UNIVERSITY OF CALIFORNIA, SAN DIEGO

Evidence of Higgs to WW to Two Leptons and Two Neutrinos at CMS

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Physics

by

Jae Hyeok Yoo

Committee in charge:

Professor Frank Würthwein, Chair
Professor Claudio Campagnari
Professor Pamela Cosman
Professor Aneesh Manohar
Professor Avraham Yagil

2014

The dissertation of Jae Hyeok Yoo is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2014

DEDICATION

Soli Deo gloria.

TABLE OF CONTENTS

Signature Page iii

Dedication iv

Table of Contents v

List of Figures ix

List of Tables xvii

Acknowledgements xxi

Vita xxiv

Abstract of the Dissertation xxv

Chapter 1 Higgs Boson in Standard Model 1
1.1 Higgs Mechanism 2
1.1.1 How particles become massive : Higgs Mechanism 2
1.2 Production and Decay of Higgs Boson 9
1.2.1 Production of Higgs Boson 9
1.2.2 Decay of Higgs Boson 13
1.3 Limits on Higgs Boson Mass 15
1.3.1 Theoretical Limits 15
1.3.2 Experimental Limits 22
1.4 The $H \to W^+W^- \to 2l2\nu$ channel 24
1.4.1 Large production rate 24
1.4.2 Angular distribution of leptons in the final state . 25
1.4.3 Kinematic variables 27
1.4.4 CMS $H \to W^+W^- \to 2l2\nu$ result before 2012 . . 29

Chapter 2 LHC and CMS Detector 30
2.1 Large Hadron Collider 31
2.2 Compact Muon Solenoid detector 34
2.2.1 Tracker 37
2.2.2 Electromagnetic Calorimeter 39
2.2.3 Hadronic Calorimeter 41
2.2.4 Magnet 44
2.2.5 Muon System 45
2.2.6 Trigger and Data Acquisition 47
2.2.7 CMS computing 47

Chapter 3 Event Reconstruction . . . 52
3.1 Tracks . . . 53
3.2 Event Primary Vertex . . . 58
3.3 Electron . . . 58
3.4 Muon . . . 62
3.5 Jet . . . 64
3.6 Missing Transverse Energy . . . 69
3.7 B-tagging . . . 73

Chapter 4 Event Selection . . . 76
4.1 Trigger . . . 78
4.1.1 Analysis Triggers . . . 79
4.1.2 Utility Triggers . . . 82
4.2 Event Primary Vertex . . . 84
4.3 Electron . . . 84
4.4 Muon . . . 91
4.5 Jet . . . 94
4.6 Missing Transverse Energy . . . 96
4.7 Top-tagging . . . 99
4.8 $Z/\gamma^* \to \ell\ell$ suppression in $ee/\mu\mu$ final states . . . 100
4.9 Additional Selections . . . 103
4.10 WW Selection . . . 103

Chapter 5 Analysis Strategy . . . 105
5.1 Cut-based Method . . . 108
5.2 Shape-based Method . . . 112

Chapter 6 Auxiliary Measurements . . . 123
6.1 Pileup Re-weighting . . . 124
6.2 Lepton Efficiency . . . 126
6.2.1 Electron Selection Efficiency . . . 127
6.2.2 Muon Selection Efficiency . . . 132
6.2.3 Trigger Efficiency . . . 135
6.3 Jet Counting Efficiency . . . 142

Chapter 7 Background Estimation . . . 145
7.1 $Z/\gamma^* \to \ell\ell$ . . . 147
7.2 Jet-induced backgrounds : W + jets . . . 151
7.3 $t\bar{t}/tW$ . . . 160
7.4 WW . . . 165
7.5 $W\gamma^*$ . . . 167
7.6 Other backgrounds . . . 170
7.6.1 $W\gamma$ . . . 170

7.6.2 $Z \to \tau\tau$ . . . 171
7.6.3 WZ/ZZ . . . 172
7.7 The result of background estimation . . . 173

Chapter 8 Statistical Interpretation . . . 175
8.1 Exclusion of SM Higgs boson . . . 176
8.2 Discovery of a New Boson . . . 180

Chapter 9 Systematic Uncertainty . . . 182
9.1 Treatment of systematic uncertainties . . . 183
9.2 Theoretical systematic uncertainties . . . 185
9.2.1 PDF+α_s . . . 185
9.2.2 Missing higher order corrections . . . 186
9.2.3 Parton shower and underlying events . . . 189
9.2.4 Jet Bin Fractions . . . 194
9.3 Instrumental Systematic Uncertainties . . . 198
9.4 Background Estimation Uncertainty . . . 204
9.5 Summary . . . 206

Chapter 10 Fit Validation . . . 207
10.1 Validation of the Fit Model Using Real Data . . . 208
10.1.1 Validation of $qq \to W^+W^-$ fit model . . . 208
10.1.2 Validation of $t\bar{t}/tW$ fit model . . . 213
10.1.3 Validation of W + jets and $W\gamma/W\gamma^*$ fit models . 214
10.2 Post-fit Analysis . . . 214
10.2.1 Post-fit result of nuisances . . . 215
10.2.2 Post-fit Result of Normalization . . . 218

Chapter 11 Results . . . 226
11.1 Cut-based Method results . . . 227
11.2 Shape-based method results . . . 227
11.3 Exclusion limit of SM Higgs boson . . . 232
11.4 Discovery of a new boson . . . 235
11.5 Measurement of Production rate($\sigma \times BR$) . . . 240

Chapter 12 Study on Spin-Parity of the New Boson . . . 244
12.1 Models . . . 245
12.2 Test Method . . . 247
12.3 Results . . . 249
12.4 Conclusion of spin-parity study . . . 251

Chapter 13 Conclusion . . . 253

Appendix A More 2-dimensional templates 259

Bibliography 273

LIST OF FIGURES

Figure 1.1: Feynman diagrams for (a) H-VV and (b) HH-VV interactions. 5
Figure 1.2: Feynman diagrams for (a) H^3 and (b) H^4 interactions. 6
Figure 1.3: Feynman diagram for Hff interaction. 9
Figure 1.4: Feynman diagrams for SM Higgs production, (a) $gg \to H$, (b) $qq \to VH$, (c) $qq \to qqH$ and (d) $gg \to ttH$. 10
Figure 1.5: Standard model Higgs production cross sections as a function of m_{H} at $\sqrt{s} = 8$ TeV for each production mode. 11
Figure 1.6: The ratio of longitudinal polarization of vector bosons as a function of $\frac{m_{\mathrm{H}}}{2m_V}$ [1] 15
Figure 1.7: Standard Model Higgs boson decay branching ratios at low $m_{\mathrm{H}}$and the total width. 16
Figure 1.8: Feynman diagrams for $W_L^+W_L^- \to W_L^+W_L^-$ scattering. 17
Figure 1.9: Feynman diagrams for Higgs boson quartic interaction. Left is tree level and the right three are one-loop correction by Higgs boson itself. 19
Figure 1.10: Upper and lower bound of m_{H} as a function of Λ_{cut}. 20
Figure 1.11: Constraint contour from fine tuning, vacuum stability, and triviality 22
Figure 1.12: Feynman diagram for 1 loop correction by Higgs boson to the W propagator. 23
Figure 1.13: $\Delta\chi^2 = \chi^2 - \chi^2_{min}$ as a function of m_{H} [2]. 24
Figure 1.14: CMS / ATLAS Higgs exclusion with 7 TeV data. 25
Figure 1.15: $\sigma \times BR$ at low m_{H}. 26
Figure 1.16: Spin configuration of Higgs boson decaying to W^+W^- in the CM frame of Higgs. Solid black arrows represent the momentum direction of Ws and grey arrows indicate the spin configuration. 27
Figure 1.17: Kinematic distributions of $p_{\mathrm{T}}^{\ell,\mathrm{max}}$, $p_{\mathrm{T}}^{\ell,\mathrm{min}}$, $\Delta\phi_{\ell\ell}$, $m_{\ell\ell}$ and m_{T}, for $m_{\mathrm{H}} = 110, 125, 145, 160$ and 200 GeV at the generator level. . . 28
Figure 1.18: Exclusion limit of SM Higgs with 2011 data($\sqrt{s} = 7$ TeV, $\mathcal{L} = 4.6$ fb^{-1}). The observed(expected) exclusion limit at CL = 95 % is m_{H}= 129 - 270(127 - 270) GeV. 29

Figure 2.1: Schematic of the CERN accelerator complex. 32
Figure 2.2: (a) The inner structure, (b) cross section and (c) magnetic field map of an LHC cryodipole. 34
Figure 2.3: Integrated luminosity delivered by LHC in 2010, 2011 and 2012 on left. Integrated luminosity delivered by LHC and recored by CMS in 2012 on right. 35
Figure 2.4: Number of interactions per bunch crossing in 2012 run period. 35
Figure 2.5: The layout of the CMS tracker [3]. 38

Figure 2.6: Material budget of CMS tracker in terms of radiation length(X_0) and interaction length(λ_0) [4]. 39
Figure 2.7: Barrel region of ECAL. 41
Figure 2.8: Endcap region of ECAL. 42
Figure 2.9: Energy resolution of ECAL supermodule measured using a test beam [5]. 43
Figure 2.10: Layout of CMS HCAL [6]. 44
Figure 2.11: Layout of a single wedge of HE(left) and HF(right) [5]. 45
Figure 2.12: Magnet system of CMS. Colored pictures are available online [5]. The top shows the yoke(red), outer vacuum tank and the coil. The left bottom shows pictorial view of the magnet system and the right bottom shows the cross section of coils. . 49
Figure 2.13: Muon system [7]. 50
Figure 2.14: Schematic of a superlayer of DT(left) and a tube(right) [8]. . . 50
Figure 2.15: Schematic of CSC. 51
Figure 2.16: Schematic of RPC [9]. 51

Figure 3.1: Tracking efficiency measured using single pion and single muon simulation as a function of p_{T} and η [10]. 57
Figure 3.2: (a) vertex resolution in x and z directions as a function of associated tracks for data and Simulation and (b) vertex reconstruction efficiency as a function of number of associated tracks in jet-enriched data and min-bias data [11]. 59
Figure 3.3: Energy resolution of reconstructed electrons as a function of generated electron energy from different information in simulation [12]. 62
Figure 3.4: Resolution of muon momemtun in $0.0 < |\eta| < 0.2$ and $1.8 < |\eta| < 2.0$. Red, blue and green are global, standalone and tracker muons, respectively [5]. 64
Figure 3.5: Factorized method for jet energy correction. 67
Figure 3.6: JEC uncertainty measured in data of $\mathcal{L} = 11$ fb^{-1} [13]. 68
Figure 3.7: The resolution of parallel(left) and perpendicular(right) components of the hadronic recoil in the data events with Z or photon as a function of reconstructed vertices [14]. 70
Figure 3.8: Distributions of PFMET and trkMET for the Drell-Yan(top) and the signal(bottom) in simulation. The events with the number of reconstructed vertices(N_{vtx}) greater than 20(blue) and less than 5(red) are drawn separately. 72
Figure 3.9: A schematic of b-tagging algorithm. 73
Figure 3.10: Distribution of the discriminating variable for the b-tagging is shown on the left. MC and data shows a good agreement. The b-tagging efficiency(x-axis) and mis-tag rate of udsg jets(y-axis) are shown on the right. 74

Figure 4.1: The cross section × branching ratio($\sigma \times BR$) for the major background processes and the SM m_{H}=125 GeV hypothesis. The branching ratio is for the leptonic decay(electron/muon + neutrino) of W or Z. 77
Figure 4.2: Schematic of $\mathrm{proj} - \mathrm{MET}$. 97
Figure 4.3: MET(black) and $\mathrm{proj} - \mathrm{MET}$(blue) in $Z/\gamma^* \to \tau\tau$(left) and SM Higgs at m_{H}=125 GeV(right). 97
Figure 4.4: $\mathrm{proj} - \mathrm{MET}$ vs. $\mathrm{proj} - \mathrm{trkMET}$ in $Z/\gamma^* \to \ell\ell$ and SM Higgs m_{H}= 125 GeV. 98
Figure 4.5: The BDT score for m_{H}= 125 GeV in red and $Z/\gamma^* \to \ell\ell$ MC in black in 0-jet(left) and 1-jet(right) categories. 102

Figure 5.1: Background composition in cut-based analysis for the m_{H}=125 GeV hypothesis in the four categories. The Higgs selection in addition to WW selection is applied. The signal contribution is overlaid. 107
Figure 5.2: WW-level plots in 0-jet $e\mu/\mu e$ channel with m_{H}=125 GeV signal overlaid. Cuts for m_{H}=125 GeV is shown with blue dotted lines and arrows. 109
Figure 5.3: WW-level plots in 0-jet $e\mu/\mu e$ channel with m_{H}=160 GeV signal overlaid. Cuts for m_{H}=160 GeV is shown with blue dotted lines and arrows. 110
Figure 5.4: WW-level plots in 0-jet $e\mu/\mu e$ channel with m_{H}=200 GeV signal overlaid. Cuts for m_{H}=200 GeV is shown with blue dotted lines and arrows. 111
Figure 5.5: 2D templates for $m_{\mathrm{H}} = 125, 160, 200$ GeV and $qq \to W^+W^-$. For visualization, each bin is divided by the area of the bin in order to avoid random peaks due to difference in the bin size. 114
Figure 5.6: S/B using different binnings. The region($60 < m_{\mathrm{T}} < 120$ GeV, $0 < m_{\ell\ell} < 100$ GeV) is shown. 116
Figure 5.7: Templates on the left and relative statistical uncertainty of the MC sample on the right of $qq \to WH$, $qq \to ZH$, $qq \to qqH$ and $gg \to H$. The templates are for $m_{\mathrm{H}} = 125$ GeV analysis in the 0-jet category. 118
Figure 5.8: Templates on the left and relative statistical uncertainty of the MC sample on the right of $qq \to W^+W^-$, $gg \to W^+W^-$, $t\bar{t}/tW$ and WZ/ZZ. The templates are for $m_{\mathrm{H}} = 125$ GeV analysis in the 0-jet category. 119
Figure 5.9: Templates on the left and relative statistical uncertainty of the MC sample on the right of $\mathrm{W}(\mathrm{e}\nu_{\mathrm{e}}) + \mathrm{jets}$, $\mathrm{W}(\mu\nu_{\mu}) + \mathrm{jets}$ and $W\gamma$. The templates are for $m_{\mathrm{H}} = 125$ GeV analysis in the 0-jet category. 120

Figure 5.10: Templates on the left and relative statistical uncertainty of the MC sample on the right of $W\gamma^*$ and $Z \to \tau\tau$. The templates are for $m_{\mathrm{H}} = 125$ GeV analysis in the 0-jet category. 121

Figure 6.1: Number of reconstructed primary vertices for data and simulation after the re-weighting procedure for the 2012 dataset. . 125
Figure 6.2: Di-lepton mass distributions for data(top) and MC(bottom) in $0.8 < |\eta| < 1.5$ and $10 < p_{\mathrm{T}} < 15$ GeV for events where the probe passes and fails the ID selection on left and right, respectively. 128
Figure 6.3: Di-lepton mass distributions for data(top) and MC(bottom) in $0.8 < |\eta| < 1.5$ and $30 < p_{\mathrm{T}} < 40$ GeV for events where the probe passes and fails the ID selection on left and right, respectively. 129
Figure 6.4: Diagrams for the cases where double-lepton triggers fail. Lepton1(2) denotes offline leptons. 136
Figure 6.5: The number of Jets observed in data (red solid dot) and MC (black line) for the Z events. 143
Figure 6.6: The leading jet p_{T}(left) and the jet veto efficiency as a function of the leading jet p_{T}(right) in data(red solid dot) and MC(black line) for the Z events. 143
Figure 6.7: The fraction of events with 0-jet(left) and 1-jet(right) events as a function of the number of vertices, comparing data (red solid dot) and MC (black solid square) for the Z events. . . . 144

Figure 7.1: Dependency of the data-driven estimations. 148
Figure 7.2: $R_{out/in}$ values as a function of BDT score for 0-jet and 1-jet categories for m_{H}=125 GeV analysis. The black is from data subtracting WZ/ZZ contribution and the red is from MC. The vertical dotted lines show the cut for signal region. 150
Figure 7.3: The p_{T} of denominator(left) and numerator(right) of (a) muon FO and (b) electron FO. The black dots represent data and solid histograms represent contributions from Drell-Yan(green) and W+jets(grey). 155
Figure 7.4: Top shows muon fake rate as a function of FO p_{T} and η, measured with the away jet $p_{\mathrm{T}} > 30$ GeV. Bottom shows electron fake rate as a function of FO p_{T} and η, measured with the away jet $p_{\mathrm{T}} > 35$ GeV. The errors are only statistical. 158
Figure 7.5: The p_{T} of muon FO (left)and electron FO(right). Plots are normalized to the unit area. On each plot, the shape of data after subtracting other backgrounds, and that of W + jets MC are compared. 159

Figure 7.6: Left plot is the electron fake rates projected on p_{T} for different away jet p_{T} thresholds, 20, 35(default) 50 GeV. Right plot is the muon fake rates projected on p_{T} for different away jet p_{T} thresholds, 15, 30(default) and 50 GeV. 159
Figure 7.7: Distribution of b-tagging discriminator for most energetic jet with $p_{\mathrm{T}} > 10$ GeV in the event in the 0-jet(left) and 1-jet(right) in top control regions where top-tagging efficiency is measured. 164
Figure 7.8: $m_{\ell\ell}$ distributions in 0-jet $e\mu/\mu e$ channel for signal events, m_{H} = 125, 200 and 400 GeV and WW after WW selection. The control region, $m_{\ell\ell} > 100$ GeV, is marked with a blue line. . . 166
Figure 7.9: $m_{\ell\ell}$ distribution for opposite-sign muon pairs after $W\gamma^*$ selection. The $W\gamma^*$ is normalized to match the data. 170
Figure 7.10: The probability for a photon to convert to a pair of electrons, and one of the electrons is selected as a good electron. 171
Figure 7.11: $m_{\ell\ell}$ and m_{T} distributions of $W\gamma$ samples. Red is after photon coverts to electrons and black is before the photon conversion with the conversion probability applied. 172

Figure 9.1: $qq \to W^+W^-$ PDF+α_s in the 0-jet category. Top are up and doen 2-dimensional shapes expressed in terms of relative difference from the nominal shape. 187
Figure 9.2: $gg \to W^+W^-$ PDF+α_s in the 0-jet category. Top are up and down 2-dimensional shapes expressed in terms of relative difference from the nominal shape. 188
Figure 9.3: $qq \to W^+W^-$ QCD scale variation in the 0-jet category. Top are up and down 2-dimensional shapes expressed in terms of relative difference from the nominal shape. 190
Figure 9.4: $qq \to W^+W^-$ Madgraph vs. MC@NLO in the 0-jet category. Top are up and down 2-dimensional shapes expressed in terms of relative difference from the nominal shape. 191
Figure 9.5: $t\bar{t}/tW$ Powheg vs. Madgraph in 1-jet category. Top are up and down 2-dimensional shapes expressed in terms of relative difference from the nominal shape. 192
Figure 9.6: Alternate shapes for lepton momentum scale and resolution for $qq \to W^+W^-$ in the 0-jet $e\mu/\mu e$ category. Top are up and down 2-dimensional shapes expressed in terms of relative difference from the nominal shape. 200
Figure 9.7: Alternate shapes for lepton efficiency for $qq \to W^+W^-$ in the 0-jet $e\mu/\mu e$ category. Top are up and down 2-dimensional shapes expressed in terms of relative difference from nominal shape. 201

Figure 9.8: Alternate shapes for MET resolution for $qq \to W^+W^-$ in the 0-jet $e\mu/\mu e$ category. Top are up and down 2-dimensional shapes expressed in terms of relative difference from nominal shape. . 202
Figure 9.9: Alternate shapes for jet energy resolution for $qq \to W^+W^-$ in the 0-jet $e\mu/\mu e$ category. Top are up and down 2-dimensional shapes expressed in terms of relative difference from the nominal shape. 203
Figure 9.10: Alternate shapes for MC statistics for $gg \to W^+W^-$ in the 0-jet $e\mu/\mu e$ category. Top are up and down 2-dimensional shapes expressed in terms of relative difference from nominal shape. . 204
Figure 9.11: Alternate shapes for $\mathrm{W}(\mu\nu_\mu) + \mathrm{jets}$. Top are up and down 2-dimensional down shapes expressed in terms of relative difference from the nominal shape. 205

Figure 10.1: Definition of signal region(SR) and two control regions(CR1 and CR2). 209
Figure 10.2: Comparison of post-fit shapes from the nominal and the full range fits. 210
Figure 10.3: $m_{\ell\ell}$(a,c) and m_{T}(b,d) distributions in CR1 at the top and CR2 at the bottom using the fit results of CR2 and CR1. 212
Figure 10.4: The post-fit $m_{\ell\ell}$ and m_{T} distributions in top-tagged region. . . 214
Figure 10.5: The post-fit $m_{\ell\ell}$ and m_{T} distributions in same-sign region. . . 215
Figure 10.6: Post-fit results on the nuisance parameters. 217

Figure 11.1: m_{T}(left) and $m_{\ell\ell}$(right) distribution for m_{H}=125 GeV analysis in 0-jet category. Top is for $e\mu/\mu e$, middle is for $ee/\mu\mu$, and the bottom is for the inclusive channel. 228
Figure 11.2: m_{T}(left) and $m_{\ell\ell}$(right) distribution for m_{H}=125 GeV analysis in 1-jet category. Top is for $e\mu/\mu e$, middle is for $ee/\mu\mu$, and the bottom is for the inclusive channel. 229
Figure 11.3: 2-dimensional templates of post-fit signal on the left and data subtracted by post backgrounds on the right. The plots show only signal region defined by $60 < m_{\mathrm{T}} < 120$ GeV and $12 < m_{\ell\ell} < 100$ GeV. 233
Figure 11.4: Stacked and data - backgrounds m_{T}(top) and $m_{\ell\ell}$(bottom) distributions using post-fit results of shape-based analysis in $e\mu/\mu e$ final states combining 7 and 8 TeV. 234
Figure 11.5: Exclusion limits of SM Higgs boson combining all categories in 7 and 8 TeV. The top is the result of the cut-based method in all categories. 236
Figure 11.6: Exclusion limit of the second SM-Higgs-like boson considering the SM Higgs at m_{H}=125 GeV as a background. 237

Figure 11.7: Observed and expected significance as a function of m_{H} for the low Higgs mass hypotheses($m_{\mathrm{H}} \leq 200$ GeV). 239
Figure 11.8: The best fit signal strength(μ) as a function of m_{H} for low Higgs mass hypotheses($m_{\mathrm{H}} \leq 200$ GeV). all categories are combined. Cut-based analysis is used in $ee/\mu\mu$ final states and shape-based analysis is used in $e\mu/\mu e$ final states 241
Figure 11.9: $-2\Delta \ln \mathrm{L}$ scan of μ with(black) and without(blue) systematic uncertainty. 242
Figure 11.10: Signal strength(μ) for individual categories for m_{H}=125 GeV. Cut-based analysis is used in $ee/\mu\mu$ final states and shape-based analysis is used in $e\mu/\mu e$ final states. 243

Figure 12.1: The azimuthal angle difference(Φ) between two leptons at the generator level. The red circle is the SM Higgs boson(0^+), the magenta square is the 0^- model, and the blue triangle is the 2^+_{min} model. 246
Figure 12.2: 2-dimensional templates for 0^+, 0^- and 2^+_{min} models. 248
Figure 12.3: Distributions of $q_{2^+_{min}}$ assuming 0^+ and 2^+_{min} models. The blue is the expected distribution assuming 2^+_{min}, and the orange is the expected distribution assuming 0^+ hypothesis. 250
Figure 12.4: Median of $q_{2^+_{min}}$ as a function of $f_{q\bar{q}}$. The blue line is the median of q assuming 2^+_{min}. Left is with $\mu = 1$ and the right is with observed μ. Right plot also shows the q calculated using data. 251
Figure 12.5: Distributions of q_{0^-} assuming 0^+ and 0^- models. Blue is the expected distribution assuming 0^-, and orange is the expected distribution assuming 0^+ hypothesis. 251

Figure 13.1: Expected(left) and observed(right) significance using data collected by July 2012. 255
Figure 13.2: Signal strength measured by the five most sensitive channels. The black line and the green bands shows the combined measurement, $\mu = 0.80 \pm 0.14$. 256
Figure 13.3: Distribution of $q_{2^+_{min}}$ assuming 0^+ and 2^+_{min} models. Blue is the expected distribution assuming 2^+_{min}, and orange is the expected distribution assuming 0^+ hypothesis. The plot shows the result using the best-fit value of the signal strength. 257

Figure A.1: Templates(left) and relative statistical uncertainty of the MC sample(right) of $qq \to WH$, $qq \to ZH$, $qq \to qqH$ and $gg \to H$. The templates are used for $m_{\mathrm{H}} = 125$ GeV analysis in the 1-jet category. 261

Figure A.2: Templates(left) and relative statistical uncertainty of the MC sample(right) of $qq \to W^+W^-$, $gg \to W^+W^-$, $t\bar{t}/tW$ and WZ/ZZ. The templates are used for $m_{\mathrm{H}} = 125$ GeV analysis in the 1-jet category. 262
Figure A.3: Templates(left) and relative statistical uncertainty of the MC sample(right) of $\mathrm{W}(\mathrm{e}\nu_{\mathrm{e}}) + \mathrm{jets}$, $\mathrm{W}(\mu\nu_{\mu}) + \mathrm{jets}$ and $W\gamma$. The templates are used for $m_{\mathrm{H}} = 125$ GeV analysis in the 1-jet category. 263
Figure A.4: Templates(left) and relative statistical uncertainty of the MC sample(right) of $W\gamma^*$ and $Z \to \tau\tau$. The templates are used for $m_{\mathrm{H}} = 125$ GeV analysis in the 1-jet category. 264
Figure A.5: Templates(left) and relative statistical uncertainty of the MC sample(right) of $qq \to qqH$ and $gg \to H$. The templates are used for $m_{\mathrm{H}} = 400$ GeV analysis in the 0-jet category. 265
Figure A.6: Templates(left) and relative statistical uncertainty of the MC sample(right) of $qq \to W^+W^-$, $gg \to W^+W^-$, $t\bar{t}/tW$ and WZ/ZZ. The templates are used for $m_{\mathrm{H}} = 400$ GeV analysis in the 0-jet category. 266
Figure A.7: Templates(left) and relative statistical uncertainty of the MC sample(right) of $\mathrm{W}(\mathrm{e}\nu_{\mathrm{e}}) + \mathrm{jets}$, $\mathrm{W}(\mu\nu_{\mu}) + \mathrm{jets}$ and $W\gamma$. The templates are used for $m_{\mathrm{H}} = 400$ GeV analysis in the 0-jet category. 267
Figure A.8: Templates(left) and relative statistical uncertainty of the MC sample(right) of $W\gamma^*$ and $Z \to \tau\tau$. The templates are used for $m_{\mathrm{H}} = 400$ GeV analysis in the 0-jet category. 268
Figure A.9: Templates(left) and relative statistical uncertainty of the MC sample(right) of $qq \to qqH$ and $gg \to H$. The templates are used for $m_{\mathrm{H}} = 400$ GeV analysis in the 1-jet category. 269
Figure A.10: Templates(left) and relative statistical uncertainty of the MC sample(right) of $qq \to W^+W^-$, $gg \to W^+W^-$, $t\bar{t}/tW$ and WZ/ZZ. The templates are used for $m_{\mathrm{H}} = 400$ GeV analysis in the 1-jet category. 270
Figure A.11: Templates(left) and relative statistical uncertainty of the MC sample(right) of $\mathrm{W}(\mathrm{e}\nu_{\mathrm{e}}) + \mathrm{jets}$, $\mathrm{W}(\mu\nu_{\mu}) + \mathrm{jets}$ and $W\gamma$. The templates are used for $m_{\mathrm{H}} = 400$ GeV analysis in the 1-jet category. 271
Figure A.12: Templates(left) and relative statistical uncertainty of the MC sample(right) of $W\gamma^*$ and $Z \to \tau\tau$. The templates are used for $m_{\mathrm{H}} = 400$ GeV analysis in the 1-jet category. 272

LIST OF TABLES

Table 1.1: $SU(2)_L \otimes U(1)_Y$ quantum numbers. 7
Table 1.2: The highest order of QCD and EW calculations. 12
Table 1.3: $\sigma \times BR$ at $m_H = 125$ GeV for most sensitive channels and the expected number of events in $\mathcal{L}_{int} = 20$ fb^{-1}. l means electrons or muons. 25

Table 2.1: Number of detectors, thickness and mean pitch of each strip, coverage in z direction, and number of layers(disks) of the four part of silicon strip detector, TIB, TOB, TID and TEC. 39

Table 3.1: Parameters used for XY shift correction for MET. 70

Table 4.1: Summary of requirements applied to electrons in the analysis triggers. The selection requirements are shown for electrons in the barrel (endcap). The abbreviation in the names means L=Loose, VL=Very Loose, T=Tight, and VT=Very Tight. . . 80
Table 4.2: Double-lepton triggers used to collect signal events. The naming convention is shown in Table 4.1. 81
Table 4.3: Single-lepton triggers used to collect signal events. The naming convention is shown in Table 4.1. 81
Table 4.4: Utility triggers for fake rate method. The identification and isolation requirements for electrons are described in Table 4.1. Jet30 in the electron triggers means that there should be at least one jet of $p_T > 30$ GeV. 83
Table 4.5: Cut values on BDT score for electron identification. Electrons with BDT score greater than the corresponding values in the table are considered as good electrons. 88
Table 4.6: Effective area used for electron isolation calculation. 90
Table 4.7: Effective areas used for muon isolation. They were calculated with Fall11 MC sample. 94
Table 4.8: Cut values on jet identification BDT scores. The BDT score is required to be greater than these values to be counted as a jet. 96
Table 4.9: Summary of WW selection. [*] For shape-based method $p_T^{\ell\ell} > 30$ GeV is applied and for cut-based method $p_T^{\ell\ell} > 45$ GeV is applied. 104

Table 5.1: m_H-dependent event selection for the cut-based analysis in the 0-jet and 1-jet categories. 112
Table 5.2: Expected significance with different binnings and range combinations. The result is from 0-jet $e\mu/\mu e$ with 12.1 fb^{-1}of data. As a reference, the expected significance from the BDT method is 1.80. 117

Table 5.3: Summary of template parameters. For the high-m_{H} templates, overflow up to $m_{\ell\ell}$=600 GeV and m_{T}=600 GeV is included in the last bin. 122

Table 6.1: Electron ID efficiency in data and simulation, and the scale factors. 130
Table 6.2: Electron ISO efficiency in data and simulation, and the scale factors. 131
Table 6.3: The additional systematic uncertainty δ_{SF} on the scale factor for the electron selection due to the N-1 factorization scheme. . 132
Table 6.4: Muon ID efficiency in data and simulation, and scale factors. . 133
Table 6.5: Muon ISO efficiency in data and simulation, and scale factors. . 134
Table 6.6: Efficiency of dZ requirement of the double-electron trigger. . . . 138
Table 6.7: Efficiency of the leading leg of the double-electron trigger. . . . 139
Table 6.8: Efficiency of the trailing leg of the double-electron trigger. . . . 139
Table 6.9: Efficiency of the single-electron trigger. 140
Table 6.10: Efficiency of dZ requirement of the double-muon trigger. 140
Table 6.11: Efficiency of the leading leg of the double-muon trigger. 141
Table 6.12: Efficiency of the trailing leg of the double-muon trigger. 141
Table 6.13: Efficiency of the single-muon trigger. 142

Table 7.1: Method of background estimation for each background. Major backgrounds are estimated by data-driven methods, and $W\gamma$ and WZ/ZZ are taken from simulation. 146
Table 7.2: The Drell-Yan estimation in the same flavor final state at WW selection level. 151
Table 7.3: The Drell-Yan estimation in $e\mu/\mu e$ channel for the cut-based selections. The dependence of $R_{out/in}$ on m_{H} is due to the m_{H}-dependent selection listed in Table 5.1. 152
Table 7.4: Summary of selection and control region definitions used in top estimation in the different jet bins. Table is taken from [15]. . . 161
Table 7.5: Monte Carlo to data scale factor for the top background contribution for 19.5 fb^{-1}. 165
Table 7.6: The result WW background estimation for cut-based and shape-based analyses. 168
Table 7.7: Expected number of signal and background events from the data-driven methods for at 8 TeV after applying $\mathrm{W^+W^-}$ selection. Only statistical uncertainties are reported. 173
Table 7.8: Expected number of signal and background events from the data-driven methods for at 8 TeV after applying $m_{\mathrm{H}} = 125$ GeV Higgs selection shown in Table 5.1. Total uncertainties including both statistical and systematic uncertainties are reported. . 174

Table 9.1: κ values of systematic uncertainty due to modeling of parton showering and underlying events. 193
Table 9.2: Table of formulae expressing the κ values for the systematic uncertainties on the jet bin fractions due to missing higher order corrections on $\sigma_{\geq 0/1/2\text{-jet}}$ in terms of the κ values for these cross sections, $\kappa_{\geq 0/1/2\text{-jet}}$, and the jet bin fractions, $f_{0/1/2\text{-jet}}$. 195
Table 9.3: κ values for the systematic uncertainties due to missing higher order corrections for the inclusive $gg \to H$ production cross section, $\sigma_{\geq 0\text{-jet}}$, $\sigma_{\geq 1\text{-jet}}$ and $\sigma_{\geq 2\text{-jet}}$. The corresponding κs are $\kappa_{\geq 0\text{-jet}}$, $\kappa_{\geq 1\text{-jet}}$ and $\kappa_{\geq 2\text{-jet}}$, respectively. 196
Table 9.4: Table of κ values for the systematic uncertainties for the jet bin fractions due to missing higher order corrections for the total inclusive Higgs cross section, the inclusive Higgs+1jet cross section, and the inclusive Higgs+2jet cross section. 197
Table 9.5: Table of κ values for the systematic uncertainties for the jet bin fractions due to missing higher order corrections for the total inclusive $qq \to W^+W^-$ cross section, the inclusive $qq \to W^+W^-$+1jet and $qq \to W^+W^-$+2jet cross sections. 198
Table 9.6: Summary of systematic uncertainties and their relative contributions to the expected total systematic uncertainty of signal strength. 206

Table 10.1: Expected signal and background yields in the two control regions(CR1 and CR2) and the full template region(Full range). . 209
Table 10.2: Comparison of normalization from nominal and full range fits. . 211
Table 10.3: Comparison of the best-fit μ values from the full range, CR1 and CR2 fits, where μ is the signal strength. 212
Table 10.4: Exclusion limit on the production rate of $H \to W^+W^- \to 2l2\nu$ at CL_{s}=95 %, significance and the observed(best-fit) signal strength using different central shapes for $qq \to W^+W^-$ process in the shape-based analysis. 213
Table 10.5: The pre-fit and post-fit normalizations in $e\mu/\mu e$ 0-jet category in 8 TeV. 219
Table 10.6: The pre-fit and post-fit normalizations in $e\mu/\mu e$ 1-jet category in 8 TeV. 220
Table 10.7: The pre-fit and post-fit normalizations in $ee/\mu\mu$ 0-jet category in 8 TeV. 221
Table 10.8: The pre-fit and post-fit normalizations in $ee/\mu\mu$ 1-jet category in 8 TeV. 222
Table 10.9: The pre-fit and post-fit normalizations in $e\mu/\mu e$ 0-jet category in 7 TeV. 223
Table 10.10: The pre-fit and post-fit normalizations in $e\mu/\mu e$ 1-jet category in 7 TeV. 224

Table 10.11: The pre-fit and post-fit normalizations in $ee/\mu\mu$ 0-jet category in 7 TeV. 225
Table 10.12: The pre-fit and post-fit normalizations in $ee/\mu\mu$ 1-jet category in 7 TeV. 225

Table 11.1: Table of yields in cut-based analysis at 7 TeV in 4.9 fb^{-1}. All final states are shown separately. Both yields and uncertainties are shown. Z+jets includes $Z \to \tau\tau$. 230
Table 11.2: Table of yields in cut-based analysis at 8 TeV in 19.5 fb^{-1}. All final states are shown separately. Both yields and uncertainties are shown. Z+jets includes $Z \to \tau\tau$. W+jets includes both $\text{W}(\text{e}\nu_{\text{e}}) + \text{jets}$ and $\text{W}(\mu\nu_{\mu}) + \text{jets}$. 231
Table 11.3: Observed and expected significances in 7 TeV. Cut-based analysis is used in $ee/\mu\mu$ final states and shape-based analysis is used in $e\mu/\mu e$ final states 238
Table 11.4: Observed and expected significances in 8 TeV. Cut-based analysis is used in $ee/\mu\mu$ final states and shape-based analysis is used in $e\mu/\mu e$ final states 238
Table 11.5: Observed and expected significances combining 7 TeV and 8 TeV results. Cut-based analysis is used in $ee/\mu\mu$ final states and shape-based analysis is used in $e\mu/\mu e$ final states. 239
Table 11.6: The observed and expected significances using different generators for the $qq \to W^+W^-$ process. Alternative generators, MC@NLO and Powheg, were used replacing the default generator, Madgraph. 240

ACKNOWLEDGEMENTS

First of all, I would like to present my deep gratitude to my advisor, professor Frank Würthwein. He always exceeded my expectation as a good advisor by a large margin and I really feel fortunate to have him as my advisor. Needless to say, he is an excellent physicist who has an exceptional talent to explain complicated things in a simple form. In addition, he is a good person who really cares for his students.

I want to thank professors Avi Yagil and Claudio Campagnari. Avi trained me to be crispier in communication, particularly, when presenting my work to others. Claudio showed me how to draw a big picture when planning a workflow. I was always surprised by his vast knowledge. I really appreciate them for the supports and opportunites they provided to me.

I want to thank my thesis committee members for their time and effort to make my thesis better. Professors Frank Würthwein, Avi Yagil, Claudio Campagnari, Aneesh Manohar and Pamela Cosman.

My gratitudes go to the members of Smurf team, Dmytro(Dima) Kovalskyi, Yanyan Gao, Dave Evans, Giuseppe Cerati, Guillelmo Gomez Ceballos Retuerto and Marco Zanetti, with whom I closely worked for my thesis work(HWW). Dima gave me a lot of opportunities to me so that I can contribute to the analysis. Yanyan and Dave shared their code which I used to do various analyses including the 2D analysis. I shared office with Yanyan during my stay at CERN, and I learned a lot of physics from her, so I call her Shifu(teacher in Chinese). Dave helped me from the first project I worked on in 2010 to the thesis work. Whenever I had problems in analysis or something else including how to fix the battery in Frank's car, I went to him and there came solutions. I started HWW analysis with Giuseppe working on DY MVA. He understood me when I knew nothing and generously helped me get up to speed. Guillelmo is probably the guy whom I bugged most in the Smurf team because we had to crosscheck results whenever results should go out. He is the most amazing guy I met in CMS because of the amount of work he did and how efficiently he managed those works on his plate.

My gratitudes go to the members of Surf'n Turf(SNT) group. Frank Golf

helped me a lot with not only technical issues but also basic knowledge about particle physics when I was a baby in the group. I am glad that we will be sharing office in Broida Hall changing only one letter in the institution name. Warren Andrew, Jacob Ribnik and Puneeth Kalavase helped me with many technical problems in my early years. Yanjun Tu and Sanjay Padhi helped me on my first project in the group, and Yanjun particularly helped me with data processing. My fellow students, Ryan Kelley, Ian MacNaill and Vince Welke, helped me understand American culture and became my English teachers. Ryan and Ian helped me with many coding issues. I have been sharing office with Vince for one and a half years, and it has been a great fun to discuss physics and "beyond" physics issues such as where the best Korean BBQ place in San Diego is. SNT is a huge group and I got helps from everyone in one way or the other. Please forgive me that I did not spell all out.

I want to thank my friends at UCSD and CERN. At UCSD, my fellow physics graduate students, Chi Yung Richard Chim, Yaojun Zhang, Ethan Cho and Xiang Zhai, and the members of Korean Soccer Freak(KSF), particularly, Jong Woo Kim who initiated KSF with me and worked in the same building until now. At CERN, Dong Ho Moon, Mi Hee Jo, Geum Bong Yu, Philip Chang, Sang Eun Lee and Hwi Dong Yoo with whom I shared valuable memories. Without them my life in US and Europe would have been dry and barren.

Last but not least, I want to thank my family. My parents and sister for always trusting my decision and fully supporting whatever I do. Even though they could, they did not demand anything from me, but waited in patience sacrificing many things in their lives. My parent-in-laws for their support and help particularly when my sons were born. My mother-in-law flew over the Pacific Ocean and spent a couple of months to take care of my wife and sons. I feel sorry for not having many opportunites to take her to good places, which she deserved. My brother-in-law for his support and cares. My wife, So Jung Kim, for her endless love and understanding. She had to travel a lot because of me, and so life was always unsettled. But, she understood all these and supported me so that I could concentrate on my work. My two sons, Timothy Jihun Yoo and Elliot Jisuk Yoo,

for the joy and the happiness they brought to our home.

Chapter 7,10, 11, 12 and 13, are reprints of the material as it appears in CMS Collaboration, "Measurement of Higgs boson production and properties in the WW decay channel with leptonic final states", J. High Energy Phys. 01 (2014) 096. The dissertation author was the primary investigator and author of this paper.

Chapter 13, is a reprint of the material as it appears in CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", Phys. Lett. B 716, 30 (2012). The dissertation author was the primary investigator and author of this paper.

VITA

2007	B.S. in Physics, Korea University
2009	M.S. in Physics, Korea University
2014	Ph.D. in Physics, University of California, San Diego

PUBLICATIONS

CMS Collaboration, "Measurement of Higgs boson production and properties in the WW decay channel with leptonic final states ", *J. High Energy Phys.* 01 (2014) 096

CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* 716, 30 (2012)

ABSTRACT OF THE DISSERTATION

Evidence of Higgs to WW to Two Leptons and Two Neutrinos at CMS

by

Jae Hyeok Yoo

Doctor of Philosophy in Physics

University of California, San Diego, 2014

Professor Frank Würthwein, Chair

In this thesis, we report the result on the search for the Standard Model(SM) Higgs boson decaying to a pair of W in the full leptonic final state using data collected by CMS detector at LHC at 7 and 8 TeV. The integrated luminosity is 4.9 fb^{-1} and 19.5 fb^{-1} at 7 and 8 TeV, respectively. The SM Higgs hypothesis is excluded at CL_{s}=95 % in $m_{\mathrm{H}} = 128 - 600$ GeV. An excess of data is observed around $m_{\mathrm{H}} = 125$ GeV which corresponds to significance of 4.0σ on the background-only hypothesis. The measured production rate normalized to the SM prediction is $0.76 \pm 0.13(stat.) \pm 0.16(syst.)$. A hypothesis test on the spin-parity nature of the new boson shows that data prefers SM Higgs boson to the graviton-like spin-2 model or spin-0 pseudo-scalar model.

Chapter 1

Higgs Boson in Standard Model

1.1 Higgs Mechanism

Properties of elementary particles in nature and their interactions(forces) with each other are described by the Standard Model(SM) [16–18] in particle physics. It is based on the gauge symmetry and the group structure of $\mathrm{SU(3)_c} \otimes \mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$, where $\mathrm{SU(3)_c}$, $\mathrm{SU(2)_L}$ and $\mathrm{U(1)_Y}$ describe color, weak iso-spin and hyper charge, respectively. The gauge symmetry requires the weak gauge bosons to be massless, but experimentally we know that Weak gauge bosons, $W^{\pm}$ and Z are massive. The cure for this is the Higgs mechanism [19–24] based on the spontaneous symmetry breaking(SSB) which breaks $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ to $\mathrm{U(1)_{EM}}$. It gives masses to weak bosons but keeps photon massless.

1.1.1 How particles become massive : Higgs Mechanism

Since Electroweak(EWK) theory is based on SU(2) symmetry, the Higgs field is given as a SU(2) doublet,

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{1.1}$$

where each element is a complex field,

$$\phi^+ = \frac{\phi_1 + i\phi_2}{\sqrt{2}} \qquad \text{and} \qquad \phi^0 = \frac{\phi_3 + i\phi_4}{\sqrt{2}}. \tag{1.2}$$

We start with the Higgs Lagrangian to understand the essence of spontaneous symmetry breaking(SSB) before making things more complicated. The full SM Lagrangian will be discussed later. The Higgs Lagrangian($\mathcal{L}_\phi$) is composed of the kinetic and the potential terms.

$$\mathcal{L}_\phi = \underbrace{(\partial_\mu \phi)^\dagger (\partial^\mu \phi)}_{\text{kinetic term}} - \underbrace{\left(\mu^2 \phi^\dagger \phi + \lambda \left(\phi^\dagger \phi\right)^2\right)}_{\text{potential}} \tag{1.3}$$

where $\mu^2 < 0$ and $\lambda > 0$. The potential term which is a function of $\phi^\dagger \phi$ is invariant under SU(2) local gauge transformation,

$$\phi(x) \rightarrow \phi(x)' = e^{i\vec{\alpha}(x)\cdot\frac{\vec{\sigma}}{2}} \phi(x), \tag{1.4}$$

where $\vec{\alpha}(x)$ is a vector of parameters and $\vec{\sigma}$ is a vector of Pauli matrices,

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \text{and} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The potential has the minimum at $\phi^\dagger\phi = -\mu^2/2\lambda = v^2/2$ where v is the vacuum expectation value of the Higgs field ϕ. Due to SU(2) symmetry, the choice of vacuum state is not definite as seen in the following equation,

$$\phi^\dagger\phi = \frac{1}{2}\left(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2\right) = \frac{v^2}{2} \tag{1.5}$$

where there are 4 real variables with only one constraint. This leads to an appropriate choice of vacuum for the physics of interest. We choose the vacuum state, ϕ_0, as

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{1.6}$$

and expand around it by $H(x)$

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \tag{1.7}$$

where $H(x)$ is the physical Higgs field. In order to make the Lagrangian invariant under SU(2) transformation, the derivative ∂_μ should be replaced by the covariant derivative $\mathcal{D}_\mu$,

$$\mathcal{D}_\mu = \partial_\mu - ig_1\frac{Y}{2}B_\mu - ig_2\frac{\vec{\sigma}}{2}\cdot\overrightarrow{W}_\mu. \tag{1.8}$$

B_μ and $\overrightarrow{W}_\mu$ are the vector fields needed for U(1) and SU(2) gauge invariance, respectively. The g_1 and g_2 are the couplings that decide the strength of the interactions associated with B_μ and $\overrightarrow{W}_\mu$. Y(weak hypercharge) and $\vec{\sigma}/2$ are the generators for U(1) and SU(2), respectively. Putting this into the Lagrangian $\mathcal{L}_\phi$, the kinetic term contains

$$\phi^\dagger\left[-ig_1\frac{Y}{2}B^\mu - ig_2\frac{\vec{\sigma}}{2}\cdot\overrightarrow{W}^\mu\right]^\dagger\left[-ig_1\frac{Y}{2}B^\mu - ig_2\frac{\vec{\sigma}}{2}\cdot\overrightarrow{W}^\mu\right]\phi. \tag{1.9}$$

In order to derive the masses of W and Z bosons, we use the vacuum state of Higgs field ϕ_0 because masses are present even without dynamical fields. With $Y = 1$ and $\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}$, and writing explicitly in 2×2 matrices, equation. (1.9) becomes

$$\begin{aligned} &\frac{1}{8} \left| \begin{pmatrix} g_1 B_\mu + g_2 W_\mu^3 & g_2(W_\mu^1 - iW_\mu^2) \\ g_2(W_\mu^1 + iW_\mu^2) & g_1 B_\mu - g_2 W_\mu^3 \end{pmatrix} \begin{pmatrix} 0 \\ v \end{pmatrix} \right|^2 \\ &\qquad = \frac{v^2}{8} \left| \begin{pmatrix} g_2(W_\mu^1 - iW_\mu^2) \\ g_1 B_\mu - g_2 W_\mu^3 \end{pmatrix} \right|^2 \\ &\qquad = \frac{v^2 g_2^2}{8} \left[\left(W_\mu^1\right)^2 + \left(W_\mu^2\right)^2 \right] + \frac{v^2}{8} \left(g_1 B_\mu - g_2 W_\mu^3\right)^2 . \end{aligned} \tag{1.10}$$

The first term can be re-written using charge states, $W^\pm = \frac{1}{\sqrt{2}} \left(W^1 \mp iW^2\right)$,

$$\frac{1}{2} \left(\frac{v g_2}{2}\right)^2 \left[\left(W_\mu^+\right)^2 + \left(W_\mu^-\right)^2 \right]. \tag{1.11}$$

Thus, we have the mass of charged W boson $m_W = \frac{v g_2}{2}$. Now we know that the second term in equation (1.10) should correspond to Z boson because that is the only remaining massive boson. Imposing the same normalization to the mixed field as the unmixed fields, the physical field for Z boson, Z_μ, is given by

$$Z_\mu = \frac{\left(g_1 B_\mu - g_2 W_\mu^3\right)}{\sqrt{g_1^2 + g_2^2}} \tag{1.12}$$

which gives its mass, $m_Z = \frac{v}{2}\sqrt{g_1^2 + g_2^2}$.

There is a massless field orthogonal to Z_μ,

$$A_\mu = \frac{\left(g_1 B_\mu + g_2 W_\mu^3\right)}{\sqrt{g_1^2 + g_2^2}} \tag{1.13}$$

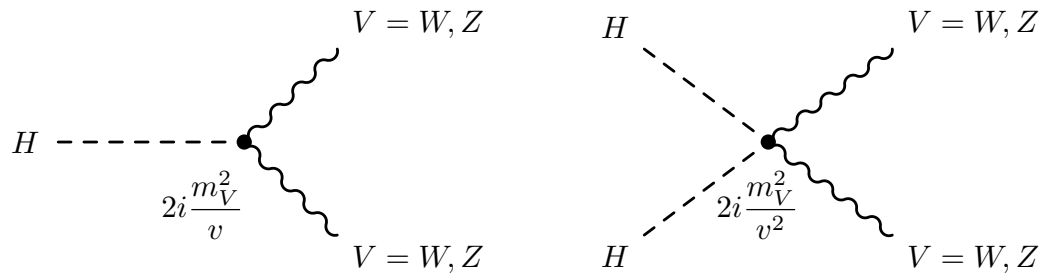
It does not appear in the Lagrangian because its mass therm is zero. This is the field that remains unbroken by SSB. So, it corresponds to the photon.

Re-writing the potential term in equation (1.3) using the physical weak boson states, $W_\mu^\pm$ and Z_μ, and their masses, we have the following terms for inter-

actions between Higgs and weak bosons,

$$2\frac{m_W^2}{v}H(x)W_\mu^+W_\mu^- + \frac{m_Z^2}{v}H(x)\left(Z_\mu\right)^2 + \frac{m_W^2}{v^2}H(x)^2W_\mu^+W_\mu^- + \frac{m_Z^2}{2v^2}H(x)^2\left(Z_\mu\right)^2. \tag{1.14}$$

For Higgs-Weak boson interactions, the couplings are proportional to the square of weak boson mass. The corresponding Feynman diagrams are shown in Figure 1.1.



(a) H-VV interaction (b) HH-VV interaction

Figure 1.1: Feynman diagrams for (a) H-VV and (b) HH-VV interactions.

Considering additional factorials due to identical particles, the vertex factors can be written as $2i\frac{m_V^2}{v}$ and $2i\frac{m_V^2}{v^2}$ for H-VV and HH-VV vertices, respectively, where V denotes W or Z.

After SSB, the Higss potential term, $\mu^2\phi^\dagger\phi + \lambda\left(\phi^\dagger\phi\right)^2$, in the Lagrangian becomes

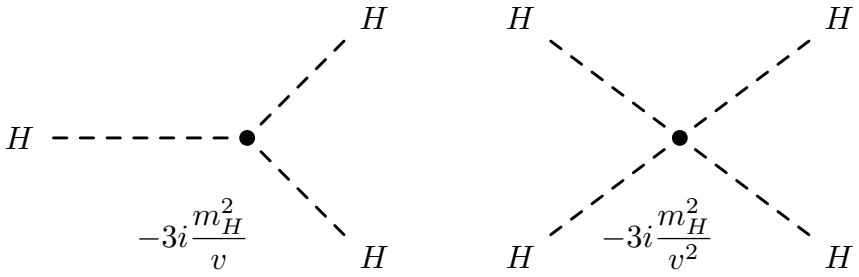
$$\mathcal{L}_{\text{Higgs Potential}} = \mu^2\phi^\dagger\phi + \lambda\left(\phi^\dagger\phi\right)^2 \tag{1.15}$$

$$= \frac{\mu^2}{2}(v+H)^2 + \frac{\lambda}{4}(v+H)^4 \tag{1.16}$$

$$= \ldots - \mu^2H^2 - \frac{\mu^2}{v}H^3 - \frac{\mu^2}{4v^2}H^4 \tag{1.17}$$

where H^0 and H^1 terms are ignored in the last line because they are irrelevant in S-matrix calculations. From equation (1.17), the Higgs mass is identified as $m_{\text{H}}^2 = -2\mu^2$. Using this definition, equation (1.17) becomes

$$\mathcal{L}_{\text{Higgs Potential}} = \ldots - \frac{1}{2}m_{\text{H}}^2H^2 - \frac{m_{\text{H}}^2}{2v}H^3 - \frac{m_{\text{H}}^2}{8v^2}H^4 \tag{1.18}$$



(a) H^3 interaction (b) H^4 interaction

Figure 1.2: Feynman diagrams for (a) H^3 and (b) H^4 interactions.

The corresponding Feynman diagrams are shown in Figure 1.2.

