconstruction software. Once the event is reconstructed, triggering algorithms determine whether or not to keep the event. The nominal processing times for the L2 trigger and EF are 40 ms and ∼4 s, but were respectively 75 ms and 1 s during 2012 operations. Because of the difference in processing times between trigger levels, events awaiting higher level decisions are kept in a pipeline. Data from accepted events are written into one or more datastreams dedicated to various objects (e.g. muons, electrons and photons, jets and missing transverse energy) for analysis.

3.2.7 Luminosity Measurement

Luminosity refers to the rate of proton–proton collisions, as discussed in Section 3.1.4. The sensitivity of analyses depends on the accuracy of the luminosity measurement. Many sub-detectors can measure the luminosity, but ATLAS uses two *forward detectors* explicitly designed for making such measurements.

The first such detector is the *LUminosity measurement using Cerenkov Integrating Detector* (LUCID).[37], located 17 m from the interaction point on both sides and designed to detect inelastic proton–proton collisions in the forward region. It can make measurements of the integrated luminosity and monitor beam conditions and instantaneous luminosity in real-time. The detector is formed from twenty aluminum tubes 1.5 m long and 15 mm in diameter, arranged around the beam pipe roughly 10 cm away and oriented pointing toward the interaction point. The tubes are set in an aluminum vessel containing C_4F_{10} gas that is 1.2–1.4 bar in pressure. PMTs are affixed to the end of each tube. Relativistic particles, exceeding the speed of light in the gas, entering the vessel emit Cerenkov photons that are measured by the PMTs, counting the number of particles in the tube. Unfortunately, the LUCID performance is expected to degrade considerably over time due to the high-radiation environment and increased LHC instantaneous luminosity.

While LUCID is able to make luminosity measurements, it can only measure the luminosity corresponding to inelastic events (See Equation 3.1). The number of events where the protons collide without fragmenting (*elastic* events) is needed to determine the total luminosity. If these events are measured, a theorem[38] relating elastic scattering to the total cross section can be used to determine the total luminosity.

The *Absolute Luminosity For ATLAS* (ALFA)[39] detector measures the scattering angles of elastic events to determine the total luminosity. The tiny elastic scattering angles (3 μrad) require special beam configurations since the beam spread regularly far exceeds this angle. β^* is increased and the emittance is decreased to set these special conditions. The ALFA

[37]Physicists are experts at taking ordinary words and making contorted acronyms out of them.
[38]The *Optical Theorem*.
[39]See the comment about the LUCID name.

consists of *Roman pot* detectors [69] located 240 m from the interaction point on both sides. A Roman pot detector consists of a detector volume called a *pot* that is attached to the beam pipe, sharing the vacuum but kept hermetic by means of a window. During stable beam operation, the pot can be moved to a 1 mm distance from the beam. Because of the beam pipe configuration, Roman pots are mounted above and below each beam pipe (such that there are four per side). The detector components consist of ten back-to-back double-sided modules containing 64 scintillating fibers arranged in a grid on each side. The thickness of each grid is 0.5 mm, and the fibers are 32 mm in length, but due to a slight trimming of the corners of each square, the effective detector area is smaller than 32×32 mm^2. The modules are connected to PMTs attached to read-out instrumentation. The detector provides a spatial resolution of 30 μm.

Figure 3.18 shows the integrated luminosity delivered by the LHC and measured by ATLAS throughout 2012.[40] The LHC delivered 22.8 fb^{-1} of proton–proton collisions, of which 21.3 fb^{-1} were recorded by ATLAS, resulting in a data-taking efficiency of ∼93%. Of this recorded data, 20.3 fb^{-1} satisfied data quality requirements and were certified for analysis. This 20.3 fb^{-1} dataset is analyzed in this dissertation.

[40] American readers should note that the dates in the horizontal axis are day/month.

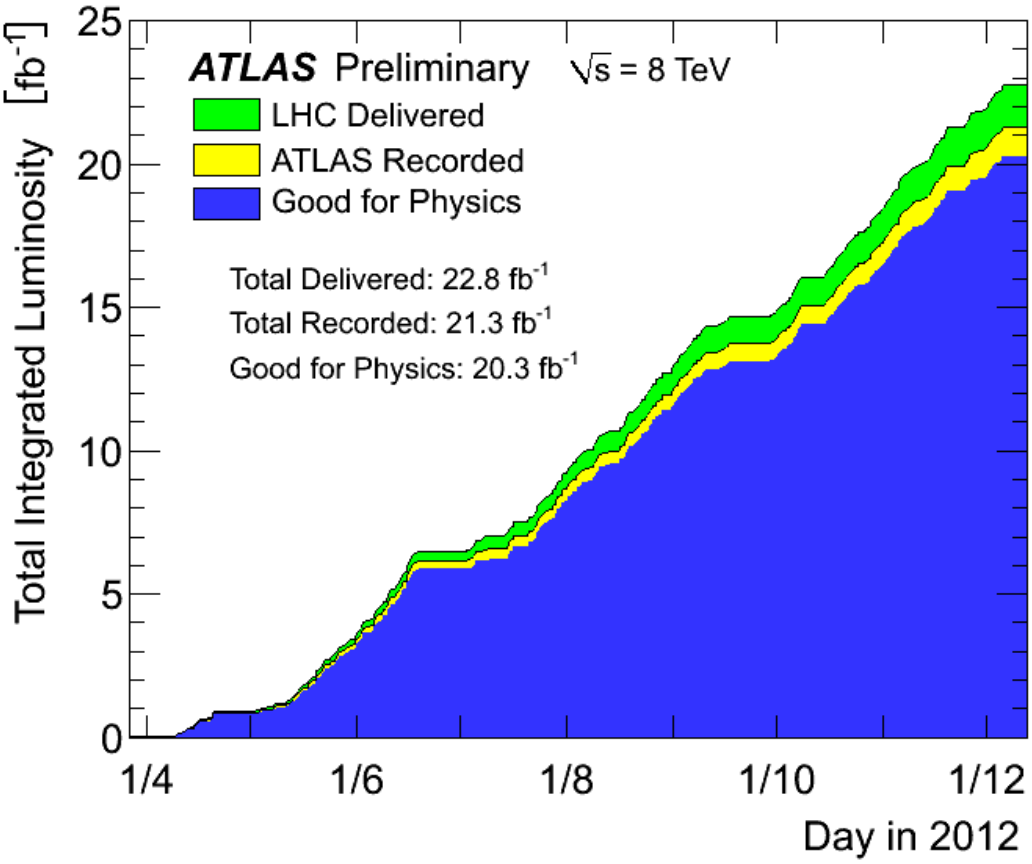


Figure 3.18: Total Integrated Luminosity in 2012 delivered by the LHC, recorded by ATLAS, and good for physics [67].

Chapter 4

Object Reconstruction

The previous chapter described how proton–proton collisions result in the millions of electronic signals measured by the detector. These signals originate from the particles described in Chapter 2 and have specific kinematic properties. The method of identifying particles and their kinematics based on detector information is called *reconstruction*. Figure 4.1 illustrates the detector signatures of some of the most common reconstructed particles, known as *physics objects* or simply *objects*.

This chapter discusses the reconstruction of those physics objects pertinent to this analysis. Since tracking information is used to seed high-level reconstruction algorithms, tracking and vertex reconstruction are first discussed. LFV Higgs boson decays consist of a single electron or muon and a tau lepton. The taus of interest decay hadronically to a neutrino and one or more hadrons (usually pions). The muon and electron are directly reconstructed from energy deposits and tracks in the detector, while the neutrino's kinematics are inferred from the *missing transverse energy* $E_{\mathrm{T}}^{\mathrm{miss}}$, an energy imbalance in the detector. Hadronic taus are not directly reconstructed but are identified from reconstructed jets. Therefore, both jet reconstruction and tau identification will be discussed.

4.1 Track and Vertex Reconstruction

The first level of ATLAS reconstruction is building *tracks* and *vertices*. Tracks indicate the paths traversed by charged particles moving through the ID and are built from energy deposits in the PIX, SCT, and TRT. The tracks are curved because of the magnetic field present in the ID, though high momentum tracks appear considerably straight. Momentum measurements are made from the magnitude of curvature of each track (a particle's charge is determined from the direction of track curvature). Consequently, tracking performance degrades as momentum increases. Reconstructed tracks can emanate from single points called *vertices*. *Primary vertices* originate from the proton–proton collisions, while secondary vertices occur where a particle decays into two or more charged particles. A full description of tracking and vertex reconstruction can be found in [64], [71], [72].

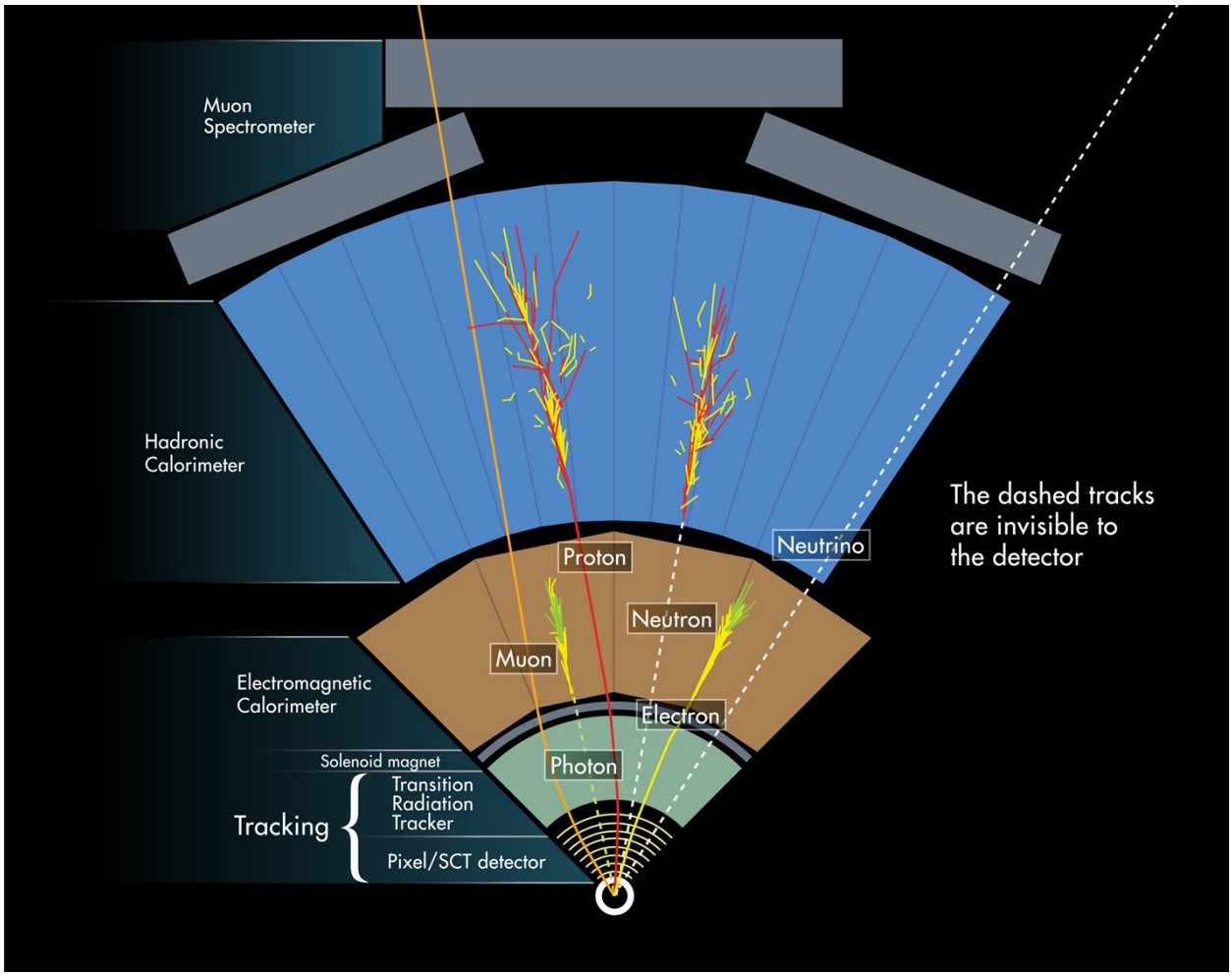


Figure 4.1: Particle detection at ATLAS [70].

4.1.1 Track Reconstruction

Track reconstruction is the most fundamental element of charged particle reconstruction, proceeding in three stages. PIX and SCT hits are first assembled into clusters, and timing information from the TRT is used to construct *drift circles*.[1] PIX and SCT clusters are then constructed into 3-D space points[2] using the detector geometry and the stereo angle difference for SCT double-hits.

Because of their high granularity, track-finding algorithms analyze the three PIX layers and first SCT layer, looking for tracks that originate from the interaction region.[3] Formations

[1] A drift circle is constructed such that its radius corresponds to closest distance a charged particle might have approached a straw. This distance is estimated from the drift time of the free electrons as they reach the cathode.

[2] 3-D space points consist of three numbers specifying the spatial location of the cluster.

[3] These tracks are called *prompt tracks*. In general, *prompt objects* are those originating from the inter-

of four space points (three from PIX, one from SCT), called *track seeds*, are produced by reconstruction algorithms. Each track seed is extended into the remaining SCT layers, forming a *track candidate*, which is then fitted using a simplified *Kalman filter*[4] [74]. Outlier clusters and fake tracks are then rejected, and ambiguous clusters[5] are resolved using a quality criteria scoring system. The tracks are then extended into the TRT so that the drift circles within a given track width[6] may be associated and ambiguities concerning whether the particle traversed the circle's left or right side resolved. All space points and drift circles are then refit and compared to tracks fit using only silicon sensors. TRT hits spoiling the fit are classified as *outliers* and removed from the fit (they are still considered as being part of the track).

In the technique of *back-tracking*, unused TRT track segments are extrapolated back into the SCT and PIX to reconstruct secondary tracks from non-prompt processes such as photon conversions into electron–positron pairs and particles that decay well inside the ID. This approach is also called *outside-in* reconstruction (in contrast to the earlier *inside-out* approach). Reconstruction of these secondary tracks improves the overall tracking efficiency.[7]

Because tracks are so fundamental to higher reconstruction, very high tracking resolution is required, specifically $\sigma_{p_\mathrm{T}}/p_\mathrm{T} = 0.05\% \cdot p_\mathrm{T} \oplus 1\%$. Another measure of tracking performance is the tracking efficiency, shown in Figure 4.2 for muons, pion, and electrons with $p_\mathrm{T} = 5$ GeV. The drop in efficiency at higher $|\eta|$ for pions and electrons is because of the increased amount of detector material traversed. Pions interact hadronically with detector nuclei, while electrons are subject to bremsstrahlung.

Information about the kinematics of the particle is obtained from the track fit. In addition to charge, momentum, η, and ϕ, the *transverse* (d_0) and *longitudinal* ($z_0 \sin\theta$) *impact parameters* specifying the distance of closest approach to the interaction point in the tranverse plane and η-planes are obtained.

4.1.2 Vertex Reconstruction

Vertices are points where two or more tracks converge. The *primary vertex* is associated with tracks emerging from the interaction point, and it distinguishes tracks of interest from those originating from softer background interactions. A full description of primary vertex reconstruction is provided in [75], but it proceeds in two stages. An algorithm for finding primary vertices first matches reconstructed tracks with potential vertices. Then, an algorithm for fitting vertices determines the vertex position and associated uncertainty. All tracks

action point.

[4]Kalman filters have wide application in science and engineering. While a description of a Kalman filter is beyond the scope of this dissertation, a delightful conceptual explanation may be found in [73].

[5]Ambiguous clusters are those that could be assigned to more than one track.

[6]These wide tracks are called *roads*, conceptually identical to those in muon triggering. See Section 3.2.6.

[7]The tracking efficiency is the number of reconstructed tracks per number of charged particles in the tracking volume.

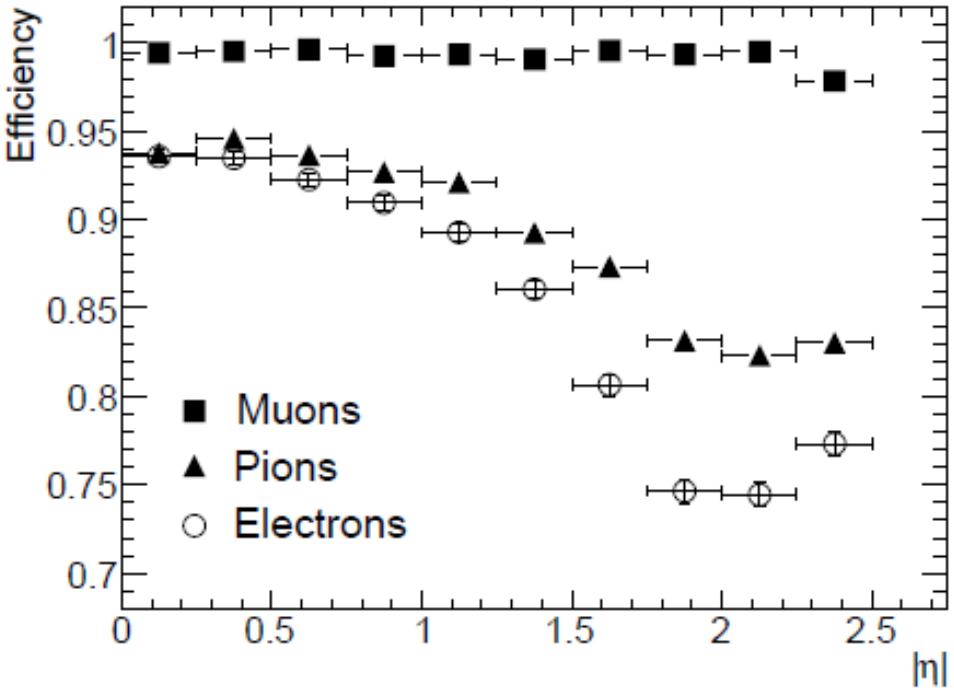


Figure 4.2: Tracking Reconstruction efficiency as a function of η for muons, pions, and electrons with $p_{\mathrm{T}} = 5$ GeV.

considered for vertex reconstruction have $p_{\mathrm{T}} > 150$ MeV and originate from the interaction point, satisfying other requirements described in [75].

The two algorithms are incorporated into an approach called *Iterative Vertex Finding*, in which a *seed vertex* is identified by making a distribution of the track z-coordinates at closest approach to the interaction point, then identifying the global maximum. An *adaptive vertex fitting algorithm* [76] locates the position of the vertex based on the seed position and surrounding tracks. This algorithm performs a χ^2 fit in which the fitted tracks closer to the seed are given preferential weighting to those further away. The fit proceeds iteratively with the weights reassessed at each pass. Tracks deemed incompatible with the vertex by more than seven standard deviations are used as new seed vertices. This compatibility is assessed using a χ^2 fit with two degrees of freedom. The procedure continues until all tracks are associated to vertices, or no other vertices are able to be located. The vertex with the highest $\sum p_{\mathrm{T}}^2$ is identified as the *primary vertex*. A vertex and its secondary vertices may be imagined as forming a "tree" of tracks. Events can have more than one such tree as a result of *pile-up*, or multiple proton–proton collisions within the same bunch crossing. The largest $\sum p_{\mathrm{T}}^2$ vertices in trees outside of the one with the primary vertex are called *pile-up vertices*.

4.2 Muon Reconstruction and Identification

Muon reconstruction is based on hits recorded in the different systems of the MS, described in [71]. Reconstructed muons have momenta from ∼3 GeV to ∼3 TeV and are based on reconstructed MS tracks. The formation of drift circles in the MDTs and clusters in the CSCs, RPCs, and TGCs through processing of the raw data is the first step in MS track reconstruction. The drift circles and clusters in the MDTs and CSCs are grouped into *track segments* that are straight lines located entirely within a single station. Seeds for this method are a pattern of drift circles or clusters, or the presence of drift circles or clusters in a RPC/TGC-defined *region of activity* having a size of $\eta \times \phi = 0.4 \times 0.4$. Track candidates are constructed from track segments using an outside-in approach, beginning with those segments in the outer and middle MS chambers, then extrapolating to other segments in the inner chambers. All track segments reasonably matching the track candidate are fitted to determine the track shape. The inner segments determine the initial track parameters. The fit also considers the detector material geometry and magnetic field structure along the track. Track parameters are corrected for energy loss in the upstream detector systems, specifically the calorimeters and ID. This method of track reconstruction is implemented by the MUONBOY algorithm [77].

Three types of muons can be reconstructed using the MS tracks. *Stand-alone muons* are reconstructed purely on the basis of MS tracks and can be reconstructed for $|\eta| < 2.7$. *Combined muons* are reconstructed by associating an ID track and a MS track and may be reconstructed for $|\eta| < 2.5$. The difference in coverage between stand-alone and combined muons is due to the ID tracking volume. Finally, *segment tag muons* are reconstructed by associating a track segment in the inner MS chambers with a track in the ID. These muons are reconstructed for $|\eta| < 2.5$.

Combined muons are constructed and identified using the STACO algorithm [77], which considers track parameters from each reconstructed track in the ID and MS as well as their covariance matrices. Such muons have improved momentum resolution in the range of 6 GeV $< p_{\mathrm{T}} <$ 100 GeV relative to other reconstructed muons. Furthermore, combined muons offer rejection of secondary muons and those from pion and kaon decays. ID tracks considered for combined muons are required to have at least one PIX hit, at least five SCT hits, less than three PIX or SCT *expected hits*,[8] and at least nine TRT hits in the region of TRT acceptance.

Segment tag muons are formed by extrapolating ID tracks to the inner MS chambers and associating them either with existing track segments or to drift circles and clusters in a cone of 100 mrad that are then used to reconstruct track segments. These muons offer special advantages over stand-alone muons. Muons with momentum below 6 GeV sometimes do not leave tracks in the middle and outer MS chambers but may be recovered using this method. There are no middle MS chambers in the barrel–end-cap transition region ($1.1 < |\eta| < 1.7$), reducing the efficiency of stand-alone reconstruction. There is

[8] An expected hit occurs when a reconstructed track traverses a sensor, but no hit is registered in that sensor.

also a reduction in MS acceptance at $\eta = 0$ due to infrastructure services needed for other detector components and at the toroid magnet feet. Segment tag muons are useful in these regions. ID tracks used for segment tag muons satisfy the same quality requirements as those used for combined muons.

The efficiency of muon reconstruction is excellent, well over 95% for most of the acceptance, as shown in Figure 4.3. The muon momentum resolution varies by $|\eta|$ from 1.7% in the central part of the detector for muons with $p_{\mathrm{T}} = \sim$10 GeV to 4% in the forward region for muons with $p_{\mathrm{T}} = \sim$100 GeV.

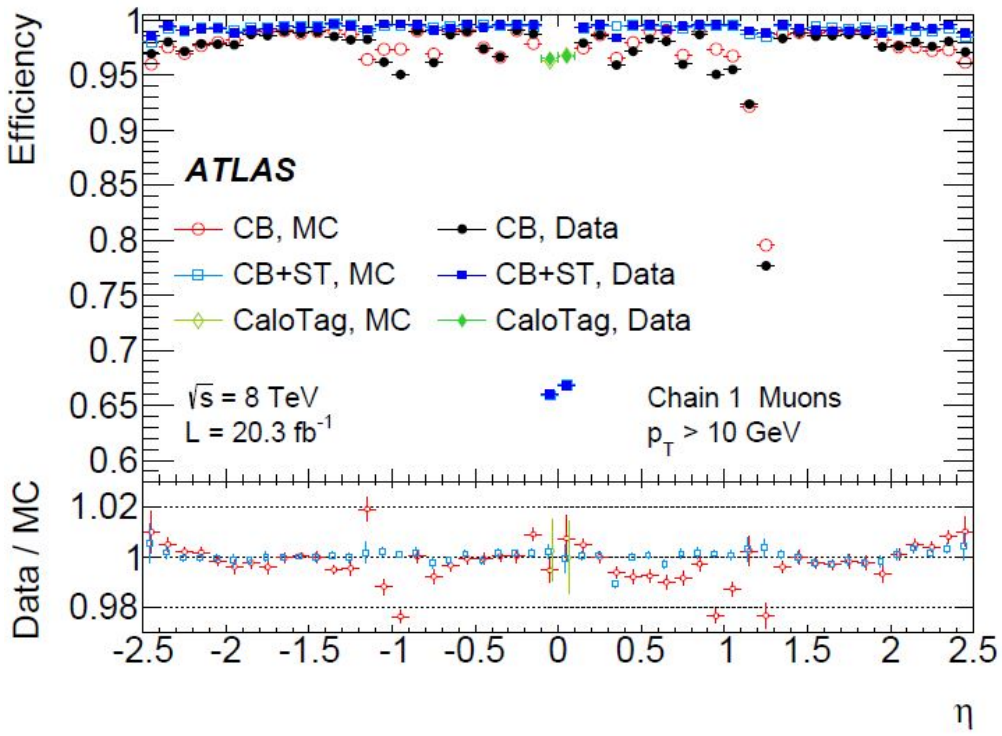


Figure 4.3: Muon reconstruction efficiency as a function of η for different reconstruction techniques [78]. The drop in acceptance for $|\eta| < 0.1$ is from the placement of detector services, resulting in fewer MS stations.

4.3 Electron Reconstruction and Identification

Electrons are reconstructed from ID tracks matched to energy deposits in the EM calorimeter. A full description of the reconstruction and identification procedures may be found in [71], [79]. Reconstructed electrons have $|\eta| < 2.47$ due to the extent of the ID tracking volume. The process starts with the reconstruction of *seed clusters*, groupings of contiguous energy deposits in the calorimeters, by partitioning the η–ϕ space of the EM calorimeter into rectangular units called *towers*, with 200 towers in η and 256 in ϕ. Each tower has a

corresponding dimension of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$, nearly identical to the granularity of the middle layer of the EM calorimeter. The energy of each tower is obtained by summing the energies of all subtended layers (including the presampler for $|\eta| < 1.8$) in the tower volume.[9] A *sliding-window algorithm* [80] identifies groups of 3×5 towers in η–ϕ with energy greater than 2.5 GeV and occurring at a local energy maximum. These groups are called *seed clusters*. A secondary sliding window algorithm then removes overlapping, duplicate seed clusters.

The next stage of electron reconstruction looks for tracks aligning with seed clusters. Tracks are reconstructed in the manner described in Section 4.1.1, but a track reconstruction using an electron-oriented pattern recognition accounting for energy lost in the detector material is also applied. In this scheme, tracks with $p_{\mathrm{T}} > 1$ GeV and three silicon hits[10] in different layers, but failing to be reconstructed as a full track, are considered. This improves electron reconstruction performance and has little overlap with the standard track reconstruction.

Tracks with four or more silicon hits are identified as being *loosely matching* if they may be extrapolated to the middle layer of the EM calorimeter from their closest approach to the primary vertex. Such tracks have to be within 0.05 in η of the cluster and 0.2 in ϕ if the track is bending toward the cluster, or within 0.05 if the track is bending away. Tracks with fewer than four silicon hits (*TRT-only tracks*) are extrapolated based on their outermost measurement point and are not subject to the η requirement. Alternatively, any track that extrapolates to the middle layer of the EM calorimeter after having its momentum rescaled to the cluster energy is considered loosely matching if it lies within 0.1 or 0.05 in ϕ, bending toward or away from the cluster, respectively. The η requirement applies to these tracks with four or more silicon hits. Loosely matching tracks are refit using a *Gaussian Sum Filter* (GSF) [81], a more generalized Kalman filter, that considers bremsstrahlung effects. Tracks that fail the GSF and TRT-only tracks are refit using a χ^2 fit.

Refitted tracks are matched to clusters using tighter requirements form *electron candidates*, described in detail in [79]. Once a matching track is found, the clusters are rebuilt using 3×7 (5×5) η–ϕ EM calorimeter cells in the barrel (end-caps), starting in the middle layer and proceeding to the remaining layers. The cluster position is recalculated in each layer considering deposited energy and the overall energy distribution in the barrel and end-caps. The cluster energy is the sum of the energies in each layer weighted by a correction factor obtained using a multi-variate algorithm incorporating different variables pertaining to the cluster position, energy distribution, and shower depth. Reconstructed electron kinematics consist of the measured energy in the calorimeters and (η, ϕ) coordinates determined by the fitted track.[11]

Reconstructed electrons may not originate from prompt electrons, but other objects such as jets or from electrons produced in photon conversions. An algorithm identifies prompt

[9] The energy of layers whose dimensions exceed the tower size and are shared by multiple towers are divided equally among all respective towers.

[10] A silicon hit is a hit in either PIX or SCT.

[11] If the track is a TRT-track, then the (η, ϕ) coordinates are taken from the cluster.

electrons and rejects those from background processes. Three classification levels, *loose*, *medium*, and *tight*, indicate progressively increasing quality criteria, or *cuts*, satisfied by the reconstructed electron. A fourth category, *multilepton*, was developed to identify low energy electrons in the $H \to ZZ^* \to 4\ell$ analysis. Each category is optimized for different values of $|\eta|$ and transverse energy due to the amount of detector material traversed by the particle and the differences in calorimeter shower shapes and tracks with increasing energy. A full description of the quality criteria for each category may be found in [79], but they broadly consist of requirements on the shower width in the strip and middle layers of the EM calorimeter; track quality requirements such as the number of PIX hits, B-layer hits, silicon hits, and impact parameter requirements; requirements on TRT hits and ratio of high threshold hits to total hits; matching between tracks and clusters; matching to photon conversions; and the amount of energy leakage into the hadronic calorimeter.

The efficiencies for the classification levels are calculated using data and simulations of $Z \to ee$ events and are shown in Figure 4.4. The drop in efficiency in the region $1.37 < |\eta| < 1.52$ is due to the increase of detector material in the transition region between the barrel and end-cap EM calorimeter.

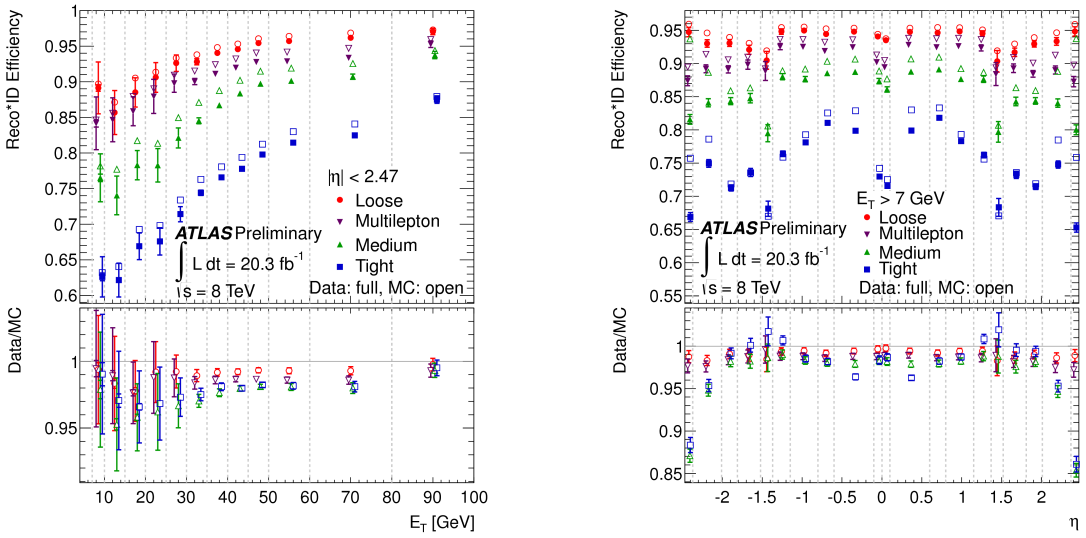


Figure 4.4: Electron Reconstruction efficiency for different electron classifications as a function of E_{T} and η [79].

4.4 Jet Reconstruction

As discussed in Chapter 2, quarks and gluons produced in proton–proton collisions hadronize,[12] forming jets composed of particles such as pions, neutrons, and protons whose signature is

[12]The top quark does not hadronize due to its short lifetime.

a densely packed array of collimated tracks and energy clusters in the calorimeters. Neutral pions decay nearly all of the time to two photons that leave clusters in the EM calorimeter, while other hadrons interact with the nuclei in the calorimeters, forming hadronic showers. Jets are important to physics analyses and play a key role in missing transverse energy reconstruction. A description of jet reconstruction may be found in [71].

4.4.1 Reconstruction

Jet reconstruction occurs in three stages: formation of energy clusters in the calorimeters, grouping of these clusters into jets, then calibration of the jet energy. *Topological clusters* [80], or neighboring groups of calorimeter cells whose energy-to-noise ratio exceeds a certain threshold, are used for jet reconstruction. Their construction involves cluster making and splitting.

The Cluster Maker algorithm first scans calorimeter cells, identifying those with an absolute energy-to-noise ratio exceeding a threshold value, $t_{\text{seed}} = 4$, as being seed cells, or *proto-clusters*, that are then added to a seed list. Noise is considered from two sources: read-out electronics noise, estimated as the rms value for the given gain and status, and pile-up events, whose contribution is estimated. The seed list is ordered according to t_{seed} in descending order. Then, neighboring cells of each proto-cluster that are not seed cells and have an absolute energy-to-noise ratio greater than a threshold $t_{\text{neighbor}} = 2$ are included on a neighbor seed list and associated with the corresponding proto-cluster. In cases where a neighbor seed is shared by multiple proto-clusters, the proto-clusters are grouped together. If a cell has absolute energy-to-noise ratio less than t_{neighbor}, then it is associated to the adjacent proto-cluster with the highest value of t_{seed}. The seed list is discarded once it is scanned. The neighbor list then becomes the seed list, and the scanning procedure is repeated until no seed cells remain. Under this approach, a cluster consists of cells with absolute energy-to-noise ratios exceeding t_{neighbor} having a perimeter of cells with absolute energy-to-noise ratio less than t_{neighbor}.

Ideally, individual clusters match to single particles, but clusters may encompass many particles whose deposited energy overlaps in constituent cells. The Cluster Splitter algorithm resolves clusters to separate particles. The algorithm first looks for local energy maxima by identifying cells with energies larger than 500 MeV that also have greater energy than neighboring cells. Such cells must have at least four neighbors to be considered. The cluster making algorithm is then applied, with the local maxima cells constituting the seed list. However, only the cells in the original clusters are considered, there are no absolute energy-to-noise thresholds, nor does any merging of cells occur. The seed list is reordered at each iteration by energy in descending order with unused neighbor cells associated to the adjacent proto-cluster. Neighbor cells bordering multiple proto-clusters are associated with the two most energetic proto-clusters and added to a shared list. Neighbor cells in the original set that are unassigned to any proto-cluster are then added to the list. Each cell is added to both proto-clusters, with each cluster receiving a weighted contribution of the cell energy. These weights are computed based on the proto-cluster energy and distance between the cell

and proto-cluster center-of-energy and are usually close to zero or one. After this step, the topological clusters are fully formed and can be interpreted as 3-D blobs of energy in the calorimeter.

The topological clusters are then grouped into jets using the *anti-k_t algorithm* [82]. For each cluster, the distance d_{iB} is computed as shown in Equation 4.1:

$$d_{iB} = p_{\mathrm{T},i}^{-2} \tag{4.1}$$

where $p_{\mathrm{T},i}$ is the p_{T} of the i-th cluster. The distance d_{ij} is then computed for every possible pair of clusters as shown in Equation 4.2:

$$d_{ij} = \min(p_{\mathrm{T},i}^{-2}, p_{\mathrm{T},j}^{-2}) \frac{\Delta R_{ij}^2}{R^2} \tag{4.2}$$

where $p_{\mathrm{T},i}$ and $p_{\mathrm{T},j}$ are the p_{T} of the i-th and j-th clusters, and ΔR_{ij} is the ΔR between the i-th and j-th clusters defined in Section 3.2.1. The parameter R is a radius parameter restricting the size of the jet. A default value of $R = 0.4$ is used at ATLAS.[13] If there is a d_{ij} that is less than d_{iB} for each cluster, then clusters i and j are merged together. Otherwise, the cluster is declared a jet and removed from further consideration.

Under this scheme, softer particles[14] are often merged with harder ones, rather than other soft particles. Isolated hard clusters form perfect cones, while hard clusters that are close to others result in a jet that is a union of cones. Thus, the harder particles determine the shape of the jet. In addition to the experimental benefit of having a well-defined jet shape, the anti-k_t algorithm offers theoretical benefits. A complication of QCD calculations is that quarks and gluons radiate arbitrary soft gluons, making calculations difficult due to the non-perturbative nature of strong coupling at low Q^2. Calculations involving gluon emission collinear to the parent quark or gluon are also difficult. The anti-k_t algorithm is resilient against soft and collinear radiation, making theoretical calculations incorporating it tractable.

Having been constructed, the jet energy is calibrated based on the calorimeter response to event conditions as described in [83]. Jets are first calibrated using the *local cluster weighting* (LCW) method in which topological clusters are labeled as hadronic or electromagnetic based on energy, density, and calorimeter location. A weight that accounts for the response of the calorimeter to EM and hadronic objects is assigned to the cluster based on its classification. These weights are obtained from simulated data. The kinematics of each cluster are recomputed with respect to the location of the primary vertex, and a pile-up correction is then applied. The jets are then calibrated to the known true energy of jets in simulation based on reconstructed jet energy. This correction is known as the *jet energy scale* (JES), and this calibration scheme is called LCW+JES. A multi-variate correction called *global sequential calibration* is then applied. Finally, to ensure a uniform calorimeter response, an η-dependent correction factor is applied to the jet energy.

[13] Wide jets with $R = 0.6$ and $R = 1.0$ are also capable of being reconstructed.

[14] Those particles having a smaller value of p_{T}.

While jets can be reconstructed within the calorimeter volume of $|\eta| < 4.9$, those with $|\eta| < 2.5$ are able to have associated tracks in the ID that may be used to identify jets initiated by b-quarks and those originating from pile-up interactions. The *jet vertex fraction* (JVF) [84] is a quantity that determines whether or not a jet originates from pile-up. It is defined as shown in Equation 4.3:

$$\text{JVF} = \frac{\sum p_{\text{T}}^{\text{jet PV tracks}}}{\sum p_{\text{T}}^{\text{all jet tracks}}} \tag{4.3}$$

where PV means primary vertex. Jets with no associated tracks have JVF $= -1$. Those with values closer to 1 are likely to have originated from the primary vertex.

4.4.2 Identification of b-quark Initiated Jets

Jets originating from b-quarks are important to many physics analyses, especially to those involving top quarks. Therefore, identification of these jets is an experimental necessity. Many *b-tagging* algorithms that exploit the physics of B meson decays have been developed to perform this identification. B hadrons tend to have longer lifetimes since the CKM mixing between b-quarks and the lighter generations is small. Thus, they travel further in the detector before decaying, on the order of 450 μm. These decays manifest as secondary vertices in reconstruction. Because b-quarks are much heavier than the first and second generation quarks, the mass of the vertex is also much larger compared to other vertex masses.

Several algorithms are used for b-tagging at ATLAS, described in [85], [86]. Quality cuts are first placed on the track impact parameters and the vertex mass to reject particles with long lifetimes (such as K_S^0 and Λ) as well as secondary interactions. The *IP3D* algorithm identifies b-jets by looking at 2-D distributions of signed transverse impact parameter significance,[15] d_0/σ_{d_0}, versus longitudinal impact parameter significance, z_0/σ_{z_0}, where σ refers to the uncertainty in the respective impact parameter. These distributions are compared with those for b- and light jets taken from simulated data using a *likelihood ratio*[16] to make a decision.

Other algorithms tag b-jets by looking for secondary vertices within the jet. The *SV1* algorithm identifies vertices sufficiently far from the primary vertex and computes the invariant mass of the tracks in the vertex, the ratio of the sum of the vertex track energies to the sum of the energies of all tracks in the jet, and the number of vertices with two tracks. The invariant mass and energy ratio are plotted against each other in a 2-D histogram, while the distribution of the number of vertices with two tracks is used individually. These two

[15] The sign of the transverse impact parameter is determined by the angle between the jet direction and the line between the point of closest approach and the primary vertex. The sign is positive for angles smaller than 90°, negative otherwise.

[16] This technique will be discussed in Chapter 10.

distributions and the ΔR between the jet axis[17] and the line connecting the primary and secondary vertices are combined in a likelihood ratio, then compared with those for b- and light quark jets taken from simulated data to render a decision.

An algorithm designed to exploit differences in topology between b- and c-quark decays called *JetFitter* is also used for b-tagging. It uses Kalman filtering to approximate the flight trajectory of the b-hadron. This trajectory consists of a line joining the primary vertex to the secondary vertices corresponding to the b- and c-hadrons. It uses similar variables as SV1 and the decay length significances between the vertices incorporated in a likelihood ratio to discriminate between b-, c-, and light jets. The capability to identify c-jets gives it broader functionality compared to SV1.

The IP3D, SV1, and JetFitter algorithms are combined in an *artificial neural network*[18] called *MV1* [87], trained using b-jets as the signal input and light jets as the background input. The output is called the *tag weight* and is interpreted as the probability that the input jet is a b-jet. Cuts on the tag weight are imposed to guarantee a minimum efficiency of correctly tagging b-jets with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$. These cuts are called *working points*. The nominal working point corresponds to a 70% tagging efficiency. Figure 4.5 shows the b-tagging efficiency for b-jets along with the mistag rates for light quark jets, less than 1% for most values of p_{T} and $|\eta|$.

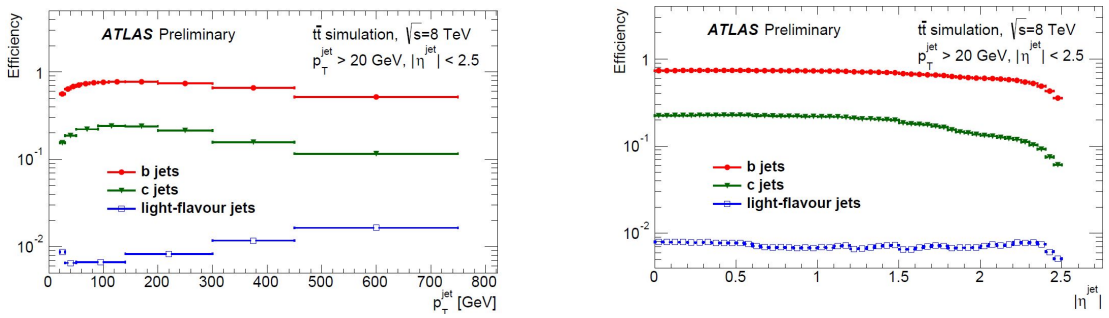


Figure 4.5: b-tagging efficiency as a function of p_{T} and $|\eta|$ [87].

[17] This is a vector pointing from the primary vertex to the weighted center-of-energy, or *barycenter*, of the jet

[18] Artificial neural networks are a type of machine learning algorithm whose operation is similar to that of the human central nervous system. They are generally used to compute outputs (e.g. decisions) based on numerous inputs whose relation to the outputs is usually not well-defined. They consist of a system of *neurons*, each with weighted connections to other neurons, which are eventually connected to the output. The weights of the connections are not known, *a priori*, and are determined through an iterative approach called *training*. Different signal and background samples are fed into the network to determine the best weighting scheme among neurons. As the weights become more settled, the network learns how to distinguish different outputs based on an increasingly complex set of inputs.

4.5 Hadronic Tau Identification

Hadronically decaying tau leptons consist largely of charged and neutral pions. Their detector signature is nearly identical to that of a light quark jet, so hadronic tau candidates are identified from reconstructed jets using various algorithms, all described in [88]. Tau candidates are seeded from reconstructed jets (known as *jet seeds*) with $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.5$ in events with primary vertices having at least three tracks. To improve the reconstruction efficiency and mitigate the impact of pile-up, the production vertex of the candidate tau is identified using the *tau vertex* (TV) *association algorithm*, which first considers tracks from the jet seed with $p_{\mathrm{T}} > 1$ GeV, satisfying quality requirements based on ID hits, and located within $\Delta R < 0.2$ of the jet seed axis. The sum of the p_{T} of tracks associated to the primary vertex and each pile-up vertex is computed, then divided by the total p_{T} of all tracks.[19] Whichever vertex has the highest fraction is chosen as the TV and used as the basis of the tau direction, track association, and coordinate system for the calculation of identification variables. The tau candidate momentum is determined from the η- and ϕ-coordinates of the barycenter of the LCW calibrated topological clusters of the jet seed. The mass of each cluster is taken to be zero. The momentum is then recalculated by only considering clusters within $\Delta R < 0.2$ of the barycenter in the coordinate system defined by the TV, and assuming the tau candidate is massless. This recalculated momentum is called the *tau axis*.

Tracks satisfying $p_{\mathrm{T}} > 1$ GeV, having two PIX hits, seven silicon hits, $|d_0| < 1.0$ mm, $|z_0 \sin\theta| < 1.5$ mm and located in the *core region*, or $\Delta R < 0.2$ of the tau axis, are considered for association to the tau candidate. These tracks are later used to classify the tau candidate as a one- or three-prong hadronic tau. Additionally, tracks located in the *isolation region*, $0.2 < \Delta R < 0.4$ of the tau axis, are associated to the tau candidate if they satisfy the core track requirements. These tracks are used for calculating identification variables.

Neutral pions are then reconstructed using a combination of algorithms that first look for features in the calorimeters, particularly in the strip layer, and input this information along with track momenta into *Boosted Decision Trees* (BDTs)[20] that decide the number of neutral pions in the event: 0, 1, or 2. Tracks and clusters from neutral pion decays are then merged together. Each neutral pion consists of two clusters located within the core region. The properties of each cluster are considered to determine whether or not the cluster originated from a neutral pion, taking into account electronic noise, pile-up, and the *underlying event*.[21] Those clusters deemed as originating from a neutral pion are associated to the tau candidate.

The tau identification process consists of the calculation of eleven variables from the associated tracks and clusters of the tau candidate, all described in [88]. These variables mostly comprise fractions of energy and momentum in given clusters and tracks, the spread of the tracks relative to the tau axis, number of isolation tracks, number of neutral pions, and invariant mass of the tracks. The variables are input into two separate BDT algorithms,

[19] Note that this fraction is analogous to the JVF defined in Section 4.4.1.

[20] A BDT is another machine learning algorithm.

[21] The underlying event refers to all processes in an event, not including pile-up, that are not associated with the process of interest.

each respectively designed to identify taus with one and three tracks. The BDTs are trained using reconstructed hadronic tau candidates in simulated $Z/Z' \to \tau\tau$ and $W \to \tau\nu$ events that are located within $\Delta R < 0.2$ of the decay products of simulated taus having $|\eta| < 2.3$.[22] Three working points are used for the training, each corresponding to a desired signal efficiency, or fraction of true hadronic taus reconstructed with the correct number of tracks and satisfying identification criteria. *Loose*, *medium*, and *tight* tau identification have respective approximate working points of 70% (45%), 55% (40%), and 40% (25%) for reconstructed taus with one (three) track(s). Events in data consisting of one or more jets were used to check the background rejection of the BDT. The background fraction with one (three) track(s) and $p_{\mathrm{T}} > 40$ GeV rejected for each working point is respectively 90% (99.1%), 95% (99.3%), and 98% (99.9%). The relationship between signal efficiency and background rejection is shown in Figure 4.6. The BDT training was performed such that the results are independent of hadronic tau p_{T} and the number of pile-up vertices.

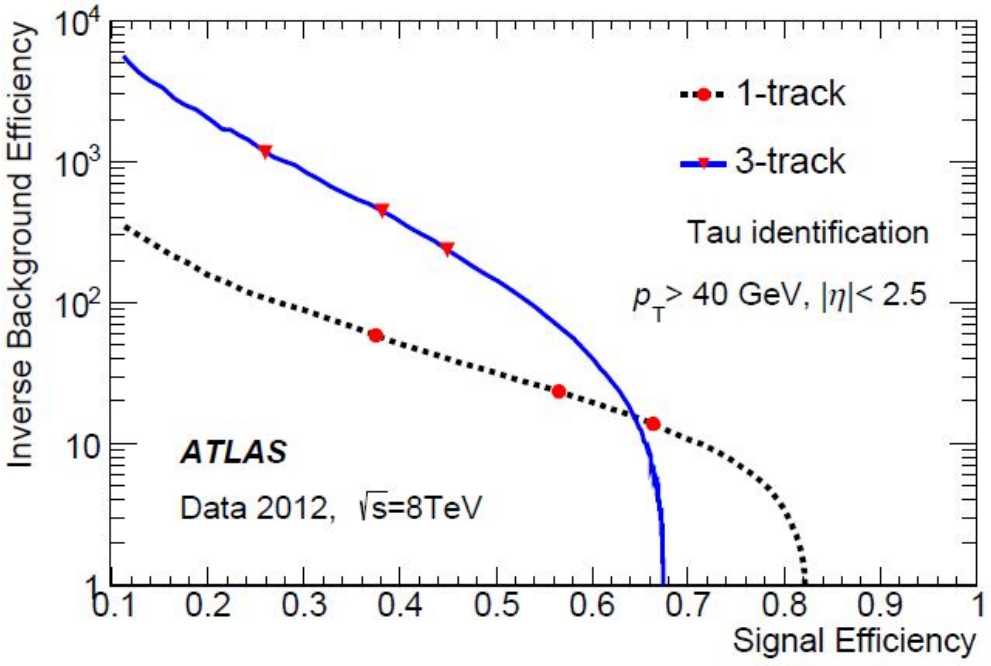


Figure 4.6: Tau identification efficiency and rejection for one and three prong taus [88].

Since jets are used to seed the tau identification process, they form the largest category of objects that may be misidentified as hadronic taus. Jets can be divided into two broad categories: those initiated by quarks and those initiated by gluons. Gluon jets tend to be more spread out and have larger numbers of tracks compared with quark jets. Consequently, the tau identification process offers better discriminating power against this category of background objects.

[22] This process is called *truth-matching*.

Electrons and muons can also be misidentified as hadronic taus with one track. To suppress these background processes, *muon* and *electron vetoes* are applied. Electrons can be distinguished from charged pions on the basis of the presence of transition radiation in the TRT (characterized by the fraction of high-threshold TRT hits), the angle between the track and the tau axis, the ratio of deposited energy in the EM calorimeter to that in all calorimeters (f_{EM}), energy leakage from the EM calorimeter into the hadronic calorimeter, and the ratio of energy in the core region having $\Delta R > 0.1$ to all energy in the core region. The electron veto algorithm is a BDT that uses these variables as input to decide whether or not to reject the tau candidate. The set of variables slightly varies over η since the TRT is only able to provide coverage for $|\eta| < 2.0$. The BDT is trained using truth-matched hadronic taus in $Z \to \tau\tau$ events as signal and hadronic taus truth-matched to electrons in $Z \to ee$ events as background. Working points corresponding to *loose*, *medium*, and *tight* selection are used, each respectively having signal efficiencies of 95%, 85%, and 75%.

Most muons are rejected as being hadronic taus by the standard muon identification algorithms, described in Section 4.2. However, muons that fail to be reconstructed in the MS can be misidentified as hadronic taus. These muons can lose most of their energy in the calorimeters or be absorbed in the calorimeter near an energetic cluster. Two variables that offer discriminating power against these types of muons are f_{EM} and the ratio of track p_{T} to cluster transverse energy, $p_{\mathrm{T}}/E_{\mathrm{T}}$. Muons depositing most of their energy in the calorimeters tend to have low f_{EM} and large $p_{\mathrm{T}}/E_{\mathrm{T}}$, while those that are absorbed near a cluster have large f_{EM} and low $p_{\mathrm{T}}/E_{\mathrm{T}}$. Thus, cuts on these variables significantly reduce the impact of these muons on tau identification, resulting in an efficiency of 96% for true hadronic taus and a rejection of 40% for misidentified muons.

Hadronic taus satisfying these identification algorithms have their energy corrected to bring them into agreement with the true energy of hadronic taus in simulation. This correction is called the *tau energy scale* (TES). While tau seeds are already calibrated to the JES, recalibration is necessary due to the following reasons: only clusters in the core region determine the hadronic tau energy, hadronic taus usually consist of one or three charged pions and zero to two neutral pions, and no correction is made for pile-up or underlying event. The correction is performed in three stages using simulated data. First, a scale factor is applied correcting the reconstructed tau energy to the true hadronic tau energy. Then, an $|\eta|$-dependent correction is applied to account for under-reconstructed cluster energies due to regions in the detector lacking instrumentation. Finally, pile-up is corrected by subtracting an increasing amount of energy based on the number of pile-up vertices. The calibrated hadronic tau energy has a resolution that is $\sim$20% for $E_{\mathrm{T}} < 50$ GeV and drops to $\sim$5% for $E_{\mathrm{T}} > 100$ GeV. The energy resolution is degraded by an additional $\sim$5% in the calorimeter barrel–end-cap transition region, $1.3 < |\eta| < 1.6$.

4.6 Missing Transverse Energy Reconstruction

Missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) refers to a momentum imbalance in the transverse plane.[23] It is an important experimental quantity, indicating neutrinos or other exotic particles that are undetectable. Accurate $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction relies on the detector being hermetic in the transverse plane.[24] A description of $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction may be found in [89].

$E_{\mathrm{T}}^{\mathrm{miss}}$ is reconstructed from energy clusters in the calorimeters and reconstructed muons[25] and has two components corresponding to each dimension in the transverse plane. Each energy cluster is calibrated according to its associated reconstructed physics object. Clusters that may be associated with multiple physics objects are calibrated according to the following order of association: electrons (e), photons (γ),[26] hadronic taus (τ), jets, and muons (μ). Equation 4.4 describes the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation:

$$E_{x(y)}^{\mathrm{miss}} = E_{x(y)}^{\mathrm{miss},e} + E_{x(y)}^{\mathrm{miss},\gamma} + E_{x(y)}^{\mathrm{miss},\tau} + E_{x(y)}^{\mathrm{miss,jets}} + E_{x(y)}^{\mathrm{miss,Soft\ Term}} + E_{x(y)}^{\mathrm{miss},\mu} \tag{4.4}$$

For each term in Equation 4.4:

$$E_{x(y)}^{\mathrm{miss,object}} = -\sum_{\text{all objects}} p_{x(y)} \tag{4.5}$$

In the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation, all jets considered have $p_{\mathrm{T}} > 20$ GeV. For combined muons, the energy lost in the ID and calorimeters that is corrected in the reconstructed muon momentum is subtracted from the calculated $E_{\mathrm{T}}^{\mathrm{miss}}$ to avoid double-counting. A soft term that accounts for jets with $p_{\mathrm{T}} < 20$ GeV and unassociated topological clusters and tracks is also included in the calculation. The unassociated topological clusters are calibrated using the LCW technique. This soft term is especially susceptible to pile-up, potentially degrading $E_{\mathrm{T}}^{\mathrm{miss}}$ performance. To suppress the effect of pile-up, the soft term is scaled by the *soft term vertex fraction* (STVF), computed as follows:

$$\mathrm{STVF} = \frac{\sum p_{\mathrm{T}}^{\text{Soft Term PV Tracks}}}{\sum p_{\mathrm{T}}^{\text{All Soft Term Tracks}}} \tag{4.6}$$

The $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution is calculated to be ∼5 GeV for events with $E_{\mathrm{T}}^{\mathrm{miss}} = 200$ GeV and ∼25 GeV for those with $E_{\mathrm{T}}^{\mathrm{miss}} = 1$ TeV.

[23] The term is a misnomer, as energy is a scalar quantity and momentum is a vector quantity. $E_{\mathrm{T}}^{\mathrm{miss}}$ is a (2-D) vector quantity.

[24] It is not possible for the detector to be hermetic longitudinally because of created particles escaping down the beam pipe.

[25] Both combined and segment tag muons are used for $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction.

[26] Reconstructed photons consist of clusters in the EM calorimeter, particularly the strip layer. Their reconstruction process shares much in common with that for electrons. See [71].

4.7 Higgs Boson Mass Reconstruction

The lifetime of the Higgs boson is so short that it decays before leaving the beam pipe, making direct detection impossible. However, its presence can be inferred by looking at the invariant mass of its decay products. The Higgs boson mass is unique, distinguishing it from other particles and making it a good discriminating variable between signal and background processes.

LFV events, depicted in Figure 5.1, have hadronic taus, neutrinos, and leptons as the final decay products. Thus

$$m_H^2 = (p_\ell^\mu + p_{\text{vis}}^\mu + p_\nu^\mu)^2 \tag{4.7}$$

where m_H is the Higgs boson mass, and p_ℓ^μ, p_{vis}^μ, and p_ν^μ are the four-vectors for the lepton, hadronic tau, and neutrino. The four-vectors for the lepton and hadronic tau correspond to their respective reconstructed objects, but only the transverse component of the neutrino four-vector may be modeled using $E_{\text{T}}^{\text{miss}}$. Knowledge of the kinematic properties of tau lepton decays can be used to determine the longitudinal component of the neutrino momentum, motivating the *Missing Mass Calculator* (MMC) technique [90] that was developed to reconstruct the mass of processes with two or more final state neutrinos (e.g. $H \to \tau\tau$, where each tau decays leaving a tau neutrino).

In this analysis, implementation of the MMC is more straightforward. Because there is only a single neutrino, an exact analytic expression for the z-component of the neutrino momentum can be obtained. If $m_\tau = 1.78$ GeV is the tau lepton mass, and $p_{\text{vis}}^\mu = (E_{\text{vis}}, \mathbf{p}_{\text{vis}})$ and $p_\nu^\mu = (E_\nu, \mathbf{p}_\nu)$, then

$$\begin{aligned} m_\tau^2 &= (p_{\text{vis}}^\mu + p_\nu^\mu)^2 \\ &= p_{\text{vis}}^\mu p_{\mu,\text{vis}} + p_\nu^\mu p_{\mu,\nu} + 2p_{\text{vis}}^\mu p_{\mu,\nu} \\ &= m_{\text{vis}}^2 + m_\nu^2 + 2p_{\text{vis}}^\mu p_{\mu,\nu} \end{aligned} \tag{4.8}$$

Making the approximation in Equation 4.8 that neutrinos are massless, defining $m_\gamma^2 = \frac{1}{2}(m_\tau^2 - m_{\text{vis}}^2)$, and noting that $E^2 = \text{p}^2 + m^2$, where $\text{p} = |\mathbf{p}|$, yields

$$\begin{aligned} m_\gamma^2 &= p_{\text{vis}}^\mu p_{\mu,\nu} = E_{\text{vis}} E_\nu - \mathbf{p}_{\text{vis}} \cdot \mathbf{p}_\nu \\ &= \text{p}_\nu E_{\text{vis}} - \mathbf{p}_{\text{vis}} \cdot \mathbf{p}_\nu \end{aligned} \tag{4.9}$$

Each of the momenta may be written as the sum of the transverse and z-components, $\mathbf{p} = \mathbf{p}_{\text{T}} + \mathbf{p}_z$. Rearranging Equation 4.9 gives:

$$\begin{aligned} m_\gamma^2 &= E_{\text{vis}} \sqrt{\text{p}_{\text{T},\nu}^2 + \text{p}_{z,\nu}^2} - \mathbf{p}_{\text{T,vis}} \cdot \mathbf{p}_{\text{T},\nu} - \text{p}_{z,\text{vis}} \text{p}_{z,\nu} \\ m_\gamma^2 + \mathbf{p}_{\text{T,vis}} \cdot \mathbf{p}_{\text{T},\nu} + \text{p}_{z,\text{vis}} \text{p}_{z,\nu} &= E_{\text{vis}} \sqrt{\text{p}_{\text{T},\nu}^2 + \text{p}_{z,\nu}^2} \end{aligned} \tag{4.10}$$

It is convenient to define the ancillary variable $E_\gamma^2 = m_\gamma^2 + \mathbf{p}_{\mathrm{T,vis}} \cdot \mathbf{p}_{\mathrm{T},\nu}$ and replace it in Equation 4.10, squaring the equation at the same time:

$$\begin{aligned}(E_\gamma^2 + \mathrm{p}_{z,\mathrm{vis}}\mathrm{p}_{z,\nu})^2 &= E_\gamma^4 + 2E_\gamma^2\mathrm{p}_{z,\mathrm{vis}}\mathrm{p}_{z,\nu} + \mathrm{p}_{z,\mathrm{vis}}^2\mathrm{p}_{z,\nu}^2 = E_{\mathrm{vis}}^2(\mathrm{p}_{\mathrm{T},\nu}^2 + \mathrm{p}_{z,\nu}^2) \\ (E_{\mathrm{vis}}^2 - \mathrm{p}_{z,\mathrm{vis}}^2)\mathrm{p}_{z,\nu}^2 - 2E_\gamma^2\mathrm{p}_{z,\mathrm{vis}}\mathrm{p}_{z,\nu} - (E_\gamma^4 - E_{\mathrm{vis}}^2\mathrm{p}_{\mathrm{T},\nu}^2) &= 0 \end{aligned} \tag{4.11}$$

Equation 4.11 is a simple quadratic equation in $\mathrm{p}_{z,\nu}$. Solving it yields

$$\mathrm{p}_{z,\nu}^{\pm} = \frac{E_\gamma^2\mathrm{p}_{z,\mathrm{vis}} \pm \sqrt{E_\gamma^4\mathrm{p}_{z,\mathrm{vis}}^2 + (E_{\mathrm{vis}}^2 + \mathrm{p}_{z,\mathrm{vis}}^2)(E_\gamma^4 - E_{\mathrm{vis}}^2\mathrm{p}_{\mathrm{T},\nu}^2)}}{E_{\mathrm{vis}}^2 - \mathrm{p}_{z,\mathrm{vis}}^2} \tag{4.12}$$

Equation 4.11 is an exact expression for the longitudinal neutrino momentum. However, the following questions may be noted:

1. How is the mass of the visible component of the hadronic tau determined, considering that hadronic taus are treated as massless in reconstruction?
2. Equation 4.11 actually lists *two* solutions. Since the neutrino can only have one longitudinal momentum, which solution is correct?
3. Detector resolution and other effects smear the true value of a quantity. It may be that the measured value is incompatible with the equation and results in no real solution. How are these effects considered?

Question 1 is addressed by considering the nature of hadronic tau decays. Hadronic taus used in this analysis are either one- or three-prong hadronic taus. The decay from the tau to the final one-prong form usually proceeds through the ρ resonance, which has a mass of $\sim$0.8 GeV. Decays to three-prong hadronic taus usually proceed through the a_1 resonance, which has a mass of $\sim$1.2 GeV. Thus, these values are used for the value of m_{vis} based on whether the reconstructed tau has one or three tracks.

Addressing Questions 2 and 3 requires some consideration of probability since there is no way *a priori* to know the true values of the measured variables (smeared by detector effects) or the true neutrino longitudinal momentum. Consider Question 3. The kinematics of the hadronic tau are taken from calorimeter clusters, while the transverse kinematics of the neutrino use the measurement of $E_{\mathrm{T}}^{\mathrm{miss}}$. The resolution of $E_{\mathrm{T}}^{\mathrm{miss}}$ is determined to be the measured quantity having the largest impact on the convergence of solutions.[27] The $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution for LFV Higgs boson events is found to be $\sigma_{\mathrm{MET}} = 5.94 + 0.84\sqrt{E_{\mathrm{T}}^{\mathrm{miss}}}$ GeV. Assuming that the distribution of $E_{\mathrm{T}}^{\mathrm{miss}}$ is Gaussian, a bivariate distribution for $E_{\mathrm{T}}^{\mathrm{miss}}$ in terms of its x- and y-components may be written as shown in Equation 4.13:

$$\mathcal{P}_{\mathrm{MET}}(X,Y) = \frac{1}{2\pi\sigma_{\mathrm{MET}}^2}\exp\left(-\frac{(X - E_{\mathrm{T},x}^{\mathrm{miss}})^2 + (Y - E_{\mathrm{T},y}^{\mathrm{miss}})^2}{2\sigma_{\mathrm{MET}}^2}\right) \tag{4.13}$$

[27] Variations of the hadronic tau energy within its resolution and the distribution of hadronic tau mass were also considered, but were found to have a negligible effect on the convergence relative to $E_{\mathrm{T}}^{\mathrm{miss}}$.

To determine the best value of $E_{\mathrm{T}}^{\mathrm{miss}}$ to use, 1600 different values (40 different values each in x and y) within a 3σ window of the peak value of the distribution in Equation 4.13 were tried in Equation 4.11. If the trial value of $E_{\mathrm{T}}^{\mathrm{miss}}$ resulted in a converging reconstructed mass, then its probability, determined from Equation 4.13, was recorded.

Each value of $E_{\mathrm{T}}^{\mathrm{miss}}$ resulting in a converging solution still has two possible solutions for the neutrino longitudinal momentum, as raised in Question 2. To resolve the ambiguity, the probability of the angle between the neutrino and reconstructed tau given the full tau momentum (see [90] for details) is calculated. This probability is multiplied by the $E_{\mathrm{T}}^{\mathrm{miss}}$ probability to yield a total probability. The solution with the highest total probability is chosen to reconstruct the Higgs boson mass in Equation 4.7.

Chapter 5

Analysis Strategy

The goal of the analysis presented in this dissertation is to detect and measure any lepton-flavor-violating Higgs boson decay processes. Chapter 2 discussed the theoretical motivations of these processes, while Chapters 3 and 4 gave a description of the detector and how physics objects are reconstructed. This chapter provides an overview of the rest of the dissertation, outlining the general analysis strategy.

To look for LFV Higgs boson decays, it is important to understand how such a decay might appear in the detector. Figure 5.1 shows a diagram of the LFV Higgs decays. Three objects are produced in the decays of interest in this analysis: a muon/electron, a hadronic tau, and a neutrino. The muon/electron and hadronic tau are reconstructed and identified using the methods described in Chapter 4. The neutrino cannot be directly observed, but its presence is inferred from the missing transverse energy measured in the event. Thus, possible LFV events should contain these objects. Additionally, since the Higgs boson and neutrino are electrically neutral, the reconstructed muon/electron and hadronic tau should have the opposite electric charge. The reconstructed objects should also have high transverse momentum since the Higgs boson is heavy.

If the LFV Higgs boson decay was the only process that produced a muon/electron, hadronic tau, and missing transverse energy, then finding one event with those objects would indicate the existence of those decays. However, other physics processes can have these objects in their final states. These processes are called *background processes* and need to be accounted in the data to determine whether or not LFV Higgs boson decays (known as the *signal processes*) exist. Accurate accounting of the data requires precise modeling of the signal and background processes, performed using simulated data and other techniques described in Chapter 6.

Once the signal and background processes are identified and modeled, data events need to be selected that have the highest sensitivity to the signal processes. Quality requirements also need to be applied to ensure that the reconstructed objects analyzed are most consistent with the signal process. This set of selection criteria is described in Chapter 7.

Selected data events are then categorized in *regions of interest*, called *signal* and *control regions*, based on whether the kinematic properties of the constituent objects are more

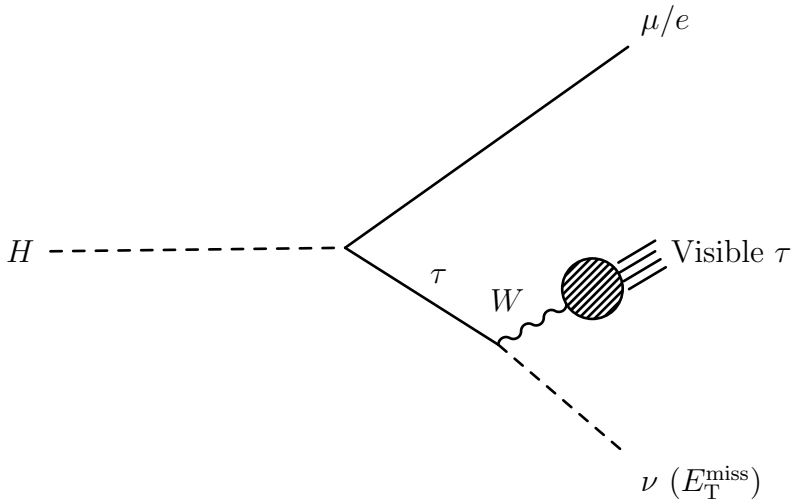


Figure 5.1: Diagram depicting the decays $H \to \mu\tau$ / $H \to e\tau$. The visible part of the tau decay, usually consisting of one or three charged pions and two or fewer neutral pions, is identified as a hadronic tau. The neutrino presence in the detector is characterized by missing transverse energy.

consistent with the signal or background. This categorization serves two purposes: it offers further improvement to the sensitivity of the analysis by identifying events that are clearly not consistent with the signal process, and it provides a means of verifying how accurately the background predictions match the data in regions where no signal is expected. These signal and control regions are described in Chapter 8.

No detector is a perfect instrument. Consequently, any measured quantity will have some degree of associated uncertainty. These *systematic uncertainties* affect the sensitivity and precision of a measurement and need to be estimated. The uncertainties considered in the analysis are described in Chapter 9.

Once the predictions of the signal and background processes and their associated systematic uncertainties are obtained, a *binned likelihood fit* of the signal and background processes to the data, using the reconstructed mass, is performed in the signal and control regions to determine the presence of the signal process. Reconstructed mass is chosen as the fitted distribution because it offers large discriminating power between signal and background processes. If no signal is found, an *exclusion limit* is computed, characterizing the degree of absence of the signal process. The structure of the binned likelihood fit and statistical techniques for determining signal presence and setting exclusion limits are described in Chapter 10. The final results of the fit are presented in Chapter 11.

Decisions about the objects and events used in the analysis have the capacity to introduce biases into the measurement. To avoid this, the analysis is performed *blinded*: the data in

the mass distribution where signal is expected are not viewed until all analysis selection requirements are finalized, and the binned likelihood fit structure is thoroughly checked and validated.

Chapter 6

Modeling of Signal and Background Processes

Data are collected through the trigger mechanisms described in Chapter 3. While a trigger mechanism can select events based on their object kinematics to prefer a particular process,[1] processes of interest (*signal processes*) generally have signatures mimicked by other processes (*background processes*). This is especially true for searches involving Higgs bosons, whose decay signatures are produced by SM processes with larger cross sections and hence, more plentiful. Figure 6.1 provides a summary of measured cross sections for different processes.

Consequently, understanding the composition of the data in terms of the background processes and accurately modeling the signal signature are critically important for a successful analysis. The signal kinematics and many of the expected backgrounds in this analysis are simulated by randomly sampling calculated kinematic distributions. This type of simulated data is called *Monte Carlo* (MC), named according to the sampling technique.

Generating an MC sample consists of several steps. The physics processes are first modeled using a *Monte Carlo generator* that simulates the proton–proton collision using some set of calculated *parton distribution functions* (PDFs), discussed in Chapter 2. The kinematics of the hard scatter process of interest and its decays are modeled by randomly sampling the calculated kinematic distributions. Finally, the generator computes the softer processes such as the underlying event, pile-up events, and the hadronization and showering of quarks and gluons. The response of the detector (that is, the electrical signals, or *hits*, produced) is then simulated using the *GEANT4* software package [91]. Physics objects are reconstructed from the simulated hits in the manner described in Chapter 4. Various corrections based on performance studies are then applied to the physics objects to bring the simulated data into better agreement with observed data. One such correction pertains to the distribution of pile-up events. MC-simulated data are usually produced before data from proton–proton collisions are recorded to expedite the progress of different physics analyses. Consequently,

[1] For example, events with Z bosons can be selected by high level dilepton triggers firing on events with two leptons of opposite charge whose invariant mass is in a window centered at 90 GeV.

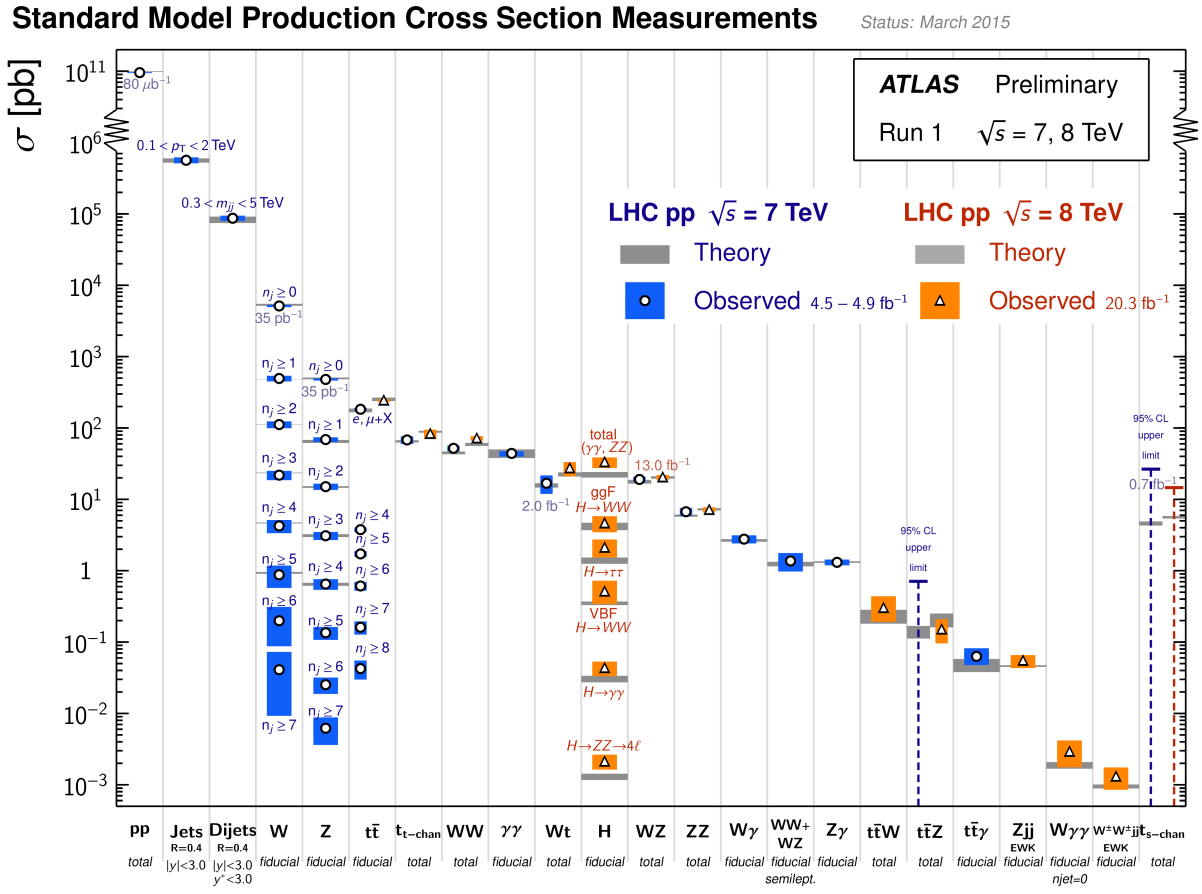


Figure 6.1: Measured cross sections for various SM processes at energies of 7 TeV and 8 TeV [67].

samples are produced with an approximate distribution of the pile-up events in observed data. As collision data are taken, MC samples are reweighted to accurately reflect the pile-up in observed data. Figure 6.2 shows the pile-up distribution from observed data used to reweight the 2012 MC samples.

MC-simulated data are typically produced with many more events than are present in observed data to reduce the sample statistical uncertainty. To bring the MC sample predictions into agreement with the observed data, it is necessary to weight each MC event by the following factor, shown in Equation 6.1:

$$S_i = \frac{L\sigma W_i}{N_{\mathrm{MC}}} \tag{6.1}$$

where S_i is the event weight, L is the integrated luminosity of the data, σ is the cross section of the process in question, W_i is an internal event weight assigned by the MC generator, and

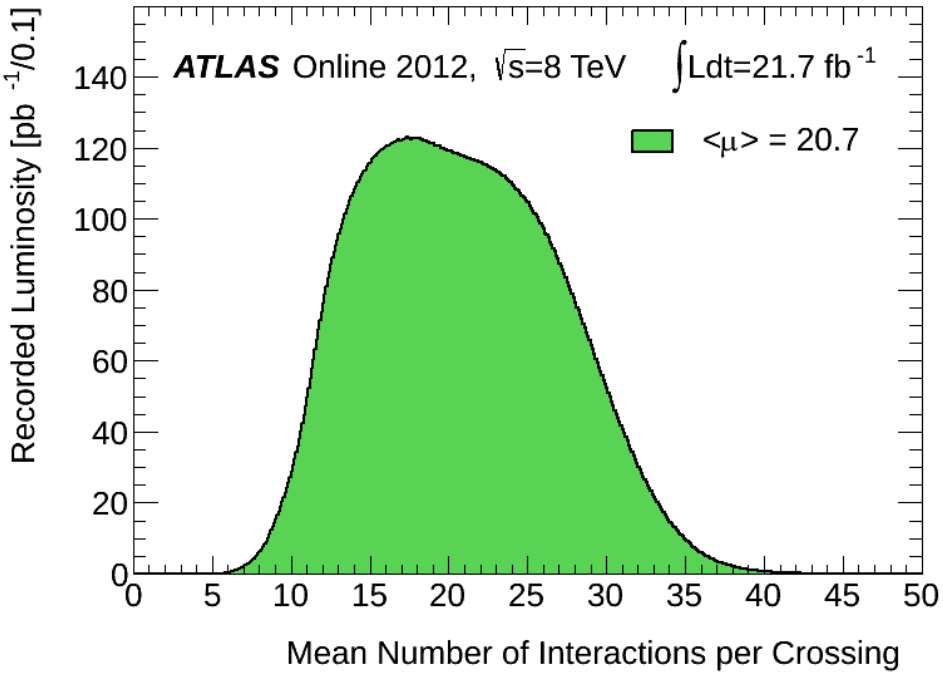


Figure 6.2: Mean Number of Interactions per Bunch Crossing [67].

N_{MC} is the total number of simulated events. Equation 6.1 is motivated by Equation 3.2 in Chapter 3.

The relevant background processes in this analysis contain hadronic tau candidates and leptons, including:

- a single W boson and zero or more jets (W + jets events)
- a single Z boson, decaying to either a hadronic and leptonic tau or two leptons, and zero or more jets ($Z \to \tau\tau/\ell\ell$ + jets events)
- one or more top quarks (top events)
- two heavy vector bosons (VV, or diboson events)
- two or more jets (multijet, or QCD events)
- a Higgs boson decaying to a hadronic and leptonic tau[2] ($H \to \tau\tau$ events)

This chapter first describes the modeling of the signal processes and which Higgs boson production modes are considered. The method for estimating the background processes is then explained. Specific details for estimating each of the listed processes are then discussed.

[2]This analysis is probably the first ATLAS analysis to consider this background, as this process was only definitively discovered in 2013.

6.1 Signal

Higgs bosons are produced at the LHC through gluon–gluon fusion, vector boson fusion, associated production with vector bosons, and associated production with top quarks. The first three production mechanisms are considered in the modeling of signal processes.

The Higgs boson kinematics in the gluon–gluon fusion and VBF signal samples are simulated using the POWHEG MC generator [92], which makes next-to-leading order (NLO) calculations, ensuring higher accuracy in the calculation. For the associated production (WH and ZH) samples, the Higgs boson kinematics are modeled only to leading order, using the PYTHIA8 MC generator [93], due to the smaller cross sections associated with these processes. The parton showering, hadronization, and underlying event in all signal samples are modeled by PYTHIA8. The CT10 parton distribution functions [94] are used for the gluon–gluon fusion and VBF samples, while the CTEQ6L1 PDFs [95] are used for the associated production samples.

The lepton-flavor-violating decays of the Higgs boson into a tau lepton and muon/electron are modeled by EvtGen 2.0 [96], using a generic phase space model.[3] The taus produced in LFV decays are assumed to be unpolarized.[4] The decays of the tau lepton are modeled by TAUOLA [97]. Photon radiation from the charged leptons is modeled using PHOTOS [98].

The signal samples are normalized, assuming a 1% branching ratio to the LFV decay, to the Higgs production cross sections calculated at next-to-next-to-leading-order (NNLO) [99]–[101] and shown in Table 2.2.

6.2 Background

Accurate modeling of the expected background contributions in data is an essential component of any analysis. This section discusses the method used to estimate the background contributions in the observed data and how each background contribution is modeled.

6.2.1 Method of Estimation

As depicted in Figure 5.1, there are two visible decay products from an LFV Higgs decay: a muon/electron and a hadronic tau. These decay products are electrically charged, and the product of the electric charges can be used to classify the event. In *opposite-sign* (OS) events,

[3]Models by EvtGen that are not generic include CP violating decays, Dalitz decays, and quark-flavor-mixing models.

[4]The tau polarization refers to any asymmetry in the production cross sections for left- and right-handed taus. Because taus are short-lived and the resultant (anti)neutrinos are always (right-)left-handed, the handedness of the tau affects the tau decay product kinematics observed in the detector. Thus, tau polarization is an important consideration for modeling decays of bosons to taus. For Higgs boson decays, the taus have opposite spins from conservation of angular momentum, while in Z boson decays, the orientation of the tau spins depends on the spin of the Z boson.

the muon/electron and hadronic tau have different electric charges, or $q_{\mu/e}q_{\tau_{\text{had}}} = -1$. *Same-sign* (SS) events are those for which the muon/electron and hadronic tau have the same electric charge, or $q_{\mu/e}q_{\tau_{\text{had}}} = 1$. Since electric charge is a conserved quantity, the visible decay products in Higgs boson decays must have opposite charge. Consequently, the signal may only be found in OS events.

Background processes can also be classified into two categories. *Electroweak* (EWK) processes are mediated by a photon, W, or Z boson. *Multijet*, or *Quantum-Chromodynamical* (QCD), processes are mediated by the strong interaction (gluons).

The number of background processes may be written as shown in Equation 6.2:

$$N_{\text{background}}^{\text{Data},OS} = N_{\text{QCD}}^{\text{Data},OS} + N_{\text{EWK}}^{\text{Data},OS} \tag{6.2}$$

An accurate estimate of the total number of background events requires an accurate estimate of the number of events from both types of backgrounds.

The number of EWK background events can be estimated using MC-simulated data. However, depending on the process, the estimated number of events can be inaccurate due to the composition of hadronic taus in the sample. Reconstructed hadronic taus fall into one of two categories: true hadronic taus or misidentified jets or leptons passing tau identification. Taus in the first category are called *true taus*, while those in the latter category are *misidentified taus*. The modeling of misidentified taus is difficult to accurately simulate, leading to inaccuracies in the estimated number of events and their kinematics.

To ensure consistency between the data and MC-simulated events, a scale factor called a *k-factor* is applied to MC predictions for a particular process, computed as $k = N(\text{Data})/N(\text{MC})$ in a dedicated control region where the process in question predominates. The number of EWK events in data can thus be determined as:

$$\begin{aligned} N_{\text{EWK}}^{\text{Data},OS} &= N_{W+\text{jets}}^{\text{Data},OS} + N_{Z\to\tau\tau}^{\text{Data},OS} + N_{Z\to\ell\ell}^{\text{Data},OS} + N_{\text{Top}}^{\text{Data},OS} + N_{VV}^{\text{Data},OS} + N_{H\to\tau\tau}^{\text{Data},OS} \\ N_{\text{EWK}}^{\text{Data},OS} &= k_{W+\text{jets}}^{OS} N_{W+\text{jets}}^{\text{MC},OS} + k_{Z\to\tau\tau}^{OS} N_{Z\to\tau\tau}^{\text{MC},OS} + k_{Z\to\ell\ell}^{OS} N_{Z\to\ell\ell}^{\text{MC},OS} + k_{\text{Top}}^{OS} N_{\text{Top}}^{\text{MC},OS} \\ &\quad + k_{VV}^{OS} N_{VV}^{\text{MC},OS} + k_{H\to\tau\tau}^{OS} N_{H\to\tau\tau}^{\text{MC},OS} \end{aligned} \tag{6.3}$$

Replacing Equation 6.3 in Equation 6.2 yields

$$\begin{aligned} N_{\text{background}}^{\text{Data},OS} = N_{\text{QCD}}^{\text{Data},OS} &+ k_{W+\text{jets}}^{OS} N_{W+\text{jets}}^{\text{MC},OS} + k_{Z\to\tau\tau}^{OS} N_{Z\to\tau\tau}^{\text{MC},OS} + k_{Z\to\ell\ell}^{OS} N_{Z\to\ell\ell}^{\text{MC},OS} \\ &+ k_{\text{Top}}^{OS} N_{\text{Top}}^{\text{MC},OS} + k_{VV}^{OS} N_{VV}^{\text{MC},OS} + k_{H\to\tau\tau}^{OS} N_{H\to\tau\tau}^{\text{MC},OS} \end{aligned} \tag{6.4}$$

Estimating the number of QCD events is more complicated. Unfortunately, MC-simulated data cannot accurately estimate the contribution from events where jets are misidentified as both the hadronic tau and lepton. Therefore, observed data are used to estimate the contribution from this background. Same-sign data events are used for this estimate since no signal is expected to be found in these events. SS data events are comprised not only of QCD events, but also EWK ones that can be estimated using MC. The estimated number of SS QCD events may be computed as follows:

$$
\begin{aligned}
N_{\text{QCD}}^{\text{Data},SS} &= N_{\text{data}}^{SS} - N_{W+\text{jets}}^{\text{Data},SS} - N_{Z\to\tau\tau}^{\text{Data},SS} - N_{Z\to\ell\ell}^{\text{Data},SS} - N_{\text{Top}}^{\text{Data},SS} - N_{VV}^{\text{Data},SS} - N_{H\to\tau\tau}^{\text{Data},SS} \\
N_{\text{QCD}}^{\text{Data},SS} &= N_{\text{data}}^{SS} - k_{W+\text{jets}}^{SS} N_{W+\text{jets}}^{\text{MC},SS} - k_{Z\to\tau\tau}^{SS} N_{Z\to\tau\tau}^{\text{MC},SS} - k_{Z\to\ell\ell}^{SS} N_{Z\to\ell\ell}^{\text{MC},SS} - k_{\text{Top}}^{SS} N_{\text{Top}}^{\text{MC},SS} \\
&\quad - k_{VV}^{SS} N_{VV}^{\text{MC},SS} - k_{H\to\tau\tau}^{SS} N_{H\to\tau\tau}^{\text{MC},SS}
\end{aligned} \tag{6.5}
$$

The number of OS QCD events can be determined if the ratio of OS to SS QCD events (r_{QCD}) is known, computed as $N_{\text{QCD}}^{Data,OS} = r_{\text{QCD}} N_{\text{QCD}}^{Data,SS}$. Substituting this expression into Equation 6.5 yields

$$
\begin{aligned}
N_{\text{background}}^{\text{Data},OS} = r_{\text{QCD}} \cdot (&N_{\text{data}}^{SS} - k_{W+\text{jets}}^{SS} N_{W+\text{jets}}^{\text{MC},SS} - k_{Z\to\tau\tau}^{SS} N_{Z\to\tau\tau}^{\text{MC},OS} - k_{Z\to\ell\ell}^{SS} N_{Z\to\ell\ell}^{\text{MC},SS} \\
&- k_{\text{Top}}^{SS} N_{\text{Top}}^{\text{MC},SS} - k_{VV}^{SS} N_{VV}^{\text{MC},SS} - k_{H\to\tau\tau}^{SS} N_{H\to\tau\tau}^{\text{MC},SS}) + k_{W+\text{jets}}^{OS} N_{W+\text{jets}}^{\text{MC},OS} \\
&+ k_{Z\to\tau\tau}^{OS} N_{Z\to\tau\tau}^{\text{MC},OS} + k_{Z\to\ell\ell}^{OS} N_{Z\to\ell\ell}^{\text{MC},OS} + k_{\text{Top}}^{OS} N_{\text{Top}}^{\text{MC},OS} + k_{VV}^{OS} N_{VV}^{\text{MC},OS} \\
&+ k_{H\to\tau\tau}^{OS} N_{H\to\tau\tau}^{\text{MC},OS}
\end{aligned} \tag{6.6}
$$

Equation 6.6 can be rewritten as

$$
\begin{aligned}
N_{\text{background}}^{\text{Data},OS} = r_{\text{QCD}} \cdot N_{\text{data}}^{SS} &+ N_{W+\text{jets}}^{OS-SS} + N_{Z\to\tau\tau}^{OS-SS} + N_{Z\to\ell\ell}^{OS-SS} + N_{\text{Top}}^{OS-SS} + N_{VV}^{OS-SS} \\
&+ N_{H\to\tau\tau}^{OS-SS}
\end{aligned} \tag{6.7}
$$

where

$$
N_{\text{background}}^{OS-SS} = k^{OS} N_{\text{background}}^{OS} - r_{\text{QCD}} \cdot k^{SS} N_{\text{background}}^{SS} \tag{6.8}
$$

The values of r_{QCD} used in this analysis were measured in the ATLAS search for $H \to \tau\tau$ decays [102]–[104]. The presented background estimation method relies on two important assumptions:

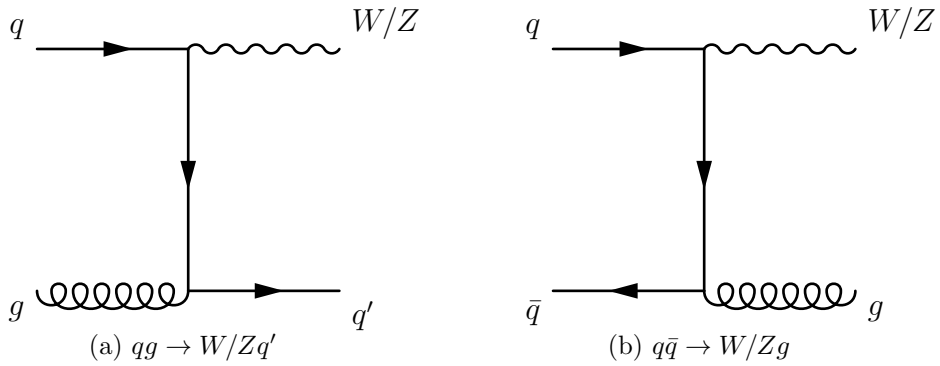
- The measured k-factors for each background process, obtained from events in dedicated control regions, are valid for corresponding events in the signal regions.
- The measured value of r_{QCD} is valid over all regions of interest.

These assumptions were checked and validated, which will be discussed in later sections.

6.2.2 W + jets

W + jets processes are the dominant background in the analysis. Leading order diagrams describing this process are shown in Figure 6.3, with the quark–gluon interaction (Figure 6.3a) occurring predominately over the quark–antiquark contribution (Figure 6.3b) due to the higher gluon luminosity in the protons. The LFV signature is mimicked by leptonic

decays of the W boson producing the light lepton and missing transverse energy and the jet misidentified as a hadronic tau.[5] Events with the quark jet and W boson final state in Figure 6.3a are expected to have a high degree of charge correlation ($N_{OS} \gg N_{SS}$), while those with the gluon jet and W boson final state in Figure 6.3b should have no charge correlation ($N_{OS} \sim N_{SS}$). Considering both contributions, the number of OS W + jets events is expected to be larger than the number of SS events ($N_{OS} > N_{SS}$).



(a) $qg \to W/Zq'$ (b) $q\bar{q} \to W/Zg$

Figure 6.3: Examples of leading order diagrams for W/Z + jets production.

The kinematics of the W + jets background are modeled using ALPGEN [105]. Leading order processes with five and fewer jets are considered. The W boson decays, parton showering, hadronization, and underlying event are modeled using PYTHIA8. The CTEQ6L1 parton distribution functions are used for the factorization. PHOTOS is used to model photon radiation from the lepton produced in the W boson decay.

OS and SS W + jets event rates are normalized to data in a dedicated control region in which W + jets is the dominant process. This control region is called the *W + jets control region* (WCR) and is discussed in Section 8.4. The k-factors used to normalize the W + jets backgrounds are obtained as shown in Equation 6.9:

$$k^{OS/SS}_{W+\text{jets}} = \frac{N^{OS/SS}_{\text{WCR Data}} - N^{OS/SS}_{\text{WCR EWK}}}{N^{OS/SS}_{\text{WCR } W+\text{jets}}} \tag{6.9}$$

where $N^{OS/SS}_{\text{WCR Data}}$, $N^{OS/SS}_{\text{WCR } W+\text{jets}}$, and $N^{OS/SS}_{\text{WCR EWK}}$ are the number of OS(SS) events respectively in observed data, W + jets, and other background processes in the WCR.[6]

Discrepancies are also observed between the data and predicted background processes in various kinematic distributions, indicating shape mismodeling in the W + jets MC-simulated

[5]The W boson can also decay to a tau lepton and the jet can be misidentified as an electron or muon, but this misidentification rate is small enough that this process may be considered negligible.

[6]In principle, this approach requires that all other background processes be properly normalized in data as well. However, this can lead to a chicken-and-egg scenario in which it is not immediately certain which background to normalize first. This ambiguity is resolved by normalizing the background whose control region has the smallest amount of contamination by other processes.

data. This is corrected by applying scale factors dependent on $p_{\mathrm{T}}(\tau_{\mathrm{had}})$, $|\Delta\eta(\mu/e, \tau_{\mathrm{had}})|$, and the number of jets in the event, using the object selection criteria outlined in Chapter 7. Shape corrections are obtained separately for OS and SS events in the WCR and for the total OS–SS W + jets prediction in events with $m_{\mu/e\tau}^{\mathrm{MMC}} > 150$ GeV in *Signal Region 1* (SR1).[7] The corrections are obtained as shown in Equation 6.10:

$$k_{W+\mathrm{jets}}^{shape} = \frac{N_{\mathrm{WCR\ Data}}^{bin} - N_{\mathrm{WCR\ EWK}}^{bin}}{N_{\mathrm{WCR}\ W+\mathrm{jets}}^{bin}} \tag{6.10}$$

where the "shape" and "bin" labels indicate that the correction depends on particular values of $p_{\mathrm{T}}(\tau_{\mathrm{had}})$, $|\Delta\eta(\mu/e, \tau_{\mathrm{had}})|$, and the number of jets in the event. An additional correction to the W + jets normalization is applied in SR1 based on those events with $m_{\mu/e\tau}^{\mathrm{MMC}} > 150$ GeV.[8]

Further details regarding the shape correction procedure and numerical values of the corrections may be found in Appendix A. The calculated k-factors and SR1 normalizations are shown in Table 6.1.

Additionally, W + jets samples modeled by ALPGEN, but with parton showering and hadronization simulated by HERWIG [106] and underlying event simulated by JIMMY [107], are used to determine systematic uncertainties associated with the extrapolation of the W + jets k-factors (see Section 9.3.4).

6.2.3 $Z \to \tau\tau$ + jets

$Z \to \tau\tau$ + jets events are another large background in the analysis. The diagrams for this process are analogous to those for W + jets and are shown in Figure 6.3. A crucial difference is that the Z boson decays to two true tau leptons, one of which decays hadronically and the other, leptonically. Consequently, a high degree of charge correlation is expected for $Z \to \tau\tau$ events due to charge conservation ($N_{OS} \gg N_{SS}$). This background is an *irreducible* background because its signature does not come from misidentified objects (that could be eliminated through improved reconstruction and modeling, in principle). Rather, the final state objects in the LFV signature also appear in the $Z \to \tau\tau$ signature (true hadronic tau, true light lepton, and true missing transverse energy). This background was a major obstacle for the ATLAS SM $H \to \tau\tau$ analysis [30], but its impact is lessened in this analysis due to the topological differences between the LFV and $Z \to \tau\tau$ processes. Specifically, an LFV event has one neutrino, resulting in better-resolved missing transverse energy compared to $Z \to \tau\tau$ events, which have three nearly back-to-back neutrinos, causing a partial cancellation of the missing transverse energy that degrades the resolution. This results in better mass separation between the $Z \to \tau\tau$ and LFV Higgs boson events.

Because $Z \to \tau\tau$ was the dominant and most difficult background in the SM $H \to \tau\tau$ analysis, it was modeled using a data-driven technique called *embedding*, precluding modeling inaccuracies in MC-simulated data that would degrade the measurement [108]. This

[7] SR1 is described in Section 8.2

[8] That is, the correction is applied to the total OS–SS background prediction, obtained using an analogous form of Equation 6.9 with OS data, and OS–SS backgrounds and W + jets events.

technique consists of identifying a pure sample of $Z \to \mu\mu$ events in data, replacing the muons with taus from a simulated $Z \to \tau\tau$ event with the simulated Z boson kinematics equivalent to that in observed data, then decaying the taus using an MC generator. The resulting hybrid dataset is called an *embedding sample* and is used in this analysis to model the $Z \to \tau\tau$ background. The advantage of this method is that the only components relying on simulation are the tau decays, which are well-modeled. The parton showering, hadronization, underlying event, missing transverse energy, Z boson kinematics, and other environmental considerations, such as multiple interactions, are taken directly from data, eliminating the systematic uncertainties associated with simulation.[9]

The embedding procedure is outlined in a flowchart, shown in Figure 6.4. $Z \to \mu\mu$ events are selected from data using dimuon and single lepton triggers with the muons satisfying various p_{T} requirements. Importantly, selected muons are required to be *isolated* in the inner detector, meaning that the sum of the p_{T} of all tracks in a cone centered around the muon, with size $\Delta R < 0.4$, has to be less than 20% of the muon p_{T}, or $I(p_{\mathrm{T}}, 0.4)/p_{\mathrm{T}}(\mu) < 0.2$. The muons are also required to have opposite charge and have an invariant mass larger than 40 GeV.

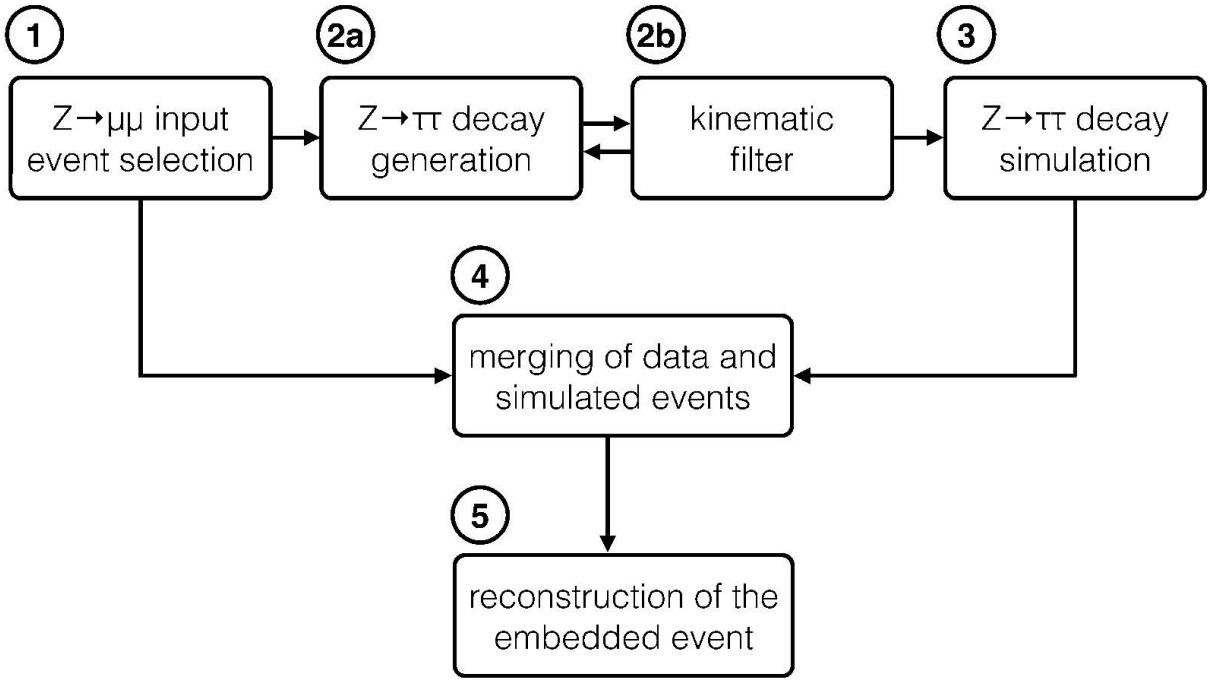


Figure 6.4: A flowchart illustrating the steps in the embedding procedure for modeling the $Z \to \tau\tau$ background [108].

[9]Ideally, $Z \to \tau\tau$ events in observed data would be used to model this background. However, it is practically difficult to identify a large, pure sample of such events in data, necessitating the use of $Z \to \mu\mu$ events.

Reconstructed muons in data are replaced with simulated tau leptons having the same energy as the reconstructed muon, but with momentum rescaled to account for the significantly larger tau mass. The taus are then decayed using TAUOLA, which properly accounts for the tau polarization. PHOTOS is used to treat the photon radiation. A kinematic filter is then applied to boost the fraction of events passing the analysis selection criteria and consists of TAUOLA performing 1000 different decays of each pair of taus for a given event in data. The first event whose decay product kinematics pass threshold criteria is kept. This threshold requirement introduces a bias into the kinematic distributions of the decay products, which is corrected through reweighting. The detector response to each decay product is then simulated without consideration of noise effects to avoid double-counting with the data.

The simulated tau decays are merged into the observed data event by removing the original muon tracks and energy deposits in the calorimeters. To remove the energy deposits in the calorimeters, $Z \to \mu\mu$ events are simulated using MC without any pile-up or underlying event. The energy deposits in the calorimeter from the simulated muons are subtracted from those in observed data. The tracks and energy deposits from the simulated tau decays are then inserted into the event (the energy deposits are added to existing calorimeter cells, including those from which the original muon energies were subtracted out). The resulting hybrid detector response is reconstructed into physics objects, using the standard ATLAS algorithms (except track reconstruction). Two studies were performed to validate the method, certifying it for use in physics analyses. The first compared $Z \to \mu\mu$ events in data with an embedding sample in which simulated muons were embedded into the event, while the second compared the tau-embedded $Z \to \tau\tau$ sample with MC-simulated $Z \to \tau\tau$ events.

Each event in the embedding sample is weighted only by the MC event weight (W_i) associated with the tau decay simulation, rather than the expression in Equation 6.1, since the embedding is derived from observed data. The number of $Z \to \tau\tau$ events in the embedding sample is adjusted to the number of events expected in observed data by directly normalizing it to the mass range corresponding to the Z boson mass peak in *Signal Region 2* (SR2), described in Section 8.3. In the muon measurement, this normalization is performed over the range 60 GeV $< m_{\mu\tau}^{\text{MMC}} <$ 110 GeV, while in the electron measurement, it is performed over the range 60 GeV $< m_{e\tau}^{\text{MMC}} <$ 90 GeV. The range in the electron measurement is narrower due to the larger presence of $Z \to ee(e \to \tau)$ events under the Z boson mass peak. The normalization is computed using the expression in Equation 6.11:

$$k_{Z\to\tau\tau}^{OS-SS} = \frac{N_{\text{SR2 Data}}^{OS} - N_{\text{SR2 Other Bkgd}}}{N_{\text{SR2 } Z\to\tau\tau}^{OS-SS}} \tag{6.11}$$

where $N_{\text{SR2 Other Bkgd}}$ refers to all OS–SS background predictions of the other processes listed at the beginning of the chapter (the contribution by SS data events is also considered). The OS–SS background predictions must be used rather than the OS predictions due to the multijet background present in SR2.

6.2.4 $Z \to \mu\mu/ee$ + jets

Z bosons also decay to pairs of light leptons, $Z \to \mu\mu/ee$, that are another background process. The diagrams for these processes are identical to the $Z \to \tau\tau$ case and are shown in Figure 6.3. The major difference between these processes and the $Z \to \tau\tau$ process is that there are no true tau leptons present in the final state. The LFV decay signature is mimicked by a real lepton from the Z boson decay, a misidentified hadronic tau, and missing transverse energy from the environment (or jets in the event). The hadronic tau can be misidentified from either a jet in the event or from the other lepton in the Z boson decay. In the former case, the second lepton from the Z boson decay is outside the tracking acceptance of the detector or is poorly reconstructed. $Z \to \mu\mu/ee$ processes in which the lepton is misidentified as a hadronic tau are expected to have a high degree of charge correlation ($N_{OS} \gg N_{SS}$), while those in which the jet is misidentified as a tau should have no charge correlation since both leptons are equally likely to be mis-reconstructed ($N_{OS} \sim N_{SS}$).

The $Z \to \mu\mu/ee$ background is modeled nearly identically to the W + jets background, with the only difference being the process. The events are generated using Alpgen, with parton showering, hadronization, and underlying event simulated using Pythia8. The CTEQ6L1 parton distribution functions are used, and PHOTOS is used for the photon radiation.

$Z \to \mu\mu/ee$ events where a jet is misidentified as a hadronic tau are normalized to data in a dedicated $Z \to \mu\mu/ee$ *control region* (ZmmCR/ZeeCR, see Section 8.6). Due to the nature of the control region, a single normalization k-factor is obtained as shown in Equation 6.12:

$$k_{Z\to\ell\ell} = \frac{N^{OS+SS}_{Z\to\ell\ell\text{CR Data}} - N^{OS+SS}_{Z\to\ell\ell\text{CR EWK}} - N^{OS+SS}_{Z\to\ell\ell\text{CR } Z\to\ell\ell(\ell\to\tau)}}{N^{OS+SS}_{Z\to\ell\ell\text{CR } Z\to\ell\ell(\text{jet}\to\tau)}} \tag{6.12}$$

where $\ell = \mu, e$ and $OS + SS$ indicates that the opposite- and same-sign events are added together. The obtained k-factors are shown in Table 6.1.

Like the W + jets background, shape mis-modeling is observed in the $Z \to \mu\mu/ee$ background. A shape correction dependent upon $p_{\text{T}}(\tau_{\text{had}})$, $|\Delta\eta(\mu/e, \tau_{\text{had}})|$, and the number of jets in the event is derived in the ZmmCR/ZeeCR and applied to $Z \to \mu\mu/ee$ events where a jet is misidentified as a hadronic tau since MC is known not to accurately model these events. The correction is determined using Equation 6.13, analogous to Equation 6.10:

$$k^{shape}_{Z\to\ell\ell} = \frac{N^{bin}_{Z\to\ell\ell\text{CR Data}} - N^{bin}_{Z\to\ell\ell\text{CR EWK}} - N^{bin}_{Z\to\ell\ell\text{CR } Z\to\ell\ell(\ell\to\tau)}}{N^{bin}_{Z\to\ell\ell\text{CR } Z\to\ell\ell(\text{jet}\to\tau)}} \tag{6.13}$$

As with Equation 6.10, the "shape" and "bin" labels indicate that the correction is a function of $p_{\text{T}}(\tau_{\text{had}})$, $|\Delta\eta(\mu/e, \tau_{\text{had}})|$, and the number of jets in the event.

The treatment of the $Z \to \mu\mu/ee$ events where the lepton is misidentified as a hadronic tau differs in the muon and electron measurements. In the muon measurement, the misidentification rate of muons as hadronic taus is small due to the efficiency of the muon reconstruction algorithms. Consequently, the corresponding $Z \to \mu\mu$ background is small. Hence,

the normalization and shape of the $Z \to \mu\mu$ background where muons are misidentified as hadronic taus are taken directly from the MC-simulated data weighted according to Equation 6.1. No further corrections are applied.

The situation is different in the electron measurement where the misidentification rate of electrons to hadronic taus is larger than for muons. Consequently, $Z \to ee$ events where the electron is misidentified as the hadronic tau are more prominent, particularly under the Z boson mass peak in SR2. There is a 1.3σ difference between observed data and the predicted number of $Z \to ee$ events under this peak due to inaccurate corrections applied to the misidentification rate of electrons as hadronic taus. Therefore, the number of $Z \to ee$ events with misidentified electrons is normalized to the observed data over the mass range 90 GeV $< m_{e\tau}^{\text{MMC}} <$ 110 GeV, calculated according to Equation 6.14:

$$k^{OS-SS}_{Z\to ee(e\to\tau)} = \frac{N^{OS}_{\text{SR2 Data}} - N_{\text{SR2 Other Bkgd}}}{N^{OS-SS}_{\text{SR2 } Z\to ee(e\to\tau)}} \tag{6.14}$$

where the other background term includes the predictions of $Z \to ee$ events where the jet is misidentified as a hadronic tau. The calculated normalization factor is shown in Table 6.1.