Now we see that the entire Higgs sector depends on only m_{H} and v. The v is calculated by $v = \left(\sqrt{2}\mathrm{G_F}\right)^{-1/2} = 246$ GeV where $\mathrm{G_F}$ is the Fermi constant which is extracted from measurement of muon lifetime. Thus, the SM Higgs sector is fully described by m_{H}. m_{H} is a function of λ and $v(m_{\mathrm{H}}^2 = 2\lambda v^2)$ and we do not know the physical meaning of λ, so the mass of Higgs boson is not predictable by theory. It's experimentalists' task to measure m_{H} and complete the Standard Model of particle physics.

The introduced Higgs field had 4 degrees of freedom, ϕ_1, ϕ_2, ϕ_3 and ϕ_4 before SSB. But, we chose the Higgs field to have only one degree of freedom, $H(x)$. Where did the three go? By breaking $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ to $U(1)_{EM}$, the three gauge bosons acquired masses. This was done by adding longitudinal components to the three gauge bosons. As a result, we have only one physical Higgs field and three massive and one massless gauge bosons, instead of four unphysical Higgs fields and four massless gauge bosons.

The fermions acquire their masses by interacting with Higgs field. Let's start a discussion with leptons because the absence of right-handed neutrinos, *i.e.* neutrinos are massless, makes the case simpler than quarks which have both right-handed and left-handed polarizations. Table 1.1 shows the quantum numbers of $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ for the left-handed electron doublet, right-handed neutrino, right-

Table 1.1: $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ quantum numbers.

	T_3	Y
$\begin{pmatrix} \nu_L \\ e_L \end{pmatrix}$	$\frac{1}{2}$	-1
ν_R	0	0
e_R	0	-2
$\begin{pmatrix} \phi_+ \\ \phi_0 \end{pmatrix}$	$\frac{1}{2}$	-1

handed electron and Higgs doublet. Electron can be replaced by muon or tau leptons. From the table, one can see that the interaction such as

$$e_R + \begin{pmatrix} \phi_+ \\ \phi_0 \end{pmatrix} \rightarrow \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} \tag{1.19}$$

conserves quantum numbers. Now the structure of the interaction is given, and we specify its strength with g_e. Including the hermitian conjugate to the Lagrangian, the lepton-Higgs interaction term becomes

$$\mathcal{L}_{int,lepton} = -g_e \left[\begin{pmatrix} \bar{\nu}_L & \bar{e}_L \end{pmatrix} \begin{pmatrix} \phi_+ \\ \phi_0 \end{pmatrix} e_R + \bar{e}_R \begin{pmatrix} \bar{\phi}_+ & \bar{\phi}_0 \end{pmatrix} \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} \right] \tag{1.20}$$

Using the chosen Higgs field in equation (1.7), the Lagrangian is calculated to be

$$\mathcal{L}_{int,lepton} = -\frac{g_e v}{\sqrt{2}} \left(\bar{e}_L e_R + \bar{e}_R e_L \right) - \frac{g_e H}{\sqrt{2}} \left(\bar{e}_L e_R + \bar{e}_R e_L \right). \tag{1.21}$$

Since $\bar{e}e = \bar{e}(P_L^2 + P_R^2)e = \bar{e}_L e_R + \bar{e}_R e_L$ where P_L and P_R are projection operators, the first term, $-\frac{g_e v}{\sqrt{2}}\bar{e}e$, corresponds to the mass term for the electron. Thus, the mass is identified to be

$$m_e = \frac{g_e v}{\sqrt{2}}. \tag{1.22}$$

Rewriting the Lagrangian in terms of m_e instead of an arbitrary g_e, we get

$$\mathcal{L}_{int} = -m_e \bar{e}e - \frac{m_e}{v}\bar{e}eH. \tag{1.23}$$

Since there isn't a physical motivation for g_e, m_e is not calculable by theory, but needs to be determined by experiments. The second term corresponds to lepton-Higgs interaction. The size of the interaction is proportional to the mass of electrons. Thus, light leptons have very weak couplings to the Higgs field. For example, electron has $m_e/v = 0.5$ MeV/246 GeV $\sim \mathcal{O}(10^{-6})$ and muon has $m_\mu/v = 106$ MeV/246 GeV $\sim \mathcal{O}(10^{-3})$.

The case for quarks is more complicated due to the presence of right-handed up-type quarks as opposed to the lepton case. In order to generate masses for up-type quarks, we need a new Higgs doublet

$$\phi_c = i\sigma_2\phi^* = \begin{pmatrix} \phi_0^* \\ -\phi_- \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} v + H(x) \\ 0 \end{pmatrix}. \tag{1.24}$$

The new Higgs field is invariant under $\mathrm{SU(2)_L}$ transformation and has Y = -1.

$$\begin{aligned} \mathcal{L}_{int,quark} &= -g_d \left[\begin{pmatrix} \bar{u}_L & \bar{d}_L \end{pmatrix} \begin{pmatrix} \phi_+ \\ \phi_0 \end{pmatrix} d_R \right] \\ &\quad -g_u \left[\begin{pmatrix} \bar{u}_L & \bar{d}_L \end{pmatrix} \begin{pmatrix} \phi_0^* \\ -\phi_+ \end{pmatrix} u_R \right] + h.c. \end{aligned} \tag{1.25}$$

$$= -\frac{g_d v}{\sqrt{2}}\bar{d}d - \frac{g_d H}{\sqrt{2}}\bar{d}d - \frac{g_u v}{\sqrt{2}}\bar{u}u - \frac{g_u H}{\sqrt{2}}\bar{u}u \tag{1.26}$$

$$= -m_d\bar{d}d - \frac{m_d}{v}\bar{d}dH - m_u\bar{u}u - \frac{m_u}{v}\bar{u}uH \tag{1.27}$$

where u and d are up-type and down-type quarks, respectively, and

$$m_d = \frac{g_d v}{\sqrt{2}} \quad \text{and} \quad m_u = \frac{g_u v}{\sqrt{2}} \tag{1.28}$$

are used as the lepton case. The h.c. means hermitian conjugate of the first two terms.

As a result, the strength of interaction depends on the fermion mass, m_f/v. Figure 1.3 shows Feynman diagram for Higgs-fermion interaction.

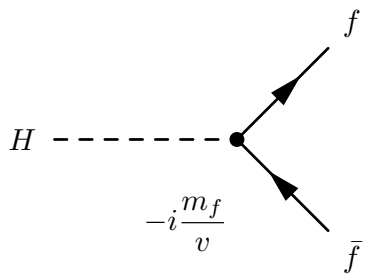


Figure 1.3: Feynman diagram for Hff interaction.

1.2 Production and Decay of Higgs Boson

1.2.1 Production of Higgs Boson

At LHC, the Standard Model Higgs boson is generated by 4 major processes, gluon-gluon fusion ($gg \to H$), vector boson fusion ($qq \to qqH$), associated production with vector bosons ($qq \to VH$) and associated production with heavy quarks($gg \to ttH$). The corresponding Fenyman diagrams are shown in Figure 1.4. Since H does not couple to gluon in ggH process it is produced via a loop of heavy quarks, *i.e.,* top and bottom quarks. At LHC $gg \to H$ has the largest production rate because gluons have the largest probability to collide which is given by parton distribution function(PDF), and top quark has very large coupling to Higgs boson.

At the hadron colliders the hadronic cross section(σ) is calculated with parton-level cross-section ($\hat{\sigma}$) convoluted with PDF,

$$\sigma(pp \to H + X) = \sum_{i,j} \int dx_1 dx_2 f_i(x_1) f_j(x_2) \hat{\sigma}\,(ij \to H + X) \tag{1.29}$$

where i and j are colliding partons, x_1 and x_2 are the longitudinal momentum fractions carried by parton i and j. Each component in the equation is subject to the following uncertainties. The partonic cross section is calculated at a given renormalization scale(μ_R) and factorization scale(μ_F). Due to possible missing higher-order QCD radiative corrections, the uncertainty is estimated by varying the scales around the central values. In the de Florian and Grazzini (dFG) calculation [ref], the central values are chosen to be $\mu_0 = m_{\mathrm{H}}$. The scales μ_R and μ_F are

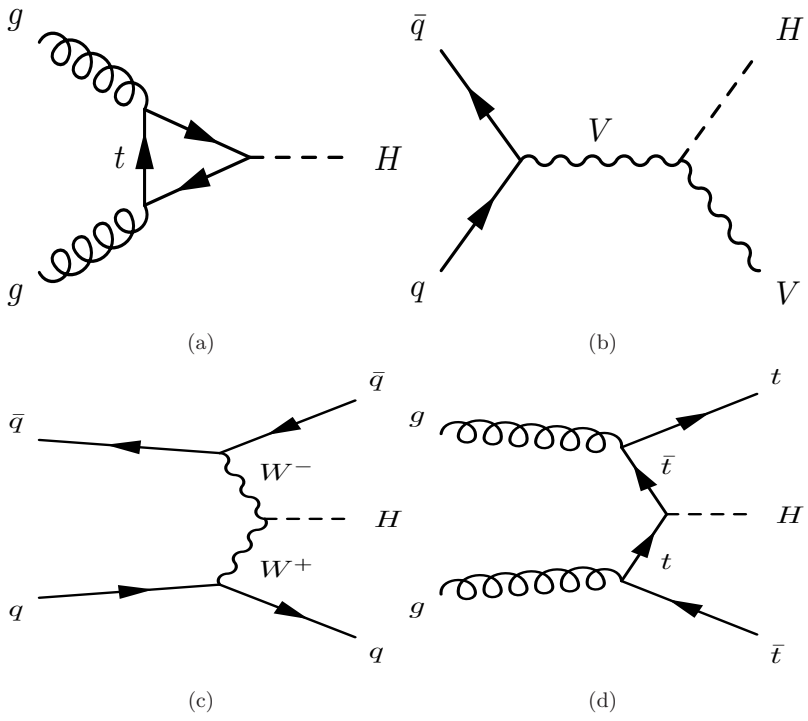


Figure 1.4: Feynman diagrams for SM Higgs production, (a) $gg \to H$, (b) $qq \to VH$, (c) $qq \to qqH$ and (d) $gg \to ttH$.

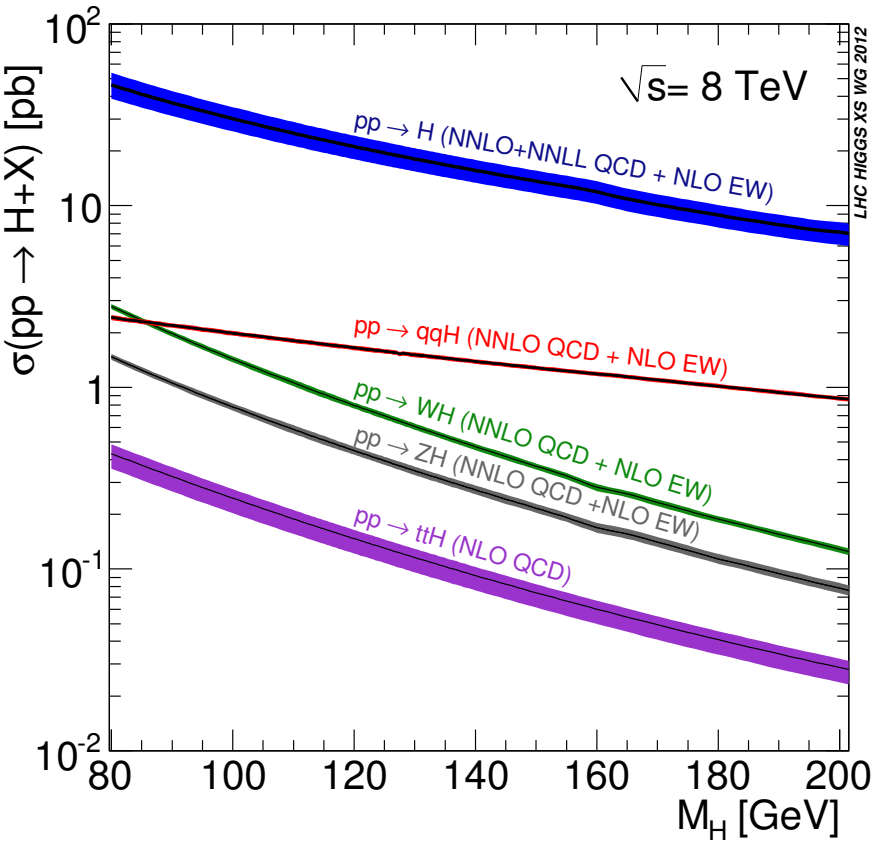


Figure 1.5: Standard model Higgs production cross sections as a function of m_{H} at $\sqrt{s} = 8$ TeV for each production mode.

varied in the range $\mu_0/2 < \mu_F, \mu_R < 2\mu_0$ with a constraint $1/2 < \mu_F/\mu_R < 2$. PDF is obtained by fitting on data measured in deep-inelastic scattering, Drell-Yan, and jet production from a wide variety of different experiments. The accuracy on those data can introduce uncertainty on PDF calculation. In addition, strong coupling constant α_s is used in DGLAP evolution [25–27] to the higher Q^2 region. Thus, its uncertainty also contributes to the total cross section. There are other uncertainties due to EWK corrections, the different choice of top and bottom quark masses, and the use of large-m_T method. But, the effect of these uncertainties to the hadronic cross section is less than a few percent [28] for $gg \to H$.

Figure 1.5 shows the hadronic cross sections as a function of m_{H} for SM Higgs production and uncertainty in different production modes. The $gg \to H$ and $qq \to qqH$ processes are calculated in complex-pole-scheme(CPS) [29], while other

WH/ZH and ttH processes are calculated in zero-width-approximation(ZWA) [30]. The highest order of QCD and EW calculations are summarized in Table 1.2. The

Table 1.2: The highest order of QCD and EW calculations.

process	QCD	EW
$pp \to H$	NNLO	NLO
$pp \to qqH$	NNLO	NLO
$pp \to WH$	NNLO	NLO
$pp \to ZH$	NNLO	NLO
$pp \to ttH$	NLO	

uncertainty is a linear combination of uncertainties from QCD scale variation and PDF+α_S. At $m_{\mathrm{H}} = 125$ GeV $gg \to H$ contributes $\sim 87\%$ to the total cross section.

1.2.2 Decay of Higgs Boson

Figure 1.1 shows that the Higgs boson can couple to a pair of weak bosons. Thus, Higgs can directly decay into W^+W^- and ZZ. Depending on m_{H}, one or two of the bosons can be off-shell. As shown in [31], at $m_{\mathrm{H}} > 2m_W$ both bosons tend to be on-shell, and at $110 < m_{\mathrm{H}} < 2m_W$ one of them is on-shell while the other is off-shell. Below 110 GeV there is a sizable contribution from the case where both bosons are off-shell, but this is not relevant to this analysis. Thus, we discuss two cases where both of them are on-shell(VV), and one is on-shell and the other is off-shell(VV^*).

Both bosons are on shell : $H \to VV$

When both bosons are on-shell($H \to VV$), the decay width at tree-level is given by [1]

$$\Gamma\left(H \to VV\right) = \frac{\mathrm{G_F} m_{\mathrm{H}}^3}{16\sqrt{2\pi}} \delta_V \sqrt{1-4\epsilon^2}\left(1-4\epsilon^2+12\epsilon^4\right) \tag{1.30}$$

where $\epsilon = \dfrac{m_V}{m_{\mathrm{H}}}$ and $\delta_W = 2$ and $\delta_Z = 1$. At high m_{H}, the decay width to WW is reduced to

$$\begin{aligned} \Gamma\left(H \to WW\right) &\to \frac{\mathrm{G_F} m_{\mathrm{H}}^3}{16\sqrt{2\pi}} \times 2 & (1.31)\\ &= 2\frac{1.16637 \times 10^{-5}\ \mathrm{GeV}^{-2} m_{\mathrm{H}}^3}{16\sqrt{2\pi}} & (1.32)\\ &\approx 0.33 m_{\mathrm{H}} \times \frac{m_{\mathrm{H}}^2}{\mathrm{TeV}^2}(\mathrm{TeV}). & (1.33) \end{aligned}$$

For example, at $m_{\mathrm{H}} = 1$ TeV, decay width for WW is 0.33 TeV. Practically, it is hard to claim a Higgs resonance at high mass regions.

The ratio of longitudinal polarization is given by [1]

$$R_L = \frac{\Gamma_L}{\Gamma_T + \Gamma_L} = \frac{1-4\epsilon^2+4\epsilon^4}{1-4\epsilon^2+12\epsilon^4} \xrightarrow{\epsilon \to 0} 1. \tag{1.34}$$

Thus, vector bosons are longitudinally polarized at high m_{H} ($\epsilon \to 0$). At the production threshold, $m_{\mathrm{H}} = 2m_V$ ($\epsilon = 1/2$), R_L is 1/3 which means that the longitudinal and the transverse polarizations are populated with a ratio of 1:2. This is shown in the Figure 1.6.

One boson is on-shell and other is off-shell : $H \to VV^* \to Vf\bar{f}$

When one boson is on-shell and the other is off-shell ($H \to VV^* \to Vf\bar{f}$), the decay width at tree-level is given by [1]

$$\Gamma\left(H \to VV^*\right) = \frac{3{G_F}^2 m_V^4}{16\pi^3} m_H \delta'_V R(\epsilon) \tag{1.35}$$

where $\delta'_W = 1$, $\delta'_Z = \frac{7}{12} - \frac{10}{9}\sin^2\theta_W + \frac{40}{9}\sin^4\theta_W$ and

$$\begin{aligned} R(\epsilon) = & \frac{3(1-8\epsilon^2+20\epsilon^4)}{(4\epsilon^2-1)^{1/2}} \arccos\left(\frac{3\epsilon^2-1}{2\epsilon^3}\right) \\ & -(1-\epsilon^2)\left(\frac{47}{2}\epsilon^2 - \frac{13}{2} + \frac{1}{\epsilon^2}\right) - 3(1-6\epsilon^2+4\epsilon^4)\ln\epsilon. \end{aligned} \tag{1.36}$$

The ratio of longitudinal polarization is given [1] by

$$R_L = \frac{\Gamma_L}{\Gamma_T + \Gamma_L} = \frac{R_L(\epsilon)}{R(\epsilon)} \tag{1.37}$$

where R_L is [1]

$$\begin{aligned} R_L(\epsilon) = & \frac{3(1-16\epsilon^2+20\epsilon^4)}{(4\epsilon^2-1)^{1/2}} \arccos\left(\frac{3\epsilon^2-1}{2\epsilon^3}\right) \\ & -(1-\epsilon^2)\left(\frac{15}{2}\epsilon^2 - \frac{13}{2} + \frac{1}{\epsilon^2}\right) - (3-10\epsilon^2+4\epsilon^4)\ln\epsilon. \end{aligned} \tag{1.38}$$

Figure 1.6 shows the fraction of longitudinal $VV(*)$ decays as a function of $m_H/2m_V(=\frac{\epsilon}{2})$. As discussed in the VV decay, the fraction goes to 1 at high m_H and to 1/3 at the production threshold, $m_H = 2m_V$. Below the threshold, the fraction of longitudinal polarization increases up to 0.6 at $m_H/2m_V \approx 0.8 \to m_H \approx 130$ GeV, and decreases slightly as m_H becomes smaller. As the fraction of decay to longitudinally polarized bosons depends on m_H, and the event kinematics depends on the fraction as well, it is important to optimize analysis for a given m_H to maximize sensitivity.

Figure 1.7 shows the branching ratios of Standard Model Higgs boson(left) and its total decay width(right). The partial widths of each decay mode are calculated with the next-leading-order(NLO) QCD and electron corrections, and all interferences at LO and NLO are taken into account [32]. The band on the left plots shows the uncertainty due to missing higher order terms.

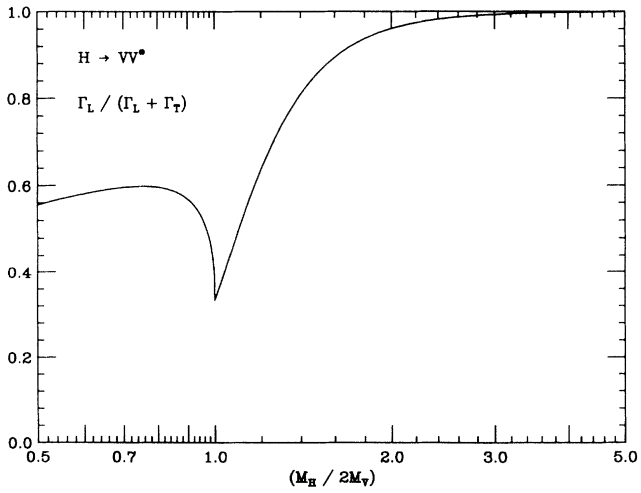


Figure 1.6: The ratio of longitudinal polarization of vector bosons as a function of $\frac{m_{\mathrm{H}}}{2m_V}$ [1]

1.3 Limits on Higgs Boson Mass

The Higgs mass is constrained by various theoretical limits and experimental measurements. This section discusses theoretical limits by perturbative unitarity, triviality and stability, and fine tuning. Then it discusses the experimental limits from indirect and direct measurements as of 2011.

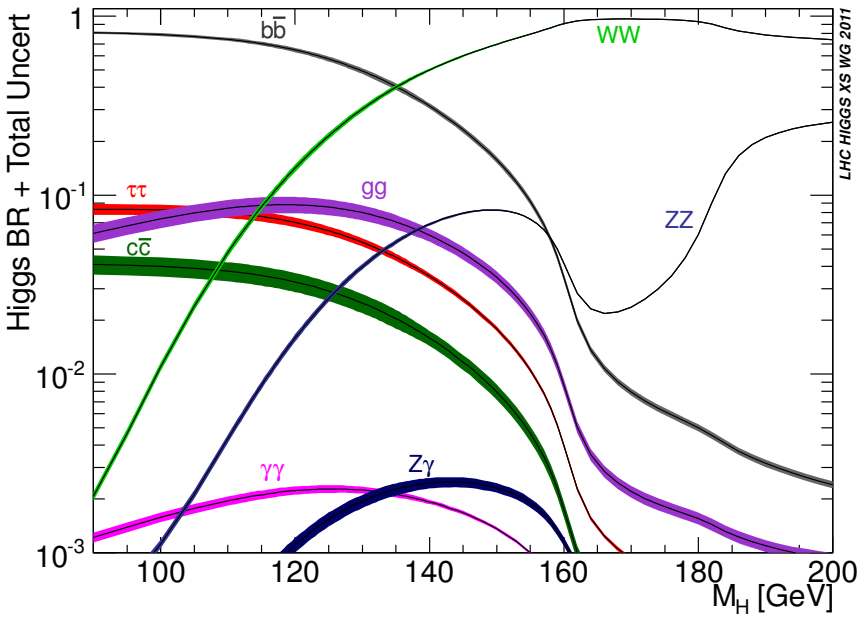
1.3.1 Theoretical Limits

Perturbative Unitarity : $W_L^+W_L^- \to W_L^+W_L^-$

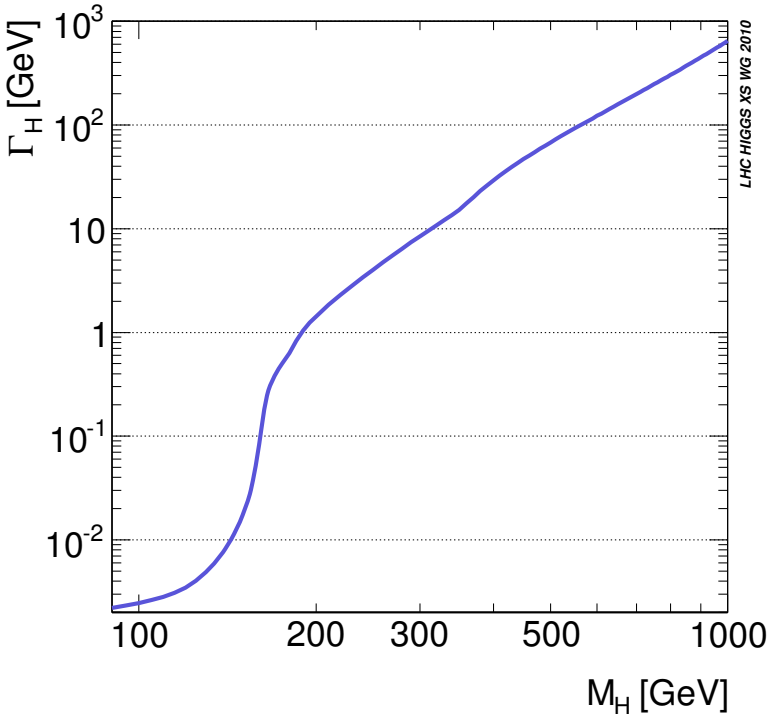
The cross section of longitudinal vector boson scattering, $V_LV_L \to V_LV_L$, increases as the energy increases, and it eventually violates unitarity, *i.e.*, probability for this process to happen is larger than unity.

Figure 1.8 shows Feynman diagrams for $W_L^+W_L^- \to W_L^+W_L^-$. There are 4-point interaction, and t-channel and s-channel interactions mediated by Z/γ^* or Higgs boson. In the high energy limit $s \gg m_{\mathrm{W}}^2$, the scattering amplitude becomes [31]

$$\mathcal{A}\left(W_L^+W_L^- \to W_L^+W_L^-\right) \sim -\frac{1}{v^2}\left(-s-t+\frac{s^2}{s-m_{\mathrm{H}}^2}+\frac{t^2}{t-m_{\mathrm{H}}^2}\right). \qquad (1.39)$$



(a) Branching ratios



(b) Total width

Figure 1.7: Standard Model Higgs boson decay branching ratios at low $m_{\mathrm{H}}$and the total width.

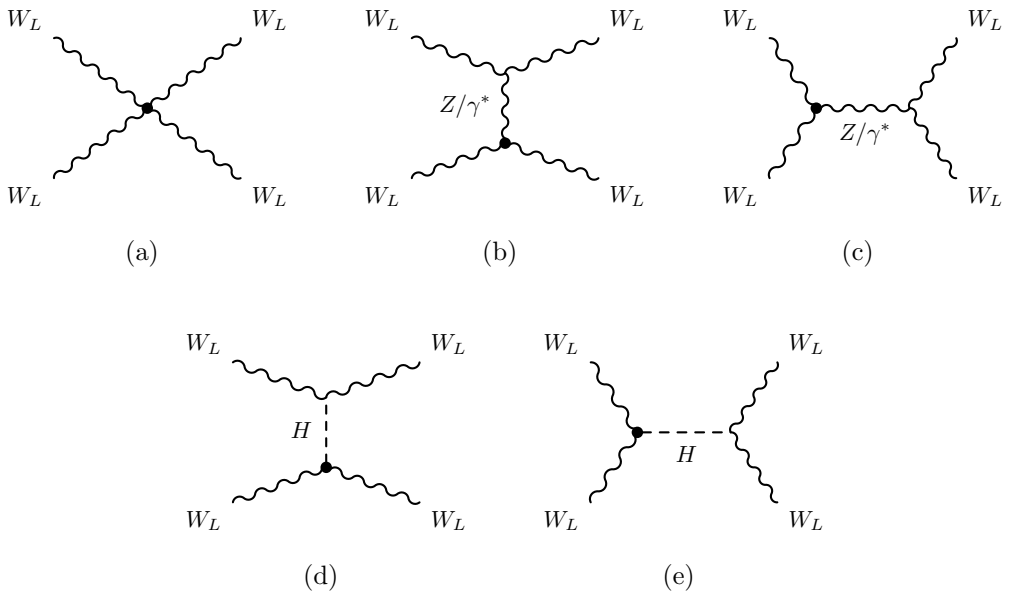


Figure 1.8: Feynman diagrams for $W_L^+W_L^- \to W_L^+W_L^-$ scattering.

According to the Electroweak Equivalence Theorem [33], at very high energy the longitudinal vector bosons can be replaced by their associated Goldstone bosons. Thus, the scattering amplitude can be written using Goldstone bosons ($w^{\pm}$)

$$\mathcal{A}\left(w^+w^- \to w^+w^-\right) = -\frac{m_{\mathrm{H}}^2}{v^2}\left(2 + \frac{m_{\mathrm{H}}^2}{s - m_{\mathrm{H}}^2} + \frac{m_{\mathrm{H}}^2}{t - m_{\mathrm{H}}^2}\right). \tag{1.40}$$

An scattering amplitude can be decomposed into partial waves a_l

$$\mathcal{A} = 16\pi \sum_{i=0}^{\infty} (2l+1)\, P_l\left(\cos\theta\right) a_l \tag{1.41}$$

where P_l is the Legendre polynomials and θ is the scattering angle. For the cross section of $2 \to 2$ processes, we have the following identity on the cross section (σ) by Optical theorem [**?**]

$$\sigma = \frac{16\pi}{s} \sum_{i=0}^{\infty} (2l+1)\left|a_l\right|^2 = \frac{1}{s} Im\left[\mathcal{A}\left(\theta = 0\right)\right] \tag{1.42}$$

which gives the unitary condition,

$$\left|a_l\right|^2 = Im\left(a_l\right) \qquad \Rightarrow \qquad Re\left(a_l\right)^2 + \left[Im\left(a_l\right) - \frac{1}{2}\right]^2 = \left(\frac{1}{2}\right)^2 \tag{1.43}$$

$$\Rightarrow \qquad \left|Re\left(a_l\right)\right| < \frac{1}{2}. \tag{1.44}$$

Then, the $l = 0$ amplitude in the limit of $s \gg m_{\rm H}^2$ becomes

$$a_0\left(w^+w^- \to w^+w^-\right) = -\frac{m_{\rm H}^2}{16\pi v^2}\left[2 + \frac{m_{\rm H}^2}{s - m_{\rm H}^2} - \frac{m_{\rm H}^2}{s}\log\left(1 + \frac{s}{m_{\rm H}^2}\right)\right] \quad (1.45)$$

$$\to -\frac{m_{\rm H}^2}{8\pi v^2} \quad (1.46)$$

The unitary condition (equation (1.44)) gives upper bound on $m_{\rm H}$,

$$|Re\left(a_0\right)| = \frac{m_{\rm H}^2}{8\pi v^2} < \frac{1}{2} \quad (1.47)$$

$$\Rightarrow \quad m_{\rm H} < 2\sqrt{\pi}v \simeq 870 \text{ GeV}. \quad (1.48)$$

Including other scattering channels,

$$Z_LZ_L,\ \ HH,\ \ Z_LH,\ \ W_L^+H,\ \ W_L^+Z_L \quad (1.49)$$

the constraint on $m_{\rm H}$ becomes tighter as [31],

$$m_{\rm H} < 710 \text{ GeV}. \quad (1.50)$$

This means that in the standard model unitariry will be violated if $m_{\rm H} > 710$ GeV unless there is a new physics that cures this problem.

So far we calculated only tree-level terms, so we can expect that adding higher order terms can solve this problem. But, including higher order terms does not guarantee that the unitary will be restored because in the high $m_{\rm H}$ regime coupling to Higgs is too large and perturbative calculation breaks down. Thus, the mass bound given in equation (1.50) can be considered the $m_{\rm H}$ regime where perturbative calculation is reliable in all s.

Triviality and Stability bounds

The variation of the Higgs quartic coupling λ is described by Renormalization Group Equation (RGE). When we consider one-loop radiation corrections by Higgs boson itself to λ as shown in Figure 1.9, the corresponding RGE is given by [31]

$$\frac{d}{dQ^2}\lambda(Q^2) = \frac{3}{4\pi^2}\lambda^2(Q^2) + \text{higher orders} \quad (1.51)$$

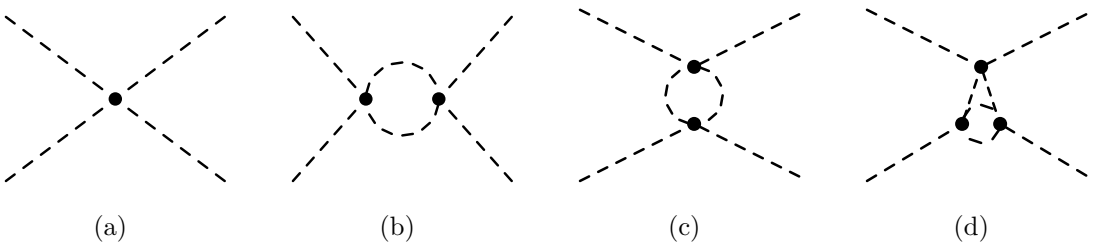


Figure 1.9: Feynman diagrams for Higgs boson quartic interaction. Left is tree level and the right three are one-loop correction by Higgs boson itself.

The solution to this equation is given by

$$\lambda(Q^2) = \frac{\lambda(v^2)}{\left[1 - \frac{3}{4\pi^2}\lambda(v^2)\log\frac{Q^2}{v^2}\right]} \tag{1.52}$$

where the EWSB scale is used as a reference energy point, $Q_0 = v$.

If the energy is much smaller than the EWSB scale, $Q^2 \ll v^2$, the quadratic coupling goes to 0, and the theory is called "trivial", which means that there is no interaction. On the other hand, if the energy is much larger than the EWSB scale, $Q^2 \gg v^2$, as Q increases the coupling will be infinite at a certain energy scale, Λ_{cut}. Using $\lambda = m_{\rm H}^2/2v^2$ and the definition of λ that it is positive, we have the following equation for denominator,

$$1 > \frac{3}{4\pi^2}\frac{m_{\rm H}^2}{2v^2}\log\frac{\Lambda_{cut}^2}{v^2} \qquad \Rightarrow \qquad m_{\rm H}^2 > \frac{8\pi^2 v^2}{\log\frac{\Lambda_{cut}^2}{v^2}}, \tag{1.53}$$

which gives a scale-dependent bound on $m_{\rm H}$. Imposing $\Lambda_{cut} = m_{\rm H}$ which means that the theory is not reliable, *i.e.*, valid scale of the theory is same as the mass of a particle, the bound on the Higgs mass is $m_{\rm H} < 640$ GeV. This result is consistent with the limit from unitarity constraint.

In the previous discussion, only one-loop correction by Higgs itself is considered. This is a proper approximation when λ is large, but when λ is small, we need to consider the contributions from fermions and vector bosons. Since, the strength of interaction with Higgs boson is proportional to the particle mass, we consider only heavy particles, vector bosons and top quarks.

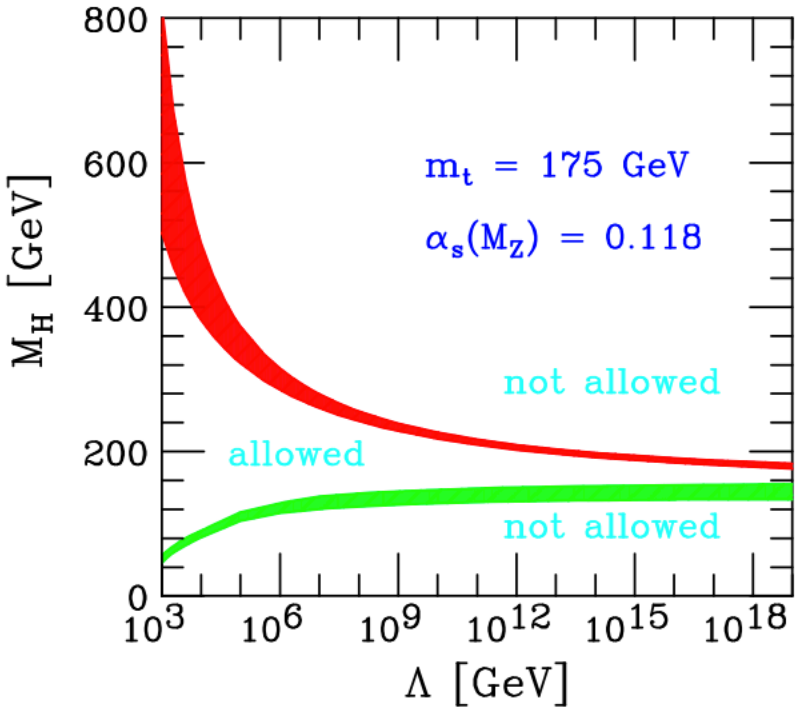


Figure 1.10: Upper and lower bound of m_{H} as a function of Λ_{cut}.

In the limit of small Higgs quartic couplings, $\lambda \ll \lambda_t, g_1, g_2$ where λ_t is the top Yukawa coupling given by $\sqrt{2}m_{\mathrm{t}}/v$, the RGE is given by [31]

$$\frac{d}{dQ^2}\lambda(Q^2) \simeq \frac{1}{16\pi^2}\left[-12\frac{m_{\mathrm{t}}^4}{v^4} + \frac{3}{16}\left(2g_2^4 + \left(g_2^2+g_1^2\right)^2\right)\right]. \tag{1.54}$$

Taking EWSB scale as the reference point, the solution to equation (1.54) is

$$\lambda\left(Q^2\right) = \lambda\left(v^2\right) + \frac{1}{16\pi^2}\left[-12\frac{m_{\mathrm{t}}^4}{v^4} + \frac{3}{16}\left(2g_2^4 + \left(g_2^2+g_1^2\right)^2\right)\right]\log\frac{Q^2}{v^2}. \tag{1.55}$$

As $\lambda(v^2)$ becomes small, the coupling can go negative, and the vacuum becomes unstable. Thus, in order to maintain the stability of vacuum, $\lambda\left(Q^2\right)$ should be positive. This requirement gives

$$m_{\mathrm{H}}^2 \;>\; \frac{v^2}{8\pi^2}\left[-12\frac{m_{\mathrm{t}}^4}{v^4} + \frac{3}{16}\left(2g_2^4 + \left(g_2^2+g_1^2\right)^2\right)\right]\log\frac{\Lambda_{cut}^2}{v^2} \tag{1.56}$$

So far the higher order contributions were taken up to 1-loop corrections. There are calculations up to 2-loops and Figure 1.10 shows lower bound (vacuum stability) and upper bound (triviality) of m_{H} as a function of cutoff scale, Λ_{cut} [31].

Fine tuning

The 1-loop radiative corrections to m_{H} when only W/Z/H and top contributions are considered are given by [31]

$$m_{\mathrm{H}}^2 = \left(m_{\mathrm{H}}^0\right)^2 + \frac{3\Lambda^2}{8\pi^2 v^2}\left[m_{\mathrm{H}}^2 + 2m_{\mathrm{W}}^2 + m_{\mathrm{Z}}^2 - 4m_{\mathrm{t}}^2\right] \tag{1.57}$$

where m_{H}^0 is the fundamental parameter of SM, Λ is the cutoff scale, and $m_{\mathrm{W}}, m_{\mathrm{Z}}, m_{\mathrm{t}}$ are W, Z, top mass, respectively. Therefore, unless Λ is in the same scale of EWSB(100 GeV $-$ 1 TeV), there should be an incredible fine-tuning between m_{H}^0 and the radiative correction to have m_{H} in the EWSB scale.

For a quantitative discussion, we first need to define what fine-tuning means. Fine-tuning is defined as the sensitivity of the weak scale to the cutoff, $|\delta m_{\mathrm{W}}^2(\Lambda)/m_{\mathrm{W}}^2|$, where δm_{W}^2 is the difference between the tree and loop values, with all other quantities held fixed [34]. So, the metric, $\mathcal{F}$, is

$$\mathcal{F} = \left|\frac{\delta m_{\mathrm{W}}^2}{m_{\mathrm{W}}^2}\right| = \left|\frac{\delta v^2}{v^2}\right| = \left|\frac{\delta \mu^2}{\mu^2}\right| = \left|\frac{\delta m_{\mathrm{H}}^2}{m_{\mathrm{H}}^2}\right| = \frac{2\Lambda^2}{m_{\mathrm{H}}^2}\left|\sum_n \log^2\left(\frac{\Lambda}{m_{\mathrm{H}}}\right)\right| \tag{1.58}$$

and $\mathcal{F} \leq 1$ represents that there is no fine-tuning.

The Figure 1.11 [34] shows two regions in $[\Lambda, m_{\mathrm{H}}]$ plane where Λ is the cutoff scale; $\mathcal{F} > 10$ in light-hatching (labeled as 10 %) and $\mathcal{F} > 100$ in thick-hatching (labeled as 1 %). In case of light Higgs scenario, the fine-tuning is required even at the low energy scale. For example, at m_{H}=130 GeV the fine-tuning of $\mathcal{F} > 10$ requires $\Lambda < 2.3$ TeV. This means that new physics should exist in the regime where LHC experiments can probe.

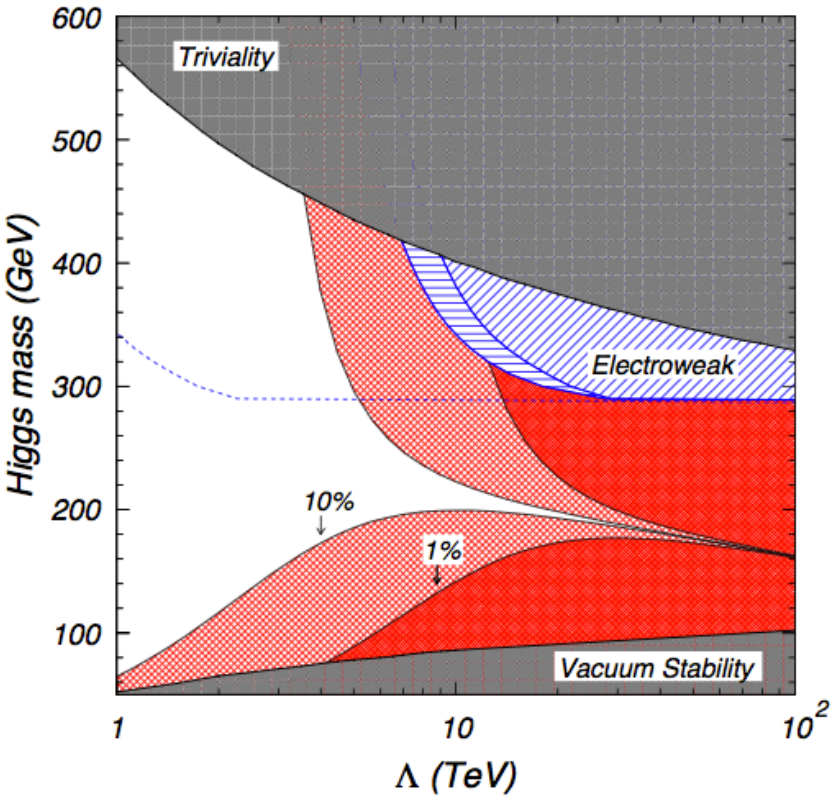


Figure 1.11: Constraint contour from fine tuning, vacuum stability, and triviality

.

1.3.2 Experimental Limits

Indirect search

There are electroweak measurements sensitive to m_{H}. For example, the mass of W boson has one-loop correction of Higgs boson as shown in Figure 1.12. Its contribution to the W mass is parametrized by Δr in the following equation [31]

$$m_{\mathrm{W}}^2 = \frac{\pi\alpha}{\sqrt{2}\mathrm{G_F}} \frac{1}{\left(1 - \frac{m_{\mathrm{W}}^2}{m_{\mathrm{Z}}^2}\right)} (1 + \Delta r)\,, \tag{1.59}$$

and the correction is

$$\Delta r \simeq \frac{\mathrm{G_F} m_{\mathrm{W}}^2}{8\sqrt{2}\pi^2} \frac{11}{3} \left(\log \frac{m_{\mathrm{H}}^2}{m_{\mathrm{W}}^2} - \frac{5}{6}\right) \tag{1.60}$$

which is dependent on m_{H} logarithmically. Thus, by measuring other quantities in the equation, we can constrain m_{H} up to the uncertainties to the measured

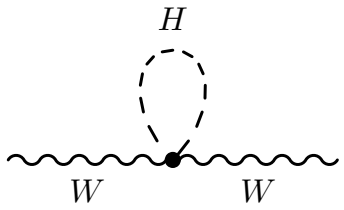


Figure 1.12: Feynman diagram for 1 loop correction by Higgs boson to the W propagator.

.

quantities. Going one step further, we can use more variables, not only m_{W}, and put them into a statistical fit [2]. A simultaneous fit is done to $\Delta\alpha^{(5)}_{had}(m^2_{\mathrm{Z}})$, $\alpha_S(m^2_{\mathrm{Z}})$, m_{Z}, m_{t}, and $\log_{10}(m_{\mathrm{H}})$ on the data collected by LEP-I/II, SLD, and Tevatron [2]. The Figure 1.13 shows $\Delta\chi^2$ curve as a function of m_{H} from EWK precision measurements assuming that Standard Model is the true theory of nature [2]. The preferred m_{H} is 94^{+29}_{-24} GeV. It also shows that the upper limit on m_{H} at C.L. = 95 % is 152 GeV.

Direct search

Before 2012, there were direct searches for SM Higgs boson by LEP, Tevatron, and LHC experiments. The LEP data showed the limit of $m_{\mathrm{H}} > 114.4$ GeV at CL_{s}= 95 % [35] and the Tevatron showed exclusion of SM Higgs hypothesis in the range of 147 GeV $< m_{\mathrm{H}} <$ 179 GeV at CL_{s}= 95 % [35]. At the end of 2011, the LHC experiments(CMS and ATLAS) showed their results with 7 TeV data on the standard model Higgs search [36, 37]. Figure 1.14 shows the 95% C.L. upper limits on $\sigma/\sigma_{\mathrm{SM}}$ as a function of m_{H} in the range of 110 - 145 GeV for CMS on the left and 110 - 150 GeV for ATLAS on the right. In both experiments, search was performed up to m_{H}= 600 GeV, but only low m_{H} region is shown on the plots. In CMS, the observed exclusion range is 127 - 600 GeV while the expected exclusion range is 118 - 543 GeV. In ATLAS, the observed exclusion range is 131-238 and 251-466 GeV while the expected exclusion range 124 - 519 GeV. Both experiments, CMS and ATLAS, showed local excess of 3.1σ and 3.5σ, respectively, around m_{H}= 125 GeV.

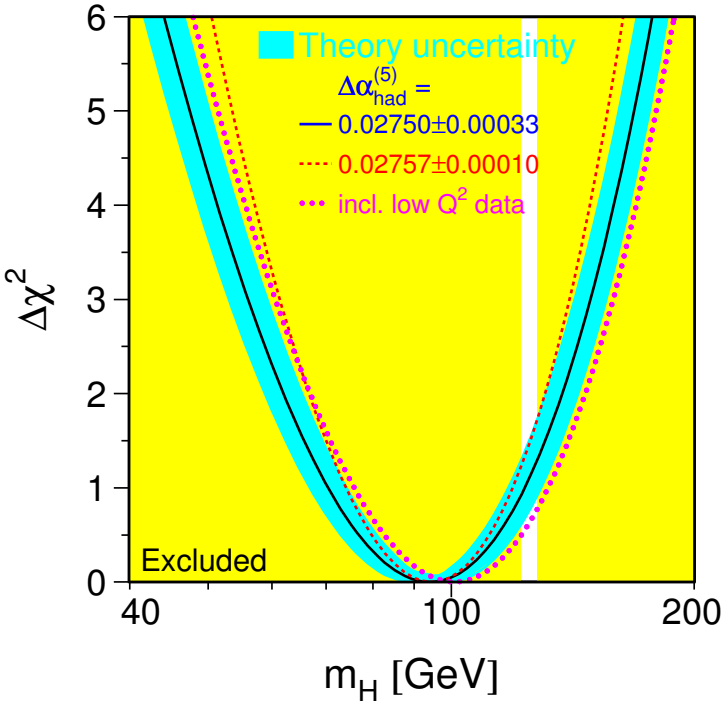


Figure 1.13: $\Delta\chi^2 = \chi^2 - \chi^2_{min}$ as a function of m_{H} [2].

1.4 The $H \to W^+W^- \to 2l2\nu$ channel

The $H \to W^+W^- \to 2l2\nu$ is an important channel because of its large production rate, which allows a good statistical power to measure the production rate. It has a special event topology due to Higgs' spin being zero and the V-A nature of leptonic W decays. This special topology results in difference between signal and backgrounds in some kinematic variables and those variables are used to separate signal from backgrounds. CMS performed a search for SM Higgs boson with 7 TeV data in $H \to W^+W^- \to 2l2\nu$ channel and no excess was found [38].

1.4.1 Large production rate

The $H \to W^+W^- \to 2l2\nu$ channel has a good statistical power to measure $\sigma \times BR$ assuming the backgrounds can be controlled. Figure 1.15 shows $\sigma \times BR$

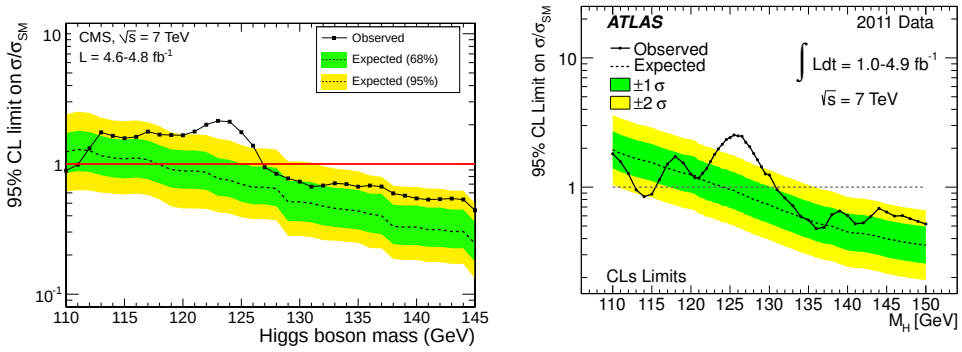


Figure 1.14: CMS / ATLAS Higgs exclusion with 7 TeV data.

at $m_{\mathrm{H}} < 250$ GeV. It shows that the $\sigma \times$ BR of $H \to W^+W^- \to 2l2\nu$ channel is large compared to the other sensitive channels, $H \to ZZ \to 4l$ and $H \to \gamma\gamma$. Table 1.3 shows $\sigma \times$ BR for the most sensitive channels, $H \to W^+W^- \to 2l2\nu$,

Table 1.3: $\sigma \times BR$ at $m_{\mathrm{H}} = 125$ GeV for most sensitive channels and the expected number of events in $\mathcal{L}_{int} = 20$ fb^{-1}. l means electrons or muons.

	$H \to WW \to 2l2\nu$	$H \to ZZ \to 4l$	$H \to \gamma\gamma$
$\sigma \times BR(pb)$	2.24×10^{-1}	2.79×10^{-3}	5.09×10^{-2}
$N_{expected}$ in $\mathcal{L}_{int} = 20$ fb^{-1}	4480	56	1018

$H \to ZZ \to 4l$ and $H \to \gamma\gamma$, and the expected signal events at the integrated luminosity, $\mathcal{L} = 20$ fb^{-1}. The expected signal events are 4480, 56 and 1018 for $H \to W^+W^- \to 2l2\nu$, $H \to ZZ \to 4l$ and $H \to \gamma\gamma$, respectively. This allows to have a good statistical power to measure the $\sigma \times$ BR with this channel.

1.4.2 Angular distribution of leptons in the final state

The spin of SM Higgs is zero, so by helicity conservation the total spin of the WW system should be zero as well. As shown in Figure 1.16, if we take the direction of W^+ momentum as z axis in the center-of-mass(CM) frame of Higgs boson, there are two cases where the spin direction is parallel to the momentum direction (transverse polarization) and one case where it is perpendicular to the momentum

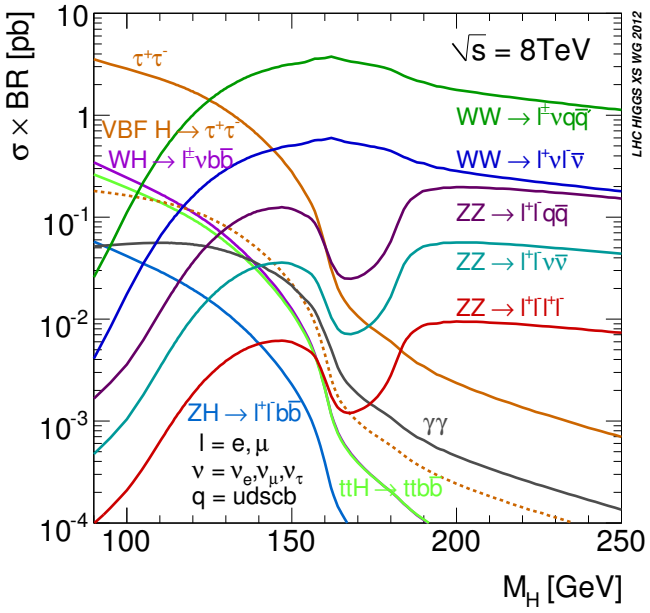


Figure 1.15: $\sigma \times BR$ at low m_{H}.

direction(longitudinal polarization). In case of transverse polarization, the leptons from Ws have strong angular dependence due to V-A nature of weak decays, i.e. neutrinos are always left-handed(anti-neutrinos are always right-handed). Let's take the case of W^+ spin in the z direction as an example. In order for the neutrino from W^+ to be left-handed, the direction of the neutrino should be in the - z direction, thus lepton should fly to z direction. In order for the anti-neutrino from W^- to be right-handed, the direction of the anti-neutrino should be in the - z direction, thus lepton should fly to z direction. Therefore, both leptons tend to move in the same direction resulting in a small angle between the two leptons. This is somewhat diluted due to boost of Higgs and Ws, but the effect is still visible and used to separate signals from non-resonant WW background. On the other hand, in case of longitudinal polarization, no specific angular correlation is present. As discussed in section 1.2.2, the fraction of longitudinal polarization depends on the Higgs mass. So, we expect that the two leptons in the final state tend to be more aligned at m_{H} where the fraction of longitudinal polarization is small.