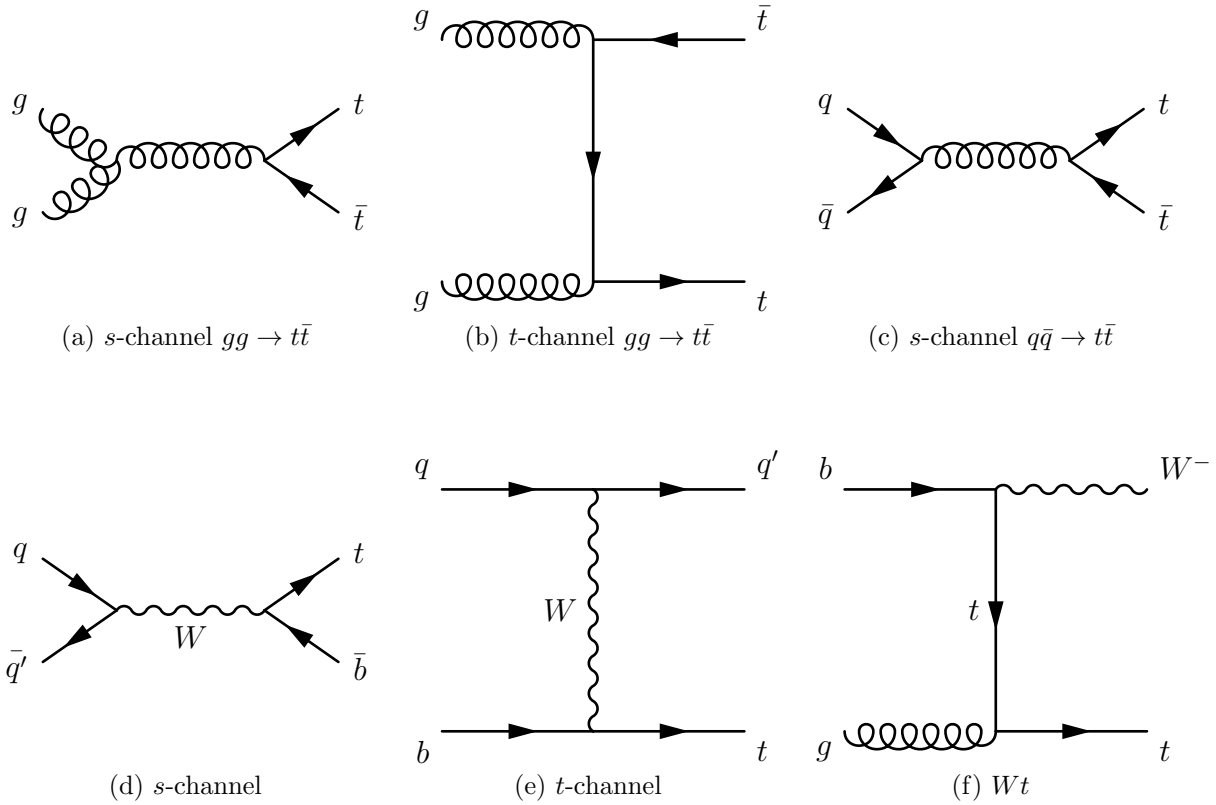
6.2.5 Top quark

Processes involving top quarks also contribute to the background, albeit less than those involving W/Z + jets. Events with top quarks are usually characterized by the presence of one or more b-tagged, high-p_{T} jets due to the mass difference between the top quark and b-quark. Top quarks are often produced in pairs ($t\bar{t}$), but are also produced individually. Diagrams of both processes are shown in Figure 6.5. The top quarks then decay into a W boson and b-quark. There are a myriad of ways that the LFV signature can be mimicked in top quark processes, as hadronic taus can be reconstructed from true taus as well as misidentified from jets. Leptons arise from the W bosons produced in heavy flavor decays. Missing transverse energy comes from the neutrinos produced in leptonic W boson decays and from the jets in the event. Top quark events are expected to have a high degree of charge correlation due to the event topology ($N_{OS} \gg N_{SS}$).

The $t\bar{t}$, s-channel single top, and Wt single top processes are modeled using Powheg. t-channel single top events are modeled using AcerMC [109]. The parton showering and hadronization are performed by Herwig, and the underlying event is modeled by Jimmy. The tau decays and photon radiation are handled respectively by TAUOLA and PHOTOS. The CT10 parton distribution functions are used for the factorization.

OS and SS top events are separately normalized to data in a dedicated *top control region* (TCR), described in Section 8.5. Each k-factor is computed as shown in Equation 6.15:

$$k^{OS/SS}_{\text{top}} = \frac{N^{OS/SS}_{\text{TCR Data}} - N^{OS/SS}_{\text{TCR EWK}}}{N^{OS/SS}_{\text{TCR Top}}} \tag{6.15}$$



(a) s-channel $gg \to t\bar{t}$ (b) t-channel $gg \to t\bar{t}$ (c) s-channel $q\bar{q} \to t\bar{t}$

(d) s-channel (e) t-channel (f) Wt

Figure 6.5: Examples of leading order diagrams for $t\bar{t}$ (upper row) and single top production (lower row).

The obtained k-factors are shown in Table 6.1.

Similarly to the W + jets background, the extrapolation of the top k-factors is validated using $t\bar{t}$, s-channel single top, and Wt single top events modeled using MC@NLO [110] (see Section 9.3.4). AcerMC t-channel single top samples are used to determine the extrapolation uncertainty.

6.2.6 Dibosons, VV

A small component of the background comes from diboson (WW, WZ, and ZZ) events. Examples of leading order diagrams for diboson production are shown in Figure 6.6. Like the top background, many different final states involving real leptons and both real and misidentified hadronic taus are possible. Strong charge correlation is expected in WW events ($N_{OS} \gg N_{SS}$), while the charge correlation in WZ and ZZ events depends on the respective boson decay products.

The ZZ and WZ processes are modeled using Herwig, while the WW process is modeled by Alpgen interfaced with Herwig. Herwig is used in all processes to model parton

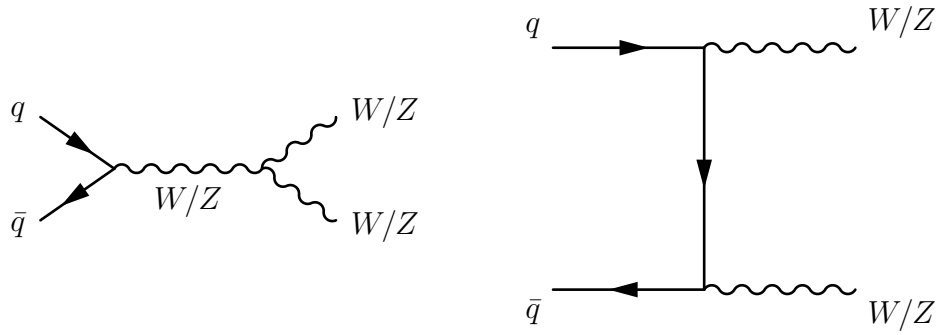


Figure 6.6: Examples of leading order diagrams for diboson production.

showering and hadronization, while JIMMY simulates the underlying event. TAUOLA and PHOTOS are used to respectively model tau decays and photon radiation, and the CTEQ6L1 parton distribution functions are used. Due to its small size, the diboson background is not normalized to data, but rather an NLO calculation of its cross section in QCD processes [111] using Equation 6.1.

6.2.7 Standard Model $H \to \tau\tau$

The tiniest background process considered in this analysis is the SM $H \to \tau\tau$ in which one tau decays hadronically while the other decays leptonically. These events have a signature similar to the $Z \to \tau\tau$ process and have high charge correlation ($N_{OS} \gg N_{SS}$).

This process is modeled almost exactly the same way as the signal Higgs boson LFV process. The only difference is that the Higgs boson decays to taus are modeled using PYTHIA8, rather than EvtGen 2.0. The $H \to \tau\tau$ samples are normalized to the NNLO Higgs boson production cross sections times $\mathrm{Br}(H \to \tau_{\mathrm{lep}}\tau_{\mathrm{had}}) = 2.88\%$ using Equation 6.1.

6.2.8 Same-Sign Data, Multijet Background, and r_{QCD}

Events with same-sign data are used to model the multijet background due to the practical difficulties in simulating events in which jets are misidentified as both hadronic taus and light leptons. The add-on technique described in Section 6.2.1 accounts for electroweak processes in the same-sign data by subtracting them out.

The use of the add-on technique is predicated upon the scale factor r_{QCD}. Rather than being a mere mathematical convenience, r_{QCD} accounts for the differences in multijet parton composition in SS versus OS events. The leading order multijet final states are $q\bar{q}$, qq', $q\bar{q}'$, qg, and gg. Illustrative diagrams are shown in Figure 6.7. OS multijet events are largely comprised of final states having an expected high degree of charge correlation: $q\bar{q}$, qq', and $q\bar{q}'$. On the other hand, SS multijet events consist largely of those for which no charge correlation is expected: qg and gg. The probability that a multijet event will be misidentified as both a hadronic tau and lepton depends on the composition of the final state partons,

leading to potential differences in the numbers and kinematics of OS and SS multijet events passing the event selection.[10] The r_{QCD} scale factor accounts for these differences.

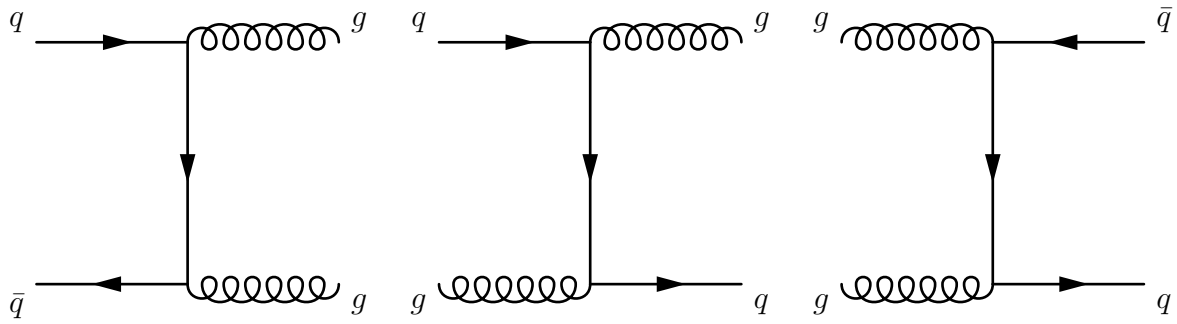


Figure 6.7: Examples of leading order diagrams for dijet production.

A dedicated measurement of r_{QCD} was performed in the ATLAS SM $H \to \tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ analysis [102]. The obtained values of r_{QCD} for both electron and muon measurements are used in this analysis since the same final states are present in both analyses.

r_{QCD} was measured in a dedicated control region with $E_{\mathrm{T}}^{\mathrm{miss}} < 15$ GeV and $m_{\mathrm{T}}(\mu/e, E_{\mathrm{T}}^{\mathrm{miss}}) < 30$ GeV.[11] Track and calorimeter isolation requirements of $I(p_{\mathrm{T}}, 0.4)/p_{\mathrm{T}}(\mu/e) < 0.06$ and $I(E_{\mathrm{T}}, 0.2)/p_{\mathrm{T}}(\mu/e) > 0.06$ are imposed. It should be noted that the E_{T} requirement actually reverses the isolation cuts. The reconstructed hadronic taus are required to pass *loose* tau identification requirements. Expected electroweak and top contributions are then subtracted from the observed data in both OS and SS events. The resulting ratio of OS to SS events yields r_{QCD}. This value was recalculated fixing the track isolation requirement and using a range of calorimeter isolation requirements ($0.06 < I(E_{\mathrm{T}}, 0.2)/p_{\mathrm{T}}(\mu) < 0.38$ and $0.06 < I(E_{\mathrm{T}}, 0.2)/p_{\mathrm{T}}(e) < 0.25$ for muon and electron events, respectively) to determine any dependence between r_{QCD} and calorimeter isolation. A linear fit was performed to extract the value of r_{QCD}. The fitted values and their statistical + systematic uncertainties are shown in Table 6.1.

To validate the values of r_{QCD} obtained from the SM $H \to \tau\tau$ analysis, *multijet control regions* (QCDCR and QCDCR2) measure r_{QCD} using the event selection requirements of this analysis. Details of the procedure are described in Section 8.7, but the values of r_{QCD} used from the $H \to \tau\tau$ analysis are found to be consistent with, and more precise than, the computed values in QCDCR and QCDCR2.

[10] Real leptons can also be produced in multijet events from heavy quark flavor decay, but these leptons are usually neither prompt nor isolated and consequently would not pass event selection.

[11] $m_{\mathrm{T}}(\mu/e, E_{\mathrm{T}}^{\mathrm{miss}}) = \sqrt{2p_{\mathrm{T}}(\mu/e)E_{\mathrm{T}}^{\mathrm{miss}}(1 - \cos\Delta\phi(\mu/e, E_{\mathrm{T}}^{\mathrm{miss}}))}$ is the *transverse mass*, or the mass of the transverse projection of the sum of the muon/electron and hadronic tau four-vectors.

Table 6.1: Summary of all scale factors used for the background estimation. The normalization factors are listed with no uncertainties because they are used as free parameters in the binned likelihood fit used to perform the measurement. The uncertainty on the r_{QCD} factor is statistical + systematic. All other uncertainties are statistical only.

Scale Factors	Values	
	Muon	Electron
r_{QCD}	1.10 ± 0.14	1.00 ± 0.13
Embedding normalization, $k_{Z\to\tau\tau}$	0.451	0.637
$k_{Z\to ee(\to\tau)}$ normalization	-	1.17
$k^{\mathrm{SR1}}_{W+\mathrm{jets}}$ normalization	0.883	0.891
$k_{Z\to\mu\mu(\to\tau)}$	1	-
$k_{Z\to\mu\mu+jet(\to\tau)}$	0.688 ± 0.032	-
$k_{Z\to ee+jet(\to\tau)}$	-	0.684 ± 0.029
k^{OS}_{W+jets}	0.771 ± 0.064	0.841 ± 0.041
k^{SS}_{W+jets}	0.90 ± 0.15	0.968 ± 0.100
k^{OS}_{top}	0.952 ± 0.039	0.908 ± 0.036
k^{SS}_{top}	1.09 ± 0.16	1.11 ± 0.14
k_{VV}	1	
$k_{H\to\tau\tau}$	1	

Chapter 7

Event Selection

Chapters 4 and 6 described the objects reconstructed from detector measurements and physics processes having those objects in their final states. Every data event at ATLAS contains some combination of reconstructed objects, but not every event is a candidate event for the LFV Higgs boson decay process. Furthermore, some events may have the objects expected from an LFV Higgs boson decay, but the object kinematics suggest that the event is more likely due to a background rather than signal process. Including these events in the analysis degrades sensitivity to the signal. This chapter describes event selection criteria that are applied to select a data sample that is most sensitive to the signal process.

7.1 Data

The data considered are from proton–proton collisions with center-of-mass energy $\sqrt{s} =$ 8 TeV, recorded by ATLAS in 2012 and subjected to various controls to ensure high quality, including ensuring that all parts of the detector were working properly at the time of collection (e.g. no malfunctions in any subsystem, no noise bursts in the calorimeters, and having a high fraction of inner detector silicon modules and TRT straws in good operation).[1] The amount of data that satisfies these data quality requirements and is good for physics analysis is 20.3 fb^{-1}.

7.2 Trigger

LFV events are expected to feature high-p_{T} muons or electrons. Consequently, the muon and electron measurements require a set of single lepton triggers to fire, indicating the presence of these high-p_{T} muons or electrons.

Events in the muon measurement are required to fire either the `EF_mu24i_tight` or `EF_mu36_tight` triggers,[2] respectively fired for muons with $p_{\mathrm{T}} > 24$ GeV and 36 GeV

[1] All data events good for physics are listed in a file called a *good runs list*.

[2] These triggers are *Event Filter* triggers. See Section 3.2.6.

and satisfying *tight* isolation requirements in the L2 trigger ($I(E_{\mathrm{T}}, 0.2) < 1.4$ GeV and $I(p_{\mathrm{T}}, 0.2) < 5.7$ GeV) [112]. An additional isolation requirement of $I(p_{\mathrm{T}}, 0.2)/p_{\mathrm{T}}(\mu) < 0.12$ is applied to muons firing the `EF_mu24i_tight` trigger.

In the electron measurement, events are required to fire the `EF_e24vhi_medium1` trigger, identifying electrons having $p_{\mathrm{T}} > 24$ GeV, satisfying $|\eta|$-dependent calorimeter energy thresholds, having *hadronic leakage*[3] energy less than 1 GeV, satisfying the isolation requirement $I(p_{\mathrm{T}}, 0.2)/E_{\mathrm{T}}(e) < 0.1$, and satisfying *medium++* classification [113] requirements.[4]

7.3 Object Definitions

Chapter 4 described the general approach for object reconstruction. However, there are many configurations for which a physics object could be reconstructed. This section describes the specific configurations chosen to reconstruct muons, electrons, jets, and hadronic taus.

7.3.1 Muons and Electrons

Reconstructed muons considered are required to be either segment tag or combined muons reconstructed using the STACO algorithm.[5] Such muons must have $p_{\mathrm{T}} > 10$ GeV, $|\eta| < 2.5$, and satisfy the following inner detector track requirements: at least one PIX hit, at least five SCT hits, fewer than three *expected hits*[6] in both PIX and SCT, and at least nine TRT hits for muons with $0.1 < |\eta| < 1.9$ [78].

Reconstructed electrons are required to satisfy *medium++* classification criteria, have $p_{\mathrm{T}} > 15$ GeV, and have $|\eta| < 2.47$. Electrons in the barrel–end-cap transition region ($1.37 < |\eta| < 1.52$) are not considered.

7.3.2 Jets

Reconstructed jets are required to have $p_{\mathrm{T}} > 30$ GeV and $|\eta| < 4.5$. Jets with $p_{\mathrm{T}} < 50$ GeV and $|\eta| < 2.4$ must satisfy $|\mathrm{JVF}| > 0.5$ (see Equation 4.3) to suppress pile-up contributions to the event.

[3] Electrons deposit much of their energy into the EM calorimeter, but some energy can "spill over" into the hadronic calorimeter. This spilled energy is called *hadronic leakage*.

[4] *medium++* classification is a re-optimized version of the *medium* classification mentioned in Section 4.3. It offers more selection variables and cuts to improve the classification performance in a high pile-up environment.

[5] STACO is nominally used for reconstructing combined muons, but by adopting loose requirements with the track fitting performed by the algorithm, segment tag muons can be reconstructed with it as well.

[6] An expected hit occurs when a track is reconstructed through a silicon module that did not register a hit.

7.3.3 Hadronic Taus

Identified hadronic taus considered satisfy *medium* tau identification and have $p_{\mathrm{T}} > 20$ GeV, $|\eta| < 2.47$, one or three tracks, and an electric charge of ± 1. The hadronic taus must satisfy muon and electron vetoes.

7.4 Object Overlap Removal

It is possible for reconstructed objects to occupy the same spatial location in the detector. Such objects are said to overlap, and only one of them is used for further analysis. Objects are considered overlapping if they are both located in a cone with radius $\Delta R < 0.2$. The following order of preference is applied to overlapping objects: muons, electrons, hadronic taus, and jets.[7]

7.5 Preselection

To reject misidentified objects and events incompatible with the topology of an LFV Higgs boson decay, additional requirements called *preselection requirements* are applied to the reconstructed objects in the data event. These requirements are as follows:

- Events are required to have at least one primary vertex with four or more associated tracks to reject events that are not the result of proton–proton collisions (for example, events due to cosmic rays or secondary particles produced from proton collisions with residual gas in the beam pipe).
- Additional requirements are applied to jets to reject those not arising from proton–proton collisions [114].
- The overlap removal is performed using reconstructed electrons meeting *loose++* classification requirements.[8] In the barrel–end-cap transition region ($1.37 < |\eta| < 1.52$), the electrons are required to satisfy *medium++* classification.
- Only one reconstructed electron or muon with $p_{\mathrm{T}} > 26$ GeV and corresponding to a trigger object may be present in the event.
- The muon or electron is required to satisfy track and calorimeter isolation criteria, respectively $I(p_{\mathrm{T}}, 0.4)/p_{\mathrm{T}}(\mu/e) < 0.06$ and $I(E_{\mathrm{T}}, 0.2)/p_{\mathrm{T}}(\mu/e) < 0.06$.
- Only one hadronic tau may be present in the event.

[7] For example, if a muon and hadronic tau are found to overlap, the muon would be kept while the hadronic tau would be discarded.

[8] *Loose++* is a re-optimized version of the *loose* classification scheme.

- The hadronic tau and muon/electron must have opposite charge.
- The reconstructed MMC mass must be larger than 50 GeV due to the 40 GeV dimuon mass requirement used in event selection for the $Z \to \tau\tau$ embedding samples.

Chapter 8

Signal and Control Regions

The preselection criteria outlined in Chapter 7 select a sample of data events consistent with the LFV Higgs boson decay processes. However, not all of these events are equally likely to originate from the signal process. A second round of criteria are applied categorizing events into different regions of interest called *signal* and *control regions*. Those events in the signal regions are deemed to be most consistent with the signal processes, while events in the control regions determine the k-factors and normalization factors associated with the modeled background processes as well as constrain systematic uncertainties in the binned likelihood fit used to measure the LFV Higgs boson branching ratio. This chapter describes the different signal and control regions considered in the analysis.

8.1 Event Categorization

The following features of LFV Higgs boson decays identify kinematic regions where signal may be better distinguished from background processes:

- A high-p_{T} electron or muon. The electron and muon p_{T} spectra peak at roughly half the Higgs boson mass and are shown in Figure 8.1.
- A single neutrino from the tau lepton decay. Since the tau lepton from the Higgs boson decay is heavily boosted due to the Higgs boson–tau lepton mass difference, the decay products of the tau lepton are tightly collimated with relatively high-p_{T}. Thus, the neutrino is expected to be roughly collinear with the hadronic tau, whose p_{T} spectrum is shown in Figure 8.1.
- Because the neutrino and hadronic tau are almost collinear, the transverse mass of the missing transverse energy and hadronic tau ($m_{\mathrm{T}}(\tau, E_{\mathrm{T}}^{\mathrm{miss}})$) will be small.
- Conversely, the large angle separation between the muon/electron and the neutrino will result in a relatively large transverse mass between the muon/electron and missing transverse energy, $m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{\mathrm{miss}})$.

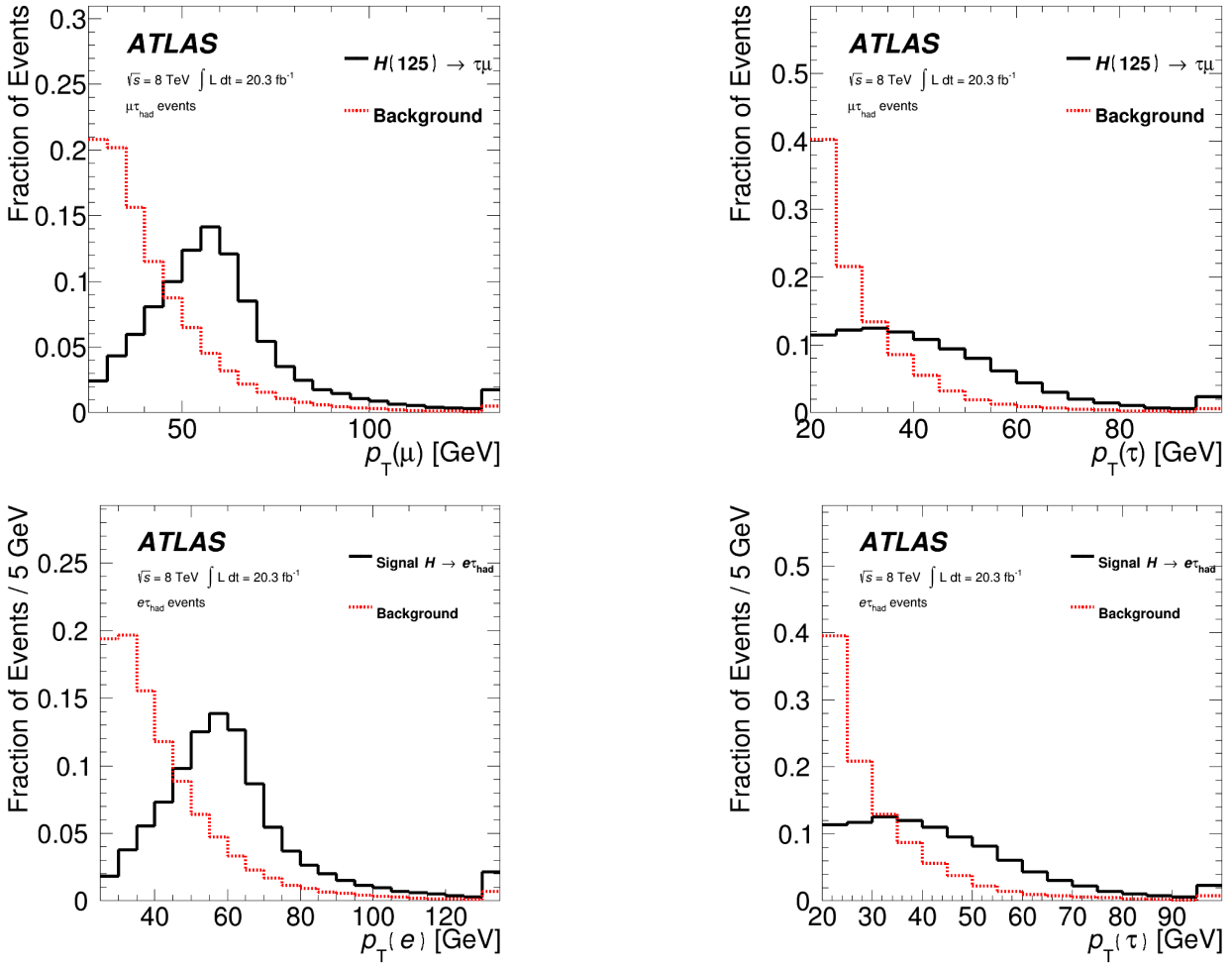


Figure 8.1: Comparisons of the relative p_{T} spectra of muons, electrons, and hadronic taus for signal LFV and background processes for the muon (top row) and electron (bottom row). These spectra are obtained after the preselection requirements outlined in Section 7.5 have been applied. All distributions are computed using MC-simulated data.

Correlations between $m_{\mathrm{T}}(\tau, E_{\mathrm{T}}^{\mathrm{miss}})$ and $m_{\mathrm{T}}(\mu/e, E_{\mathrm{T}}^{\mathrm{miss}})$ for the signal, W + jets, and $Z \to \tau\tau$ backgrounds define *Signal Region 1* (SR1), *Signal Region 2* (SR2, which also serves as a control region for the $Z \to \tau\tau$ background), and the *W + jets Control Region* (WCR). The requirements on $m_{\mathrm{T}}(\tau, E_{\mathrm{T}}^{\mathrm{miss}})$ and $m_{\mathrm{T}}(\mu/e, E_{\mathrm{T}}^{\mathrm{miss}})$ are indicated in Table 8.1. These correlations are shown for both muon and electron measurements in Figures 8.2 and 8.3, with SR1, SR2, and WCR outlined in the correlations plot for observed data.

An optimization study is performed to determine the kinematic cuts that maximize signal sensitivity. Two sets of $m_{\mu\tau}^{\mathrm{MMC}}$ mass histograms are constructed for the signal and total predicted background: one histogram is populated with events satisfying tested $p_{\mathrm{T}}(\tau)$ and

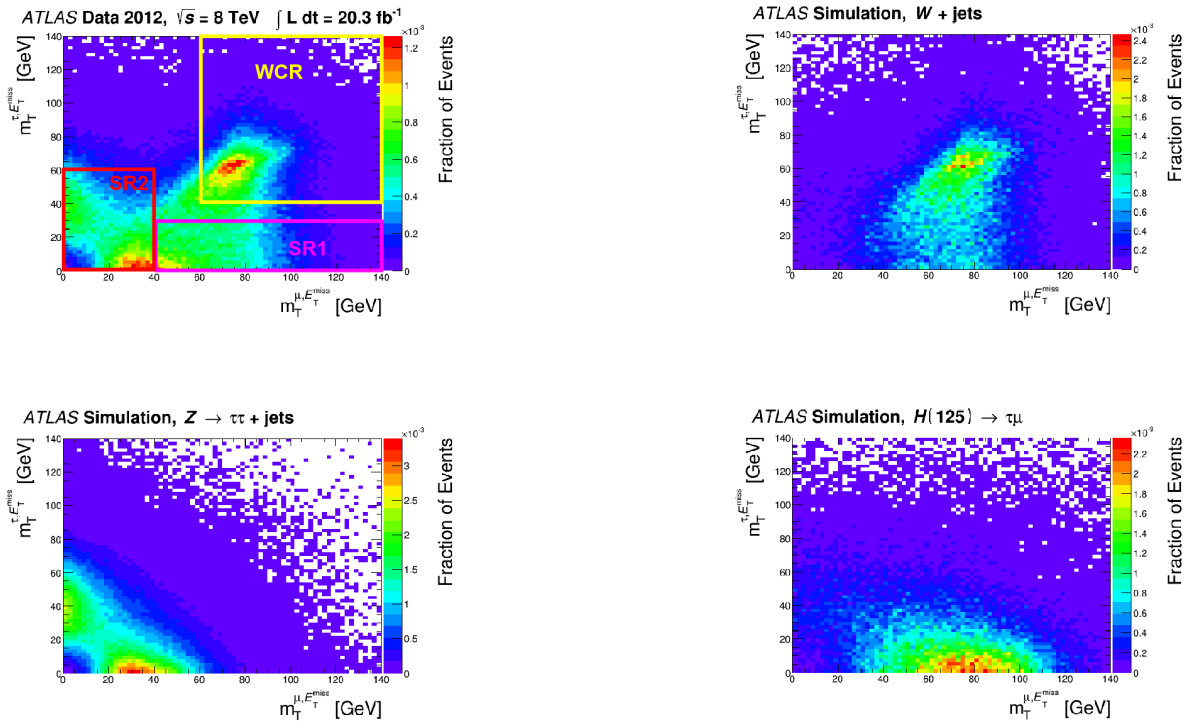


Figure 8.2: Two-dimensional transverse mass distributions for opposite-sign data, W + jets, $Z \to \tau\tau$, and signal events in the muon measurement [115]. All events pass the preselection criteria outlined in Section 7.5. The boxed areas in the data plot outline the W + jets Control Region and signal regions.

$p_{\mathrm{T}}(\mu)$ requirements, while the other is populated with the events not passing those requirements. The *significance* (Ξ) is calculated as shown in Equation 8.1, using histogram bins i in the signal mass window $100 < m_{\mu\tau}^{\mathrm{MMC}} < 150$ GeV:

$$\Xi = \sqrt{2 \sum_{i} (s_i + b_i) \ln(1 + \frac{s_i}{b_i} - s_i)} \tag{8.1}$$

where s_i and b_i refer to the signal and total background events in Bin i. The total significance is computed as $\Xi_{\mathrm{total}} = \sqrt{\Xi_{\mathrm{pass}}^2 + \Xi_{\mathrm{fail}}^2}$, where Ξ_{pass} and Ξ_{fail} respectively refer to events satisfying or failing the tested kinematic requirements. The kinematic criteria maximizing Ξ_{total} are applied to the objects used in the analysis. The study is performed for the muon measurement and finds that $p_{\mathrm{T}}(\mu) > 26$ GeV and $p_{\mathrm{T}}(\tau) > 45$ GeV provide the best sensitivity for distinguishing signal from background. The large obtained value of $p_{\mathrm{T}}(\tau)$ is due to W + jets being the dominant background in the signal mass window. The W + jets background is characterized by jets misidentified as hadronic taus that tend to have a softer $p_{\mathrm{T}}(\tau)$ spectrum

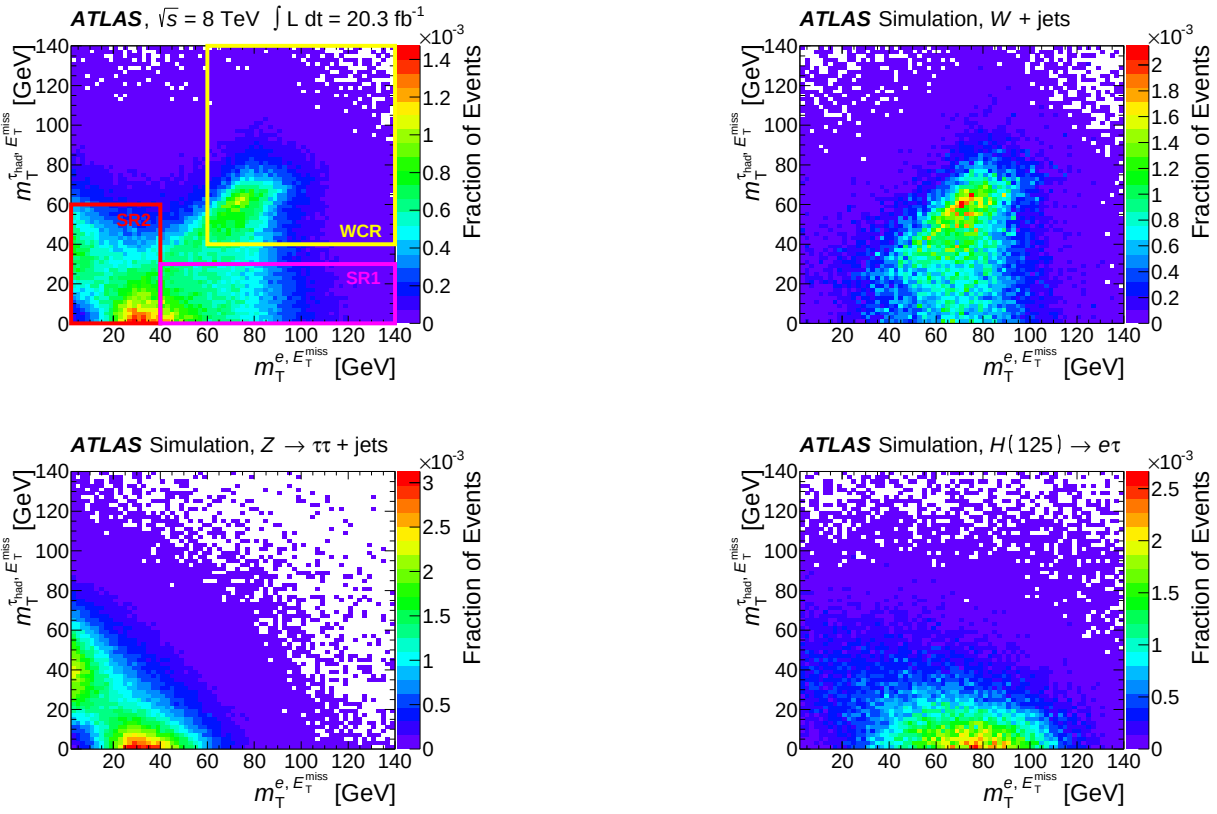


Figure 8.3: Two-dimensional transverse mass distributions for opposite-sign data, W + jets, $Z \to \tau\tau$, and signal events in the electron measurement [34]. All events pass the preselection criteria outlined in Section 7.5. The boxed areas in the data plot outline the W + jets Control Region and signal regions.

than the signal process. The optimized p_{T} requirements are used with the corresponding objects in the electron measurement due to the similarity of the signal and background processes (i.e. $p_{\mathrm{T}}(\tau) > 45$ GeV and $p_{\mathrm{T}}(e) > 26$ GeV).

Two additional kinematic requirements are applied to reduce backgrounds in the signal and control regions. To suppress the top background, no events with b-tagged jets are included in the signal or control region definitions. However, a *Top Control Region* (TCR) is defined as those events that have at least two jets, including at least one b-tagged jet. The other requirement is that the hadronic tau and muon/electron in each event satisfy $|\Delta\eta(\tau, \mu/e)| < 2.0$ and is applied to suppress background from dijet and W + jets processes. The final state objects in these processes tend to be back-to-back, occupying different hemispheres of the detector.

Two *QCD Control Regions* are defined with $|\Delta\eta(\tau, \mu/e)| > 2.0$ to verify the value of r_{QCD} used in each measurement. The first region (QCDCR) consists of all events passing

the selection for SR1 and SR2 (aside from the $|\Delta\eta(\tau, \mu/e)|$ requirement), while the second (QCDCR2) consists of events satisfying SR2 criteria with the exception that 20 GeV $< p_{\mathrm{T}}(\tau) <$ 45 GeV.

A control region (ZmmCR/ZeeCR) is used for measuring the k-factors for the $Z \to \mu\mu/ee$ processes where a jet is misidentified as a hadronic tau. Events categorized into the ZmmCR/ZeeCR are required to have two muons/electrons with the invariant mass of the muon/electron pair between 80 GeV and 100 GeV. A reconstructed (though misidentified) hadronic tau is also required to be present in the event.

The kinematic requirements of all signal and control regions are listed in Table 8.1.

Criterion	SR1	SR2	WCR	TCR
$p_{\mathrm{T}}(\ell)$ [GeV]	> 26	> 26	> 26	> 26
$p_{\mathrm{T}}(\tau_{\mathrm{had}})$ [GeV]	> 45	> 45	> 45	> 45
$\lvert\Delta\eta(\ell, \tau)\rvert$	< 2	< 2	< 2	< 2
$m_{\mathrm{T}}^{\ell, E_{\mathrm{T}}^{\mathrm{miss}}}$ [GeV]	> 40	< 40	> 60	–
$m_{\mathrm{T}}^{\tau_{\mathrm{had}}, E_{\mathrm{T}}^{\mathrm{miss}}}$ [GeV]	< 30	< 60	> 40	–
N_{jet}	–	–	–	≥ 2
$N_{b-\mathrm{jet}}$	0	0	0	≥ 1
$\ell\ell$ veto?	Yes	Yes	Yes	Yes

Criterion	Z$\ell\ell$CR	QCDCR	QCDCR2
$p_{\mathrm{T}}(\ell)$ [GeV]	> 26	> 26	> 26
$p_{\mathrm{T}}(\tau_{\mathrm{had}})$ [GeV]	> 45	> 45	< 45
$\lvert\Delta\eta(\ell, \tau)\rvert$	< 2	> 2	> 2
$m_{\mathrm{T}}^{\ell, E_{\mathrm{T}}^{\mathrm{miss}}}$ [GeV]	–	*	< 40
$m_{\mathrm{T}}^{\tau_{\mathrm{had}}, E_{\mathrm{T}}^{\mathrm{miss}}}$ [GeV]	–	*	< 60
N_{jet}	–	–	–
$N_{b-\mathrm{jet}}$	0	0	0
$\ell\ell$ veto?	No	Yes	Yes

Table 8.1: Event selection criteria defining the signal and control regions in both muon and electron measurements. The symbol ℓ indicates a muon or electron. Asterisks (*) indicate that the transverse mass requirements for QCDCR are the union of those for SR1 and SR2.

8.2 Signal Region 1

The criteria for events categorized into SR1 are listed in Table 8.1. Kinematic distributions of the p_{T} spectra of hadronic taus and muons/electrons and missing transverse energy are shown in Figures 8.4 and 8.5, respectively for the muon and electron measurements. Event yields for SR1 are shown in Tables 8.3 and 8.4.

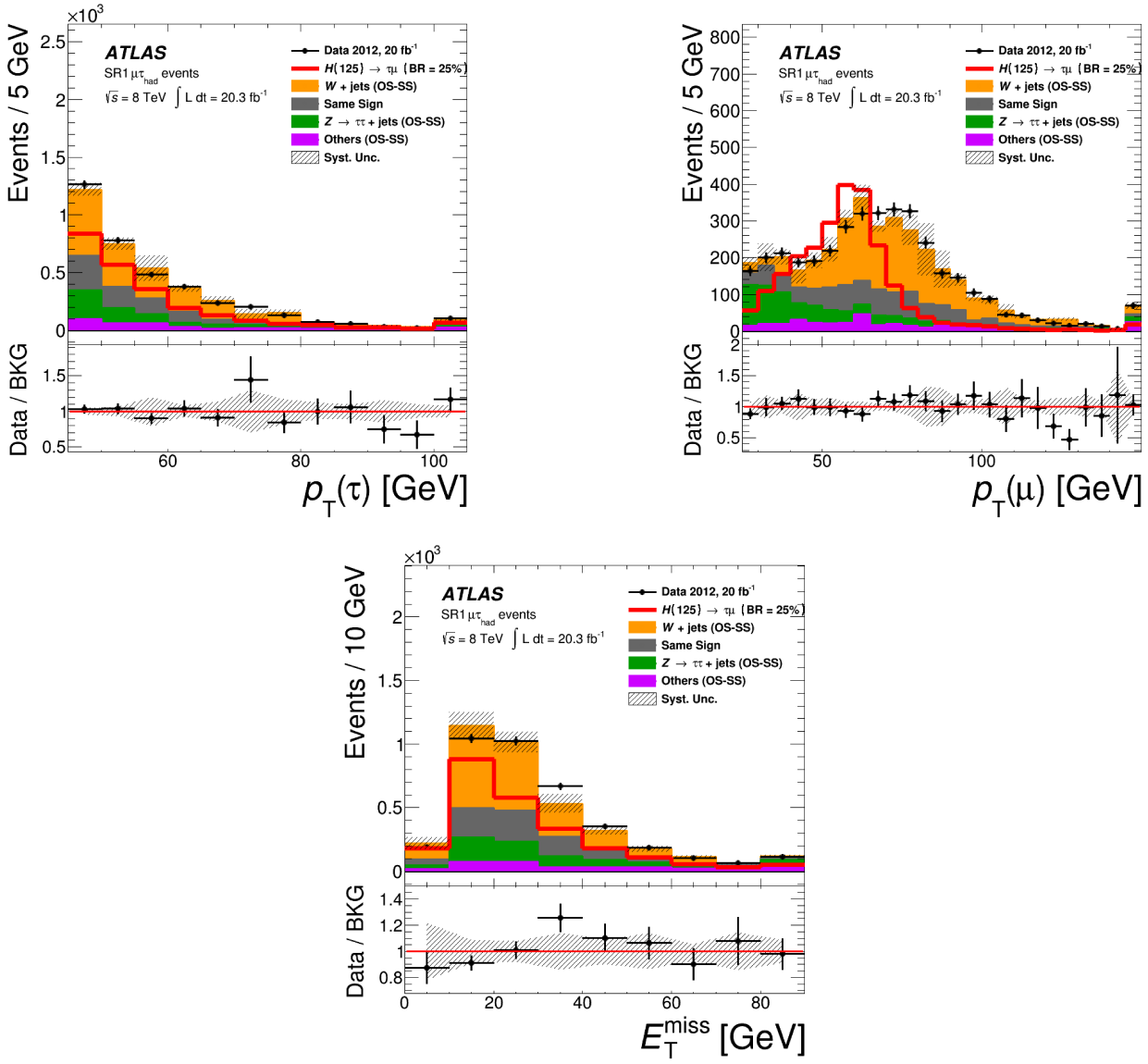


Figure 8.4: $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in SR1 for the muon measurement.

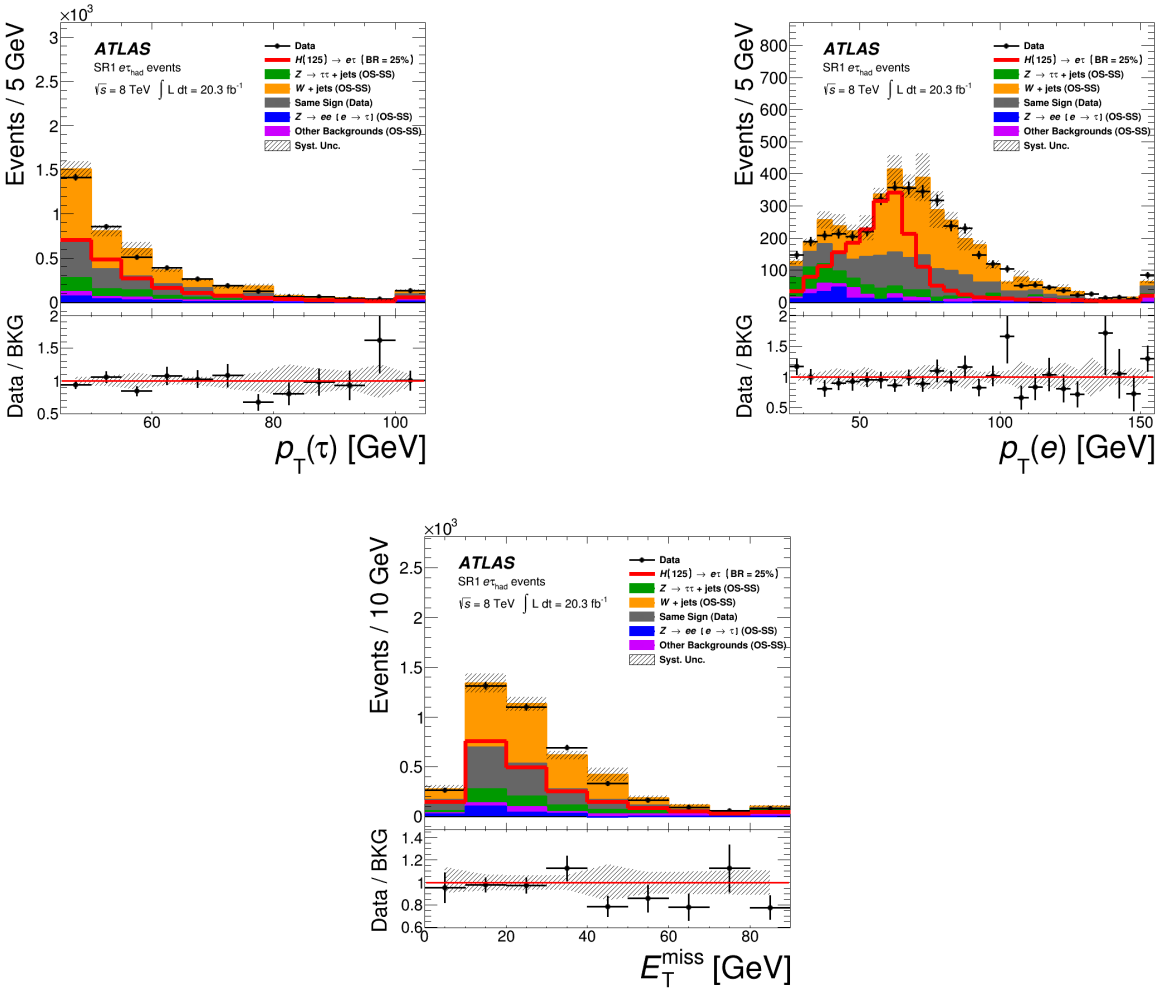


Figure 8.5: $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in SR1 for the electron measurement.

8.3 Signal Region 2

The criteria for events categorized into SR2 are listed in Table 8.1. Kinematic distributions of the p_{T} spectra of hadronic taus and muons/electrons and missing transverse energy are shown in Figures 8.6 and 8.7, respectively for the muon and electron measurements. Event yields for SR2 are shown in Tables 8.3 and 8.4.

8.4 W + jets Control Region

The criteria for events categorized into WCR are listed in Table 8.1. Kinematic distributions of the p_{T} spectra of hadronic taus and muons/electrons and missing transverse energy are

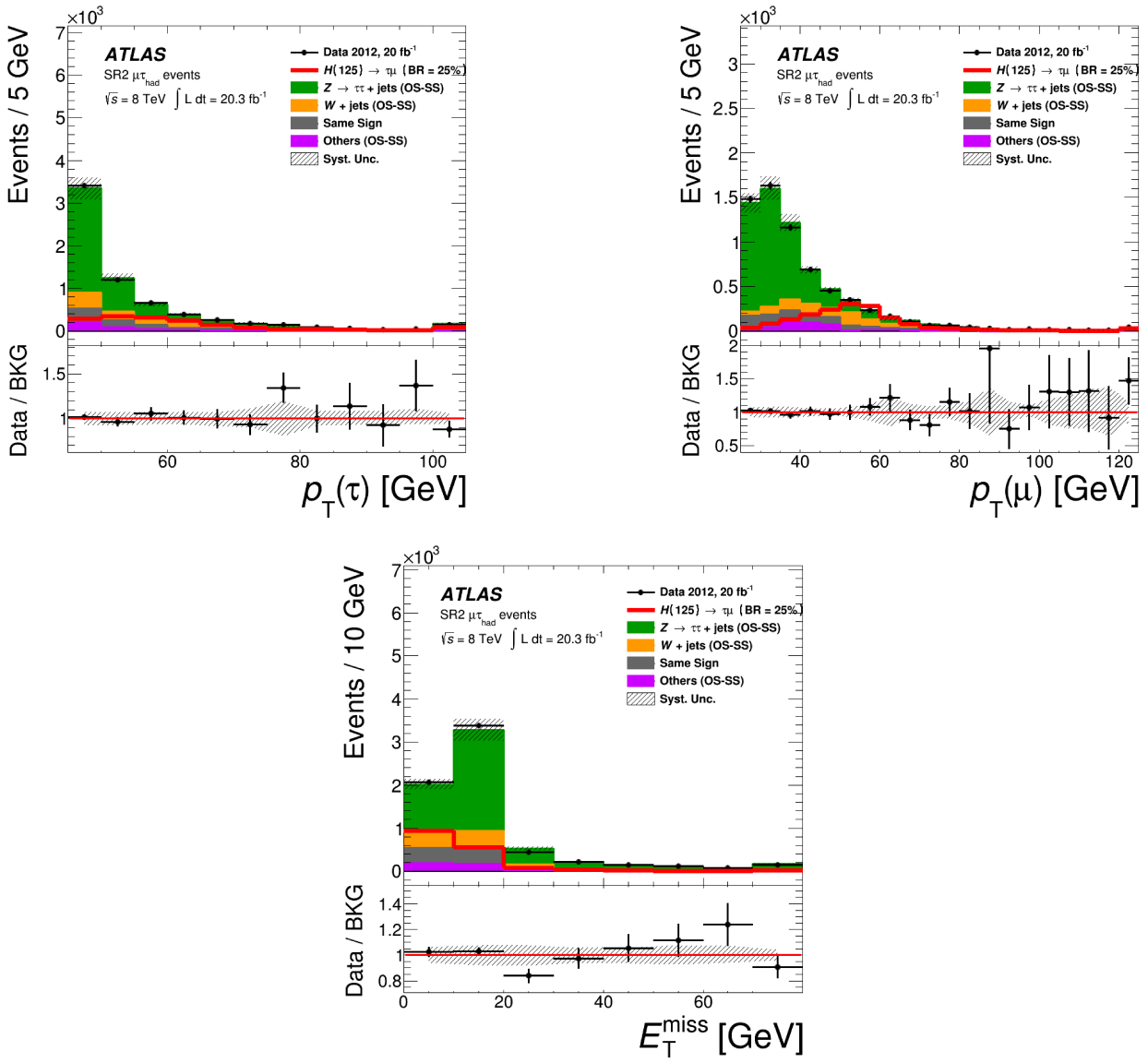


Figure 8.6: $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in SR2 for the muon measurement.

shown in Figures 8.8 and 8.9, respectively for the muon and electron measurements. Event yields for WCR are shown in Tables 8.3 and 8.4.

8.5 Top Control Region

The criteria for events categorized into TCR are listed in Table 8.1. Kinematic distributions of the p_{T} spectra of hadronic taus and muons/electrons and missing transverse energy are

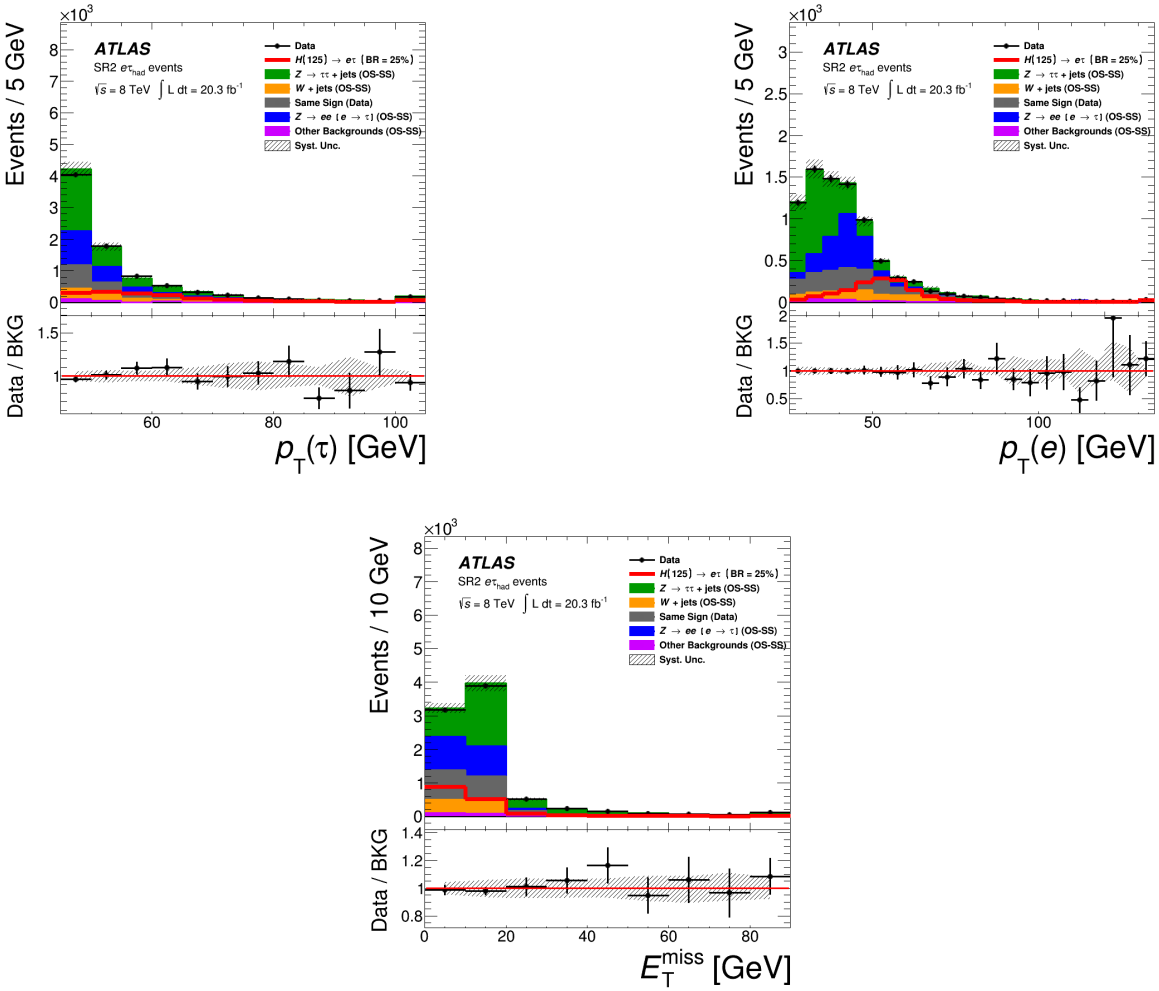


Figure 8.7: $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in SR2 for the electron measurement.

shown in Figures 8.10 and 8.11, respectively for the muon and electron measurements. Event yields for TCR are shown in Tables 8.3 and 8.4.

8.6 $Z \to \mu\mu/ee$ Control Region

The criteria for events categorized into ZmmCR/ZeeCR are listed in Table 8.1. Kinematic distributions of the p_{T} spectra of hadronic taus and muons/electrons and missing transverse energy are shown in Figures 8.12 and 8.13, respectively for the muon and electron measurements.