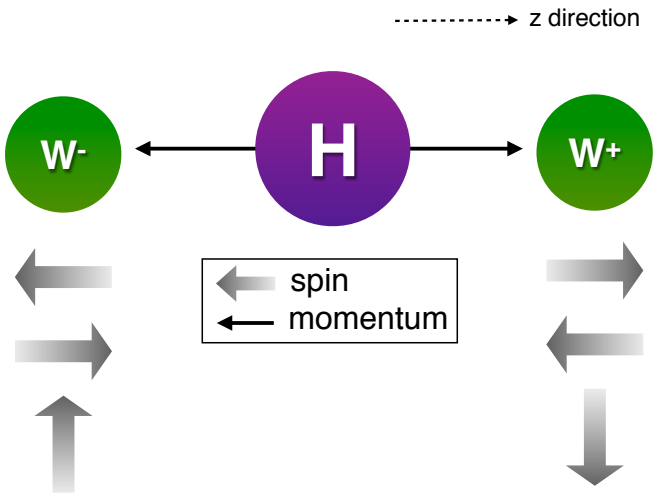


Figure 1.16: Spin configuration of Higgs boson decaying to W^+W^- in the CM frame of Higgs. Solid black arrows represent the momentum direction of Ws and grey arrows indicate the spin configuration.

1.4.3 Kinematic variables

Figure 1.17 shows distributions of kinematic variables for multiple Higgs hypotheses, $m_{\mathrm{H}} = 110, 125, 145, 160$ and 200 GeV. The plotted variables are leading and trailing lepton $p_{\mathrm{T}}(p_{\mathrm{T}}^{\ell,\mathrm{max}}, p_{\mathrm{T}}^{\ell,\mathrm{min}})$, azimuthal angle difference between the two leptons ($\Delta\phi_{\ell\ell}$), di-lepton invariant mass ($m_{\ell\ell}$), and Higgs transverse mass (m_{T}) which is defined as

$$m_{\mathrm{T}} = \sqrt{2p_{\mathrm{T}}^{\ell\ell}\mathrm{MET}(1 - cos(\Delta\phi_{\ell\ell-\mathrm{MET}}))} \tag{1.61}$$

where $p_{\mathrm{T}}^{\ell\ell}$ is transverse momentum of the di-lepton system, MET is missing transverse momentum, and $\Delta\phi_{\ell\ell-\mathrm{MET}}$ is the angle between the momentum of di-lepton system and MET in the transverse plane. Most of the events have $p_{\mathrm{T}}^{\ell,\mathrm{max}}$ greater than 20 GeV for all m_{H} hypotheses. $p_{\mathrm{T}}^{\ell,\mathrm{min}}$ is quite populated at low $p_{\mathrm{T}}^{\ell,\mathrm{min}}$ region, especially for low m_{H} hypotheses. In case of m_{H}= 125 GeV, approximately 25 % of events are rejected by requiring $p_{\mathrm{T}}^{\ell,\mathrm{min}} > 10$ GeV. $\Delta\phi_{\ell\ell}$ shows non-straightforward trend. The angle tends to get smaller as m_{H} increases up to 160 GeV, and the angle becomes wide after m_{H}= 160 GeV. This behavior was expected in the Fig-

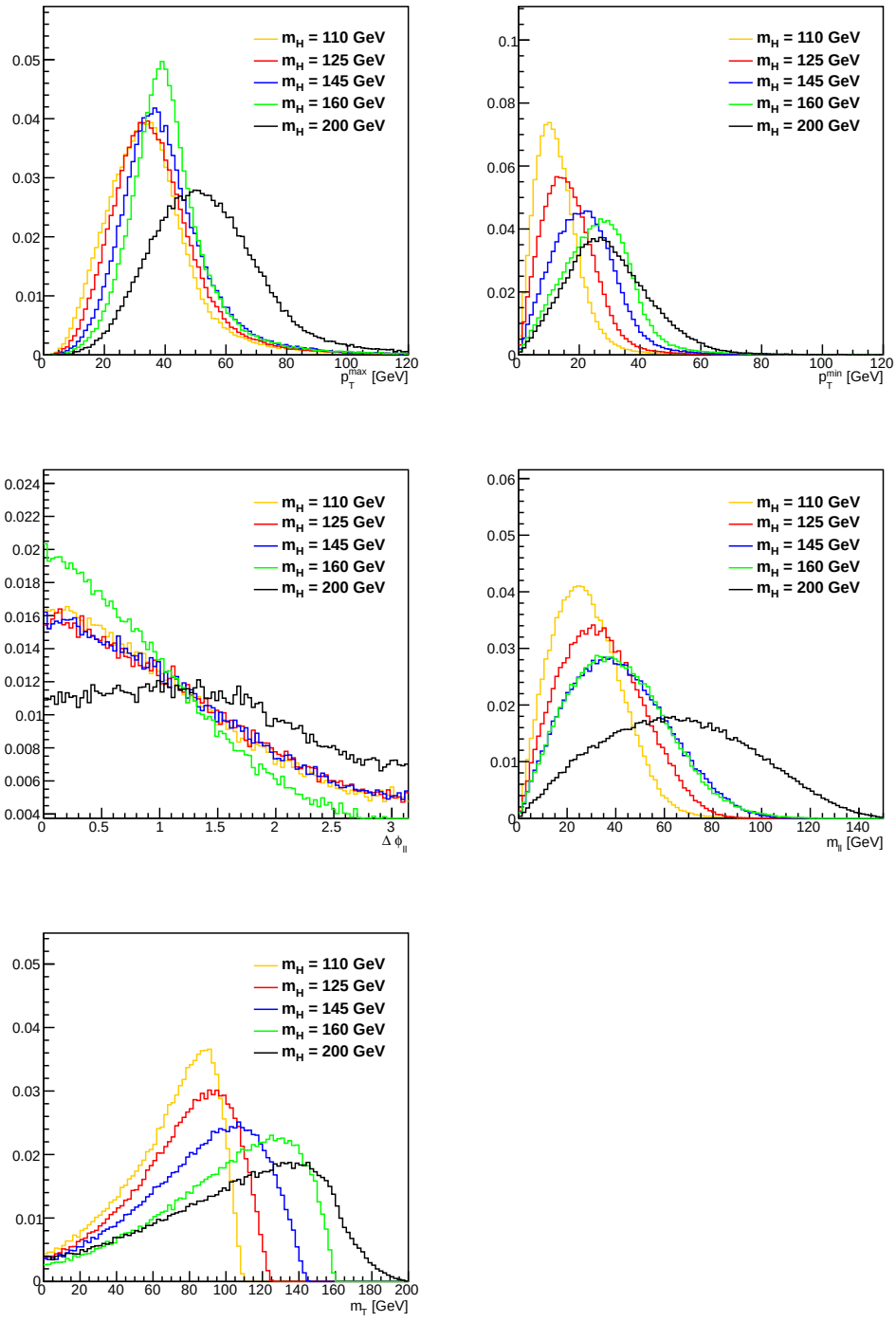


Figure 1.17: Kinematic distributions of $p_{\mathrm{T}}^{\ell,\mathrm{max}}$, $p_{\mathrm{T}}^{\ell,\mathrm{min}}$, $\Delta\phi_{\ell\ell}$, $m_{\ell\ell}$ and m_{T}, for $m_{\mathrm{H}} = 110, 125, 145, 160$ and 200 GeV at the generator level.

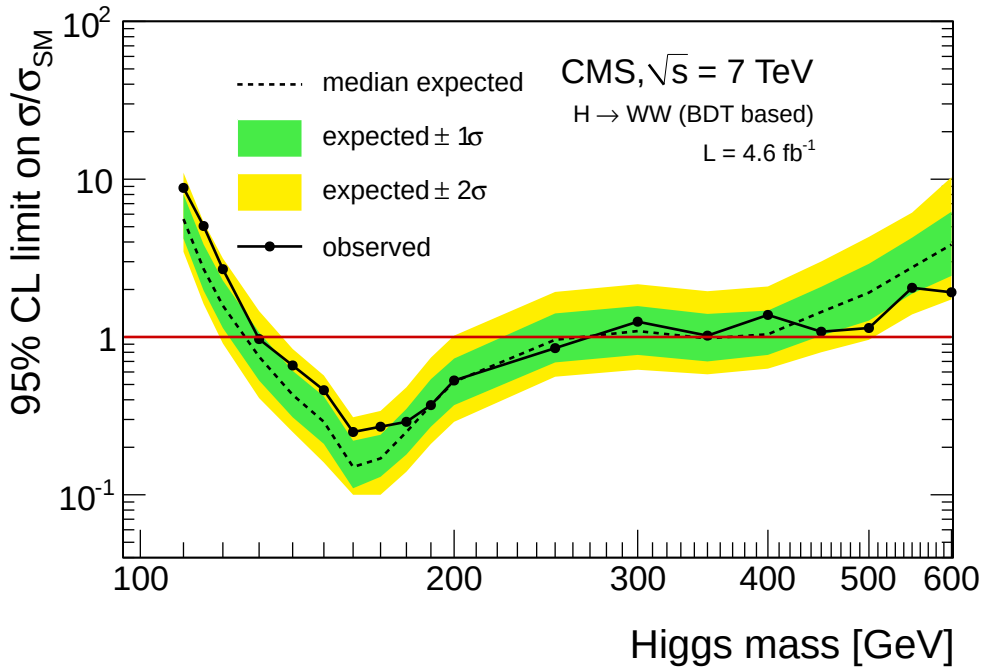


Figure 1.18: Exclusion limit of SM Higgs with 2011 data($\sqrt{s}$ = 7 TeV, $\mathcal{L}$ = 4.6 fb^{-1}). The observed(expected) exclusion limit at CL = 95 % is m_{H}= 129 - 270(127 - 270) GeV.

ure 1.6 where the fraction of the longitudinal polarization is at the minimum at m_{H}= $2m_{\mathrm{W}}$ which is about 160 GeV. Since small $\Delta\phi_{\ell\ell}$ yields small $m_{\ell\ell}$ we expect small $m_{\ell\ell}$ for low m_{H} hypotheses. The Higgs transverse mass m_{T} shows clear drop at m_{H}.

1.4.4 CMS $H \to W^+W^- \to 2l2\nu$ result before 2012

Before 2012, 4.6 fb^{-1} of data at $\sqrt{s} = 7$ TeV collected by CMS detector was analyzed for SM Higgs search in $H \to W^+W^- \to 2l2\nu$ channel [38]. Figure 1.18 shows exclusion limit of SM Higgs boson using a multivariate technique based on the boosted decision tree(BDT) algorithm [38]. The observed exclusion limit at CL_{s}= 95 % is m_{H} = 129 - 270 GeV with expected limit m_{H} = 127 - 270 GeV. There is a slight overall excess in low m_{H} region which might indicate existence of SM Higgs boson at low m_{H}.

Chapter 2

LHC and CMS Detector

This chapter describes hardware aspects of the Large Hadron Collider(LHC) and the Compact Muon Solenoid(CMS) detector. The content is heavily based on the chapter 1 of CMS Technical Design Report(TDR) volume 1 [5].

2.1 Large Hadron Collider

In order to answer the key question in particle physics, the origin of mass, physicists constructed a high-energy proton-proton(hadrons) collider at CERN. The circumference of the accelerator is about 27 km, which is large. Thus, the collider is called "Large Hadron collider(LHC)".

The protons are accelerated to the desired collision energy going through multiple steps. Figure 2.1 shows a schematic of the LHC complex. The protons are made by applying electric field to hydrogen gas (Hydrogen atom is composed of one proton and one electron) in a metal cylinder. Then, the protons and electrons that constituted hydrogen atoms are separated. The separated protons leave the metal cylinder at the energy 90 keV to be sent to Radio Frequency Quadrupole(QRF). QRF not only accelerates the protons to 750 keV, but also provides a transverse focusing of the beam. The next destination is a linear accelerator(LINAC2) where the protons gain energy up to 50 MeV($v = 0.3c$). The protons are then injected to Proton Synchrotron Booster(PSB) which has 157 m circumference. The PSB accelerates the protons to 1.4 GeV before sending them to Proton Synchrotron (PS) of 628 m circumference. PS is responsible for making 81 proton bunches with 25 ns spacing and energy of 25 GeV($v = 0.87c$), and send them to Super Proton Synchrotron(SPS) with a circumference of 7 km. The SPS accelerates the proton bunches to 450 GeV and finally injects them to the LHC, and they are accelerated to the desired collision energy($v = 0.999999c$). The beam lifetime which is the time interval during which the intensity of the beam becomes 1/e is around 10 hours, and after that the protons are dumped to prepare for the next proton fill [39].

To reach very high collision energy with a fixed-size accelerator, the magnetic field that bends the protons should be very high. In order to operate at the design proton energy of LHC (7 TeV), the magnetic field should be 8.33 T.

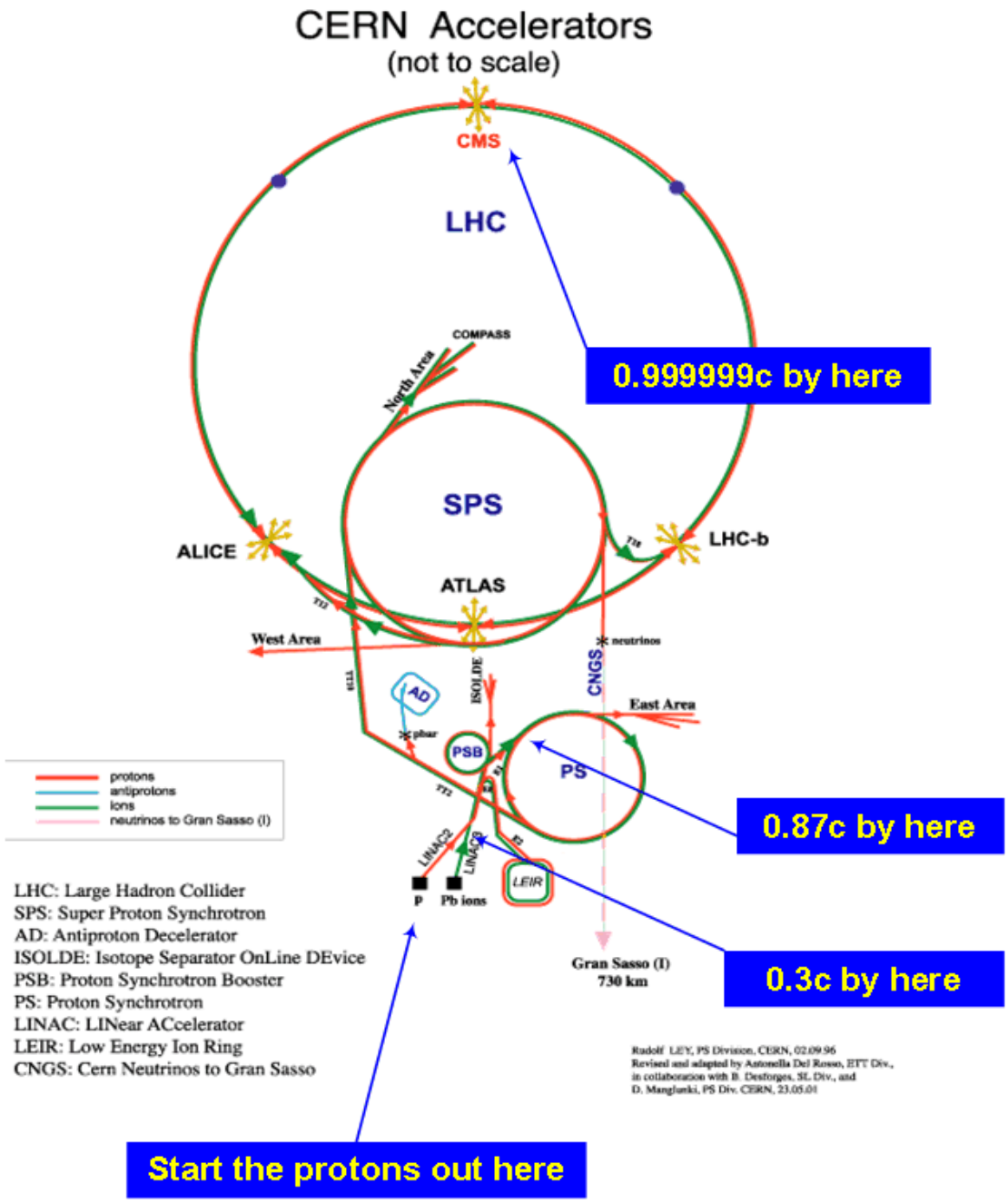


Figure 2.1: Schematic of the CERN accelerator complex.

This magnetic field is too high, LHC uses dipole magnets using electromagnets. For an electromagnet to have a high magnetic field, the current that runs in the rounding coil should be high. But, practically this is not easy because the coils will burn at high current due to its resistance. LHC dipole magnets use niobium-

titanium(NbTi) cables with a current 11850 A at the temperature 1.9 K. The extremely low temperature, 1.9 K, at which the coil becomes a superconductor is achieved using superfluid helium. Figure 2.2 shows the inner structure, cross section and magnetic field map of an LHC cryodipole.

The luminosity is given by

$$\mathcal{L} = \frac{\gamma f k_B N_p^2}{4\pi\epsilon_n \beta^*} F \tag{2.1}$$

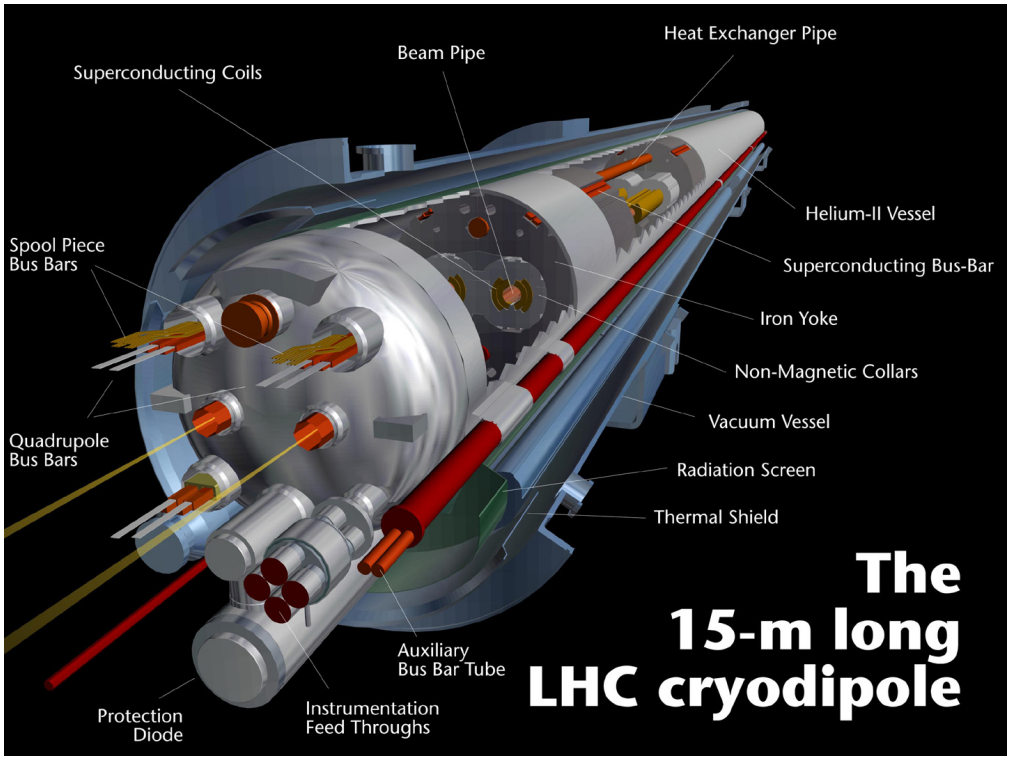
where γ is the Lorentz factor, f is the revolution frequency, k_B is the number of bunches, N_p is the number of protons per bunch, ϵ_n is the normalized transverse emittance, β^* the amplitude function at the interaction point, and F is the reduction factor due to the crossing angle. The left figure in Figure 2.3 shows the integrated luminosity delivered by LHC in 2010, 2011 and 2012 as a function of time each year. The delivered integrated luminosity is 6.1 fb^{-1} at $\sqrt{s} = 7$ TeV in 2011 and 23.3 fb^{-1} at $\sqrt{s} = 8$ TeV in 2012.

While a perfect detector can record the delivered luminosity with 100 % efficiency, in reality there are lost collisions due data acquisition system being busy and temporary unavailability of detector subsystems. The right figure in Figure 2.3 shows the delivered and the recorded luminosity integrated in the 2012 run period. Of the 23.30 fb^{-1}of the delivered luminosity, 21.79 fb^{-1}is recorded by CMS detector.

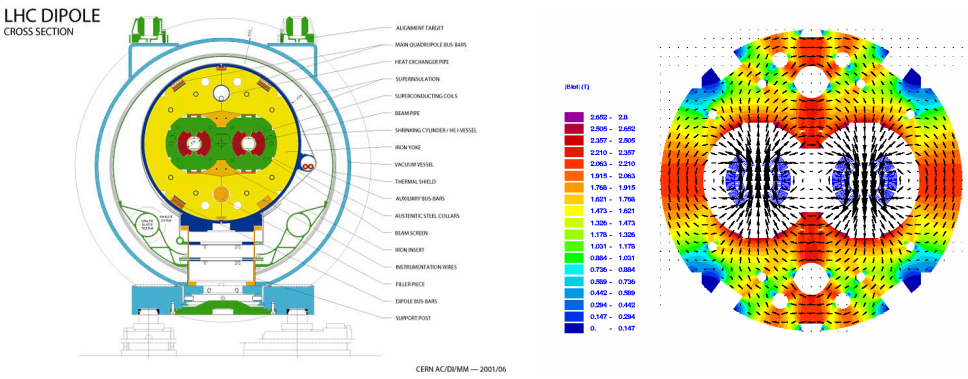
As the proton beams are very squeezed in a proton bunch, there are multiple proton-to-proton interactions in one bunch crossing. This multiple interaction is called pileup. The pileup can be calculated by

$$N_{\text{PU}} = \sigma_{\text{min bias}} \times \mathcal{L}_{\text{bunch crossing}} \tag{2.2}$$

where N_{PU} is the number of pileup events, $\sigma_{\text{min bias}}$ is the inelastic p-p cross section, and $\mathcal{L}_{bunchcrossing}$ is the luminosity per bunch crossing. In 2011 and 2012 the proton bunches crossed every 50 ns and the peak instantaneous luminosity is 7.7×10^{33} cm^{-2}s^{-1}. Figure 2.4 shows the pileup distribution of the data recored by CMS detector in 2012 run period. The average pileup is 21.



(a)



(b) (c)

Figure 2.2: (a) The inner structure, (b) cross section and (c) magnetic field map of an LHC cryodipole.

2.2 Compact Muon Solenoid detector

There are two multi-purpose detectors built at the points where the two beams collide. Compact Muon Solenoid(CMS) detector [5] is one of them. The

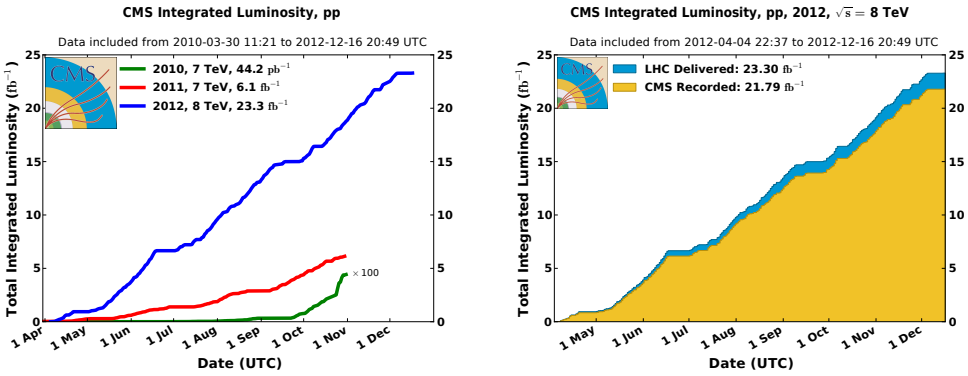


Figure 2.3: Integrated luminosity delivered by LHC in 2010, 2011 and 2012 on left. Integrated luminosity delivered by LHC and recored by CMS in 2012 on right.

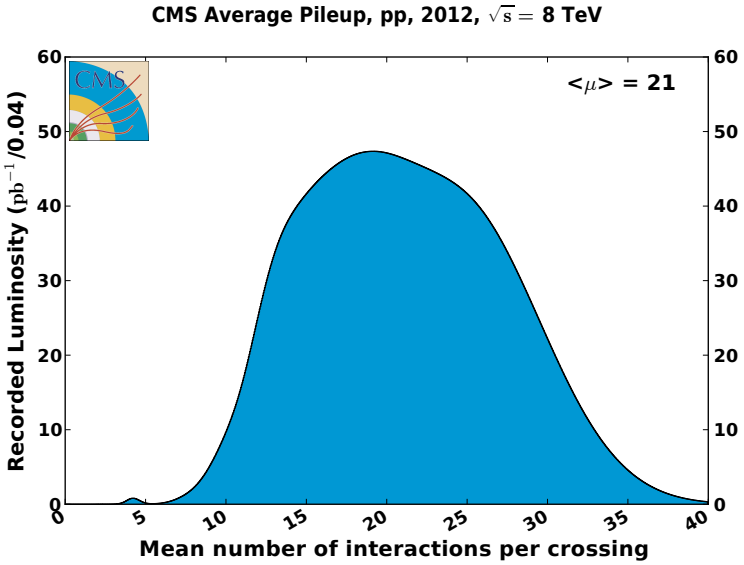


Figure 2.4: Number of interactions per bunch crossing in 2012 run period.

design of the CMS detector is based on the detection of SM Higgs boson in the wide mass range, especially, $H \to \gamma\gamma$ in the low m_{H} range and $H \to ZZ \to 4l$ in the medium m_{H} range. To achieve this the detector needs to have a good muon identification and momentum resolution, good charged particle momentum resolution and reconstruction efficiency, good electromagnetic energy resolution, and good MET and di-jet mass resolution where MET is the transverse momentum imbalance due to non-interacting particles such as neutrinos and jet is a collimated

spray of particles originated from quarks and gluons. The design of the CMS detector described in the following sections is to meet these requirements.

The momentum and location of physics objects is expressed with respect to the origin centered at the collision point in the detector. The x-axis points to the center of the collider, the y-axis points upward, and the z-axis goes clockwise along the beam line. CMS uses cylindrical coordinate system due to its cylindrical shape. The azimuthal angle, ϕ, is the angle in the x-y plane and the polar angle, θ, is measured with respect to the z-axis. Pseudo-rapidity is defined as $\eta = -\ln\tan(\theta/2)$. Advantage of pseudo-rapidity is that its difference is Lorentz invariant under longitudinal boost [1]. Because the initial momentum in z direction is not known due to movement of partons inside a proton, and momentum x and y direction is almost zero, the momentum and energy of an object is expressed in terms of transverse quantities, p_{T} and E_T, calculated in the transverse plane using only x and y components.

The overall layout of the CMS detector is as follows. Starting from the collision point to outward, there is the inner tracker system(Tracker) that is com-

[1] In frame O,

$$y_1 = \frac{1}{2}\ln\frac{E_1 + p_{1L}}{E_1 - p_{1L}} \text{ and } y_2 = \frac{1}{2}\ln\frac{E_2 + p_{2L}}{E_2 - p_{2L}} \tag{2.3}$$

$$\triangle y = y_1 - y_2 = \frac{1}{2}\ln\frac{(E_1 + p_{1L})(E_2 - p_{2L})}{(E_1 - p_{1L})(E_2 + p_{2L})} \tag{2.4}$$

In frame O' which is boosted to z direction with velocity $\beta = v/c$,

$$\triangle y' = y'_1 - y'_2 = \frac{1}{2}\ln\frac{(E'_1 + p'_{1L})(E'_2 - p'_{2L})}{(E'_1 - p'_{1L})(E'_2 + p'_{2L})} \tag{2.5}$$

where

$$E'_i = \gamma(E_i - \beta p_i) \text{ and } p'_i = \gamma(-\beta E_i + p_i) \qquad (i = 1, 2) \tag{2.6}$$

Thus we have

$$\begin{aligned}
\triangle y' &= \frac{1}{2}\ln\frac{(\gamma(E_1 - \beta p_{1L}) + \gamma(-\beta E_1 + p_{1L})')(\gamma(E_2 - \beta p_{2L}) - \gamma(-\beta E_2 + p_{2L}))}{(\gamma(E_1 - \beta p_{1L}) - \gamma(-\beta E_1 + p_{1L}))(\gamma(E_2 - \beta p_{2L}) + \gamma(-\beta E_2 + p_{2L}))} && (2.7)\\
&= \frac{1}{2}\ln\frac{(E_1 + p_{1L})(1-\beta)((E_2 - p_{2L})(1+\beta)}{(E_1 - p_{1L})(1+\beta)((E_2 + p_{2L})(1-\beta)} && (2.8)\\
&= \frac{1}{2}\ln\frac{(E_1 + p_{1L})(E_2 - p_{2L})}{(E_1 - p_{1L})(E_2 + p_{2L})} && (2.9)\\
&= \triangle y && (2.10)
\end{aligned}$$

The rapidity difference is invariant under Lorentz boost along the beam axis.

posed of pixel detector and silicon strips and covers $|\eta| < 2.5$, Electromagnetic calorimeter(ECAL) that covers $|\eta| < 3.0$, Hadronic calorimeter(HCAL) that covers $|\eta| < 5.0$, the magnet system, and the muon system that covers $|\eta| < 2.4$. The details of each sub-detector system are described in the following sections.

2.2.1 Tracker

The purpose of the tracker is to measure momentum of charged particles. Based on the charged particle flux, the tracker volume can be divided into 3 regions in radial direction. The first region is $r < 20$ cm, closest to the interaction vertex. Due to highest particle flux, the pixel detectors each of which has a size of 100×150 μm^2 are installed. The occupancy is about 0.01 % per pixel per bunch crossing. The second region is $20 < r < 55$ cm where particle flux is low enough for silicon micro-strip detectors to be able used. The minimum size of a strip is 10 cm$\times$80 μm. The occupancy is about $2 - 3$ % per bunch crossing. The third region is $55 < r <$ 110 cm where larger-pitch silicon micro-strips are used. The maximum size of a strip is 25 cm $\times$ 180 μm. The occupancy is about 1 % per crossing.

The layout of the tracker is shown in Figure 2.5. The total size is 1.1 m in radius and 5.4 m in length. In the barrel region, there 3 layers of hybrid pixel detectors at r = 4.4, 7.3 and 10.2 cm. The silicon micro-strips are placed in $20 < r < 110$ cm. The barrel region is further separated into Inner Barrel(TIB) with 4 layers and Outer Barrel(TOB) with 6 layers. In the endcap region, there are 2 pixel layers, Tracker Inner Disks(TID) with 3 layers and Tracker End Cap(TEC) with 9 layers. The TID is installed to fill the gap between TIB and TEC. The tracker system provides a coverage up to $|\eta| = 2.5$.

The pixel tracker is composed of 3 layers in barrel and 2 disks in both endcap regions. The barrel layers have a length of 53 cmand the endcap disks are located at $|z| = 34.5$ cm and 46.5 cm. The dimension of each pixel is 100 μm in r and ϕ, and 150 μm^2 in z. The size is chosen to take into account Lorentz drift [1] and to maintain charge sharing between multiple pixels. Each endcap disk is

[1] Charge carriers in the pixel detector are deflected by the magnetic field perpendicular to the electric field.

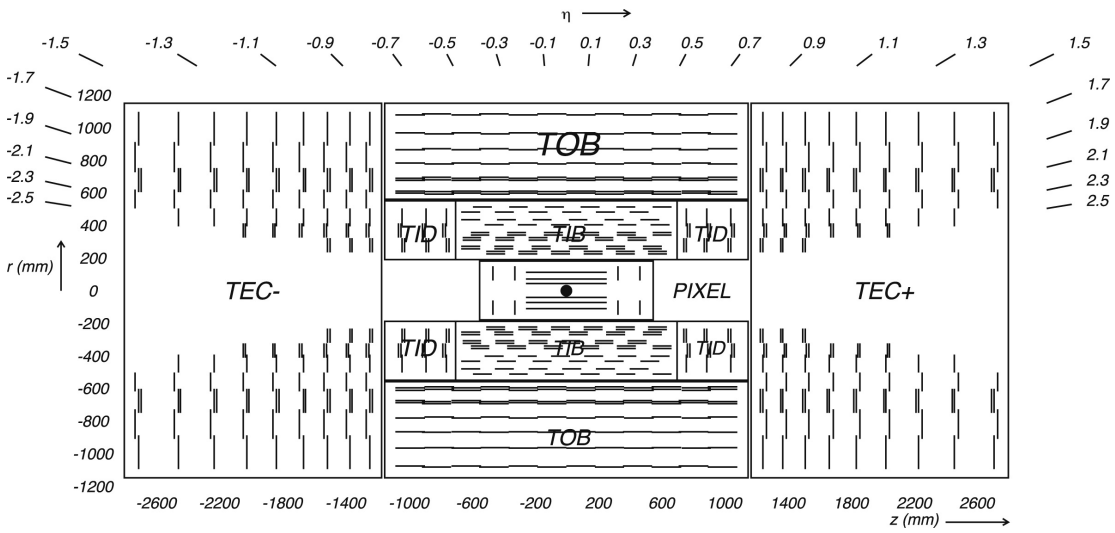


Figure 2.5: The layout of the CMS tracker [3].

composed of 24 blades assembled to form a turbine-like geometry and the blades are rotated by 20 ° considering Lorentz effect. The pixel detector provides spatial resolution of 10 and 20 μm for $r-\phi$ and z measurements, respectively.

The strip tracker is composed of TIB and TOB in barrel region and TEC and TID in the endcap region. The Number of detectors, thickness, and mean pitch is shown in the tab. 2.1. The first and the second layers of TIB are made with stereo modules with angle 100 mrad providing single-point resolution of 23-34 μm in the $r-\phi$ direction and 230 μm in z direction. As TIB, the first and the second layers of TOB are made with stereo modules with angle 100 mrad providing single-point resolution of 35-52 μm in the $r-\phi$ direction and 530 μm in z direction. In the endcap region, the first and the second layers of TID, and the first, second and the fifth layers of TEC are stereo modules that provide 3-dimensional measurements from $r-\phi$ and $r-z$ measurements.

Ideal tracker would have no energy loss of charged particles while they cross the tracker so that they deposit their initial energies in calorimeter. However, real trackers can not be made that way due to materials used to build the tracker such as electrical cables, cooling services, support structures, electronics and beam-pipe. Figure 2.6 shows the material budget of tracker in the units of radiation length(X_0) and interaction length(λ_0).

Table 2.1: Number of detectors, thickness and mean pitch of each strip, coverage in z direction, and number of layers(disks) of the four part of silicon strip detector, TIB, TOB, TID and TEC.

Part	Number of detectors	thickness (μm)	mean pitch (μm)	coverage	layers (disks)
TIB	2724	320	81-118	$\|z\| < 65$ cm	4
TOB	5208	500	81-183	$\|z\| < 110$ cm	6
TID	816	320	97-143	$65 < \|z\| < 120$ cm	3
TEC	6400	320-500	96-183	$120 < \|z\| < 280$ cm	9

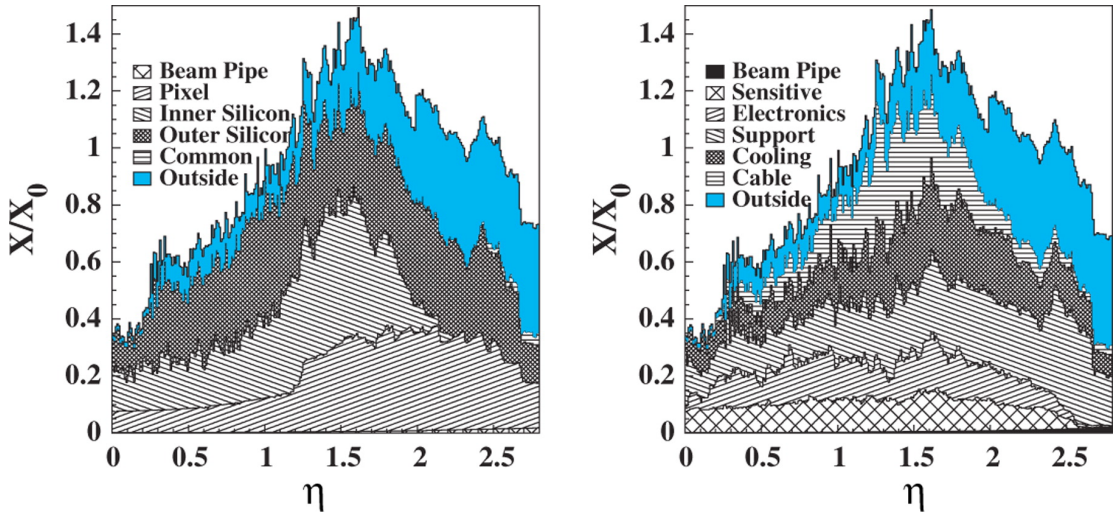


Figure 2.6: Material budget of CMS tracker in terms of radiation length(X_0) and interaction length(λ_0) [4].

2.2.2 Electromagnetic Calorimeter

The Electromagnetic Calorimeter(ECAL) is crucial to measure the energy of electrons and photons. The CMS ECAL is composed of 61200 and 7324 lead tungstate($PbWO_4$) scintillating crystals in barrel and endcap regions, respectively. $PbWO_4$ crystals have short radiation length(0.89 cm) and small Moliere radius(2.2 cm). It is fast in emitting light and radiation-resistant. But, it produces relatively

low light yields, which requires amplification of light signal using photo-detectors. Silicon avalanche photo-diodes(APD) are used in barrel and vacuum photon triodes(VPT) are used in endcap. The crystals and APDs are sensitive with temperature, so the stability of temperature is required for operation of ECAL.

In the barrel region(EB), there are 36 supermodules and each covers the region, $0 < |\eta| < 1.479$. As shown in Figure 2.7, one supermodule contains 4 modules and each module has 40 - 50 submodules which are composed of $10(2 \times 5)$ sub-units(one crystal + one capsule). Each crystal covers $0.0174 \times 0.0174 (\approx 1\ ^\circ \times 1\ ^\circ)$ in $\Delta\phi \times \Delta\eta$ and is tilted at $3\ ^\circ$ with respect to the line from nominal vertex position. The length of a crystal is 230 mm which corresponds to $25.8X_0$.

In the endcap region(EE) as shown in Figure 2.8, the basic unit is "supercrystal" which is a collection of 5×5 crystals. Each endcap region is covered by two D-shape structures that consist of the supercrystals. One crystal has the front face size of 28.6×28.6 mm^2 and length of 220 mm$(24.7X_0)$. Each crystal is off-pointing the nominal vertex as the barrel crystals. In the endcap region, the granularity of the ECAL is not good enough to distinguish two photons when they are very close in space. Therefore, a preshower device is placed in front of the endcap crystals. It is made of 2 planes of silicon strip detectors with a pitch of 1.9 mm located behind the disks of lead absorber of depth $2X_0$ and $3X_0$. Due to its finer structure, it provides better spacial resolution that enables to reject false photons.

The performance of the supermodule in the barrel region is measured using a test beam. The electrons in the test beam were incident on the central crystal of 3×3 crystals. Figure 2.9 shows its energy resolution$(\sigma(E)/E)$ as a function of beam energy(E). The energy resolution of calorimeters is parametrized by

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \tag{2.11}$$

where S is the stochastic term that reflects the fact that the development of showers is a statistical process, N is to account for instrumental effects such as noise and pedestals, and C is the constant term for calibration errors such as non-uniformity of detectors. The figure shows corresponding values for two curves that use different trigger conditions.

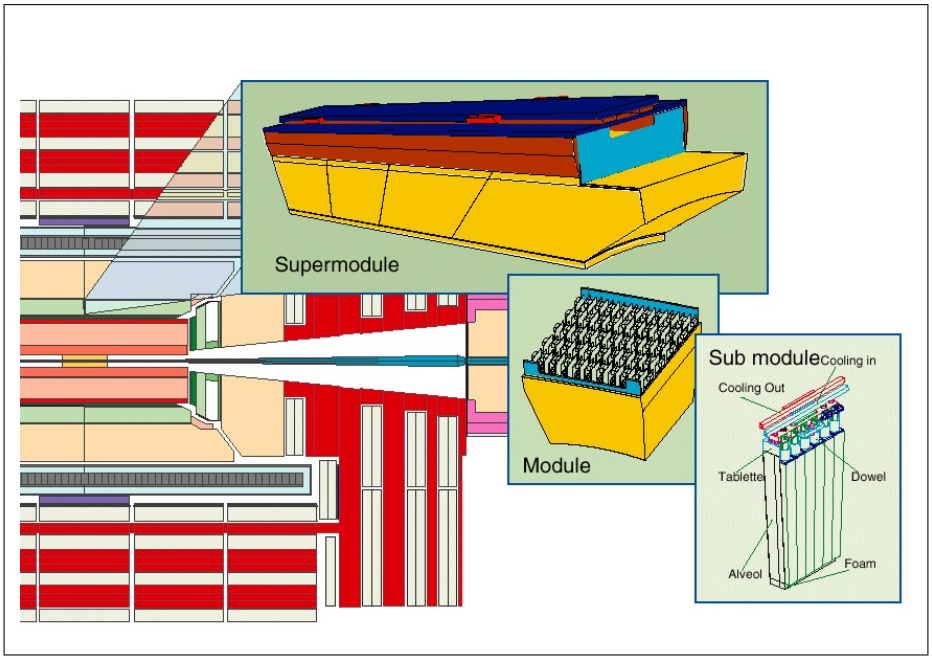


Figure 2.7: Barrel region of ECAL.

2.2.3 Hadronic Calorimeter

The hadronic calorimeter(HCAL) is made to measure energy of jets. For the design of HCAL it is important to minimize the Gaussian tails in the energy resolution, and to achieve good containment of energy deposit and hermeticity for MET calculation. CMS HCAL is a sampling detector that is composed of alternating layers of an absorber and a scintillator. As a hadronic particle hits an absorber plate, interactions occur to produce secondary particles, and these produced particles interact with the material in the next layer of absorber, resulting in a shower of hadronic particles. When these particles cross the active scintillating layers, they cause them to emit light which can be detected by optical devices. The design of HCAL is influenced by the magnet parameters because the most of the calorimeter system in CMS is located in the solenoid. This motivated to choose a material with short interaction length, and brass is chosen as the HCAL material for this reason. The HCAL is organized by Hadron Barrel(HB), Hadron Outer(HO), Hadron Endcap(HE) and Hadron Forward(HF). Figure 2.10 shows the

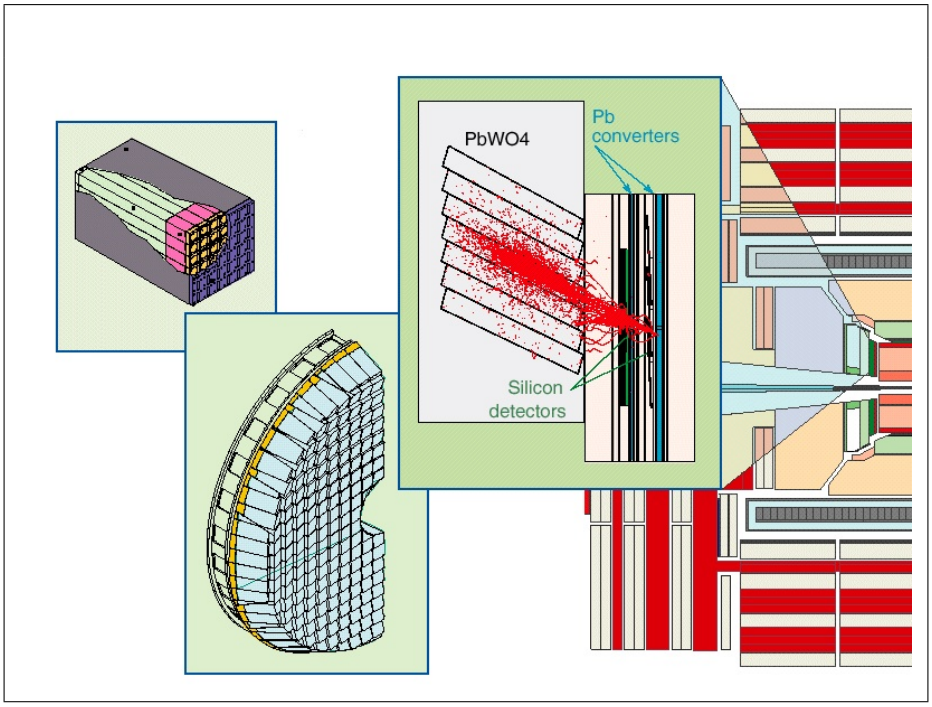


Figure 2.8: Endcap region of ECAL.

layout of CMS HCAL where HF is not shown.

The HB consists of 2304(32×72 in $\eta \times \phi$) towers of size $\Delta\phi \times \Delta\eta = 0.087 \times 0.087$. One HCAL tower has the same $\Delta\phi \times \Delta\eta$ coverage as 5×5 ECAL crystals. There are 15 brass plates with thickness 5 cm and 2 stainless steel for mechanical support. The first scintillating plate placed before the first brass plate has width of 9 mm while the other 16 scintillating plates have width of 3.7 mm.

In the barrel, HO which covers $|\eta| < 1.26$ is placed to complement the short length of HB that may not be enough to contain all particles. The escaping showering particles is the cause of tail in the energy resolution. Adding HO effectively increases the interaction length over 10, thus energy resolution is enhanced. This is very important for MET resolution calculated using calorimeter information. HO is divided into 5 sections in η, resulting 5 rings each of which covers 2.5 m in z direction. The central ring has two layers of scintillator placed at $r = 3.85$ m and $r = 4.097$ m with an iron absorber of thickness 18 cm between them. The other

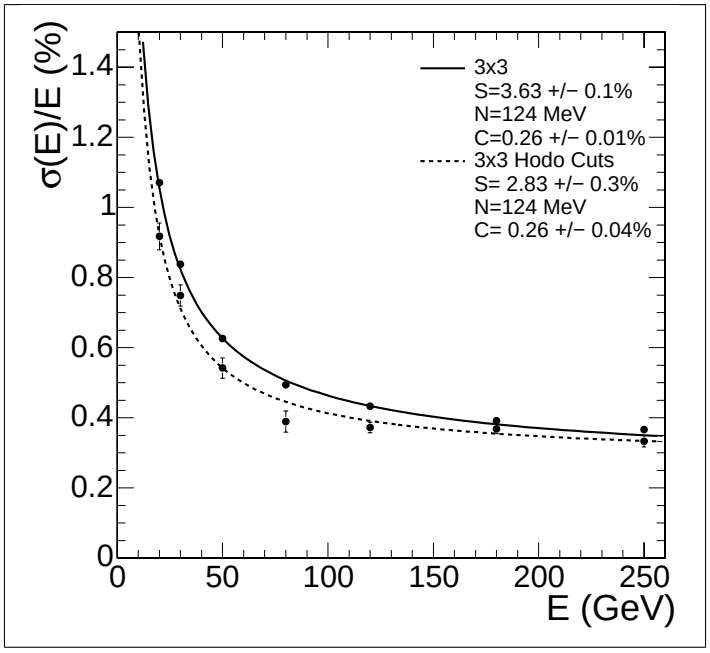


Figure 2.9: Energy resolution of ECAL supermodule measured using a test beam [5].

rings have one scintillator layer at $r = 4.097$ m.

The HE is composed of 2304 towers covering $1.3 < |\eta| < 3.0$. As shown in left figure of Figure 2.11, the 5 outermost towers have ϕ segmentation of 5 ° and η segmentation of 0.087. The 8 innermost towers have ϕ segmentation of 10 ° and varying η segmentations of 0.09 - 0.35.

The HF is located at $z = \pm 11.2$ m covering $3.0 < |\eta| < 5.0$. It provides improvement in measurement of MET and reconstruction of forward jets which can be used to identify very interesting processes such as Vector Boson Fusion(VBF). Because HF receive the bulk of particle energy of the collision, the material should be resistive to radiation. HF modules are made of steel blocks with quartz fibers. The particles crossing the fibers emit Cherenkov lights and the lights are collected by photomultipliers connected to the fibers. The right figure of Figure 2.11 shows a $\Delta\phi = 20$ ° wedge of HF. The ϕ segmentation is 10 ° for all towers except for the the two innermost towers which have $\Delta\phi = 20$ °. The η segmentation is 0.175 except for the innermost and the outermost towers which have $\Delta\eta = 0.3$ and 0.1,

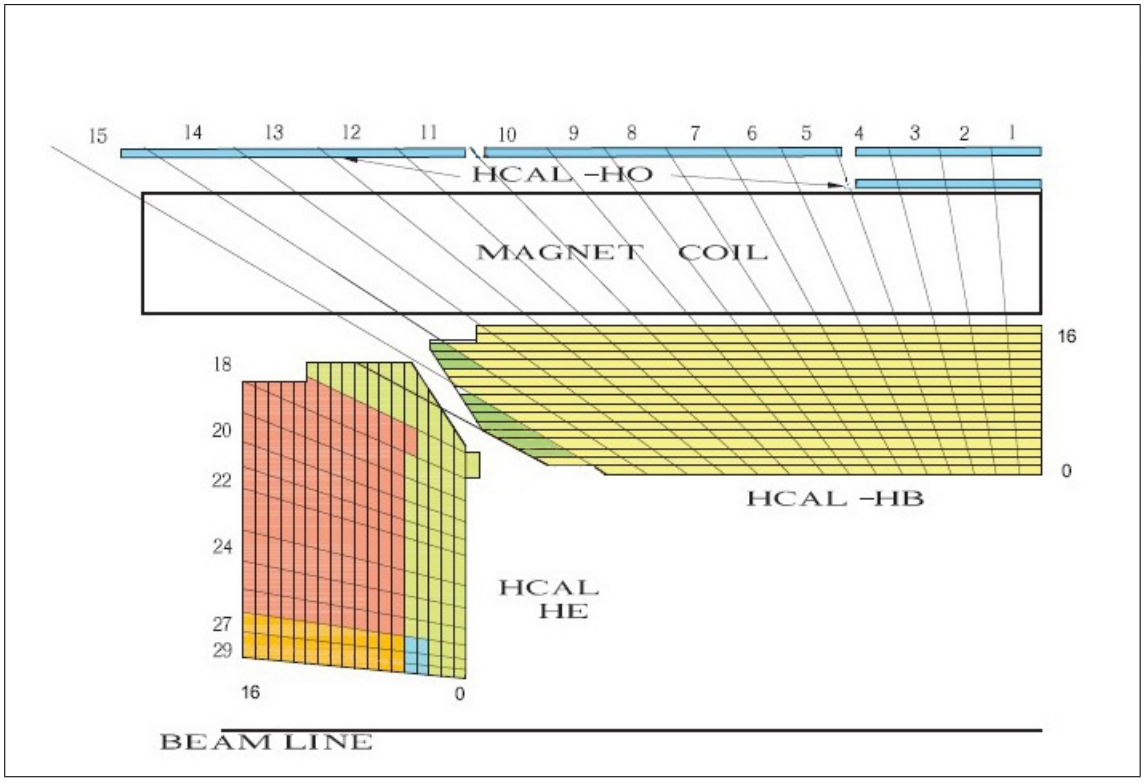


Figure 2.10: Layout of CMS HCAL [6].

respectively.

2.2.4 Magnet

The momentum of a charge particle can be determined by measuring its curvature in magnetic field. Stronger magnetic field bends the trajectory more, thus allows better measurement of momentum. CMS(as its names indicates) uses superconducting solenoid which produces a uniform magnetic field 3.8(4.0 in design) T in z direction. The solenoid has 2168 turns and the current to generate 3.8 T magnetic field is around 18 kA, giving a stored energy of 2.3 GJ. The size of the magnet system is 12.9 m in length and 5.9 m in diameter. The three layers of return yokes that guide the magnetic field back to the solenoid are installed outside of solenoid, interleaved with the muon system. The magnet system also

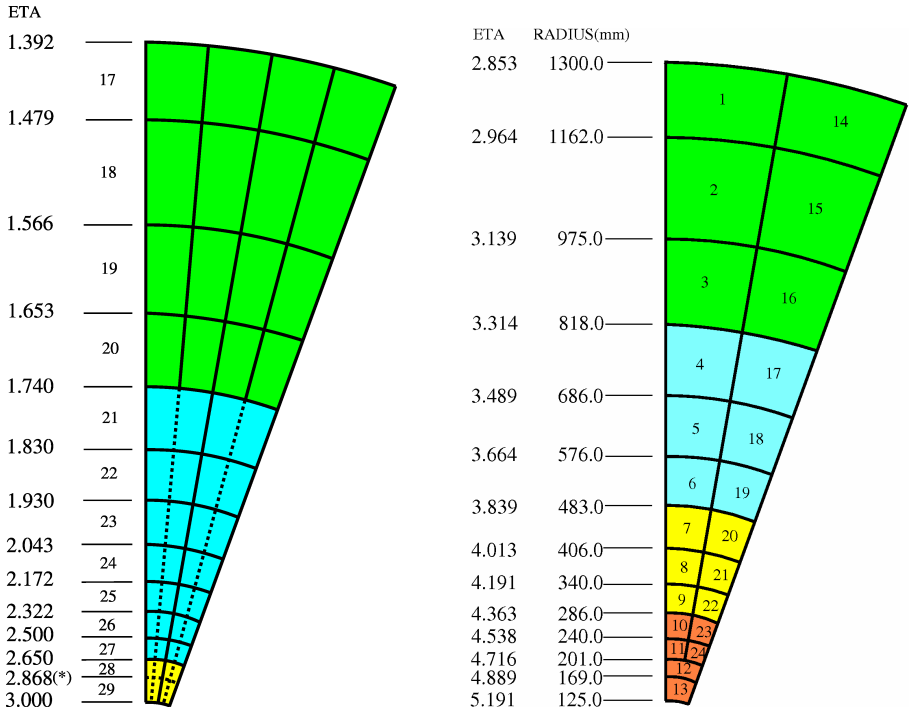


Figure 2.11: Layout of a single wedge of HE(left) and HF(right) [5].

provides mechanical support of the detector because of its strength and tolerance to its own magnetic field. Figure 2.12 shows the photos of CMS magnet system : return yoke, outer vacuum tank and coil.