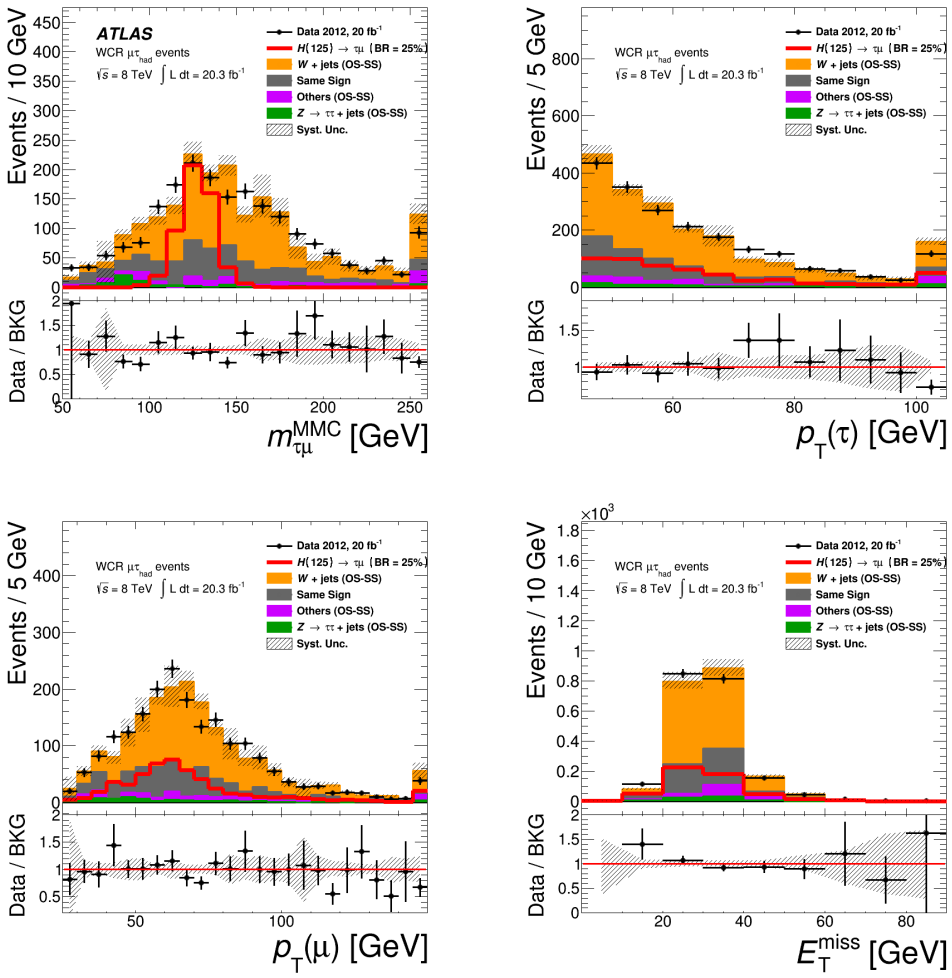


Figure 8.8: $m_{\mu\tau}^{\mathrm{MMC}}$, $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in WCR for the muon measurement.

8.7 Multijet Control Regions

Section 6.2.8 described how the values of r_{QCD} were measured in the SM $H \to \tau\tau$ analysis. To ensure that the measured values of r_{QCD} are still valid for this analysis, they are remeasured in QCDCR and QCDCR2, which are dominated by multijet processes. r_{QCD} is computed as shown in Equation 8.2:

$$r_{\mathrm{QCD}} = \frac{N^{OS}_{\mathrm{QCDCR/2\ Data}} - N^{OS}_{\mathrm{QCDCR/2\ EWK}}}{N^{SS}_{\mathrm{QCDCR/2\ Data}} - N^{SS}_{\mathrm{QCDCR/2\ EWK}}} \tag{8.2}$$

where $N_{\mathrm{QCDCR/2\ EWK}}$ refers to the total number of electroweak (and top quark) events modeled using the samples discussed in Chapter 6. The measured values of r_{QCD} are found to be consistent with the previously measured values, and are shown in Table 8.2:

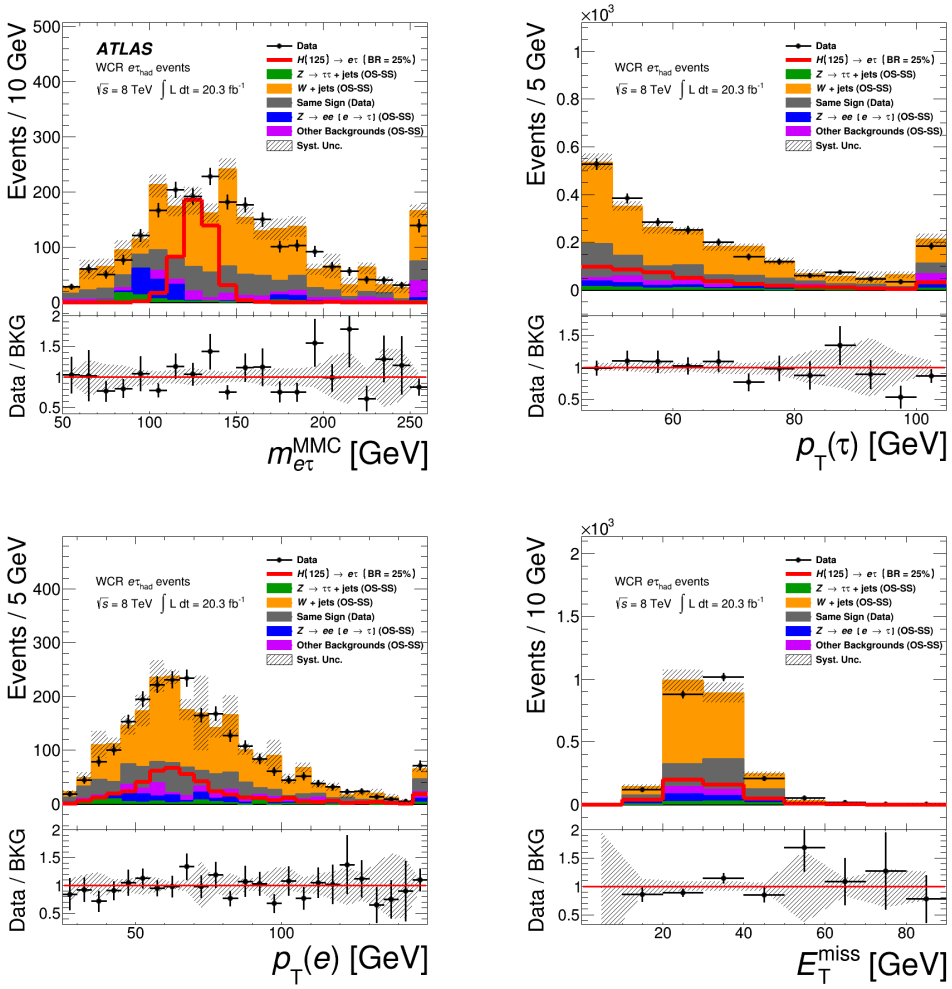


Figure 8.9: $m_{e\tau}^{\mathrm{MMC}}$, $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in WCR for the electron measurement.

8.7.1 Multijet Control Region 1

The criteria for events categorized into QCDCR are listed in Table 8.1. Kinematic distributions of the p_{T} spectra of hadronic taus and muons/electrons and missing transverse energy are shown in Figures 8.14 and 8.15, respectively for the muon and electron measurements.

8.7.2 Multijet Control Region 2

The criteria for events categorized into QCDCR2 are listed in Table 8.1. Kinematic distributions of the p_{T} spectra of hadronic taus and muons/electrons and missing transverse energy are shown in Figures 8.16 and 8.17, respectively for the muon and electron measurements.

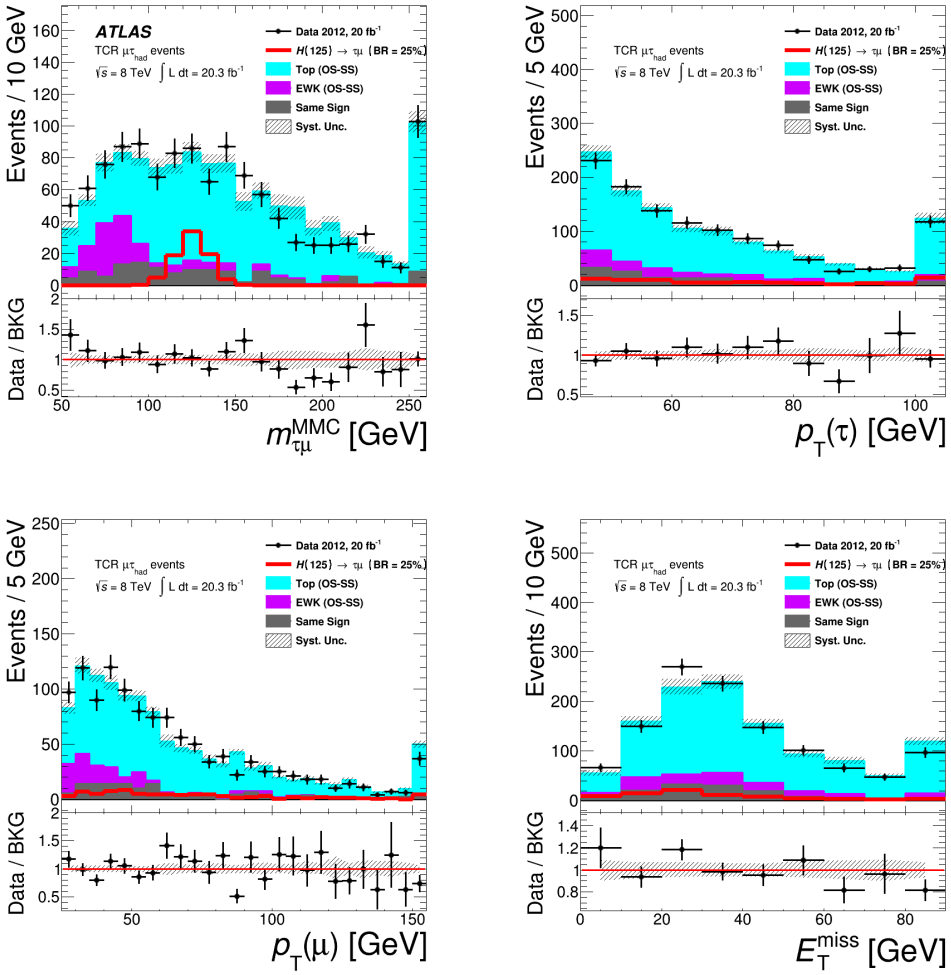


Figure 8.10: $m_{\mu\tau}^{\mathrm{MMC}}$, $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in TCR for the muon measurement.

Measurement	Muon	Electron
QCDCR	1.25 ± 0.39	1.15 ± 0.38
QCDCR2	1.13 ± 0.09	1.05 ± 0.05

Table 8.2: Measured values of r_{QCD} in QCDCR and QCDCR2 for both muon and electron measurements. The listed uncertainties are statistical only. The measured numbers are in excellent agreement with the values of r_{QCD} obtained in the SM $H \to \tau\tau$ analysis: $r_{\mathrm{QCD}} = 1.10 \pm 0.14$ and 1.00 ± 0.13, respectively, for the muon and electron measurements.

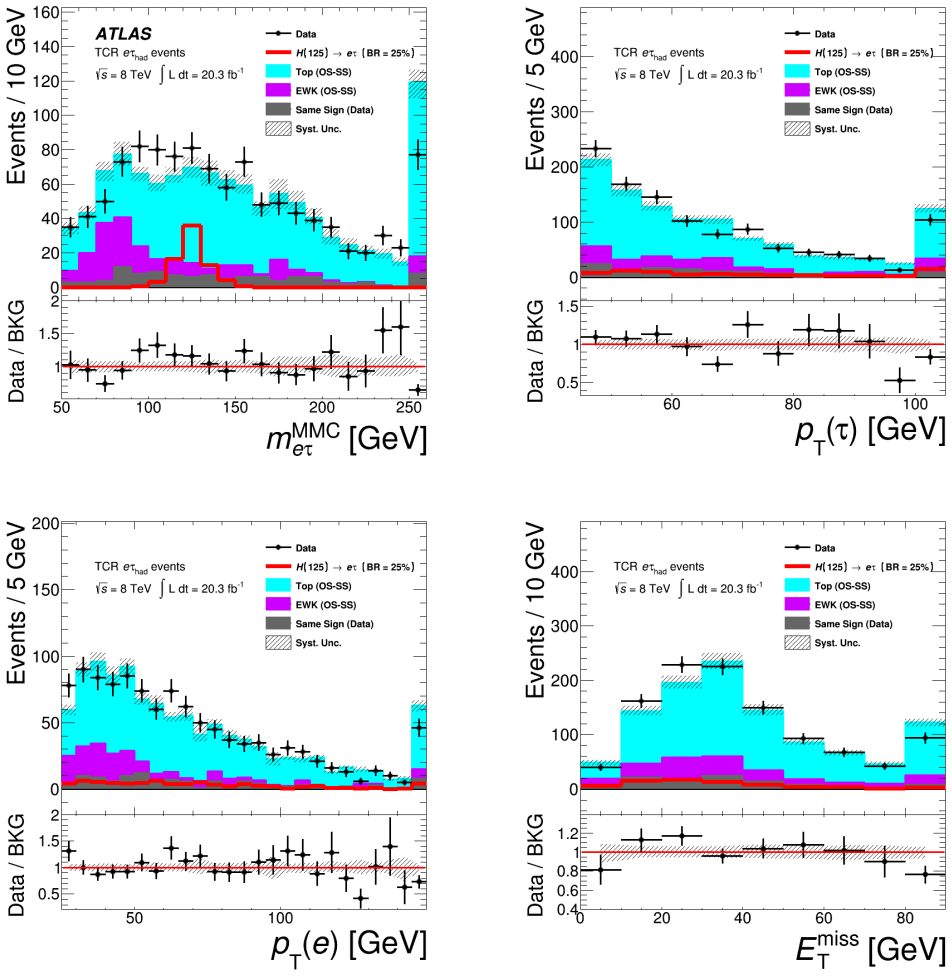


Figure 8.11: $m_{e\tau}^{\mathrm{MMC}}$, $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in TCR for the electron measurement.

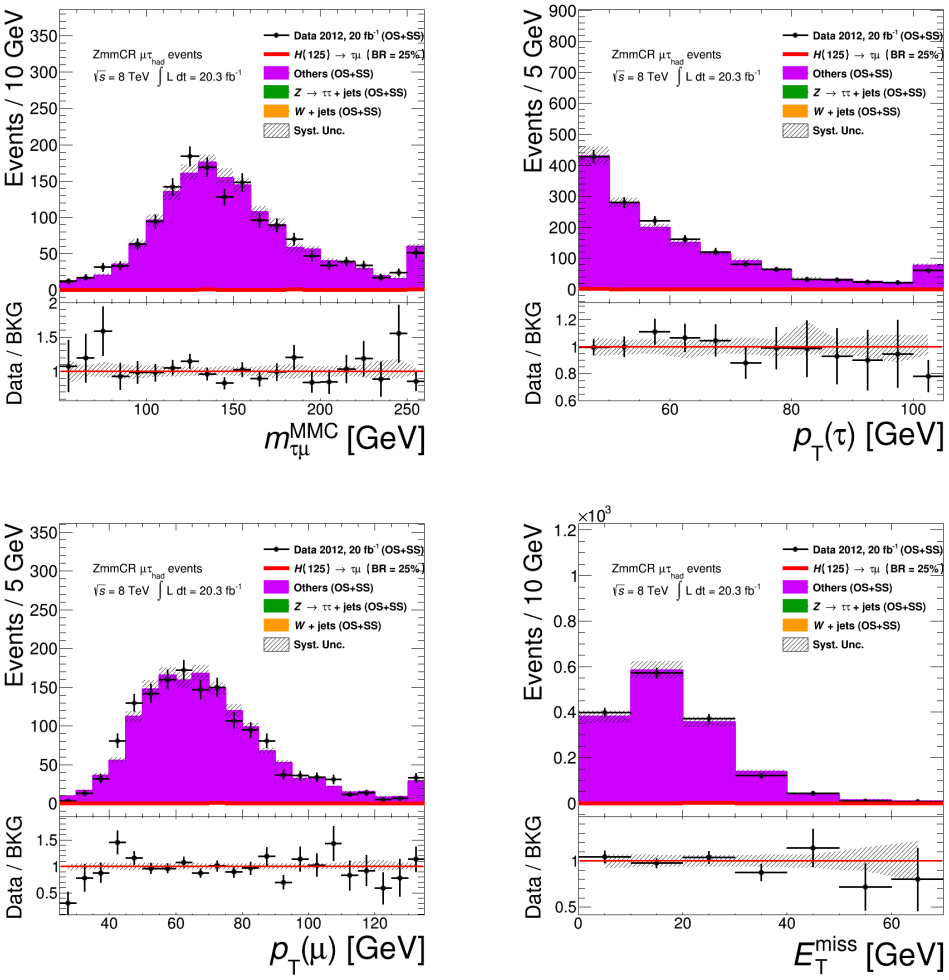


Figure 8.12: $m_{\mu\tau}^{\mathrm{MMC}}$, $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in ZmmCR for the muon measurement.

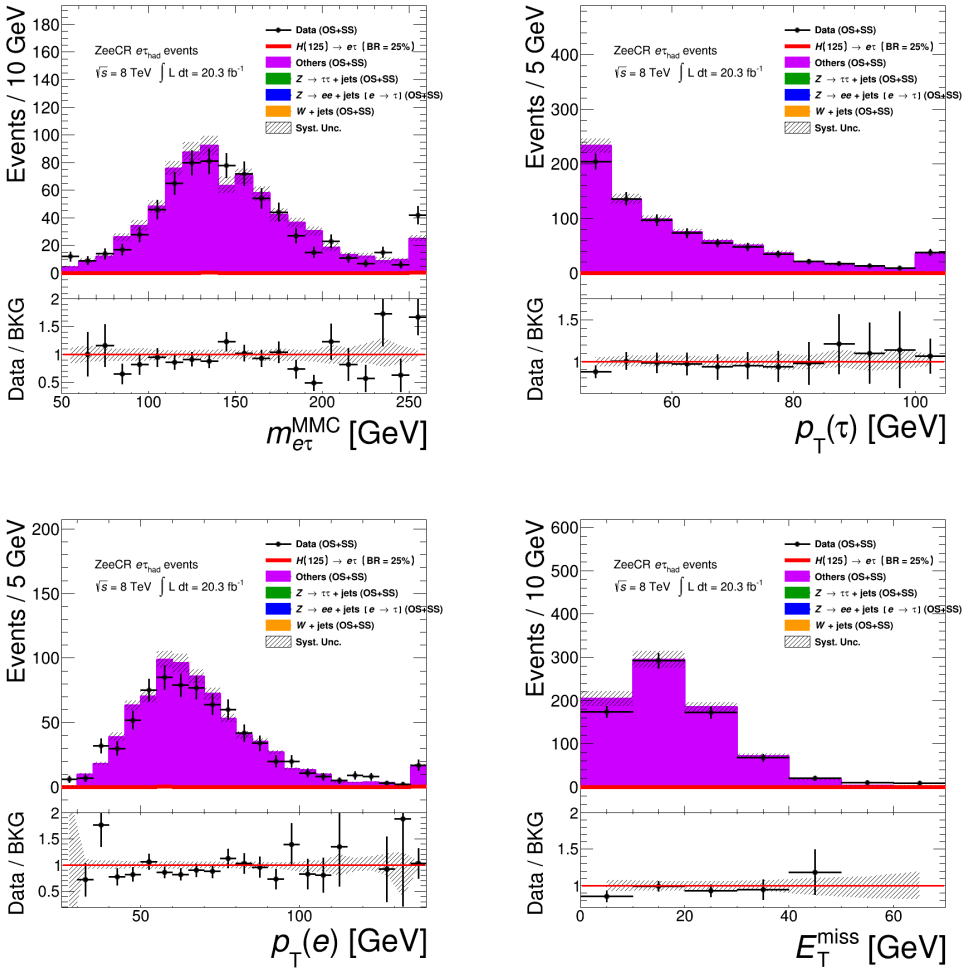


Figure 8.13: $m_{e\tau}^{\mathrm{MMC}}$, $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in ZeeCR for the electron measurement.

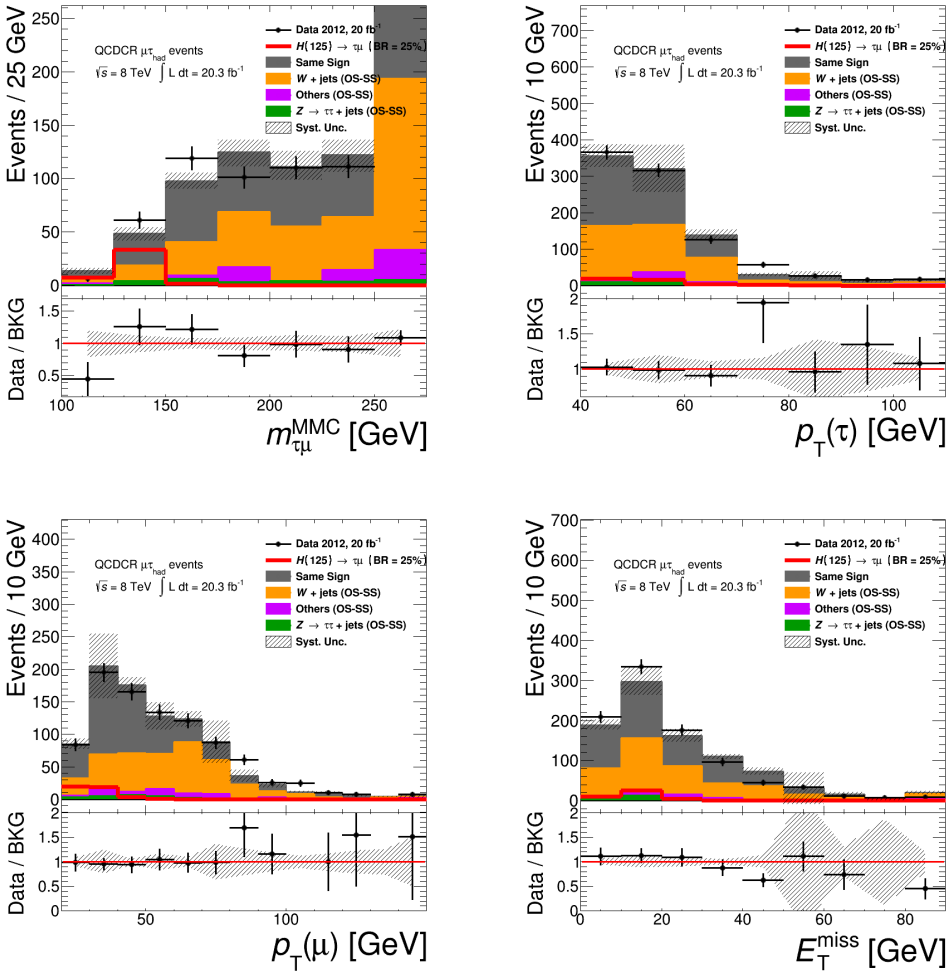


Figure 8.14: $m_{\mu\tau}^{\mathrm{MMC}}$, $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in QCDCR for the muon measurement.

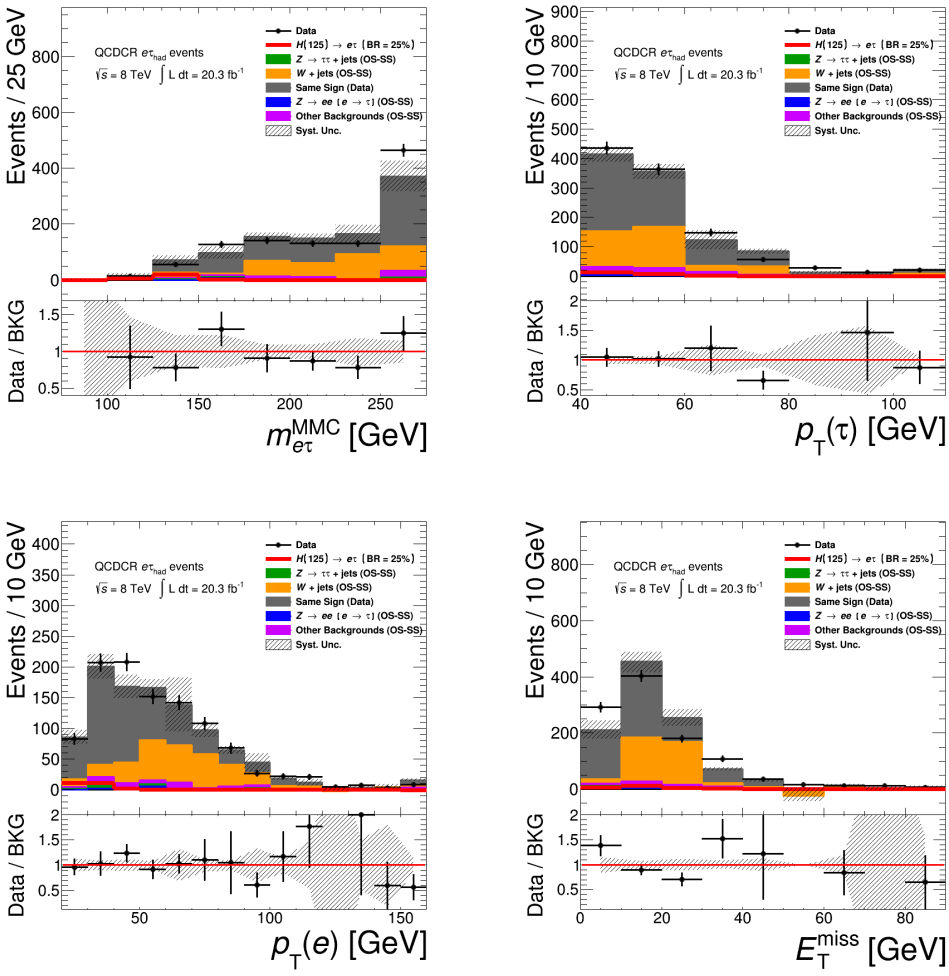


Figure 8.15: $m_{e\tau}^{\mathrm{MMC}}$, $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in QCDCR for the electron measurement.

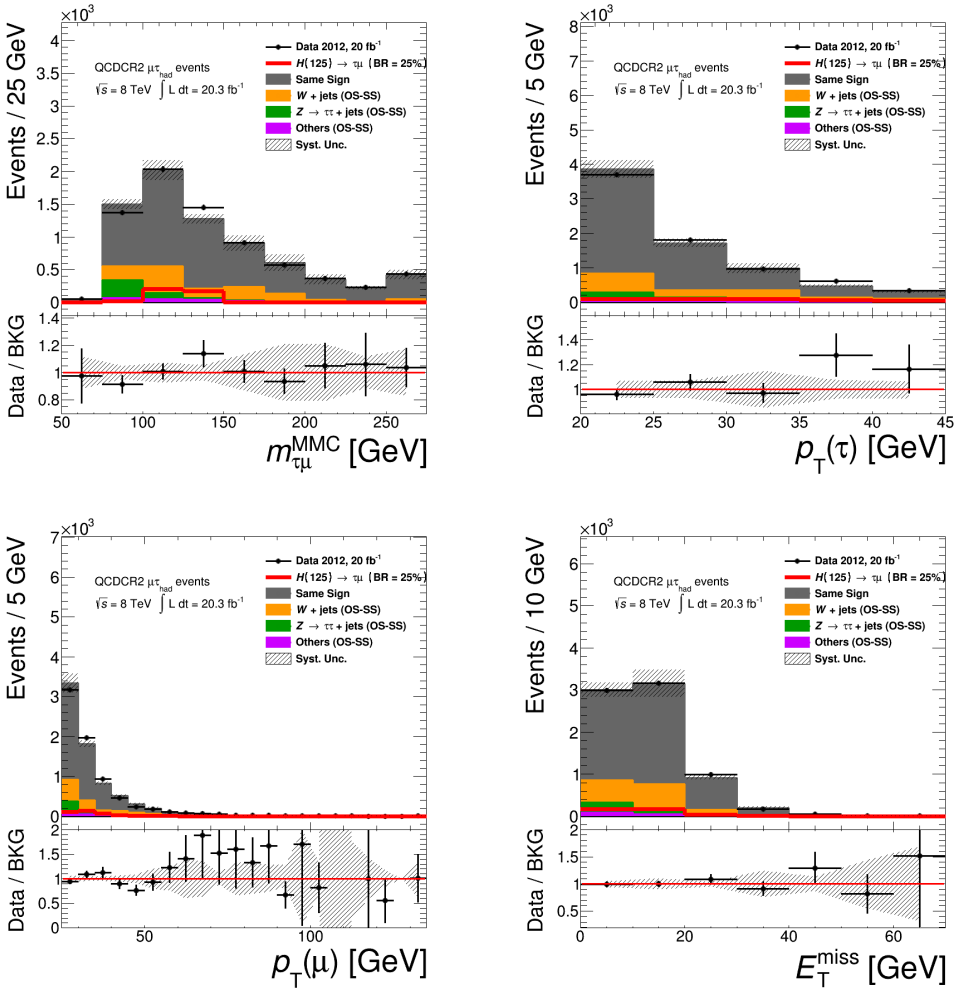


Figure 8.16: $m_{\mu\tau}^{\mathrm{MMC}}$, $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in QCDCR2 for the muon measurement.

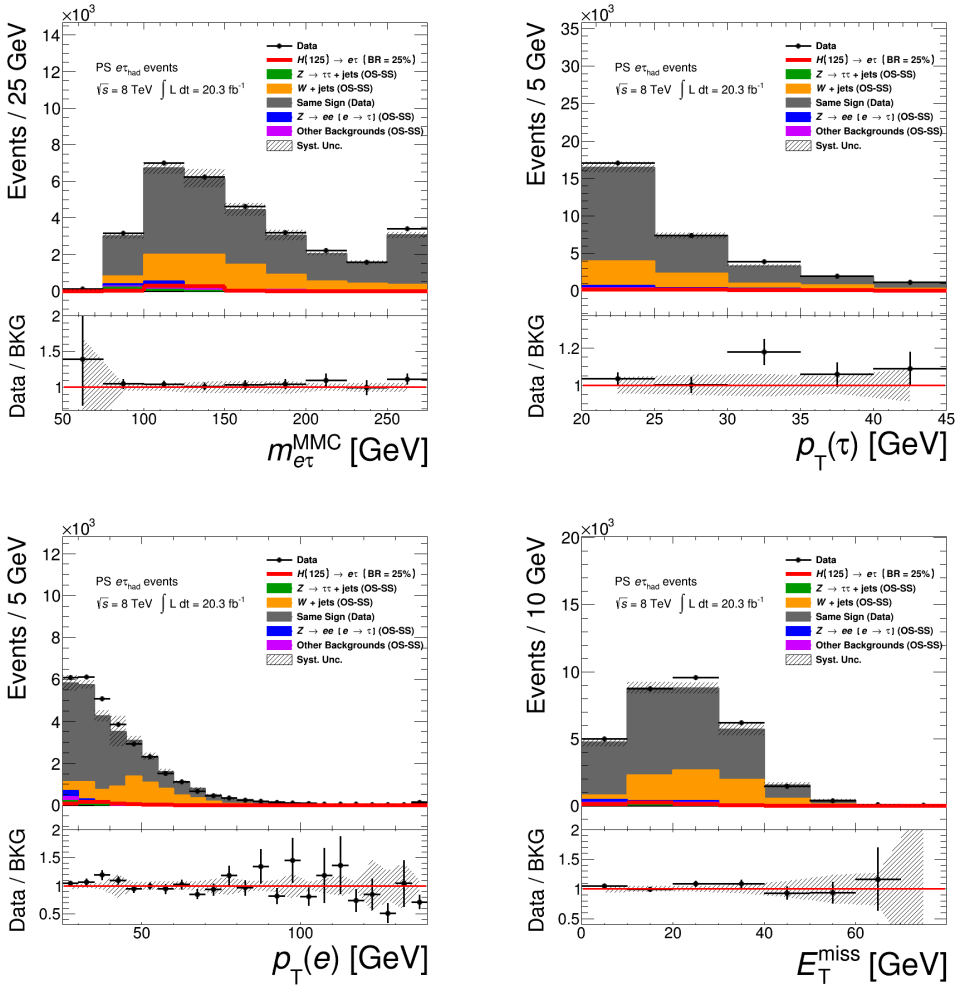


Figure 8.17: $m_{e\tau}^{\mathrm{MMC}}$, $p_{\mathrm{T}}(\tau)$, $p_{\mathrm{T}}(\mu)$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in QCDCR2 for the electron measurement.

Table 8.3: Event yields for signal and background predictions in SR1, SR2, WCR, and TCR after applying the optimized selection cuts. The signal predictions are obtained assuming a branching ratio of $\mathrm{Br}(H \to \mu\tau) = 1.0\%$. Statistical and systematic uncertainties are respectively quoted for each process. The systematic uncertainties on the total background predictions account for correlations between the uncertainties for different background components.

Process	SR1	SR2	WCR	TCR
Signal $H \to \mu\tau$				
ggF	$82.4 \pm 1.2 \pm 10.6$	$58.6 \pm 1.0 \pm 8.2$	$18.3 \pm 0.6 \pm 2.6$	$1.72 \pm 0.18 \pm 0.25$
VBF	$10.29 \pm 0.12 \pm 0.80$	$5.46 \pm 0.09 \pm 0.48$	$1.86 \pm 0.05 \pm 0.35$	$0.728 \pm 0.032 \pm 0.081$
WH	$2.33 \pm 0.06 \pm 0.17$	$1.34 \pm 0.05 \pm 0.11$	$0.538 \pm 0.031 \pm 0.092$	$0.339 \pm 0.023 \pm 0.040$
ZH	$1.24 \pm 0.03 \pm 0.10$	$0.741 \pm 0.026 \pm 0.081$	$0.359 \pm 0.019 \pm 0.048$	$0.464 \pm 0.021 \pm 0.039$
Total Signal	$96.3 \pm 1.2 \pm 10.7$	$66.1 \pm 1.0 \pm 8.2$	$21.1 \pm 0.6 \pm 2.6$	$3.25 \pm 0.18 \pm 0.27$
Background				
$H \to \tau\tau$	$12.7 \pm 0.3 \pm 1.5$	$55.9 \pm 0.7 \pm 7.0$	$1.61 \pm 0.10 \pm 0.23$	$1.55 \pm 0.09 \pm 0.20$
$VV{+}Z \to \mu\mu$	$204 \pm 29 \pm 29$	$307 \pm 39 \pm 66$	$70 \pm 15 \pm 19$	$6.5 \pm 2.2 \pm 2.5$
$Z \to \tau\tau$	$639 \pm 14 \pm 83$	$4412 \pm 33 \pm 782$	$75.1 \pm 5.0 \pm 10.3$	$124.3 \pm 6.6 \pm 10.5$
W + jets	$1852 \pm 93 \pm 231$	$919 \pm 76 \pm 139$	$1255 \pm 82 \pm 176$	$0.6 \pm 8.9 \pm 19.2$
Top	$147.8 \pm 9.2 \pm 14.4$	$52.3 \pm 5.8 \pm 9.5$	$74.4 \pm 7.0 \pm 17.2$	$920 \pm 18 \pm 54$
SS Data	$846 \pm 31 \pm 118$	$787 \pm 29 \pm 110$	$520 \pm 24 \pm 73$	$130 \pm 12 \pm 18$
Total Background	$3688 \pm 103 \pm 241$	$6478 \pm 97 \pm 812$	$1994 \pm 87 \pm 173$	$1181 \pm 25 \pm 59$
Data	3754	6606	1996	1184

Table 8.4: Event yields for signal and background predictions in SR1, SR2, WCR, and TCR after applying the optimized selection cuts. The signal predictions are obtained assuming a branching ratio of $\mathrm{Br}(H \to e\tau) = 1.0\%$. Statistical and systematic uncertainties are respectively quoted for each process. The systematic uncertainties on the total background predictions account for correlations between the uncertainties for different background components.

Process	SR1	SR2	WCR	TCR
Signal $H \to e\tau$				
ggF	$67.9 \pm 1.1 \pm 8.7$	$55.0 \pm 0.9 \pm 7.9$	$16.0 \pm 0.5 \pm 2.5$	$1.58 \pm 0.16 \pm 0.33$
VBF	$8.95 \pm 0.11 \pm 0.67$	$5.344 \pm 0.085 \pm 0.461$	$1.710 \pm 0.049 \pm 0.277$	$0.653 \pm 0.031 \pm 0.076$
WH	$2.022 \pm 0.079 \pm 0.162$	$1.188 \pm 0.060 \pm 0.099$	$0.542 \pm 0.042 \pm 0.102$	$0.310 \pm 0.031 \pm 0.062$
ZH	$1.116 \pm 0.032 \pm 0.097$	$0.708 \pm 0.025 \pm 0.052$	$0.293 \pm 0.017 \pm 0.049$	$0.412 \pm 0.019 \pm 0.034$
Total Signal	$80.0 \pm 1.1 \pm 8.7$	$62.2 \pm 1.0 \pm 7.9$	$18.5 \pm 0.5 \pm 2.5$	$2.96 \pm 0.17 \pm 0.34$
Background				
$H \to \tau\tau$	$10.1 \pm 0.3 \pm 1.2$	$44.5 \pm 0.6 \pm 5.6$	$1.33 \pm 0.09 \pm 0.30$	$1.16 \pm 0.08 \pm 0.13$
$VV{+}Z \to ee(\text{jet} \to \tau)$	$120 \pm 40 \pm 53$	$142 \pm 45 \pm 82$	$85.3 \pm 19.3 \pm 35.9$	$11.7 \pm 3.8 \pm 1.8$
$Z \to ee(\to \tau)$	$185 \pm 25 \pm 46$	$1958 \pm 102 \pm 237$	$144 \pm 20 \pm 20$	$9.33 \pm 2.77 \pm 2.71$
$Z \to \tau\tau$	$505 \pm 17 \pm 82$	$3562 \pm 37 \pm 705$	$67.6 \pm 6.2 \pm 13.5$	$127 \pm 8 \pm 17$
W + jets	$2089 \pm 121 \pm 320$	$894 \pm 78 \pm 207$	$1402 \pm 94 \pm 196$	$28.5 \pm 7.2 \pm 10.5$
Top	$111 \pm 8 \pm 16$	$40.9 \pm 4.8 \pm 7.4$	$82.8 \pm 7.2 \pm 16.8$	$803 \pm 16 \pm 62$
SS Data	$1212 \pm 35 \pm 158$	$1714 \pm 41 \pm 223$	$546 \pm 23 \pm 71$	$109 \pm 10 \pm 14$
Total Background	$4222 \pm 136 \pm 365$	$8311 \pm 147 \pm 734$	$2327 \pm 102 \pm 185$	$1088 \pm 23 \pm 70$
Data	4079	8275	2305	1103

Chapter 9

Systematic Uncertainties

> *[T]he future truths of physical science are to be looked for in the sixth place of decimals.* —A. A. Michelson, speech at the University of Chicago, 1894

To make a perfect measurement, a measuring apparatus would not only have to measure the true value of the quantity, but also make the measurement to infinite precision. Thus, the concept of a perfect measurement is an idealization. Not only are measurements limited by precision, but they may also be biased such that the measured quantity is offset from the true value by an amount exceeding the measurement precision.[1] These imperfections in the measurement are called *systematic uncertainties* and are usually mitigated through calibration and use of high-precision instruments. Despite the ideal nature of a perfect measurement, scientists aspire to high-precision measurements, as they can provide indirect clues about unobserved physical processes.

Systematic uncertainties affect the predicted number of background and signal events as well as the shape of the kinematic distributions. The uncertainties considered fall into one of three categories. *Detector uncertainties* pertain to the reconstructed physics objects described in Chapter 4, arising from the limitations of the ATLAS detector, described in Chapter 3. The second class of uncertainties are *theory uncertainties*, consisting of uncertainties in the predicted cross sections of the signal processes and those background processes not normalized to data. The final category of uncertainties pertains to the analysis methodology. This chapter describes each class of uncertainties.

It is customary to provide a table listing the dominant sources of uncertainty for the signal and background processes in a measurement. However, the application of the signal and background uncertainties in the binned likelihood fit described in Chapter 10 precludes such a table. Systematic variations and their correlations are estimated for each signal and background process, then constrained using data in control regions. The values mentioned in the following sections are estimates before the fit is performed.

[1] An example of a biased measurement apparatus is a scale that consistently under-weighs by 5 lbs.

9.1 Detector Uncertainties

The five physics objects described in Chapter 4 are muons, electrons, jets, hadronic taus, and missing transverse energy, each reconstructed from energy deposits in the detector. The uncertainties in measured energies and the reconstruction process lead to uncertainties in the reconstructed object kinematics. This section describes the systematic uncertainties considered for each object. In addition, the systematic uncertainty on the measured luminosity is also discussed.

9.1.1 Muon Uncertainties

Detailed studies by the ATLAS Muon Combined Performance group determined the systematic uncertainties associated with reconstructed muons [78]. Two different systematic effects for the muons are considered in this analysis. The first effect pertains to the determination of muon momentum resolution and momentum scale. The momentum scale factor corrects for energy lost by the muon as it traverses the detector and is determined using an analytic expression that depends on parameters obtained from fits to the dimuon mass distribution in $Z \to \mu\mu$ events in data and MC-simulation. This fit is performed in the Z boson mass window of 76–96 GeV. Systematic uncertainties on this expression are obtained by running this fit under many different configurations (e.g. size of the mass window, using MC samples with no corrections applied to the detector alignment), considering additional detector material, and deactivating the toroid magnets to determine the quality of the detector alignment. It is found that the muon energy scale uncertainty is about 0.2%, with negligible uncertainty on the muon momentum resolution.

The systematic uncertainties associated with the reconstructed muon efficiency are also considered. The efficiency is measured using the *tag and probe* method with $Z \to \mu\mu$ events in data [78].[2] The tag and probe muons are required to be isolated and satisfy $p_{\mathrm{T}} > 25$ GeV and 10 GeV, respectively, in addition to other requirements. The probe muons are used to determine the efficiencies of the muon identification algorithm. The same procedure is performed using MC-simulated $Z \to \mu\mu$ samples. The two sets of efficiencies are compared to derive correction factors bringing MC samples in agreement with data. Systematic uncertainties on these scale factors arise from many sources including the background subtraction method used in the probe sample, the cone size used to match ID and MS tracks, and other

[2]Tag and probe is a data-driven technique that exploits the expected kinematic properties of physics objects based on their parent particles to produce a clean sample of those objects. The basic procedure is as follows: a physics object (the *tag*) is selected that satisfies strict criteria, ensuring that it is a true object (as opposed to a fake). A second physics object (the *probe*) is then selected with very loose criteria. If the invariant mass of the tag and probe objects falls within the mass window of the parent particle (e.g. W and Z bosons, J/ψ meson), then the probe is retained. The sample of probes can be used for efficiency studies by determining the fraction that remain after a given cut is applied. The tag and probe need not be the same physics object, and the technique is not immune from background processes, which are usually removed by fitting the mass distribution with a template of the expected distribution for the given parent process, then subtracting the excess.

biases in the tag and probe method. The total systematic uncertainty on the muon efficiency results in a ∼2% uncertainty on the predicted number of signal and background events.

9.1.2 Electron Uncertainties

Three different systematic effects are considered for reconstructed electrons, obtained from detailed studies performed by the ATLAS electron/photon performance group [79], [116]. The first systematic effect pertains to the measurement of the *electron energy scale*, the corrective factor applied to the reconstructed electron energy, whose predominant source of uncertainty is unaccounted material, located in front of the EM calorimeters, that is not modeled in MC-simulated data used for calibration. Since the calibration is based on $Z \to ee$ events in data, electrons with energies around 40 GeV are unaffected by this systematic, while low energy electrons are more adversely affected. To estimate this effect, the energy scale calibration is performed using a dedicated MC $Z \to ee$ sample utilizing a detector geometry model incorporating extra material in front of the EM calorimeters. The difference in scale factors between the two detector geometries is taken as the systematic uncertainty, ranging from -2.0% to 1.2% depending on reconstructed electron E_{T} and η.

The second dominant source of uncertainty arises from unaccounted material in front of the presampler, affecting the calibration of measured energy used to correct that of electrons in the transition region $1.5 < |\eta| < 1.8$. This calibration is performed using $Z \to ee$ events and checked using $W \to e\nu$ events in MC-simulation and data. Based on studies with $W \to e\nu$ events, this systematic effect has an upper limit of 1.4% on the energy scale uncertainty, depending on E_{T} and $|\eta|$. Other systematic effects on the energy scale were studied but found to be small (∼0.1%). The maximum energy scale uncertainty for electrons used in this dissertation is 1.6%.

The second systematic effect for electrons concerns uncertainties in the measured electron energy resolution. As discussed in Chapter 3, the fractional resolution of the EM calorimeter can be modeled as $\sigma_E/E = a/\sqrt{E} \oplus c\%$, where a is the *sampling term* and c is the constant term. At high energies, the resolution is dominated by the constant term. The parameters are extracted from fits to the mass distribution of electrons in $Z \to ee$ processes in data and MC-simulation. The dominant contribution to the uncertainty in the fitted constant term is the uncertainty from the sampling term. The systematic uncertainty is obtained from fitting to MC-simulated data with the sampling term increased by 10%. This yields an uncertainty ranging from 0.3% to 1.8% in the constant term for electrons considered in this analysis, taken to be the energy resolution systematic uncertainty.

Systematic uncertainties on the reconstructed electron efficiency are also considered. This efficiency is a product of the efficiencies for different aspects of electron reconstruction, including the calorimeter isolation efficiency, trigger efficiency, and identification efficiency. Each is measured using the tag and probe technique with $Z \to ee$ events in data. The data efficiencies are then compared to those obtained from MC-simulation to produce scale factors correcting the electron reconstruction efficiency in MC, which is limited by MC mis-modeling. Systematic uncertainties on these scale factors arise from the method of background selection

and varying calorimeter isolation requirements, the size of the mass selection window, and tag/probe requirements. The total systematic uncertainty on the reconstructed electron efficiency, incorporating systematics due to calorimeter isolation, trigger, and identification, is about 1–2%.

9.1.3 Jet Uncertainties

Jets are considered in this analysis only for defining the regions of interest. However, they are an important component of the measured missing transverse energy. Thus, their systematic uncertainties are also considered. Most of the systematic sources pertain to measurements of the *jet energy scale* (JES), the scale factor applied to correct mis-measurements of the jet energy. Systematic effects pertaining to the *jet energy resolution* (JER) and b-tagging are also considered.

9.1.3.1 Energy Scale and Resolution

Detailed studies have been performed measuring the JES and JER and their associated systematic uncertainties [83]. The JES obtained from comparison of data to MC-simulation relies on calibration of the jet energy in data. Two techniques are used for this calibration, both of which utilize momentum balance to correct the measured jet energy. The first method is called *direct balance* and compares the measured jet momentum against that of a photon or leptonically decaying Z boson (known as *reference bosons*), measured with relatively high precision. The second method also considers the momentum of a reference boson, but the balance is performed against the *hadronic recoil*, the vector difference between the missing transverse energy and reference boson momentum, $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}} - \mathbf{p}_{\mathrm{T}}^{Z,\gamma}$. In both of these techniques, the average value of the distribution of $p_{\mathrm{T}}^{\mathrm{jet}}/p_{\mathrm{T}}^{Z/\gamma}$ is taken as the calibration factor. Both methods are applied to events with a single high-p_{T} jet ($p_{\mathrm{T}} > 10/12$ GeV for balance against Z bosons/photons) and all jets having JVF > 0.25 for jets with $p_{\mathrm{T}} < 50$ GeV and $|\eta| < 2.4$ to reject pile-up. The leading jet is required to be roughly back-to-back with the reference boson ($\Delta\phi < 2.8/2.9$ for Z/γ calibration). Sub-leading jets are also required to satisfy p_{T} requirements dependent on the energy of the reference boson. A correction factor is also applied, accounting for radiation (e.g. underlying event) outside of the jet cone. Each method offers a different probe into the calorimeter response and has varying sensitivity to effects such as pile-up and soft QCD radiation.

Many systematic effects for the JES are considered. Variations of the event and object selection criteria affect the JES, as different physics processes can have more of an impact. The requirements on sub-leading jet p_{T} and $\Delta\phi$ between the leading jet and reference boson are varied to determine the sensitivity of the measurement to additional QCD radiation present in the event. JVF requirements are varied to determine the effect of pile-up on the JES. The sensitivity of the JES to the number of primary vertices in the event and the average number of interactions is also considered. The effects of the electron energy scale and resolution and corresponding muon momentum scale and resolution are considered, but

found to be negligible. A relatively significant source of uncertainty pertains to the modeling of jets in MC-simulated data. The difference between measured data-to-MC scale factors between events and jets, modeled using POWHEG and Sherpa [117] samples, is taken as a systematic uncertainty. The statistical uncertainties in the MC-simulated data are considered as a systematic. Also considered is the uncertainty on the correction factor due to radiation outside of the jet cone, but this systematic effect was negligible for jets calibrated using the LCW+JES scheme. For jets calibrated using photons, the systematic effects of photon energy scale and contamination by events in which a jet fakes a photon are considered. The uncertainty in the quark–gluon composition in data as well as the varying response of the calorimeter to different quark flavor jets is also considered. Overall, the JES systematic uncertainty is small, contributing ∼1% uncertainty on the predicted number of signal and background events in the analysis.

The jet energy resolution is determined from MC-simulated data using the direct balance method and considering the width of the $p_{\mathrm{T}}^{\mathrm{jet}}/p_{\mathrm{T}}^{Z/\gamma}$ distribution. Reconstructed jets are matched to a truth jet based on how closely the energy of truth particles in the reconstructed jet cone matches the energy of the truth jet. The width of the $p_{\mathrm{T}}^{\mathrm{jet}}/p_{\mathrm{T}}^{Z/\gamma}$ distribution obtained with the matched reconstructed jet is corrected to account for the nonzero width of the $p_{\mathrm{T}}^{\mathrm{jet}}/p_{\mathrm{T}}^{Z/\gamma}$ distribution obtained with truth jets. A second method for determining the resolution is measuring the width of the $p_{\mathrm{T}}^{\mathrm{jet}}/p_{\mathrm{T}}^{truth-jet}$ distribution, obtained with the matched truth and reconstructed jets. The difference in widths between the two methods is taken as a systematic uncertainty. Another source of JER uncertainty is found by comparing widths using different MC generators. Finally, all systematic uncertainties considered for the JES are propagated into the JER to determine their systematic effect. The JER uncertainty is found to have about a ∼0–3% effect on the predicted number of signal and background events, depending on signal or control region.

9.1.3.2 *b*-Tagging

Systematic uncertainties pertaining to the *b*-tagging efficiency are considered in the analysis. Numerous uncertainties are considered and are described in [86], [87]. Broadly, the uncertainty includes the propagation of JES and JER uncertainties through the *b*-tagging algorithms, the modeling of the jets in MC-simulated data, the normalization of the MC samples used, and uncertainty in the mis-tag rate. The effect of the systematic uncertainty due to *b*-tagging was found to be negligible for the predicted signal and total background distributions in the signal regions.

9.1.4 Hadronic Tau Uncertainties

Hadronic taus are identified from reconstructed jets and are consequently subject to their associated systematic uncertainties. In addition, systematic uncertainties are determined for the identification efficiency and tau energy scale (TES) [88]. The tau identification efficiency

consists of two separate efficiencies: the offline tau ID efficiency[3] and the electron veto efficiency.

The offline efficiency is calculated using the tag and probe technique with $Z \to \tau_{\text{lep}}\tau_{\text{had}}$ events in data. It is practically impossible to have a pure sample of these events, so a cut on the sum of tracks in the core region ($\Delta R < 0.2$) and selected tracks in the outer region ($0.2 < \Delta R < 0.6$) of the tau axis is applied to suppress pile-up and underlying event contributions to the background. An additional advantage of this cut is that the distribution of background events is well-separated from that of signal $Z \to \tau_{\text{lep}}\tau_{\text{had}}$ events. The muons/electrons from the τ_{lep} decay with $p_{\text{T}} > 26$ GeV and satisfying $|\eta|$ and calorimeter isolation requirements are selected as the tag objects. Hadronic tau candidates with one or three core tracks, electric charge $= \pm 1$, $p_{\text{T}} > 15$ GeV, $|\eta| < 2.5$, and satisfying muon and electron overlap removal requirements are selected. Electron and muon vetos are applied for one-prong hadronic taus, and each tau must pass a very loose BDT score requirement to reject jets. Hadronic taus passing these requirements and having the opposite electric charge with respect to the tag leptons populate the probe sample. Other kinematic cuts are applied to suppress background contributions from other Z/W + jets processes. The efficiency is measured by fitting the data to signal and background templates and determining the ratio of the fitted signal events to the total probe sample. Correction factors are then obtained to bring the efficiency from MC-simulation into agreement with that from data.

Systematic uncertainties on these correction factors arise from many sources. The largest signal uncertainty is the modeling of underlying event and parton showering, estimated by comparing the correction factors obtained using ALPGEN+JIMMY and PYTHIA signal samples. Large background uncertainties include the W + jets template shape, estimated by comparing templates obtained in signal and control regions. The statistical uncertainties on jet misidentification are propagated through the efficiency calculation, yielding another systematic uncertainty.

A tag and probe analysis with $Z \to ee$ events determines the efficiency of electrons misidentified as hadronic taus passing the electron veto algorithm. An electron with $p_{\text{T}} > 35$ GeV and located in the hemisphere opposite the misidentified hadronic tau is chosen as the tag electron. The p_{T} cut reduces the $Z \to \tau_{\text{lep}}\tau_{\text{had}}$ background. The corresponding misidentified hadronic tau serves as the probe and is required not to overlap with any reconstructed electrons. The efficiency is computed by taking the ratio of the probe population after and before a veto algorithm is applied. Correction factors are applied to ensure agreement between MC-simulated data and observed data.

Systematic uncertainties on these scale factors largely come from the statistical uncertainties of the remaining probe sample after the veto has been applied. Systematic uncertainties are also assigned based on varying the probe and tag selection criteria as well as the uncertainty in the estimated background. Consideration of these systematics and those associated with the offline tau identication results in a total tau identification uncertainty of 2–3%

[3]This is the algorithm that identifies one- and three-prong taus and has the classifications *loose*, *medium*, and *tight*.

(4–5%) for one-(three-)prong hadronic taus.

Many sources of uncertainty are considered for measuring the tau energy scale, whose measurement was discussed in Section 4.5. Two approaches are used to evaluate these uncertainties, the first being a *deconvolution method* in which the energy difference between reconstructed and truth hadronic taus is determined as well as energy differences between hadronic taus in data and those in simulation (i.e. the modeling of hadronic taus in simulation). This method breaks up a hadronic tau into its constituent particles (e.g. pions) and determines the detector response to each particle (either through test-beam studies, simulation, or other measurements). The uncertainties in the single-particle tests are propagated back to the hadronic tau through the use of pseudo-experiments consisting of varying the particle energies according to both systematic and statistical uncertainties, then determining the effect on the reconstructed tau energy. The TES uncertainty in a pseudo-experiment is found by comparing the reconstructed tau energy obtained through variations of systematic and statistical uncertainties with that obtained from varying only statistical uncertainties. The mean TES shift found from many pseudo-experiments is taken as a systematic uncertainty. Other effects considered in the deconvolution method include pile-up, the modeling of the detector in simulation, underlying event, modeling of parton showering, and post-calibration error between the reconstructed and truth hadronic taus.

An *in-situ method* is also used to determine the TES uncertainty. This method examines $Z \to \tau_\mu \tau_{\text{had}}$ events in data and MC-simulation and compares their *visible mass* distributions, $m_{\tau\tau}^{\text{vis}}$.[4] The reconstructed hadronic tau energy in MC-simulated data is shifted until the peak aligns with the peak in observed data. Some kinematic cuts are placed on the reconstructed hadronic taus and muons to ensure a high purity sample of hadronic taus. Systematic effects due to possible fit bias, missing transverse energy resolution and scale uncertainties, muon momentum resolution and trigger efficiency uncertainties, and uncertainties in the multijet background are considered.

The uncertainties associated with both the deconvolution and in-situ methods yield a total TES systematic uncertainty of 2–4%.

9.1.5 Missing Transverse Energy Uncertainties

Missing transverse energy ($E_{\text{T}}^{\text{miss}}$) is constructed using all tracks and energy clusters, weighted according to their associated objects. It may be computed as shown in Equation 4.4. The systematic uncertainties associated with each term in Equation 4.4 are propagated into the final systematic uncertainty on the $E_{\text{T}}^{\text{miss}}$. Systematic effects unique to the $E_{\text{T}}^{\text{miss}}$ calculation are present in the soft term accounting for low-p_{T} jets and unassociated clusters and tracks.

Uncertainties on the energy scale and resolution of the soft term arise primarily from pile-up and modeling in MC-simulated data and can be estimated using two methods in-

[4]The visible mass is the invariant mass of the reconstructed hadronic tau and muon. The neutrino (represented by the missing transverse energy) is invisible and is thus not part of the calculation.

volving $Z \to \mu\mu$ data and MC-simulated events [118].[5] The first method considers events having no jets with $p_{\mathrm{T}} > 20$ GeV. In these events, the $E_{\mathrm{T}}^{\mathrm{miss}}$ contributions come from the muons and soft term. The data–MC ratio of the distributions of missing transverse energy projected along the Z boson direction is computed. If the scale were measured perfectly, this distribution would sharply peak at unity. Thus, the average deviation from unity is taken as the uncertainty in the scale. The uncertainty on the resolution is evaluated similarly by looking for deviations from unity in the data–MC ratios of E_x^{miss} and E_y^{miss} resolution, which are functions of total transverse energy.[6]

The other method for determining the scale and resolution uncertainties consists of comparing the soft term against the transverse momentum of all the "hard" objects in the event. The hard momentum may be calculated in the manner shown in Equation 9.1:

$$p_{x(y)}^{\mathrm{hard}} = -(E_{x(y)}^{\mathrm{miss}} - E_{x(y)}^{\text{Soft Term}}) + E_{x(y)}^{\mathrm{miss,true}} \tag{9.1}$$

The projections of the missing energy soft term in the transverse (perpendicular) and longitudinal (anti-parallel) directions of the hard transverse momentum determine the uncertainty of the energy scale and resolution. The average length of the longitudinal soft term projection and the resolutions of both longitudinal and transverse soft term projections are determined for different bins of $p_{\mathrm{T}}^{\mathrm{hard}}$. The longitudinal projection is sensitive to the energy scale since it balances directly against $p_{\mathrm{T}}^{\mathrm{hard}}$. The corresponding quantities are obtained from data, and the data–MC ratio can be used to obtain the uncertainty in the soft term energy scale and resolution.[7] The calculations are also performed as a function of the number of primary vertices to determine the effect of pile-up.

The uncertainty in the soft term energy scale and resolution has a small effect on the predicted number of signal and background events, approximately 0–1%.

9.1.6 Luminosity

The measurement of the integrated luminosity is discussed briefly in Chapter 3 and at length in [63]. Many techniques and algorithms measure and calibrate the luminosity. However, several systematic effects result in uncertainty in the measurement, including inaccurate beam centering for the van der Meer (vdM) scans, beam jitter or random fluctuations of the beam position, emittance growth during the course of a vdM scan, variations in the measured visible cross section, fits of vdM to determine the beam spot parameters, background subtraction, variations in concurrently measured beam parameters, inaccurate modeling of the inner detector geometry, electromagnetic interactions between the beams, inability to

[5] As may be inferred from the energy scales described in previous sections, the $E_{\mathrm{T}}^{\mathrm{miss}}$ energy scale characterizes any systematic disagreement between measured and true $E_{\mathrm{T}}^{\mathrm{miss}}$ values.

[6] The $E_{x(y)}^{\mathrm{miss}}$ resolution is computed as the rms width of the distribution $E_{x(y)}^{\mathrm{miss}} - E_{x(y)}^{\mathrm{miss,true}}$, where $E_{x(y)}^{\mathrm{miss,true}}$ is the true missing energy in the x-(y-)direction. For data, the true value is taken to be zero. Hence, the resolution is simply the rms width of the x- and y- components of the missing energy.

[7] The quantity $p_{\mathrm{T}}^{\mathrm{hard}}$ is obtained for data by finding the total transverse momentum vector of all reconstructed objects in the event.

factorize the beam into horizontal and vertical components, and dependence on the number of interactions per bunch crossing. The total uncertainty on the 20.3 fb^{-1} 2012 dataset is found to be 2.8%. This uncertainty is only applied to signal and background processes not normalized to the data.

9.2 Theory Uncertainties

The cross sections of the signal $H \to \mu\tau/e\tau$, $H \to \tau\tau$, diboson, and $Z \to \mu\mu(\mu \to \tau)$ processes are determined from calculations, rather than normalization to data. Thus, uncertainties on the calculated cross sections are considered for these processes, specifically uncertainties on the QCD renormalization and factorization scales, uncertainties on the calculated parton distribution functions, and an uncertainty due to pile-up reweighting.

9.2.1 QCD Scale Uncertainties

To determine the cross sections and kinematics of a process, an infinite number of calculations would be needed, accounting for every possibility of the process occurring. The Feynman diagrams shown for the signal and background processes in Chapters 2 and 6 are examples of leading order processes. Other processes account for radiative corrections and have respective Feynman diagrams featuring loops. Depending on the location of these loops, they account for particle self-interactions and vacuum polarization. Calculating loop diagrams is notoriously difficult, as the related integral expressions are divergent.[8] To produce a finite expression, the integrals are *regularized* by introducing a parameter (the *regulator*) resulting in a convergent expression and yielding the desired answer when set to zero.[9] A widely used regularization scheme is *dimensional regularization* in which the integrals are computed over $4 - \epsilon$ spacetime dimensions, where ϵ is the regulator. In the case $\epsilon = 0$, the calculation reduces to the standard four-dimensional spacetime. Using a *renormalization scheme*, the regulator may be rewritten in terms of another parameter relating to an experimentally observed quantity. A commonly used scheme is the *modified minimal subtraction scheme* ($\overline{\text{MS}}$) that replaces the regulator by a parameter called the *renormalization scale* (μ_R). The choice of this parameter is such that $\mu_R = 1$ corresponds to an experimental observable measured at a conventional energy scale.[10] In principle, the laws of physics are independent of renormalization scale, but any predictive calculation chooses a scale due to the impossibility of calculating an infinite number of diagrams. The cross sections of LHC processes depend on the (renormalized) value of α_S because they occur through the interactions of quarks and gluons in colliding protons.

[8] If all (infinitely many) diagrams were calculated, these infinities would presumably cancel out.

[9] In practice, setting the regulator to zero usually results in some form of infinity, but the parameter can be used to understand how the expression diverges.

[10] The implication of this statement is that some physical constants are actually dependent upon the energy scale in which they are measured. For example, the charge of the electron actually increases in higher energy interactions. Higher energy interactions are equivalently interactions over shorter distances.

In proton–proton collisions, the exact fraction of proton momentum x among the constituent partons is unknown, but can be modeled with a parton distribution function, which is essential for computing cross sections involving protons. While the parton distribution functions depend on x, they also rely on the momentum transferred between the colliding partons, parametrized using the *factorization scale* (μ_F). While the dependence on the factorization scale can be mitigated by performing higher order calculations, it is always present.

For Higgs bosons produced through gluon–gluon fusion, the energy transfer is roughly on the order of the Higgs boson mass, so the factorization and renormalization scales are chosen to be $\mu_R, \mu_F = \sqrt{m_H^2 + (p_{\mathrm{T}}^H)^2}$, where $m_H = 125$ GeV is the Higgs boson mass and p_{T}^H is its transverse momentum. For VBF and associated production, the momentum transferred is roughly on the order of the W boson mass. The factorization and renormalization scales are chosen to be $\mu_R, \mu_F = m_W/2$. The uncertainty due to choice of scales is determined by varying each scale by a factor of two in both direction (while obeying the constraint $0.5 \leq \mu_R/\mu_F \leq 2$) and determining the change in production cross section. These are computed to be 7.8% and 1.0% on the inclusive gluon–gluon fusion and VBF/associated production cross sections, respectively, by the LHC Higgs Cross Section Working Group [119].

A different uncertainty on the cross section is necessary for gluon–gluon fusion samples in Signal Region 1 due to a bias in the distribution of jets introduced by the selection cuts, seen in Figure 9.1. The uncertainty is corrected using a procedure developed by Stewart and Tackmann based on the fraction of events having zero jets, one jet, and more than one jet [120]. Upon applying this procedure, the QCD scale uncertainty on the gluon–gluon fusion production cross section is 10.1%.

QCD scale uncertainties are also calculated for the diboson and $Z \to \mu\mu(\mu \to \tau)$ background predictions [121]. These uncertainties are respectively 5% and 1%.

9.2.2 Parton Distribution Function Uncertainties

Parton distribution functions (PDFs) have been obtained experimentally through the work of different groups and collaborations. Each set differs in choice of dataset, statistical treatment, and parametrization based on partons and the value and uncertainty of α_S, [119]. Consequently, the predicted cross sections vary depending on choice of PDF. This variation is a systematic uncertainty that is assessed by taking the largest difference between predictions using MSTW2008NLO [22], NNPDF [122], and variations of the free parameters in the CT10 PDF [94]. The overall uncertainty for Higgs boson processes is found to be 7.5% for gluon–gluon fusion and 2.8% for VBF and associated production.

The PDF uncertainty is also considered for diboson and $Z \to \mu\mu(\mu \to \tau)$ background predictions [121]. An uncertainty of 4% is applied to the predictions of these backgrounds.

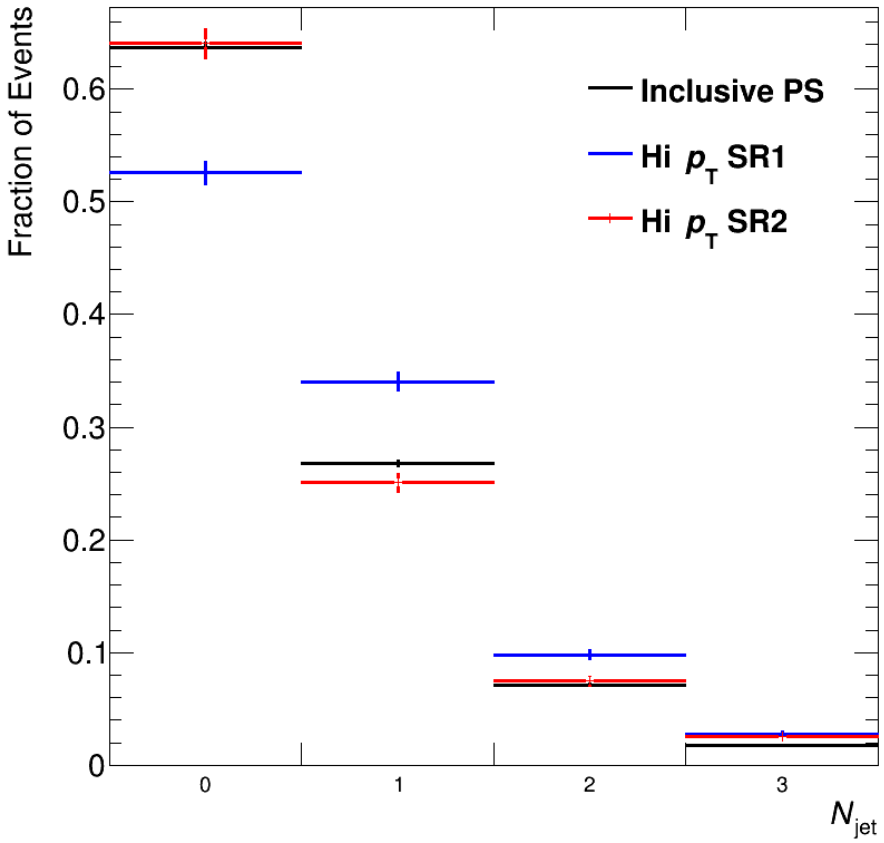


Figure 9.1: Distributions of N_{jet} for LFV Higgs boson signal events produced through gluon–gluon fusion in Signal Region 1, Signal Region 2, and after preselection (inclusive).

9.2.3 Pile-up Reweighting Uncertainty

MC-simulated data samples are often produced before actual data are collected. Consequently, the modeled pile-up distribution is an approximation to the actual one. After data have been taken, a reweighting is applied to the MC-simulated data to bring it into agreement with the observed data. A systematic uncertainty is assigned to account for this procedure. The uncertainty is found to be less than 1% on the relevant signal and background predictions.

9.2.4 $H \to \tau\tau$ Branching Ratio Uncertainty

The uncertainty on the calculated $H \to \tau\tau$ cross section was studied by the LHC Higgs Cross Section Working Group [123]. An uncertainty of 5.7% is assigned to the predicted number of $H \to \tau\tau$ events.

9.3 Methodological Uncertainties

A third class of uncertainties pertaining to the methodology of the analysis are considered. These uncertainties are found to have the largest impact on the final result and consist of uncertainties associated with the production of $Z \to \tau\tau$ embedding samples, the measurement of r_{QCD}, shape corrections obtained for the W + jets and $Z \to \mu\mu/ee(\mathrm{jet} \to \tau)$ samples, extrapolation uncertainties associated with the normalization and modeling of W + jets and top backgrounds, and uncertainties on the obtained k-factors.

9.3.1 Uncertainties with the $Z \to \tau\tau$ Embedding Technique

The embedding technique used to model the $Z \to \tau\tau$ background was described in Section 6.2.3 and consists of replacing the reconstructed muons in $Z \to \mu\mu$ events in data with taus taken from MC-simulated data, then decaying the taus and reconstructing the detector signature to obtain a $Z \to \tau\tau$ sample, precluding the need to model pile-up, underlying event, missing transverse energy, and other aspects of the event environment difficult to model accurately using MC-simulated data.

Two systematic uncertainties are considered for the embedding technique [108]. The first uncertainty is obtained by varying the inner detector isolation requirements used in the selection of the muons in data. One variation consists of removing the isolation requirement entirely, while the other consists of tightening the requirements such that $I(p_{\mathrm{T}}, 0.4)/p_{\mathrm{T}}(\mu) < 0.06$ and $I(E_{\mathrm{T}}, 0.2)/p_{\mathrm{T}}(\mu) < 0.04$. These variations determine the effects that muon isolation has on the event environment.

The other systematic uncertainty pertains to the subtraction of the reconstructed muon energy deposits in the calorimeters. This subtraction is based on the deposits predicted in MC-simulated data. There are large uncertainties in these predictions, and the systematic effect was studied by scaling the predicted deposits by $\pm 20\%$ to determine the effect on the background predictions.

Both systematic variations result in a 4–7% uncertainty on the predicted number of $Z \to \tau\tau$ events.

9.3.2 r_{QCD} Measurement Uncertainty

The measurement of r_{QCD} in a dedicated study in the ATLAS SM $H \to \tau\tau$ search was discussed in Section 6.2.8. Several systematic effects on the measured r_{QCD} values are considered

[102], including eliminating the lepton track isolation requirement ($I(p_{\mathrm{T}}, 0.4)/p_{\mathrm{T}}(\ell)$), performing the study with *medium*-BDT hadronic taus, varying the upper end of the calorimeter isolation ($I(E_{\mathrm{T}}, 0.2)/p_{\mathrm{T}}(\ell)$) fit range by one bin in each direction, and fixing the calorimeter isolation and performing the fit over the track isolation. These systematic variations are combined with the statistical uncertainty to obtain a 13% uncertainty on the measured r_{QCD} value in both muon and electron measurements.

9.3.3 Shape Uncertainties

Shape corrections to the $p_{\mathrm{T}}(\tau)$, $|\Delta\eta(\ell, \tau)|$, and N_{jet} distributions are obtained for the W + jets and $Z \to \mu\mu/ee(\mathrm{jet} \to \tau)$ backgrounds due to Pythia mis-modeling. A systematic effect is assigned to each of these backgrounds to account for this correction.

Two sets of shape corrections are obtained for the W + jets background in the W control region and $m_{\mu/e\tau}^{\mathrm{MMC}} > 150$ GeV region of Signal Region 1. An estimate of the shape correction systematic in SR1 is found by obtaining the W + jets predictions using the shape correction obtained in the WCR, then taking half the difference between the WCR and SR1 shape predictions. The choice to take half of the difference between the shapes is made because the true shape distribution is unknown, and any attempt to assign a systematic is an educated guess, at best.

In Signal Region 2, no shape corrections are applied. However, a systematic on the shape from the Pythia prediction is determined by obtaining predictions of W + jets events in SR2 using the shape obtained in SR1, then taking half of the difference between the default and SR1 shapes.

The systematic uncertainties due to shape corrections to the W + jets background are the largest sources of uncertainty in the analysis, with 10/12% uncertainties on the predicted number of W + jets events in the signal window of $110 < m_{\mu/e\tau}^{\mathrm{MMC}} < 150$ GeV.

A systematic uncertainty on the shape corrections to the $Z \to \mu\mu/ee(\mathrm{jet} \to \tau)$ backgrounds is determined by comparing the predictions with the correction to those without and taking half of the difference. This systematic uncertainty had a negligible effect on the result of the binned likelihood fit used to extract the signal (See Chapter 10).

9.3.4 Extrapolation Uncertainties

Correction factors (k-factors) correct the predicted number of events in MC-simulated data to observed data for both W + jets and top backgrounds. These are obtained in dedicated control regions and applied in signal regions. However, there is no guarantee that these factors remain valid outside of their dedicated control regions. To account for the possibility of inaccuracies in the k-factor modeling, *extrapolation uncertainties* are determined for the W + jets and top backgrounds.[11] These uncertainties are only applied in SR1 and SR2.

[11] An extrapolation uncertainty was not considered for the $Z \to \mu\mu/ee(\mathrm{jet} \to \tau)$ background due to its tiny size.

To determine the extrapolation uncertainty for W + jets, background predictions in SR1 and SR2 are obtained using MC-simulated data based on the ALPGEN+HERWIG generators.[12] These predictions are then compared with those made using the default ALPGEN+PYTHIA W + jets samples, with the difference taken as a systematic uncertainty.

The normalization of the $Z \to \tau\tau$ background is dependent upon the predicted number of W + jets events. Thus, it can be affected by the extrapolation uncertainty associated with the W + jets background. The normalization was recomputed incorporating the predictions made by ALPGEN+HERWIG W + jets in SR2. The difference in the number of $Z \to \tau\tau$ events using the default and recomputed normalizations is taken as a systematic uncertainty on $Z \to \tau\tau$ events.[13] These uncertainties were initially estimated to be 6.5(1.8)% and 0.6(18)% for SR1 and SR2 in the muon (electron) measurement, but they are constrained by the data in the binned likelihood fit (see Chapter 10).

The extrapolation uncertainty on the top background is determined in a similar manner. Predictions of the number of top events are obtained in SR1 and SR2 using MC@NLO. The difference in MC@NLO and default POWHEG predictions is taken as a systematic uncertainty for the top background. These uncertainties are found to be 7.2(3.0)% and 3.7(15)% in SR1 and SR2 in the muon (electron) measurement.

9.3.5 Uncertainties on k-Factors

The statistical uncertainties on each k-factor in Table 6.1 are propagated through the analysis to determine the uncertainty in the predicted number of corresponding background events. A systematic uncertainty for each k-factor is assigned to the corresponding background.

[12] k-Factors for the ALPGEN+HERWIG W + jets are obtained in the WCR.

[13] This uncertainty was also applied to $Z \to \tau\tau$ events in the WCR.

Chapter 10

Signal Extraction

The previous chapters described how the electronic signatures produced by particles passing through the detector are reconstructed into various physics objects, collectively referred as *data*. In a proton–proton collision, every kinematic distribution of data is a composite of the effects of different physics processes. However, the exact composition of the data in terms of the constituent processes is unknown and may only be estimated using simulated data or data-driven techniques. In searches for new physics, a (generally) large fraction of the data is comprised of background processes, while the signal process manifests as an excess of data above the predicted total background distribution. The process of determining the amount of signal in the data is called *signal extraction*.

Effective signal extraction requires a *discriminating variable* in which the behavior of the signal is different from the behavior of the background. The contributions of the signal and background processes in the observed data are estimated in the distributions of the discriminating variable, given systematic and statistical uncertainties on the signal and background processes. The distributions of the discriminating variable for the signal and background processes and their systematic uncertainties in given signal and control regions can be assembled together in a mathematical function called a *binned likelihood function*, also known as a *fit model*.

Statistical methods are applied to the fit model to determine whether or not the data are more consistent with signal + background or background-only hypotheses. If the data are consistent with the signal + background hypothesis, the amount of signal is measured and compared with the predicted value. Generally, the presence of signal is characterized by an excess of data events above the predicted number of background events. If the data are not consistent with the presence of the signal, the degree of signal absence is characterized using an interval called an *exclusion limit*.

The goal of this dissertation is to determine whether or not the observed 125 GeV Higgs boson has LFV decays involving hadronic taus. The $m_{\mu/e\tau}^{\mathrm{MMC}}$ distribution is chosen as the discriminating variable since the Higgs boson mass distribution peaks at 125 GeV, while the mass distributions for other processes peak at lower values. If LFV Higgs boson decays are observed, then the branching ratio of LFV process will be measured. If the decays are not

observed, then the upper exclusion limit on the branching ratio will be determined.

This chapter discusses the statistical methodology necessary to determine the consistency of the signal and data.[1] The structure of the fit model and the utilized software will also be discussed.

10.1 Statistical Tools and Techniques

In any sort of experimental study, data are gathered and analyzed to answer the question posed in the problem statement. A standard aspect of experimental design is the statement of the *hypothesis*, or a testable conjecture answering the experimental question. Scientific hypotheses are translated into *statistical hypotheses* consisting of statements about the nature of one or more parameters of the data,[2] which are estimated using *statistics*, or quantities computed from the data. The statistical procedure of *hypothesis testing* determines whether or not the data are consistent with a given hypothesis by using a *test statistic* comparing an estimate of the parameter of interest with the hypothesized true value and having a known *sampling distribution*.[3]

The sampling distribution of a test statistic is also useful for setting *exclusion limits*, intuitively understood as a one-sided bound (i.e. upper or lower) on the true value of a parameter, though the exact interpretation is more nuanced and depends on method of statistical inference. In *frequentist inference*, the probability of a given outcome is interpreted as its relative frequency to the other outcomes over an ensemble of an infinite number of trials [125].[4] For example, a frequentist would interpret the probability of a tossed fair six-sided die landing on "3" being $\frac{1}{6}$ as the relative fraction, or *frequency*, of an infinite number of tosses of fair six-sided dice landing on "3". A consequence of the frequentist interpretation is that the measurement of the true parameters of the data would require an ensemble of an infinite number of datasets. Thus, the parameters can only be approximated from the observed dataset. In 1937, J. Neyman developed a method for constructing statements called *confidence intervals* (CIs) that are *guaranteed* to contain the true value of the parameter in at least some fraction $1-\alpha$ of the samples of data taken [127]. The number $1-\alpha$ is called the *confidence level*.[5] The CI for a given parameter is constructed using the sampling distribution of the test statistic and the parameters measured from a given dataset. An exclusion limit

[1] The statistical ideas presented in this chapter are discussed in more detail in [124].

[2] Some examples of parameters are the mean and variance of a dataset.

[3] The sampling distribution of a test statistic refers to the distribution of test statistics evaluated for representative samples taken from a dataset.

[4] The other method of statistical inference is called *Bayesian inference* in which probability is interpreted as the "degree of belief" in a particular outcome rather than the relative frequency of a particular outcome. There are situations when Bayesian inference is crucial, such as evaluating risk. One would not want to determine whether or not it is safe to cross a road by making many trials and measuring the fraction of times struck by a car during the crossing! Bayesian methods of data analysis are the focus of [126] and good discussion can be found in [125].

[5] The meaning of α will be discussed in the next section.

usually consists of a one-sided CI, subject to the interpretation that values of the parameter outside of the interval are not consistent with observed data.

In this dissertation, the branching ratio on the LFV Higgs boson decays to hadronic taus and light leptons is considered to be the parameter of interest. The statistical techniques and tools used are the frequentist methods agreed upon by ATLAS and CMS for combinations of Higgs boson searches [128].[6]

10.1.1 Hypothesis Testing

The technique of statistical hypothesis testing presents two mutually exclusive statements called *hypotheses* about a parameter in the model [129]. The default statement is usually formulated as the parameter having a specified value and is called the *null hypothesis* H_0. The second hypothesis is called the *alternative hypothesis* H_1. The alternative hypothesis can either be constructed as *two-sided*, where the parameter is stated not to equal the value specified in the null hypothesis, or as *one-sided*, in which the parameter is greater or less than the value specified in the null case.

In principle, the true value of a parameter for a distribution cannot be known since it would require an ensemble of an infinite number of datasets. However, it is possible to estimate the parameter from a single dataset using a function called an *estimator*. For example, the mean of a distribution can be estimated by adding all the data-points together, then dividing by the total number of data-points. A good estimator should be *unbiased*, meaning that its average over different datasets should equal the true value of the parameter.[7] Additionally, the estimator should have as small a variance as possible.[8] Finally, a good estimator should be *consistent*: in the limit of an infinite number of data-points, the estimator converges on the true value of the parameter.

Once the estimator has been chosen, a *test statistic* compares the data to the value specified in the null hypothesis. By observing where the computed test statistic falls in its sampling distribution, a decision can be made whether or not to reject the null hypothesis in favor of the alternative hypothesis by defining a *critical region* in the sampling distribution and choosing to reject the null hypothesis if the test statistic falls within this region. An important consideration is that the size of the critical region is arbitrary, but its choice affects the probability that two types of errors could occur. The first error type is called a *type I error* and occurs when the null hypothesis is rejected when it is actually true. It has probability α of occurring.[9] The second type of error occurs when the hypothesis test fails to reject the null hypothesis when it is false. This is called a *type II error*, and the probability of its occurrence is labeled as β.

The sampling distribution of a test statistic is a probability density function. Therefore, the area under its curve is equal to one, by definition. Thus, the probability of a type I error

[6] It should be noted that the technique of hypothesis testing is also a frequentist method.

[7] In other words, the expectation value of the estimator should be equal to the true value of the parameter.

[8] Such an estimator is called a *minimum variance unbiased estimator*.

[9] This is the same α from the definition of the CI.

α equals the area under the curve in the critical region.[10] Thus, reducing the size of the critical region lowers α. In practice, α is chosen to be 0.05, but may be reduced as demanded by the situation. An important consideration is that a smaller critical region increases the probability that a type II error may occur, should the true value of the parameter be different from what is being tested in the null hypothesis.[11]

At this point, to test a hypothesis, a test statistic is computed and critical region determined based on a desired significance, α. If the test statistic falls inside the critical region, then the null hypothesis is rejected in favor of the alternative hypothesis. Due to the arbitrary nature of the significance, it is useful to know whether or not the null hypothesis was barely or broadly rejected. This knowledge is encoded in a quantity called the *p-value*, which is the minimum significance required to reject the null hypothesis for a given test statistic. Another interpretation of the p-value is that it is the probability that another independent, identically distributed dataset would yield a test statistic larger than the one obtained given the null hypothesis. The smaller the p-value, the more improbable this outcome, and hence the more likely it is that the data are inconsistent with the null hypothesis. Unlike the significance, the p-value is not an arbitrary quantity. Consequently, it is customary to report it along with the decision to reject the null hypothesis. The p-value is usually reported as the number of standard deviations the computed test statistic is located from the central value of the sampling distribution.

In addition to estimates about the value of a parameter, it is also possible to estimate the range of possible values the true value of the parameter may take.[12] Under frequentist inference, the $100(1-\alpha)\%$ *confidence interval* (CI) may be constructed. Using the sampling distribution of the estimator, the CI is a statement of the form $l(\hat{\mu}) \leq \mu \leq u(\hat{\mu})$ such that $P(l(\hat{\mu}) \leq \mu \leq u(\hat{\mu})) \geq 1-\alpha$, where α is the significance level. $l(\hat{\mu})$ and $u(\hat{\mu})$ are the lower and upper ends of the interval determined from the estimated value ($\hat{\mu}$) of the true parameter (μ). It is important to note that the CI is unique to a given estimator for a particular dataset. Estimators computed on other identically distributed datasets will have different CIs, leading to the interpretation of the CI: if CIs are computed for each dataset of an infinite ensemble of datasets, then $100(1-\alpha)\%$ of them will contain the true value of the parameter. While lowering α increases the probability that the CI contains the true value of the parameter, it has the disadvantage of making the interval longer (increasing the uncertainty in the value of μ).

The confidence intervals described are called *two-sided confidence intervals*. One-sided $100(1-\alpha)\%$ *confidence limits* (or *confidence bounds*) on a parameter that are statements of the form $l(\hat{\mu}) \leq \mu$ and $\mu \leq u(\hat{\mu})$ for lower and upper limits may also be computed. They

[10] The value of α is also called the *significance level*.

[11] The ability of a statistical test to distinguish between the hypothesized and true values of a parameter is called its *power*, computed as $1-\beta$, and depends on the separation between the two hypotheses and the number of data-points. Increasing the power of a statistical test is one of the reasons why it is preferable to have "high stats."

[12] The estimates about the single value of the parameter are called *point estimates*, while those of the range of values are called *interval estimates*.

respectively correspond to the probability statements $P(l(\hat{\mu}) \leq \mu) \geq 1-\alpha$ and $P(l(\mu \leq u(\hat{\mu})) \geq 1-\alpha$.

To illustrate these concepts, consider a simple example: Suppose that a company produces wooden pencils. To properly fit inside their boxes without rattling around or being too long, the pencils must be 10 cm in length. The standard deviation of the pencil length is known to be $\sigma = 0.5$ cm. A random sampling of 25 pencils has a mean length of $\bar{x} = 10.3$ cm. Do the pencils meet the specification? What is the 95% confidence interval on the mean length? Would the test be able to tell if the true mean of the pencils was 10.3 cm?

The first thing to do is specify the hypotheses. If μ is the true mean pencil length, then $H_0 : \mu = 10$ cm and $H_1 : \mu \neq 10$ cm. Since the length of each pencil is a random variable, the mean pencil length, $\bar{x}$, is a random variable with an approximately Gaussian distribution, according to the central limit theorem. Thus, the following test statistic may be used for the hypothesis test:

$$z_0 = \frac{\bar{x} - \mu}{\sigma/\sqrt{N}} \tag{10.1}$$

where μ is the hypothesized true value, σ is the known standard deviation, and N is the size of the sample. The sampling distribution for this test is the *standard normal distribution*, which has the functional form $f(x) = \frac{1}{\sqrt{2\pi}} \exp(\frac{-x^2}{2})$. Computing the test statistic yields $z_0 = (10.3 \text{ cm} - 10 \text{ cm})/((0.5 \text{ cm})/\sqrt{25}) = 3.00$. Since this test is two-sided,[13] the p-value is found by calculating $p = 2(1-\Phi(z_0))$, where $\Phi(z_0)$ is the *cumulative distribution function* (CDF) of the standard normal distribution, evaluated at z_0.[14] In this case, $p = 2(1-\Phi(3.00)) = 0.0027$. Thus, at a significance level of 0.05, the null hypothesis may be rejected in favor of the alternative hypothesis. In other words, there is a statistically significant difference between the sample mean and hypothesized true mean values. Note that the choice of significance 0.05 corresponds to a critical region of $|z| > z_{\alpha/2}$ or $|z| > z_{0.05/2} = 1.96$. Figure 10.1 illustrates the standard normal distribution. The critical region is shaded in red, and the observed test statistic is indicated by the blue line.

Based on Figure 10.1, The $100(1-\alpha)\%$ CI on the mean length corresponds to a statement of the form $P(-z_{\alpha/2} \leq z_0 \leq z_{\alpha/2}) = 1-\alpha$. Thus, the desired $100(1-\alpha)\%$ CI is simply $-z_{\alpha/2} \leq z_0 \leq z_{\alpha/2}$. Using Equation 10.1, this can be rearranged as $\bar{x} - \frac{z_{\alpha/2}\sigma}{\sqrt{N}} \leq \mu \leq \bar{x} + \frac{z_{\alpha/2}\sigma}{\sqrt{N}}$. Substituting in the computed values yields a 95% CI of $10.3 \text{ cm} - (1.96)(2.0 \text{ cm})/\sqrt{25} \leq \mu \leq 10.3 \text{ cm} + (1.96)(2.0 \text{ cm})/\sqrt{25}$, or $10.1 \text{ cm} \leq \mu \leq 10.5$ cm.

To determine whether or not the test is sensitive to a true mean of 10.3 cm, it is sufficient to calculate the power of the test with an alternative hypothesis of $\mu = 10.3$ cm. It can be shown [129] that the sampling distribution for such an alternative hypothesis is a Gaussian distribution with unit variance and a mean of $\frac{\delta\sqrt{N}}{\sigma}$, where δ is the difference in parameter

[13] The alternative hypothesis specifies that the true value of μ could be greater or less than 10 cm, hence being two-sided.

[14] The CDF returns the area underneath the probability density function from negative infinity to the value of the argument.

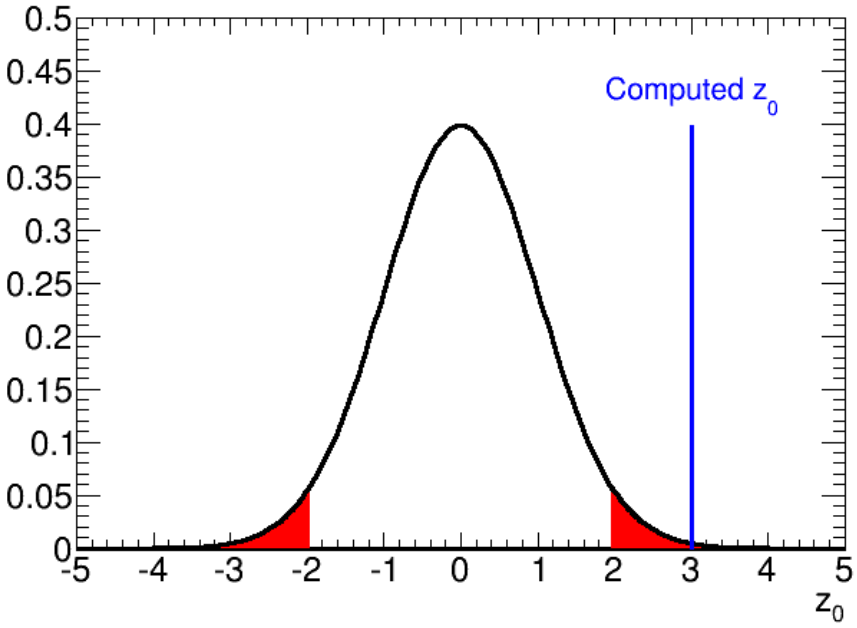


Figure 10.1: The sampling distribution for z_0, the standard normal distribution. The red shaded areas are the critical region, bounded by $|z_0| > 1.96$ and chosen to comprise 5% of the total area of the distribution, indicating a significance level of 0.05. The line corresponds to the observed z_0 computed in the example.

values between the alternative and null hypotheses. For this example, the mean of the alternative hypothesis sampling distribution is (10.3 cm−10 cm)$\sqrt{25}$/(2.0 cm) = 3. The sampling distributions for the null and alternative hypotheses are shown in Figure 10.2. If the null hypothesis is tested at a significance level of 0.05, then the area of the red region in the figure is the probability of making a type II error (β) since the null hypothesis would not be rejected in favor of the (true) alternative hypothesis. The value of beta is found by evaluating $\beta = \Phi(-z_\beta) = \Phi(z_{\alpha/2} - \frac{\delta\sqrt{N}}{\sigma}) = \Phi(1.96 - 3) = \Phi(-1.04)$.[15] Evaluating this yields $\beta = 0.149$. Thus, the power of the test is $1 - \beta = 0.851$, indicating reasonable sensitivity to a true mean pencil length of 10.3 cm.

10.1.2 Parameter Estimation

The previous section introduced the fundamental concepts of hypothesis testing, one of whose key ingredients is an unbiased estimator of the parameter of interest having minimal

[15]The reader may notice that the area under the null hypothesis to the left of $z_0 = -1.96$ has not been subtracted out, though the null hypothesis would be rejected in favor of the alternative hypothesis in this region. In most cases, the area of this region is so small that it may be safely ignored without compromising the calculation.

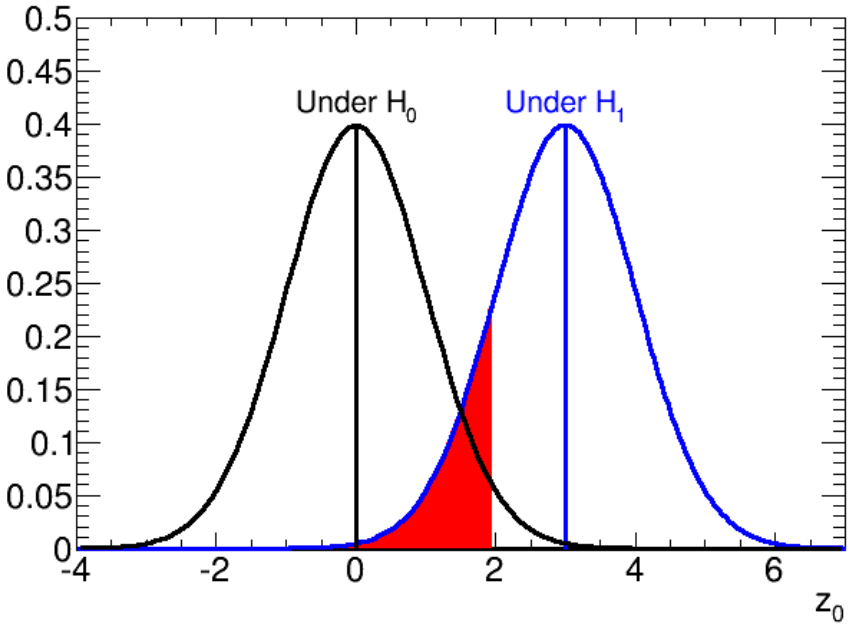


Figure 10.2: The sampling distributions for the null and alternative hypotheses used in the example, centered respectively at zero and three. The red shaded area equals the probability of making a type II error since the null hypothesis would not be rejected in favor of the (true) alternative hypothesis. This area is chosen under the assumption that the null hypothesis is being tested at a significance level of 0.05.

variance. In general, it is not obvious how an estimator for a parameter may be obtained, especially in models where there may be correlations between different parameters. Yet, techniques have been developed to obtain estimators that are (approximately) unbiased and satisfy the requirements of minimal variance. One of the more widely used methods is the *method of maximum likelihood*, developed by the eminent statistician Sir R. A. Fisher in early twentieth century.

The method of maximum likelihood works as follows: Consider a dataset $\mathbf{x}$ with datapoints $x_1, x_2, \ldots, x_N$. Suppose that the dataset depends on an unknown parameter of interest θ. If the data are distributed according to some probability distribution $f(x|\theta)$, then the *likelihood function* may be written as shown in Equation 10.2:

$$\mathcal{L}(\theta) = \prod_{i=1}^{N} f(x_i|\theta) \tag{10.2}$$

Equation 10.2 is the probability of obtaining the observed dataset given a specific value of the parameter θ. The value of θ maximizing the likelihood is called the *maximum likelihood*

estimator of θ, which may be found by solving for the value of θ such that $\frac{d\mathcal{L}}{d\theta} = 0$. In practice, it is more convenient to consider the natural logarithm of the likelihood function (known as the *log-likelihood*, $\ln \mathcal{L}$) and maximize it with respect to θ: $\frac{d \ln \mathcal{L}}{d\theta} = 0$.

Suppose that the data depend on several parameters, $\theta_1, \theta_2, \ldots, \theta_k$.[16] The likelihood function then becomes a function of those parameters ($\mathcal{L}(\theta_1, \theta_2, \ldots, \theta_k)$), and the estimator for θ_i may be found by finding the particular value maximizing the likelihood function, keeping all other parameters constant.

Maximum likelihood estimators have many useful properties, particularly in the *asymptotic limit* as the number of data-points approaches infinity. The estimators for each parameter are approximately unbiased and consistent, satisfy the minimal variance requirement, and have Gaussian sampling distributions, provided that the true value of the parameter does not lie on the boundary of the parameter space [130].

10.1.3 Nuisance Parameters

In any search for new physics, one has a histogram $\mathbf{n}$ with N bins and n_i data events in the i-th bin. The data are comprised of background and possible signal events whose exact composition is the goal of the study. The predicted number of events in each bin can be written as $\mu s_i + b_i$, where s_i and b_i are the signal and background events, and μ is a parameter called the *signal strength*. If there is no signal, then $\mu = 0$ and the data are consistent with the estimated number of background events. Otherwise, $\mu = 1$ implies that data are consistent with the estimated number of signal and background events.

It is desired to know the value of μ to determine whether or not new physics may be present. This can be done by constructing a likelihood function, assuming that the data in each bin are Poisson distributed with the predicted number of events equal to the expected value.[17] Thus, the likelihood function may be written as shown in Equation 10.3:

$$\mathcal{L}(\mu) = \prod_{i=1}^{N} \frac{(\mu s_i + b_i)^{n_i} e^{-(\mu s_i + b_i)}}{n_i!} \tag{10.3}$$

If the number of predicted signal and background events were known precisely with no uncertainties, then the maximum likelihood estimate for μ (known as $\hat{\mu}$) would simply be the value of μ maximizing the function in Equation 10.3.[18] Unfortunately, the precision of the predicted number of events is degraded by statistical and systematic uncertainties, which are modeled as additional parameters in the likelihood function called *nuisance parameters* (the desired parameter μ is called the *parameter of interest*). Letting $\boldsymbol{\theta}$ represent all nuisance

[16] The k-dimensional space of all possible values of the parameters θ_i is called the *parameter space*.

[17] A Poisson distribution is a discrete probability distribution modeling the probability that k random events will occur in an interval of time $1/\lambda$, given that they happen on average at a rate of λ. The functional form is $\text{Poisson}(k|\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$.

[18] It is easy to show that for a histogram of one bin, $\hat{\mu} = \frac{N-B}{S}$, where N, S, and B are the number of data, signal, and background events.

parameters, the likelihood function in Equation 10.3 can be rewritten as shown in Equation 10.4:

$$\mathcal{L}(\mu, \boldsymbol{\theta}) = \prod_{i=1}^{N} \frac{(\mu s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta}))^{n_i} e^{-(\mu s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta}))}}{n_i!} \tag{10.4}$$

Unlike Equation 10.3, there may be multiple solutions maximizing the likelihood function in Equation 10.4 since the nuisance parameters can change the estimated values of signal and background. Furthermore, there may be correlations between the nuisance parameters and the parameter of interest. The presence of nuisance parameters degrades the precision of the estimated parameter of interest as well as possibly biasing its estimate. These issues are mitigated by introducing bins from a control region where little to no signal is expected. The observed data can constrain the nuisance parameters not unique to the signal process, reducing the range of their possible values. If $\mathbf{m}$ is a control histogram with M bins, each bin populated with m_j data events against an estimated b_j background events, then the likelihood function can be modified as in Equation 10.5:

$$\mathcal{L}(\mu, \boldsymbol{\theta}) = \prod_{i=1}^{N} \frac{(\mu s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta}))^{n_i} e^{-(\mu s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta}))}}{n_i!} \prod_{j=1}^{M} \frac{b_j(\boldsymbol{\theta})^{m_j} e^{-b_j(\boldsymbol{\theta}))}}{m_j!} \tag{10.5}$$

Fundamentally, the problem still remains that the desired likelihood should only be a function of the parameter of interest. Any computed p-value testing the parameter of interest should be independent of the values of any of the nuisance parameters. A solution is simply to replace the nuisance parameters by their values maximizing the likelihood function for a given value of the parameter of interest.[19] This technique is called *profiling* and leads to the construction of an object called the *profile likelihood ratio* (λ), defined in Equation 10.6:

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \breve{\boldsymbol{\theta}}(\mu))}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})} \tag{10.6}$$

where $\breve{\boldsymbol{\theta}}$ is the *conditional maximum likelihood estimator* (CMLE) of $\boldsymbol{\theta}$ and is a function of the parameter of interest μ. $\hat{\boldsymbol{\theta}}$ and $\hat{\mu}$ are the (unconditional) maximum likelihood estimators of $\boldsymbol{\theta}$ and μ. $\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})$ refers to the absolute maximum value of the likelihood function. The CMLEs broaden the shape of λ about $\hat{\mu}$, resulting in larger uncertainties in the estimated value of μ, consistent with the expected effect of the nuisance parameters.

10.1.4 Test Statistics and Exclusion Limits

By construction, the possible values of the profile likelihood ratio range from zero to one, with values close to one indicating that the data are consistent with the tested value of μ,

[19] See the review of statistics in [10] for more discussion.

while those close to zero suggest otherwise. A convenient test statistic [131] exploiting the behavior of the profile likelihood ratio is shown in Equation 10.7:

$$t_\mu = -2 \ln \lambda(\mu) \tag{10.7}$$

The values of t_μ range from zero to infinity. Large values of t_μ suggest increasing disagreement between data and the tested value of μ. Thus, the p-value can be used to exclude possible values of the signal strength. To know the p-value, the sampling distribution for t_μ, given a particular value of μ, must be known. From Wilks' theorem, this distribution in the asymptotic limit is a χ^2 distribution with one degree of freedom [132].

For determining a discovery or placing exclusion limits, modified versions of t_μ are required, as it is sensitive to both upward and downward fluctuations in data. Consider the test statistic q_0, used for testing discovery and shown in Equation 10.8:

$$q_0 = \begin{cases} -2 \ln \lambda(0) & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases} \tag{10.8}$$

This statistic tests the background-only model (with $\mu = 0$) against the observed data. A low p-value from the observed q_0 statistic indicates that the data are not consistent with the background-only hypothesis, implying the presence of signal. This statistic is constructed so that only upward fluctuations of data, leading to positive values of $\hat{\mu}$, are counted against the background-only hypothesis. A downward fluctuation of data results in $\hat{\mu} < 0$, which does not imply the presence of signal as a negative signal strength μ is unphysical.[20] In the asymptotic limit, it can be shown [131] that the sampling distribution for q_0 ($f(q_0|0)$) is a χ^2 distribution with one degree of freedom, having observed p-value $p_0 = \int_{q_0^{obs}}^{\infty} f(q_0|0) dq_0 = 1 - \Phi(\sqrt{q_0})$, corresponding to a significance of $z_0 = \sqrt{q_0}$. This is the significance that is quoted when an excess of data events are observed above the predicted number of background events.

Another test statistic called q_μ is used for determining exclusion limits on the signal strength and is given by Equation 10.9 [131]:

$$q_\mu = \begin{cases} -2 \ln \lambda(\mu) & \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases} \tag{10.9}$$

This test statistic is similar to q_0 in Equation 10.8, but there are some important differences. Under the discovery test statistic q_0, downward fluctuations of the data are not counted against the background-only hypothesis. For the statistic q_μ, the null hypothesis is that

[20] To understand this better, consider this example: suppose one has ten observed data events. If six background events and one signal event with negligible uncertainty are predicted, then the data are most consistent with four signal events, leading to a positive estimate of the signal strength $\hat{\mu} = 4$. On the other hand, if one predicts thirteen background events and one signal event with negligible uncertainty, then the estimated signal strength is negative, $\hat{\mu} = -3$. The second case does not imply that there is negative signal!

there is both signal and background. Thus, upward fluctuations of the data (resulting in nonzero values for $\hat{\mu}$) are not counted against the signal + background hypothesis, hence $q_\mu = 0$ for $\hat{\mu} > \mu$. For setting exclusion limits, it is necessary to compute p-values assuming both signal + background and background-only hypotheses. The most general form of the sampling distribution, $f(q_\mu|\mu')$, is given by Equation 10.10:

$$f(q_\mu|\mu') = \Phi\left(\frac{\mu'-\mu}{\sigma}\right)\delta(q_\mu) + \frac{1}{\sqrt{8\pi q_\mu}}\exp\left[-\frac{1}{2}\left(\sqrt{q_\mu} - \frac{\mu'-\mu}{\sigma}\right)^2\right] \tag{10.10}$$

where $\delta(q_\mu)$ is the Dirac delta function, and σ is the standard deviation of the distribution of the estimator $\hat{\mu}$, found by calculating the covariance matrix of the model parameters, $V_{ij} = \text{cov}(\hat{\theta}_i, \hat{\theta}_j)$. Under this notation, the parameter of interest $\mu = \theta_0$ and $\sigma^2 = V_{00}$. The inverse of the covariance matrix may be computed from the likelihood function as $V_{ij}^{-1} = -E\left[\frac{\partial^2 \ln \mathcal{L}}{\partial\theta_i \partial\theta_j}\right]$. In the case where $\mu' = \mu$, Equation 10.10 simplifies, and the p-value may be computed as $p_\mu = 1 - \Phi(\sqrt{q_\mu})$.

Another important test statistic is considered for models in which a negative signal strength is unphysical (that is $\mu \geq 0$). Known as $\tilde{q}_\mu$, it has the form given in Equation 10.11 [131]:

$$\tilde{q}_\mu = \begin{cases} -2\ln\frac{\mathcal{L}(\mu, \breve{\boldsymbol{\theta}}(\mu))}{\mathcal{L}(0, \breve{\boldsymbol{\theta}}(0))} & \hat{\mu} < 0 \\ -2\ln\frac{\mathcal{L}(\mu, \breve{\boldsymbol{\theta}}(\mu))}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})} & 0 \leq \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases} \tag{10.11}$$

where $\breve{\boldsymbol{\theta}}(\mu)$ and $\breve{\boldsymbol{\theta}}(0)$ are the CMLEs of $\boldsymbol{\theta}$ evaluated for signal + background and background-only hypotheses. The sampling distribution is similar to Equation 10.10, given in Equation 10.12:

$$f(\tilde{q}_\mu|\mu') = \Phi\left(\frac{\mu'-\mu}{\sigma}\right)\delta(\tilde{q}_\mu) + \begin{cases} \frac{1}{\sqrt{8\pi\tilde{q}_\mu}}\exp\left[-\frac{1}{2}\left(\sqrt{\tilde{q}_\mu} - \frac{\mu'-\mu}{\sigma}\right)^2\right] & 0 < \tilde{q}_\mu \leq \mu^2/\sigma^2 \\ \frac{1}{\sqrt{8\pi(\mu^2/\sigma^2)}}\exp\left[-\frac{1}{2}\frac{(\tilde{q}_\mu - (\mu^2 - 2\mu\mu')/\sigma^2)^2}{4(\mu^2/\sigma^2)}\right] & \tilde{q}_\mu > \mu^2/\sigma^2 \end{cases} \tag{10.12}$$

In practice, since both q_μ and $\tilde{q}_\mu$ rely on asymptotic approximations, there is virtually no difference between exclusion limits computed with either statistic.

When performing hypothesis testing for discovery or setting an exclusion limit on the value of the signal strength, it is possible that the signal + background model may be difficult to distinguish from the background-only model. This is particularly true in cases where the signal strength is very nearly zero. In these cases, the hypothesis test is said to have no *sensitivity* to the tested signal strength. Philosophically, it should not be possible to exclude a signal strength for which the model is not sensitive. This concept relates to the *power* of a hypothesis test, mentioned in Section 10.1.1. Rather than using the p-values of either

the background-only or signal + background hypotheses to make a decision, they may be combined using a technique called the CL_s *method* [133]–[135]. The ratio CL_s is constructed as shown in Equation 10.13:

$$CL_s(\mu) = \frac{p_\mu}{1 - p_b} \tag{10.13}$$

where $p_\mu = \int_{\tilde{q}_\mu^{obs}}^{\infty} f(\tilde{q}_\mu|\mu)d\tilde{q}_\mu$ and $p_b = \int_{\tilde{q}_\mu^{obs}}^{\infty} f(\tilde{q}_\mu|0)d\tilde{q}_\mu$ are the p-values for the signal + background and background-only hypotheses. Both hypotheses use the same test statistic, $\tilde{q}_\mu$, but with different assumptions of the signal strength for the sampling distributions. An example is shown in Figure 10.3:

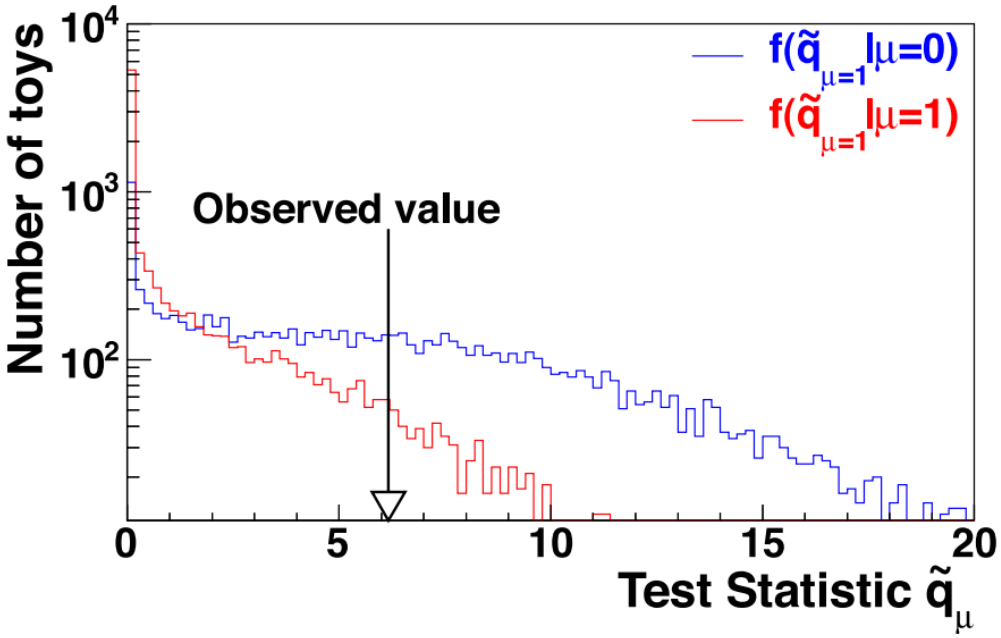


Figure 10.3: An example of two sampling distributions for $\tilde{q}_\mu$ assuming signal + background ($\mu = 1$) and background-only ($\mu = 0$) hypotheses. The p-values for each hypothesis equal the area under the respective curves to the right of the observed value [128].