2.2.5 Muon System

The Muon system of CMS is composed of three gaseous detectors, Drift Tube chambers(DT), Cathode Strip Chambers(CSC) and Resistive Plate Chambers(RPC) that cover $0 < |\eta| < 1.2$, $0.9 < |\eta| < 2.4$ and $0 < |\eta| < 1.6$, respectively. Figure 2.13 shows the layout of a quarter of muon system. Muon Barrel region(MB) has four layers of muon station interleaved with three layers of return yoke. Each layer has a cylindrical shape around the beam axis. There are 5 segments in z direction as the return yoke. Muon Endcap region(ME) also has four layers(disks) of muon stations placed perpendicular to the beam axis. The innermost disk(ME1)

has 3 concentric rings and the other disks have 2 rings.

DT chamber consists of 250 chambers constructed in 4 layers at r = 4.0, 4.9, 5.9 and 7.0 m from the beam axis. Each chamber has 2 superlayers in ϕ and 1 superlayer in z direction. Each superlayer has 4 layers of drift tubes. Each drift tube has the width of $\approx$ 4 cm and the height of $\approx$ 1 cm, and there is a stretched wire(anode) in the middle of tube filled with a mixture of Ar and CO_2 gas. When a muon passes through a tube, it knocks electrons off the atoms of the gas, and they are drifted to the anode. Each tube provides 2-dimensional measurement. One is given by the position of the central wire, and the other is given by the drift time of electrons divided by drift speed. Figure 2.15 shows a schematic of a movement of electrons in a tube and an illustration of a single tube with the electric field lines. Each station provides ϕ precision better than 100 μm in position and 1 mrad in direction.

CSC is composed of 6 gas gaps filled with a mixture of Ar, CO_2 and CF_4, and each gap has a plane of radial cathode and a plane of anode wires perpendicular to the strips. The gap between the anode wires is about 3 mm and the width of a strip is 3-16 mm. Figure 2.15 shows a schematic of a CSC on the left and an illustration of what happens when a muon passes through a gas gap. When a muon traverses in the gas, it knocks out electrons from the gas atoms. Then, an avalanche of electrons is created and moves to the wires. The ionized positive atoms move toward the strips and make a charge pulse. Because the strips are laid perpendicular to each other, each gas gap provides 2-dimensional measurements. By weighting by charge distribution, a precise spatial measurements can be made. Each CSC provides spatial resolution of 200 μm using strips and ϕ resolution of order of 10 mrad. Because of fast drift time, CSC is used in the Level-1 trigger.

RPC has one anode plate and one cathode plate organized in parallel as shown in Figure 2.16. The plates are separated by a gas chamber of thickness 2 mm filled with a mixture of $C_2H_2F_4$ and i-C_2H_{10}. When a muon passes through the chamber, an avalanche of electrons is formed, and the electrons are collected by external metallic strips. The pattern of the strip hits are then translated to the momentum measurement of the muon. Though the spatial resolution of RPC is

not as good as DT or CSC, its time resolution is very good(3 ns), thus it is able to identify correct bunch crossing without ambiguities.

2.2.6 Trigger and Data Acquisition

The design bunch crossing rate of 40 MHz yields $\sim 10^9$ events per second at the design luminosity. However, only $\sim$ 100 Hz of events can be recorded on tape. So, the trigger system should attain about an order of 10^6 reduction of events. CMS trigger and DAQ system is composed of detector electronics, Level-1(L1) triggers, readout networks and High-Level triggers(HLT).

The total time allocated to L1 trigger decision and data transit is 3.2 μm. During this time, the data collected by detectors are kept in buffers until the decision is made. The decision is made based on the presence of trigger primitives such as photons, electrons, muons and jets in the kinematic region of interests. It also employs global sums of E_T and MET. Custom hardware processors are used for L1 decision. L1 reduces the crossing rate by an order of 1000 to 100 kHz.

Once the L1 decision is made, the data on the buffer is transferred to the front-end memories for access by the DAQ system. The data for an event is sent to a processor in a computing farm with $\mathcal{O}(10^3)$ processors. Each processor runs the same HLT software to reduce the L1 rate 100 kHz down to 300 Hz. The HLT software uses partial event reconstruction, and makes decision combining information from multiple virtual trigger levels, for example, L2 for calorimeter and muon and L3 for tracking. The use of HLT after L1 gives a maximal flexibility because it gives freedom in selection.

2.2.7 CMS computing

Even after the rate of data recording is reduced to 300 Hz by HLT, it is still a huge amount of data to store and process. CMS thus employed highly distributed computing model(grid system) with Tier-0 center at CERN supplemented by Tier-1 and Tier-2 computing centers all around the world.

Tier-0 center repacks RAW data into primary datasets using trigger in-

formation and sends them to Tier-1 centers. It also does prompt reconstruction that produces RECO and Analysis Object Data(AOD), and distributes them to Tier-1. Tier-0 is not accessible by analysers, but performs only scheduled activities. Tier-1 centers store a subset of RAW data as backup, provides CPU power for re-reconstruction, skimming, calibration and AOD extraction, and stores and distributes the produced datasets to Tier-0 or Tier-2. Tier-2 centers participate in MC production organized by Tier-1 and send the produced MC samples to Tier-1 for distribution in the CMS collaboration. Other than this task, Tier-2 serves analysers by providing local computing resources as well as grid-based analysis support for the whole experiment.

I gratefully acknowledge the CMS collaboration and CERN because any of my thesis work would have not been possible without the contribution of each member of the CMS collaboration and CERN.

Figure 2.12: Magnet system of CMS. Colored pictures are available online [5]. The top shows the yoke(red), outer vacuum tank and the coil. The left bottom shows pictorial view of the magnet system and the right bottom shows the cross section of coils.

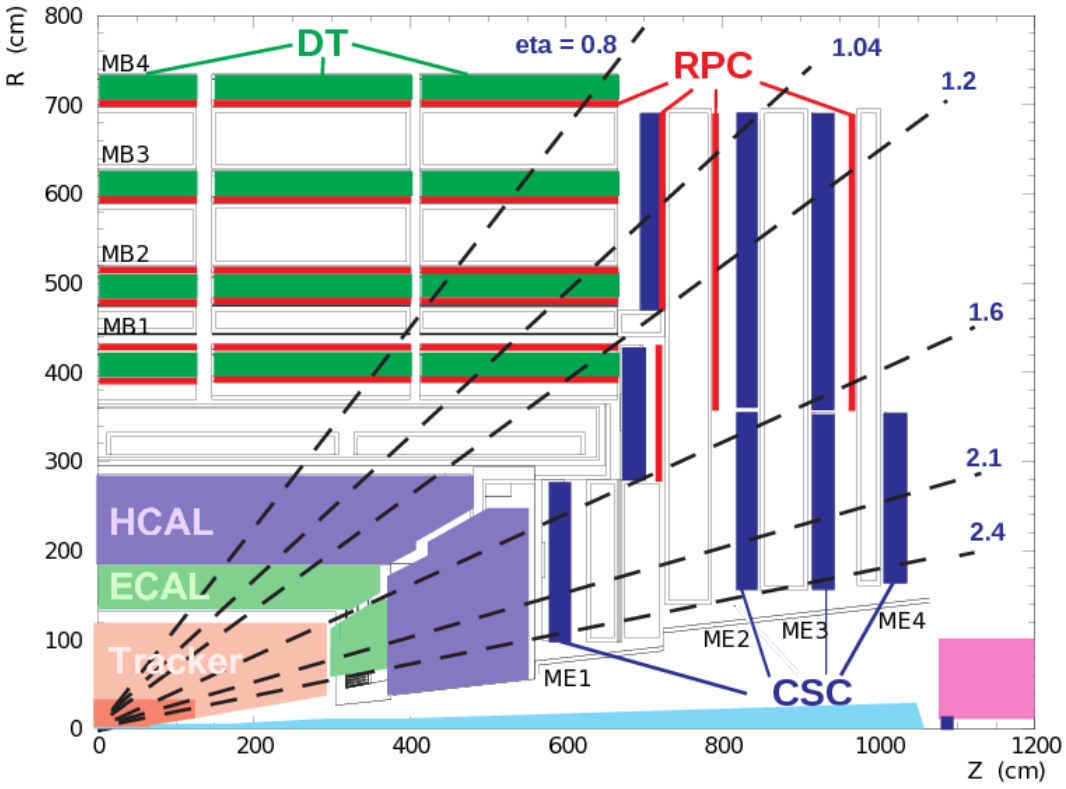


Figure 2.13: Muon system [7].

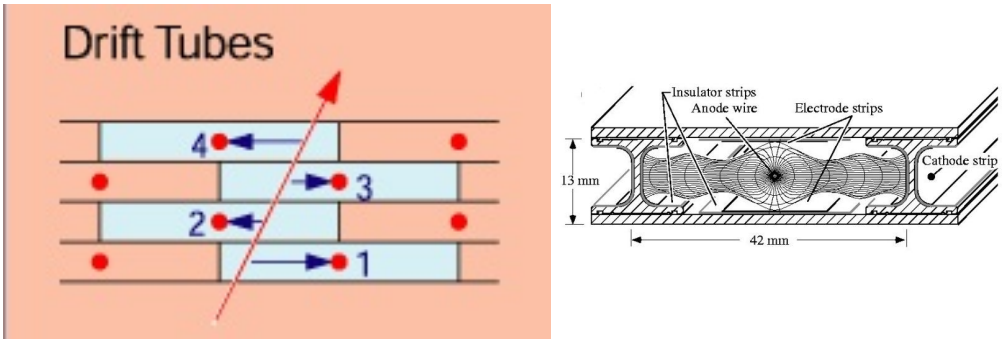


Figure 2.14: Schematic of a superlayer of DT(left) and a tube(right) [8].

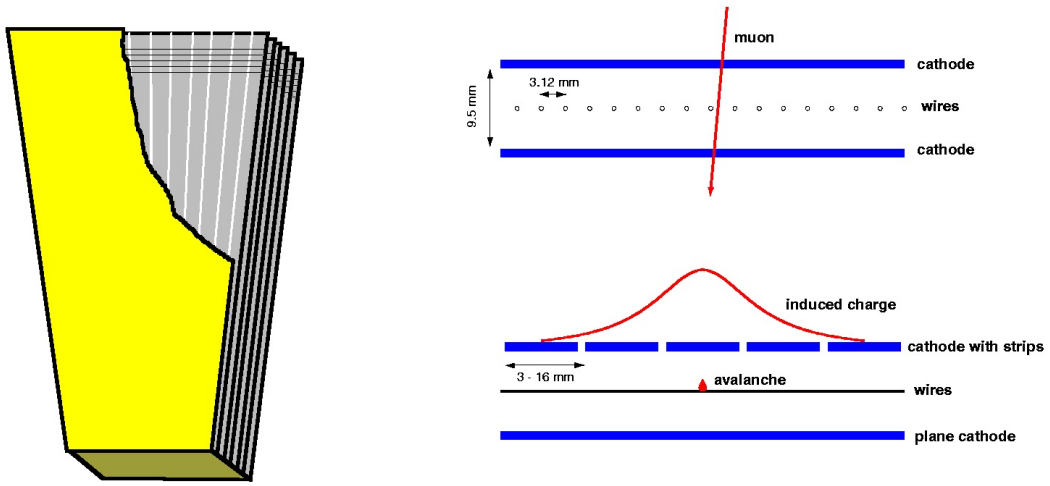


Figure 2.15: Schematic of CSC.

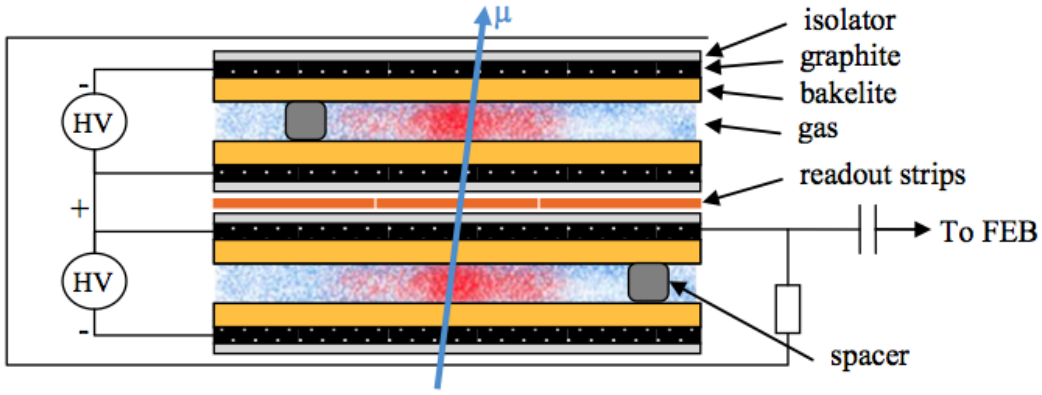


Figure 2.16: Schematic of RPC [9].

Chapter 3

Event Reconstruction

The information such as hits in the tracker and the muon system, and the energy deposit in the calorimeters which is collected by sub-detectors is used to reconstruct objects. Using the available information, we can reconstruct stable particles such as electron, muon, pion, kaon and photon, and other objects used to select and distinguish signal events from backgrounds. This chapter discusses how these objects are reconstructed.

3.1 Tracks

When charged particles($e^{\pm}, \mu^{\pm}, \pi^{\pm}, ...$) traverse in the tracker, it leaves hits in the pixel and the silicon layers. By collecting those hits that are consistent with each other, we can reconstruct the trajectory of the particle, and eventually calculate its momentum. The reality is not that simple though, because there are multiple particles in an event, and a random combination of hits can fake a particle track. Interactions with the material in the tracker can change the direction and the magnitude of the momentum of particles. In addition, there are uncertainties to the sensors that can result in no hits in a silicon layer even though a charge particle passed it. Therefore, we should consider these points when reconstructing tracks from the hits in the tracker.

CMS uses the Combinatorial Track Finder(CTF) [40] as a general tracking algorithm, and this section describes the procedure in detail.

Generation of seeds : finding starting points

The reconstruction of a track starts with seeding, *i.e.* finding the initial values of the track parameters. The seeds can be obtained from the tracking system itself or from external systems. CMS takes an approach to obtain the seed from the innermost layer of the tracker, and find track candidates starting from inside to outside. The most important reason for this choice is that the charged particles have interactions as they fly out, and a non-negligible amount of materials in the tracking system that the particle should go through changes their initial momentum. Therefore, in order to reduce the effect of materials as much as

possible, we start from the innermost layer for track finding. This means that the seed needs to be obtained before the innermost layer of the tracker. Most of the charged particles cross the 3 layers [1] of pixel detector, so the seeds are obtained from the triplets of the pixel measurements. However, because of the geometry of the pixel detector and the inefficiency of the pixel readouts, the seeds are also obtained from combining information from the pixel layers, the vertex/beam spot and the hits in the strip layers [5].

Pattern recognition : finding track candidates

Using the obtained seed, the Kalman filter [41] is used to find tracks. From the seed layer, the filter proceeds to the next layer using coarse track parameters obtained from the seed. Starting from the seed, this is done layer after layer updating the trajectory information after each layer. At each layer, the filter is executed in the following steps.

First, with given the trajectory state, it is determined which of the adjacent layers are expected to be crossed by extrapolating the trajectory. Second, the detectors that are compatible with the trajectory are determined. Third, for each compatible detector, compatible measurements are searched according to the expected trajectory of a charged particle in the magnetic field, considering multiple scattering and energy loss in the tracker material. Fourth, if the measurement is compatible with the trajectory, the track candidate includes this measurement, and the track parameters are updated. When there are multiple measurements compatible with the trajectory, several new trajectories are created. In addition, to account for the case where a charged particle passes the layer without leaving any hits(invalid hit), one additional track is created without adding position measurements. Due to an exponential increase of the number of track candidates, the maximum number of track candidates at each step is limited using χ^2, and the number of valid and invalid hits.

This procedure is repeated to the next tracker layer until either the outer-

[1] At least 3 hits or 2 hits plus a beam spot are needed to fully determine the five track parameters.

most tracker layer is reached or the track does not satisfy the requirements on the number of added layers and uncertainty on the track parameters. As the iteration proceeds, the uncertainty on the position measurements in the $r-\phi$ plane improves due to having larger lever arm as the filter goes inside-out. But, the uncertainty in the $r-z$ plane becomes larger due to the geometry of double-strip layers (longer overlap in the z direction) and absence of z measurements in the single-strip layers where the length of the strip (10-20 cm) is used to constrain the track.

When many hits are present in the tracker, there is a large probability that the same track is reconstructed multiple times (duplicate tracks) from different seeds. So, it is important to select only one track candidate among them in order to avoid reconstructing the same track multiple times. Two tracks are considered duplicate if the fraction of the shared hits (f_{shared}) is greater than 50%. f_{shared} is defined as

$$f_{shared} = \frac{N_{shared}^{hits}}{min(N_1^{hits}, N_2^{hits})} \tag{3.1}$$

where N_{shared} is the number of shared hits and $N_{1(2)}^{hits}$ is the number of hits for track candidate 1(2). If f_{shared} is greater than 0.5, the track candidate with larger number of hits is selected. In case there are same number of hits in both tracks, the track with higher χ^2 is selected. This cleaning continues until none of the track pairs share more than 50 % of their hits.

Track fit : filtering and smoothing

From the pattern recognition explained in the previous subsection, we obtained a list of hits and track parameters with associated covariance matrix for each trajectory. Because the track parameters can be biased by the constraints imposed at the seeding stage, the track is refitted using a combination of the Kalman filter and the smoother which work inside-out and outside-in, respectively.

The Kalman filter starts at the innermost layer with the trajectory parameters obtained from the seed, using the covariance matrix scaled by a large factor. At a given layer, the track parameters and the covariance matrix are updated accounting for the energy loss and multiple scattering in the material. The procedure continues to the last hit.

The smoother starts from the outermost hit, and move backward to the center of the detector. It uses the track parameters from the Kalman filter as the starting condition using the covariance matrix scaled by a large factor. At a given hit, the updated parameters of smoother is combined (as a weighted mean) with the predicted parameters of the Kalman filter(excluding the current hit, the extrapolation starts from the innermost hit).

Track selection : rejecting bad guys

In the average LHC events with jets, the track finder procedure yields a significant fraction of fake tracks, *i.e.*, a random combination of hits that results in a track. These are rejected by applying quality cuts on track η, p_{T}, χ^2 per n.d.f., impact parameters, significance of impact parameters and the number of crossed layers [42].

Iterative tracking

In order to attain a high tracking efficiency and a low fake rate, CMS developed an iterative tracking procedure [43]. The basic idea is to select the high quality tracks first and to find tracks that do not make the high quality tracks but still are good tracks that should be reconstructed. There are 6 iterations in CMS, and the main difference is the initial seeding. For example, the zeroth iteration uses the triplet of pixel detector, and the first iteration uses pixel-pair seeding to recover tracks that failed to make the pixel triplet requirement. In addition, the track parameters are different to maximize the performance.

Performance

Figure 3.1 shows the tracking efficiency measured in single muon and single pion simulations as a function of p_{T} and η. Top left plot shows the efficiency of muon track reconstruction as a function of η for muon p_{T} = 1, 10 and 100 GeV. Top right plot shows the efficiency of muon track reconstruction as a function of p_{T} for different η regions. Bottom left plot shows the efficiency of pion track reconstruction as a function of η for pion p_{T} = 1, 10 and 100 GeV. Bottom right

plot shows the efficiency of pion track reconstruction as a function of p_{T} for different η regions. They show that the tracking efficiency for muons is very close to 100 % in the region of phase space we are interested in for this analysis ($p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.4$). Efficiency for pion is lower than the one for muon because pions decay in flight($\pi^+ \to \mu^+ \nu_\mu$).

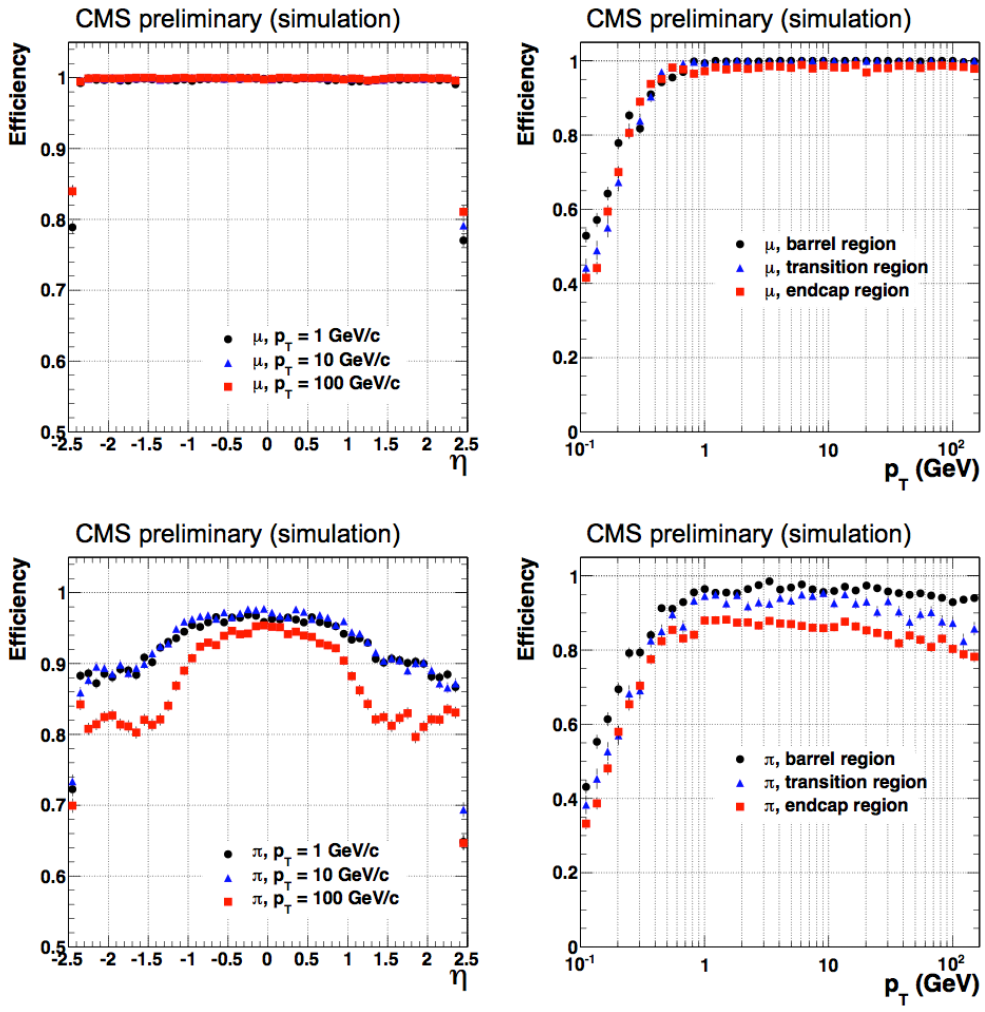


Figure 3.1: Tracking efficiency measured using single pion and single muon simulation as a function of p_{T} and η [10].

3.2 Event Primary Vertex

The primary vertex, the space point where a hard interaction takes place, is reconstructed using the reconstructed tracks described in the previous section. The tracks are selected based on

- compatibility with interaction region : the transverse impact parameter significance with respect to the beam line should be less than 5,
- number of hits in the tracker : more than 4(2) hits should be in the silicon strips(pixel detector)
- track fit quality : χ^2/ndof should be less than 20.

The selected tracks are clustered using the Deterministic Annealing algorithm [44]. At first, only z information is used at the point of the closest approach(PCA) with respect to the beam line. Then, an adaptive vertex fit [45] is performed using the clustered tracks for each primary vertex which has at least two associated tracks. The fit calculates the best estimates of the vertex parameters such as position and covariant matrices. The vertex is retained if the distance between the vertex and the beam line is less than 1 cm.

Figure 3.2 shows the vertex reconstruction efficiency as a function of number of tracks(N_{track}) for data and simulation, and the transverse in the x direction and the longitudinal vertex resolution as a function of N_{track} in the jet-enriched data and the min-bias data. The reconstruction efficiency is close to 100 % with $N_{track} > 2$. The transverse and longitudinal resolutions measured with the min-bias data are less than 30 and 40 μm, respectively, with $N_{track} = 30$.

3.3 Electron

Electrons are reconstructed using the information from the tracker and the ECAL because an electron makes hits in the tracker and makes energy deposit in the ECAL. To reconstruct an electron, ECAL clustering is done first to collect energy spread including bremsstrahlung(the collection is called "supercluster"), and

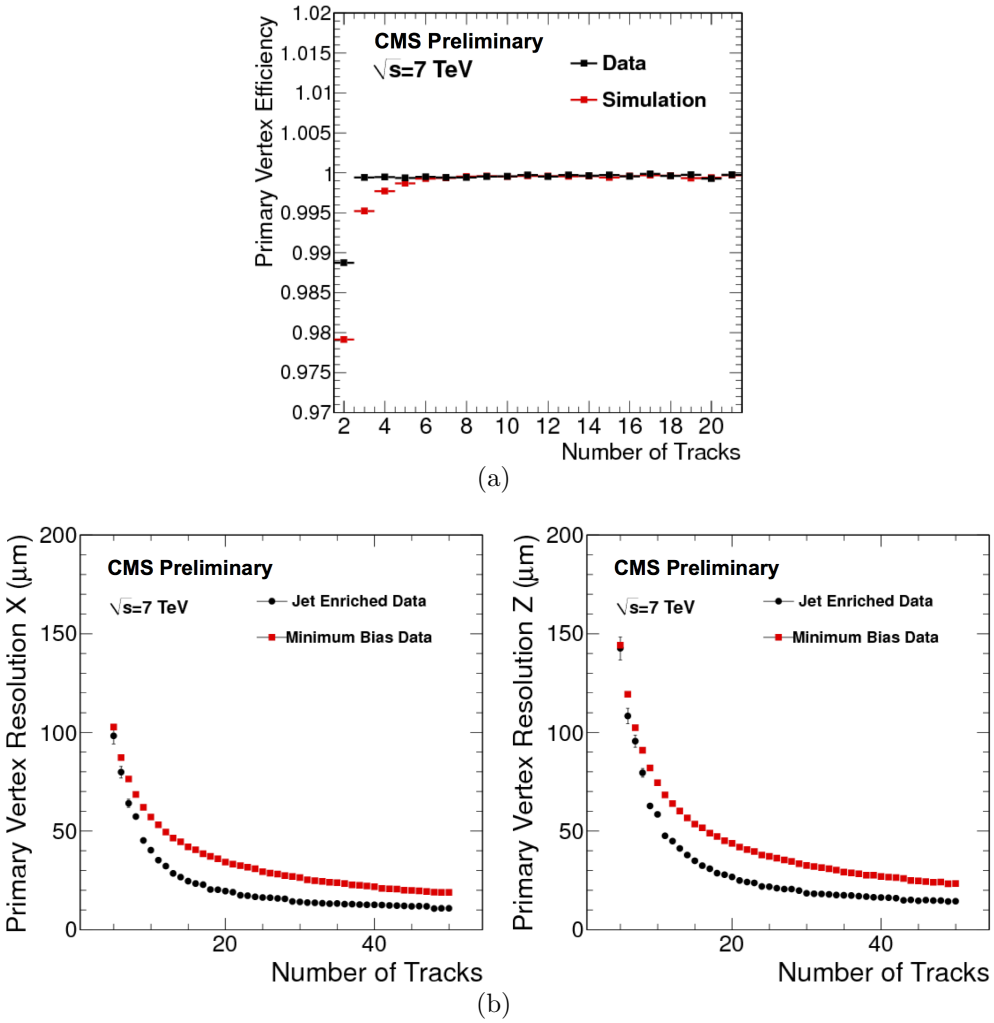


Figure 3.2: (a) vertex resolution in x and z directions as a function of associated tracks for data and Simulation and (b) vertex reconstruction efficiency as a function of number of associated tracks in jet-enriched data and min-bias data [11].

the track reconstruction is performed using a pixel seed found by the supercluster-driven method [46].

The electrons and photons radiated off the electrons form an electromagnetic shower, and make their energy deposit in the ECAL. But, when an electron travels in the tracker which is placed in a strong magnetic field, it radiates photons by bremsstrahlung, and the energy deposit in the ECAL has a spread in the ϕ direction. The size of the energy loss due to bremsstrahlung is significant enough to be included for electron energy calculation. For example, for electrons of energy 10, 30 and 50 GeV, about 35 % of electrons lose more than 70 % of their initial energy via bremsstrahlung, and about 10 % of them lose more than 95 % [46]. Therefore, in order to obtain the initial electron energy, it is critical to collect all bremsstrahlung photons. The algorithm for this purpose is called super-clustering algorithm [46].

CMS employs two algorithms, hybrid for barrel region and island for endcap region [46]. The hybrid algorithm forms a domino of 3 or 5 crystals in η, and dynamically searches for dominos in ϕ separated by a domino with energy less than 100 MeV. The island algorithm starts with making a cluster from a seed crystal with energy deposit above a certain threshold, and collect crystals around it in ϕ and then η direction. The resultant clusters in narrow η-window and wider ϕ-window are then used to construct a supercluster.

The position of the shower(x) is measured as a weighted mean of position of crystals in a cluster [5].

$$x = \frac{\sum_{i}^{N_{crystals}} x_i \cdot \omega_i}{\sum_{i} \omega_i} \tag{3.2}$$

where x_i is the position of the crystal i, and ω_i is the weight for the crystal i given by

$$\omega_i = \omega_0 + \log \frac{E_i}{\sum_j E_j}. \tag{3.3}$$

The logarithmic form of the weight is motivated by the fact that the energy density decreases exponentially in the lateral direction from the shower core.

Once the energy is collected, electron tracks are reconstructed. The first step is to generate seeds to start the tracking algorithm. The energy-weighted mean supercluster position is extrapolated to the interaction point(beam spot) to find compatible hits in the pixel detector assuming both charge hypotheses. The innermost layer is looked for first with loose $\Delta\phi$ and Δz window. If compatible hits are not found in the first layer, search goes on to the next layer. If a compatible hit is found, the z coordinate of the primary vertex is calculated, and the predicted trajectory is used to find compatible hit(s) in the next pixel layer(s). Using the selected seed, compatible hits in the next silicon layer are looked for, and extrapolation is done to the next layer using Bethe-Heitler modeling of electron bremsstrahlung [47] and Gaussian Sum Filter(GSF) [48] which assumes that the *pdf* of the Bethe-Heitler model is a Gaussian mixture. The procedure is continued to the last layer unless two consecutive hits are not found. At each layer, the trajectory state is updated using the weighed mean of the measurement and the prediction. When there are multiple compatible hits, the two most compatible ones from χ^2 test are kept. Finally, a track is created if there are at least five hits.

Figure 3.3 shows the resolution of reconstructed electron energy as a function of the true electron energy measured with simulation [12]. Blue reverse triangle is measured using only tracker information and green upright triangle is measured using only ECAL information. The red star is a combination of the tracker and ECAL measurements. The resolution is estimated the half minimum width that contains 68.3%(Gaussian 1σ width) in the energy distribution as denoted as effective resolution. The red circle corresponds to the width of Guassian fit in the core of the energy distribution. As shown in the figure, precision is dominated by the information from the tracker(blue reverse triangle) and the ECAL(green upright triangle) at low and high energy, respectively. The final resolution comes from the combination of the information weighted by their errors(red star). These errors are evaluated by the half minimum width that contains 68.3%(Gaussian 1σ width) in the energy distribution. The red circle corresponds to the width of the Gaussian

fit in the core of the energy distribution.

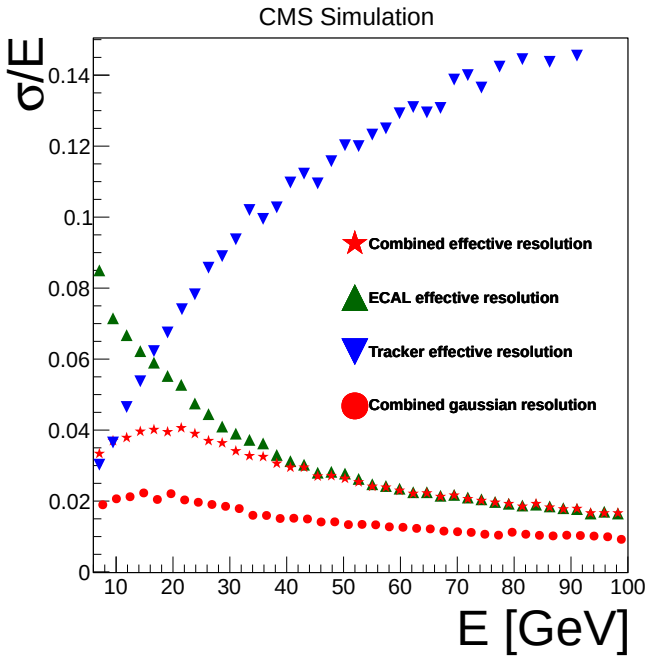


Figure 3.3: Energy resolution of reconstructed electrons as a function of generated electron energy from different information in simulation [12].

3.4 Muon

In CMS there are three types of muons depending on the information used in the reconstruction [5]. They are standalone, tracker and global muons.

The *Standalone* muon reconstruction uses information from the muon system(DT, CSC and RPC), *i.e.* the inner tracker information is not used. It starts with the reconstruction of the track segments in the muon chambers. The digitized electronic signals in DT, CSC and RPC are used to reconstruct hits. Then, the hits in DT and CSC are matched to form the segments. The information of these segments, such as momentum, at the innermost muon chamber is used as a seed to construct a muon track using the Kalman-filter algorithm. The Kalman-filter is an iterative algorithm that updates the track parameters iteratively as it goes to the next station. At each step of the track parameters estimation, if no matching

segment is found then the search continues to the next station taking into account the detector effects such as multiple scattering and energy loss in the material. The procedure goes until the outermost station, updating the track parameters at each step. Then, the Kalman-filter is applied from the outermost to the innermost station, and the track parameters are defined at the innermost station. Finally, the measured muon track is extrapolated to the interaction point and a vertex-constrained fit is performed to obtain the final track parameters.

The *Tracker* muon reconstruction uses information from the inner tracker, *i.e.* the muon system information is not used for the momentum measurement. This approach considers all tracks as potential muon candidates, and checks their compatibility with the muon system. All tracker tracks with $p_{\mathrm{T}} > 0.5$ GeV and $p > 2.5$ GeV are extrapolated to the muon system considering the expected detector effects such as magnetic field, multiple scattering and energy loss in the material. If there is at least one muon segment matched to the extrapolated track, this muon is considered as a Tracker muon. The tracker muon gives good momentum measurement and identification for the low p_{T} muons which are hard to be reconstructed by the muon system because they do not leave enough track segments in the muon system.

The *Global* muon reconstruction uses information from both the tracker and the muon system. For a standalone muon track obtained by the way explained already, a matching with a tracker track is done. The muon track at the innermost station is extrapolated to the last layer of the tracker considering the expected detector effects. If the matching is successful, the Kalman-filter is used to reconstruct the tracks. After that, all reconstructed tracks are fitted again without constraints on the beam spot, using the hits associated with the standalone muons and the hits in the silicon strips. A fit is done again using the tracker hits and the hits in the innermost muon station, and the fit quality is compared with that of the tracker-only fit. This is to detect muon bremsstrahlung or any loss of energy before reaching the muon station.

In summary, there are three approaches for muon reconstruction in CMS. Having multiple algorithms provides more reliable muon reconstruction, and physics

analysis can choose algorithms of their interests. Figure 3.4 shows the momen-

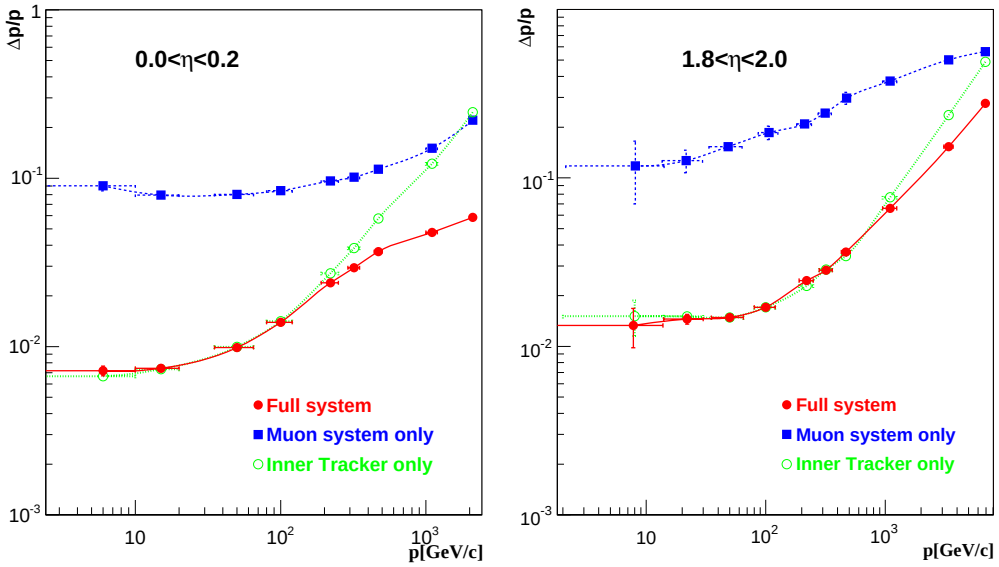


Figure 3.4: Resolution of muon momemtun in $0.0 < |\eta| < 0.2$ and $1.8 < |\eta| < 2.0$. Red, blue and green are global, standalone and tracker muons, respectively [5].

tum resolution of different muon reconstruction algorithms as a function of muon p_{T} in the barrel(left) and endcap(right) [5]. The resolution of standalone muons is dominated by multiple scattering in the material before the muon station at $p_{\mathrm{T}} < 200$ GeV and by the spacial resolution of the muon chambers at $p_{\mathrm{T}} > 200$ GeV. The tracker muons give much better resolution at low p_{T} but the resolution goes up to the same level as the standalone muons at very high p_{T}. For muons with $p_{\mathrm{T}} < 200$ GeV, the resolution is better than 3 %.

3.5 Jet

The existence of gluons and quarks in the event is manifested as a spray of hadrons, "jet". The initial gluons and quarks are hadronized to hadrons, the hadrons decay to another hadrons. This process continues until there is not enough energy to decay to other hadrons. These hadrons in the shower make energy deposit primarily in HCAL. Thanks to the fine granularity of the calorimeters and high precision of tracking in CMS, individual stable particles(electron, muon,

photon, charged and neutral hadron) can be reconstructed by Particle-Flow(PF) algorithm [49]. This algorithm uses all available information from sub-detectors to optimally determine the type of particles, momenta and energies. These particles are used to reconstruct jets.

The jets are reconstructed by clustering individual particles that are considered to originate from the same quark/gluon. There are multiple jet reconstruction algorithms, but they can be classified to two categories, cone-based and sequential algorithms. The cone-based algorithms(Midpoint cone [50], Iterative cone [51, 52], SIS(Seedless Infrared-Safe) cone [53]) start with a seed cone[2] to start the algorithm and calculate the energy and the momentum sum of the particles inside the cone. Then, it continuously merge other particles outside of the cone, recalculate the energy and momentum. The algorithm continues until the direction of the cone does not change, i.e. the cone becomes stable. These algorithms(except SIS cone) have problems that the stable cone changes, *e.g.* two cones are merged or an existing cone disappears, by adding an extra-soft particle(Infrared safety) or splitting a particle into multiple particles(collinear safety).

The sequential algorithms are designed to be insensitive to these problems. The algorithms define a distance between particles(which does not necessarily have to be a geometrical distance), and repeatedly combine the closest pair until some stopping conditions are satisfied. The algorithms start with defining two distances, the distance between particle i and j(d_{ij}) and the distance between particle i and the beam(d_{iB}). If the minimum of d_{ij} and d_{iB} is d_{ij} then the two particles are combined and used as a new particle to be paired in the next iteration, and if the minimum is d_{iB} then the particle i is called a jet and removed from the list of particles because it is considered as a radiation from the beam. The algorithms continue combining particles until all particles are clustered into jets. The distances are defined as

$$d_{ij} = min\left(k_{ti}^{2p}, k_{tj}^{2p}\right)\frac{\Delta_{ij}^2}{R^2}, \tag{3.4}$$

$$d_{iB} = k_{ti}^{2p} \tag{3.5}$$

[2] As the name says, SIS cone algorithm does not start with a seed cone, but tests all possible cones.

where k_{ti} is the transverse momentum of particle i, $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ with y_i and ϕ_i being the rapidity and the azimuthal angle of particle i, respectively, and the parameter R is the scale that determines the distance of the reconstructed jets. The parameter p controls the relative weight between the energy(k_t) and the geometrical scales(Δ_{ij}), and this is the parameter that distinguishes three different algorithms, $k_t(p = 2)$ [54], Cambridge-Aachen($p = 0$) [55] and anti-$k_t(p = -2)$ [56]. The clustering uses FastJet algorithm [57] which significantly improves the timing of calculation, and provides the jet area used for subtraction of contribution from pileup. The jets used in this analysis are reconstructed using anti-k_t algorithm with $R = 0.5$.

Due to the non-linear calorimeter response of the CMS detector, the measured jet energy can be different from the energy of the true parton which initiated the jet. CMS employs a jet energy correction(JEC) method [58] factorized into multiple levels, the L1, L2, L3 and the residual L2L3 corrections for data as shown in Figure 3.5.

The L1 correction is the pileup correction, *i.e.* to remove the offset energy produced by pileup. The correction factor(C_{L1}) is defined by [58]

$$C_{L1} = 1 - \frac{(\rho - \langle \rho_{UE+noise} \rangle) \cdot A_{jet}}{p_T^{raw}} \tag{3.6}$$

where ρ is the per-event energy density, $\langle \rho_{UE+noise} \rangle$ is the average energy density of Underlying Event(UE) and noise which is measured using events that contain only one reconstructed vertex, *i.e.*, no pileup, A_{jet} is the jet area and p_T^{raw} is the uncorrected transverse momentum of the jet. The L2 correction is to make the jet energy response flat in η. At a given η the response is corrected so that it becomes the same level with the central region, $|\eta| < 1.3$. So, it is a relative correction. The correction factors are derived either from MC or using data-driven method (di-jet balance technique [58]). The L3 correction is to make the jet energy response flat in p_{T}. The central region, $|\eta| < 1.3$, is used as a reference for the correction. Apart from the L2 correction, L3 correction is an absolute correction such that the corrected jet p_{T} is same as the p_{T} of the parton that initiated the jet. The correction factors are derived either from MC or using data-driven method(Z/γ^*+jet balance

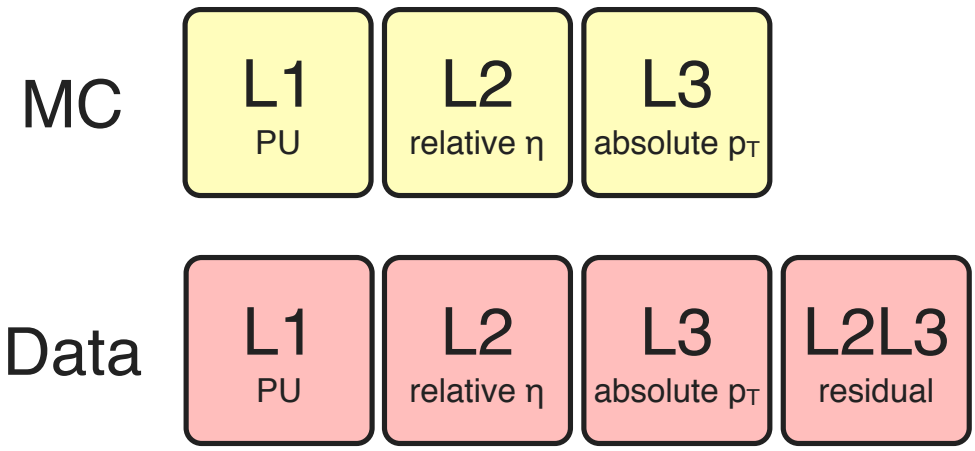


Figure 3.5: Factorized method for jet energy correction.

technique [58]). For L2 and L3 corrections, residual corrections are applied to data in order to account for the small differences between data and MC.

Figure 3.6 shows the uncertainty on the JEC factor as a function of jet p_{T} at $|\eta_{jet}| = 0, 2.0,$ and 2.7 They show that the uncertainty is less than 3 % for jets with $p_{\mathrm{T}} > 30$ GeV at $|\eta_{jet}| = 0$ and 2.0. The uncertainty becomes larger in the forward region, *e.g.* it is as large as 8 % for jets with $p_{\mathrm{T}} > 30$ GeV at $|\eta_{jet}| = 2.7$.

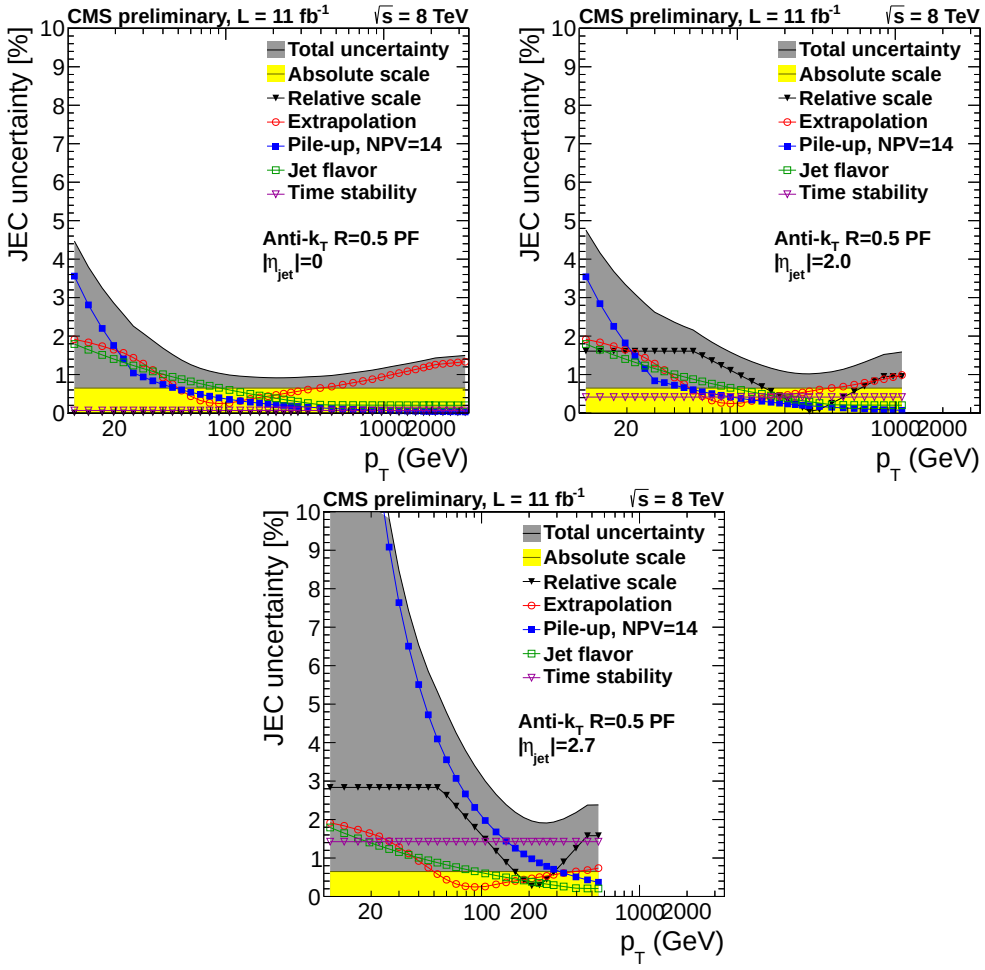


Figure 3.6: JEC uncertainty measured in data of $\mathcal{L} = 11$ fb^{-1} [13].

3.6 Missing Transverse Energy

When neutrinos are produced at colliders, they do not leave any signatures in the detector, so they can not be reconstructed. However, we can infer the existence of them, or any weakly-interacting particles, by computing the imbalance in the vector sum of transverse momenta of the reconstructed particles. The transverse momentum of the initial particles is zero [3]. By momentum conservation, the total momentum of the particles produced after collisions should be 0 as well. So, if the transverse momenta of all particles in the final state are summed up, the negative value of the vector sum should correspond to the transverse momentum sum of neutrinos. Thus, we define "Missing Transverse Energy(MET)" as the negative value of the sum of all PF candidates momenta,

$$\overrightarrow{\mathrm{MET}} = -\sum_{i}^{\text{All PF candidates}} \overrightarrow{p_{\mathrm{T}}}(i). \tag{3.7}$$

The ϕ distribution of true MET should be flat because of the rotational symmetry of collisions with respect to the beam axis. However, possibly due to the anisotropic detector response, inactive calorimeter response, detector misalignments, and displacement of the beam spot, the ϕ distributions of MET in both MC and data are not a flat, but a sinusoidal shape with a period 2π. Thus, we correct this by shifting the origin of the coordinates in the transverse momentum plane for x and y components individually. Since the size of the shift increases linearly as a function of number of reconstructed primary vertices, the form of correction is given by

$$\alpha + \beta N_{\text{vertex}} \text{ (GeV)} \tag{3.8}$$

where α and β shown in Table 3.1 are constants and N_{vertex} is the number of primary vertices.

The performance of MET reconstruction is severely degraded in the high luminosity environment because of the random contribution of paricles from pileup to the MET calculation. Figure 3.7 shows the resolution of parallel and perpendicular

[3] It is not perfectly zero because partons have transverse movements inside the proton. But, the energy of their motion is at most a few hundred MeV which is much less than the resolution of measurements.

Table 3.1: Parameters used for XY shift correction for MET.

		α(GeV)	β(GeV)
MC	correction for X	-3.00×10^{-2}	-6.62×10^{-2}
	correction for Y	3.71×10^{-1}	-1.49×10^{-1}
Data	correction for X	3.54×10^{-1}	2.65×10^{-1}
	correction for Y	1.89×10^{-1}	1.66×10^{-1}

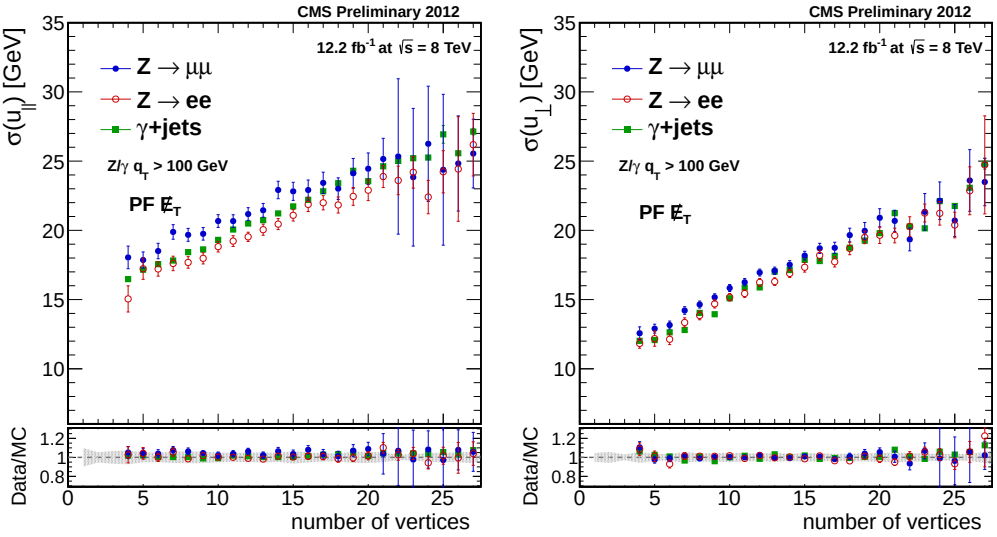


Figure 3.7: The resolution of parallel(left) and perpendicular(right) components of the hadronic recoil in the data events with Z or photon as a function of reconstructed vertices [14].

components of the hadronic recoil in the data events with Z or photon as a function of reconstructed vertices [14]. These events do not have genuine source of MET, *i.e.* no neutrinos in the final state, so MET originates from the mis-measurement of objects. Since electrons, muons and photons are well-measured, the dominant contribution comes from the momentum mis-measurement of the recoiling jets. The plots clearly show that the resolution of the recoiling jet momentum increases as the number of vertices increases. So, we use another definition of MET which is calculated with only charged PF candidates associated with the event primary vertex. Because the particles from other vertices than the event primary vertex are excluded in the MET calculation, the MET calculated using this method is independent of number of pileups. This MET definition is called "trkMET" and the exact definition is

$$\overrightarrow{\text{trkMET}} = -\overrightarrow{p_{\text{T}}}(\ell_1) - \overrightarrow{p_{\text{T}}}(\ell_2) - \sum_{i}^{\text{All charged PF candidates}} \overrightarrow{p_{\text{T}}}(i) \tag{3.9}$$

where $\overrightarrow{p_{\text{T}}}(\ell_1)$ and $\overrightarrow{p_{\text{T}}}(\ell_2)$ are the transverse momenta of the leptons. The charged PF candidates must meet the following requirements.

- The longitudinal impact parameter of the track matched to the PF candidate with respect to the event primary vertex should be less than 0.1 cm.
- ΔR between the track matched to the PF candidate and the leptons should be larger than 0.1 in order to avoid counting leptons twice.

Figure 3.8 shows the distributions of PFMET and trkMET for the Drell-Yan and the signal in simulation. The events with the number of reconstructed vertices(N_{vtx}) greater than 20 and less than 5 are drawn separately. In case of Drell-Yan process, PFMET increases significantly as N_{vtx} increases, while trkMET does not depend on N_{vtx}. On the other hand, in case of signal process, the dependence of both PFMET and trkMET on N_{vtx} is small. This indicates that if we use PFMET as a cut variable, the rejection power of $Z/\gamma^* \to \ell\ell$ background will be weaker as N_{vtx} increases while trkMET does not show this dependence. However, trkMET has a shortcoming that it has a longer tail than PFMET in $Z/\gamma^* \to \ell\ell$ events. Therefore, we use both MET variables to suppress $Z/\gamma^* \to \ell\ell$.