If $CL_s(\mu = 1) \leq \alpha$, then the signal is excluded at $100(1 - \alpha)\%$ CL_s confidence level.[21] The $100(1 - \alpha)\%$ confidence level upper limit on μ may be found by solving $CL_s(\mu) = \alpha$ for μ. This comprises the exclusion limit on the signal strength.

[21]The CL_s confidence level is lower than the actual confidence level due to its construction. See [128] for details.

10.1.5 Expected Limits

Suppose that an excess of data events was observed over the predicted number of background events. It is natural to question whether or not the excess is due to a statistical fluctuation of data or if there is truly signal present. To address this question, it must be known whether or not the measurement is sensitive to the signal strength corresponding to the excess. The sensitivity of an experiment can be determined by knowing the *expected significance* to claim discovery and *expected exclusion limits* on the signal strength.

The sensitivity of an experiment may be assessed using a special dataset called an *Asimov dataset* [131], which has the property that all maximum likelihood estimators of all parameters yield their true values. The dataset can be constructed simply from the total predicted number of signal and background events. In other words, if $\mathbf{A}$ is the Asimov data histogram, then the number of entries in the i-th bin is $A_i(\mu, \boldsymbol{\theta}) = \mu s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta})$, where $s_i(\boldsymbol{\theta})$ and $b_i(\boldsymbol{\theta})$ are the predictions of the signal and background events, subject to systematic uncertainties represented as nuisance parameters $\boldsymbol{\theta}$. Notably, the Asimov dataset is constructed assuming no statistical uncertainties on the signal or background.

The test statistics q_0, q_μ, and $\tilde{q}_\mu$ and the ratio CL_s may be constructed using the Asimov dataset. The expected significance to claim discovery and the expected limits on signal strength may be found using the formulae and methods in Section 10.1.4. Since the Asimov dataset does not contain statistical variations, it is useful to know how the exclusion limits might change in response to statistical variations in the data. This is assessed by computing the expected limit for ± 1 and ± 2 standard deviations of the signal strength. These expected limits then form the ± 1 and ± 2 *error bands* on the expected limit.

While one procedure for calculating the standard deviation on the signal strength involves determining the covariance matrix, a computationally easier procedure is used in practice. A suitable approximation is $\sigma^2 = \frac{\mu^2}{q_\mu}$, where μ is the signal strength, and q_μ is computed using the Asimov dataset ($\tilde{q}_\mu$ may be used as well) [131].

10.2 The `HistFactory` Tool

When different measurements of a parameter of interest are made, it is useful to combine measurements to make a stronger statement about any exclusion limits or observed significances concerning a parameter of interest. Such combinations require a high degree of consistency between different measurements to ensure proper accounting of nuisance parameters. The ATLAS and CMS collaborations agreed to use a statistical framework called *RooStats* [136] to facilitate the combination of results in Higgs boson searches and measurements [128]. Under this framework, the distributions of signal and background events, their associated nuisance parameters, parameters of interest, and observed data are encoded in a formatted ROOT file called a *workspace*. The `HistFactory` tool [137] was developed to facilitate the construction of workspaces.

`HistFactory` forms likelihood functions from provided template histograms and information about systematics and model parameters, then stores them in the workspace. It is designed such that the program input has a modular structure to facilitate combinations of measurements. The basic structure in the `HistFactory` input is the *channel*, which matches a region of interest in an analysis (i.e. a signal or control region). Each channel is comprised of constituents called *samples*, which refer to either signal or background processes. Three pieces of information are provided by the user for each sample:

- A histogram with the estimated number of events in data assuming no systematic effects, called the *nominal distribution* or *template*.
- Any free parameters on the overall normalization of sample events, called `NormFactor` elements. These are typically used to indicate the parameter of interest in the signal samples but are also used in any background samples directly normalized to data.
- Information concerning the systematic uncertainties. `HistFactory` allows for two types of systematic uncertainties:
    1. The first type of systematic uncertainty is called `OverallSys`, for *overall systematic*, and is an uncertainty on the total number of sample events. The input consists of the systematic name and the relative variation ($N_{\text{sys}}/N_{\text{nominal}}$) of the number of sample events for $+1$ and -1σ variations of the systematic source. `HistFactory` allows for asymmetric systematic variations. For example, a systematic effect may result in a 7% variation downward of the predicted number of sample events for a $+1\sigma$ variation of the systematic and a 3% variation upward for a -1σ variation of the systematic.
    2. The second type of systematic uncertainty is called `HistoSys`, for *histogram systematic*, but is also known as a *shape systematic*. These systematics affect the distribution of events (or *shape*) within the nominal sample histogram while preserving the overall number of events. Like `OverallSys`, the required input consists of the systematic name and a pair of histograms with the altered nominal distribution corresponding to $+1\sigma$ and -1σ variations of the systematic source. These histograms are required to be normalized to their affiliated nominal sample histogram. The systematic shape variations do not have to be symmetric. `HistFactory` utilizes an interpolation algorithm for shifting the shapes between the systematic variations and the nominal distribution.

`HistFactory` offers these options to provide a high degree of flexibility in constructing the model, particularly with regard to the systematic uncertainties. Each name assigned to the `OverallSys` and `HistoSys` elements is represented in the model as a nuisance parameter. `OverallSys` and `HistoSys` elements can also have the same name, representing a source of systematic uncertainty having both overall and shape systematic effects. Corresponding elements for different samples in different channels can share the same name, indicating that

a nuisance parameter is *correlated* among these samples. Nuisance parameters are correlated among all samples and channels by default, unless there is some reason to decorrelate the parameter for certain samples and/or channels. Decorrelated nuisance parameters are varied separately in the fit.

Some care and attention to detail are necessary to format properly the input histograms. `HistFactory` requires that all histograms have a uniform bin size. Furthermore, the bin contents of each sample histogram must be positive to maintain the validity of the Poisson distribution.[22] In practice, each bin of the input histogram maps to a non-uniform range of values of the discriminating variable to ensure compliance with the requirements on uniform, positive sample bins while improving the sensitivity of the model.

`HistFactory` requires the configuration for each channel be specified in an *Extensible Markup Language* (XML) file indicating which samples are present in a channel and which `NormFactor`, `OverallSys`, and `HistoSys` elements apply to each sample. Furthermore, the nominal histograms for the samples and data, as well as histograms associated with the `HistoSys` elements, (all of which are stored in a ROOT file) are indicated in the configuration. The user also provides `HistFactory` with a top-level XML file indicating the channels used in the *measurement* and specifying the parameter of interest.

Once `HistFactory` has the required input, it creates likelihood functions that are stored in the workspace. There is much flexibility in how the likelihood functions may be constructed. For this dissertation, the likelihood functions have the form shown in Equation 10.14:

$$\mathcal{L}(\mu, \boldsymbol{\theta}, \gamma) = \prod_{c \in \text{channels}} \prod_{i \in \text{bins}} \text{Poisson}(n_{ci} | \mu s_{ci}(\boldsymbol{\theta}) + b_{ci}(\boldsymbol{\theta}, \gamma_{ci})) \text{Poisson}(\tau_{ci} | \gamma_{ci} \tau_{ci}) \prod_{\theta \in \{\boldsymbol{\theta}\}} \text{Gaussian}(\theta | \hat{\theta}) \tag{10.14}$$

where n_{ci}, s_{ci}, and b_{ci} are the number of data events in the i-th bin of the c-th channel, μ is the parameter of interest, and $\tau_{ci} = (b_{ci}^{\text{nom}}/\delta_{ci})^2$ is the expected number of background events. b_{ci}^{nom} is the nominal number of background events, and δ_{ci} is the associated statistical uncertainty. $\boldsymbol{\theta}$ is the total set of nuisance parameters. The number of signal and background events can be written as $s_{ci}(\boldsymbol{\theta}) = \prod_{j \in \text{samples}} \eta_{cij}(\boldsymbol{\theta}) \sigma_{cij}(\boldsymbol{\theta})$ and $b_{ci}(\boldsymbol{\theta}) = \prod_{j \in \text{samples}} \gamma_{ci} \phi_{cij}(\boldsymbol{\theta}) \eta_{cij}(\boldsymbol{\theta}) \sigma_{cij}(\boldsymbol{\theta})$, where γ_{ci} is the nuisance parameter for the statistical uncertainty, $\phi_{cij}(\boldsymbol{\theta})$ is the product of any `NormFactor` elements associated with the j-th sample, $\eta_{cij}(\boldsymbol{\theta})$ is the variation due to the `OverallSys` elements, and $\phi_{cij}(\boldsymbol{\theta})$ represents the bin contents of the nominal distributions along with the variations due to the `HistoSys` elements.

The likelihood function in Equation 10.14 is composed of two types of Poisson distributions and a Gaussian distribution. The first Poisson distribution evaluates the likelihood based on agreement between the data and predicted signal and background events. This function is undefined for negative values of the predicted number of events. `HistFactory` consequently requires every sample to have a positive bin count, though technically, only

[22] This will be explained later.

the sample sum needs to be positive. The second function evaluates the likelihood of the statistical nuisance parameters (centered at 1) having a particular value, given the statistical uncertainty in the total background. The Gaussian distribution is a unit variance distribution used to evaluate the likelihood of a given nuisance parameter θ having a particular value compared to the maximum likelihood estimator $\hat{\theta}$.[23] For an Asimov dataset, $\hat{\theta} = 0$ and $\theta = \pm 1$ represents $\pm 1\sigma$ variations of the systematic source. If the estimator $\hat{\theta} \neq 0$, then the nuisance parameter is said to be *pulled*.

10.3 The Structure of the Fit Model

Two separate fit models are constructed for analysis of the data in the muon and electron measurements using the `HistFactory` tool. Four channels consisting of regions of interest are used in each measurement: Signal Region 1 (SR1), Signal Region 2 (SR2), the W + jets control region (WCR), and the top control region (TCR). The $m_{\mu/e\tau}^{\text{MMC}}$ distribution is chosen as the discriminating variable. To satisfy `HistFactory` requirements on positive sample values and improve the sensitivity of the analysis, the binning scheme shown in Table 10.1 is utilized. The signal window is defined as 110 GeV $< m_{\mu/e\tau}^{\text{MMC}} <$ 150 GeV. The bins outside of this window are useful for constraining the values of the nuisance parameters associated with the background samples. By construction, the binning in SR1 and SR2 exploits shape

Region	N_{bins}	$m_{\mu/e\tau}^{\text{MMC}}$ Bin Boundaries [GeV]
SR1	12	50, 80, 90, 100, 110, 120, 130, 140, 150, 160, 180, 250, 2000
SR2	12	50, 60, 70, 80, 90, 100, 110, 120, 130, 145, 175, 250, 2000
WCR	1	50, 2000
TCR	1	50, 2000

Table 10.1: The binning scheme used in the fit model. The second column indicates the number of bins in each channel histogram. The third column lists the range of $m_{\mu/e\tau}^{\text{MMC}}$ values corresponding to a given bin.

differences between the signal and background distributions. The purpose of the TCR and WCR channels is to constrain the nuisance parameters associated with the top and W + jets backgrounds. A single bin is used due to the negligible signal present in these regions.

The following samples are used in the fit model:

[23] The specific functional form is $\text{Gaussian}(\theta|\hat{\theta}) = \frac{1}{\sqrt{2\pi\sigma_{\hat{\theta}}^2}} \exp\left(\frac{-(\theta-\hat{\theta})^2}{2\sigma_{\hat{\theta}}^2}\right)$, with the uncertainty on $\hat{\theta}$ being unity. However, this uncertainty might be different for the observed data if the nuisance parameter is not accurately estimated. If $\hat{\theta} < 1$, then the nuisance parameter is said to be *constrained*. If it is greater than one, then it is *underconstrained*.

- **LFV Higgs Boson Signal,** $H \to \mu\tau/e\tau$: Four different MC-simulated signal samples are used corresponding to the gluon–gluon fusion, VBF, and associated (ZH and WH) production modes. Each sample is normalized to its respective production cross section shown in Table 2.2, assuming a 1% branching ratio. $H \to \mu\tau$ ($H \to e\tau$) is the signal process used in the muon (electron) measurement.

- W **+ jets**, $Z \to \tau\tau$ **+ jets**, **Same-Sign Data**, and **Top**: These samples are modeled as described in Chapter 6.

- **SM** $H \to \tau\tau$: This background was not divided by production mode due to its tiny size ($< 1\%$ of the total background) and broad distribution.

- $Z \to ee(e \to \tau)$ **+ jets**: this background was only considered in the electron measurement. Because it is directly normalized to data in SR2, it is considered separately from $Z \to ee$ events where a jet is misidentified as a hadronic tau.

- **Other Backgrounds**: This background consists of the diboson (VV) and $Z \to \mu\mu$ ($Z \to ee(\text{jet} \to \tau)$) samples in the muon (electron) measurement.

The input histograms are checked for negative bins arising from the OS–SS background estimation technique to ensure compliance with `HistFactory` requirements that all sample bin contents be positive (i.e. larger than zero). If any negative bins are found in the nominal distributions, the absolute value of the bin contents is used instead.[24] In the case of the systematic variations, the bin contents are set to 0.05 (the minimum value allowed by `HistFactory`).

Many parameters are used in the fit model and are described in the remainder of the section. Numerous tests and checks of the fit model are also performed to ensure the quality of the results and are described in Appendix B.

10.3.1 Parameters of Interest

The parameter of interest is the signal strength of the LFV Higgs boson processes. The LFV Higgs boson signal sample predictions assume a branching ratio of 1%. Consequently, the signal strength is interpreted directly as the branching ratio of the LFV Higgs boson processes. Since there are two LFV processes under study, the muon and electron measurements use different parameters of interest.

10.3.2 Free Parameters

`NormFactor` elements are used with background samples which were directly normalized to data. They are listed in Table 10.2:

[24] Such cases usually occur in the tails of distributions and are small compared to the overall number of sample events.

`NormFactor` Element	Description
ATLAS_norm_LH12_ZttEmb	$Z \to \tau\tau$ sample normalization. Constrained by data in the Z boson mass window.
ATLAS_norm_LH12_WjSR1hi	Applied to W + jets sample only in SR1 channel. Constrained by data for $m_{\mu/e\tau}^{\mathrm{MMC}} > 150$ GeV.
ATLAS_norm_LH12_ZeeET	Normalization for $Z \to ee(e \to \tau)$ sample. Only present in electron measurement. Constrained by data in the Z boson mass window.

Table 10.2: The floating parameters (`NormFactor` elements) used in the fit model.

10.3.3 Nuisance Parameters

Over fifty nuisance parameters are incorporated into the fit model to account for the systematic uncertainties described in Chapter 9. A list is shown in Table 10.3.

Table 10.3: Nuisance parameters used in the fit model.

Nuisance Parameter	Description
ATLAS_EMB_ISOL	$Z \to \tau\tau$ embedding isolation uncertainty
ATLAS_EMB_MFS	$Z \to \tau\tau$ embedding muon cell energy subtraction uncertainty
ATLAS_rQCD	r_{QCD} uncertainty
ATLAS_Top_OS_k	Normalization uncertainty for OS top events
ATLAS_Top_SS_k	Normalization uncertainty for SS top events
ATLAS_Wj_OS_k	Normalization uncertainty for OS W + jets events
ATLAS_Wj_SS_k	Normalization uncertainty for SS W + jets events
ATLAS_WJ_SR1_SHAPE	Shape uncertainty due to W + jets reweighting in SR1
ATLAS_WJ_SR2_SR1_SHAPE	Shape uncertainty due to W + jets reweighting in SR2
ATLAS_TOP_EXTRAPOLATION	Extrapolation uncertainty for top background
ATLAS_WJ_EXTRAPOLATION	Extrapolation uncertainty for W + jets background
ATLAS_Z_SHAPE	Reweighting systematic for $Z \to \ell\ell$ (jet $\to \tau_{\mathrm{had}}^{fake}$) events
ATLAS_BTag_BEFF	Uncertainty on b-tagging efficiency (b-jets)
ATLAS_BTag_CEFF	Uncertainty on b-tagging efficiency (c-jets)
ATLAS_BTag_LEFF	Uncertainty on b-tagging efficiency (light jets)
ATLAS_EL_EFF	Electron identification efficiency uncertainty
ATLAS_EL_EFF_Emb	Electron identification efficiency uncertainty for $Z \to \tau\tau$ embedding events
ATLAS_EL_RES	Electron energy resolution uncertainty
ATLAS_EL_SCALE	Electron energy scale uncertainty
ATLAS_MU_EFF	Muon trigger and identification efficiency uncertainty
ATLAS_MU_EFF_Emb	Muon trigger and identification efficiency uncertainty for $Z \to \tau\tau$ embedding events
ATLAS_MU_SCALE	Muon momentum scale uncertainty
ATLAS_LUMI_2012	Uncertainty on measured integrated luminosity in 2012
ATLAS_TAU_ID_2012	Uncertainty for tau identification efficiency
ATLAS_TES_TOTAL_2012	Uncertainty on tau energy scale for misidentified hadronic taus
ATLAS_TES_INSITU_2012	Uncertainty on *in situ* tau energy scale
ATLAS_TES_TRUTH_2012	TES uncertainty component for true hadronic taus
ATLAS_JER_2012	Uncertainty on jet energy resolution
ATLAS_JES_2012_Detector1	JES uncertainty due to detector effects
ATLAS_JES_2012_Eta_StatMethod	JES uncertainty pertaining to statistical uncertainties in $\|\eta\|$-dependent scale factors
ATLAS_JES_2012_Modelling1	JES uncertainty pertaining to jet modeling in MC-simulated data

Table 10.3: Nuisance parameters used in the fit model. (continued)

Nuisance Parameter	Description
ATLAS_JES_2012_PileRho_TAU_GG	JES uncertainty pertaining to pile-up correction, applies to top and ggF Higgs boson processes
ATLAS_JES_2012_PileRho_TAU_QG	JES uncertainty pertaining to pile-up correction, applies to W + jets and Z + jets processes
ATLAS_JES_2012_PileRho_TAU_QQ	JES uncertainty pertaining to pile-up correction, applies to VBF Higgs boson, VV, WH, and ZH processes
ATLAS_JES_2012_Statistical1	JES uncertainty due to statistical uncertainties in MC-simulated data
ATLAS_JES_Eta_Modelling	JES uncertainty pertaining to $\|\eta\|$-dependence in scale factors
ATLAS_JES_FlavComp_TAU_G	JES uncertainty pertaining to flavor composition, applies to ggF Higgs boson, W + jets, and Z + jets proceses
ATLAS_JES_FlavComp_TAU_Q	JES uncertainty pertaining to flavor composition, applies to top, VBF Higgs, VV, WH, and ZH processes
ATLAS_JES_FlavResp	JES uncertainty pertaining to calorimeter response to different flavor jets
ATLAS_JES_Flavb	JES uncertainty pertaining to b-jets
ATLAS_JES_Mu	JES uncertainty due to the average number of proton–proton interactions
ATLAS_JES_NPV	JES uncertainty due to the number of primary vertices
ATLAS_MET_RESOSOFT	$E_{\mathrm{T}}^{\mathrm{miss}}$ resolution uncertainty on the soft term
ATLAS_MET_SCALESOFT	$E_{\mathrm{T}}^{\mathrm{miss}}$ scale uncertainty on the soft term
ATLAS_PU_rescaling	Pile-up reweighting uncertainty
ATLAS_BR_tautau	Uncertainty on $H \to \tau\tau$ branching ratio
QCDscale_V	QCD scale uncertainty for VV and $Z \to \mu\mu(\mu \to \tau)$ processes
QCDscale_VH	QCD scale uncertainty for associated Higgs boson production processes
QCDscale_ggH	QCD scale uncertainty for ggF Higgs boson production processes
QCDscale_qqH	QCD scale uncertainty for VBF Higgs boson production processes
pdf_Higgs_gg	PDF uncertainty for ggF Higgs boson production processes
pdf_Higgs_qq	PDF uncertainty on VBF Higgs boson production processes
pdf_Higgs_VH	PDF uncertainty for associated Higgs boson production processes
pdf_qq	PDF uncertainty for VV and $Z \to \mu\mu(\mu \to \tau)$ processes

The systematic variations for many of the nuisance parameters in Table 10.3 were computed for the ATLAS SM $H \to \tau\tau$ analysis and are able to be used for the LFV Higgs boson analysis. The systematic variations for the JER, extrapolation uncertainties, and embedding isolation uncertainty are inherently one-sided and are symmetrized to produce two-sided variations.[25] Statistical limitations in the MC samples lead to statistical fluctuations in the produced systematic variations. It is necessary to implement a procedure to determine whether a systematic effect is significant or simply the result of a statistical fluctuation, as the latter case can compromise the integrity of the fit. This procedure is called *pruning*, and is implemented in the following manner:

1. For a nuisance parameter to be implemented as an `OverallSys` element for a particular sample, the effect of the systematic variation must be larger than 0.5% of the number of nominal sample events.

2. To be considered as a `HistoSys` element, the significance $S_i = \frac{|u_i - d_i|}{\sigma_i}$, is computed for the i-th bin of the pair of histograms corresponding to the systematic effect, where

[25] Consider a nominal histogram $\mathbf{n}$ with a one-sided variation $\mathbf{v}$. The i-th bin in $\mathbf{n}$ and $\mathbf{v}$ may be related as $v_i = n_i + \sigma_i$, where σ_i is the difference in histogram bin contents. It is desired to have a symmetrized variation histogram $\mathbf{s}$, such that $s_i = n_i - \sigma_i$. It can be shown that such a histogram may be obtained by computing $\mathbf{s} = 2\mathbf{n} - \mathbf{v}$.

u_i and d_i are the bins from the histograms representing the upward and downward variations of the systematic source. The quantity σ_i represents the total signal or background statistical uncertainty, depending on whether or not the sample is a signal or background sample. If the quantity $S_i > 0.2$ for at least two bins with 100 or more nominal events, then the systematic variation is implemented as a `HistoSys` for the sample.

`HistoSys` nuisance parameters are only applied in the channels SR1 and SR2 since the WCR and TCR channels consist of only a single bin. To improve fit model stability further, the following additional considerations are applied to the implementation of nuisance parameters:

- The pruning procedure for shape systematics is not applied to the nuisance parameters ATLAS_rQCD, ATLAS_Top_OS_k, ATLAS_Top_SS_k, ATLAS_Wj_OS_k, and ATLAS_Wj_SS_k because these variations arise naturally through shape differences between OS and SS events in the respective samples. These variations are automatically considered as shape systematics in their respective samples.[26]

- Because of the method of background estimation (Section 6.2.1), the statistical uncertainty on the predicted W + jets background is larger relative to other processes due to the significant fraction of W + jets SS events. Thus, the systematic variations for W + jets are especially susceptible to statistical fluctuations. To mitigate these effects, the bins of each upward and downward systematic variation histogram are computed according to the formula in Equation 10.15:

$$N_{\text{sys},i} = \frac{N^{OS}_{\text{sys},i} + N^{SS}_{\text{sys},i}}{N^{OS}_{\text{nom},i} + N^{SS}_{\text{nom},i}} N^{OS-SS}_{\text{nom},i} \tag{10.15}$$

  where N_i refers to the number of events in the i-th bin for the systematic (sys) variations and nominal (nom) distribution.[27] The factor $N^{OS-SS}_{\text{nom},i}$ is the predicted number of W + jets events. This procedure is not applied for the ATLAS_rQCD, ATLAS_Wj_OS_k, and ATLAS_Wj_SS_k nuisance parameters.

- All nuisance parameters for the $H \to \tau\tau$ and signal associated production samples are considered as `OverallSys`-only due to the tiny size of these samples.

- The ATLAS_JES_2012_Detector1, ATLAS_JES_Flavb, ATLAS_JES_2012_Statistical1, ATLAS_JES_2012_Eta_StatMethod, ATLAS_JES_NPV, and ATLAS_JES_2012_PileRho_{QQ, QG, GG} uncertainties are treated as `OverallSys` because no shape variations are expected for these systematic variations.

[26] An exception is that ATLAS_rQCD is not applied as a shape systematic to same-sign data since no OS component exists for that sample.

[27] This procedure would be invalid if there were significant shape differences between OS and SS W + jets events, but these shape differences are found to be minimal.

- Due to the small size of the constituent backgrounds in the Other Background sample, all systematic uncertainties are treated as `OverallSys`-only except for ATLAS_rQCD, which is assigned a `HistoSys` uncertainty.

- Due to the broad nature and tiny size of the top background, all nuisance parameters are implemented as `OverallSys` with the exceptions of ATLAS_rQCD, ATLAS_Top_OS_k, and ATLAS_Top_SS_k, which are also implemented as `HistoSys` uncertainties.

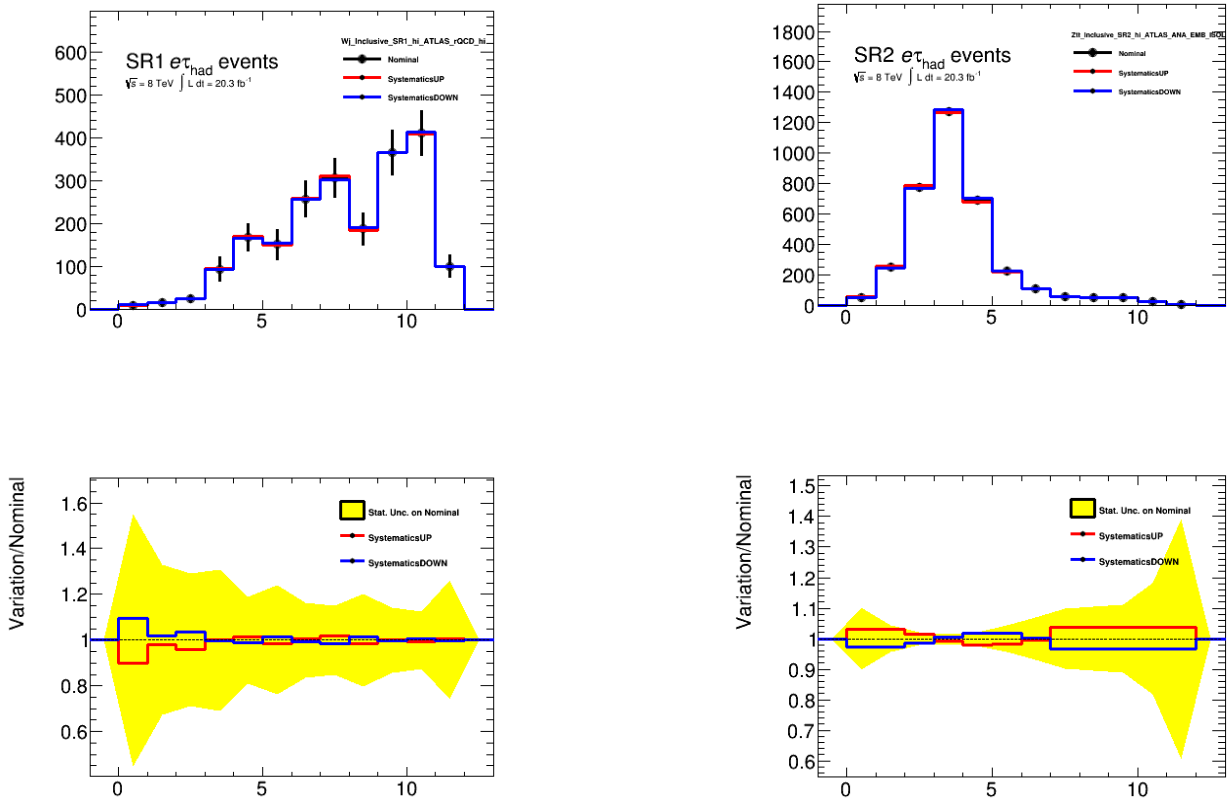
- The ATLAS_WJ_EXTRAPOLATION uncertainty is implemented as `OverallSys`-only since no statistically significant shape differences are found between ALPGEN and HERWIG W + jets samples.

- A final procedure called *smoothing* is applied to mitigate statistical fluctuations in the systematic variations of the samples. Each variation is divided by the nominal distribution to produce a ratio histogram. The ROOT method `TH1::Smooth(1)` is then applied, and the ratio histogram is multiplied by the nominal distribution to produce the smoothed systematic variation. In addition, the signal systematic distributions are "pre-treated" in the following manner:

  1. The tails of the signal distribution are identified as being those bins outside of the largest bin $\pm$ 2 bins.
  2. For each tail, the ratio $r = N(\text{sys})/N(\text{nom})$ is computed, where $N(\text{sys})$ and $N(\text{nom})$ are the total number of respective systematic and nominal events in the tail.
  3. The i-th bin contents of each tail are replaced by the quantity $r \cdot N_i(\text{nom})$, where $N_i(\text{nom})$ is the number of nominal signal events for that bin.

  This smoothing procedure is not applied to histograms associated with the ATLAS_rQCD, ATLAS_Top_OS_k, ATLAS_Top_SS_k, ATLAS_Wj_OS_k, and ATLAS_Wj_SS_k nuisance parameters.

Tables 10.4, 10.5, 10.6, and 10.7 list the types of systematic variations considered in the SR1 and SR2 channels for the muon and electron measurements. Two examples of systematic variations used in the fit model are provided in Figure 10.4.



(a) ATLAS_rQCD for W + jets in SR1

(b) ATLAS_EMB_ISOL for $Z \to \tau\tau$ in SR2

Figure 10.4: Examples of systematic variations used in the fit model, both taken from the electron measurement.

Table 10.4: Nuisance parameters used in the fit model for SR1 in the muon measurement following the described pruning procedure. The labels "O Sys", "HSys", and "OHSys" indicate that an `OverallSys` element, `HistoSys` element, or both, respectively, are implemented for the given nuisance parameter. The label "ATLAS_" is omitted for brevity.

Nuisance Parameter	$H \to \mu\tau$ ggF	$H \to \mu\tau$ VBF	$H \to \mu\tau$ WH	$H \to \mu\tau$ ZH	$H \to \tau\tau$	$Z \to \tau\tau$	W + jets	Top	Same-Sign Data	Other Bkgd
MU_SCALE	HSys	-	O Sys	-	-	-	-	-	-	O Sys
TES_TRUTH_2012	OHSys	O Sys	O Sys	O Sys	O Sys	O Sys	-	O Sys	-	O Sys
TES_SINGLEPART_2012	OHSys	O Sys	O Sys	O Sys	O Sys	O Sys	-	O Sys	-	O Sys
TES_TOTAL_2012	-	-	-	-	-	-	OHSys	O Sys	-	O Sys
JER_2012	OHSys	O Sys	O Sys	O Sys	O Sys	-	-	O Sys	-	O Sys
JES_Flavb	-	-	-	-	-	-	-	O Sys	-	-
JES_2012_Detector1	-	O Sys	O Sys	O Sys	-	-	O Sys	O Sys	-	O Sys
JES_2012_Eta_StatMethod	-	-	O Sys	-	-	-	-	O Sys	-	-
JES_Eta_Modelling	OHSys	OHSys	O Sys	O Sys	O Sys	-	-	O Sys	-	O Sys
JES_FlavComp_TAU_Q	-	OHSys	O Sys	O Sys	O Sys	-	-	O Sys	-	O Sys
JES_FlavComp_TAU_G	OHSys	-	-	-	O Sys	-	O Sys	-	-	O Sys
JES_FlavResp	OHSys	OHSys	O Sys	O Sys	O Sys	-	OHSys	O Sys	-	O Sys
JES_2012_Modelling1	OHSys	OHSys	O Sys	O Sys	O Sys	-	O Sys	O Sys	-	O Sys
JES_Mu	-	-	-	-	-	-	-	O Sys	-	O Sys
JES_NPV	-	-	-	-	O Sys	-	-	-	-	O Sys
JES_2012_PileRho_TAU_QQ	-	-	O Sys	O Sys	-	-	-	-	-	-
JES_2012_PileRho_TAU_QG	-	-	-	-	-	-	-	-	-	O Sys
JES_2012_PileRho_TAU_GG	-	-	-	-	O Sys	-	-	O Sys	-	-
JES_2012_Statistical1	-	-	-	-	-	-	-	O Sys	-	-
MET_RESOSOFT	HSys	-	-	O Sys	-	-	-	O Sys	-	O Sys
MET_SCALESOFT	-	-	-	-	O Sys	-	-	O Sys	-	O Sys
MU_EFF	O Sys	O Sys	O Sys	O Sys	O Sys	-	O Sys	O Sys	-	O Sys
MU_EFF_Emb	-	-	-	-	-	O Sys	-	-	-	-
PU_RESCALE_2012	OHSys	O Sys	O Sys	-	-	-	-	-	-	-
TAU_ID_2012	O Sys	O Sys	O Sys	O Sys	O Sys	O Sys	-	O Sys	-	O Sys
BTag_BEFF	-	-	-	-	-	-	-	O Sys	-	-
BTag_CEFF	-	-	O Sys	-	-	-	-	-	-	-
TOP_EXTRAPOLATION	-	-	-	-	-	-	-	O Sys	-	-
WJ_EXTRAPOLATION_hi	-	-	-	-	-	O Sys	O Sys	-	-	-
WJ_SR1_SHAPE	-	-	-	-	-	-	HSys	-	-	-
ANA_EMB_ISOL	-	-	-	-	-	O Sys	-	-	-	-
ANA_EMB_MFS	-	-	-	-	-	O Sys	-	-	-	-
BR_tautau	-	-	-	-	O Sys	-	-	-	-	-
pdf_Higgs_gg	O Sys	-	-	-	O Sys	-	-	-	-	-
QCDscale_ggH	O Sys	-	-	-	O Sys	-	-	-	-	-
pdf_Higgs_qq	-	O Sys	O Sys	O Sys	O Sys	-	-	-	-	-
QCDscale_qqH	-	O Sys	-	-	-	-	-	-	-	-

Table 10.4: Nuisance parameters used in the fit model for SR1 in the muon measurement following the described pruning procedure. The labels "O Sys", "HSys", and "OHSys" indicate that an `OverallSys` element, `HistoSys` element, or both, respectively, are implemented for the given nuisance parameter. The label "ATLAS_" is omitted for brevity. (continued)

Nuisance Parameter	$H \to \mu\tau$ ggF	$H \to \mu\tau$ VBF	$H \to \mu\tau$ WH	$H \to \mu\tau$ ZH	$H \to \tau\tau$	$Z \to \tau\tau$	W + jets	Top	Same-Sign Data	Other Bkgd
QCDscale_VH	-	-	O Sys	O Sys	-	-	-	-	-	-
pdf_qq	-	-	-	-	-	-	-	-	-	O Sys
QCDscale_VV	-	-	-	-	-	-	-	-	-	O Sys
rQCD_hi	-	-	-	-	-	-	OHSys	OHSys	O Sys	OHSys
LUMI_2012	O Sys	O Sys	O Sys	O Sys	O Sys	-	-	-	-	O Sys
Wj_OS_k	-	-	-	-	-	-	OHSys	-	-	-
Wj_SS_k	-	-	-	-	-	-	OHSys	-	-	-
Top_OS_k	-	-	-	-	-	-	-	OHSys	-	-
Top_SS_k	-	-	-	-	-	-	-	OHSys	-	-

Table 10.5: Nuisance parameters used in the fit model for SR2 in the muon measurement following the described pruning procedure. The labels "O Sys", "HSys", and "OHSys" indicate that an `OverallSys` element, `HistoSys` element, or both, respectively, are implemented for the given nuisance parameter. The label "ATLAS_" is omitted for brevity.

Nuisance Parameter	$H \to \mu\tau$ ggF	$H \to \mu\tau$ VBF	$H \to \mu\tau$ WH	$H \to \mu\tau$ ZH	$H \to \tau\tau$	$Z \to \tau\tau$	W + jets	Top	Same-Sign Data	Other Bkgd
MU_SCALE	HSys	-	O Sys	-	-	-	-	O Sys	-	O Sys
TES_TRUTH_2012	OHSys	O Sys	O Sys	O Sys	O Sys	OHSys	-	O Sys	-	O Sys
TES_SINGLEPART_2012	HSys	OHSys	O Sys	O Sys	O Sys	OHSys	-	O Sys	-	-
TES_TOTAL_2012	-	-	-	-	-	-	OHSys	O Sys	-	O Sys
JER_2012	OHSys	O Sys	O Sys	O Sys	O Sys	-	OHSys	O Sys	-	O Sys
JES_Flavb	-	-	-	-	-	-	-	O Sys	-	-
JES_2012_Detector1	-	-	O Sys	-	-	-	-	O Sys	-	O Sys
JES_2012_Eta_StatMethod	-	-	-	O Sys	-	-	-	O Sys	-	-
JES_Eta_Modelling	OHSys	OHSys	-	O Sys	O Sys	-	O Sys	O Sys	-	-
JES_FlavComp_TAU_Q	-	-	O Sys	O Sys	-	-	-	O Sys	-	O Sys
JES_FlavComp_TAU_G	HSys	-	-	-	O Sys	-	O Sys	-	-	O Sys
JES_FlavResp	HSys	-	O Sys	O Sys	O Sys	-	O Sys	O Sys	-	O Sys
JES_2012_Modelling1	HSys	-	O Sys	O Sys	O Sys	-	O Sys	O Sys	-	O Sys
JES_Mu	-	-	-	O Sys	-	-	-	O Sys	-	-
JES_NPV	-	-	O Sys	-	-	-	O Sys	O Sys	-	-
JES_2012_PileRho_TAU_QQ	-	-	O Sys	O Sys	-	-	-	-	-	O Sys
JES_2012_PileRho_TAU_QG	-	-	-	-	-	-	O Sys	-	-	O Sys

Table 10.5: Nuisance parameters used in the fit model for SR2 in the muon measurement following the described pruning procedure. The labels "O Sys", "HSys", and "OHSys" indicate that an `OverallSys` element, `HistoSys` element, or both, respectively, are implemented for the given nuisance parameter. The label "ATLAS_" is omitted for brevity. (continued)

Nuisance Parameter	$H \to \mu\tau$ ggF	$H \to \mu\tau$ VBF	$H \to \mu\tau$ WH	$H \to \mu\tau$ ZH	$H \to \tau\tau$	$Z \to \tau\tau$	W + jets	Top	Same-Sign Data	Other Bkgd
JES_2012_PileRho_TAU_GG	-	-	-	-	-	-	-	O Sys	-	-
JES_2012_Statistical1	-	-	O Sys	-	-	-	-	O Sys	-	O Sys
MET_RESOSOFT	O Sys	-	-	-	-	-	O Sys	O Sys	-	O Sys
MET_SCALESOFT	-	-	O Sys	-	-	-	-	O Sys	-	O Sys
MU_EFF	O Sys	O Sys	O Sys	O Sys	O Sys	-	O Sys	O Sys	-	O Sys
MU_EFF_Emb	-	-	-	-	-	O Sys	-	-	-	-
PU_RESCALE_2012	OHSys	O Sys	O Sys	-	O Sys	-	-	-	-	O Sys
TAU_ID_2012	O Sys	O Sys	O Sys	O Sys	O Sys	O Sys	-	O Sys	-	O Sys
BTag_BEFF	-	-	-	-	-	-	-	O Sys	-	-
BTag_CEFF	-	-	O Sys	O Sys	-	-	-	-	-	-
TOP_EXTRAPOLATION	-	-	-	-	-	-	-	O Sys	-	-
WJ_EXTRAPOLATION_hi	-	-	-	-	-	O Sys	O Sys	-	-	-
WJ_SR2_SR1_SHAPE	-	-	-	-	-	-	HSys	-	-	-
ANA_EMB_ISOL	-	-	-	-	-	OHSys	-	-	-	-
ANA_EMB_MFS	-	-	-	-	-	OHSys	-	-	-	-
BR_tautau	-	-	-	-	O Sys	-	-	-	-	-
pdf_Higgs_gg	O Sys	-	-	-	O Sys	-	-	-	-	-
QCDscale_ggH	O Sys	-	-	-	O Sys	-	-	-	-	-
pdf_Higgs_qq	-	O Sys	O Sys	O Sys	-	-	-	-	-	-
QCDscale_qqH	-	O Sys	-	-	-	-	-	-	-	-
QCDscale_VH	-	-	O Sys	O Sys	-	-	-	-	-	-
QCDscale_V	-	-	-	-	-	-	-	-	-	O Sys
pdf_qq	-	-	-	-	-	-	-	-	-	O Sys
QCDscale_VV	-	-	-	-	-	-	-	-	-	O Sys
rQCD_hi	-	-	-	-	-	-	OHSys	OHSys	O Sys	OHSys
LUMI_2012	O Sys	O Sys	O Sys	O Sys	O Sys	-	-	-	-	O Sys
Wj_OS_k	-	-	-	-	-	-	OHSys	-	-	-
Wj_SS_k	-	-	-	-	-	-	OHSys	-	-	-
Top_OS_k	-	-	-	-	-	-	-	OHSys	-	-
Top_SS_k	-	-	-	-	-	-	-	OHSys	-	-

Table 10.6: Nuisance parameters used in the fit model for SR1 in the electron measurement following the described pruning procedure. The labels "O Sys", "HSys", and "OHSys" indicate that an `OverallSys` element, `HistoSys` element, or both, respectively, are implemented for the given nuisance parameter. The label "ATLAS_" is omitted for brevity.

Nuisance Parameter	$H \to e\tau$ ggF	$H \to e\tau$ VBF	$H \to e\tau$ WH	$H \to e\tau$ ZH	$H \to \tau\tau$	$Z \to \tau\tau$	$Z \to ee$ $(e \to \tau)$	W + jets	Top	Same-Sign Data	Other Bkgd
EL_SCALE	O Sys	O Sys	O Sys	O Sys	O Sys	O Sys	O Sys	O Sys	O Sys	-	O Sys
EL_RES	HSys	O Sys	O Sys	O Sys	O Sys	O Sys	O Sys	HSys	O Sys	-	O Sys
TES_TRUTH_2012	OHSys	O Sys	O Sys	O Sys	O Sys	O Sys	-	-	O Sys	-	O Sys
TES_SINGLEPART_2012	OHSys	O Sys	O Sys	O Sys	O Sys	O Sys	-	-	O Sys	-	O Sys
TES_TOTAL_2012	-	-	-	-	-	-	-	O Sys	O Sys	-	O Sys
TES_TOTAL_2012_El	-	-	-	-	-	-	O Sys	-	-	-	-
JER_2012	OHSys	OHSys	O Sys	-	O Sys	-	O Sys	HSys	O Sys	-	O Sys
JES_Flavb	-	-	-	-	-	-	-	-	O Sys	-	-
JES_2012_Detector1	O Sys	O Sys	O Sys	O Sys	O Sys	-	O Sys	-	O Sys	-	O Sys
JES_2012_Eta_StatMethod	-	-	O Sys	-	O Sys	-	O Sys	-	O Sys	-	O Sys
JES_Eta_Modelling	OHSys	OHSys	O Sys	O Sys	O Sys	-	O Sys	HSys	O Sys	-	O Sys
JES_FlavComp_TAU_Q	-	OHSys	O Sys	O Sys	O Sys	-	-	-	O Sys	-	O Sys
JES_FlavComp_TAU_G	OHSys	-	-	-	O Sys	-	O Sys	O Sys	-	-	O Sys
JES_FlavResp	OHSys	OHSys	O Sys	O Sys	O Sys	-	O Sys	-	O Sys	-	O Sys
JES_2012_Modelling1	OHSys	OHSys	O Sys	O Sys	O Sys	-	O Sys	-	O Sys	-	O Sys
JES_Mu	-	-	O Sys	O Sys	-	-	O Sys	-	O Sys	-	-
JES_NPV	-	-	O Sys	-	O Sys	-	O Sys	-	O Sys	-	O Sys
JES_2012_PileRho_TAU_QQ	-	-	O Sys	-	-	-	-	-	-	-	O Sys
JES_2012_PileRho_TAU_QG	-	-	-	-	-	-	O Sys	-	-	-	O Sys
JES_2012_PileRho_TAU_GG	O Sys	-	-	-	O Sys	-	-	-	O Sys	-	-
JES_2012_Statistical1	-	-	O Sys	-	-	-	O Sys	-	O Sys	-	O Sys
MET_RESOSOFT	-	-	-	-	O Sys	-	O Sys	-	-	-	O Sys
MET_SCALESOFT	-	-	-	-	-	-	O Sys	-	O Sys	-	O Sys
EL_EFF	O Sys	O Sys	O Sys	O Sys	O Sys	-	O Sys	O Sys	O Sys	-	O Sys
EL_EFF_Emb	-	-	-	-	-	O Sys	-	-	-	-	-
PU_RESCALE_2012	OHSys	O Sys	O Sys	O Sys	-	-	-	-	-	-	-
TAU_ID_2012	O Sys	O Sys	O Sys	O Sys	O Sys	O Sys	-	-	O Sys	-	O Sys
BTag_BEFF	-	-	-	-	-	-	-	-	O Sys	-	-
BTag_CEFF	-	-	O Sys	-	-	-	-	-	-	-	-
TOP_EXTRAPOLATION	-	-	-	-	-	-	-	-	O Sys	-	-
WJ_EXTRAPOLATION_hi	-	-	-	-	-	O Sys	-	O Sys	-	-	-
WJ_SR1_SHAPE	-	-	-	-	-	-	-	HSys	-	-	-
ANA_EMB_ISOL	-	-	-	-	-	O Sys	-	-	-	-	-
ANA_EMB_MFS	-	-	-	-	-	O Sys	-	-	-	-	-
BR_tautau	-	-	-	-	O Sys	-	-	-	-	-	-
pdf_Higgs_gg	O Sys	-	-	-	O Sys	-	-	-	-	-	-
QCDscale_ggH	O Sys	-	-	-	O Sys	-	-	-	-	-	-

Table 10.6: Nuisance parameters used in the fit model for SR1 in the electron measurement following the described pruning procedure. The labels "O Sys", "HSys", and "OHSys" indicate that an `OverallSys` element, `HistoSys` element, or both, respectively, are implemented for the given nuisance parameter. The label "ATLAS_" is omitted for brevity. (continued)

Nuisance Parameter	$H \to e\tau$ ggF	$H \to e\tau$ VBF	$H \to e\tau$ WH	$H \to e\tau$ ZH	$H \to \tau\tau$	$Z \to \tau\tau$	$Z \to ee$ $(e \to \tau)$	W + jets	Top	Same-Sign Data	Other Bkgd
pdf_Higgs_qq	-	O Sys	-	-	-	-	-	-	-	-	-
QCDscale_qqH	-	O Sys	-	-	-	-	-	-	-	-	-
pdf_Higgs_VH	-	-	O Sys	O Sys	-	-	-	-	-	-	-
QCDscale_VH	-	-	O Sys	O Sys	-	-	-	-	-	-	-
pdf_qq	-	-	-	-	-	-	-	-	-	-	O Sys
QCDscale_VV	-	-	-	-	-	-	-	-	-	-	O Sys
rQCD_hi	-	-	-	-	-	-	OHSys	OHSys	OHSys	O Sys	OHSys
LUMI_2012	O Sys	O Sys	O Sys	O Sys	O Sys	-	-	-	-	-	O Sys
Zee_k	-	-	-	-	-	-	-	-	-	-	O Sys
Wj_OS_k	-	-	-	-	-	-	-	OHSys	-	-	-
Wj_SS_k	-	-	-	-	-	-	-	OHSys	-	-	-
Top_OS_k	-	-	-	-	-	-	-	-	OHSys	-	-
Top_SS_k	-	-	-	-	-	-	-	-	OHSys	-	-

Table 10.7: Nuisance parameters used in the fit model for SR2 in the electron measurement following the described pruning procedure. The labels "O Sys", "HSys", and "OHSys" indicate that an `OverallSys` element, `HistoSys` element, or both, respectively, are implemented for the given nuisance parameter. The label "ATLAS_" is omitted for brevity.

Nuisance Parameter	$H \to e\tau$ ggF	$H \to e\tau$ VBF	$H \to e\tau$ WH	$H \to \tau$ ZH	$H \to \tau\tau$	$Z \to \tau\tau$	$Z \to ee$ $(e \to \tau)$	W + jets	Top	Same-Sign Data	Other Bkgd
EL_SCALE	O Sys	O Sys	O Sys	O Sys	O Sys	O Sys	O Sys	O Sys	O Sys	-	O Sys
EL_RES	OHSys	OHSys	O Sys	O Sys	O Sys	-	OHSys	OHSys	O Sys	-	O Sys
TES_TRUTH_2012	OHSys	OHSys	O Sys	O Sys	O Sys	OHSys	-	O Sys	O Sys	-	O Sys
TES_SINGLEPART_2012	HSys	OHSys	O Sys	-	O Sys	OHSys	-	-	-	-	O Sys
TES_TOTAL_2012	-	-	-	-	-	-	-	OHSys	O Sys	-	O Sys
TES_TOTAL_2012_El	-	-	-	-	-	-	OHSys	-	-	-	-
JER_2012	HSys	O Sys	O Sys	O Sys	O Sys	-	O Sys	OHSys	O Sys	-	O Sys
JES_Flavb	-	-	-	-	-	-	-	-	O Sys	-	-
JES_2012_Detector1	-	O Sys	O Sys	O Sys	-	-	-	O Sys	O Sys	-	O Sys
JES_2012_Eta_StatMethod	-	-	O Sys	-	-	-	-	O Sys	O Sys	-	O Sys
JES_Eta_Modelling	HSys	OHSys	O Sys	-	O Sys	-	-	O Sys	O Sys	-	O Sys
JES_FlavComp_TAU_Q	-	OHSys	O Sys	-	-	-	-	-	O Sys	-	O Sys
JES_FlavComp_TAU_G	OHSys	-	-	-	O Sys	-	O Sys	O Sys	-	-	O Sys

Table 10.7: Nuisance parameters used in the fit model for SR2 in the electron measurement following the described pruning procedure. The labels "O Sys", "HSys", and "OHSys" indicate that an `OverallSys` element, `HistoSys` element, or both, respectively, are implemented for the given nuisance parameter. The label "ATLAS_" is omitted for brevity. (continued)

Nuisance Parameter	$H \to e\tau$ ggF	$H \to e\tau$ VBF	$H \to e\tau$ WH	$H \to \tau$ ZH	$H \to \tau\tau$	$Z \to \tau\tau$	$Z \to ee$ $(e \to \tau)$	W + jets	Top	Same-Sign Data	Other Bkgd
JES_FlavResp	OHSys	-	-	-	O Sys	-	-	O Sys	O Sys	-	O Sys
JES_2012_Modelling1	OHSys	OHSys	O Sys	O Sys	O Sys	-	-	O Sys	O Sys	-	O Sys
JES_Mu	-	-	-	-	-	-	-	O Sys	O Sys	-	O Sys
JES_NPV	-	-	O Sys	O Sys	-	-	O Sys	O Sys	O Sys	-	O Sys
JES_2012_PileRho_TAU_QG	-	-	-	-	-	-	-	O Sys	-	-	O Sys
JES_2012_PileRho_TAU_GG	-	-	-	-	O Sys	-	-	-	-	-	-
JES_2012_Statistical1	-	-	O Sys	-	-	-	-	O Sys	O Sys	-	O Sys
MET_RESOSOFT	-	-	O Sys	-	-	-	O Sys	O Sys	O Sys	-	O Sys
MET_SCALESOFT	-	-	O Sys	-	-	-	O Sys	O Sys	-	-	O Sys
EL_EFF	O Sys	O Sys	O Sys	O Sys	O Sys	-	O Sys	O Sys	O Sys	-	-
EL_EFF_Emb	-	-	-	-	-	O Sys	-	-	-	-	-
PU_RESCALE_2012	OHSys	O Sys	O Sys	O Sys	O Sys	-	-	-	-	-	-
TAU_ID_2012	O Sys	O Sys	O Sys	O Sys	O Sys	O Sys	-	-	O Sys	-	O Sys
BTag_BEFF	-	-	-	-	-	-	-	-	O Sys	-	-
BTag_CEFF	-	-	O Sys	-	-	-	-	-	-	-	-
Z_SHAPE	-	-	-	-	-	-	-	-	-	-	O Sys
TOP_EXTRAPOLATION	-	-	-	-	-	-	-	-	O Sys	-	-
WJ_EXTRAPOLATION_hi	-	-	-	-	-	O Sys	-	O Sys	-	-	-
WJ_SR2_SR1_SHAPE	-	-	-	-	-	-	-	HSys	-	-	-
ANA_EMB_ISOL	-	-	-	-	-	OHSys	-	-	-	-	-
ANA_EMB_MFS	-	-	-	-	-	OHSys	-	-	-	-	-
BR_tautau	-	-	-	-	O Sys	-	-	-	-	-	-
pdf_Higgs_gg	O Sys	-	-	-	O Sys	-	-	-	-	-	-
QCDscale_ggH	O Sys	-	-	-	O Sys	-	-	-	-	-	-
pdf_Higgs_qq	-	O Sys	-	-	-	-	-	-	-	-	-
QCDscale_qqH	-	O Sys	-	-	-	-	-	-	-	-	-
pdf_Higgs_VH	-	-	O Sys	O Sys	-	-	-	-	-	-	-
QCDscale_VH	-	-	O Sys	O Sys	-	-	-	-	-	-	-
pdf_qq	-	-	-	-	-	-	-	-	-	-	O Sys
QCDscale_VV	-	-	-	-	-	-	-	-	-	-	O Sys
rQCD_hi	-	-	-	-	-	-	OHSys	OHSys	OHSys	O Sys	OHSys
LUMI_2012	O Sys	O Sys	O Sys	O Sys	O Sys	-	-	-	-	-	O Sys
Zee_k	-	-	-	-	-	-	-	-	-	-	O Sys
Wj_OS_k	-	-	-	-	-	-	-	OHSys	-	-	-
Wj_SS_k	-	-	-	-	-	-	-	OHSys	-	-	-
Top_OS_k	-	-	-	-	-	-	-	-	OHSys	-	-
Top_SS_k	-	-	-	-	-	-	-	-	OHSys	-	-

Chapter 11

Results

Upon validation of the fit model, the data in the signal window are unblinded, and the binned likelihood fit described in the previous chapter is performed. No signal is observed. Consequently, expected and observed limits on the branching ratios $\mathrm{Br}(H \to \mu\tau)$ and $\mathrm{Br}(H \to e\tau)$ are computed. This chapter presents these results as well as a combined result with the complementary LFV analysis using leptonic taus, described in Section 2.2.2.1.