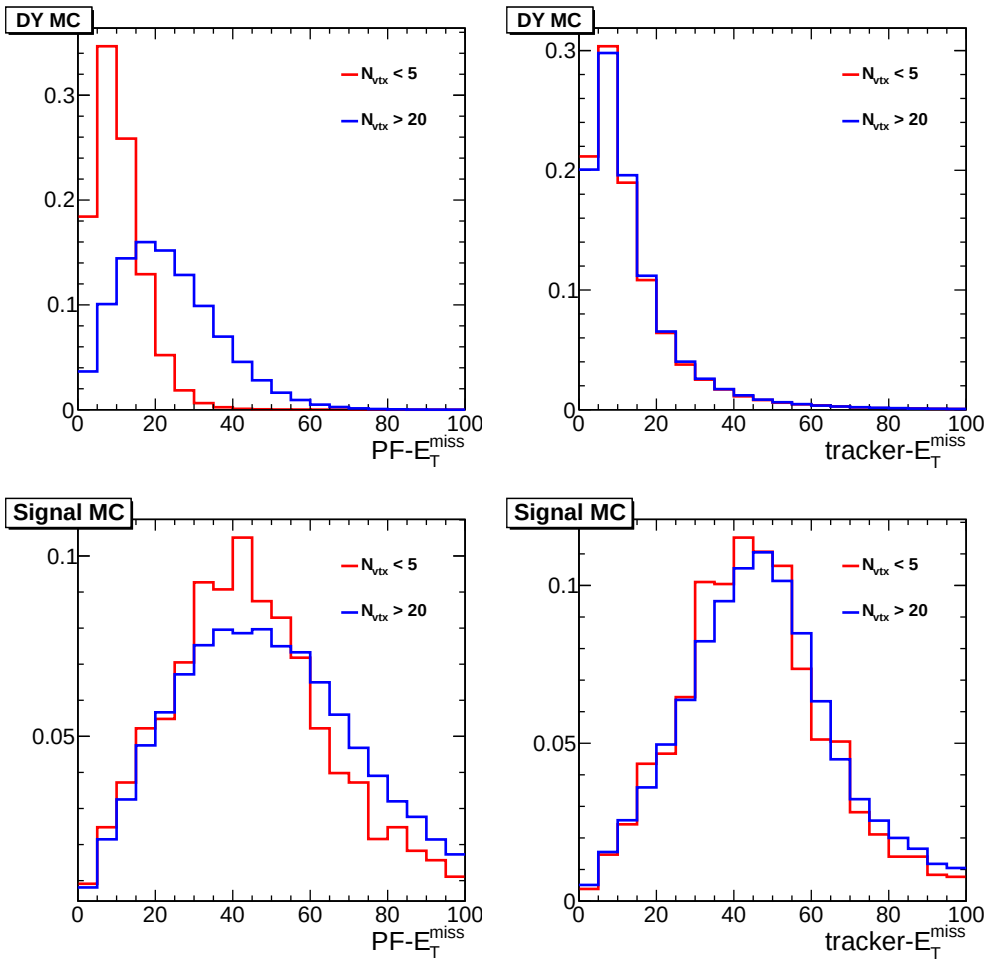


Figure 3.8: Distributions of PFMET and trkMET for the Drell-Yan(top) and the signal(bottom) in simulation. The events with the number of reconstructed vertices(N_{vtx}) greater than 20(blue) and less than 5(red) are drawn separately.

3.7 B-tagging

The presence of a bottom quark in an event is manifested by the presence a displaced secondary vertex. The bottom quark is hadronized to a b-hadron($B^0, B^{\pm}, \Lambda_b^0, ...$), and traverses over a measurable distance($c\tau \approx 500\ \mu$m) before decaying into other particles. Therefore, some b-tagging algorithms use the distance between the primary vertex and the secondary vertex. In addition, using the fact that the impact parameter of the displaced tracks with respect to the primary vertex is larger than the one of the tracks from the primary vertex, some algorithms use the information of the impact parameter(IP). In this analysis, we use an algorithm that uses the IP information.

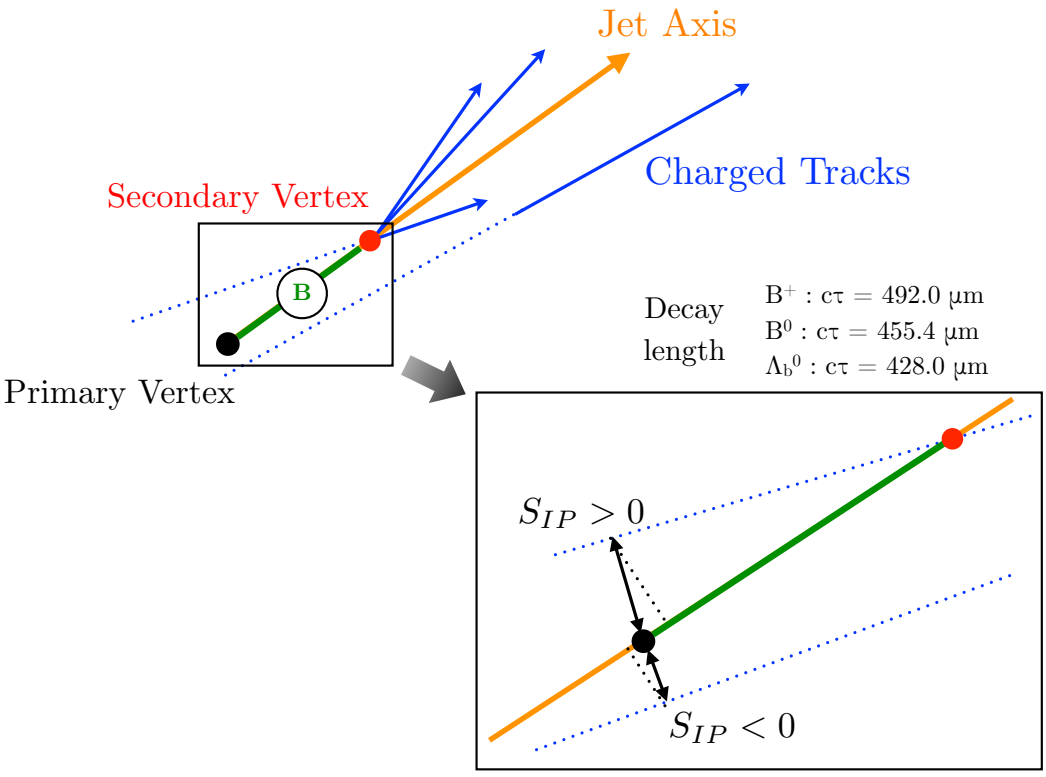


Figure 3.9: A schematic of b-tagging algorithm.

The IP is calculated in 3D thanks to the good resolution of the pixel detector in z direction. The IP can be signed(+ or −) depending on the position of the

associated track. The sign is obtained from the sign of the scalar product of IP vector from the primary vertex and the direction vector of the jet to which the track belongs as shown in Figure 3.9. Ideally, for the decays with a sizable lifetime the IP should be large and positive, but it is not always positive in case the real direction of the B meson is different from the direction of the jet. But, they still tend to be positive. For the decays with very short lifetime or random tracks, the IP is small and symmetric around 0. These are well shown in the left plot of Figure 3.10.

The b-tagging algorithm used in this analysis is "Track Counting High Efficiency(TCHE) [59]". This algorithm uses the impact parameter significance, $S_{IP} = IP/\sigma_{IP}$ where σ_{IP} is the uncertainty of the IP measurement, as a discriminating variable. The algorithm requires at least 2 tracks to have S_{IP} above a given threshold. Thus, the discriminator is the S_{IP} of the jet which has the second highest σ_{IP}.

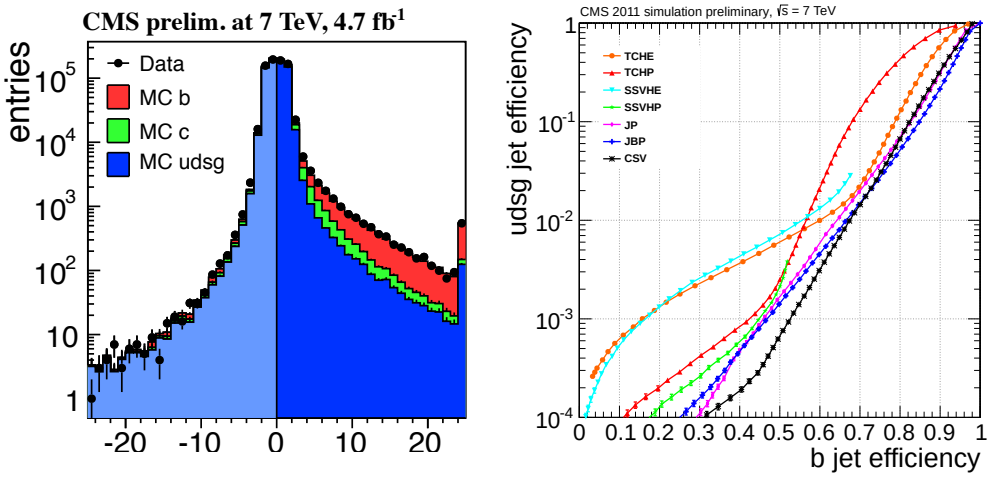


Figure 3.10: Distribution of the discriminating variable for the b-tagging is shown on the left. MC and data shows a good agreement. The b-tagging efficiency(x-axis) and mis-tag rate of udsg jets(y-axis) are shown on the right.

Figure 3.10 shows the distribution of the discriminator of the TCHE algorithm in MC and data, and the b-tagging efficiency(x-axis) and mis-tag rate of udsg jets(y-axis). The tagging efficiency of TCHE at the same mis-tag rate is not the best as shown on the right plot. At the working point of this analysis which is

close to 8 % mistag rate, b-tagging efficiency is about 10 % lower than the best-performing tagger. But, this tagger was selected for this analysis because it is the best-performing tagger of the ones that show good agreement between MC and data.

Chapter 4

Event Selection

Using the reconstructed objects discussed in the previous chapter, we can select the Higgs events by selecting events with two leptons and MET. But, the selected events contain not only signal, but also backgrounds. A challege is that the production rate of the backgrounds is much larger than that of signal. Fig. 4.1

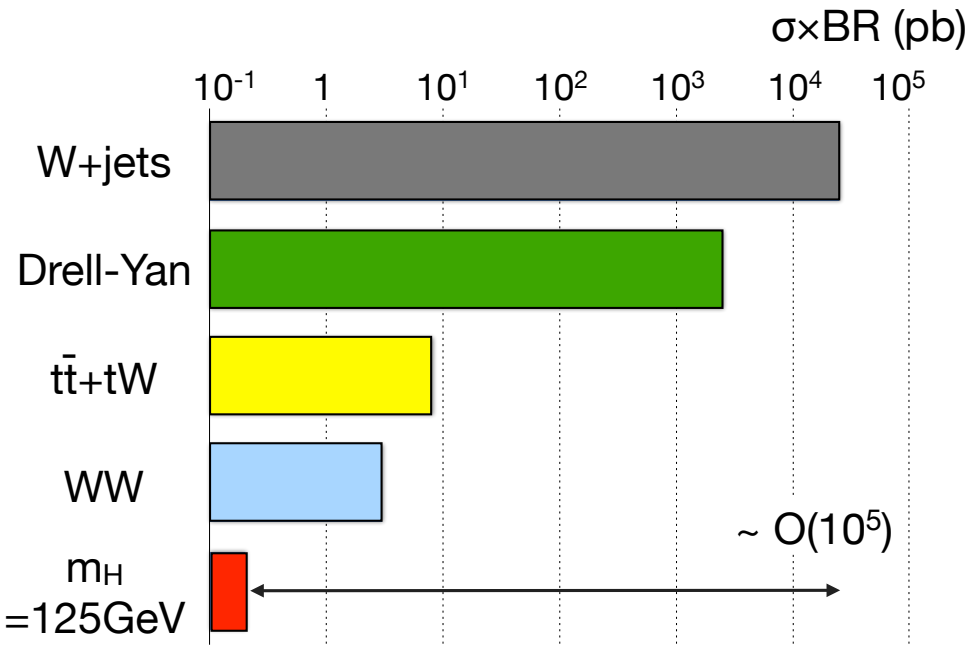


Figure 4.1: The cross section × branching ratio($\sigma \times BR$) for the major background processes and the SM m_{H}=125 GeV hypothesis. The branching ratio is for the leptonic decay(electron/muon + neutrino) of W or Z.

shows the cross section × branching ratio($\sigma \times BR$) for the major background processes and the SM m_{H}=125 GeV hypothesis. The production rate of backgrounds is much larger than the signal. For example, the $\sigma \times BR$ of the W + jets is a factor of $\mathcal{O}(10^5)$ larger than that of signal. Therefore, we need to apply stringent requirements which select the signal events with high efficiency, and suppress background events with a high rejection rate.

The backgrounds can be divided into two categories depending on the way they can be suppressed. The first kind is reducible backgrounds which can be

suppressed by tighting the requirements on the object selection. The reducible backgrounds are $Z/\gamma^* \to \ell\ell$, $t\bar{t}/tW$, W + jets, $W\gamma$, $W\gamma^*$, WZ/ZZ and $Z \to \tau\tau$ where for WZ/ZZ there are more than 2 leptons in the final state. The other kind is irreducible backgrounds which have the exactly same final states as signal, therefore they can not be suppressed by tightening object selections, but by using event kinematics. The irreducible backgrounds are WW and WZ/ZZ where for WZ/ZZ there are only 2 leptons in the final state, *e.g.*, missing leptons.

This chapter describes the selection criteria to suppress the reducible backgrounds. The requirements are imposed to trigger, vertex, electron, muon, jets, and top-tagging selections. In addition, there are requirements to suppress particular backgrounds. All these requirements are designed to select events containing a pair of Ws to make a subset of sample with a reasonable signal-to-background ratio for signal extraction. The next chapter describes the selection to suppress irreducible backgrounds to extract the signal events.

4.1 Trigger

As discussed in section 1.4.3, $H \to W^+W^- \to 2l2\nu$ events have trailing lepton whose transverse momentum goes down very low for low m_{H} hypotheses. But, triggering low p_{T} leptons is very challenging because of large background events. Therefore, in order to record signal events with high efficiency, we need to trigger on the leading lepton, or on both leptons. The option to trigger on the leading lepton is not possible because the requirements should be very tight to maintain a sustainable bandwidth. Thus, we trigger on both leptons. The double-lepton triggers we designed for this analysis have high efficiency for signal events, but are loose enough to collect events in the several control regions we use for various studies. We also use control region triggers that allow fake rate and lepton selection efficiency measurements which are described in section 7.2 and 6.2, respectively, with a precision good enough for this analysis.

4.1.1 Analysis Triggers

The analysis triggers, the triggers used to select signal and control region events, impose tight cuts to maintain the rate. For the electron HLT objects there are requirements on the kinematics(p_{T} and η), the shower shapes, the track-to-cluster matching and the track/calorimeter isolation. For the muon HLT objects there are requirements on the kinematics(p_{T} and η). The naming convention and the corresponding cut variables with their cut values are listed in Table 4.1. The H/E is the ratio of energy deposit in HCAL to that of ECAL. The $\sigma_{\eta\eta}$ is the weighted sum of η difference between the seed crystal and the 5x5 crystals surrounding it. The $|\Delta\eta|(|\Delta\phi|)$ is the difference in absolute value between the center of the supercluster and the direction of the track trajectory in $\eta(\phi)$ direction. The $|\frac{1}{E} - \frac{1}{p}|$ is the difference between the reciprocal of supercluster energy and the reciprocal of the track momentum. The ECalIso/$\mathrm{E_T}$, HCalIso/$\mathrm{E_T}$ and TrkIso/$\mathrm{E_T}$ are the sum of the transverse energy within $dR < 0.3$ around the center of the energy deposit or the track trajectory divided by the transverse energy, $\mathrm{E_T}$. The details of these variables are discussed in section 4.3.

In this analysis we use double-lepton triggers as shown in Table 4.2 and single-lepton triggers as shown in Table 4.3. The double-lepton triggers require two HLT objects to be present, and each of them is required to match an L1 seed. The offline lepton p_{T} requirement is 20/10 GeV, so the online lepton p_{T} requirement is a bit looser, 17/8 GeV, in order to be avoid loosing events by online selection. In addition, the longitudinal distance between the two vertices of the leptons is required to be less than 0.2 cm in order to ensure that the two leptons are coming from the same interaction point. The requirement of the single lepton triggers is tighter than that of double-lepton triggers to maintain the bandwidth. Single-lepton triggers help enhance the overall trigger efficiency by selecting events that the double-lepton triggers missed.

Because online variables are constructed using more simplified algorithms than the offline variables, they do not exactly correspond to the offline ones. To account for this, we measure the trigger efficiency with respect to the offline selection, and apply corrections accordingly. The details on this can be found in 6.2.3

Table 4.1: Summary of requirements applied to electrons in the analysis triggers. The selection requirements are shown for electrons in the barrel (endcap). The abbreviation in the names means L=Loose, VL=Very Loose, T=Tight, and VT=Very Tight.

name	criterion
CaloId_T	H/E $<$ 0.15(0.10)
	$\sigma_{\eta\eta}$$<$ 0.011 (0.031)
CaloId_VT	H/E $<$ 0.05(0.05)
	$\sigma_{\eta\eta}$$<$ 0.011 (0.031)
TrkId_VL	$\|\Delta\eta\|<$ 0.01 (0.01)
	$\|\Delta\phi\|<$ 0.15 (0.10)
TrkId_T	$\|\Delta\eta\|<$ 0.008 (0.008)
	$\|\Delta\phi\|<$ 0.07 (0.05)
CaloIso_VL	ECalIso/E_T $<$ 0.2 (0.2)
	HCalIso/E_T $<$ 0.2 (0.2)
CaloIso_T	ECalIso/E_T $<$ 0.15 (0.075)
	HCalIso/E_T $<$ 0.15 (0.075)
CaloIso_VT	ECalIso/E_T $<$ 0.05 (0.05)
	HCalIso/E_T $<$ 0.05 (0.05)
TrkIso_VL	TrkIso/E_T $<$ 0.2 (0.2)
TrkIso_T	TrkIso/E_T $<$ 0.15 (0.075)
TrkIso_VT	TrkIso/E_T $<$ 0.05 (0.05)
WP80	H/E $<$ 0.10(0.05)
	$\sigma_{\eta\eta}$$<$ 0.01 (0.03)
	$\|\Delta\eta\|<$ 0.007 (0.007)
	$\|\Delta\phi\|<$ 0.06 (0.03)
	$\|\frac{1}{E}-\frac{1}{p}\|<$ 0.05 (0.05)
	ECalIso/E_T $<$ 0.15 (0.10)
	HCalIso/E_T $<$ 0.10 (0.10)
	TrkIso/E_T $<$ 0.05 (0.05)

Table 4.2: Double-lepton triggers used to collect signal events. The naming convention is shown in Table 4.1.

Double-lepton trigger name	L1 seed
HLT_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_	L1_DoubleEG_13_7
Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_v[15-19]	
HLT_Mu17_Mu8_v[16-22]	L1_DoubleMu_10_Open
HLT_Mu17_TkMu8_v[9-14]	OR L1_DoubleMu_10_3p5
HLT_Mu17_Ele8_CaloIdT_CaloIsoVL_	L1_Mu12_EG7
TrkIdVL_TrkIsoVL_v[4-9]	
HLT_Mu8_Ele17_CaloIdT_CaloIsoVL_	L1_MuOpen_EG12
TrkIdVL_TrkIsoVL_v[4-9]	OR L1_Mu3p5_EG12

Table 4.3: Single-lepton triggers used to collect signal events. The naming convention is shown in Table 4.1.

Single-lepton trigger name	L1 seed
HLT_Ele27_WP80_v[8-11]	L1_SingleEG20 OR L1_SingleEG22
HLT_IsoMu24_eta2p1_v[11-15]	L1_SingleMu16er

where the measurement on the trigger efficiency is discussed.

4.1.2 Utility Triggers

The efficiency measurements of the lepton selection are performed by Tag-And-Probe method [60] using $Z/\gamma^* \to \ell\ell$ events. In order to use this method, we select a pure sample of $Z/\gamma^* \to \ell\ell$ events to reduce bias due to selecting non-prompt leptons from other background processes or pileup. Apart from the analysis triggers, for this measurement we do not have to select all available events, but a pure sample with an adequate statistics. The single lepton triggers used to collect signal events, listed in Table 4.3, also can be used to select $Z/\gamma^* \to \ell\ell$ events. The leading lepton is likely to be triggered, making the trailing lepton unbiased sample that covers a wide range of kinematic region that stretches to low lepton p_{T}.

In order to estimate jet-induced backgrounds such as $\mathrm{W}+\mathrm{jets}$ that have a non-prompt lepton that passes the full lepton selection, we use "fake rate" method. The details of this method are discussed in 7.2. In this method we define a loose lepton selection, and calculate the ratio, "fake rate", for a lepton that pass a loose selection to pass the full selection, using a data sample of single-lepton events dominated by QCD processes. The assumption of this method is that the jets in $\mathrm{W}+\mathrm{jets}$ events and QCD events are not different, and the fake rate measured using the two samples is same. This is possible only if the kinematics, particularly p_{T}, of the progenitor of the jets are same [1]. In data, we do not have a handle to control the p_{T} of the progenitor, so we use the p_{T} of the jet that is on the other side of the lepton(away jet). This is justified by the fact that QCD events are dominated by di-jet events where the two jets are likely to be back-to-back. By choosing an appropriate p_{T} cut of the away jet, we can select relevant event samples to measure fake rate and the systematic uncertainty due to limitation of controlling the progenitor p_{T} is covered by varying the away jet p_{T} threshold.

[1] Consider a fake lepton of p_{T}=20 GeV. The fake rate in case that the progenitor p_{T}=25 GeV(a) is different from the case that the progenitor p_{T}=100 GeV(b). (b) has more extra energy (100 - 20 = 80 GeV) than (a), so it has larger probability to be un-isolated giving lower fake rate. Other issue is the composition of the progenitor in the two samples, *e.g.* gluon-quark ratio and quark flavor, but this is a second-order effect.

Because the leptons in the data sample collected by the analysis triggers are also trigger objects, the loosest possible "loose" definition is the trigger requirement of the analysis triggers. We devised a set of single-lepton triggers that have a loose or the same requirements on leptons as the double-lepton triggers. In order to cover large range of lepton p_{T}, we use several triggers with different lepton p_{T} thresholds. In addition, as fake rate is measured with a p_{T} cut on the away jet we can obtain more events in the relevant phase space by applying p_{T} cut on the away jet. The single lepton triggers used for fake rate measurement are listed in Table 4.4. These triggers provide sufficient statistics for measurement of fake rate and estimation of its systematic uncertainty.

Table 4.4: Utility triggers for fake rate method. The identification and isolation requirements for electrons are described in Table 4.1. Jet30 in the electron triggers means that there should be at least one jet of $p_{\mathrm{T}} > 30$ GeV.

Trigger name	L1 seed
HLT_Ele8_CaloIdT_TrkIdVL_v[2-5]	L1_SingleEG5
HLT_Ele8_CaloIdT_CaloIsoVL_	L1_SingleEG7
TrkIdVL_TrkIsoVL_v[12-15]	
HLT_Ele17_CaloIdT_CaloIsoVL_	L1_SingleEG12
TrkIdVL_TrkIsoVL_v[3-6]	
HLT_Ele8_CaloIdT_CaloIsoVL_	L1_SingleEG7
TrkIdVL_TrkIsoVL_Jet30_v[3-7]	
HLT_Ele17_CaloIdT_CaloIsoVL_	L1_SingleEG12
TrkIdVL_TrkIsoVL_Jet30_v[3-7]	
HLT_Mu8_v[16-18]	L1_SingleMu3
HLT_Mu17_v[3-5]	L1_SingleMu12

4.2 Event Primary Vertex

The offline primary vertices are required to be within 24 cm from the center of the detector in z direction. It should be within 2 cm from the beam spot in the radial direction. The degrees of freedom of the vertex fit should be 4 or larger.

At high luminosity collisions, there are multiple proton-proton interactions in the same bunch crossing. In those interactions there is usually only one interaction that is of interest for our analysis, which triggered that event. These interactions tend to be associated with energetic objects, while the other interactions are mostly inelastic scatterings that produce soft objects. Therefore, we choose the event primary vertex by selecting the primary vertex with the largest scalar sum of p_{T}^2 of tracks associated with the vertex.

4.3 Electron

An electron candidate is reconstructed if there is a track and a supercluster energy deposit compatible with the track momentum. There are multiple sources of electrons which do not originate from hard interactions. These electrons are called "non-prompt electrons" as opposed to the "prompt electrons" originated from hard interactions. Of the non-prompt electrons, those that originate from jets is called "fake electrons or fakes". We can get fake electrons if

- Early conversion : π^0 decays to two photons, and one of the two photons undergoes an asymmetric conversion to an electron-positron pair, *i.e.*, one of the them carries most of the photon momentum,
- Charge exchange : $\pi^{\pm}$ is converted to π^0 in the ECAL, and the π^0 decays to a pair of photons,
- Random combination : $\pi^{\pm}$ overlaps π^0(a track of $\pi^{\pm}$ randomly matches a supercluster energy deposit by π^0), and
- Heavy flavor decay : B or D hadron decays semi-leptonically.

In order to suppress fake electrons, we apply selections composed of requirements on the identification, isolation and impact parameter. Other source of non-prompt electrons is a photon conversion to a pair of electron and a positron in the material. If the conversion is asymmetric, *i.e.* one particle carries most of the photon momentum, that particle can be selected as an electron. Thus, we impose requirements for conversion rejection.

For the electron identification we use a BDT-based multivariate approach [61]. The Boosted Decision Tree(BDT) [62] is a multivariate algorithm that uses distributions of multiple variables and their correlations to optimally distinguish one hypothesis from the other. The training is done with 2011 data; $Z/\gamma^* \to \ell\ell$ events for signal and QCD-dominated events collected by the fake rate triggers listed in section 4.1. In order to increase the separation, and to mitigate a possible bias due to the trigger selections, a set of pre-selection cuts that are as tight as the trigger requirements is applied as follows :

- $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.5$
- $\sigma_{i\eta i\eta} < 0.01/0.03$ (barrel/endcap)
- $|\Delta\phi_{in}| < 0.15/0.10$ (barrel/endcap)
- $|\Delta\eta_{in}| < 0.007/0.009$ (barrel/endcap)
- H/E $< 0.12/0.10$ (barrel/endcap)
- $\dfrac{\sum_{\text{tracks with dR<0.3}} E_{\mathrm{T}}}{p_{\mathrm{T}}} < 0.2$
- $\dfrac{\left(\sum_{\text{ECAL with dR<0.3}} E_{\mathrm{T}}\right) - 1}{p_{\mathrm{T}}} < 0.2$
- $\dfrac{\sum_{\text{HCAL with dR<0.3}} E_{\mathrm{T}}}{p_{\mathrm{T}}} < 0.2$

The definition of the variables is already discussed in section 4.1.

The input variables to the BDT are the following. They are categorized by their characteristics.

- kinematics : p_{T}, η

  The electrons in the signal events which decay from W tend to be more energetic than the ones from jets. So, p_{T} gives a good discriminating power.

- shower shape : $\sigma_{i\eta i\eta}$, $\phi_{i\eta i\eta}$, $\Delta\phi_{SC}$, $\Delta\eta_{SC}$, $E_{3\times3}/E_{5\times5}$(R9), $E_{1\times5}/E_{5\times5}$, $E_{\mathrm{PS}}/E_{\mathrm{SC}}$

  The $\sigma_{i\eta i\eta}(\phi_{i\eta i\eta})$ is the weighted sum of $\eta(\phi)$ difference between the seed crystal which contains the highest energy, and the 5×5 crystals surrounding it. The $\Delta\phi_{SC}(\Delta\eta_{SC})$ is the width of average distance between the seed crystal and the 5×5 crystals surrounding it. The $E_{3\times3}/E_{5\times5}$(R9) and $E_{1\times5}/E_{5\times5}$ are the ratio of energy deposit in the $3\times3(1\times5)$ crystals to the 5×5 crystals centered at the seed crystal. All these variables are related to the lateral shower shape. $E_{\mathrm{PS}}/E_{\mathrm{SC}}$ is the ratio of the energy deposit in the preshower detector and the energy deposit in the supercluster. This variable is related to the longitudinal shower shape.

  The hadron showers are longer and broader, and subject to larger fluctuation than the electromagnetic shower. But, electrons have a large shower width in ϕ direction due to bremsstrahlung. The amount of bremsstrahlung depends on the p_{T} of electron and the material that the electron goes through which depends on η. For example, electrons bremsstrahlung more at low p_{T} because it bends more in the magnetic field, and at high η because there are more materials that enhances the probability to radiate photons.

- track fit quality : χ^2(GSF)/ndof, χ^2(CTF)/ndof

  The χ^2(GSF)/ndof and χ^2(CTF)/ndof are the variables to measure the quality of GSF and CTF track fits, respectively.

  These variables reject poorly reconstructed tracks and conversion in the tracker.

- cluter-track matching (geometry) : $\Delta\phi_{in}$, $\Delta\eta_{in}$, $\Delta\eta_{out}$

  The $\Delta\phi_{in}(\Delta\eta_{in})$ is the distance in $\eta(\phi)$ direction between the supercluster position and the track trajectory extrapolated from the interaction point to the supercluster. $\Delta\eta_{out}$ is the distance in η direction between the supercluster position and the track trajectory extrapolated from the supercluster to the interaction point.

  This variables reject random combination of a track and an ECAL energy deposit.

- cluter-track matching (energy-momentum) : E_{SC}/p, E_C/p_{out}, $1/E_{\rm SC} - 1/p$

  The E_{SC}/p is the supercluster energy divided by the track momentum at the point of closest approach (PCA) to the beam spot. The E_C/p_{out} is the electron cluster energy divided by the track momentum at the PCA to the electron cluster, extrapolated from the outermost track state. The $1/E_{\rm SC} - 1/p$ is the difference between the reciprocals of the supercluster energy and the track momentum.

  These variables reject random combination of a track and an ECAL energy deposit.

- fraction of energy carried away by bremsstrahlung : f_{brem}

  The f_{brem} is the ratio of the difference between the momenta measured at the vertex and the outermost state to the momentum measured at the vertex. This variable shows the fraction of momentum loss by bremsstrahlung.

  When a $\pi^{\pm}$ is converted to π^0 in ECAL through charge exchange process, all energy of the initial energy of $\pi^{\pm}$ is stored in the ECAL. In this case, the

momentum of the track and the energy in the ECAL match perfectly giving $f_{brem} \sim 1$. But, most electrons undergo bremsstrahlung giving $f_{brem} < 1$. Therefore, this variable is used to reject charge exchange process.

- ratio of hadronic energy to EM energy : H/E

  The H/E is the ratio of the energy deposit in the HCAL tower behind the electromagnetic seed cluster to the energy of the seed cluster.

  Electrons make most of their energy deposit only in ECAL while hadrons make them in both ECAL and HCAL. Therefore, H/E is smaller for electrons than hadrons.

- impact parameter : transverse and 3D impact parameters with respect to the primary vertex

  These variables rejects electrons produced at a displaced vertex, particularly, decay from heavy flavor hadrons and photon conversion which also has dedicated selection and random combination of a track and an ECAL energy deposit as well as electrons from pileup vertices.

For selecting good electrons, we finally require that the BDT score be greater than the values depending on the kinematic region as shown in Table 4.5.

Table 4.5: Cut values on BDT score for electron identification. Electrons with BDT score greater than the corresponding values in the table are considered as good electrons.

	$0 < \|\eta\| < 0.8$	$0.8 < \|\eta\| < 1.479$	$1.479 < \|\eta\| < 2.5$
10 GeV $< p_{\mathrm{T}} <$ 20 GeV	0.0	0.1	0.62
$p_{\mathrm{T}} >$ 20 GeV	0.94	0.85	0.92

The prompt leptons are produced in a quite environment, *i.e.*, no energetic particles around it, while non-prompt leptons are accompanied by a number of

energetic particles that come from hadronic shower. Therefore, by requiring little activities around a lepton candidate, we can significantly reduce the contribution of non-prompt leptons.

For electrons, the isolation requirements are imposed by computing an isolation variable using PF candidates. In the high luminosity environment there are random contribution from pileup to the isolation calculation, so we need to correct for this to prevent a degradation of the isolation requirement. To reduce the contributions from the random charged PF candidates, they are required to be close to the event primary vertex. The contribution from the random neutral PF candidates is corrected by subtracting the expected contribution. The variable is defined as a scalar sum of the p_{T} of the PF candidates satisfying the following requirements.

- $\Delta R < 0.4$ to the electron candidate in the $\eta \times \phi$ plane
- Other PF electrons and PF muons in the isolation cone are vetoed
- Gamma PF candidates are required to be in the region $\Delta R > 0.08$ from the electron candidate
- Charged hadron PF candidates are required to be in region $\Delta R < 0.015$ from the electron candidate
- Charged hadron PF candidates are required to be associated with the event primary vertex : their closest vertex should be the event primary vertex
- Neutral components are corrected by subtracting pileup contribution which is calculated by $\rho \times \mathrm{A_{eff}}$,

where ρ is the event-by-event energy density [**?**] calculated using `kt6PFJets` algorithm [**?**], and $\mathrm{A_{eff}}$ is the effective area as shown in Table 4.6. The correction is applied to only neutral particles because charged particles are required to be from the event primary vertex.

The isolation variable we cut on is calculated as

$$\frac{\mathrm{Iso_{PF}}}{p_{\mathrm{T}}} = [\mathrm{Iso_{charged\,hadron}} + \{\mathrm{Iso_{gamma}} + \mathrm{Iso_{neutral\,hadron}} - \rho \times \mathrm{A_{eff}}\}] \times \frac{1}{p_{\mathrm{T}}} \qquad (4.1)$$

Table 4.6: Effective area used for electron isolation calculation.

$\|\eta\|$	0 - 1.0	1.0 - 1.479	1.479 - 2.0	2.0 - 2.2	2.2 - 2.3	2.3 - 2.4
A_{eff}	0.19	0.25	0.12	0.21	0.27	0.44

where $\mathrm{Iso}_{\mathrm{charged\,hadron}}$, $\mathrm{Iso}_{\mathrm{gamma}}$, and $\mathrm{Iso}_{\mathrm{neutral\,hadron}}$ are the scalar sum of the p_{T} of the charged hadron, gamma and neutral hadron PF candidates, respectively, in the isolation cone of 0.4 around the electron. We require $\frac{\mathrm{Iso_{PF}}}{p_{\mathrm{T}}}$ to be less than 0.15 in both barrel and endcap.

In order to reject the electrons from a conversion from a photon, we reject the electrons if there is a reconstructed conversion vertex where one of the two tracks match with the electron, the probability of the conversion vertex fit is greater than 10^{-6}, and the distance between the conversion vertex and the point of closest approach to the event primary vertex is greater than 2 cm. The electron candidate is also rejected if there is any missing hit in the electron track between the conversion vertex and the event primary vertex.

The impact parameter requirements are such that the transverse(longitudinal) impact parameter between the electron track and the event primary vertex is less than 0.02 cm(0.1 cm). These requirements reject electrons produced at a displaced vertex, particularly, decay from heavy flavor hadrons and photon conversion, and random combination of a track and an ECAL energy deposit as well as electrons from pileup vertices.

The efficiency of the full electron selection measured with 20/10 GeV requirement in MC is about 80 % for electrons in the m_{H}=125 GeV events and about 5 % for electrons whose mothers are not Ws in W+jets. The efficiency is measured with respect to the trigger selection that is discussed at the beginning of this section.

4.4 Muon

A muon candidate is reconstructed if there is a track in the tracker and/or are segments in the muon system. There are multiple sources of muons which do not originate from gauge boson decays.

- decay-in-flight : a charged hadron decays to muon in the tracker,
- punch-through : a charged hadron survive the HCAL, and leaves a track at the early stage of the muon system, and
- Heavy flavor decay : B or D hadron decays semi-leptonically.

In order to reject muons from these sources, we apply a muon selection which is composed of requirements on the identification, isolation and impact parameter.

The identification requirements are as follows.

- The muon should be identified as PF muon
- $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.4$
- The number of tracker layers where the muon track made hits must be greater than 5. This requirement is to guarantee a good p_{T} measurement, but it also rejects decay-in-flights.
- The number of pixel hits of the muon track must be greater than 0. This requirement is to reject decay-in-flights.
- The relative resolution of the muon p_{T} must be less than 0.1. This requirement is to guarantee a good p_{T} measurement, but it also rejects decay-in-flights.
- The χ^2/ndof of the kink finder algorithm which finds muons from decay-in-flights must be less than 20. This requirement is to reject decay-in-flights.
- The muon should be global muon or tracker muons

If muons are global muon, χ^2/ndof of the global fit must be less than 10 to reject hadronic punch through and decay-in-flights, there must be at least one muon hit matching the global fit to reject hadronic punch through and decay-in-flights, and there must be at least two muon segments in different muon stations to reject hadronic punch through and accidental track-to-segment matches. If the muon is not a global muon, it can be a tracker muon satisfying that at least two muon segments one of which is in the outermost muon station are matched to reject bad momentum measurements and hadronic punch through.

The prompt leptons are produced in a quite environment, *i.e.*, no energetic particles around it, while non-prompt leptons are accompanied by a number of energetic particles that come from hadronic shower. Therefore, by requiring little activities around a lepton candidate, we can significantly reduce the contribution of non-prompt leptons.

For muons, the BDT-based isolation variable [**?**] is used. It uses the energy deposits of PF candidates of three categories, charged hadron, gamma and neutral hadrons in the concentric isolation cones of size $\Delta R = 0 - 0.1,\ 0.1 - 0.2,\ 0.2 - 0.3,\ 0.3 - 0.4$ and $0.4 - 0.5$. The basic idea of dividing a conventional cone into rings is that the shape of the p_{T} sum of particles in a given ring is different for prompt and non-prompt muons, and we can obtain better separation by using that information. Neutral components are corrected by subtracting the pileup contribution which is calculated by $\rho \times \mathrm{A_{eff}}$ where ρ (`kt6PFJets`) is the event-by-event energy density and $\mathrm{A_{eff}}$ is the effective area. Effective areas are from Fall 11 simulation (`kMuEAFall11MC`), and values are shown in Table 4.7. Exact definition of input variables to the BDT is the following :

- PF charged hadron
  - minimum of $\mathrm{Iso}_{\mathrm{PF\ charged,\ 01}}/p_{\mathrm{T}}$ and 2.5
  - minimum of $\mathrm{Iso}_{\mathrm{PF\ charged,\ 12}}/p_{\mathrm{T}}$ and 2.5
  - minimum of $\mathrm{Iso}_{\mathrm{PF\ charged,\ 23}}/p_{\mathrm{T}}$ and 2.5
  - minimum of $\mathrm{Iso}_{\mathrm{PF\ charged,\ 34}}/p_{\mathrm{T}}$ and 2.5
  - minimum of $\mathrm{Iso}_{\mathrm{PF\ charged,\ 45}}/p_{\mathrm{T}}$ and 2.5

- PF gamma : If negative, 0.0 is assigned
  - minimum of $[\mathrm{Iso}_{\mathrm{PF\ gamma,\ 01}} - \rho \times \mathrm{A_{eff}}] / p_{\mathrm{T}}$ and 2.5
  - minimum of $[\mathrm{Iso}_{\mathrm{PF\ gamma,\ 12}} - \rho \times \mathrm{A_{eff}}] / p_{\mathrm{T}}$ and 2.5
  - minimum of $[\mathrm{Iso}_{\mathrm{PF\ gamma,\ 23}} - \rho \times \mathrm{A_{eff}}] / p_{\mathrm{T}}$ and 2.5
  - minimum of $[\mathrm{Iso}_{\mathrm{PF\ gamma,\ 34}} - \rho \times \mathrm{A_{eff}}] / p_{\mathrm{T}}$ and 2.5
  - minimum of $[\mathrm{Iso}_{\mathrm{PF\ gamma,\ 45}} - \rho \times \mathrm{A_{eff}}] / p_{\mathrm{T}}$ and 2.5
- PF neutral hadron : If negative, 0.0 is assigned
  - minimum of $[\mathrm{Iso}_{\mathrm{PF\ neutral,\ 01}} - \rho \times \mathrm{A_{eff}}] / p_{\mathrm{T}}$ and 2.5
  - minimum of $[\mathrm{Iso}_{\mathrm{PF\ neutral,\ 12}} - \rho \times \mathrm{A_{eff}}] / p_{\mathrm{T}}$ and 2.5
  - minimum of $[\mathrm{Iso}_{\mathrm{PF\ neutral,\ 23}} - \rho \times \mathrm{A_{eff}}] / p_{\mathrm{T}}$ and 2.5
  - minimum of $[\mathrm{Iso}_{\mathrm{PF\ neutral,\ 34}} - \rho \times \mathrm{A_{eff}}] / p_{\mathrm{T}}$ and 2.5
  - minimum of $[\mathrm{Iso}_{\mathrm{PF\ neutral,\ 45}} - \rho \times \mathrm{A_{eff}}] / p_{\mathrm{T}}$ and 2.5

where we define

$$\mathrm{Iso}_{\mathrm{PF,XY}} = \sum_{\substack{\text{PF candidates in the cone of} \\ \text{0.X}<\Delta R^{\mu-\text{PF candidate}}<\text{0.Y}}} p_{\mathrm{T}} \tag{4.2}$$

We require that the BDT score be greater than 0.82 (0.86) in $10 < p_{\mathrm{T}} < 20$ GeV and 0.86 (0.82) in $p_{\mathrm{T}} > 20$ GeV. The cut values correspond to the ones in barrel (endcap).

In addition, we require transverse/longitudinal impact parameters to be associated with the event primary vertex. The transverse impact parameter is required to be less than 0.01 cm for $p_{\mathrm{T}} < 20$ GeV and 0.02 cm for $p_{\mathrm{T}} > 20$ GeV. The longitudinal impact parameter is required to be less than 0.1 cm. These requirements reject decay-in-flight, heavy flavor decay and muons from pileup vertices.

The efficiency of the full muon selection measured with 20/10 GeV requirement in MC is about 90 % for muons in the m_{H}=125 GeV events and about 11 % for muons whose mothers are not Ws in W+jets.

Table 4.7: Effective areas used for muon isolation. They were calculated with Fall11 MC sample.

PF gamma						
$\lvert\eta\rvert$	$0.0-1.0$	$1.0-1.479$	$1.479-2.0$	$2.0-2.2$	$2.2-2.3$	$2.3-2.4$
$0.0 < dR < 0.1$	0.004	0.002	0.003	0.009	0.003	0.011
$0.1 < dR < 0.2$	0.012	0.008	0.006	0.012	0.019	0.024
$0.2 < dR < 0.3$	0.026	0.020	0.012	0.022	0.027	0.034
$0.3 < dR < 0.4$	0.042	0.033	0.022	0.036	0.059	0.068
$0.4 < dR < 0.5$	0.060	0.043	0.036	0.055	0.092	0.115
PF neutral hadron						
$\lvert\eta\rvert$	$0.0-1.0$	$1.0-1.479$	$1.479-2.0$	$2.0-2.2$	$2.2-2.3$	$2.3-2.4$
$0.0 < dR < 0.1$	0.002	0.004	0.004	0.004	0.010	0.014
$0.1 < dR < 0.2$	0.005	0.007	0.009	0.009	0.015	0.017
$0.2 < dR < 0.3$	0.009	0.015	0.016	0.018	0.022	0.026
$0.3 < dR < 0.4$	0.013	0.021	0.026	0.032	0.037	0.042
$0.4 < dR < 0.5$	0.017	0.026	0.035	0.046	0.063	0.135

4.5 Jet

In the high pileup environment, the issue of jet selection is to reject random jets from pileup vertices. This can be done by requiring tracks in the jets to be associated with the event primary vertex. But, since this requirement uses tracks, it can not be applied outside of the tracker volume($|\eta| > 2.5$). Outside of the tracker, we should rely on the information in the calorimeters. One characteristic of jets from pileup vertices is that they are softer than the jets from hard interaction, and thus need to be overlaid to pass the p_{T} threshold which is 30 GeV in this analysis. Since they are overlaid, the shower shape is broader than that of jets from hard interaction.

In this analysis, we apply a BDT-based technique to suppress jets originated from pileup [63] using variables related to above-mentioned characteristics. The following variables are used in the BDT-based suppression technique

- Number of reconstructed primary vertices in the event

- Kinematics of the jet : p_{T}, η and ϕ
- Longitudinal and transverse impact parameters of the most energetic charged PF candidates in the jet : this requirement reduces pileup jets in the tracker volume
- Fraction of charged PF candidates associated with the event primary vertex(PV)
  - sum p_{T} of PF candidates with $d_Z(PV) < 0.2$ divided by sum p_{T} of all PF candidates : value of this variable is close to 1 for jets from hard interaction and 0 for jets from pileup
  - sum p_{T} of PF candidates with d_Z(all vertices not PV) < 0.2 divided by sum p_{T} of all PF candidates : value of this variable is close to 0 for jets from hard interaction and 1 for jets from pileup
- Number of neutral and charged PF candidates in the jet : because jets from pileup tend to be overlaid by multiple soft jets, they have higher multiplicity
- p_{T}-weighted mean of ΔR of all PF candidates within $\Delta R < 0.5$ in the jet : this variable is a measure of shower shape in the cone. The value of this variable is smaller for jets from hard interaction than those from pileup because the distribution of particles in the cone is more widespread.
- Fraction of sum p_{T} of all PF candidate in the rings of ΔR = 0.0 - 0.1, 0.1 - 0.2, 0.2 - 0.3, 0.3 - 0.4 and 0.4 - 0.5 : The particles in the jets from hard interaction are centered around the jet axis, thus most of its energy is concentrated around the jet axis. However, the particles in the jets from pileup are more spread because multiple jets are overlaid, thus its energy is less centralized than the hard interaction jets. These variables make use of this difference. For example, the fraction of energy(sum p_{T}) in the ring of ΔR = 0.0 - 0.1 is higher for jet from hard interaction.

We select jets of which BDT scores are greater than the values in Table 4.8. On top of the BDT requirement, jets are required not to overlap selected electrons and

Table 4.8: Cut values on jet identification BDT scores. The BDT score is required to be greater than these values to be counted as a jet.

p_{T}(GeV)	$0 < \|\eta\| < 2.5$	$2.5 < \|\eta\| < 2.75$	$2.75 < \|\eta\| < 3.0$	$3.0 < \|\eta\| < 4.7$
- 10	0.0	0.0	0.0	0.2
10 - 20	-0.4	-0.4	-0.4	0.4
20 - 30	0.0	0.0	0.2	0.6
30 -	0.0	0.0	0.6	0.2

muons in order to avoid cases where leptons are reconstructed as jets. If the ΔR between a lepton and the jet is less than 0.3, the jet is likely to be a lepton, thus it is not considered as a jet.

Of the selected jets, the high p_{T} jets($p_{\mathrm{T}} > 30$ GeV and $|\eta| < 4.7$) are used for counting the number of jets, and the low p_{T} jets($10 < p_{\mathrm{T}} < 30$ GeV] and $|\eta| < 4.7$) are used for rejecting top events.

4.6 Missing Transverse Energy

The MET is used to reject processes which do not have true source of missing energy. Target processes are QCD and $Z/\gamma^* \to \ell\ell$ where MET is caused by mis-measurement of jet momentum and pileup. However, in case of $Z/\gamma^* \to \tau\tau$ the τ can decay into lepton and neutrinos, giving true missing momentum. Because of large mass difference between Z and τ, τ is generally boosted significantly with its decay products flying in the same direction of the τ momentum. Therefore, the missing energy component perpendicular to the lepton momentum is a better measure for true MET which originate from W decays. To realize this we define "projected MET" (proj − MET) as

$$\mathrm{proj-MET} = \begin{cases} \mathrm{MET} & \text{if } \Delta\phi_{min} > \frac{\pi}{2}, \\ \mathrm{MET}\sin(\Delta\phi_{min}) & \text{if } \Delta\phi_{min} < \frac{\pi}{2} \end{cases} \tag{4.3}$$

where

$$\Delta\phi_{min} = min(\Delta\phi(\ell_1, \mathrm{MET}), \Delta\phi(\ell_2, \mathrm{MET})). \tag{4.4}$$

The $\Delta\phi(\ell_i, \text{MET})$ is the angle between MET and the lepton i on the transverse plane. Figure 4.2 shows the schematic of proj $-$ MET. Figure 4.3 shows the MET

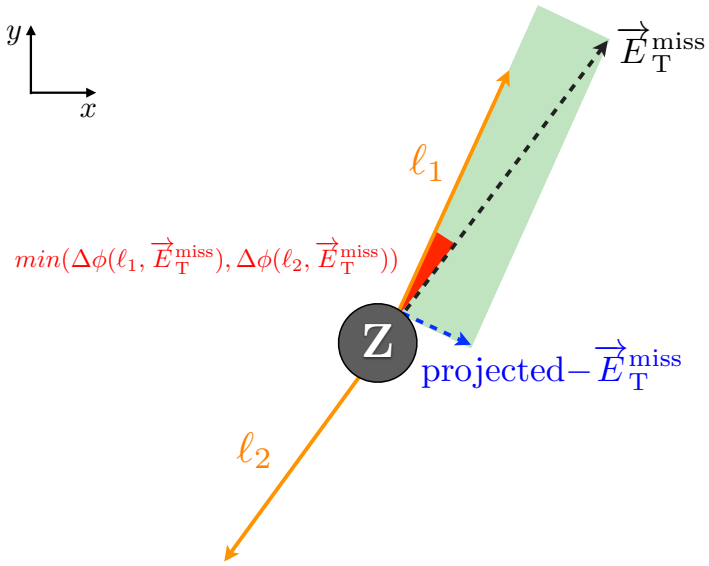


Figure 4.2: Schematic of proj $-$ MET.

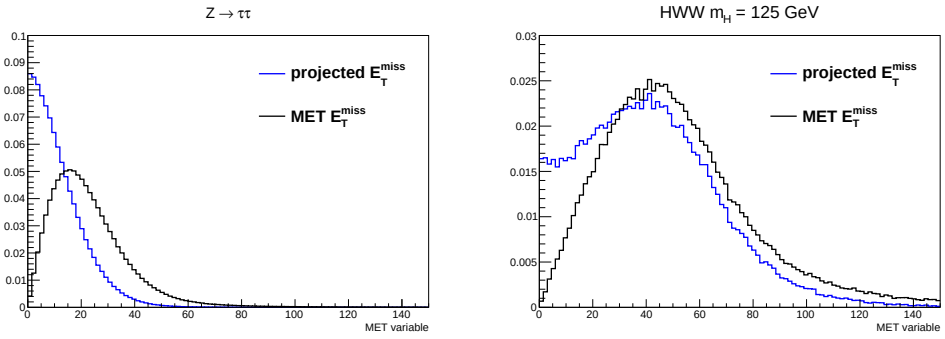


Figure 4.3: MET(black) and proj $-$ MET(blue) in $Z/\gamma^* \to \tau\tau$(left) and SM Higgs at m_{H}=125 GeV(right).

and the proj $-$ MET distributions in black and red, respectively, in $Z/\gamma^* \to \tau\tau$ and SM Higgs at m_{H}=125 GeV on left and right, respectively. We can see that proj $-$ METis more squeezed to the lower values in case of $Z/\gamma^* \to \tau\tau$, giving better rejection power.

The trkMET is a variable insensitive to the pileup, but a drawback is that its tail is longer than PFMET in the $Z/\gamma^* \to \ell\ell$ events. By isospin symmetry the

average number of neutral hadrons should be same as that of the charged hadrons. But, if there is an imbalance between neutral and charged components in the jets, the trkMET can be calculated arbitrarily high. For example, for a di-jet event with back-to-back 30 GeV jets where one jet is composed of 10 GeV of charged components and 20 GeV of neutral components, and the other jet has the opposite composition, the PFMET is 0 while trkMET is 10 GeV. Therefore, we use the minimum of PFMET and trkMET

$$\text{min} - \text{proj} - \text{MET} = min\left(\text{proj} - \text{MET}, \text{proj} - \text{trkMET}\right), \tag{4.5}$$

as a MET variable to protect trkMET from those fluctuations. The other advantage of taking the minimum is that the the correlation between the two MET definitions is strong in the events with true MET, while weak in the events with fake MET as shown in Fig. 4.4. So, by taking the minimum of them, we can get additional rejection of $Z/\gamma^* \to \ell\ell$ backgrounds without a harm to the signal.

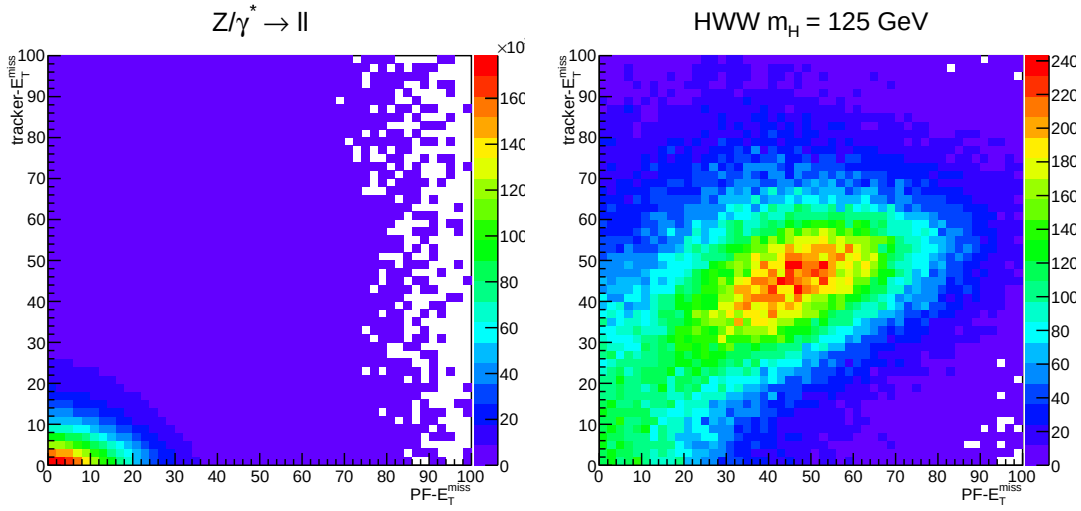


Figure 4.4: proj − MET vs. proj − trkMET in $Z/\gamma^* \to \ell\ell$ and SM Higgs m_{H}= 125 GeV.