11.1 Muon and Electron Measurements

In the muon measurement, the expected upper limit at 95% confidence on $\mathrm{Br}(H \to \mu\tau)$ is found to be $1.24^{+0.50}_{-0.35}\%$. The observed upper limit is found to be 1.85%, corresponding to a 1.3σ excess of data events over the predicted number of background events. The best-fit value of $\mathrm{Br}(H \to \mu\tau)$[1] is found to be $0.77 \pm 0.62\%$. Two fits are also performed with SR1, SR2, WCR, and TCR with the signal strength parameters decorrelated in SR1 and SR2 to determine the relative contribution of each signal region to the result. All results may be found in Table 11.3.

The same fits are performed in the electron measurement. The expected upper limit at 95% confidence on $\mathrm{Br}(H \to e\tau)$ is found to be $2.07^{+1.06}_{-0.58}\%$. The observed upper limit is found to be 1.81%, indicating a deficit of data events compared to the predicted number of background events. The best-fit value of $\mathrm{Br}(H \to e\tau)$ is found to be $-0.47^{+1.08}_{-1.18}\%$. These results, along with those for decorrelated signal strength parameters in SR1 and SR2, may be found in Table 11.3.

Figures 11.1 and 11.3 show the background predictions after the fit is performed in SR1 and SR2, respectively for the muon and electron measurements.[2] Excellent agreement between the data and predicted background is observed in all figures. The data and background

[1] This is $\hat{\mu}$, discussed in the previous chapter.

[2] The fit adjusts the nuisance parameters to their best-fit values, increasing or reducing the number of background events, accordingly.

predictions in SR1 and SR2 are combined in Figures 11.2 and 11.4 for the muon and electron measurements, respectively.

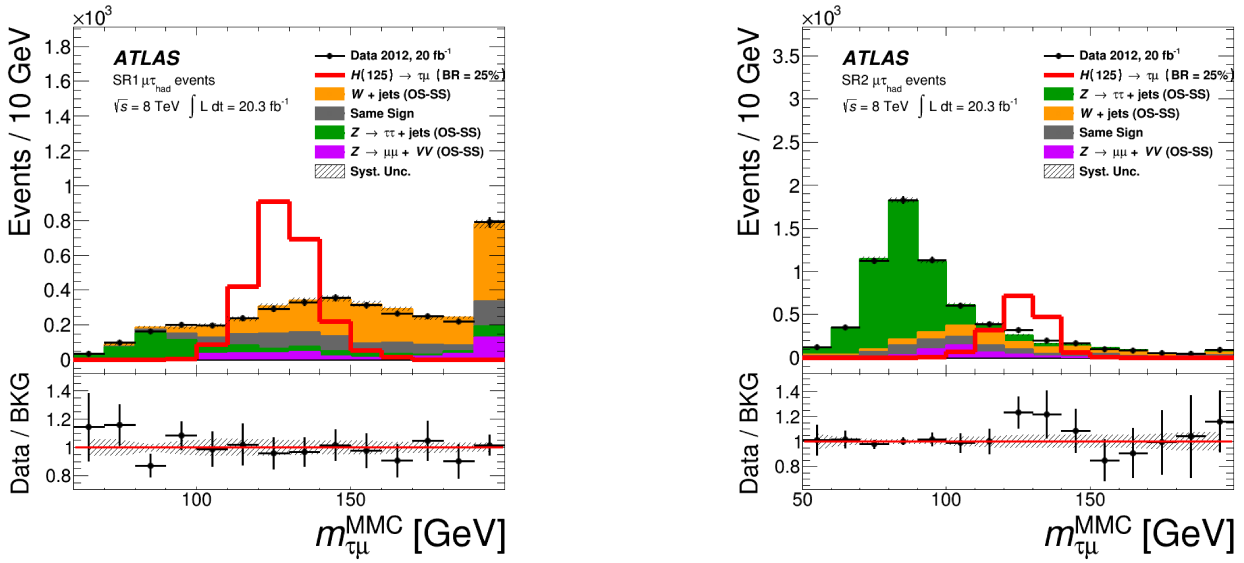


Figure 11.1: Post-fit $m_{\mu\tau}^{\text{MMC}}$ distributions in SR1 and SR2 in the muon measurement. The red histogram indicates the predicted $H \to \mu\tau$ distribution, assuming $\text{Br}(H \to \mu\tau) = 25\%$ [115].

Post-fit event yields with uncertainties in the signal mass region are computed for SR1 and SR2. The muon measurement yields are shown in Table 11.1, and the electron measurement yields are shown in Table 11.2.

11.2 Combination with the Leptonic Tau Analysis

The results from the search for LFV Higgs boson decays with leptonic taus [34], described in Section 2.2.2.1, are combined with the results in the previous section to set more stringent expected and observed limits in the muon and electron measurements. The leptonic tau channels are two signal regions defined respectively by the presence or lack of jets with $|\eta| < 2.4$. Only the highest- and second-highest-p_{T} leptons are considered, each required to have $p_{\text{T}} > 35$ GeV and 12 GeV and have $|\eta| < 2.4$. Additional requirements are placed on the $\Delta\phi$ between each combination of leptons and $E_{\text{T}}^{\text{miss}}$. Requirements on the p_{T} difference between the leptons are also imposed. The dominant background in the analysis is non-prompt leptons occurring from misidentified jets or from secondary decays of hadrons. A binned likelihood fit using collinear mass as the discriminating variable is performed to determine the LFV signal strength.

In the muon measurement, the leptonic tau search determined an expected upper limit on $\text{Br}(H \to \tau\mu)$ at 95% confidence of $1.73^{+0.74}_{-0.49}\%$ with an observed limit of 1.79%, corresponding

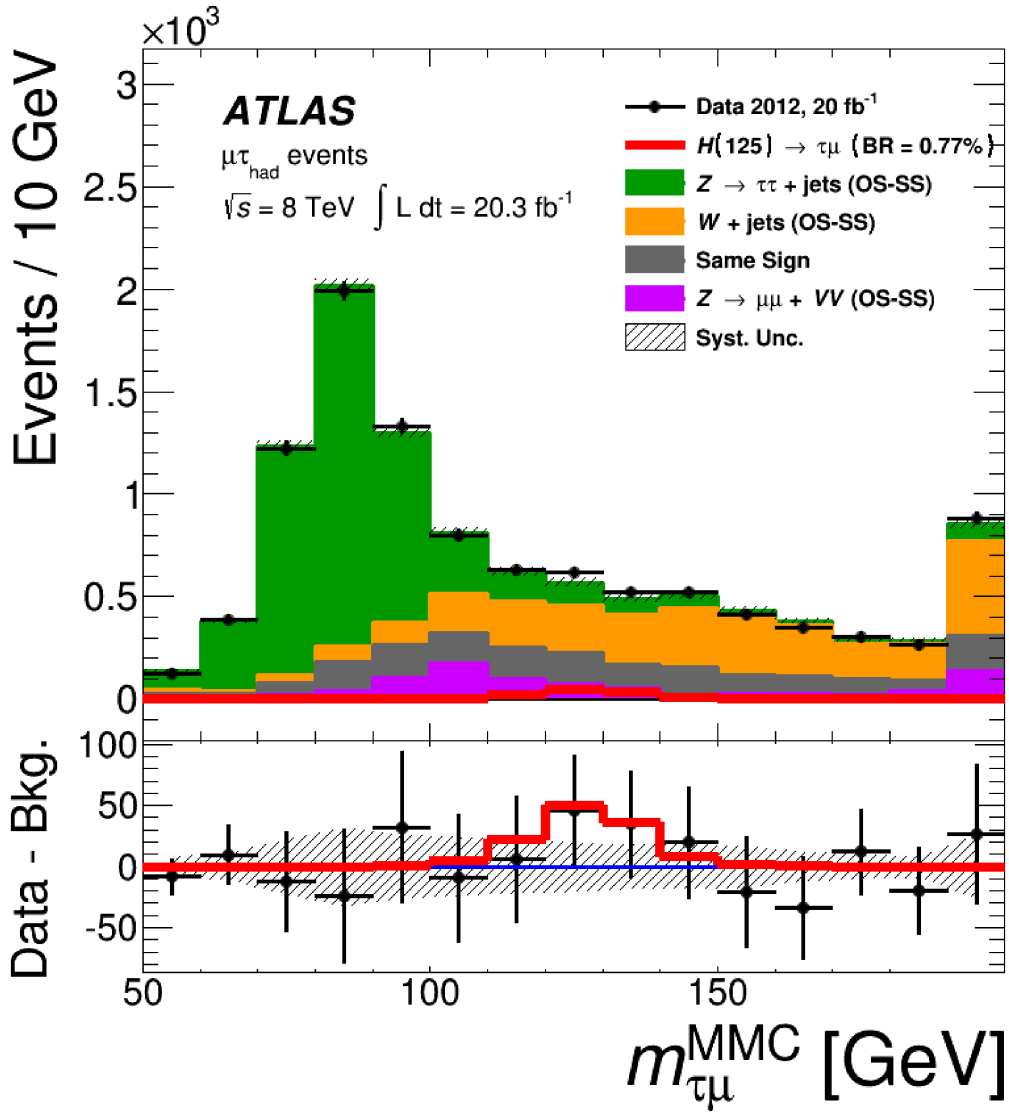


Figure 11.2: Combined Post-fit $m_{\mu\tau}^{\mathrm{MMC}}$ distribution for the muon measurement. The red histogram indicates the predicted $H \to \mu\tau$ distribution for the best-fit $\mathrm{Br}(H \to \mu\tau) = 0.77\%$ [115].

to a best-fit value of the branching ratio equal to $0.03^{+0.88}_{-0.86}\%$. An expected limit on $\mathrm{Br}(H \to e\tau)$ was found to be $1.48^{+0.60}_{-0.42}\%$ with an observed limit of 1.36%, corresponding to a best-fit value of the branching ratio equal to $-0.26^{+0.79}_{-0.82}\%$.

The workspaces for the hadronic and leptonic tau channels are merged together and refit to obtain the combined limits. The combined expected limit on $\mathrm{Br}(H \to \mu\tau)$ is $1.01^{+0.40}_{-0.29}\%$ with an observed limit of 1.43%, corresponding to a best-fit value of the branching ratio equal to $0.53^{+0.51}_{-0.51}\%$. The combined expected limit on $\mathrm{Br}(H \to e\tau)$ is $1.21^{+0.49}_{-0.34}\%$ with an observed limit of 1.04%, corresponding to a best-fit value of the branching ratio equal to

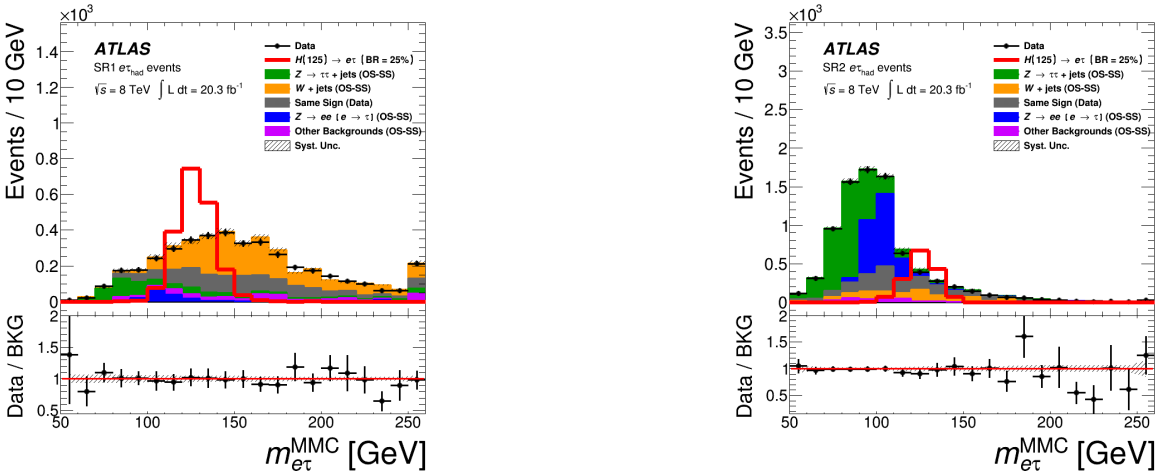


Figure 11.3: Post-fit $m_{e\tau}^{\text{MMC}}$ distributions in SR1 and SR2 for the electron measurement. The red histogram indicates the predicted $H \to e\tau$ distribution, assuming $\text{Br}(H \to e\tau) = 25\%$ [34].

$-0.34^{+0.64}_{-0.66}\%$. All results are shown in Table 11.3.

From these measurements, exclusion plots are constructed highlighting the LFV branching ratios inconsistent with observed data. These plots are shown in Figures 11.5 and 11.6 and include the expected and observed limits associated with each channel to highlight the relative contribution of the channel to the result.

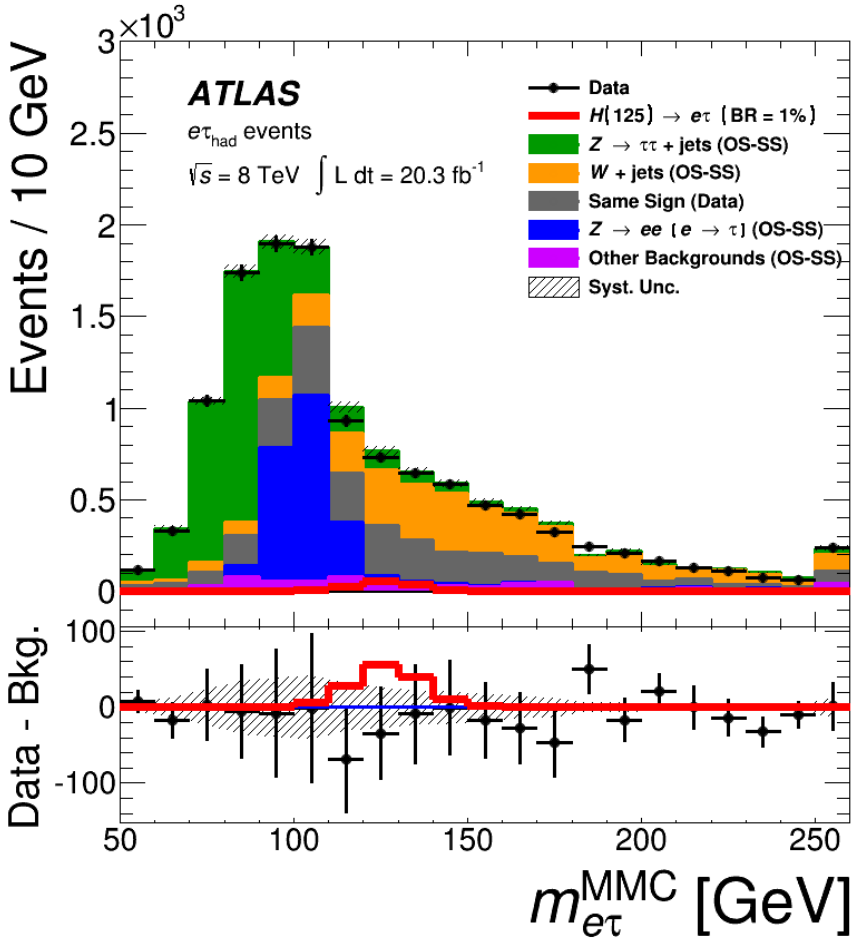


Figure 11.4: Combined Post-fit $m_{e\tau}^{\mathrm{MMC}}$ distribution for the electron measurement. The red histogram indicates the predicted $H \to e\tau$ distribution, assuming $\mathrm{Br}(H \to e\tau) = 1\%$ [34].

Table 11.1: Events yields in the 110 GeV $< m_{\mu\tau}^{\mathrm{MMC}} <$ 150 GeV region in the muon measurement [115]. The signal yields assume $\mathrm{Br}(H \to \mu\tau) = 0.77\%$. Event yields for the background processes are obtained after a combined fit to SR1, SR2, WCR, and TCR. The uncertainties on each process are respectively statistical and systematic. The systematic uncertainty on the total background prediction accounts for all correlations between nuisance parameters.

	SR1	SR2
Signal	$69.1 \pm 0.8 \pm 9.2$	$48.5 \pm 0.8 \pm 7.5$
$Z \to \tau\tau$	$133.4 \pm 6.9 \pm 9.1$	$262.6 \pm 9.7 \pm 18.6$
W+jets	$619 \pm 54 \pm 55$	$406 \pm 42 \pm 34$
Top	$39.5 \pm 5.3 \pm 4.7$	$19.6 \pm 3.1 \pm 3.3$
Same–Sign events	$335 \pm 19 \pm 47$	$238 \pm 16 \pm 34$
$VV + Z \to \mu\mu$	$90 \pm 21 \pm 16$	$81 \pm 22 \pm 17$
$H \to \tau\tau$	$6.82 \pm 0.21 \pm 0.97$	$13.7 \pm 0.3 \pm 1.9$
Total background	$1224 \pm 62 \pm 63$	$1021 \pm 51 \pm 49$
Data	1217	1075

Table 11.2: Events yields in the 110 GeV $< m_{e\tau}^{\mathrm{MMC}} <$ 150 GeV region in the electron measurement. The signal yields assume $\mathrm{Br}(H \to e\tau) = 1.0\%$. Event yields for the background processes are obtained after a combined fit to SR1, SR2, WCR, and TCR. The uncertainties on each process are respectively statistical and systematic. The systematic uncertainty on the total background prediction accounts for all correlations between nuisance parameters.

	SR1	SR2
LFV signal ($\mathrm{Br}(H \to e\tau) = 1.0\%$)	$75 \pm 1 \pm 8$	$59 \pm 1 \pm 8$
W+jets	$740 \pm 80 \pm 110$	$370 \pm 60 \pm 70$
Same-Sign events	$390 \pm 20 \pm 60$	$570 \pm 30 \pm 80$
$Z \to \tau\tau$	$116 \pm 8 \pm 11$	$245 \pm 11 \pm 20$
$VV + Z \to ee(jet \to \tau_{\mathrm{had}}^{\mathrm{misid}})$	$71 \pm 31 \pm 30$	$60 \pm 20 \pm 40$
$Z \to ee(e \to \tau_{\mathrm{had}}^{\mathrm{misid}})$	$69 \pm 17 \pm 11$	$320 \pm 40 \pm 40$
Top	$18 \pm 5 \pm 4$	$10.2 \pm 2.6 \pm 2.2$
$H \to \tau\tau$	$4.6 \pm 0.2 \pm 0.7$	$10.5 \pm 0.3 \pm 1.5$
Total background	$1410 \pm 90 \pm 70$	$1590 \pm 80 \pm 70$
Data	1397	1501

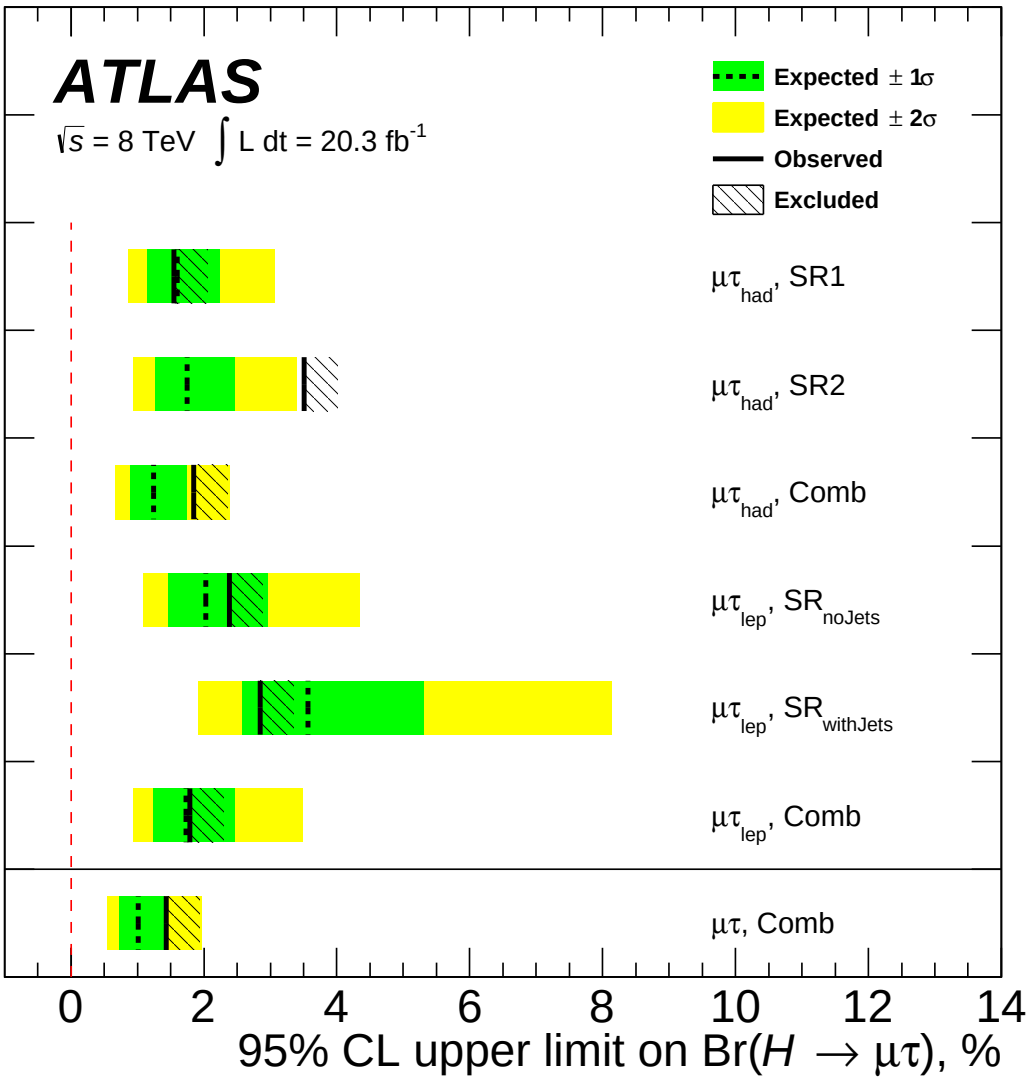


Figure 11.5: Expected and Observed Limits in the muon measurement for both hadronic and leptonic tau analyses [34].

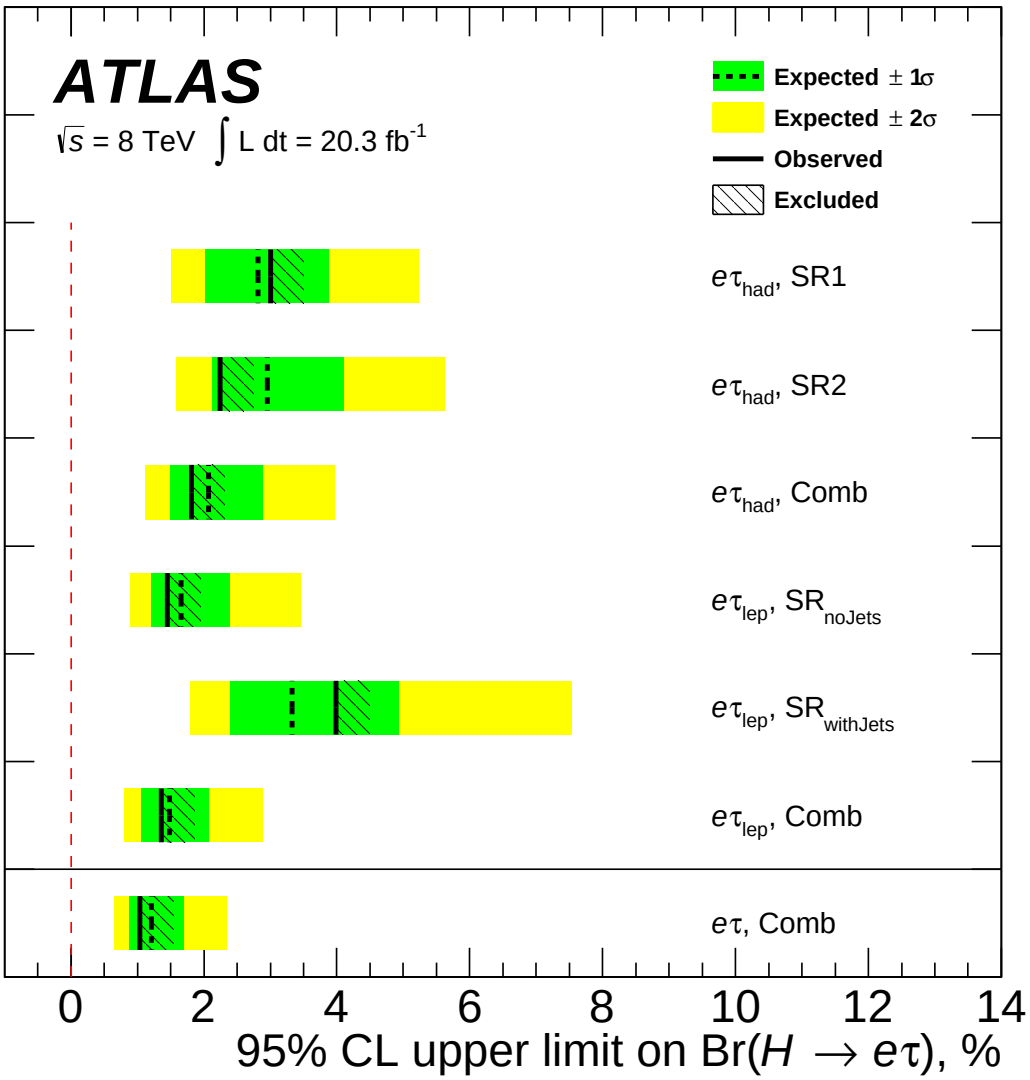


Figure 11.6: Expected and Observed Limits in the electron measurement for both hadronic and leptonic tau analyses [34].

Table 11.3: Results of the searches for LFV $H \to e\tau$ and $H \to \mu\tau$ decays. The limits are computed assuming that either $\mathrm{Br}(H \to \mu\tau) = 0$ or $\mathrm{Br}(H \to e\tau) = 0$. The expected and observed upper limits at 95% confidence level and the best-fit values of the branching ratios for the individual categories and their combination are listed. [34]

Channel	Category	Expected limit [%]	Observed limit [%]	Best fit Br [%]
$H \to e\tau_{\mathrm{had}}$	SR1	$2.81^{+1.06}_{-0.79}$	3.0	$0.33^{+1.48}_{-1.59}$
	SR2	$2.95^{+1.16}_{-0.82}$	2.24	$-1.33^{+1.56}_{-1.80}$
	Combined	$2.07^{+0.82}_{-0.58}$	1.81	$-0.47^{+1.08}_{-1.18}$
$H \to e\tau_{\mathrm{lep}}$	$\mathrm{SR}_{\mathrm{noJets}}$	$1.66^{+0.72}_{-0.46}$	1.45	$-0.45^{+0.89}_{-0.97}$
	$\mathrm{SR}_{\mathrm{withJets}}$	$3.33^{+1.60}_{-0.93}$	3.99	$0.74^{+1.59}_{-1.62}$
	Combined	$1.48^{+0.60}_{-0.42}$	1.36	$-0.26^{+0.79}_{-0.82}$
$H \to e\tau$	Combined	$1.21^{+0.49}_{-0.34}$	1.04	$-0.34^{+0.64}_{-0.66}$
$H \to \mu\tau_{\mathrm{had}}$	SR1	$1.60^{+0.64}_{-0.45}$	1.55	$-0.07^{+0.81}_{-0.86}$
	SR2	$1.75^{+0.71}_{-0.49}$	3.51	$1.94^{+0.92}_{-0.89}$
	Combined	$1.24^{+0.50}_{-0.35}$	1.85	$0.77^{+0.62}_{-0.62}$
$H \to \mu\tau_{\mathrm{lep}}$	$\mathrm{SR}_{\mathrm{noJets}}$	$2.03^{+0.93}_{-0.57}$	2.38	$0.31^{+1.06}_{-0.99}$
	$\mathrm{SR}_{\mathrm{withJets}}$	$3.57^{+1.74}_{-1.00}$	2.85	$-1.03^{+1.66}_{-1.82}$
	Combined	$1.73^{+0.74}_{-0.49}$	1.79	$0.03^{+0.88}_{-0.86}$
$H \to \mu\tau$	Combined	$1.01^{+0.40}_{-0.29}$	1.43	$0.53^{+0.51}_{-0.51}$

Chapter 12

Conclusions

A search for lepton-flavor-violating Higgs boson decays to muons/electrons and hadronic taus using the full 2012 dataset of proton–proton collisions at center-of-mass energy $\sqrt{s} = 8$ TeV recorded at the ATLAS experiment has been presented. No such decays are observed. The expected upper limit at 95% confidence on $\mathrm{Br}(H \to \mu\tau)$ ($\mathrm{Br}(H \to e\tau)$) is $1.24^{+0.50}_{-0.35}\%$ ($2.07^{+1.06}_{-0.58}\%$.). The observed upper limit is 1.85% (1.81%). These limits are improved after combination with a complementary analysis searching for LFV Higgs boson decays with leptonic taus in the final state.

Although no LFV decays are observed, some interesting features are present in the results. A small excess in data over the predicted number of background events is present in the muon measurement and is similar in size to the excess observed in the CMS analysis [32]. Furthermore, the deficit of data events compared to the predicted background suggests that LFV Higgs boson decays to electrons do not occur, consistent with the statement in Chapter 2 that either $H \to \mu\tau$ or $H \to e\tau$ may be observed, but not both. Given that the Higgs boson coupling to fermions is proportional to the fermion mass, it is not surprising that Nature would favor $H \to \mu\tau$ decays over $H \to e\tau$ ones. In spite of these facts, the observed excess comes from two bins in SR2. Thus, the result is more consistent with a statistical fluctuation, rather than a true excess. Hopefully the ATLAS/CMS LFV Higgs boson analyses of data collected in proton–proton collisions at $\sqrt{s} = 13$ TeV will definitively resolve the issue.

There are a few ways that a future analysis could improve upon the analysis presented:

- Unlike the SM $H \to \tau\tau$ analysis, no categories were used exploiting the unique topology of Higgs bosons produced through vector boson fusion or at high p_{T}. Such categorization in a future analysis could improve the sensitivity substantially. A category optimized for Higgs bosons produced through associated production, which offer another unique topology, may also be considered.

- The major source of uncertainty in this analysis is due to the modeling of the W + jets background. To avoid the modeling inaccuracies contributing to the systematic uncertainties, a data-driven W + jets modeling approach may be considered.

- The electron measurement uses the optimized selection criteria obtained in the muon measurement. A separate optimization might improve the sensitivity of that analysis.
- The arguably best way to improve the sensitivity is incorporating multivariate techniques such as BDTs or neural networks into the selection and categorization processes, rather than using fixed selection and categorization criteria.

While the next decades are expected to be dedicated to precision Higgs boson measurements, it is hoped that searches for new physics will produce further evidence of physics beyond the Standard Model, offering explanations for the neutrino mass, lepton flavor violation among neutrinos, and quantum gravity, as well as discover phenomena unimagined.

Bibliography

[1] F. J. Hasert et al. "Observation of Neutrino-Like Interactions Without Muon Or Electron in the Gargamelle Neutrino Experiment". In: *Phys. Lett.* B46 (1973), pp. 138–140. DOI: 10.1016/0370-2693(73)90499-1.

[2] G. Arnison et al. "Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV". In: *Phys. Lett.* B122 (1983). [,611(1983)], pp. 103–116. DOI: 10.1016/0370-2693(83)91177-2.

[3] M. Banner et al. "Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN $\bar{p}p$ Collider". In: *Phys. Lett.* B122 (1983), pp. 476–485. DOI: 10.1016/0370-2693(83)91605-2.

[4] G. Arnison et al. "Experimental Observation of Lepton Pairs of Invariant Mass Around 95 GeV/c^2 at the CERN SPS Collider". In: *Phys. Lett.* B126 (1983), pp. 398–410. DOI: 10.1016/0370-2693(83)90188-0.

[5] P. Bagnaia et al. "Evidence for $Z^0 \to e^+e^-$ at the CERN $\bar{p}p$ Collider". In: *Phys. Lett.* B129 (1983), pp. 130–140. DOI: 10.1016/0370-2693(83)90744-X.

[6] F. Abe, H. Akimoto, A. Akopian, et al. "Observation of Top Quark Production in $\overline{p}p$ Collisions with the Collider Detector at Fermilab". In: *Phys. Rev. Lett.* 74 (14 Apr. 1995), pp. 2626–2631. DOI: 10.1103/PhysRevLett.74.2626. URL: http://link.aps.org/doi/10.1103/PhysRevLett.74.2626.

[7] S. Abachi, B. Abbott, M. Abolins, et al. "Observation of the Top Quark". In: *Phys. Rev. Lett.* 74 (14 Apr. 1995), pp. 2632–2637. DOI: 10.1103/PhysRevLett.74.2632. URL: http://link.aps.org/doi/10.1103/PhysRevLett.74.2632.

[8] Georges Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Lett.* B716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[9] Serguei Chatrchyan et al. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Phys. Lett.* B716 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[10] K. A. Olive et al. "Review of Particle Physics". In: *Chin. Phys.* C38 (2014), p. 090001. DOI: 10.1088/1674-1137/38/9/090001.

[11] Kaustubh Agashe, Andrew E. Blechman, and Frank Petriello. "Probing the Randall-Sundrum geometric origin of flavor with lepton flavor violation". In: *Phys. Rev.* D74 (2006), p. 053011. DOI: 10.1103/PhysRevD.74.053011. arXiv: hep-ph/0606021 [hep-ph].

[12] Eugene D. Commins. "Electron spin and its history". In: *Ann. Rev. Nucl. Part. Sci.* 62 (2012), pp. 133–157. DOI: 10.1146/annurev-nucl-102711-094908.

[13] David Griffiths. *Introduction to Elementary Particles*. 2nd ed. Weinheim Germany: Wiley-VCH, 2008. ISBN: 978-3-527-40601-2.

[14] Nicola Cabibbo. "Unitary Symmetry and Leptonic Decays". In: *Phys. Rev. Lett.* 10 (12 June 1963), pp. 531–533. DOI: 10.1103/PhysRevLett.10.531. URL: http://link.aps.org/doi/10.1103/PhysRevLett.10.531.

[15] Makoto Kobayashi and Toshihide Maskawa. "CP-Violation in the Renormalizable Theory of Weak Interaction". In: *Progress of Theoretical Physics* 49.2 (1973), pp. 652–657. DOI: 10.1143/PTP.49.652. eprint: http://ptp.oxfordjournals.org/content/49/2/652.full.pdf+html. URL: http://ptp.oxfordjournals.org/content/49/2/652.abstract.

[16] J. H. Christenson, J. W. Cronin, V. L. Fitch, et al. "Evidence for the 2π Decay of the K_2^0 Meson". In: *Phys. Rev. Lett.* 13 (4 July 1964), pp. 138–140. DOI: 10.1103/PhysRevLett.13.138. URL: http://link.aps.org/doi/10.1103/PhysRevLett.13.138.

[17] S. L. Glashow. "Partial Symmetries of Weak Interactions". In: *Nucl. Phys.* 22 (1961), pp. 579–588. DOI: 10.1016/0029-5582(61)90469-2.

[18] Steven Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (21 Nov. 1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264. URL: http://link.aps.org/doi/10.1103/PhysRevLett.19.1264.

[19] Abdus Salam. "Weak and Electromagnetic Interactions". In: *8th Nobel Symposium Lerum, Sweden, May 19-25, 1968*. Vol. C680519. 1968, pp. 367–377.

[20] Christopher Tully. *Elementary Particle Physics in a Nutshell*. Princeton: Princeton University Press, 2011. ISBN: 978-0-691-13116-0.

[21] Roel Aaij et al. "Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays". In: *Phys. Rev. Lett.* 115 (2015), p. 072001. DOI: 10.1103/PhysRevLett.115.072001. arXiv: 1507.03414 [hep-ex].

[22] A. D. Martin, W. J. Stirling, R. S. Thorne, et al. "Parton distributions for the LHC". In: *Eur. Phys. J.* C63 (2009), pp. 189–285. DOI: 10.1140/epjc/s10052-009-1072-5. arXiv: 0901.0002 [hep-ph].

[23] Wim de Boer. "The Discovery of the Higgs Boson with the CMS Detector and its Implications for Supersymmetry and Cosmology". In: *Time and Matter 2013 (TAM2013) Venice, Italy*. 2013. arXiv: 1309.0721 [hep-ph]. URL: https://inspirehep.net/record/1252561/files/arXiv:1309.0721.pdf.

[24] LHC Higgs Cross Section Working Group. *LHC Higgs Cross Section Working Group*. URL: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG.

[25] Raymond Davis, Don S. Harmer, and Kenneth C. Hoffman. "Search for Neutrinos from the Sun". In: *Phys. Rev. Lett.* 20 (21 May 1968), pp. 1205–1209. DOI: 10.1103/PhysRevLett.20.1205. URL: http://link.aps.org/doi/10.1103/PhysRevLett.20.1205.

[26] Y. Fukuda, T. Hayakawa, E. Ichihara, et al. "Evidence for Oscillation of Atmospheric Neutrinos". In: *Phys. Rev. Lett.* 81 (8 Aug. 1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562. URL: http://link.aps.org/doi/10.1103/PhysRevLett.81.1562.

[27] Q. R. Ahmad, R. C. Allen, T. C. Andersen, et al. "Measurement of the Rate of $\nu_e + d \rightarrow p + p + e^-$ Interactions Produced by ^8B Solar Neutrinos at the Sudbury Neutrino Observatory". In: *Phys. Rev. Lett.* 87 (7 July 2001), p. 071301. DOI: 10.1103/PhysRevLett.87.071301. URL: http://link.aps.org/doi/10.1103/PhysRevLett.87.071301.

[28] Q. R. Ahmad, R. C. Allen, T. C. Andersen, et al. "Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory". In: *Phys. Rev. Lett.* 89 (1 June 2002), p. 011301. DOI: 10.1103/PhysRevLett.89.011301. URL: http://link.aps.org/doi/10.1103/PhysRevLett.89.011301.

[29] Ziro Maki, Masami Nakagawa, and Shoichi Sakata. "Remarks on the Unified Model of Elementary Particles". In: *Progress of Theoretical Physics* 28.5 (1962), pp. 870–880. DOI: 10.1143/PTP.28.870. eprint: http://ptp.oxfordjournals.org/content/28/5/870.full.pdf+html. URL: http://ptp.oxfordjournals.org/content/28/5/870.abstract.

[30] Georges Aad et al. "Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector". In: *JHEP* 04 (2015), p. 117. DOI: 10.1007/JHEP04(2015)117. arXiv: 1501.04943 [hep-ex].

[31] Serguei Chatrchyan et al. "Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons". In: *JHEP* 05 (2014), p. 104. DOI: 10.1007/JHEP05(2014)104. arXiv: 1401.5041 [hep-ex].

[32] Vardan Khachatryan et al. "Search for Lepton-Flavour-Violating Decays of the Higgs Boson". In: *Phys. Lett.* B749 (2015), pp. 337–362. DOI: 10.1016/j.physletb.2015.07.053. arXiv: 1502.07400 [hep-ex].

[33] R. Keith Ellis, I. Hinchliffe, M. Soldate, et al. "Higgs Decay to $\tau^+\tau^-$: A Possible Signature of Intermediate Mass Higgs Bosons at the SSC". In: *Nucl. Phys.* B297 (1988), p. 221. DOI: 10.1016/0550-3213(88)90019-3.

[34] Georges Aad et al. "Search for lepton-flavour-violating decays of the Higgs and Z bosons with the ATLAS detector". In: *Eur. Phys. J.* C76 (2016). arXiv: 1604.07730 [hep-ex]. Submitted.

[35] Shikma Bressler, Avital Dery, and Aielet Efrati. "Asymmetric lepton-flavor violating Higgs boson decays". In: *Phys. Rev.* D90.1 (2014), p. 015025. DOI: 10.1103/PhysRevD.90.015025. arXiv: 1405.4545 [hep-ph].

[36] Robert H. Bernstein and Peter S. Cooper. "Charged Lepton Flavor Violation: An Experimenter's Guide". In: *Phys. Rept.* 532 (2013), pp. 27–64. DOI: 10.1016/j.physrep.2013.07.002. arXiv: 1307.5787 [hep-ex].

[37] J. Adam et al. "New constraint on the existence of the $\mu^+ \to e^+\gamma$ decay". In: *Phys. Rev. Lett.* 110 (2013), p. 201801. DOI: 10.1103/PhysRevLett.110.201801. arXiv: 1303.0754 [hep-ex].

[38] U. Bellgardt et al. "Search for the Decay $\mu^+ \to e^+e^+e^-$". In: *Nucl. Phys.* B299 (1988), p. 1. DOI: 10.1016/0550-3213(88)90462-2.

[39] Bernard Aubert et al. "Searches for Lepton Flavor Violation in the Decays $\tau^\pm \to e^\pm\gamma$ and $\tau^\pm \to \mu^\pm\gamma$". In: *Phys. Rev. Lett.* 104 (2010), p. 021802. DOI: 10.1103/PhysRevLett.104.021802. arXiv: 0908.2381 [hep-ex].

[40] K. Hayasaka et al. "Search for Lepton Flavor Violating Tau Decays into Three Leptons with 719 Million Produced $\tau^+\tau^-$ Pairs". In: *Phys. Lett.* B687 (2010), pp. 139–143. DOI: 10.1016/j.physletb.2010.03.037. arXiv: 1001.3221 [hep-ex].

[41] Wilhelm H. Bertl et al. "A Search for muon to electron conversion in muonic gold". In: *Eur. Phys. J.* C47 (2006), pp. 337–346. DOI: 10.1140/epjc/s2006-02582-x.

[42] C. Dohmen et al. "Test of lepton flavor conservation in $\mu \to e$ conversion on titanium". In: *Phys. Lett.* B317 (1993), pp. 631–636. DOI: 10.1016/0370-2693(93)91383-X.

[43] J. Kaulard et al. "Improved limit on the branching ratio of $\mu^- \to e^+$ conversion on titanium". In: *Phys. Lett.* B422 (1998), pp. 334–338. DOI: 10.1016/S0370-2693(97)01423-8.

[44] Gianluca Blankenburg, John Ellis, and Gino Isidori. "Flavour-Changing Decays of a 125 GeV Higgs-like Particle". In: *Phys. Lett.* B712 (2012), pp. 386–390. DOI: 10.1016/j.physletb.2012.05.007. arXiv: 1202.5704 [hep-ph].

[45] Roni Harnik, Joachim Kopp, and Jure Zupan. "Flavor Violating Higgs Decays". In: *JHEP* 03 (2013), p. 026. DOI: 10.1007/JHEP03(2013)026. arXiv: 1209.1397 [hep-ph].

[46] Stephen P. Martin. "A Supersymmetry primer". In: (1997). [Adv. Ser. Direct. High Energy Phys.18,1(1998)]. DOI: 10.1142/9789812839657_0001,10.1142/9789814307505_0001. arXiv: hep-ph/9709356 [hep-ph].

[47] M. Arana-Catania, E. Arganda, and M. J. Herrero. "Non-decoupling SUSY in LFV Higgs decays: a window to new physics at the LHC". In: *JHEP* 09 (2013). [Erratum: JHEP10,192(2015)], p. 160. DOI: 10.1007/JHEP10(2015)192,10.1007/JHEP09(2013)160. arXiv: 1304.3371 [hep-ph].

[48] P. T. Giang, L. T. Hue, D. T. Huong, et al. "Lepton-Flavor Violating Decays of Neutral Higgs to Muon and Tauon in Supersymmetric Economical 3-3-1 Model". In: *Nucl. Phys.* B864 (2012), pp. 85–112. DOI: 10.1016/j.nuclphysb.2012.06.008. arXiv: 1204.2902 [hep-ph].

[49] Yuji Omura, Eibun Senaha, and Kazuhiro Tobe. "Lepton-flavor-violating Higgs decay $h \to \mu\tau$ and muon anomalous magnetic moment in a general two Higgs doublet model". In: *JHEP* 05 (2015), p. 028. DOI: 10.1007/JHEP05(2015)028. arXiv: 1502.07824 [hep-ph].

[50] R. N. Mohapatra. "Mechanism for Understanding Small Neutrino Mass in Superstring Theories". In: *Phys. Rev. Lett.* 56 (1986), pp. 561–563. DOI: 10.1103/PhysRevLett.56.561.

[51] E. Arganda, M. J. Herrero, X. Marcano, et al. "Imprints of massive inverse seesaw model neutrinos in lepton flavor violating Higgs boson decays". In: *Phys. Rev.* D91.1 (2015), p. 015001. DOI: 10.1103/PhysRevD.91.015001. arXiv: 1405.4300 [hep-ph].

[52] E. Arganda, M. J. Herrero, X. Marcano, et al. "Radiatively-induced LFV Higgs Decays from Massive ISS Neutrinos". In: *International Conference on High Energy Physics 2014 (ICHEP 2014) Valencia, Spain, July 2-9, 2014*. 2014. arXiv: 1410.5779 [hep-ph]. URL: https://inspirehep.net/record/1323102/files/arXiv:1410.5779.pdf.

[53] Lisa Randall and Raman Sundrum. "A Large mass hierarchy from a small extra dimension". In: *Phys. Rev. Lett.* 83 (1999), pp. 3370–3373. DOI: 10.1103/PhysRevLett.83.3370. arXiv: hep-ph/9905221 [hep-ph].

[54] Gilad Perez and Lisa Randall. "Natural Neutrino Masses and Mixings from Warped Geometry". In: *JHEP* 01 (2009), p. 077. DOI: 10.1088/1126-6708/2009/01/077. arXiv: 0805.4652 [hep-ph].

[55] Lyndon R. Evans and Philip Bryant. "LHC Machine". In: *J. Instrum.* 3 (2008). This report is an abridged version of the LHC Design Report (CERN-2004-003), S08001. 164 p. URL: http://cds.cern.ch/record/1129806.

[56] Katherine McAlpine. *Large Hadron Rap*. YouTube. 2008. URL: https://www.youtube.com/watch?v=j50ZssEojtM.

[57] Michael Benedikt, Paul Collier, John Mertens V and Poole, et al. *LHC Design Report*. Geneva: CERN, 2004. URL: https://cds.cern.ch/record/823808.

[58] Jean-Luc Caron. "CERN Accelerator Complex (operating and approved projects). Chaine des accelerateurs du CERN (en fonctionnement et avec les projets aprouves)." AC Collection. Legacy of AC. Pictures from 1992 to 2002. June 1991. URL: https://cds.cern.ch/record/841493.

[59] AC Team. "Diagram of an LHC dipole magnet. Schéma d'un aimant dipôle du LHC". June 1999. URL: https://cds.cern.ch/record/40524.

[60] Frank Zimmermann. *LHC: The Machine*. SLAC Summer Institute 2012. 2012. URL: https://www-conf.slac.stanford.edu/ssi/2012/Presentations/Zimmermann.pdf.

[61] Oliver Sim Brüning, Paul Collier, P Lebrun, et al. *LHC Design Report*. Geneva: CERN, 2004. URL: https://cds.cern.ch/record/782076.

[62] Mike Lamont. "Status of the LHC". In: *J. Phys. Conf. Ser.* 455 (2013), p. 012001. DOI: 10.1088/1742-6596/455/1/012001.

[63] Georges Aad et al. "Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC". In: *Eur. Phys. J.* C73.8 (2013), p. 2518. DOI: 10.1140/epjc/s10052-013-2518-3. arXiv: 1302.4393 [hep-ex].

[64] G. Aad et al. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.

[65] M Capeans, G Darbo, K Einsweiller, et al. *ATLAS Insertable B-Layer Technical Design Report*. Tech. rep. CERN-LHCC-2010-013. ATLAS-TDR-19. Geneva: CERN, Sept. 2010. URL: https://cds.cern.ch/record/1291633.

[66] G. Aad et al. "ATLAS pixel detector electronics and sensors". In: *JINST* 3 (2008), P07007. DOI: 10.1088/1748-0221/3/07/P07007.

[67] ATLAS Collaboration. *ATLAS Experiment - Public Results*. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/.

[68] Nikiforos Nikiforou. *Performance of the ATLAS Liquid Argon Calorimeter after three years of LHC operation and plans for a future upgrade*. Tech. rep. arXiv:1306.6756. Comments: 12 pages, 25 figures, Proceedings of talk presented in "Advancements in Nuclear Instrumentation Measurement Methods and their Applications", Marseille, 2013. June 2013. URL: http://cds.cern.ch/record/1558820.

[69] U. Amaldi et al. "The Real Part of the Forward Proton Proton Scattering Amplitude Measured at the CERN Intersecting Storage Rings". In: *Phys. Lett.* B66 (1977), p. 390. DOI: 10.1016/0370-2693(77)90022-3.

[70] Joao Pequenao and Paul Schaffner. "An computer generated image representing how ATLAS detects particles". Jan. 2013. URL: https://cds.cern.ch/record/1505342.

[71] G. Aad et al. "Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics". In: (2009). arXiv: 0901.0512 [hep-ex].

[72] T. Cornelissen, M. Elsing, W. Gavrilenko I. and Liebig, et al. "The new ATLAS track reconstruction (NEWT)". In: *J. Phys. Conf. Ser.* 119 (2008), p. 032014. DOI: 10.1088/1742-6596/119/3/032014.

[73] Tim Babb. *How a Kalman filter works, in pictures.* URL: http://www.bzarg.com/p/how-a-kalman-filter-works-in-pictures/.

[74] R. Fruhwirth. "Application of Kalman filtering to track and vertex fitting". In: *Nucl. Instrum. Meth.* A262 (1987), pp. 444–450. DOI: 10.1016/0168-9002(87)90887-4.

[75] *Performance of primary vertex reconstruction in proton-proton collisions at $\sqrt{s} = 7$ TeV in the ATLAS experiment.* Tech. rep. ATLAS-CONF-2010-069. Geneva: CERN, July 2010. URL: https://cds.cern.ch/record/1281344.

[76] R. Fruhwirth, W. Waltenberger, and P. Vanlaer. "Adaptive vertex fitting". In: *J. Phys.* G34 (2007), N343. DOI: 10.1088/0954-3899/34/12/N01.

[77] S. Hassani, L. Chevalier, E. Lancon, et al. "A muon identification and combined reconstruction procedure for the ATLAS detector at the LHC using the (MUONBOY, STACO, MuTag) reconstruction packages". In: *Nucl. Instrum. Meth.* A572 (2007), pp. 77–79. DOI: 10.1016/j.nima.2006.10.340.

[78] Georges Aad et al. "Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton–proton collision data". In: *Eur. Phys. J.* C74.11 (2014), p. 3130. DOI: 10.1140/epjc/s10052-014-3130-x. arXiv: 1407.3935 [hep-ex].

[79] *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data.* Tech. rep. ATLAS-CONF-2014-032. Geneva: CERN, June 2014. URL: http://cds.cern.ch/record/1706245.

[80] W. Lampl, S. Laplace, D. Lelas, et al. *Calorimeter Clustering Algorithms: Description and Performance.* Tech. rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003. Geneva: CERN, Apr. 2008. URL: https://cds.cern.ch/record/1099735.

[81] *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung.* Tech. rep. ATLAS-CONF-2012-047. Geneva: CERN, May 2012. URL: https://cds.cern.ch/record/1449796.

[82] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The Anti-k(t) jet clustering algorithm". In: *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].

[83] *Determination of the jet energy scale and resolution at ATLAS using Z/γ–jet events in data at $\sqrt{s} = 8$ TeV.* Tech. rep. ATLAS-CONF-2015-057. Geneva: CERN, Oct. 2015. URL: http://cds.cern.ch/record/2059846.

[84] *Pile-up subtraction and suppression for jets in ATLAS.* Tech. rep. ATLAS-CONF-2013-083. Geneva: CERN, Aug. 2013. URL: http://cds.cern.ch/record/1570994.

[85] *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*. Tech. rep. ATLAS-CONF-2011-102. Geneva: CERN, July 2011. URL: https://cds.cern.ch/record/1369219.

[86] *Calibration of b-tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment*. Tech. rep. ATLAS-CONF-2014-004. Geneva: CERN, Feb. 2014. URL: http://cds.cern.ch/record/1664335.

[87] *Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 ATLAS data*. Tech. rep. ATLAS-CONF-2014-046. Geneva: CERN, July 2014. URL: http://cds.cern.ch/record/1741020.

[88] Georges Aad et al. "Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in *pp* collisions at $\sqrt{s}$=8 TeV". In: *Eur. Phys. J.* C75.7 (2015), p. 303. DOI: 10.1140/epjc/s10052-015-3500-z. arXiv: 1412.7086 [hep-ex].

[89] *Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton-Proton Collisions recorded in 2012 at 8 TeV*. Tech. rep. ATLAS-CONF-2013-082. Geneva: CERN, Aug. 2013. URL: http://cds.cern.ch/record/1570993.

[90] A. Elagin, P. Murat, A. Pranko, et al. "A New Mass Reconstruction Technique for Resonances Decaying to di-tau". In: *Nucl. Instrum. Meth.* A654 (2011), pp. 481–489. DOI: 10.1016/j.nima.2011.07.009. arXiv: 1012.4686 [hep-ex].

[91] S. Agostinelli et al. "GEANT4: A Simulation toolkit". In: *Nucl. Instrum. Meth.* A506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[92] Simone Alioli, Paolo Nason, Carlo Oleari, et al. "NLO Higgs boson production via gluon fusion matched with shower in POWHEG". In: *JHEP* 04 (2009), p. 002. DOI: 10.1088/1126-6708/2009/04/002. arXiv: 0812.0578 [hep-ph].

[93] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. "A Brief Introduction to PYTHIA 8.1". In: *Comput. Phys. Commun.* 178 (2008), pp. 852–867. DOI: 10.1016/j.cpc.2008.01.036. arXiv: 0710.3820 [hep-ph].

[94] Hung-Liang Lai, Marco Guzzi, Joey Huston, et al. "New parton distributions for collider physics". In: *Phys. Rev.* D82 (2010), p. 074024. DOI: 10.1103/PhysRevD.82.074024. arXiv: 1007.2241 [hep-ph].

[95] J. Pumplin, D. R. Stump, J. Huston, et al. "New generation of parton distributions with uncertainties from global QCD analysis". In: *JHEP* 07 (2002), p. 012. DOI: 10.1088/1126-6708/2002/07/012. arXiv: hep-ph/0201195 [hep-ph].

[96] D. J. Lange. "The EvtGen particle decay simulation package". In: *Nucl. Instrum. Meth.* A462 (2001), pp. 152–155. DOI: 10.1016/S0168-9002(01)00089-4.

[97] Z. Was. "TAUOLA the library for tau lepton decay, and KKMC / KORALB / KORALZ /... status report". In: *Nucl. Phys. Proc. Suppl.* 98 (2001). [,96(2000)], pp. 96–102. DOI: 10.1016/S0920-5632(01)01200-2. arXiv: hep-ph/0011305 [hep-ph].

[98] N. Davidson, T. Przedzinski, and Z. Was. "PHOTOS Interface in C++: Technical and Physics Documentation". In: *Comput. Phys. Commun.* 199 (2016), pp. 86–101. DOI: 10.1016/j.cpc.2015.09.013. arXiv: 1011.0937 [hep-ph].

[99] Charalampos Anastasiou and Kirill Melnikov. "Higgs boson production at hadron colliders in NNLO QCD". In: *Nucl. Phys.* B646 (2002), pp. 220–256. DOI: 10.1016/S0550-3213(02)00837-4. arXiv: hep-ph/0207004 [hep-ph].

[100] V. Ravindran, J. Smith, and W. L. van Neerven. "NNLO corrections to the total cross-section for Higgs boson production in hadron hadron collisions". In: *Nucl. Phys.* B665 (2003), pp. 325–366. DOI: 10.1016/S0550-3213(03)00457-7. arXiv: hep-ph/0302135 [hep-ph].

[101] Paolo Bolzoni, Fabio Maltoni, Sven-Olaf Moch, et al. "Higgs production via vector-boson fusion at NNLO in QCD". In: *Phys. Rev. Lett.* 105 (2010), p. 011801. DOI: 10.1103/PhysRevLett.105.011801. arXiv: 1003.4451 [hep-ph].

[102] D. Alvarez Piqueras, A. Andreazza, S. Banerjee, et al. *Measurement of the Higgs boson couplings in the $\tau\tau$ final state with the ATLAS detector - Supporting Note*. Tech. rep. ATL-COM-PHYS-2014-170. Geneva: CERN, Mar. 2014. URL: https://cds.cern.ch/record/1666539.

[103] *Search for the Standard Model Higgs boson in $H \to \tau^+\tau^-$ decays in proton-proton collisions with the ATLAS detector*. Tech. rep. ATLAS-CONF-2012-160. Geneva: CERN, Nov. 2012. URL: https://cds.cern.ch/record/1493624.

[104] Georges Aad et al. "Search for the Standard Model Higgs boson in the $H \to \tau^+\tau^-$ decay mode in $\sqrt{s} = 7$ TeV pp collisions with ATLAS". In: *JHEP* 09 (2012), p. 070. DOI: 10.1007/JHEP09(2012)070. arXiv: 1206.5971 [hep-ex].

[105] Michelangelo L. Mangano, Mauro Moretti, Fulvio Piccinini, et al. "ALPGEN, a generator for hard multiparton processes in hadronic collisions". In: *JHEP* 07 (2003), p. 001. DOI: 10.1088/1126-6708/2003/07/001. arXiv: hep-ph/0206293 [hep-ph].

[106] G. Corcella, I. G. Knowles, G. Marchesini, et al. "HERWIG 6.5: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)". In: *JHEP* 01 (2001), p. 010. DOI: 10.1088/1126-6708/2001/01/010. arXiv: hep-ph/0011363 [hep-ph].

[107] J. M. Butterworth, Jeffrey R. Forshaw, and M. H. Seymour. "Multiparton interactions in photoproduction at HERA". In: *Z. Phys.* C72 (1996), pp. 637–646. DOI: 10.1007/s002880050286. arXiv: hep-ph/9601371 [hep-ph].

[108] Georges Aad et al. "Modelling $Z \to \tau\tau$ processes in ATLAS with τ-embedded $Z \to \mu\mu$ data". In: *JINST* 10.09 (2015), P09018. DOI: 10.1088/1748-0221/10/09/P09018, 10.1088/1748-0221/2015/9/P09018. arXiv: 1506.05623 [hep-ex].

[109] Borut Paul Kersevan and Elzbieta 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". In: *Comput. Phys. Commun.* 184 (2013), pp. 919–985. DOI: 10.1016/j.cpc.2012.10.032. arXiv: hep-ph/0405247 [hep-ph].

[110] Stefano Frixione and Bryan R. Webber. "Matching NLO QCD computations and parton shower simulations". In: *JHEP* 06 (2002), p. 029. DOI: 10.1088/1126-6708/2002/06/029. arXiv: hep-ph/0204244 [hep-ph].

[111] John M. Campbell, R. Keith Ellis, and Ciaran Williams. "Vector boson pair production at the LHC". In: *JHEP* 07 (2011), p. 018. DOI: 10.1007/JHEP07(2011)018. arXiv: 1105.0020 [hep-ph].

[112] "Performance of the ATLAS muon trigger in 2011". In: ATLAS-CONF-2012-099 (July 2012). URL: https://cds.cern.ch/record/1462601.

[113] "Performance of the ATLAS Electron and Photon Trigger in p-p Collisions at $\sqrt{s} =$ 7 TeV in 2011". In: ATLAS-CONF-2012-048 (May 2012). URL: https://cds.cern.ch/record/1450089.

[114] Georges Aad et al. "Characterisation and mitigation of beam-induced backgrounds observed in the ATLAS detector during the 2011 proton-proton run". In: *JINST* 8 (2013), P07004. DOI: 10.1088/1748-0221/8/07/P07004. arXiv: 1303.0223 [hep-ex].

[115] Georges Aad et al. "Search for lepton-flavour-violating $H \to \mu\tau$ decays of the Higgs boson with the ATLAS detector". In: *JHEP* 11 (2015), p. 211. DOI: 10.1007/JHEP11(2015)211. arXiv: 1508.03372 [hep-ex].

[116] Georges Aad et al. "Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data". In: *Eur. Phys. J.* C72 (2012), p. 1909. DOI: 10.1140/epjc/s10052-012-1909-1. arXiv: 1110.3174 [hep-ex].

[117] T. Gleisberg, S. Hoeche, F. Krauss, et al. "Event generation with SHERPA 1.1". In: *JHEP* 02 (2009), p. 007. DOI: 10.1088/1126-6708/2009/02/007. arXiv: 0811.4622 [hep-ph].

[118] *Performance of Missing Transverse Momentum Reconstruction in ATLAS with 2011 Proton-Proton Collisions at $\sqrt{s} = 7$ TeV*. Tech. rep. ATLAS-CONF-2012-101. Geneva: CERN, July 2012. URL: https://cds.cern.ch/record/1463915.

[119] S. Dittmaier et al. "Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables". In: (2011). DOI: 10.5170/CERN-2011-002. arXiv: 1101.0593 [hep-ph].

[120] Iain W. Stewart and Frank J. Tackmann. "Theory Uncertainties for Higgs and Other Searches Using Jet Bins". In: *Phys. Rev.* D85 (2012), p. 034011. DOI: 10.1103/PhysRevD.85.034011. arXiv: 1107.2117 [hep-ph].

[121] *Combined Standard Model Higgs boson searches with up to 2.3 fb^{-1} of pp collisions at $\sqrt{s}$ = 7 TeV at the LHC*. Tech. rep. ATLAS-CONF-2011-157. Geneva: CERN, Nov. 2011. URL: https://cds.cern.ch/record/1399599.

[122] Richard D. Ball, Luigi Del Debbio, Alberto Forte Stefano and Guffanti, et al. "A first unbiased global NLO determination of parton distributions and their uncertainties". In: *Nucl. Phys.* B838 (2010), pp. 136–206. DOI: 10.1016/j.nuclphysb.2010.05.008. arXiv: 1002.4407 [hep-ph].

[123] S. Dittmaier et al. "Handbook of LHC Higgs Cross Sections: 2. Differential Distributions". In: (2012). DOI: 10.5170/CERN-2012-002. arXiv: 1201.3084 [hep-ph].

[124] Glen Cowan. *Statistical Data Analysis*. Oxford University Press, 1998. ISBN: 0198501552.

[125] Harrison B. Prosper. "Practical Statistics for Particle Physicists". In: *Proceedings, 2012 European School of High-Energy Physics (ESHEP 2012): La Pommeraye, Anjou, France, June 06-19, 2012*. 2014, pp. 195–216. DOI: 10.5170/CERN-2014-008.195. arXiv: 1504.00945 [stat.ME]. URL: http://inspirehep.net/record/1357983/files/arXiv:1504.00945.pdf.

[126] D. S. Sivia and J. Skilling. *Data Analysis: A Bayesian Tutorial*. 2nd ed. Oxford University Press, 2006. ISBN: 0198568320.

[127] J. Neyman. "Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability". In: *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* 236.767 (1937), pp. 333–380. ISSN: 00804614. URL: http://www.jstor.org/stable/91337.

[128] *Procedure for the LHC Higgs boson search combination in Summer 2011*. Tech. rep. CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11. Geneva: CERN, Aug. 2011. URL: https://cds.cern.ch/record/1379837.

[129] Douglas C. Montgomery, George C. Runger, and Norma F. Hubele. *Engineering Statistics*. 4th ed. Wiley, 2006. ISBN: 0471735574.

[130] Douglas C. Montgomery and George C. Runger. *Applied Statistics and Probability for Engineers*. 4th ed. Wiley, 2006. ISBN: 0471745898.

[131] Glen Cowan, Kyle Cranmer, Eilam Gross, et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *Eur. Phys. J.* C71 (2011). [Erratum: Eur. Phys. J.C73,2501(2013)], p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0, 10.1140/epjc/s10052-013-2501-z. arXiv: 1007.1727 [physics.data-an].

[132] S. S. Wilks. "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses". In: *The Annals of Mathematical Statistics* 9.1 (1938), pp. 60–62. ISSN: 00034851. URL: http://www.jstor.org/stable/2957648.

[133] Alexander L. Read. "Modified frequentist analysis of search results (The CL(s) method)". In: *Workshop on confidence limits, CERN, Geneva, Switzerland, 17-18 Jan 2000: Proceedings*. 2000. URL: http://weblib.cern.ch/abstract?CERN-OPEN-2000-205.

[134] Alexander L. Read. "Presentation of search results: The CL(s) technique". In: *J. Phys.* G28 (2002). [,11(2002)], pp. 2693–2704. DOI: 10.1088/0954-3899/28/10/313.

[135] Thomas Junk. "Confidence level computation for combining searches with small statistics". In: *Nucl. Instrum. Meth.* A434 (1999), pp. 435–443. DOI: 10.1016/S0168-9002(99)00498-2. arXiv: hep-ex/9902006 [hep-ex].

[136] Lorenzo Moneta, Kevin Belasco, Kyle S. Cranmer, et al. "The RooStats Project". In: *PoS* ACAT2010 (2010), p. 057. arXiv: 1009.1003 [physics.data-an].

[137] Kyle Cranmer, George Lewis, Lorenzo Moneta, et al. *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*. Tech. rep. CERN-OPEN-2012-016. New York: New York U., Jan. 2012. URL: https://cds.cern.ch/record/1456844.

Appendix A

Shape Corrections

To account for mis-modeling of the W + jets and $Z \to \mu\mu/ee$ + jets backgrounds, scale factors dependent on $p_{\mathrm{T}}(\tau)$, $|\Delta\eta(\mu/e,\tau)|$, and the number of jets computed using Equation 6.10 are applied to these backgrounds to bring them into agreement with the observed data. This appendix describes the procedure for obtaining these scale factors. The binning scheme of $p_{\mathrm{T}}(\tau)$, $|\Delta\eta(\mu/e,\tau)|$, and the number of jets are listed in Table A.1 for each correction.

Shape Correction	$p_{\mathrm{T}}(\tau)$ [GeV]
W + jets SR1	20–25, 25–35, 35–45, 45–60, 60–80, ≥ 80
W + jets WCR	20–25, 25–35, 35–45, ≥ 45
$Z \to \mu\mu/ee$ + jets	20–25, 25–25, 35–45, ≥ 45

Shape Correction	$\lvert\Delta\eta(\mu/e,\tau)\rvert$	N_{jet}
W + jets SR1	0–0.5, 0.5–1.0, 1.0–2.0, ≥ 2.0	0, ≥ 1
W + jets WCR	0–0.5, 0.5–1.0, ≥ 1.0	0, ≥ 1
$Z \to \mu\mu/ee$ + jets	0–0.5, 0.5–1.0, ≥ 1.0	0, ≥ 1

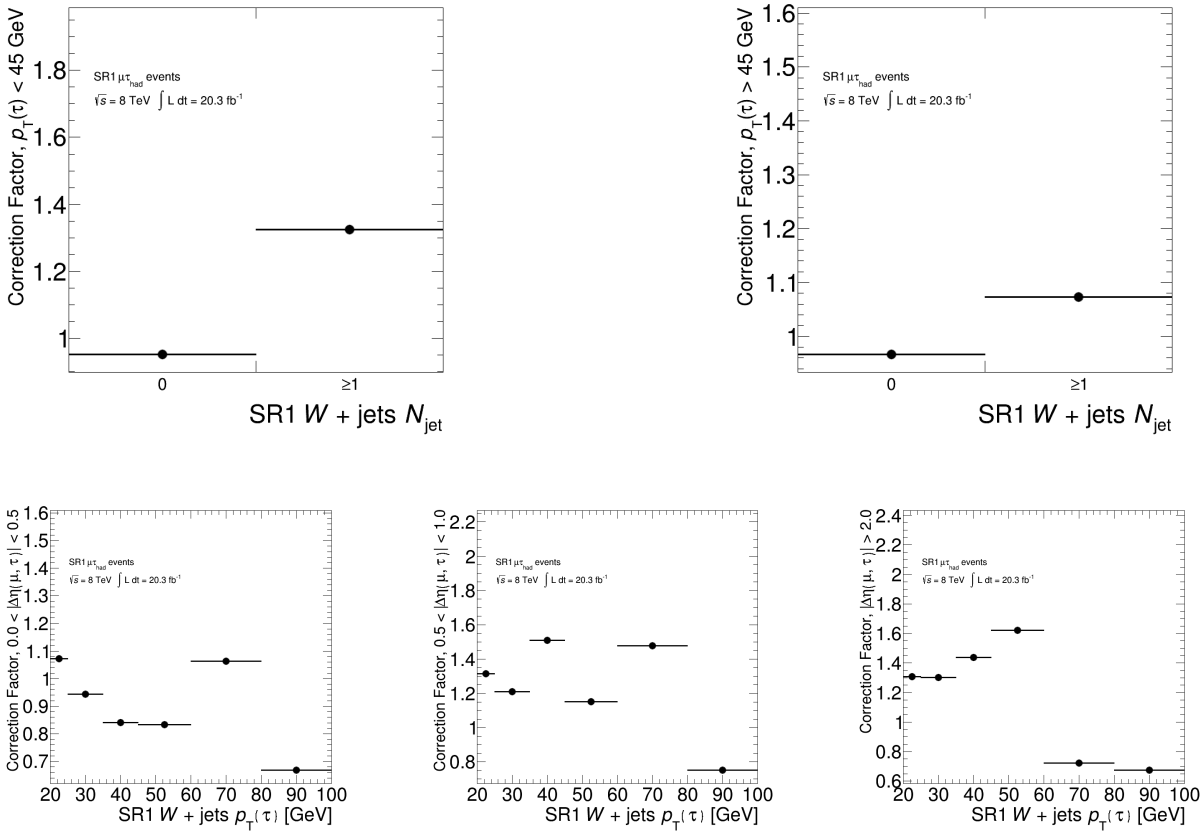
Table A.1: Binning scheme used to obtain shape corrections for the W + jets and $Z \to \mu\mu/ee$ + jets backgrounds. The symbol ℓ indicates a muon or electron.

W + jets

Two sets of shape corrections are measured for the W + jets background. One set is obtained in WCR, while the other is obtained in SR1. The SR1 shape corrections are determined in two stages due to the limited number of W + jets events with $m^{\mathrm{MMC}}_{\mu/e\tau}$ mass larger than 150 GeV. First, a correction due to the number of jets is applied, followed by a two-dimensional correction dependent on $|\Delta\eta(\mu/e,\tau)|$ and $p_{\mathrm{T}}(\tau)$. The N_{jet} correction is dependent on whether or not $p_{\mathrm{T}}(\tau) > 45$ GeV. For the W + jets shape in the WCR, a three-dimensional correction

dependent on $p_{\mathrm{T}}(\tau)$, $|\Delta\eta(\mu/e,\tau)|$, and the number of jets is obtained. The SR1 shape corrections for the muon and electron measurements are shown respectively in Figures A.1 and A.2. Those for the WCR are shown in Figures A.3, A.4, A.5, and A.6. The WCR shape corrections are applied to W + jets samples in the control regions, except for QCDCR, which uses the corrections obtained in SR1, and QCDCR2, which uses the default sample shape.

Figure A.1: Shape corrections applied to the W + jets background in SR1 and QCDCR in the muon measurement. The correction is applied to both OS and SS events.



$Z \to \mu\mu/ee$ + jets

The shape corrections for $Z \to \mu\mu/ee$ + jets are measured in ZmmCR/ZeeCR and applied in all signal and control regions to those $Z \to \mu\mu/ee$ events where the jet is misidentified as a hadronic tau. The correction is a three-dimensional correction dependent on $p_{\mathrm{T}}(\tau)$, $|\Delta\eta(\mu/e,\tau)|$, and the number of jets. Figures A.7 and A.8 show the shape corrections for the muon and electron measurements, respectively.

Figure A.2: Shape corrections applied to the W + jets background in SR1 and QCDCR in the electron measurement. The correction is applied to both OS and SS events.

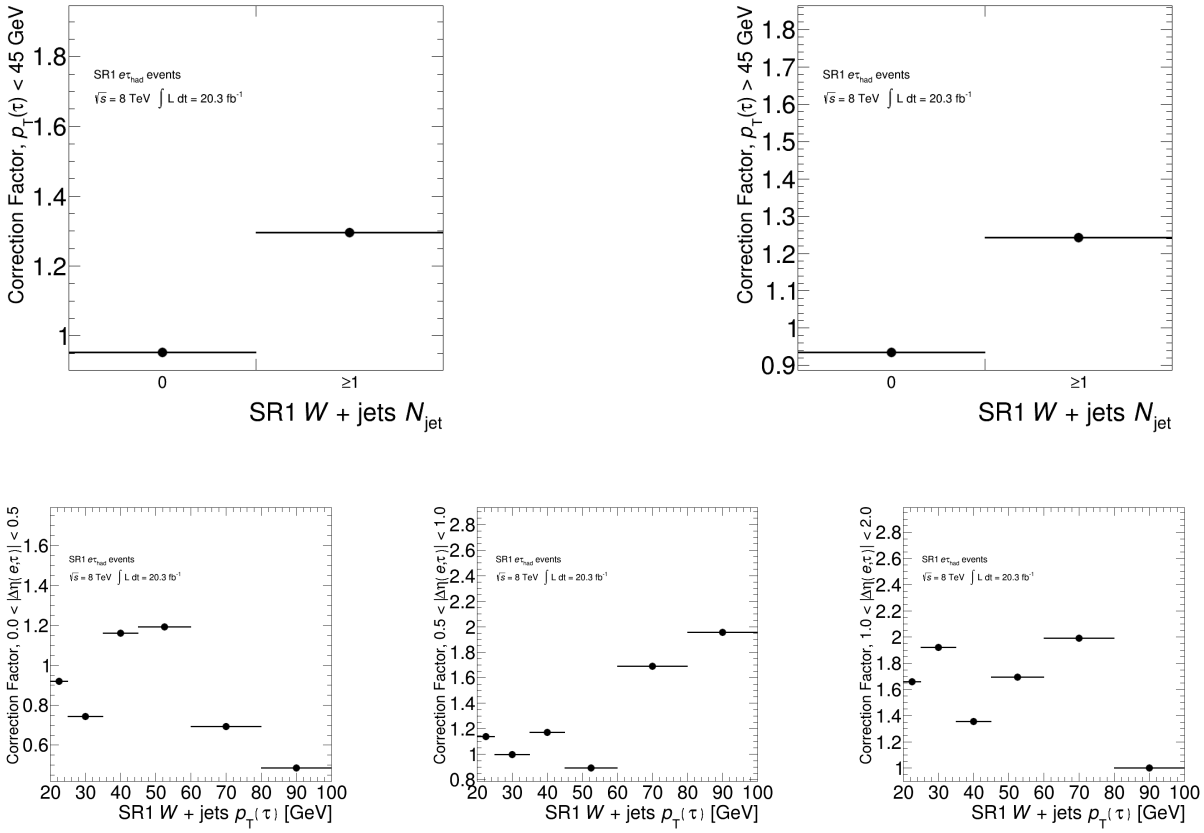


Figure A.3: Shape corrections applied to the W + jets background for OS events in WCR, TCR, and ZmmCR in the muon measurement.

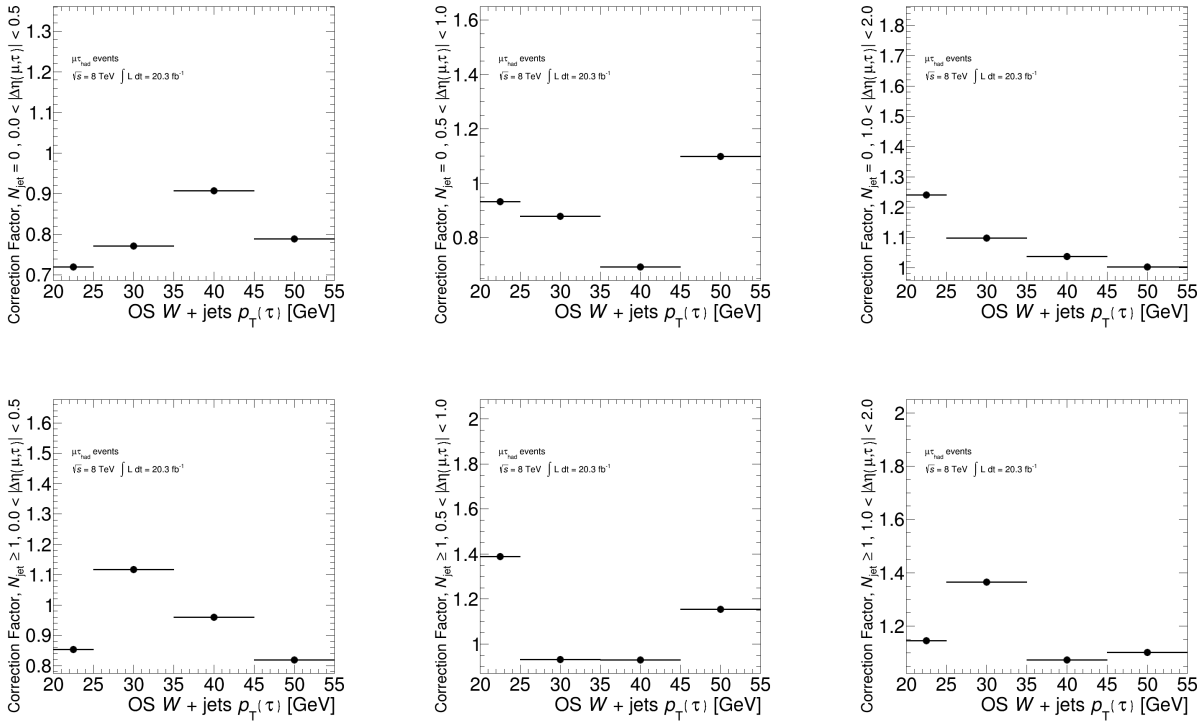


Figure A.4: Shape corrections applied to the W + jets background for SS events in WCR, TCR, and ZmmCR in the muon measurement.

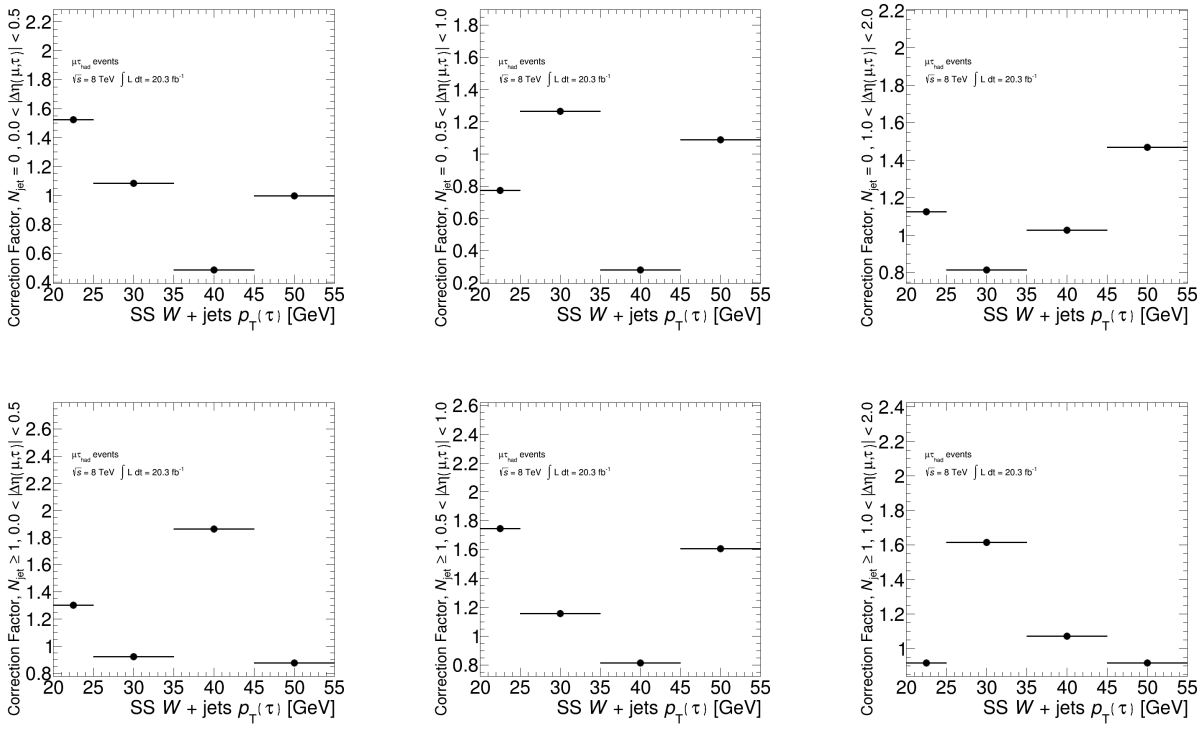


Figure A.5: Shape corrections applied to the W + jets background for OS events in WCR, TCR, and ZeeCR in the electron measurement.

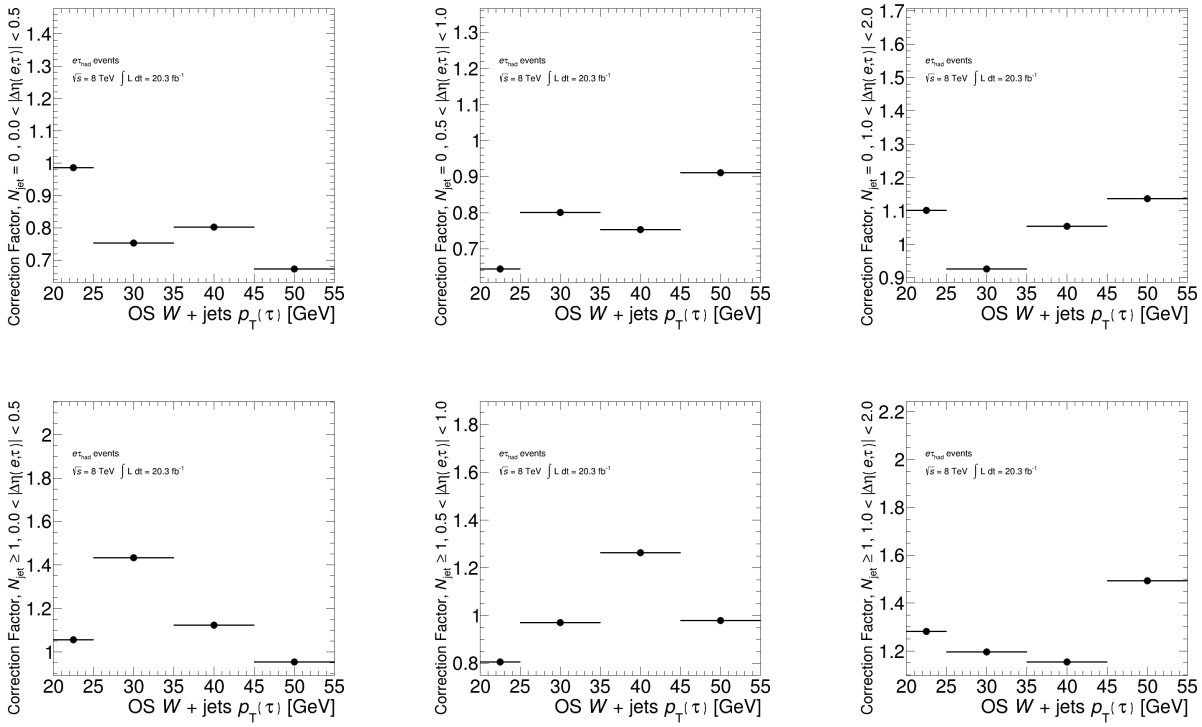


Figure A.6: Shape corrections applied to the W + jets background for SS events in WCR, TCR, and ZeeCR in the electron measurement.

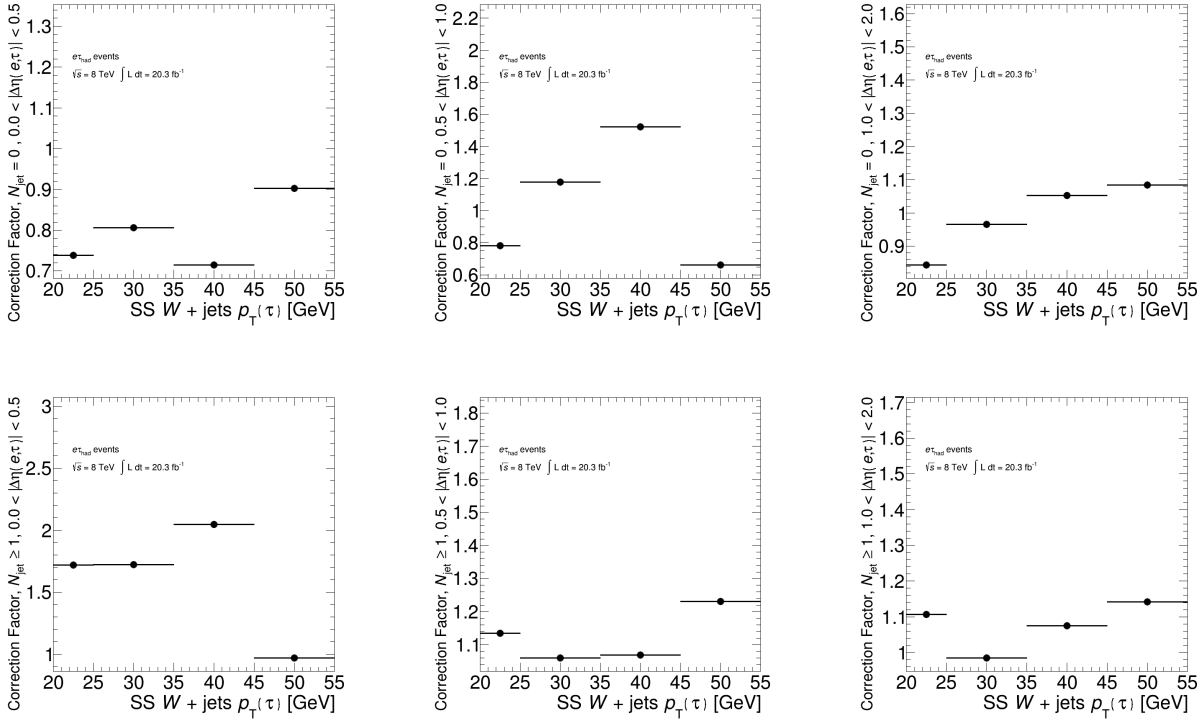


Figure A.7: Shape corrections applied to $Z \to \mu\mu$ + jets events where a jet is misidentified as a hadronic tau. The correction is applied in all signal and control regions in the muon measurement.

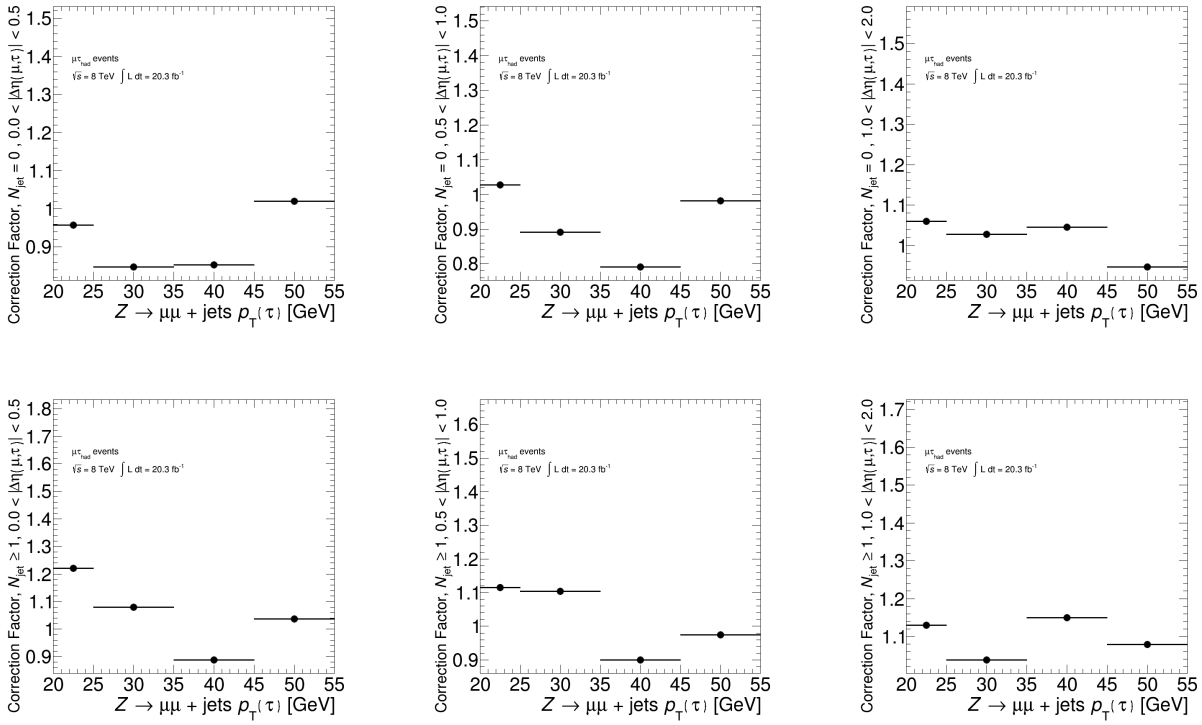
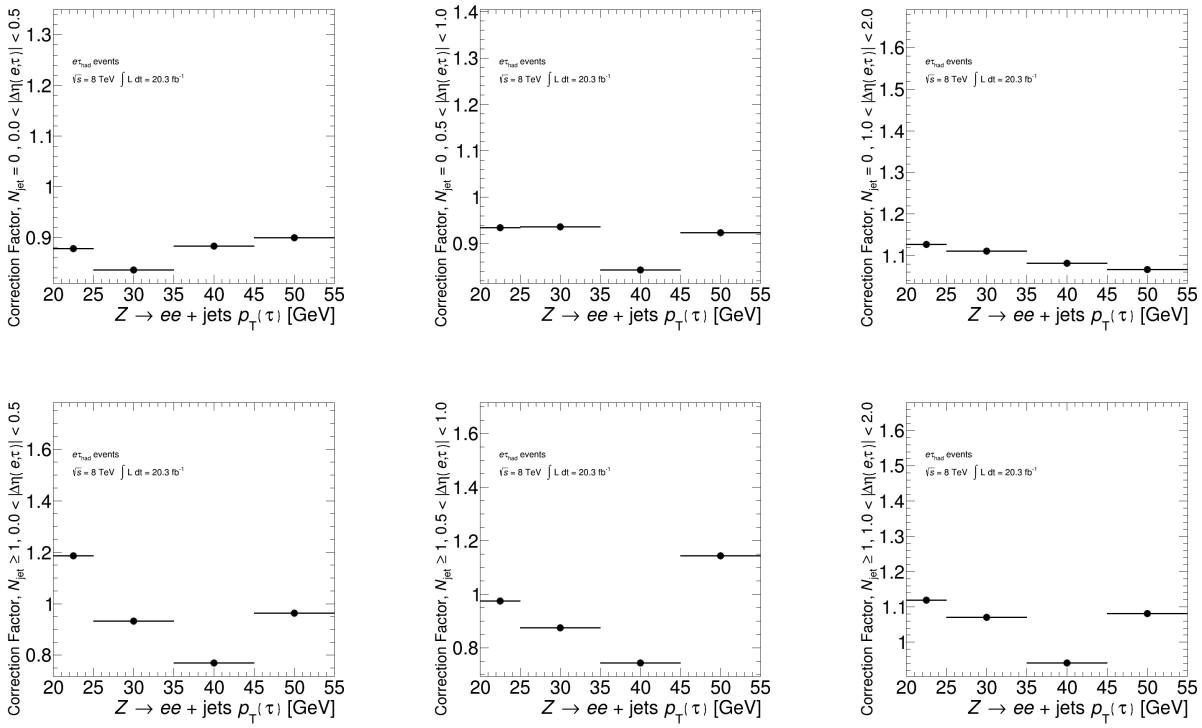


Figure A.8: Shape corrections applied to $Z \to \mu\mu$ + jets events where a jet is misidentified as a hadronic tau. The correction is applied in all signal and control regions in the electron measurement.



Appendix B

Tests of the Fit Model

Before unblinding, several checks are performed using an Asimov dataset constructed from the sum of the predicted background distributions to ensure that the fit model described in Chapter 10 is well-behaved and will provide valid results. In addition to the default exclusion limit, limits are computed considering only statistical uncertainties, systematic uncertainties inflated to twice their values, and systematic uncertainties at half their values. Limits are also computed using a hybrid dataset consisting of observed data with Asimov data in the blinded signal region. The results of these show that the expected and observed limits in all cases agree, improving or worsening as the systematic uncertainties are reduced or inflated.

The behavior of the nuisance parameters is also checked before and after unblinding. If they are modeled correctly, then in the asymptotic limit, the shape of the likelihood function should be Gaussian around its maximum value. If the negative log-likelihood is considered, each nuisance parameter should have a parabolic shape around the maximum value. The shape of each nuisance parameter negative log-likelihood is checked, as any significant non-parabolic behavior (particularly the presence of local minima) in the profiling of the nuisance parameter about the maximum might indicate potential issues with the fit. It is found that before and after unblinding, all nuisance parameters have the desired parabolic shape.

The nuisance parameters are also ranked by their impact on the fit of the signal strength, assessed by computing the change in the best fit of the signal strength for $\pm 1\sigma$ variations of each nuisance parameter. Ranking plots for the muon and electron measurements are shown in Figures B.1 and B.2. Profiles of the top six ranked nuisance parameters for each measurement are shown in Figures B.3 and B.4.

The post-fit values of the `NormFactor` parameters are also checked for consistency with unity, as any inconsistency indicates an issue with the modeling. All `NormFactor` parameters are found to be consistent in both muon and electron measurements. The fitted values for each `NormFactor` parameter are shown in Figure B.5.

Figure B.1: Top ranked nuisance parameters in the muon measurement.

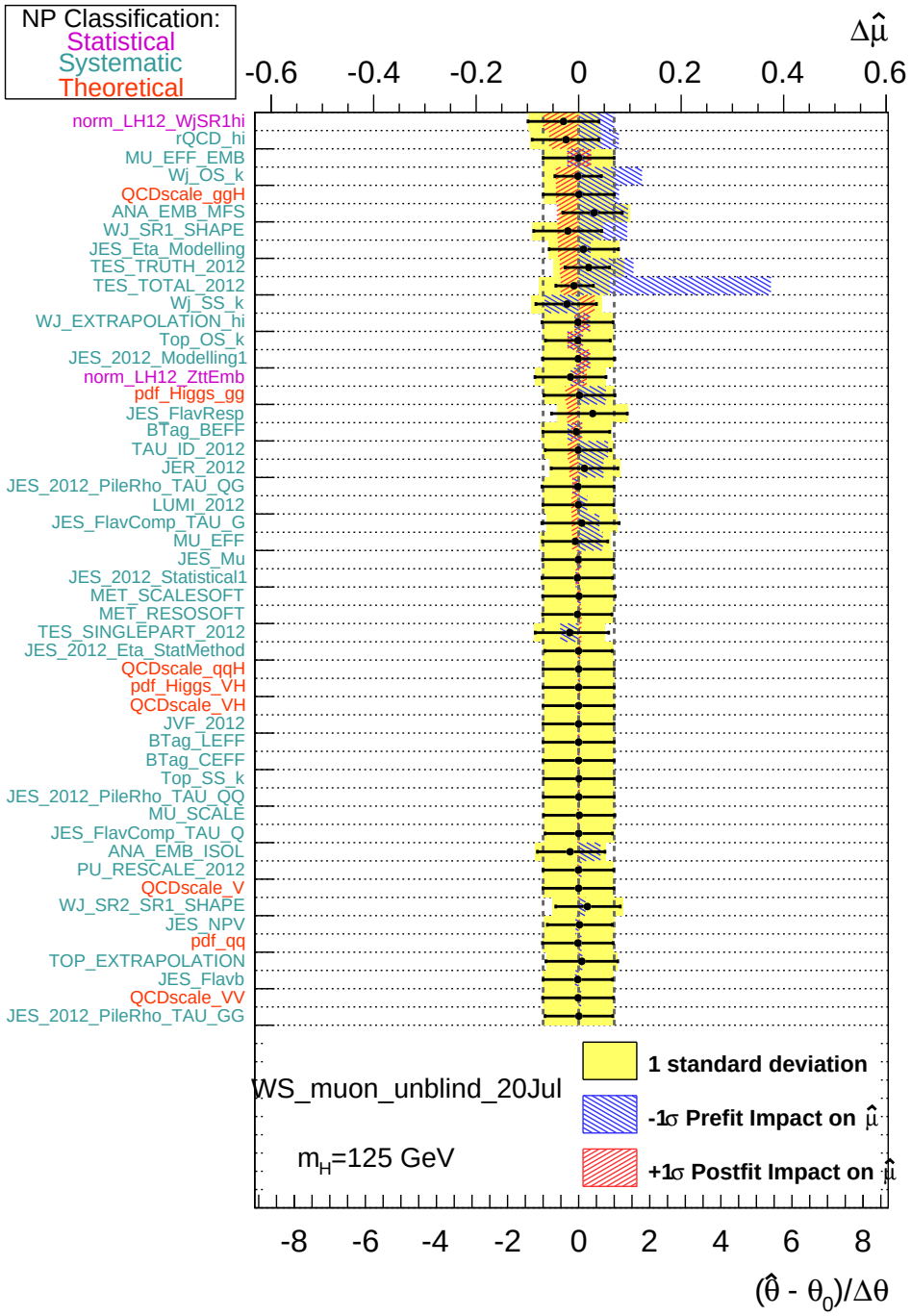


Figure B.2: Top ranked nuisance parameters in the electron measurement.

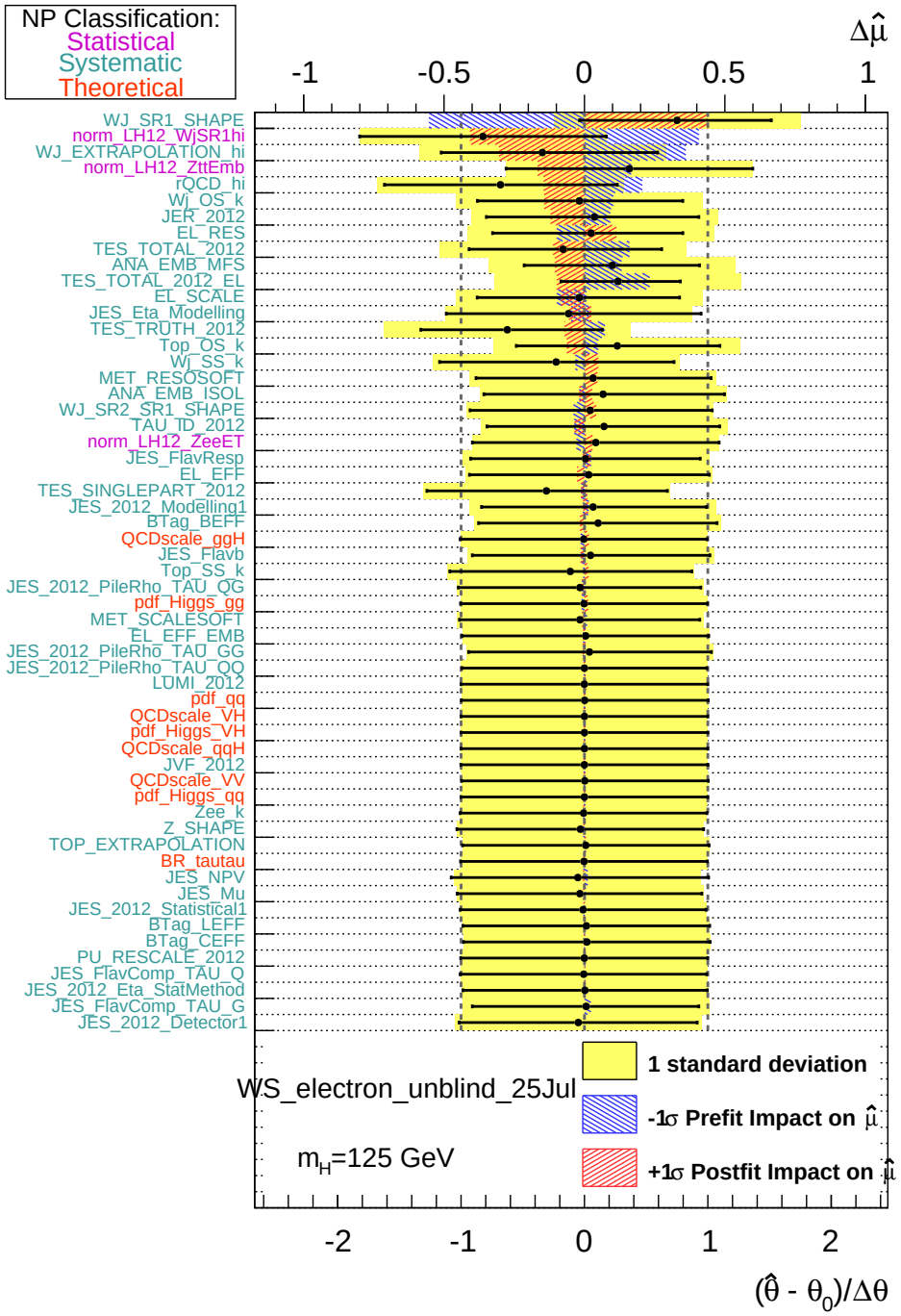


Figure B.3: Profiles of the top six ranked nuisance parameters in the muon measurement.

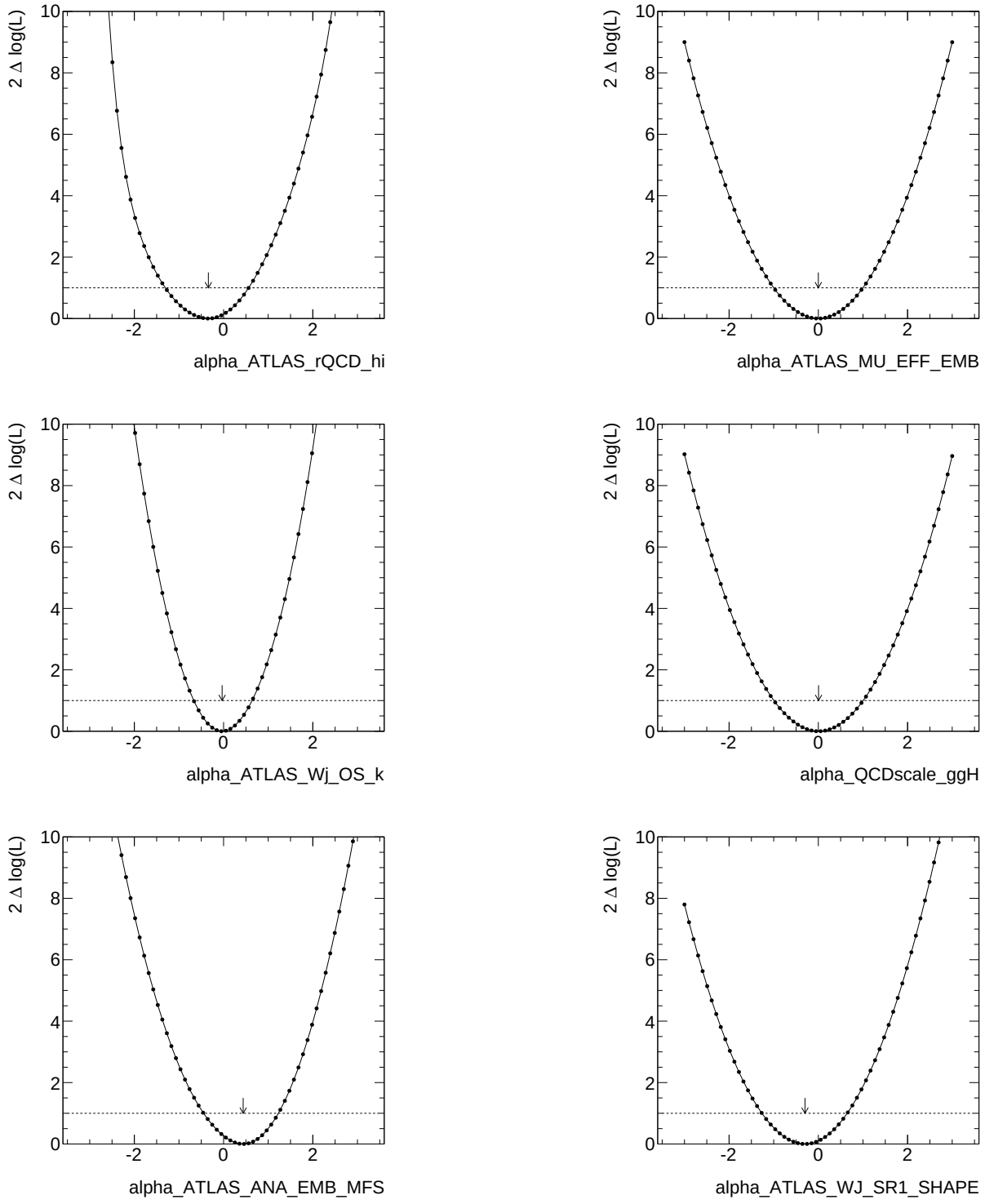


Figure B.4: Profiles of the top six ranked nuisance parameters in the electron measurement.

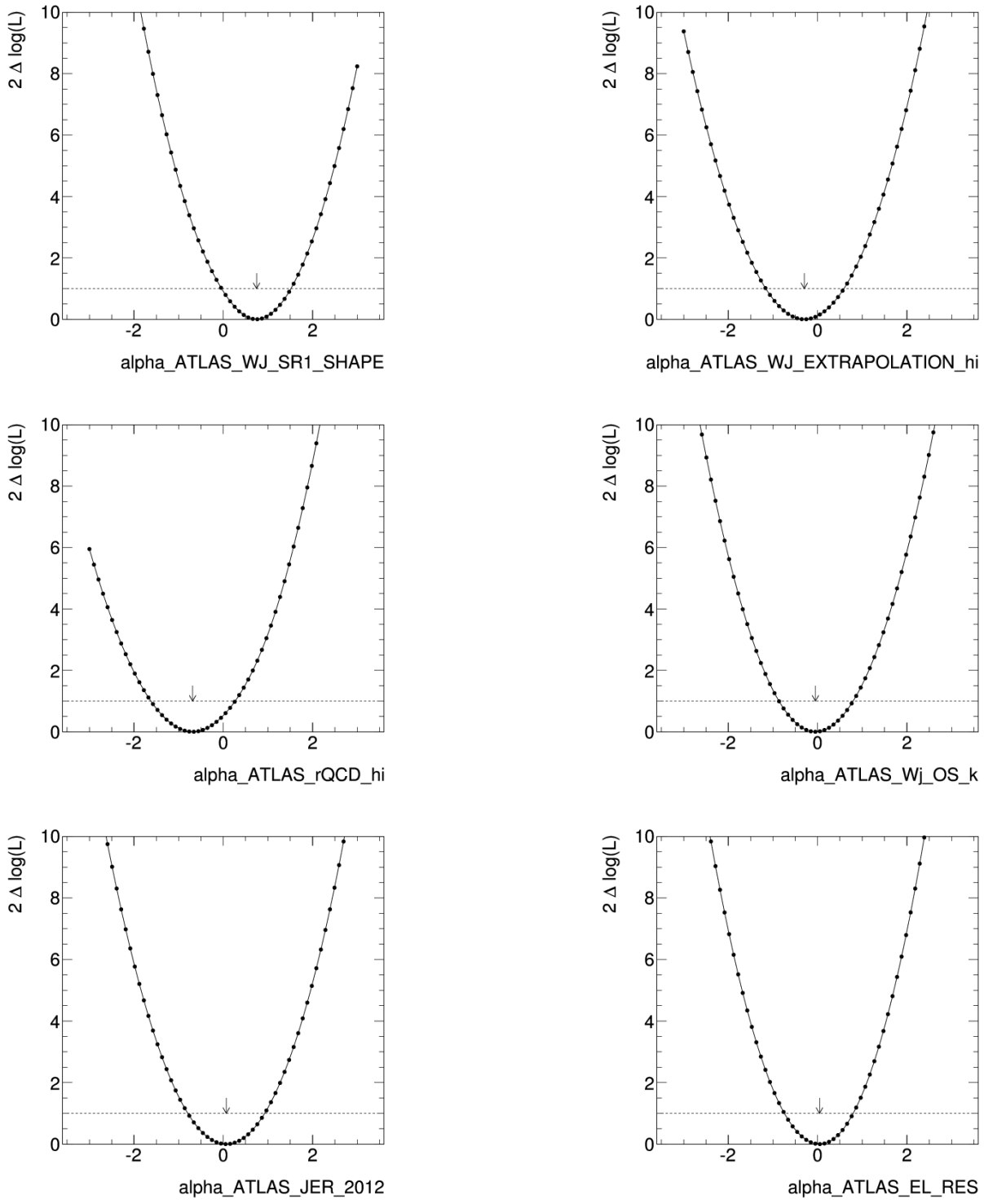
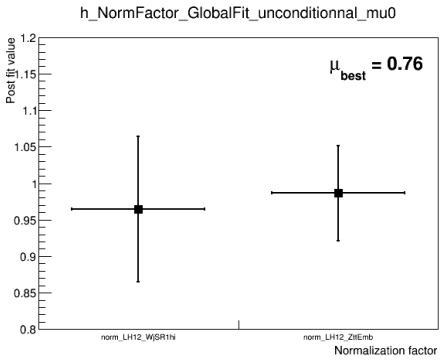


Figure B.5: Fitted values of the `NormFactor` parameters used in the muon and electron measurements. All values are consistent with unity, indicating no underlying issues with the background modeling.

(a) Muon Measurement



(b) Electron Measurement