We require min − proj − MET to be greater than 20 GeV as a baseline selection. For further rejection of $Z/\gamma^* \to \ell\ell$ background, we apply BDT-based technique which will be discussed in detail in section 4.8.

4.7 Top-tagging

The $t\bar{t}/tW$ events contain additional b quarks on top of the two leptons and MET. The b quarks are then hadronized to B mesons, and develop hadronic showers. Experimentally, they can be identified by finding a displaced secondary vertex, B-tagging, or a soft muon decayed from the B meson.

The section 3.7 described how the discriminator for b-tagging is constructed. If the b-tagging discriminator(TCHE) of a jet that passes the jet selection criteria described in section 4.5 is greater than 2.1, this jet is selected as a b-tagged jet. The event is rejected if there is at least one b-tagged jet.

We enhance the rejection of $t\bar{t}/tW$ backgrounds by removing events that contain non-isolated soft muons from heavy flavor decays. The following requirements are imposed to select soft muons.

- $p_{\mathrm{T}} > 3$ GeV
- The muon should be a tracker muon,
- The muon should have at least two muon segments one of which is in the outermost muon stations are matched,
- $N_{\text{track layers}} > 5$,
- $|d_0| < 0.2$ cm,
- $|d_z| < 0.2$ cm, and
- $\mathrm{Iso}_{\mathrm{Total}}/p_{\mathrm{T}} > 0.1$ if $p_{\mathrm{T}} > 20$ GeV

where $N_{\text{track layers}}$ is the number of tracker layers where the muon track has hits, $d_0(d_z)$ are the transverse(longitudinal) impact parameter with respect to the event primary vertex, and $\mathrm{Iso}_{\mathrm{Total}}$ is the sum of E_{T} inside of the cone with $dR < 0.3$ in ECAL and HCAL, and p_{T} of tracker. Adding soft muon requirement especially helps in the 0-jet category where the top rejection increases about 50 % compared to using b-tagging only.

The total top rejection using both methods when applied to $t\bar{t}/tW$ MC is about 50 % in the 0-jet category and 80 % in the 1-jet category.

4.8 $Z/\gamma^* \to \ell\ell$ suppression in $ee/\mu\mu$ final states

In the $ee/\mu\mu$ channel, $Z/\gamma^* \to \ell\ell$ is one of the dominant backgrounds, and it is suppressed by rejecting events that have di-lepton mass around the Z mass. So, we veto events with $|m_{\ell\ell} - m_{\mathrm{Z}}| < 15$ GeV. However, since the production rate of $Z/\gamma^* \to \ell\ell$ process is very large, even after rejecting those events, there remains a significant contribution. To further reject this background, a tight MET requirement is typically imposed because the MET in $Z/\gamma^* \to \ell\ell$ tends to be smaller than the MET in the signal process. This is because there is no genuine source of MET in the $Z/\gamma^* \to \ell\ell$ events. In addition to MET there are some difference in event kinematics. For example, in $Z/\gamma^* \to \ell\ell$, the di-lepton system and the leading jet tend to be back-to-back while in the signal events they do not have such correlation because there is additional degree of freedom from neutrinos.

Usually, $Z/\gamma^* \to \ell\ell$ background is suppressed by applying a strong MET selection. But, this can lead to a significant loss in signal. So, we developed a BDT-based $Z/\gamma^* \to \ell\ell$ suppression technique [64] to recover the loss in the signal efficiency. The training is done with a combination of signal samples, m_{H}= 125 and 200 GeV, for signal and $Z/\gamma^* \to \ell\ell$ MC for background. The motivation of using a combination of the two m_{H} points is to use one training for all m_{H} points. The $e\mu/\mu e$ events are used in the training because there should not be any difference in the training variables between $ee/\mu\mu$ and $e\mu/\mu e$ events. As a background sample, a combination of the Madgraph [65] and the Powheg [66] samples is used to maximize statistics. For the training, we apply the WW selection which is defined in section 4.10) for the cut-based analysis which is discussed in sec 4.10. The training is done separately in the 0-jet and 1-jet categories.

The input variables to the BDT are the following.

- MET-related variables
  - proj – MET
  - proj – trkMET
  - MET significance(MET/$\sqrt{\sum_i^{\text{All objects}} E_T(i)}$ where $E_T(i)$ is the transverse energy of the object i used for MET calculation) : this variable is

a measure of goodness of MET measurement, *i.e.*, being small means good measurement. This variable is small for signal where real MET is present and large for $Z/\gamma^* \to \ell\ell$ where MET comes from primarily mis-measurement of jet energy.

- kinematic variables
  - Di-lepton $p_{\mathrm{T}}(p_{\mathrm{T}}^{\ell\ell})$: this variable tends to larger for $Z/\gamma^* \to \ell\ell$ than signal because p_{T} of the trailing lepton is softer in signal because it comes from virtual W with much smaller mass.
  - Higgs transverse mass($m_{\mathrm{T}} = \sqrt{2p_{\mathrm{T}}^{\ell\ell}\mathrm{MET}(1 - cos(\Delta\phi_{\ell\ell-\mathrm{MET}}))}$) : signal tends to have larger value because it has larger MET and the azimuthal angle difference between di-lepton system and MET is smaller due to spin-0ness of Higgs boson and the V-A nature of the W decay which is discussed in subsection 1.4.2.
  - leading jet p_{T}: this variable is larger for $Z/\gamma^* \to \ell\ell$ because di-lepton system and the leading jet are more likely to be back-to-back and the cut on $p_{\mathrm{T}}^{\ell\ell}(> 45$ GeV) selects events with higher jet p_{T}.
  - recoil of the di-lepton + MET system(the magnitude of the vector sum of PFMET and the di-lepton system) : this variable is smaller for signal because the azimuthal angle difference between di-lepton system and MET is smaller.
- azimuthal angle differences
  - Azimuthal angle difference between di-lepton system and leading jet($p_{\mathrm{T}} >$ 15 GeV) : this variable is close to π for $Z/\gamma^* \to \ell\ell$ because leading jet recoils off the Z boson which decays into two leptons. For signal it is less back-to-back because of additional degrees of freedom from neutrinos.
  - Azimuthal angle difference between leading jet ($p_{\mathrm{T}} > 15$ GeV) and MET : this variable is close to 0 for $Z/\gamma^* \to \ell\ell$ because MET mostly comes from mis-measurement of jet energy. For signal it is more broad because

MET and jet do not have topological correlation due to additional degrees of freedom from leptons.

- Azimuthal angle difference between di-lepton system and MET : This variable is close to π for both signal and $Z/\gamma^* \to \ell\ell$ because for signal the azimuthal angle difference between di-lepton system and MET is large and for $Z/\gamma^* \to \ell\ell$ MET is aligned with jet p_{T} and it is back-to-back with $p_{\mathrm{T}}^{\ell\ell}$. But, it has different correlations for signal and $Z/\gamma^* \to \ell\ell$ with other input variables such as azimuthal angle difference between leading jet and MET.

- other variable

  - Number of reconstructed vertices(N_{vtx}) : This variable has different correlations for signal and $Z/\gamma^* \to \ell\ell$ with other input variables. For example, for $Z/\gamma^* \to \ell\ell$ MET has a large correlation with N_{vtx} because the more interaction are present, the worse jet energy resolution becomes, thus MET becomes large. But, for signal, MET is less affected by pileup because the magnitude of additional contribution from pileup events is very small compared to the real MET caused by neutrinos.

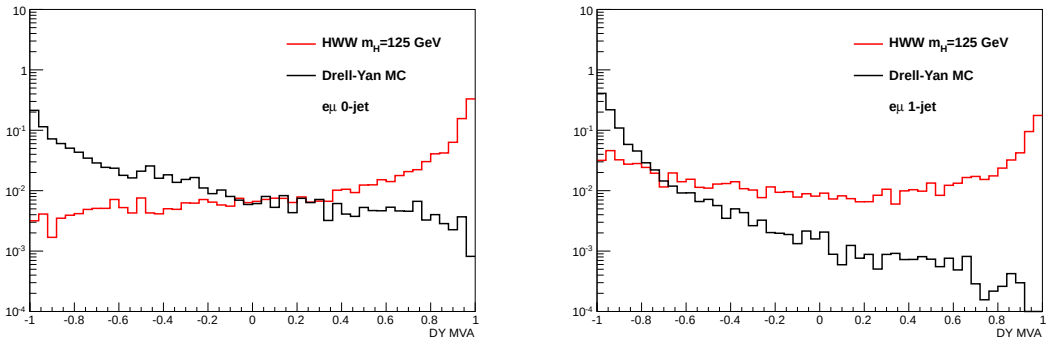


Figure 4.5: The BDT score for m_{H}= 125 GeV in red and $Z/\gamma^* \to \ell\ell$ MC in black in 0-jet(left) and 1-jet(right) categories.

Figure 4.5 shows the BDT scores for m_{H}= 125 GeVin red and $Z/\gamma^* \to \ell\ell$ MC in black in 0-jet(left) and 1-jet(right) categories. We finally require that the BDT

score be greater than 0.88(0.84) in 0-jet(1-jet) category. This requirement is applied to only $ee/\mu\mu$ category. The signal(m_{H}=125 GeV) efficiency of this selection is 50 % and 35 % in 0-jet and 1-jet categories, respectively, and the background($Z/\gamma^* \to \ell\ell$ MC) efficiency is 0.5 % and 0.1 % in 0-jet and 1-jet categories

4.9 Additional Selections

To reject events with more than two leptons such as WZ/ZZ where both bosons decay leptonically, we veto events if they contain a third lepton with $p_{\mathrm{T}} >$ 10 GeV. This requirement rejects about 30 % of events containing more then two leptons in WZ/ZZ.

At low $m_{\ell\ell}$ region, there are multiple resonances such as $\Upsilon(m_{\Upsilon} \sim 10$ GeV) and $J/\psi(m_{J/\psi} \sim 3$ GeV). In order to reject these resonances, we apply $m_{\ell\ell} >$ 12 GeV .

W + jets background tends to have small $p_{\mathrm{T}}^{\ell\ell}$ because the lepton from W and the recoiling jet are likely to be back-to-back. So, in order to suppress W + jets further, $p_{\mathrm{T}}^{\ell\ell} > 30(45)$ GeV is applied for shape-based (cut-based) approaches.

4.10 WW Selection

All requirements imposed so far are designed to select events containing a pair of W. This baseline selection is called "WW selection" and the stage of selection where the WW selection is applied is called "WW level". By applying WW selection, we can reach better signal-to-background ratio(S/B), therefore a reasonable extraction of signal becomes possible. Table 4.9 summarizes the requirements of the WW selection.

Table 4.9: Summary of WW selection. [*] For shape-based method $p_{\mathrm{T}}^{\ell\ell} > 30$ GeV is applied and for cut-based method $p_{\mathrm{T}}^{\ell\ell} > 45$ GeV is applied.

Selection	$e\mu/\mu e$	$ee/\mu\mu$
$p_{\mathrm{T}}^{\ell,\mathrm{max}}$	> 20 GeV	> 20 GeV
$p_{\mathrm{T}}^{\ell,\mathrm{min}}$	> 10 GeV	> 10 GeV
Lepton selection	applied	applied
Number of jet selection	applied	applied
Third lepton veto	applied	applied
opposite-sign requirement	applied	applied
$m_{\ell\ell}$	> 12 GeV	> 12 GeV
$\lvert m_{\ell\ell} - m_Z \rvert > 15$ GeV	not applied	applied
$\mathrm{min} - \mathrm{proj} - \mathrm{MET}$	> 20 GeV	> 20 GeV
BDT-based Drell-Yan suppression	not applied	applied
Top veto	applied	applied
$p_{\mathrm{T}}^{\ell\ell}$	> 30 GeV[*]	> 45 GeV

Chapter 5

Analysis Strategy

Chapter 4 described the event selection to suppress reducial backgrounds. This chapter discusses methods to separate signal component from irreducible backgrounds.

After selecting events containing two Ws, we extract the signal component from them. The events can be categorized based on the lepton flavor($e\mu/\mu e$ or $ee/\mu\mu$) and the number of jets(0 or 1). The advantage of dividing events into different categories is that the background composition is different in the different categories. Figure 5.1 shows the background composition after the Higgs selection which will be discussed in the following section. The total background is normalized to 1, so the plot shows only relative contributions from each background. To have a sense of the size of the signal, m_{H}=125 GeV contribution(red) is overlaid. In the 0-jet category, WW is the dominant background in both $e\mu/\mu e$ and $ee/\mu\mu$ categories, and W + jets and $Z/\gamma^* \to \ell\ell$ follow in the $e\mu/\mu e$ and $ee/\mu\mu$ categories, repectively. In the 1-jet category, WW and $t\bar{t}/tW$ are almost equally dominant in both $e\mu/\mu e$ and $ee/\mu\mu$ categories, and W + jets and $Z/\gamma^* \to \ell\ell$ follow in the $e\mu/\mu e$ and $ee/\mu\mu$ categories, repectively.

This analysis is dedicated to search for $gg \to H$ process, so the events with 2 or more jets are not considered because this region of phase space is used to study the $qq \to qqH$ process. We thus have two categories in the number of jets(0-jet or 1-jet) and two categories in the lepton flavors($e\mu/\mu e$ or $ee/\mu\mu$). Therefore, we have four categories and the dedicated analysis is performed in each category. This allows better constraints on the backgrounds and thus results in better sensitivity to the discovery.

The extraction of signal component is done by two methods. The first method which is simpler but less sensitive to discovery is the "cut-based" method. This method is a simple cut-and-count approach that uses a subset of the selected WW events with enhanced S/B at a given m_{H}. The selections are optimized for a given m_{H}, accounting for the kinematic difference between different m_{H} hypotheses which was discussed in section 1.4.3, This method is applied to all lepton flavor final states. The other method which is more complicated and sophiscated but more sensitive to discovery is the "shape-based" method which uses binned

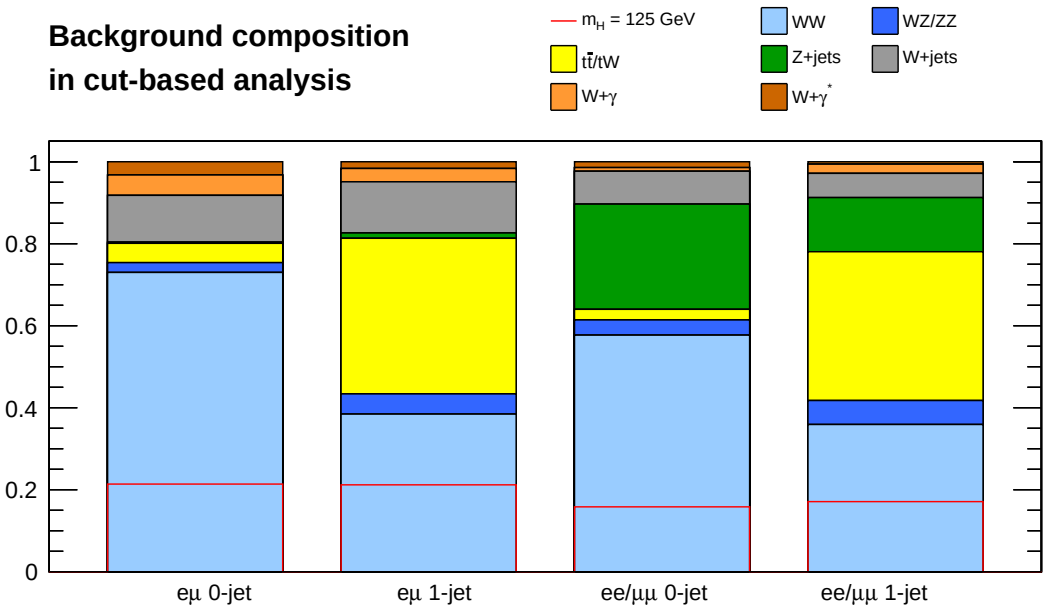


Figure 5.1: Background composition in cut-based analysis for the m_{H}=125 GeV hypothesis in the four categories. The Higgs selection in addition to WW selection is applied. The signal contribution is overlaid.

2-dimensional templates of m_{T} and $m_{\ell\ell}$. A binned maximum likelihood fit to the 2-dimensional templates is performed to extract the signal component. The shape-based method uses the difference in shape between signal and background, *i.e.*, more information than just normalization, so can provide better separation between them. In addition, because the fit can constrain backgrounds further than the estimated uncertainties, the performance of this method is better than the cut-based method. In this method most of the templates are taken from simulation, so it is important to ensure that the shapes taken from simulation are consistent with what is observed in data. In the $ee/\mu\mu$ channel, the $Z/\gamma^* \to \ell\ell$ is one of the leading backgrounds, and the shape of the $Z/\gamma^* \to \ell\ell$ taken from simulation is not reliable because of the poor modeling of MET and the poor statistics in the sample. So, we apply this method to only $e\mu/\mu e$ category.

5.1 Cut-based Method

The "cut-based" method is a simple and robust cut-and-count analysis approach. In this method, for a given m_{H} the expected signal and background yields are calculated in the region of phase space where S/B is enhanced. The yields and the uncertainties of each process are used to determine the signal component by a maximum likelihood fit. Because the estimated uncertainties are used, the precise estimation of the uncertainties is very important. This method is simpler than the other method because it uses only yields(no worries about the shapes), and most of the backgrounds are estimated by data-driven methods. The issue is the accuracy(normalization) and the precision(uncertainty) of the measurements.

As described in section 1.4.3, the kinematics of the signal events vary depending on m_{H}. Therefore, in addition to the WW selection, a dedicated selection for each m_{H} point is applied to enhance S/B at the given m_{H}. We use the following five variables to improve S/B,

- lepton transverse momentum : $p_{\mathrm{T}}^{\ell,\mathrm{max}}$ and $p_{\mathrm{T}}^{\ell,\mathrm{min}}$
- di-lepton mass : $m_{\ell\ell}$
- the azimuthal angle difference between the leptons : $\Delta\phi_{\ell\ell}$
- Higgs transverse mass :

$$m_{\mathrm{T}} = \sqrt{2p_{\mathrm{T}}^{\ell\ell}\mathrm{MET}(1 - cos(\Delta\phi_{\ell\ell-\mathrm{MET}}))} \tag{5.1}$$

  where $p_{\mathrm{T}}^{\ell\ell}$ is the transverse momentum of the di-lepton system, MET is the missing transverse momentum and $\Delta\phi_{\ell\ell-\mathrm{MET}}$ is the angle between di-lepton direction and MET in the transverse plane.

Figure 5.2-5.4 show the signal and the background distributions of the five variables at WW-level for three m_{H} hypotheses at 125, 160 and 200 GeV. The signal is overlaid on the background, and the cut values are indicated on the plots. The cut values for all m_{H} points studied by this analysis is shown in Table 5.1. Applying these cuts, S/B is improved from 2 - 4 % to 15 - 20 % depending on the categories.

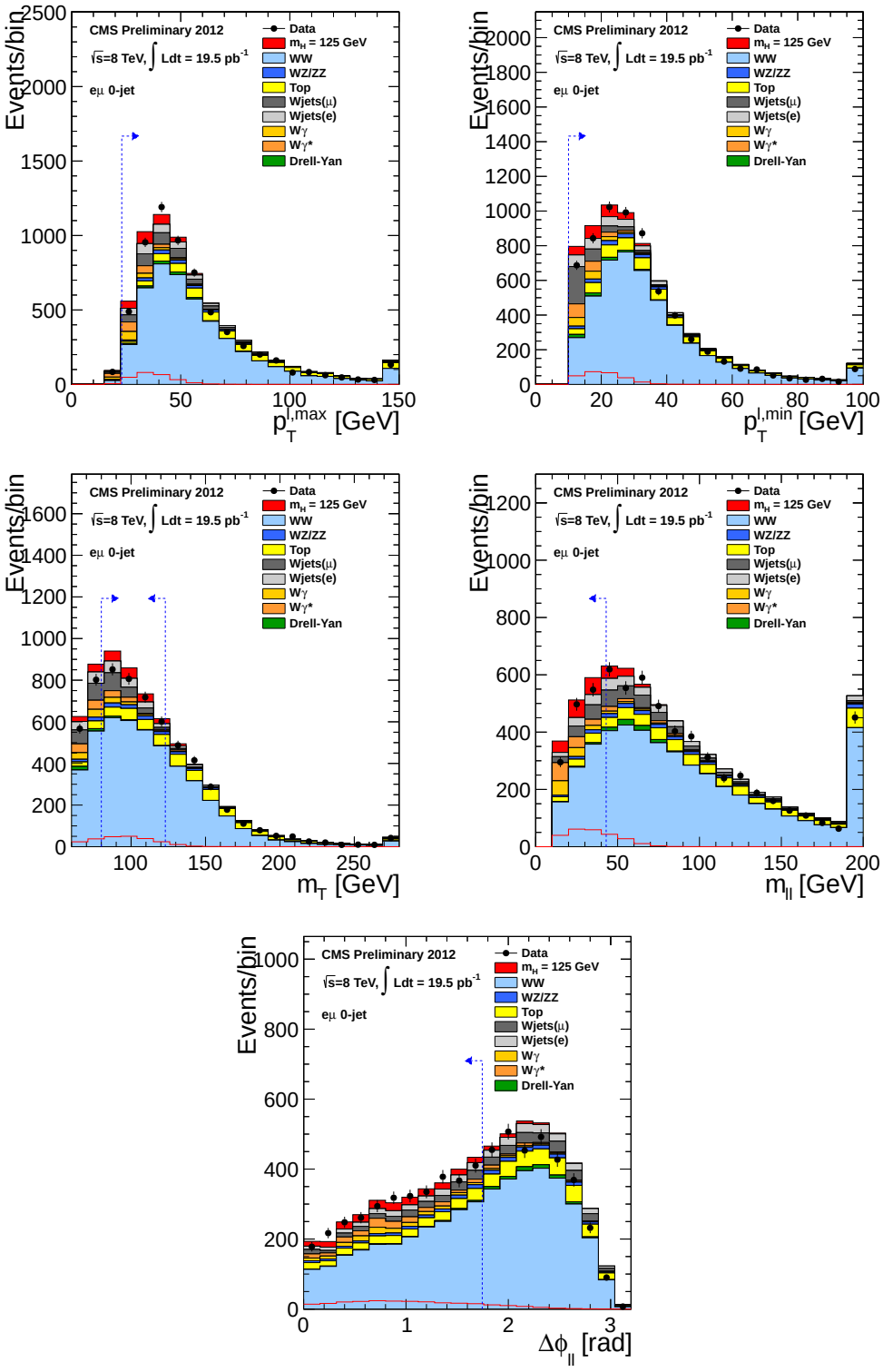


Figure 5.2: WW-level plots in 0-jet $e\mu/\mu e$ channel with m_{H}=125 GeV signal overlaid. Cuts for m_{H}=125 GeV is shown with blue dotted lines and arrows.

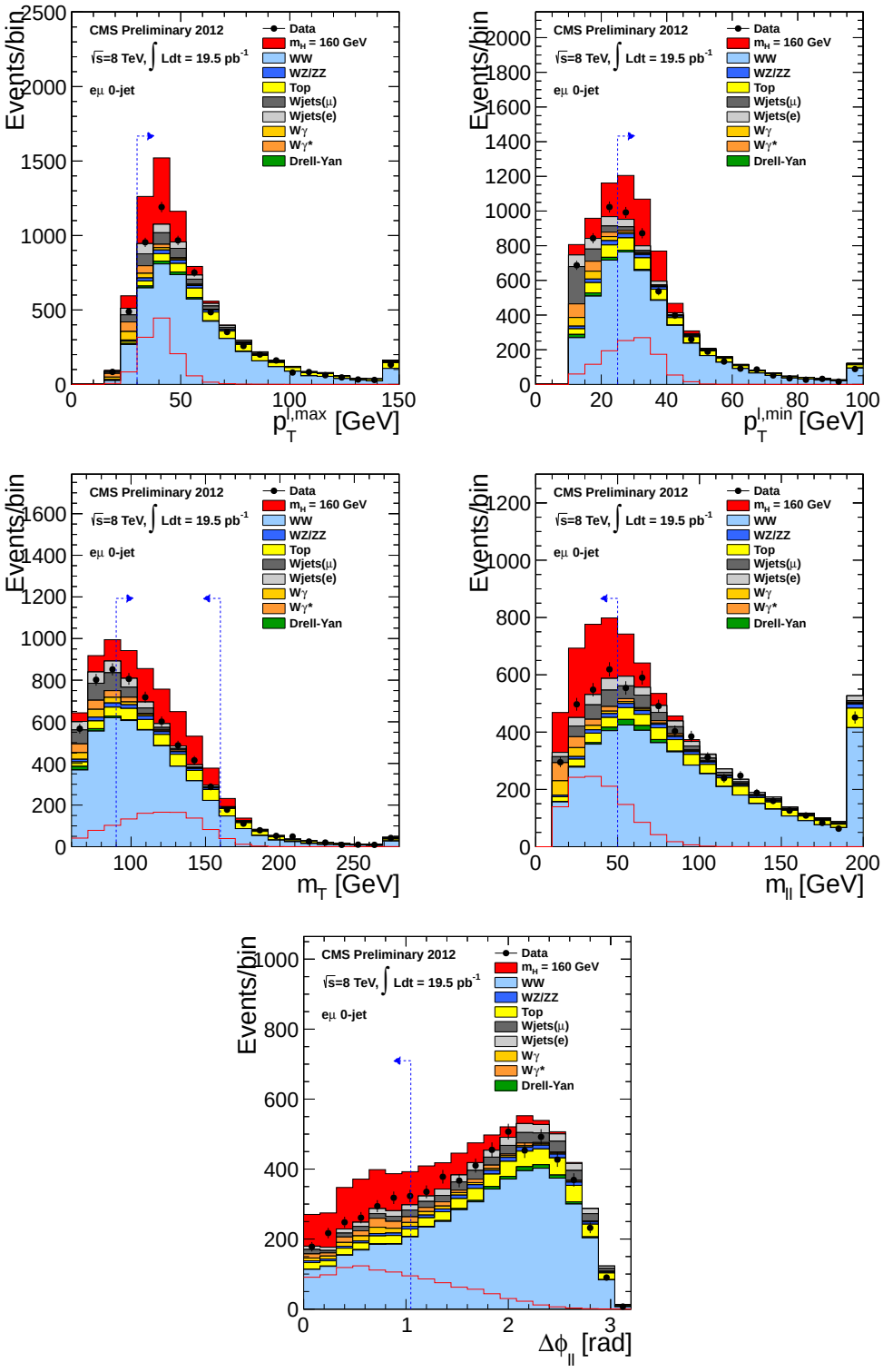


Figure 5.3: WW-level plots in 0-jet $e\mu/\mu e$ channel with m_{H}=160 GeV signal overlaid. Cuts for m_{H}=160 GeV is shown with blue dotted lines and arrows.

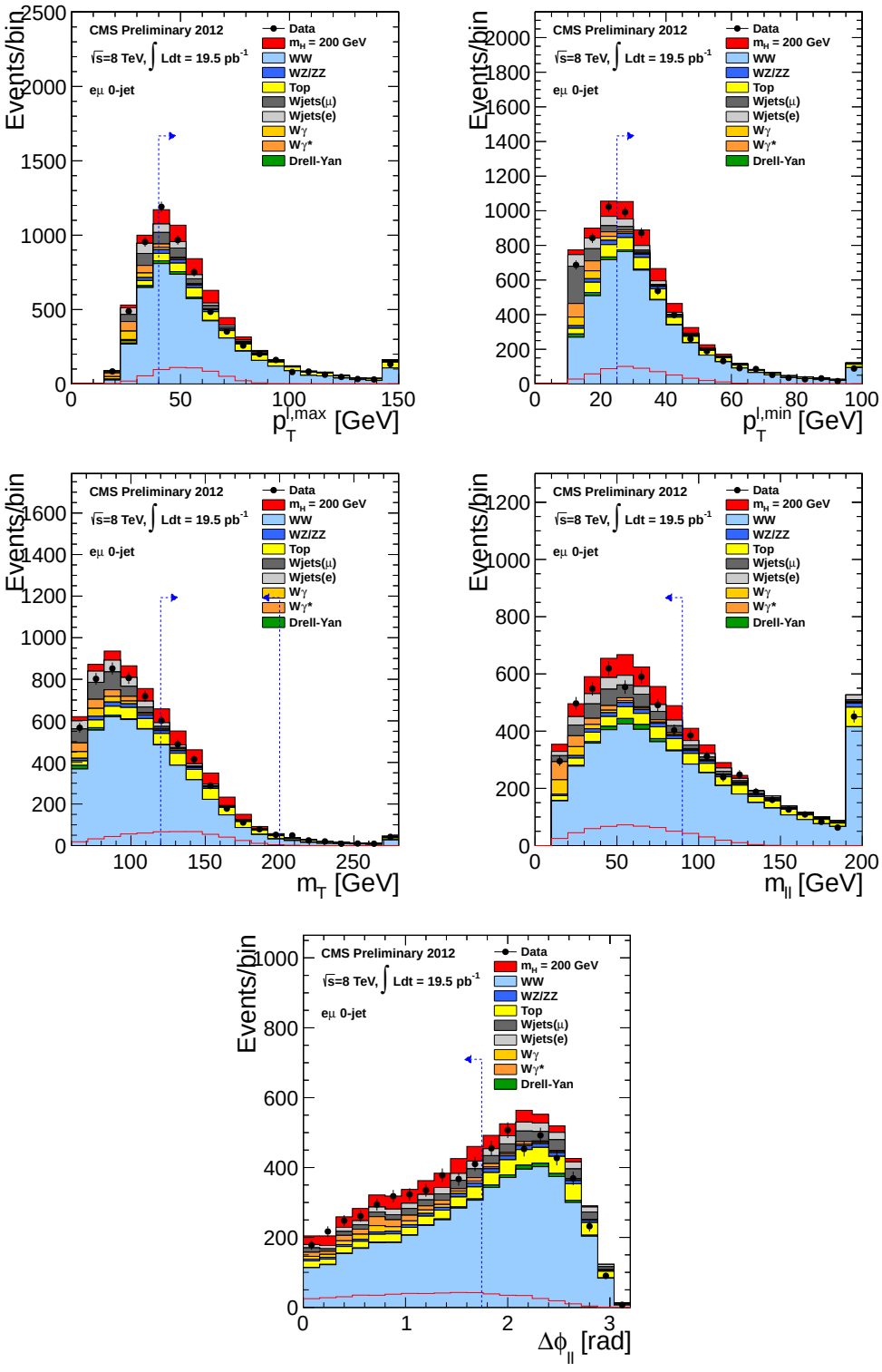


Figure 5.4: WW-level plots in 0-jet $e\mu/\mu e$ channel with m_{H}=200 GeV signal overlaid. Cuts for m_{H}=200 GeV is shown with blue dotted lines and arrows.

Table 5.1: m_{H}-dependent event selection for the cut-based analysis in the 0-jet and 1-jet categories.

m_{H} [GeV]	$p_{\mathrm{T}}^{\ell,\mathrm{max}}$ [GeV]	$p_{\mathrm{T}}^{\ell,\mathrm{min}}$ [GeV]	$m_{\ell\ell}$ [GeV]	$\Delta\phi_{\ell\ell}$ [$\cdot$]	m_{T} [GeV]
	$>$	$>$	$<$	$<$	
115	20	10	40	115	80 - 110
120	20	10	40	115	80 - 120
125	23	10	43	100	80 - 123
130	25	10	45	90	80 - 125
135	25	12	45	90	80 - 128
140	25	15	45	90	80 - 130
145	25	15	45	90	80 - 130
150	27	25	50	90	80 - 150
160	30	25	50	60	90 - 160
170	34	25	50	60	110 - 170
180	36	25	60	70	120 - 180
190	38	25	80	90	120 - 190
200	40	25	90	100	120 - 200
250	55	25	150	140	120 - 250
300	70	25	200	175	120 - 300
350	80	25	250	175	120 - 350
400	90	25	300	175	120 - 400
450	110	25	350	175	120 - 450
500	120	25	400	175	120 - 500
550	130	25	450	175	120 - 550
600	140	25	500	175	120 - 600

5.2 Shape-based Method

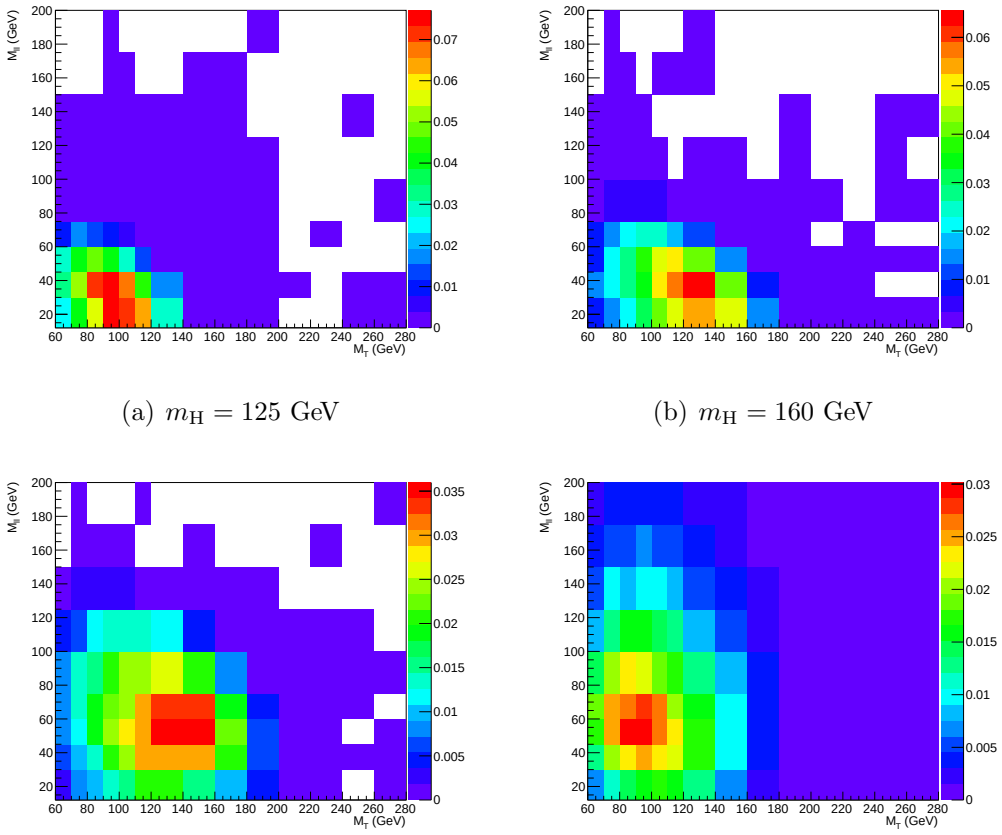
The "shape-based method" uses binned 2-dimensional templates of m_{T} and $m_{\ell\ell}$. This method is more complicated than the cut-based method because it uses the shape and the additional uncertainties to the shape on top of normalization

should be taken into account. To extract the signal component, the 2-dimensional binned shape is fitted to data. The shapes for most of the background are taken from simulation except the W + jets which is taken from data subtracted by simulation. Therefore, it is critical to make sure that the fit model correctly describes what is observed in data. There was a huge effort to validate this using pseudo-data and data-control region. Chapter 10 describes the details.

There are two main motivations for developing this analysis method. Before this method was employed as the main analysis method for the public result of this analysis at CMS, the BDT-based multivariate method had been the main analysis method [67]. However, because of the non-trivial dependence of the discriminator(BDT score) on the input variables, it was very difficult to have a physical interpretation of the observed result with this method. For example, if there is a fluctuation in data, we do not know where this comes from, *i.e.* which input variable is responsible for this unexpected behavior. In addition, because a single template definition is used for multiple m_{H} hypotheses, it avoids unnecessary statistical fluctuations between data selected for adjacent hypotheses. On the practical side, the implementation of the analysis became much simpler compared to the BDT method because of using the same background template across different m_{H} hypotheses while the BDT method used m_{H}-dependent selection similar to the cut-based method. This not only simplified the work flow but also allowed to draw the 2-dimensional log likelihood scan in the plane of signal strength and m_{H}, which was not possible with the m_{H}-dependent selections for some technical reasons.

The modeling of MET in $Z/\gamma^* \to \ell\ell$ background is quite poor in the tail of high MET. This is because high MET in this process can be obtained when jet energy is poorly measured. This corresponds to the tail of jet energy response distribution which is not well modelled by a Gaussian function which is used to estimate the resolution in the bulk of the distribution. Given that the Monte Carlo simulation is based on random sampling, this part is not particularly well modelled. In addition, the statistics of the available MC sample is limited resulting in huge weights per event. Therefore, we can not rely on $Z/\gamma^* \to \ell\ell$ simulation, and the shape-based method is applied to only $e\mu/\mu e$ channels.

The choice of the variables is motivated by the fact that the best variable that discriminates signal from background is the Higgs mass. But, due to the neutrinos in the final state, the Higgs mass can not be reconstructed. Therefore, we use the two m_{H}-like variables, the Higgs transverse mass(m_{T}) and the di-lepton mass($m_{\ell\ell}$) to construct the 2-dimensional templates. These two variables are weakly correlated in the main background processes.



(a) $m_{\mathrm{H}} = 125$ GeV

(b) $m_{\mathrm{H}} = 160$ GeV

(c) $m_{\mathrm{H}} = 200$ GeV

(d) $qq \to W^+W^-$

Figure 5.5: 2D templates for $m_{\mathrm{H}} = 125, 160, 200$ GeV and $qq \to W^+W^-$. For visualization, each bin is divided by the area of the bin in order to avoid random peaks due to difference in the bin size.

For the determination of the range and the binning, following points were considered :

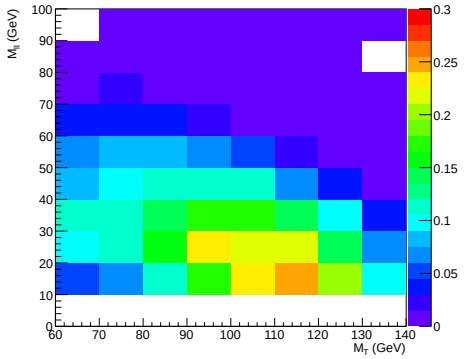
- Range should cover multiple m_{H}

- Binning should allow enough granularity to distinguish the signal from the background in the region where signal is populated : sideband region can have coarser binning in case of low statistics in that region
- statistical uncertainty of the templates should be small with respect to the total background : otherwise the templates will not be reliable
- There should not be any empty bin when all backgrounds are summed : otherwise any excess in data will claim infinite significance, *i.e.* something is observed when nothing expected
- The data events in the bins should be reasonably populated so that the nuisances can be constrained by data

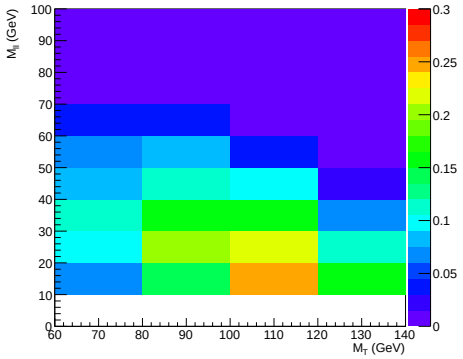
Due to large difference in the kinematics of signal events between high and low m_{H}, we decided to use two template definitions, one for the low Higgs masses, m_{H}= 110 - 250 GeV, and the other for the high Higgs masses, m_{H}= 300 - 600 GeV. In order to further enhance S/B, $p_{\mathrm{T}}^{\ell,\mathrm{max}} > 50$ GeV is applied for the high m_{H} templates. So, for the low(high) m_{H} hypotheses, we use $60 < m_{\mathrm{T}} < 280$ GeV($60 < m_{\mathrm{T}} < 380$ GeV) and $0 < m_{\ell\ell} < 280$ GeV($0 < m_{\ell\ell} < 450$ GeV). The events with high m_{H} hypothesis tend to have high m_{T} and high $m_{\ell\ell}$ but this is the phase space where background processes are not populated. Thus, for the high m_{H} templates, the top and the right bins are overflow bins allowing events up to $m_{\ell\ell} < 600$ GeV and $m_{\mathrm{T}} < 600$ GeV, which covers m_{H} hypothesis up to 600 GeV.

The width of the m_{T} and $m_{\ell\ell}$ distribution in signal events is 20 - 25 GeV for low m_{H} events, so we take 20-25 GeVas a maximum size of a bin for low m_{H} templates. Figure 5.6 shows the S/B with 4 different binnings in the region where m_{H}=125 GeV signal is populated($60 < m_{\mathrm{T}} < 120$ GeV, $0 < m_{\ell\ell} < 100$ GeV) in the 0-jet $e\mu/\mu e$category ; bin size of [m_{T}, $m_{\ell\ell}$] = [10 GeV, 10 GeV], [20 GeV, 10 GeV], [10 GeV, 20 GeV], and [20 GeV, 20 GeV]. The S/B with the bin size of 10 - 20 GeV does not change S/B, so we can use any bin size in that range as long as each bin contains enough statistics.

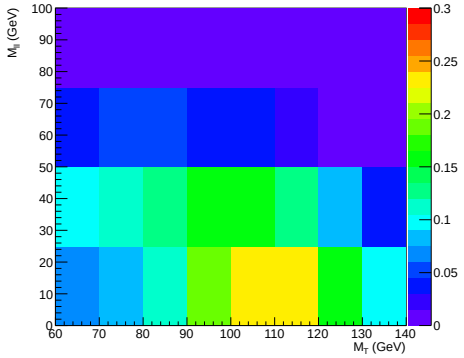
This method gives about 25 % better sensitivity at m_{H}=125 GeV in terms of expected significance compared to the state-of-the-art analysis method(BDT



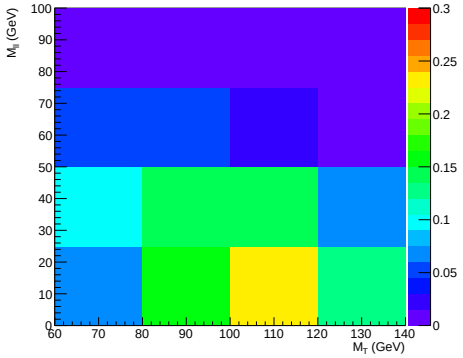
(a) $m_{\mathrm{T}} = 10$ GeV, $m_{\ell\ell} = 10$ GeV



(b) $m_{\mathrm{T}} = 20$ GeV, $m_{\ell\ell} = 10$ GeV



(c) $m_{\mathrm{T}} = 10$ GeV, $m_{\ell\ell} = 20$ GeV



(d) $m_{\mathrm{T}} = 20$ GeV, $m_{\ell\ell} = 25$ GeV

Figure 5.6: S/B using different binnings. The region($60 < m_{\mathrm{T}} < 120$ GeV, $0 < m_{\ell\ell} < 100$ GeV) is shown.

Table 5.2: Expected significance with different binnings and range combinations. The result is from 0-jet $e\mu/\mu e$ with 12.1 fb^{-1}of data. As a reference, the expected significance from the BDT method is 1.80.

binning $m_{\mathrm{T}} \times m_{\ell\ell}$	range [GeV]	expected significance	binning $m_{\mathrm{T}} \times m_{\ell\ell}$	range [GeV]	expected significance
2×2	m_{T} : 80 - 125 $m_{\ell\ell}$: 12 - 80	1.76	2×4	m_{T} : 80 - 120 $m_{\ell\ell}$: 0 - 100	1.89
3×3	m_{T} : 80 - 125 $m_{\ell\ell}$: 12 - 80	1.85	5×4	m_{T} : 80 - 180 $m_{\ell\ell}$: 0 - 100	2.06
4×4	m_{T} : 80 - 125 $m_{\ell\ell}$: 12 - 80	1.88	8×6	m_{T} : 80 - 240 $m_{\ell\ell}$: 0 - 150	2.19

method) [67] when this method was developed. The improvement comes from using expanded phase space which can be used to constrain backgrounds further. Table 5.2 shows the expected significance in the 0-jet $e\mu/\mu e$ category varying the bin size and the range. The left part of the table shows the result when the bin size is varied while the range is fixed to the selection that BDT method used. The expected significance is consistent in all cases. The right part of the table shows the result when the bin size is fixed while the range is gradually expanded to high m_{t} and $m_{\ell\ell}$ region. The expected significance increases significantly as the range is expanded.

Considering all these as well as a necessity of fine bins in the signal region for the spin-parity hypothesis separation test which will be discussed in chapter 12, we use the binning and the ranges as shown in Table 5.3.

The Figure 5.7 - 5.10 show the templates in the $e\mu/\mu e$ 0-jet category at 8 TeV and the relative statistical uncertainty of the template with respect to the total background for each process. There is not any background process that has large statistical uncertainty. More templates can be found in the Appendix **??**.

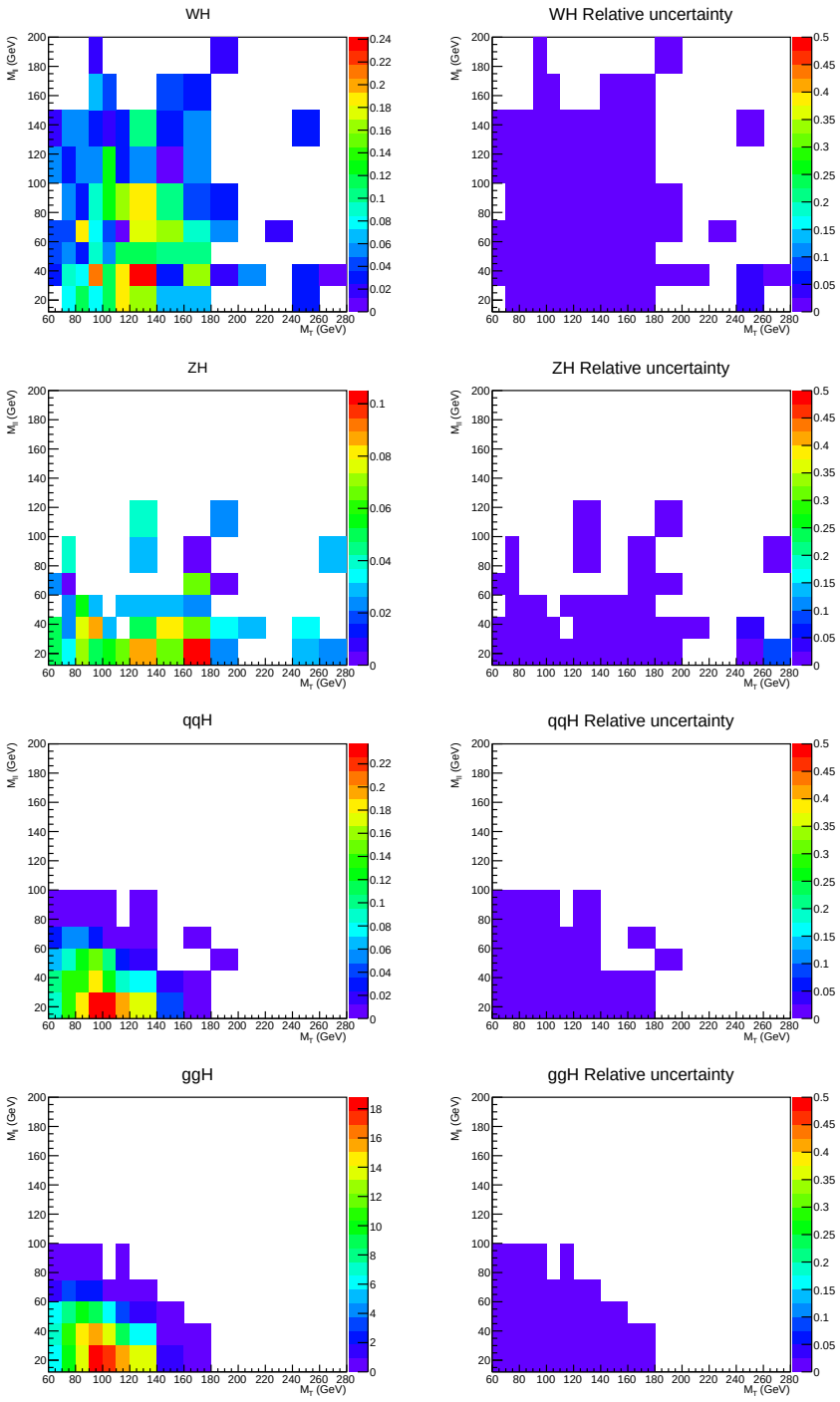


Figure 5.7: Templates on the left and relative statistical uncertainty of the MC sample on the right of $qq \to WH$, $qq \to ZH$, $qq \to qqH$ and $gg \to H$. The templates are for $m_{\mathrm{H}} = 125$ GeV analysis in the 0-jet category.

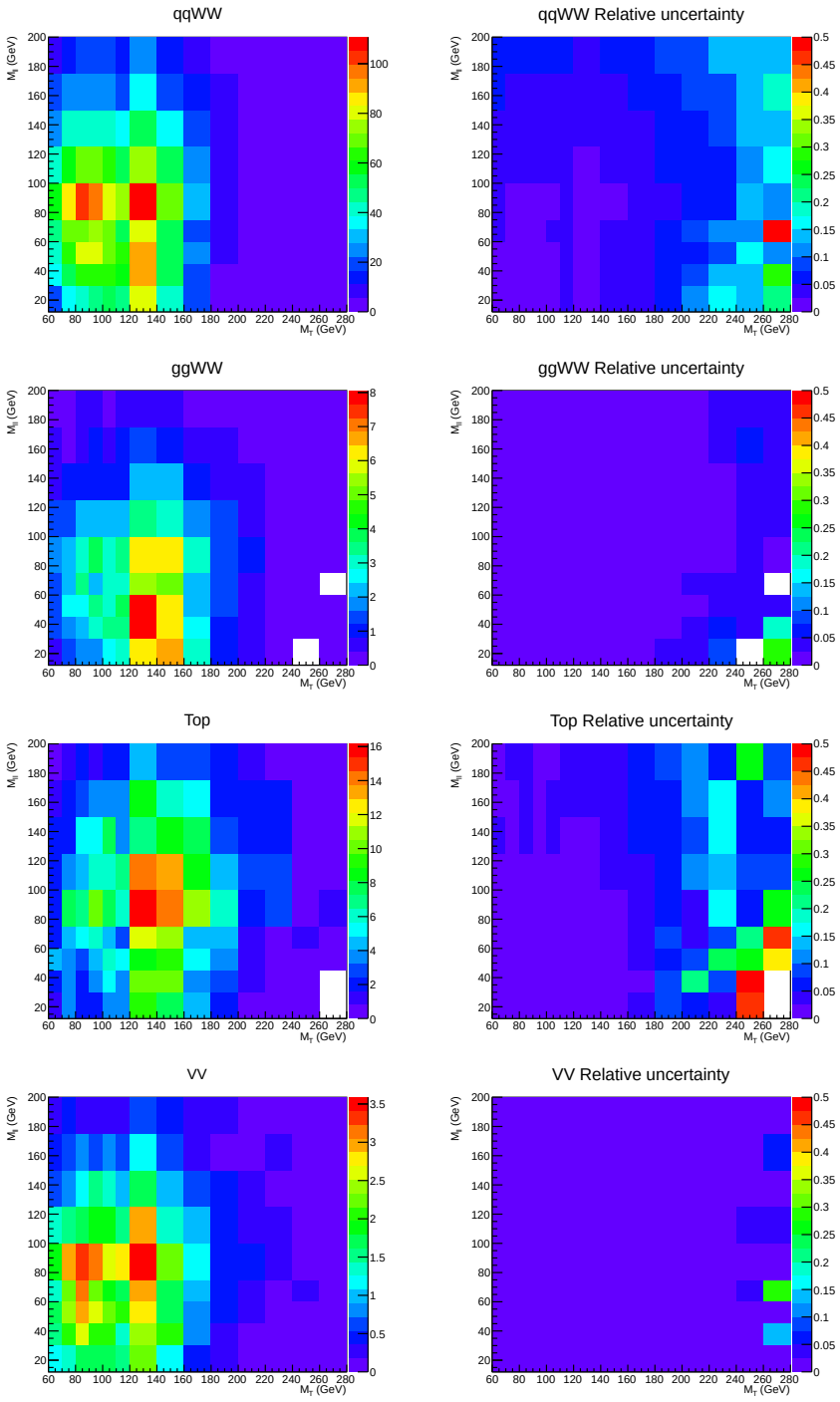


Figure 5.8: Templates on the left and relative statistical uncertainty of the MC sample on the right of $qq \to W^+W^-$, $gg \to W^+W^-$, $t\bar{t}/tW$ and WZ/ZZ. The templates are for $m_{\mathrm{H}} = 125$ GeV analysis in the 0-jet category.

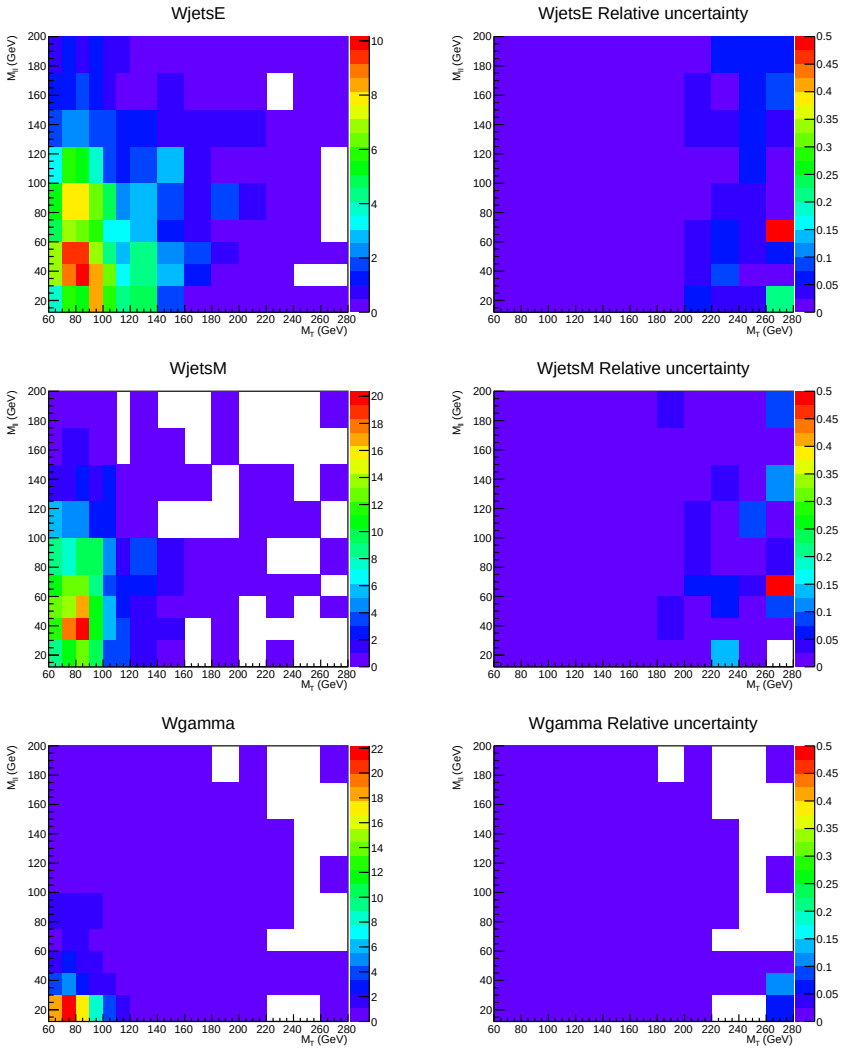


Figure 5.9: Templates on the left and relative statistical uncertainty of the MC sample on the right of $\mathrm{W}(\mathrm{e}\nu_{\mathrm{e}}) + \mathrm{jets}$, $\mathrm{W}(\mu\nu_{\mu}) + \mathrm{jets}$ and $W\gamma$. The templates are for $m_{\mathrm{H}} = 125$ GeV analysis in the 0-jet category.

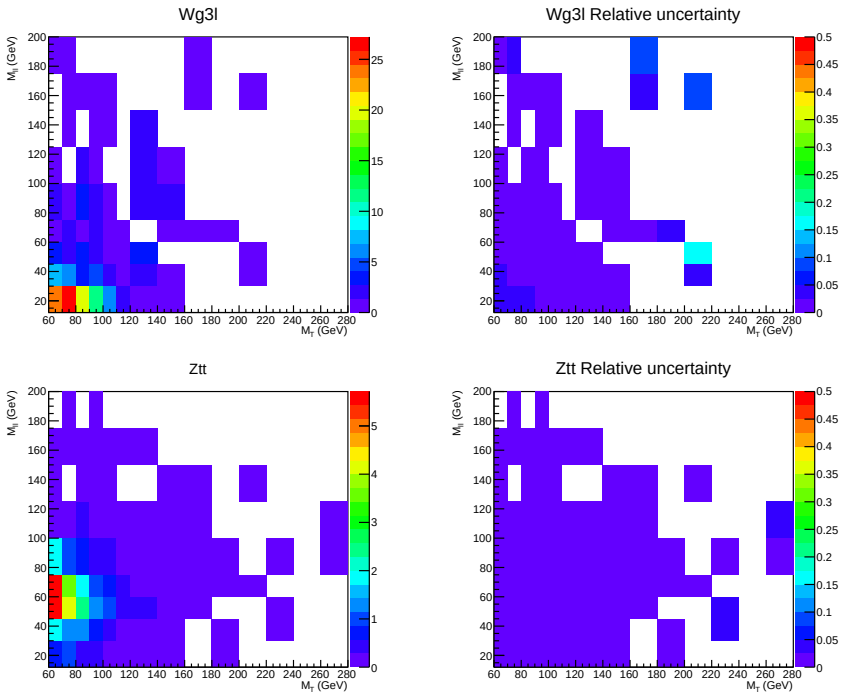


Figure 5.10: Templates on the left and relative statistical uncertainty of the MC sample on the right of $W\gamma^*$ and $Z \to \tau\tau$. The templates are for $m_{\mathrm{H}} = 125$ GeV analysis in the 0-jet category.

Table 5.3: Summary of template parameters. For the high-m_{H} templates, overflow up to $m_{\ell\ell}$=600 GeV and m_{T}=600 GeV is included in the last bin.

m_{H}(GeV)	Variable	Binning
110 - 250	m_{T} 14 bins	60, 70, 80, 90, 100, 110, 120, 140, 160, 180, 200, 220, 240, 260, 280
	$m_{\ell\ell}$ 9 bins	12, 30, 45, 60, 75, 100, 125, 150, 175, 200
300 - 600	m_{T} 10 bins	80, 110, 140, 170, 200, 230, 260, 290, 320, 350, 380
	$m_{\ell\ell}$ 8 bins	0, 56.25, 112.5, 168.75, 225, 281.25, 337.5, 393.75, 450

Chapter 6

Auxiliary Measurements

Due to the rapid rise in the instantaneous luminosity at LHC, the data taking conditions have changed rapidly. In particular it is difficult to exactly reproduce the number of pileup between data and simulation, and thus there will be differences in the number of reconstructed primary vertices. We correct this disagreement by reweighting the simulation to match the number of pileup events in data.

When the collisions happen, events are recorded when they pass the trigger requirements. Thus, per event trigger efficiency needs to be measured using data. To select identified and isolated leptons from Higgs events, we require that each lepton meet strict offline criteria. To account for the possible difference of offline selection efficiencies in simulation and data, we need to measure it in simultion and data, and apply corrections to the simulation.

The analysis is optimized in the different categories of number of jets. But, jet counting efficiency which is defined as the fraction of events in the certain category can be different in simulation and data. This is important for the signal because we rely on the simulation to calculate the jet counting efficiency. So, we correct the efficiency of $H \to W^+W^- \to 2l2\nu$ in simulation by a data/simulation scale factor measured in simulation and data using $Z/\gamma^* \to \ell\ell$ events.

This chapter describes these auxiliary measurements.

6.1 Pileup Re-weighting

The target pileup distribution for data is generated using the instantaneous luminosity per bunch crossing for each luminosity section, and the total pp inelastic cross section of 69.4 mb, integrated over the full data-taking period. The source distribution is taken from $Z/\gamma^* \to \ell\ell$ simulation without applying any cuts.

Figure 6.1 shows the number of reconstructed primary vertices from $Z/\gamma^* \to \ell\ell$ events in data and simulation after this re‘weighting procedure. The residual differences reflect the size of the systematic uncertainty in the determination of the amount of pileup present in the data. Since the dependence of the efficiencies for selecting signal and background events on pileup are small, this systematic

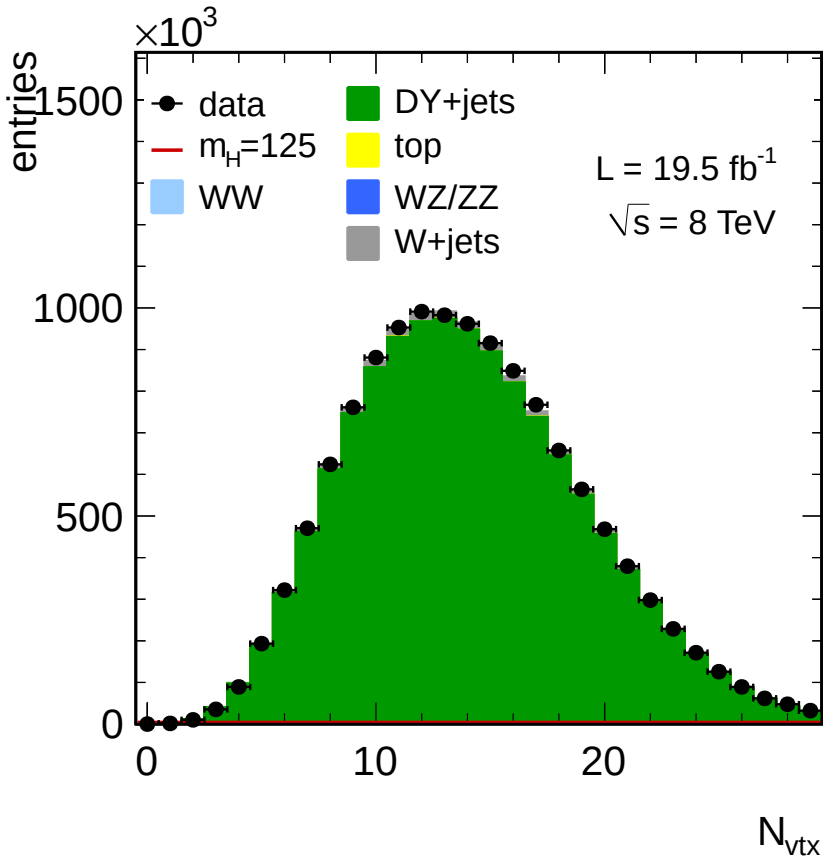


Figure 6.1: Number of reconstructed primary vertices for data and simulation after the re-weighting procedure for the 2012 dataset.

uncertainty due to the pileup is small when propagated to the final result.

6.2 Lepton Efficiency

The lepton selection efficiencies are measured using the Tag-And-Probe method [60]. It uses $Z/\gamma^* \to e^+e^-$ or $Z/\gamma^* \to \mu^+\mu^-$ events that pass the single-lepton trigger to select an unbiased sample of prompt leptons. One lepton which is called "tag" is required to pass the single lepton trigger and the full offline selection. By requiring the full selection on the tag, we can enhance the purity of the sample. The other lepton which is called "probe" is required to pass a set of selection criteria that enhances the purity of the sample while leaving the criteria under study unbiased. The fact that the tag passed single-lepton trigger guarantees that the probe is not biased by triggers. Both legs can be used as a tag in the offline lepton selection efficiency measurement while only one of them can be used in trigger efficiency measurement due to the correlation between the two online lepton objects in the double-lepton triggers.

The offline selection efficiency is composed of identification(ID) and isolation(ISO) parts. We use the N-1 method where ID and ISO efficiencies are measured with respect to the other, and multiplied afterwards. When measuring the efficiency of the ID part (ϵ_{ID}) the probe is required to pass the full isolation requirement, and when measuring the efficiency of the ISO part (ϵ_{ISO}) the probe is required to pass the full identification requirement. The efficiency is measured in bins of p_{T} and η in order to account for the dependences on the kinematics and the material budget of the detector.

Requiring tight selections to the tag, and using the N-1 method gives a high-purity sample of Z/γ^* events in data. However, there are residual contributions from non-prompt leptons from W+jets and QCD events especially at low p_T. Thus, we extract yields by fitting the $m_{\ell\ell}$ distribution of the tag and probe pair with "Gaussian $\oplus$ (exponential $\oplus$ error function)". The Gaussian function which is to describe the Z/γ^* invariant mass peak is modeled by simulation, and Gaussian smearing is applied to account for the difference in lepton momentum resolutions

between data and simulation. The (exponential $\oplus$ error function) is for the backgrounds. To allow enough sidebands for estimating backgrounds, the $m_{\ell\ell}$ range is chosen to be 60 GeV $< m_{\ell\ell} <$ 120 GeV. In case of simulation, efficiencies are calculated by counting events in 81 GeV $< m_{\ell\ell} <$ 101 GeV. In addition, in order to select the leptons from Z/γ^* only, the dR between the probe and the closest, same flavor, generator-level lepton after final-state-radiation(FSR) is required to be within 0.2.

The data/simulation scale factors per event are then calculated by multiplying per lepton efficiencies measured in data and simulation,

$$\epsilon_{offline}^{event} = \epsilon_{ID}^{lepton1} \times \epsilon_{ISO}^{lepton1} \times \epsilon_{ID}^{lepton2} \times \epsilon_{ISO}^{lepton2}. \tag{6.1}$$

The trigger efficiency($\epsilon_{trigger}$) is measured with respect to the full offline selection. Finally, the simulation is corrected by applying the offline selection scale factor and the trigger efficiency,

$$\frac{\epsilon_{offline,data}^{event}}{\epsilon_{offline,simulation}^{event}} \times \epsilon_{trigger}. \tag{6.2}$$

In the N-1 method, a possible bias due to the correlation between the ID and the ISO requirements is estimated by comparing efficiencies measured by N-1 method and combined measurement of ID+ISO efficiency in simulation. The difference which ranges from 0 to 8 % depending on the kinematic bins of the probe is assigned as systematic uncertainty of the method. In addition, the uncertainty of lepton reconstruction efficiency is taken as a systematic uncertainty.

6.2.1 Electron Selection Efficiency

The electron efficiency is composed of two factors, the reconstruction efficiency and the selection efficiency. The reconstruction efficiency is the probability for a super-cluster energy deposit to be matched to a reconstructed ECAL-driven GSF electron. In this analysis, we use the result measured with 2010 data [68]. The efficiency is measured as a function of p_{T} and η of a lepton. From the study, we assume that the reconstruction efficiency for an electron is unity with an uncertainty at the level of 2 %. Some representative plots of di-lepton mass distribution

are shown in Figure 6.2 and 6.3. Figure 6.2 shows di-lepton mass distributions for data(top) and MC(bottom) in $0.8 < |\eta| < 1.5$ and $10 < p_{\mathrm{T}} < 15$ GeV for events where the probe passes and fails ID selection on left and right, respectively. In data, there is a considerable contribution from backgrounds, particularly in the failing case(right). This is accounted in the fit. The blue and red lines in the data plots are for signal and background, respectively. Figure 6.3 shows di-lepton mass distributions for data(top) and MC(bottom) in $0.8 < |\eta| < 1.5$ and $30 < p_{\mathrm{T}} < 40$ GeV for events where the probe passes and fails ID selection on left and right, respectively. The blue and red lines in the data plots are for signal and background, respectively. Apart from the low p_{T} case in Figure 6.2, the there is negligible contribution from backgrounds.

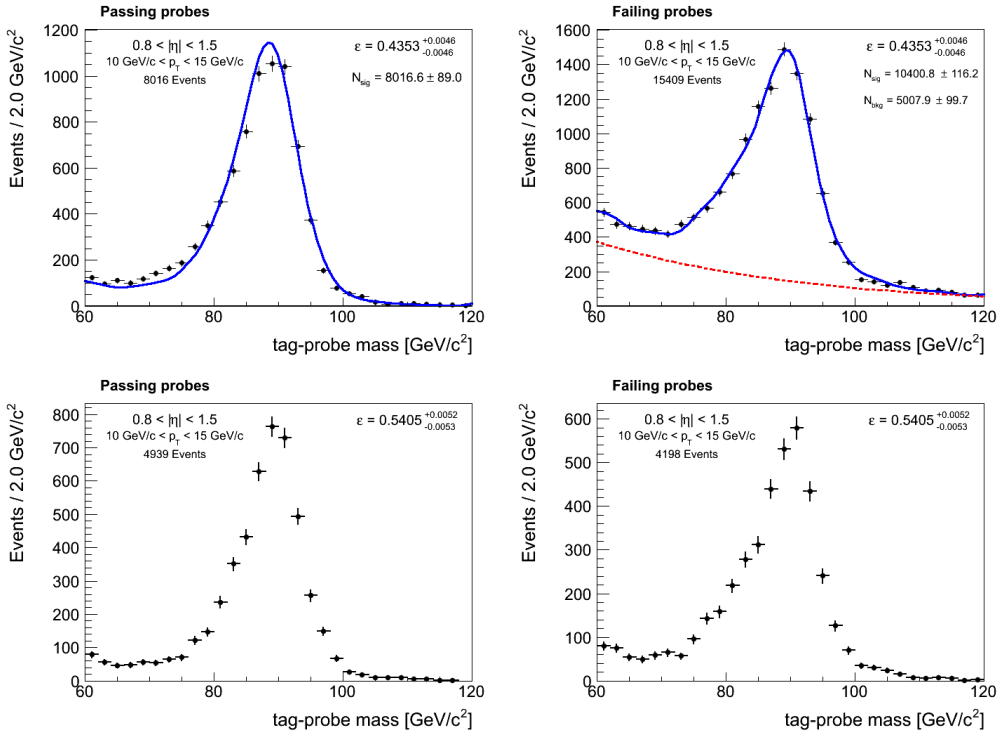


Figure 6.2: Di-lepton mass distributions for data(top) and MC(bottom) in $0.8 < |\eta| < 1.5$ and $10 < p_{\mathrm{T}} < 15$ GeV for events where the probe passes and fails the ID selection on left and right, respectively.

Table 6.1 and 6.2 show the electron N-1 efficiencies in data and simulation, and data/simulation scale factors for ID and ISO selections, respectively. The scale

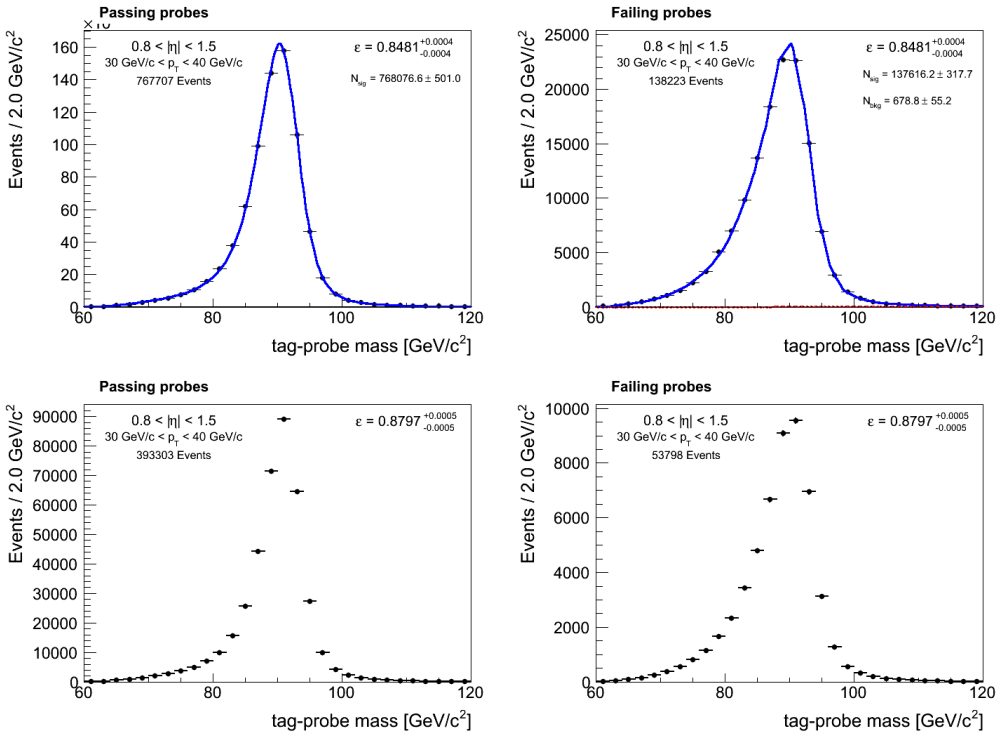


Figure 6.3: Di-lepton mass distributions for data(top) and MC(bottom) in $0.8 < |\eta| < 1.5$ and $30 < p_{\mathrm{T}} < 40$ GeV for events where the probe passes and fails the ID selection on left and right, respectively.

Table 6.1: Electron ID efficiency in data and simulation, and the scale factors.

	$0 < \|\eta\| < 0.8$	$0.8 < \|\eta\| < 1.479$	$1.479 < \|\eta\| < 2$	$2 < \|\eta\| < 2.5$
N-1 Efficiencies in data				
$10 < p_T < 15$	0.3289 ± 0.0049	0.4353 ± 0.0046	0.1551 ± 0.0040	0.1059 ± 0.0026
$15 < p_T < 20$	0.5981 ± 0.0026	0.6330 ± 0.0028	0.3140 ± 0.0033	0.2379 ± 0.0030
$20 < p_T < 30$	0.7208 ± 0.0009	0.7457 ± 0.0010	0.5147 ± 0.0011	0.4609 ± 0.0021
$30 < p_T < 40$	0.8293 ± 0.0002	0.8481 ± 0.0004	0.6780 ± 0.0015	0.5962 ± 0.0015
$40 < p_T < 50$	0.8623 ± 0.0004	0.8840 ± 0.0005	0.7558 ± 0.0007	0.6573 ± 0.0009
$50 < p_T < 7000$	0.8745 ± 0.0004	0.8936 ± 0.0005	0.7859 ± 0.0012	0.6891 ± 0.0015
N-1 Efficiencies in simulation				
$10 < p_T < 15$	0.4583 ± 0.0055	0.5405 ± 0.0053	0.2267 ± 0.0058	0.1890 ± 0.0056
$15 < p_T < 20$	0.6671 ± 0.0030	0.7126 ± 0.0031	0.3904 ± 0.0045	0.3426 ± 0.0047
$20 < p_T < 30$	0.7594 ± 0.0010	0.7943 ± 0.0011	0.5762 ± 0.0019	0.5090 ± 0.0021
$30 < p_T < 40$	0.8492 ± 0.0005	0.8797 ± 0.0005	0.7215 ± 0.0010	0.6291 ± 0.0012
$40 < p_T < 50$	0.8774 ± 0.0004	0.9091 ± 0.0004	0.7843 ± 0.0007	0.6915 ± 0.0010
$50 < p_T < 7000$	0.8893 ± 0.0007	0.9169 ± 0.0007	0.8039 ± 0.0013	0.7149 ± 0.0017
data/simulation scale factors				
$10 < p_T < 15$	0.7177 ± 0.0138	0.8053 ± 0.0117	0.6842 ± 0.0249	0.5602 ± 0.0214
$15 < p_T < 20$	0.8966 ± 0.0056	0.8882 ± 0.0055	0.8045 ± 0.0126	0.6943 ± 0.0128
$20 < p_T < 30$	0.9491 ± 0.0017	0.9388 ± 0.0019	0.8933 ± 0.0035	0.9056 ± 0.0056
$30 < p_T < 40$	0.9766 ± 0.0006	0.9641 ± 0.0007	0.9396 ± 0.0024	0.9477 ± 0.0030
$40 < p_T < 50$	0.9828 ± 0.0006	0.9724 ± 0.0007	0.9637 ± 0.0013	0.9507 ± 0.0018
$50 < p_T < 7000$	0.9834 ± 0.0009	0.9746 ± 0.0009	0.9776 ± 0.0021	0.9639 ± 0.0032

Table 6.2: Electron ISO efficiency in data and simulation, and the scale factors.

	$0 < \|\eta\| < 0.8$	$0.8 < \|\eta\| < 1.479$	$1.479 < \|\eta\| < 2$	$2 < \|\eta\| < 2.5$
N-1 Efficiencies in data				
$10 < p_T < 15$	0.7827 ± 0.0042	0.7973 ± 0.0029	0.8009 ± 0.0199	0.8546 ± 0.0110
$15 < p_T < 20$	0.8167 ± 0.0007	0.8360 ± 0.0034	0.8155 ± 0.0145	0.8776 ± 0.0041
$20 < p_T < 30$	0.8798 ± 0.0011	0.8815 ± 0.0007	0.8787 ± 0.0017	0.9246 ± 0.0091
$30 < p_T < 40$	0.9391 ± 0.0000	0.9398 ± 0.0007	0.9337 ± 0.0006	0.9598 ± 0.0004
$40 < p_T < 50$	0.9710 ± 0.0001	0.9721 ± 0.0002	0.9704 ± 0.0002	0.9802 ± 0.0002
$50 < p_T < 7000$	0.9816 ± 0.0002	0.9815 ± 0.0008	0.9815 ± 0.0003	0.9873 ± 0.0004
N-1 Efficiencies in simulation				
$10 < p_T < 15$	0.7724 ± 0.0061	0.8015 ± 0.0052	0.7856 ± 0.0109	0.8516 ± 0.0110
$15 < p_T < 20$	0.8214 ± 0.0027	0.8362 ± 0.0028	0.8213 ± 0.0052	0.8761 ± 0.0053
$20 < p_T < 30$	0.8850 ± 0.0008	0.8906 ± 0.0009	0.8768 ± 0.0016	0.9091 ± 0.0017
$30 < p_T < 40$	0.9464 ± 0.0003	0.9473 ± 0.0003	0.9360 ± 0.0006	0.9484 ± 0.0007
$40 < p_T < 50$	0.9757 ± 0.0002	0.9768 ± 0.0002	0.9708 ± 0.0003	0.9720 ± 0.0004
$50 < p_T < 7000$	0.9843 ± 0.0003	0.9842 ± 0.0003	0.9827 ± 0.0005	0.9809 ± 0.0006
data/simulation scale factors				
$10 < p_T < 15$	1.0133 ± 0.0096	0.9948 ± 0.0074	1.0195 ± 0.0290	1.0036 ± 0.0183
$15 < p_T < 20$	0.9943 ± 0.0034	0.9999 ± 0.0053	0.9930 ± 0.0188	1.0017 ± 0.0076
$20 < p_T < 30$	0.9941 ± 0.0016	0.9898 ± 0.0013	1.0021 ± 0.0027	1.0170 ± 0.0102
$30 < p_T < 40$	0.9923 ± 0.0003	0.9921 ± 0.0008	0.9976 ± 0.0009	1.0121 ± 0.0008
$40 < p_T < 50$	0.9951 ± 0.0002	0.9952 ± 0.0003	0.9996 ± 0.0004	1.0084 ± 0.0005
$50 < p_T < 7000$	0.9973 ± 0.0003	0.9973 ± 0.0008	0.9988 ± 0.0005	1.0065 ± 0.0008

factors are close to unity except the low p_{T} bins for ID efficiency.

Table 6.3: The additional systematic uncertainty δ_{SF} on the scale factor for the electron selection due to the N-1 factorization scheme.

	$0 < \|\eta\| < 0.8$	$0.8 < \|\eta\| < 1.479$	$1.479 < \|\eta\| < 2$	$2 < \|\eta\| < 2.5$
$10 < p_T < 15$	0.075	0.043	0.089	0.091
$15 < p_T < 20$	0.020	0.018	0.045	0.041
$20 < p_T < 30$	0.005	0.007	0.016	0.006
$30 < p_T < 40$	0.001	0.001	0.003	0.001
$40 < p_T < 50$	0.000	0.000	0.000	0.001
$50 < p_T$	0.000	0.000	0.000	0.000

Table 6.3 shows the systematic uncertainty of the N-1 method. The uncertainty is estimated by comparing measured efficiencies using N-1 method and the measuring the full (ID and ISO) efficiency at the same time in simulation.

6.2.2 Muon Selection Efficiency

As in the electron case, muon efficiency is composed of two factors, the reconstruction efficiency and the selection efficiency. The reconstruction efficiency is the probability for a well-reconstructed track in the muon chamber to be matched to a reconstructed track in the inner tracker. In this analysis, we use the result measured with 2010 data [68]. The efficiency is measured as a function of p_{T} and η of a lepton. From the study, we assume that the reconstruction efficiency for a muon is unity with uncertainty at the level of 1.5 %.

Table 6.4 and 6.5 show the muon N-1 efficiencies in data and simulation, and data-to-simulation scale factors for ID and ISO selections, respectively. The scale factors are close to unity in all bins.

In case of muon, the systematic uncertainty due to N-1 method is negligible($<$ 0.2 %) compared to the reconstruction efficiency which is at the level of 1.5 %.

Table 6.4: Muon ID efficiency in data and simulation, and scale factors.

	$0 < \|\eta\| < 0.8$	$0.8 < \|\eta\| < 1.2$	$1.2 < \|\eta\| < 2.4$
N-1 Efficiencies in data			
$10 < p_T < 15$	0.9650 ± 0.0023	0.9576 ± 0.0023	0.9352 ± 0.0014
$15 < p_T < 20$	0.9652 ± 0.0005	0.9500 ± 0.0016	0.9389 ± 0.0008
$20 < p_T < 30$	0.9687 ± 0.0004	0.9565 ± 0.0006	0.9497 ± 0.0011
$30 < p_T < 40$	0.9720 ± 0.0002	0.9611 ± 0.0006	0.9536 ± 0.0025
$40 < p_T < 50$	0.9732 ± 0.0001	0.9640 ± 0.0004	0.9599 ± 0.0005
$50 < p_T < 7000$	0.9675 ± 0.0004	0.9546 ± 0.0006	0.9331 ± 0.0003
N-1 Efficiencies in simulation			
$10 < p_T < 15$	0.9774 ± 0.0015	0.9750 ± 0.0020	0.9537 ± 0.0015
$15 < p_T < 20$	0.9808 ± 0.0008	0.9738 ± 0.0013	0.9580 ± 0.0010
$20 < p_T < 30$	0.9828 ± 0.0003	0.9739 ± 0.0005	0.9624 ± 0.0004
$30 < p_T < 40$	0.9844 ± 0.0001	0.9766 ± 0.0003	0.9659 ± 0.0002
$40 < p_T < 50$	0.9849 ± 0.0001	0.9778 ± 0.0002	0.9696 ± 0.0002
$50 < p_T < 7000$	0.9819 ± 0.0002	0.9715 ± 0.0004	0.9493 ± 0.0004
data/simulation scale factors			
$10 < p_T < 15$	0.9872 ± 0.0028	0.9821 ± 0.0031	0.9806 ± 0.0021
$15 < p_T < 20$	0.9841 ± 0.0009	0.9755 ± 0.0021	0.9801 ± 0.0013
$20 < p_T < 30$	0.9857 ± 0.0005	0.9821 ± 0.0008	0.9868 ± 0.0012
$30 < p_T < 40$	0.9874 ± 0.0002	0.9841 ± 0.0007	0.9873 ± 0.0026
$40 < p_T < 50$	0.9881 ± 0.0002	0.9859 ± 0.0004	0.9899 ± 0.0005
$50 < p_T < 7000$	0.9854 ± 0.0005	0.9826 ± 0.0008	0.9829 ± 0.0005

Table 6.5: Muon ISO efficiency in data and simulation, and scale factors.

	$0 < \|\eta\| < 0.8$	$0.8 < \|\eta\| < 1.2$	$1.2 < \|\eta\| < 2.4$
N-1 Efficiencies in data			
$10 < p_T < 15$	0.6693 ± 0.0037	0.6776 ± 0.0040	0.7590 ± 0.0017
$15 < p_T < 20$	0.7447 ± 0.0021	0.7615 ± 0.0030	0.8347 ± 0.0000
$20 < p_T < 30$	0.8903 ± 0.0008	0.8932 ± 0.0006	0.9044 ± 0.0006
$30 < p_T < 40$	0.9606 ± 0.0007	0.9636 ± 0.0007	0.9659 ± 0.0002
$40 < p_T < 50$	0.9837 ± 0.0001	0.9855 ± 0.0000	0.9886 ± 0.0001
$50 < p_T < 7000$	0.9875 ± 0.0002	0.9879 ± 0.0002	0.9910 ± 0.0002
N-1 Efficiencies in simulation			
$10 < p_T < 15$	0.6556 ± 0.0037	0.6832 ± 0.0048	0.7322 ± 0.0027
$15 < p_T < 20$	0.7519 ± 0.0020	0.7691 ± 0.0030	0.8139 ± 0.0017
$20 < p_T < 30$	0.8954 ± 0.0006	0.8962 ± 0.0009	0.8847 ± 0.0006
$30 < p_T < 40$	0.9642 ± 0.0002	0.9664 ± 0.0003	0.9584 ± 0.0002
$40 < p_T < 50$	0.9857 ± 0.0001	0.9878 ± 0.0002	0.9872 ± 0.0001
$50 < p_T < 7000$	0.9885 ± 0.0002	0.9902 ± 0.0003	0.9899 ± 0.0002
data/simulation scale factors			
$10 < p_T < 15$	1.0209 ± 0.0081	0.9917 ± 0.0092	1.0365 ± 0.0044
$15 < p_T < 20$	0.9904 ± 0.0038	0.9901 ± 0.0055	1.0256 ± 0.0021
$20 < p_T < 30$	0.9944 ± 0.0011	0.9966 ± 0.0012	1.0222 ± 0.0010
$30 < p_T < 40$	0.9962 ± 0.0007	0.9971 ± 0.0008	1.0078 ± 0.0003
$40 < p_T < 50$	0.9980 ± 0.0001	0.9977 ± 0.0002	1.0014 ± 0.0002
$50 < p_T < 7000$	0.9990 ± 0.0003	0.9977 ± 0.0003	1.0011 ± 0.0002

6.2.3 Trigger Efficiency

The analysis uses a combination of the single-lepton and the double-lepton triggers. For the double-lepton triggers, there is a requirement on dZ, the longitudinal distance between two lepton vertices, on top of the requirement on each leg. This requirement is imposed to select events from hard interactions, not from pileup events. Only events that pass both requirements are recorded in the data sample under study. Thus, there is 100 % correlation between the two leptons, *i.e.* if one lepton passed the per lepton requirement, the other lepton passed it as well, otherwise trigger objects have not been stored at all in the data samples. This introduces a slight change in the Tag-And-Probe method such that only one tag is selected in an event, and we do it by selecting the tag randomly.

The dZ requirement for the double-lepton triggers is designed to be highly efficient. However, in the early part of data in 2012 there was a technical issue that caused inefficiency of dZ requirement for the double-muon triggers at the level of 15 %. The inefficiency is absorbed by using single-lepton triggers to a negligible level. Thus, in the per event trigger efficiency calculation we assume dZ efficiency is 100 %.

The efficiency of the each leg of the double-lepton triggers is measured separately because there are different requirements imposed to them. For $e\mu$ triggers, we assume that the efficiency of both legs can be modelled by measurements of per leg efficiency of double-electron and double-muon triggers.This assumption was validated by measuring $e\mu$ trigger efficiency using $t\bar{t}$ events with MET > 20 GeV. In order to avoid possible bias, the muon leg efficiency was measured using events passing single-electron triggers and the electron leg efficiency was measured using events passing single-muon triggers. The result turned out to be consistent with our model using per leg efficiency from measurements of double-lepton trigger within statistical uncertainties.

Using the per leg efficiency of the double-lepton triggers and the single-lepton triggers, the per event trigger efficiency is calculated. The requirement of the single-lepton trigger is tighter than the requirements applied to each leg in double-lepton trigger. So, there are only three cases where an event passes the

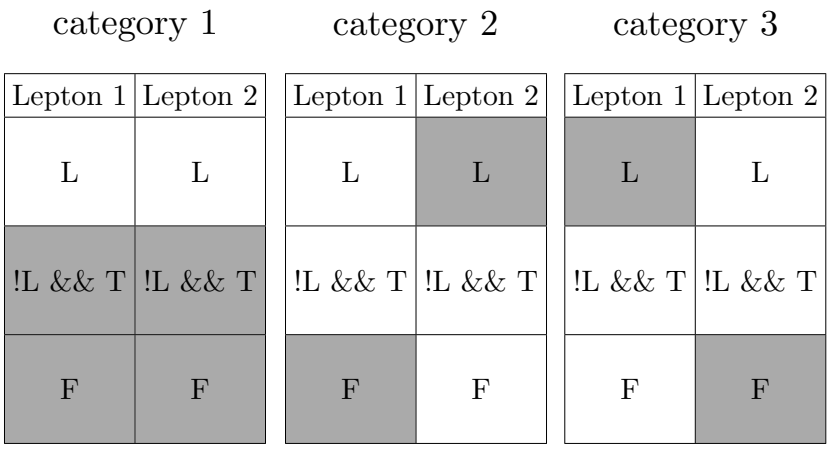


Figure 6.4: Diagrams for the cases where double-lepton triggers fail. Lepton1(2) denotes offline leptons.

L means leading leg requirement of double-lepton trigger. T means trailing leg requirement of double-lepton trigger. F means failing of trailing lepton requirement, thus leading lepton requirement,of double-lepton trigger. Shade means that the corresponding offline lepton falls into that online requirement.

trigger requirements.

1. Both leptons pass the requirement on each leg in the double-lepton trigger and dZ

2. Both leptons pass the requirement on each leg in the double-lepton trigger, but failed dZ. At least one of the leptons pass the single-lepton trigger.

3. One of the leptons fails the double-lepton trigger requirements, but the other lepton passes the single-lepton trigger.

As mentioned above, dZ efficiency is assumed to be 100 % in our calculation, so (2) is not included in the per event efficiency calculation. Figure 6.4 shows the failing cases of the double-lepton trigger. Online, there are three cases for each lepton

: passing leading leg requirement(L), passing trailing lepton requirement(T), and passing none of them(F). Shade means that the corresponding offline lepton falls into that online requirement. There are 9(3×3) possible combinations in total, of which there are 6 cases where the trigger fails. The category 1(4 combinations) is the case where none of the two leptons pass the leading leg requirement. The category 2 and 3 are the cases where one lepton passes the leading leg requirement, but the other leg fails the trailing leg requirement. Converting this into an equation, we have

$$\begin{aligned}
\epsilon_{double-lepton}\left(p_{T1},\eta_1,p_{T2},\eta_2\right) &= 1-\Big[&& (6.3)\\
&\underbrace{\left(1-\epsilon_{DL}\left(p_{T1},\eta_1\right)\right)\left(1-\epsilon_{DL}\left(p_{T2},\eta_2\right)\right)}_{\text{category 1}} && (6.4)\\
&+\underbrace{\epsilon_{DL}\left(p_{T2},\eta_2\right)\left(1-\epsilon_{DT}\left(p_{T1},\eta_1\right)\right)}_{\text{category 2}} && (6.5)\\
&+\underbrace{\epsilon_{DL}\left(p_{T1},\eta_1\right)\left(1-\epsilon_{DT}\left(p_{T2},\eta_2\right)\right)}_{\text{category 3}}\Big] && (6.6)
\end{aligned}$$

where $p_{T1(2)}$ and $\eta_{1(2)}$ are p_{T} and η of offline leptons, $\epsilon_{DL}(\epsilon_{DT})$ is the efficiency for the leading(trailing) leg of the double-lepton trigger. Each term corresponds to the 3 categories in Figure 6.4. In case the double-lepton trigger fails because one of the legs fails, the other leg which passed double-lepton trigger requirement might pass the single lepton trigger. This way, inefficiency of double-lepton trigger can be recovered by the single-lepton trigger. The recovery of efficiency by the single-lepton trigger is

$$\begin{aligned}
\epsilon_{single-lepton}\left(p_{T1},\eta_1,p_{T2},\eta_2\right) &= \underbrace{\left(1-\epsilon_{DT}\left(p_{T1},\eta_1\right)\right)\epsilon_S\left(p_{T2},\eta_2\right)}_{\text{category 4}} && (6.7)\\
&+\underbrace{\left(1-\epsilon_{DT}\left(p_{T2},\eta_2\right)\right)\epsilon_S\left(p_{T1},\eta_1\right)}_{\text{category 5}} && (6.8)
\end{aligned}$$

where ϵ_S is the efficiency of the single-lepton trigger. Therefore, the total per-event

trigger efficiency is given by

$$\begin{aligned}
\epsilon_{event}\left(p_{T1},\eta_1,p_{T2},\eta_2\right) &= 1-\Big[\left(1-\epsilon_{DL}\left(p_{T1},\eta_1\right)\right)\left(1-\epsilon_{DL}\left(p_{T2},\eta_2\right)\right) && (6.9)\\
&\quad +\epsilon_{DL}\left(p_{T2},\eta_2\right)\left(1-\epsilon_{DT}\left(p_{T1},\eta_1\right)\right) && (6.10)\\
&\quad +\epsilon_{DL}\left(p_{T1},\eta_1\right)\left(1-\epsilon_{DT}\left(p_{T2},\eta_2\right)\right)\Big] && (6.11)\\
&\quad +\left(1-\epsilon_{DT}\left(p_{T1},\eta_1\right)\right)\epsilon_S\left(p_{T2},\eta_2\right) && (6.12)\\
&\quad +\left(1-\epsilon_{DT}\left(p_{T2},\eta_2\right)\right)\epsilon_S\left(p_{T1},\eta_1\right) && (6.13)
\end{aligned}$$

Table 6.6: Efficiency of dZ requirement of the double-electron trigger.

	$0<\lvert\eta\rvert<0.8$	$0.8<\lvert\eta\rvert<1.479$	$1.479<\lvert\eta\rvert<2$	$2<\lvert\eta\rvert<2.5$
$10<p_T<12.5$	1.0000 ± 0.0045	0.9946 ± 0.0043	0.9853 ± 0.0191	0.9724 ± 0.0213
$12.5<p_T<15$	1.0000 ± 0.0011	0.9934 ± 0.0024	1.0000 ± 0.0036	0.9773 ± 0.0095
$15<p_T<17.5$	0.9982 ± 0.0010	0.9971 ± 0.0011	0.9949 ± 0.0030	0.9886 ± 0.0048
$17.5<p_T<20$	0.9990 ± 0.0005	0.9969 ± 0.0008	0.9952 ± 0.0020	0.9884 ± 0.0034
$20<p_T<22.5$	0.9993 ± 0.0004	0.9980 ± 0.0006	0.9970 ± 0.0012	0.9873 ± 0.0024
$22.5<p_T<25$	0.9997 ± 0.0002	0.9980 ± 0.0004	0.9962 ± 0.0010	0.9911 ± 0.0016
$25<p_T<27.5$	0.9994 ± 0.0002	0.9982 ± 0.0003	0.9965 ± 0.0007	0.9905 ± 0.0013
$27.5<p_T<30$	0.9993 ± 0.0001	0.9985 ± 0.0002	0.9975 ± 0.0005	0.9908 ± 0.0011
$30<p_T<35$	0.9995 ± 0.0001	0.9986 ± 0.0001	0.9972 ± 0.0003	0.9920 ± 0.0005
$35<p_T<40$	0.9996 ± 0.0000	0.9987 ± 0.0001	0.9972 ± 0.0002	0.9921 ± 0.0004
$40<p_T<50$	0.9997 ± 0.0000	0.9990 ± 0.0000	0.9973 ± 0.0001	0.9925 ± 0.0003
$50<p_T<7000$	0.9997 ± 0.0000	0.9992 ± 0.0001	0.9978 ± 0.0002	0.9922 ± 0.0005

The efficiency of dZ, leading and trailing leg requirements for double-electron triggers is shown in Table 6.6 - 6.8. The efficiency of the single-electron trigger is shown in Table 6.9.

The efficiency of dZ, leading and trailing leg requirements for double-muon triggers is shown in Table 6.10 - 6.12. The efficiency of the single-muon trigger is shown in Table 6.13.

Table 6.7: Efficiency of the leading leg of the double-electron trigger.

	$0 < \|\eta\| < 0.8$	$0.8 < \|\eta\| < 1.479$	$1.479 < \|\eta\| < 2$	$2 < \|\eta\| < 2.5$
$10 < p_T < 12.5$	0.0000 ± 0.0041	0.0000 ± 0.0021	0.0000 ± 0.0102	0.0000 ± 0.0112
$12.5 < p_T < 15$	0.0000 ± 0.0011	0.0000 ± 0.0009	0.0092 ± 0.0062	0.0021 ± 0.0049
$15 < p_T < 17.5$	0.0437 ± 0.0035	0.0460 ± 0.0034	0.2456 ± 0.0128	0.2300 ± 0.0148
$17.5 < p_T < 20$	0.8312 ± 0.0044	0.6617 ± 0.0057	0.8570 ± 0.0080	0.8365 ± 0.0098
$20 < p_T < 22.5$	0.9618 ± 0.0019	0.9560 ± 0.0021	0.9768 ± 0.0027	0.9685 ± 0.0035
$22.5 < p_T < 25$	0.9709 ± 0.0013	0.9721 ± 0.0014	0.9843 ± 0.0018	0.9785 ± 0.0024
$25 < p_T < 27.5$	0.9784 ± 0.0009	0.9764 ± 0.0010	0.9879 ± 0.0013	0.9859 ± 0.0016
$27.5 < p_T < 30$	0.9823 ± 0.0006	0.9809 ± 0.0008	0.9884 ± 0.0010	0.9869 ± 0.0012
$30 < p_T < 35$	0.9849 ± 0.0003	0.9842 ± 0.0004	0.9901 ± 0.0005	0.9869 ± 0.0006
$35 < p_T < 40$	0.9880 ± 0.0002	0.9863 ± 0.0003	0.9925 ± 0.0003	0.9907 ± 0.0004
$40 < p_T < 50$	0.9900 ± 0.0001	0.9903 ± 0.0001	0.9945 ± 0.0002	0.9912 ± 0.0003
$50 < p_T < 7000$	0.9910 ± 0.0002	0.9925 ± 0.0002	0.9958 ± 0.0003	0.9911 ± 0.0005

Table 6.8: Efficiency of the trailing leg of the double-electron trigger.

	$0 < \|\eta\| < 0.8$	$0.8 < \|\eta\| < 1.479$	$1.479 < \|\eta\| < 2$	$2 < \|\eta\| < 2.5$
$10 < p_T < 12.5$	0.9101 ± 0.0157	0.8313 ± 0.0135	0.7598 ± 0.0362	0.8841 ± 0.0306
$12.5 < p_T < 15$	0.9633 ± 0.0051	0.9284 ± 0.0060	0.9316 ± 0.0126	0.9382 ± 0.0132
$15 < p_T < 17.5$	0.9685 ± 0.0030	0.9554 ± 0.0034	0.9572 ± 0.0066	0.9595 ± 0.0076
$17.5 < p_T < 20$	0.9673 ± 0.0022	0.9665 ± 0.0023	0.9716 ± 0.0041	0.9774 ± 0.0044
$20 < p_T < 22.5$	0.9695 ± 0.0017	0.9699 ± 0.0018	0.9762 ± 0.0028	0.9786 ± 0.0030
$22.5 < p_T < 25$	0.9731 ± 0.0012	0.9745 ± 0.0013	0.9764 ± 0.0021	0.9758 ± 0.0025
$25 < p_T < 27.5$	0.9771 ± 0.0009	0.9779 ± 0.0010	0.9831 ± 0.0015	0.9831 ± 0.0017
$27.5 < p_T < 30$	0.9810 ± 0.0007	0.9807 ± 0.0008	0.9829 ± 0.0012	0.9842 ± 0.0013
$30 < p_T < 35$	0.9828 ± 0.0003	0.9831 ± 0.0004	0.9830 ± 0.0006	0.9840 ± 0.0007
$35 < p_T < 40$	0.9850 ± 0.0002	0.9843 ± 0.0003	0.9861 ± 0.0004	0.9879 ± 0.0005
$40 < p_T < 50$	0.9870 ± 0.0001	0.9874 ± 0.0002	0.9883 ± 0.0002	0.9885 ± 0.0003
$50 < p_T < 7000$	0.9882 ± 0.0003	0.9893 ± 0.0003	0.9900 ± 0.0004	0.9888 ± 0.0006

Table 6.9: Efficiency of the single-electron trigger.

	$0 < \|\eta\| < 0.8$	$0.8 < \|\eta\| < 1.479$	$1.479 < \|\eta\| < 2$	$2 < \|\eta\| < 2.5$
$10 < p_T < 12.5$	0.0000 ± 0.0021	0.0000 ± 0.0010	0.0000 ± 0.0051	0.0000 ± 0.0057
$12.5 < p_T < 15$	0.0000 ± 0.0005	0.0000 ± 0.0004	0.0000 ± 0.0017	0.0000 ± 0.0020
$15 < p_T < 17.5$	0.0000 ± 0.0002	0.0000 ± 0.0002	0.0000 ± 0.0008	0.0000 ± 0.0010
$17.5 < p_T < 20$	0.0000 ± 0.0001	0.0000 ± 0.0001	0.0002 ± 0.0005	0.0000 ± 0.0006
$20 < p_T < 22.5$	0.0000 ± 0.0001	0.0000 ± 0.0001	0.0005 ± 0.0004	0.0007 ± 0.0005
$22.5 < p_T < 25$	0.0006 ± 0.0002	0.0006 ± 0.0002	0.0118 ± 0.0011	0.0250 ± 0.0018
$25 < p_T < 27.5$	0.0255 ± 0.0007	0.0251 ± 0.0007	0.1320 ± 0.0026	0.1636 ± 0.0032
$27.5 < p_T < 30$	0.6009 ± 0.0016	0.4072 ± 0.0018	0.4926 ± 0.0031	0.4710 ± 0.0035
$30 < p_T < 35$	0.8905 ± 0.0005	0.8634 ± 0.0007	0.6775 ± 0.0015	0.6602 ± 0.0018
$35 < p_T < 40$	0.9171 ± 0.0004	0.9012 ± 0.0004	0.7285 ± 0.0011	0.7103 ± 0.0013
$40 < p_T < 50$	0.9361 ± 0.0002	0.9239 ± 0.0003	0.7618 ± 0.0007	0.7298 ± 0.0009
$50 < p_T < 7000$	0.9471 ± 0.0004	0.9402 ± 0.0005	0.7808 ± 0.0012	0.7374 ± 0.0017

Table 6.10: Efficiency of dZ requirement of the double-muon trigger.

	$0 < \|\eta\| < 0.8$	$0.8 < \|\eta\| < 1.2$	$1.2 < \|\eta\| < 2.1$	$2.1 < \|\eta\| < 2.4$
$10 < p_T < 12.5$	0.9710 ± 0.0045	0.9687 ± 0.0045	0.9572 ± 0.0029	0.9682 ± 0.0044
$12.5 < p_T < 15$	0.9714 ± 0.0028	0.9692 ± 0.0032	0.9612 ± 0.0022	0.9617 ± 0.0040
$15 < p_T < 17.5$	0.9794 ± 0.0017	0.9702 ± 0.0026	0.9710 ± 0.0016	0.9688 ± 0.0031
$17.5 < p_T < 20$	0.9907 ± 0.0009	0.9838 ± 0.0016	0.9819 ± 0.0011	0.9764 ± 0.0023
$20 < p_T < 22.5$	0.9912 ± 0.0007	0.9862 ± 0.0012	0.9823 ± 0.0009	0.9755 ± 0.0020
$22.5 < p_T < 25$	0.9896 ± 0.0006	0.9873 ± 0.0009	0.9828 ± 0.0007	0.9785 ± 0.0016
$25 < p_T < 27.5$	0.9897 ± 0.0005	0.9886 ± 0.0007	0.9815 ± 0.0006	0.9784 ± 0.0014
$27.5 < p_T < 30$	0.9900 ± 0.0004	0.9874 ± 0.0006	0.9813 ± 0.0005	0.9794 ± 0.0011
$30 < p_T < 35$	0.9911 ± 0.0002	0.9869 ± 0.0004	0.9810 ± 0.0003	0.9805 ± 0.0006
$35 < p_T < 40$	0.9918 ± 0.0001	0.9862 ± 0.0003	0.9790 ± 0.0003	0.9789 ± 0.0006
$40 < p_T < 50$	0.9934 ± 0.0001	0.9855 ± 0.0002	0.9767 ± 0.0002	0.9777 ± 0.0004
$50 < p_T < 7000$	0.9937 ± 0.0002	0.9855 ± 0.0003	0.9747 ± 0.0004	0.9799 ± 0.0009

Table 6.11: Efficiency of the leading leg of the double-muon trigger.

	$0 < \|\eta\| < 0.8$	$0.8 < \|\eta\| < 1.2$	$1.2 < \|\eta\| < 2.1$	$2.1 < \|\eta\| < 2.4$
$10 < p_T < 12.5$	0.0005 ± 0.0012	0.0114 ± 0.0030	0.0034 ± 0.0010	0.0085 ± 0.0025
$12.5 < p_T < 15$	0.0005 ± 0.0006	0.0188 ± 0.0026	0.0068 ± 0.0010	0.0092 ± 0.0021
$15 < p_T < 17.5$	0.2500 ± 0.0049	0.2363 ± 0.0060	0.2688 ± 0.0040	0.2442 ± 0.0069
$17.5 < p_T < 20$	0.9696 ± 0.0015	0.9169 ± 0.0033	0.9036 ± 0.0023	0.8007 ± 0.0055
$20 < p_T < 22.5$	0.9714 ± 0.0011	0.9243 ± 0.0026	0.9138 ± 0.0018	0.8176 ± 0.0046
$22.5 < p_T < 25$	0.9717 ± 0.0009	0.9269 ± 0.0021	0.9236 ± 0.0015	0.8479 ± 0.0037
$25 < p_T < 27.5$	0.9717 ± 0.0007	0.9311 ± 0.0017	0.9221 ± 0.0012	0.8506 ± 0.0031
$27.5 < p_T < 30$	0.9712 ± 0.0006	0.9280 ± 0.0014	0.9233 ± 0.0010	0.8558 ± 0.0026
$30 < p_T < 35$	0.9706 ± 0.0003	0.9289 ± 0.0008	0.9198 ± 0.0006	0.8692 ± 0.0014
$35 < p_T < 40$	0.9722 ± 0.0002	0.9286 ± 0.0006	0.9206 ± 0.0005	0.8796 ± 0.0012
$40 < p_T < 50$	0.9726 ± 0.0002	0.9320 ± 0.0004	0.9215 ± 0.0003	0.8889 ± 0.0009
$50 < p_T < 7000$	0.9725 ± 0.0003	0.9337 ± 0.0007	0.9216 ± 0.0006	0.9016 ± 0.0018

Table 6.12: Efficiency of the trailing leg of the double-muon trigger.

	$0 < \|\eta\| < 0.8$	$0.8 < \|\eta\| < 1.2$	$1.2 < \|\eta\| < 2.1$	$2.1 < \|\eta\| < 2.4$
$10 < p_T < 12.5$	0.9838 ± 0.0035	0.9752 ± 0.0041	0.9779 ± 0.0021	0.9371 ± 0.0057
$12.5 < p_T < 15$	0.9850 ± 0.0021	0.9735 ± 0.0030	0.9814 ± 0.0016	0.9403 ± 0.0047
$15 < p_T < 17.5$	0.9846 ± 0.0015	0.9770 ± 0.0023	0.9817 ± 0.0013	0.9391 ± 0.0040
$17.5 < p_T < 20$	0.9824 ± 0.0012	0.9772 ± 0.0018	0.9807 ± 0.0011	0.9422 ± 0.0033
$20 < p_T < 22.5$	0.9831 ± 0.0009	0.9791 ± 0.0014	0.9812 ± 0.0009	0.9388 ± 0.0029
$22.5 < p_T < 25$	0.9824 ± 0.0007	0.9782 ± 0.0012	0.9826 ± 0.0007	0.9468 ± 0.0024
$25 < p_T < 27.5$	0.9837 ± 0.0006	0.9797 ± 0.0010	0.9822 ± 0.0006	0.9438 ± 0.0020
$27.5 < p_T < 30$	0.9831 ± 0.0005	0.9792 ± 0.0008	0.9829 ± 0.0005	0.9440 ± 0.0017
$30 < p_T < 35$	0.9827 ± 0.0002	0.9797 ± 0.0004	0.9814 ± 0.0003	0.9482 ± 0.0010
$35 < p_T < 40$	0.9840 ± 0.0002	0.9792 ± 0.0003	0.9825 ± 0.0002	0.9501 ± 0.0008
$40 < p_T < 50$	0.9846 ± 0.0001	0.9800 ± 0.0002	0.9831 ± 0.0002	0.9551 ± 0.0006
$50 < p_T < 7000$	0.9851 ± 0.0002	0.9804 ± 0.0004	0.9829 ± 0.0003	0.9563 ± 0.0013

Table 6.13: Efficiency of the single-muon trigger.

	$0 < \|\eta\| < 0.8$	$0.8 < \|\eta\| < 1.2$	$1.2 < \|\eta\| < 2.1$	$2.1 < \|\eta\| < 2.4$
$10 < p_T < 12.5$	0.0000 ± 0.0005	0.0000 ± 0.0005	0.0001 ± 0.0002	0.0000 ± 0.0004
$12.5 < p_T < 15$	0.0000 ± 0.0002	0.0001 ± 0.0003	0.0002 ± 0.0002	0.0000 ± 0.0003
$15 < p_T < 17.5$	0.0000 ± 0.0001	0.0004 ± 0.0003	0.0001 ± 0.0001	0.0000 ± 0.0002
$17.5 < p_T < 20$	0.0000 ± 0.0001	0.0015 ± 0.0004	0.0004 ± 0.0001	0.0000 ± 0.0002
$20 < p_T < 22.5$	0.0004 ± 0.0001	0.0030 ± 0.0004	0.0028 ± 0.0003	0.0000 ± 0.0001
$22.5 < p_T < 25$	0.4031 ± 0.0018	0.3699 ± 0.0027	0.3912 ± 0.0019	0.0000 ± 0.0001
$25 < p_T < 27.5$	0.8847 ± 0.0010	0.8076 ± 0.0018	0.7733 ± 0.0014	0.0001 ± 0.0001
$27.5 < p_T < 30$	0.8955 ± 0.0008	0.8172 ± 0.0015	0.7881 ± 0.0011	0.0002 ± 0.0001
$30 < p_T < 35$	0.9100 ± 0.0004	0.8267 ± 0.0008	0.7968 ± 0.0006	0.0001 ± 0.0000
$35 < p_T < 40$	0.9230 ± 0.0003	0.8368 ± 0.0006	0.8048 ± 0.0005	0.0001 ± 0.0000
$40 < p_T < 50$	0.9350 ± 0.0002	0.8480 ± 0.0004	0.8161 ± 0.0003	0.0001 ± 0.0000
$50 < p_T < 7000$	0.9408 ± 0.0003	0.8526 ± 0.0007	0.8207 ± 0.0006	0.0002 ± 0.0001

6.3 Jet Counting Efficiency

The jet counting efficiency of $H \to W^+W^- \to 2l2\nu$ events($\epsilon_{H\to W^+W^-\to 2l2\nu}$) is calculated by

$$\epsilon_{H\to W^+W^-\to 2l2\nu} = \epsilon^{MC}_{H\to W^+W^-\to 2l2\nu} \times \frac{\epsilon^{Data}_{Z/\gamma^*\to\ell\ell}}{\epsilon^{MC}_{Z/\gamma^*\to\ell\ell}} \tag{6.14}$$

where $\epsilon^{MC}_{H\to W^+W^-\to 2l2\nu}$ is the jet counting efficiency of $H \to W^+W^- \to 2l2\nu$ in simulation, and $\frac{\epsilon^{Data}_{Z/\gamma^*\to\ell\ell}}{\epsilon^{MC}_{Z/\gamma^*\to\ell\ell}}$ is the data/simulation scale factor measured using $Z/\gamma^* \to \ell\ell$ events. The efficiencies of $Z/\gamma^* \to \ell\ell$ events are evaluated using Drell-Yan events with di-lepton mass within 7.5 GeV of the Z peak. Figure 6.5 shows the distribution of the number of jets for data and MC. In all bins a good agreement with difference less than 1 % is observed.

The systematic uncertainty to jet counting efficiency comes from the statistics when measuring the data/simulation scale factor and the theoretical uncertainty of jet counting in simulation. The former is less than 1 % being negligible compared to the theoretical uncertainty which is at the level of 15 %. This uncertainty will be discussed in detail in chapter 9.

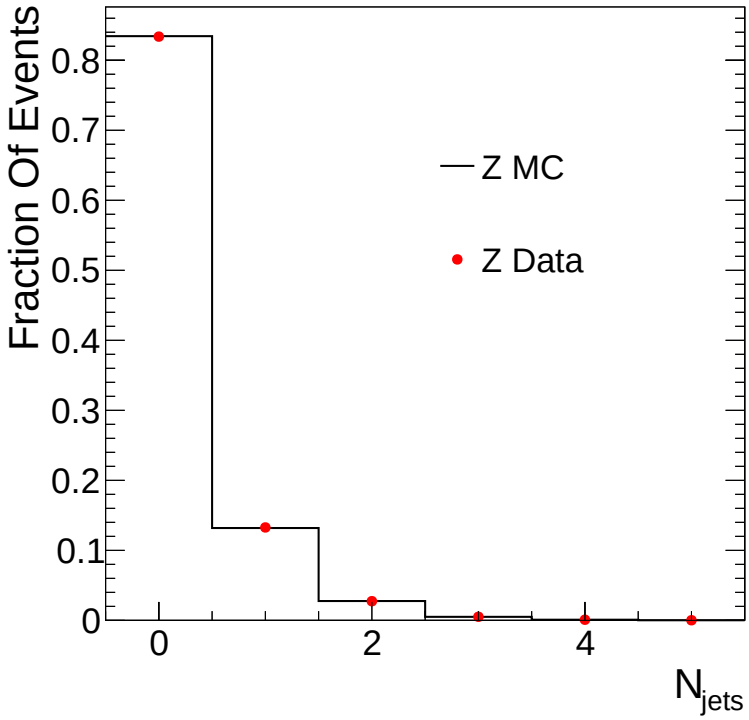


Figure 6.5: The number of Jets observed in data (red solid dot) and MC (black line) for the Z events.

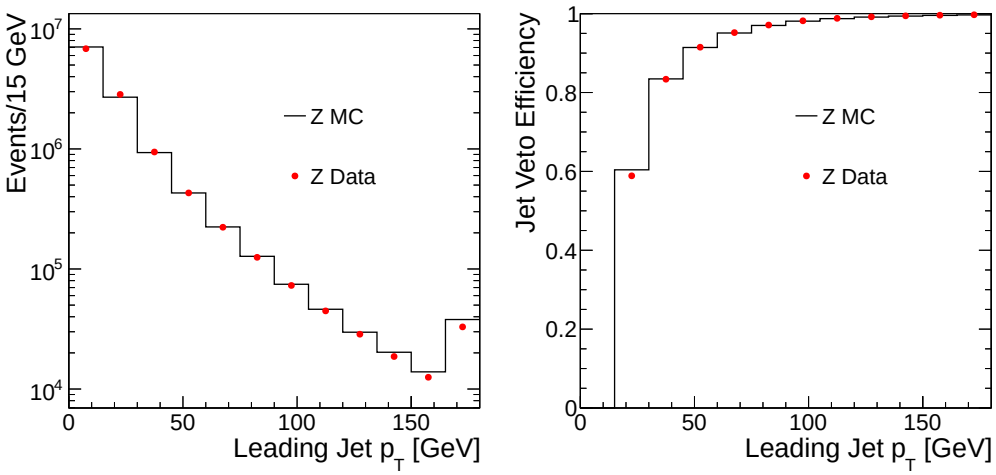


Figure 6.6: The leading jet p_{T}(left) and the jet veto efficiency as a function of the leading jet p_{T}(right) in data(red solid dot) and MC(black line) for the Z events.

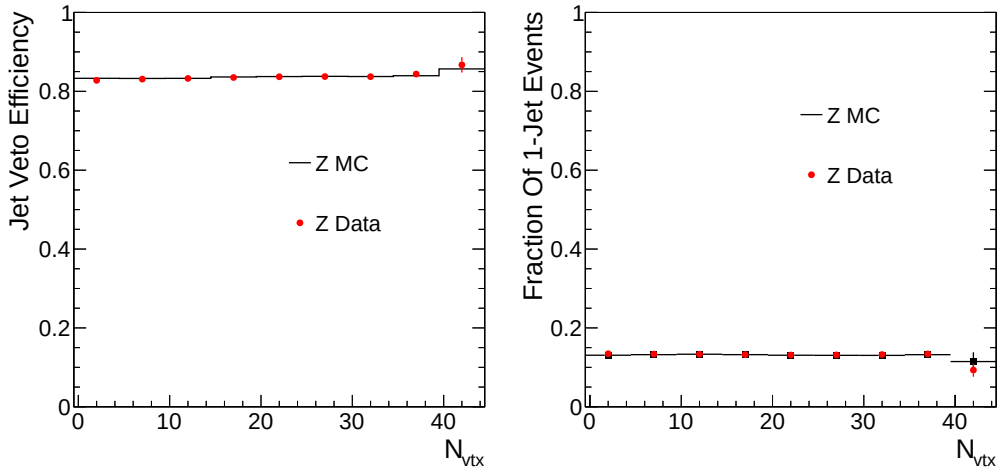


Figure 6.7: The fraction of events with 0-jet(left) and 1-jet(right) events as a function of the number of vertices, comparing data (red solid dot) and MC (black solid square) for the Z events.

Figure 6.6 shows the p_{T} distribution of the leading jet(left) and the jet veto efficiency as a function of the leading jet p_{T}(right). With the jet p_{T} threshold 30 GeV, the agreement between data and MC is very good. Figure 6.7 shows fraction of events with 0-jet and 1-jet as a function of the number of reconstructed vertices(N_{vtx}). Data and MC show very good agreement. The dependence on N_{vtx} is small in both data and simulation. Thus, with the chosen jet identification, the jet counting efficiency is not affected by the pileup correction.

Chapter 7

Background Estimation

Even after applying WW selection and the m_{H}-dependent selection to suppress backgrounds, there are events that survive these selections. In order to extract the signal component from data, a precise estimation of these residual backgrounds along with their uncertainty is essential.

The methods to estimate the contribution of each background depend on the process. The best way is to measure it using data control samples. This method is called the "data-driven method" which will be discussed in more detail in the next paragraph. For most of the main backgrounds, we do data-driven estimations using dedicated data control samples. For the rest of the process, we take it from simulation because they are small and well modelled by simulation(WZ/ZZ) or hard to measure in data($W\gamma$). Table 7.1 shows which background is estimated by the data-driven methods and which is taken from simulation.

Table 7.1: Method of background estimation for each background. Major backgrounds are estimated by data-driven methods, and $W\gamma$ and WZ/ZZ are taken from simulation.

Background	$qq \to W^+W^-$	$t\bar{t}/tW$	$Z/\gamma^* \to \ell\ell$	W + jets	$W\gamma^*$
Method	data	data	data	data	data

Background	$gg \to W^+W^-$	$W\gamma$	WZ/ZZ
Method	data	simulation	simulation

The basic idea of the data-driven method is the following :

- define a control region which is obtained by inverting or loosening a subset of the requirements in the final selection(WW or Higgs selection). The phase space selected by the final selection is called "signal region(SR)" and the inverted or loosened selection is called "control region(CR)".

- measure the ratio(ϵ) of the number of events passing the final selection ($N_{SR}^{independent}$) to the number of events passing the control region requirement ($N_{CR}^{independent}$) using a data sample independent of the data sample used

to extract the signal events. The independent sample can be either data or MC.

- apply the ratio(ϵ) to the number of events in the control sample passing the control region selection($N_{CR}^{control}$). The control sample is obtained by simply applying the control region selection to the data sample used to extract signal events.

This procedure can be written as an equation,

$$N_{SR} = \epsilon \times N_{CR}^{control}. \tag{7.1}$$

When measuring ϵ or $N_{CR}^{control}$, we need to subtract the contributions from other processes to measure only process we are interested in. This results in dependencies between estimations of the backgrouns processes. Figure 7.1 shows the dependency of the data-driven background estimation. The $Z/\gamma^* \to \ell\ell$ and W + jets do not depend on the other data-driven estimations. The estimation of $W\gamma^*$ depends on the estimation of W + jets. The estimation of $t\bar{t}/tW$ depends on $W\gamma^*$, $Z/\gamma^* \to \ell\ell$ and W + jets. The estimation of $qq \to W^+W^-$ depends on $W\gamma^*$, $Z/\gamma^* \to \ell\ell$, W + jets and $t\bar{t}/tW$. These dependencies natually decide in which order background processes need to be estimated. The $Z/\gamma^* \to \ell\ell$ and W + jets are done first, and $W\gamma^*$, $t\bar{t}/tW$ and $qq \to W^+W^-$ follow. This chapter discusses the details following the above order.

7.1 $Z/\gamma^* \to \ell\ell$

The $Z/\gamma^* \to \ell\ell$ is one of the main backgrounds in the $ee/\mu\mu$ category. The $Z/\gamma^* \to \ell\ell$ events do not have intrinsic MET. But, any mis-measurements can lead to a fake(instrumental) MET, of which the mis-measurement of jet momentum dominates. The expected contribution of the $Z/\gamma^* \to \ell\ell$ in the signal region is estimated by counting events near the Z mass ($|m_{\ell\ell} - m_Z| < 7.5$ GeV [1]) in data, subtracting the contributions from other processes, and scaling it by the

[1] Tighter mass window than Z veto requirement in the WW selection is chosen to reduce other background contributions

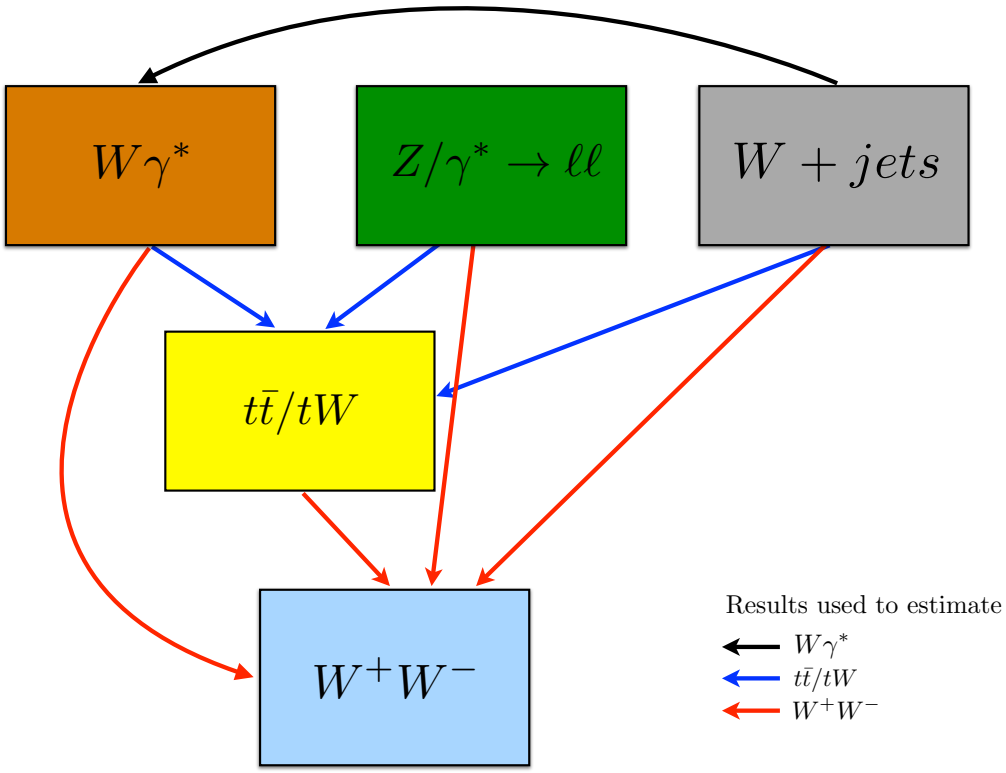


Figure 7.1: Dependency of the data-driven estimations.

ratio($R_{out/in}$) which is defined as the ratio of events outside of Z peak to the events inside of Z peak with looser BDT score. In equation, this can be written by

$$N(ll)^{DY}_{SR} = \left(N(ll)^{data}_{CR} - 0.5 \times N(e\mu)^{data}_{CR} \times k_{ll} - N^{WZ/ZZ,MC}_{CR}\right) \times R(ll)_{out/in} \quad (7.2)$$

where

- ll : ee or $\mu\mu$
- $N(ll)^{DY}_{SR}$: expected $Z/\gamma^* \to \ell\ell$ contribution in the signal region(SR)
- $N(ll)^{data}_{CR}$: number of ll data events in the control region(CR) which is under the Z peak ($|m_{\ell\ell} - m_Z| < 7.5$ GeV)
- $N(e\mu)^{data}_{CR}$: number of $e\mu$ data events in the control region(CR) which is under the Z peak ($|m_{\ell\ell} - m_Z| < 7.5$ GeV)

- k_{ll} : efficiency correction for $e\mu$ events,
  $k_{ee} = \sqrt{\frac{N(ee)^{loose}_{in}}{N(\mu\mu)^{loose}_{in}}}$ and $k_{\mu\mu} = \sqrt{\frac{N(\mu\mu)^{loose}_{in}}{N(ee)^{loose}_{in}}}$, measured with loose selection $(\mathrm{min-proj-MET} > 20\ \mathrm{GeV})$
- $N^{WZ/ZZ,MC}_{CR}$: contribution from WZ/ZZ events estimated in MC
- $R(ll)_{out/in}$: $R_{out/in}$ value measured in data with looser BDT score

The number of events in the CR($N(ll)^{data}_{CR}$) is corrected by subtracting WZ/ZZ and $e\mu$ contribution($N(e\mu)^{data}_{CR}$) which is dominated by $t\bar{t}/tW$, with an efficiency correction between $ee/\mu\mu$ and $e\mu/\mu e$. The lepton efficiency does not depend on the BDT score, so we can measure it without the BDT score requirement to get more statistics. The statistical uncertainty on the lepton efficiency measured with the events under the Z peak is much smaller(< 1 %) than the dominant systematic uncertainty. The WZ/ZZ contribution is taken from MC because they have real MET, and real MET is well-modelled by simulation. For the WZ/ZZ contribution in the CR, we assign 10 % of systematic uncertainty.

The $R_{out/in}$ is measured in data with a looser BDT score. The assumption is that the $R_{out/in}$ does not change as a function of the BDT score, which is expected because the BDT score is dominantly dependent on MET, and MET in $Z/\gamma^* \to \ell\ell$ events mostly comes from the jet momentum mis-measurement, which is weakly correlated to the lepton energy/momentum from which $m_{\ell\ell}$ is calculated. So, we divide the BDT score into 4 bins ([-0.9,-0.85], [-0.85,-0.60], [-0.60,WP], [WP,1.0]) where WP is 0.88 for 0-jet and 0.84 for 1-jet, and measure the $R_{out/in}$ using the bin closest to the signal region in order to get the sample with similar kinematics as the signal region events. Figure 7.2 shows $R_{out/in}$ divided in 4 bins for 0-jet and 1-jet for m_{H}= 125 GeV analysis. The fourth bin, located after the vertical dotted lines, corresponds to the signal region. The results using data subtracting WZ/ZZ component, and using MC are drawn as black and red, respectively. The results show that $R_{out/in}$ is almost flat as a function of the BDT score.

The systematic uncertainty comes from lack of "flatness" of $R_{out/in}$, the largest difference between the third bin and the other bins. In addition, there is an alternate method to estimate $Z/\gamma^* \to \ell\ell$ using the γ + jets data sample. The

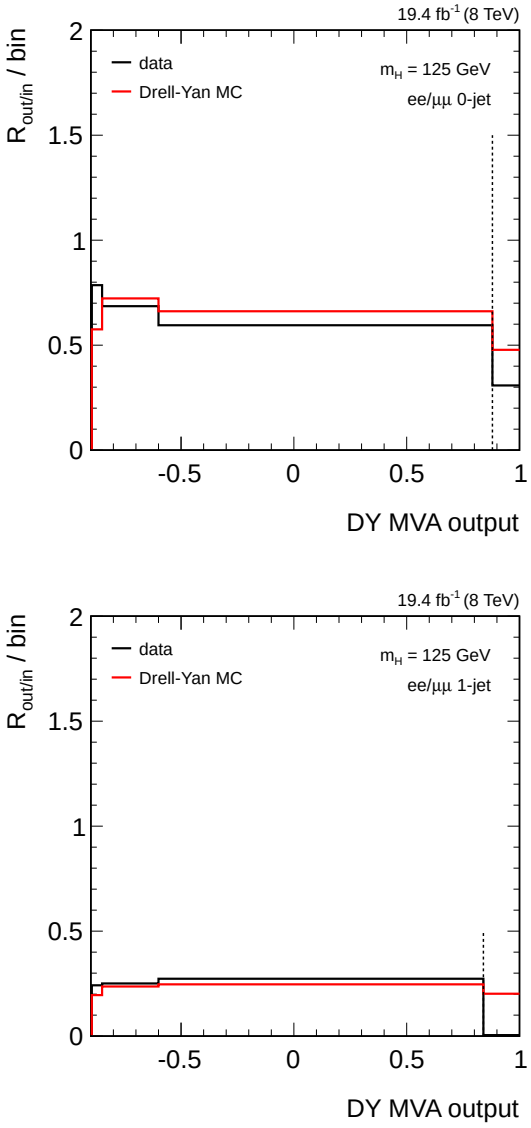


Figure 7.2: $R_{out/in}$ values as a function of BDT score for 0-jet and 1-jet categories for m_{H}=125 GeV analysis. The black is from data subtracting WZ/ZZ contribution and the red is from MC. The vertical dotted lines show the cut for signal region.

difference between this method and $R_{out/in}$ method in the extrapolation from CR to SR is about 30%. We take the maximum of the lack of flatness and the 30 % as the final systematics of the estimation.

At the end we merge ee and $\mu\mu$ to get more statistics. The table 7.2 shows the final estimation at WW level. The table 7.3 shows the final estima-

Table 7.2: The Drell-Yan estimation in the same flavor final state at WW selection level.

N_{jet}	N_{in}(data)	$R_{out/in}$	N_{out}(data)	N_{out} (MC)
0	775.16 ± 74.47	$0.28 \pm 0.01 \pm 0.08$	$218.64 \pm 21.53 \pm 65.59$	34.64 ± 9.56
1	350.69 ± 38.00	$0.25 \pm 0.01 \pm 0.08$	$89.20 \pm 9.83 \pm 26.76$	21.05 ± 7.16

tion of $Z/\gamma^* \to \ell\ell$ at all $m_{\rm H}$ hypotheses up to 300 GeV. After 300 GeV, the $N_{out}(data)/N_{out}(MC)$ scale factor for WW selection is applied.

7.2 Jet-induced backgrounds : W + jets

As discussed in chapter 4, jets can fake leptons(electron or muon) in various ways, as a result, W + jets is the second largest background in the most sensitive channel, 0-jet $e\mu/\mu e$. Most of the jet-induced background comes from W + jets where one jet fakes a lepton(single fake) and a small contribution comes from QCD events where two jets fake two leptons(double fakes). Even though the cross section of QCD process is huge, the probability for a jet to fake a lepton is very small, being in the order of $\mathcal{O}(10^{-3}) \sim \mathcal{O}(10^{-5})$ depending on the kinematics. In addition, QCD events do not have a source of true MET so the contribution is suppressed dramatically by the tight MET requirement. In the end, the contribution of the double fakes is negligible compared to the systematic uncertainty of the method, so it is not explicitly taken into account in this analysis.

The estimation of the jet-induced background starts from measuring the "fake rate(FR)" that is defined as the probability for a lepton with loose selection to pass the full selection. The leptons that pass the loose selection are called "fakable object(FO)". The fake rate is measured in the data events that is dominated by QCD di-jet events collected by pre-scaled low $p_{\rm T}$ single-lepton triggers. The measured FR is applied to the data sample where the lepton selection of one of the leptons is loosened. The details will be discussed in the following subsections.

Table 7.3: The Drell-Yan estimation in $e\mu/\mu e$ channel for the cut-based selections. The dependence of $R_{out/in}$ on m_{H} is due to the m_{H}-dependent selection listed in Table 5.1.

0-jet				
mass	N_{in}(data)	$R_{out/in}$	N_{out}(data)	N_{out} (MC)
115 GeV	145.94 ± 15.08	$0.31 \pm 0.01 \pm 0.09$	$45.07 \pm 4.81 \pm 13.52$	8.70 ± 5.08
120 GeV	263.31 ± 20.83	$0.31 \pm 0.01 \pm 0.09$	$81.33 \pm 6.80 \pm 24.40$	13.57 ± 5.88
125 GeV	154.60 ± 16.13	$0.60 \pm 0.02 \pm 0.19$	$92.22 \pm 9.98 \pm 29.33$	16.64 ± 6.63
130 GeV	119.10 ± 14.17	$0.87 \pm 0.03 \pm 0.26$	$103.61 \pm 12.74 \pm 31.08$	16.64 ± 6.63
135 GeV	112.30 ± 14.57	$0.83 \pm 0.03 \pm 0.25$	$93.52 \pm 12.49 \pm 28.06$	14.53 ± 6.29
140 GeV	108.72 ± 14.48	$0.74 \pm 0.02 \pm 0.22$	$80.72 \pm 11.09 \pm 24.22$	17.17 ± 6.82
150 GeV	91.78 ± 14.57	$0.40 \pm 0.02 \pm 0.12$	$37.01 \pm 6.20 \pm 11.10$	7.60 ± 4.47
160 GeV	21.03 ± 8.64	$0.90 \pm 0.06 \pm 0.27$	$18.83 \pm 7.85 \pm 5.65$	7.60 ± 4.47
170 GeV	9.70 ± 8.14	$0.82 \pm 0.06 \pm 0.25$	$7.92 \pm 6.68 \pm 2.38$	7.60 ± 4.47
180 GeV	7.06 ± 9.35	$0.62 \pm 0.05 \pm 0.19$	$4.34 \pm 5.76 \pm 1.34$	5.71 ± 4.05
190 GeV	64.23 ± 15.78	$0.34 \pm 0.02 \pm 0.10$	$21.63 \pm 5.51 \pm 6.49$	10.31 ± 5.20
200 GeV	94.35 ± 21.88	$0.21 \pm 0.01 \pm 0.06$	$20.10 \pm 4.83 \pm 6.03$	7.67 ± 4.47
250 GeV	193.85 ± 36.84	$0.06 \pm 0.00 \pm 0.02$	$11.17 \pm 2.24 \pm 3.35$	7.08 ± 4.09
300 GeV	89.23 ± 27.97	$0.13 \pm 0.01 \pm 0.04$	$11.16 \pm 3.62 \pm 3.35$	7.08 ± 4.09
1-jet				
mass	N_{in}(data)	$R_{out/in}$	N_{out}(data)	N_{out} (MC)
115 GeV	28.00 ± 8.66	$0.19 \pm 0.00 \pm 0.06$	$5.28 \pm 1.64 \pm 1.58$	2.16 ± 2.16
120 GeV	71.36 ± 12.41	$0.19 \pm 0.00 \pm 0.06$	$13.45 \pm 2.36 \pm 4.04$	4.35 ± 3.08
125 GeV	53.67 ± 10.57	$0.27 \pm 0.01 \pm 0.08$	$14.69 \pm 2.92 \pm 4.41$	4.35 ± 3.08
130 GeV	39.38 ± 9.49	$0.36 \pm 0.01 \pm 0.11$	$14.21 \pm 3.44 \pm 4.26$	4.35 ± 3.08
135 GeV	45.07 ± 9.97	$0.34 \pm 0.01 \pm 0.10$	$15.32 \pm 3.41 \pm 4.60$	4.35 ± 3.08
140 GeV	46.46 ± 10.18	$0.31 \pm 0.01 \pm 0.09$	$14.32 \pm 3.16 \pm 4.30$	4.35 ± 3.08
150 GeV	59.54 ± 12.09	$0.20 \pm 0.01 \pm 0.06$	$11.79 \pm 2.43 \pm 3.54$	2.19 ± 2.19
160 GeV	20.90 ± 6.94	$0.42 \pm 0.02 \pm 0.13$	$8.80 \pm 2.95 \pm 2.64$	0.00 ± 0.00
170 GeV	16.48 ± 6.89	$0.39 \pm 0.02 \pm 0.12$	$6.49 \pm 2.73 \pm 1.95$	0.00 ± 0.00
180 GeV	22.86 ± 8.18	$0.33 \pm 0.01 \pm 0.10$	$7.51 \pm 2.71 \pm 2.25$	0.00 ± 0.00
190 GeV	70.87 ± 13.43	$0.22 \pm 0.01 \pm 0.07$	$15.71 \pm 3.04 \pm 4.71$	0.00 ± 0.00
200 GeV	99.31 ± 16.50	$0.17 \pm 0.01 \pm 0.05$	$16.60 \pm 2.83 \pm 4.98$	0.00 ± 0.00
250 GeV	128.66 ± 21.79	$0.09 \pm 0.00 \pm 0.03$	$12.08 \pm 2.11 \pm 3.63$	2.65 ± 2.65
300 GeV	71.21 ± 17.81	$0.11 \pm 0.01 \pm 0.03$	$7.94 \pm 2.03 \pm 2.38$	5.00 ± 3.54

Definition of fakable objects

The definition of FO is limited by the analysis triggers used to collect the signal events. We can not define FO to be looser than the trigger requirement because FR can be biased due to tighter FO selection caused by the trigger requirement. Therefore, the loosest selection that we can afford is the trigger requirement. For electron, we use the following definition which is basically an offline version of the trigger selection with conversion rejection added :

- $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.5$
- $\sigma_{i\eta i\eta} < 0.01/0.03$ (barrel/endcap)
- $|\Delta\phi_{in}| < 0.15/0.10$ (barrel/endcap)
- $|\Delta\eta_{in}| < 0.007/0.009$ (barrel/endcap)
- H/E $< 0.12/0.10$ (barrel/endcap)
- $\dfrac{\sum_{\text{tracks with dR<0.3}} E_{\mathrm{T}}}{p_{\mathrm{T}}} < 0.2$
- $\dfrac{\left(\sum_{\text{ECAL with dR<0.3}} E_{\mathrm{T}}\right) - 1}{p_{\mathrm{T}}} < 0.2$
- $\dfrac{\sum_{\text{HCAL with dR<0.3}} E_{\mathrm{T}}}{p_{\mathrm{T}}} < 0.2$
- conversion rejection

The background caused by photon conversions can have different FR, so it is not estimated by this method. By excluding the contribution from the conversion rejection, we construct a sample of FO with similar sources.

The definition of muon FO is given by relaxing the transverse impact parameter and the isolation requirements :

- $|d_0| < 0.2$ cm
- Isolation MVA output > -0.6

The definition for muon is simpler because of the loose trigger requirement.

Measurement of fake rate

The FR is measured in data dominated by QCD di-jet events. These events are collected by single-lepton triggers listed in Table **??**. The events selected by these triggers are dominated by leptons that originate from quarks and gluons.

Though the sample is dominated by QCD events, there are contributions from Electroweak(EWK) processes such as $\mathrm{W}+\mathrm{jets}$ and $Z/\gamma^* \to \ell\ell$. In order to suppress $\mathrm{W}+\mathrm{jets}$ contribution, MET > 20 GeV and $m_{\mathrm{T}} < 15$ GeV are required. The residual contribution is estimated using simulation normalized by a data/MC scale factor calculated using the bulk of $\mathrm{W}+\mathrm{jets}$ events defined as MET > 30 GeV, $60 < m_{\mathrm{T}} < 90$ GeV and the full lepton selection. The deviation of scale factor from unity is taken as a systematic uncertainty for FR measurement due to the $\mathrm{W}+\mathrm{jets}$ contribution. The $Z/\gamma^* \to \ell\ell$ contribution is suppressed by rejecting events that contain an additional lepton that makes an opposite-sign lepton pair with di-lepton mass within 15 GeV of the Z mass. The residual contribution is estimated using simulation normalized by a data/MC scale factor calculated using events within 15 GeV of the Z mass. The deviation of scale factors from unity is taken as a systematic uncertainty for FR measurement due to the $Z/\gamma^* \to \ell\ell$ contribution .

Another EWK contribution comes from di-electron events when the second electron is outside of the tracker acceptance. This electron is not reconstructed as an electron because it does not have a track, but can be reconstructed as a jet. The result is an event with one lepton and low MET, which can fall into the FO selection. In order to suppress this, the $|\eta|$ of the jet that recoils against FO is required to be less than 2.5. For di-muon events this does not happen because if a muon is outside of tracker acceptance, it is not reconstructed as muon, giving a large MET in the event. This event is not selected for FO because of high MET.

Figure 7.3 shows the p_{T} distributions of the denominator and the numerator of electron and muon FO. The data is shown with black dots and the EWK contributions from $Z/\gamma^* \to \ell\ell$ and $\mathrm{W}+\mathrm{jets}$ are shown in green and grey, respectively. At higher p_{T} the relative EWK contribution is larger. Because FR is higher in $\mathrm{W}+\mathrm{jets}$ events than the QCD di-jet events, if EWK contribution were not sub-

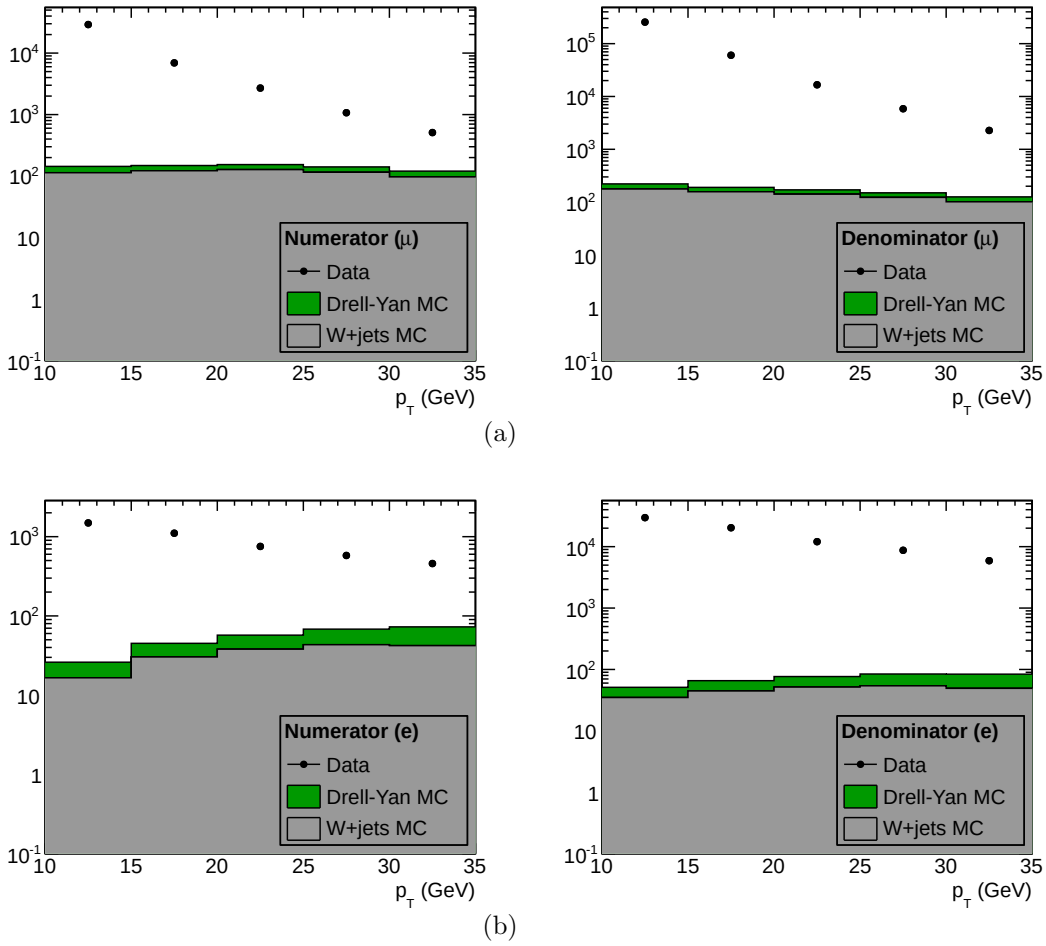


Figure 7.3: The p_{T} of denominator(left) and numerator(right) of (a) muon FO and (b) electron FO. The black dots represent data and solid histograms represent contributions from Drell-Yan(green) and W+jets(grey).

tracted, the FR would be measured higher than it should be because FR of EWK sample is higher than the FR of QCD di-jet sample. The subtraction of the EWK contribution reduces FR in the last p_{T} bins by up to 30 %.

A fake lepton originates from a parton, and the fake rates have strong correlation with the p_{T} of the parton from which a fake lepton originates. So, it is very important to use similar kinematic region when measuring FR with that of the region where the FR is applied. The best handle to select a relevant kinematic region is the p_{T} of the progenitor parton, but this information is not available in data. So, we use the p_{T} of the jet which is separated by ΔR(FO, leading jet) > 1.0 as an approximate handle of the progenitor p_{T}. This jet is called "away jet". Given that the FR is measured with the sample dominated by QCD di-jet events, the FO and the leading jet tend to be back-to-back in the $r - \phi$ plane. Therefore, by gauging the p_{T} of the away jet, we can gauge the p_{T} of the progenitor of the FO. Ideally this works, but reality is more complicated because the p_{T} of the away jet does not necessarily represent the p_{T} of the progenitor of the FO. The away jet p_{T} can be mis-measured and the direction of the jet can be off back-to-back direction. Therefore, it is not straightforward to decide the p_{T} of the away jet.

We use same-sign events to make sure that we select the relevant kinematic range of the progenitor for the W + jets sample. These events are dominated by W + jets and $W\gamma(^*)$. Since we can control $W\gamma(^*)$ relatively well, we can use these events to confirm that the measured FR predicts the data correctly. But, even though the prediction is consistent with observation, we can not conclusively claim that the measured FR is correct because the composition of the fakes can be different in the same-sign and the opposite-sign events. For example, same-sign events can not originate from W+q events, which result in an opposite-sign lepton pair[2]. Therefore, the result using the same-sign events can give a rough idea, but we can not conclusively decide the away jet $p_{\mathrm{T}}$from them.

The away jet p_{T} thresholds are determined based on multiple information. The first information is the comparison in simulation of the p_{T} spectrum of the

[2]If W has (+1) charge, q has (-1/3) charge which makes up a (-1) charge meson, and it subsequently decays to a (-) charge lepton. The ratio of the opposite-sign to the same-sign events measured in W + jets MC is 6:4

jets in the W + jets sample and the jets in the QCD sample that envelop the FO. The other information is the comparison of the isolation variable of FO in the QCD events and the W + jets events in data. These show that the best match is attained when using 35 GeV for electrons and 30 GeV for muons. Figure 7.4 shows the FR for muons and electrons using these thresholds. The expected same-sign events in the 0-jet category is $794 \pm 18(stat.)$ while the number of observed data events is 752 which is consistent with the expectation.

Application of fake rate

The measured FR as a function of p_{T} and η of the FO is used to predict the contribution of W + jets in the signal region. In data we first select events in which one lepton passes the full lepton selection while the other lepton passes the FO selection but not the full selection(Tight+Loose sample). The contribution from other background processes are subtracted in order to get a pure W + jets sample. Figure 7.5 shows the p_{T} distribution of the muon and electron FO on left and right, respectively. The plots are normalized to the unit area. The data with other backgrounds subtracted and the W + jets simulation show good agreement in shape giving confidence that the selected data sample is dominated by W + jets process. Now we use the measured FR as a weight to the Tight+Loose sample in the following way.

$$N_{\mathrm{W+jets}}^{\mathrm{prediction}} = \sum_{i}^{N_{\mathrm{Tight+Loose}}} \frac{\mathrm{FR}(p_{\mathrm{T}i}, \eta_i)}{1 - \mathrm{FR}(p_{\mathrm{T}i}, \eta_i)} \tag{7.3}$$

where $N_{\mathrm{W+jets}}^{\mathrm{prediction}}$ is the prediction of W + jets background in the signal region, $N_{\mathrm{Tight+Loose}}$ is the number of Tight+Loose events and $p_{\mathrm{T}i}$ and η_i are the p_{T} and η of the Loose lepton in the i^{th} Tight+Loose event.

Systematics

As mentioned before, the away jet p_{T} is a handle for the QCD di-jet events to give relevant kinematic range for FO in W + jets. But, it has a limited control of the kinematics of the parton from which the FO originates. We, therefore,

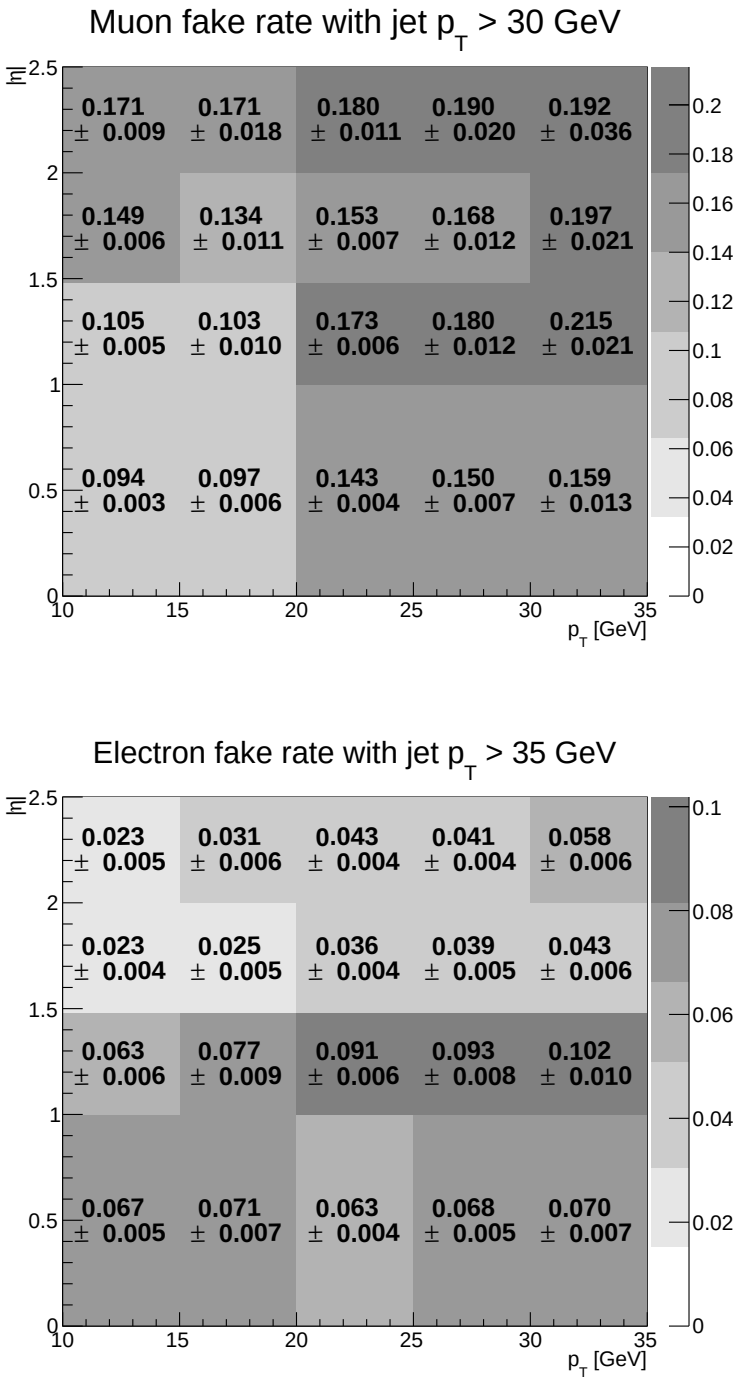


Figure 7.4: Top shows muon fake rate as a function of FO p_{T} and η, measured with the away jet $p_{\mathrm{T}} > 30$ GeV. Bottom shows electron fake rate as a function of FO p_{T} and η, measured with the away jet $p_{\mathrm{T}} > 35$ GeV. The errors are only statistical.

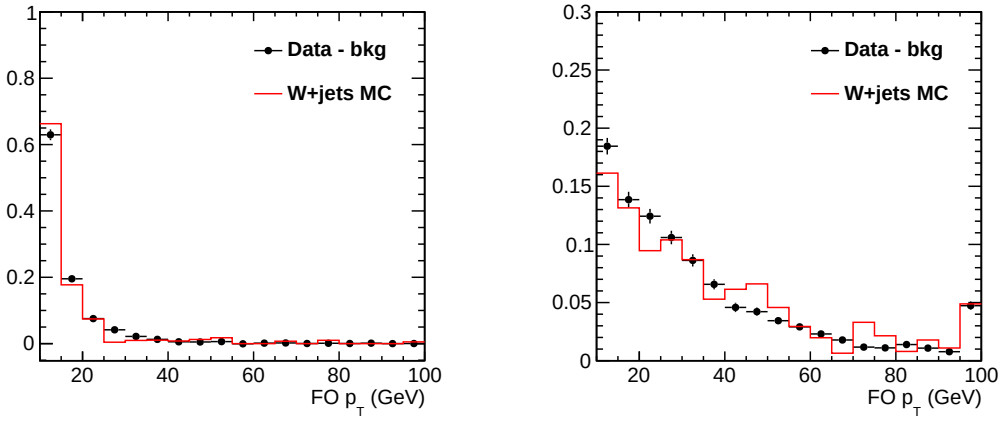


Figure 7.5: The p_{T} of muon FO (left)and electron FO(right). Plots are normalized to the unit area. On each plot, the shape of data after subtracting other backgrounds, and that of W + jets MC are compared.

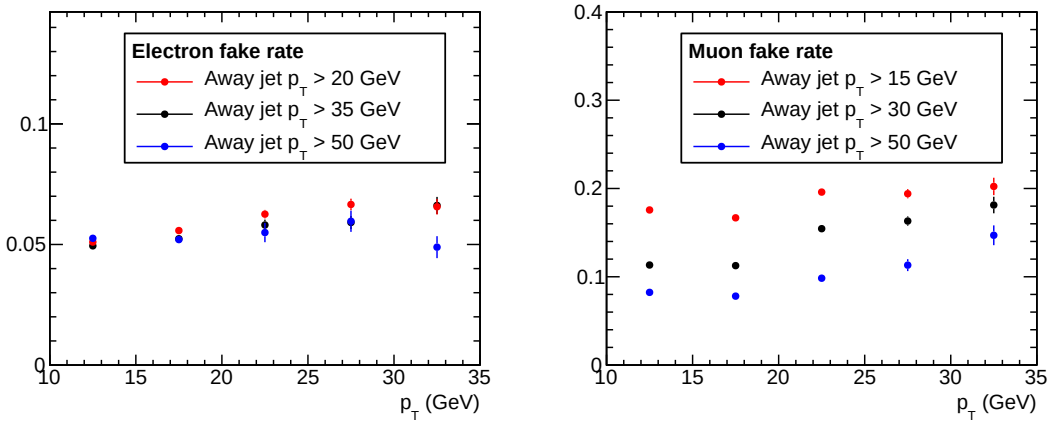


Figure 7.6: Left plot is the electron fake rates projected on p_{T} for different away jet p_{T} thresholds, 20, 35(default) 50 GeV. Right plot is the muon fake rates projected on p_{T} for different away jet p_{T} thresholds, 15, 30(default) and 50 GeV.

assign a systematic uncertainty to account for our limitation of controlling the progenitor p_{T} by varying the away jet p_{T} threshold to cover the relevant kinematic range of the W + jets sample. Figure 7.6 shows the dependence of FR on the away jet p_{T} thresholds. Left plots shows the dependence of electron FR, and the right plot shows the dependence of muon FR. The alternative away jet p_{T} thresholds are 20 and 50 GeV for electrons and 15 and 50 GeV for muons. The prediction using the alternative p_{T} thresholds differ by 30 % compared to the nominal results when electrons and muons are combined, and this difference is taken as systematic uncertainty. The dependence on the jet p_{T} is larger for muons than electrons. This is because of difference in the FO definition. The electron FR is a dominated by ID requirement which has small dependence on the away jet p_{T}, while the muon FR is dominated by ISO requirement which has strong dependence on the away jet p_{T}. In addition, the composition of the sample can be different in the QCD di-jet sample and the W + jets sample. So, we do a closure test in simulation assuming that the parton composition is well-modelled by simulation. The result shows that the difference between the prediction and the observation is $\sim$ 20 % for both electrons and muons. The two systematic uncertainties are combined in quadrature, resulting 36 % total systematic uncertainty for both electrons and muons.

7.3 $t\bar{t}/tW$

The backgrounds induced by top quarks come from $t\bar{t}/tW$ processes where the W boson decayed from a top quark decays leptonically. These processes result in two isolated leptons with considerable MET, which is the signature of signal events, and b-tagged jets which can be used to suppress this background. However, in case the b-jet goes out of the tracker coverage or it is not tagged as a b-jet, these events can survive the signal selection. To estimate its contribution, We use a data-driven method applied to 0-jet and 1-jet separately.

The top background is estimated after the WW selection and a common data/simulation scale factor for $t\bar{t}/tW$ is calculated in 0-jet and 1-jet categories.

Then, the simulation is normalized by the measured scale factor. The basic idea of the top estimation is to re-weight the top-tagged control region using the top-tagging efficiency obtained from an independent sample. This can be expressed as the following equation,

$$N_{\text{top-veto}} = N_{\text{top-tagged}} \times \frac{1 - \epsilon_{top-tagging}}{\epsilon_{top-tagging}} \tag{7.4}$$

where $N_{\text{top-veto}}$ is the prediction for the top-vetoed events in the signal region, $N_{\text{top-tagged}}$ is the number of the top-tagged events which can be obtained by inverting the top-veto requirement, and $\epsilon_{top-tagging}$ is the top-tagging efficiency measured in an independent data sample. The top-tagging efficiency is measured in a different way in the 0-jet and the 1-jet categories. Table 7.4 shows the definition of the control region and the tagging efficiency in the 0-jet and the 1-jet categories.

Table 7.4: Summary of selection and control region definitions used in top estimation in the different jet bins. Table is taken from [15].

jet bin	signal region (top veto)	control region (top tag)	ϵ_{tag}	
			denominator	numerator
0	no soft muon, no low p_{T} b-jets	either a soft muon or a low p_{T} b-jet	1-jet bin, leading jet is btagged	denom. + top tag
1	no soft muon, no low/high p_{T} b-jets	leading jet is b-tagged	2-jet bin, sub-leading jet is btagged	denom. + top tag

0-jet category method

We start with measuring the "per-leg" tagging efficiency, which is the probability for a b quark to be tagged. In the $t\bar{t}$ events, there are two b quarks in the final states, so we can explore the second b quark after requiring that the first b quark be tagged. A caveat is that the tW contribution should be subtracted because those events have only one b quark in the final state, and the second jet is from FSR or ISR regardless of jet flavors. Some of the FSR or ISR jets can be b-tagged as well, and this makes tW indistinguishable from $t\bar{t}$. Therefore, this should be accounted for when calculating the "per-event" top-tagging efficiency.

For the denominator the "per-leg" efficiency is defined by requiring exactly one b-tagged jet with $p_{\mathrm{T}} > 30$ GeV. This requirement is to select a sample dominated by $t\bar{t}/tW$. Of the selected events, a subset of events containing at least one b-tagged jet with $10 < p_{\mathrm{T}} < 30$ GeV or one soft-muon becomes the numerator. The ratio of the yields in the numerator to the denominator after subtracting contributions from other backgrounds as well as tW is the "per-leg" top-tagging efficiency, $\epsilon^{data}_{per-leg}$.

The "per-event" top-tagging efficiency is then calculated by the following equation,

$$\begin{aligned}\epsilon^{data}_{\text{top-tag, per-event}} = & \left(1-\left(1-\epsilon^{data}_{per-leg}\right)^2\right)\left(f^{MC}_{t\bar{t}} + x\left(1-f^{MC}_{t\bar{t}}\right)\right) \\ & + \left(1-f^{MC}_{t\bar{t}}\right)(1-x)\,\epsilon^{data}_{per-leg}\end{aligned} \tag{7.5}$$

where the first term corresponds to the tagging efficiency for the events with two taggable legs, and the second term corresponds to the tagging efficiency for the events with one taggable jet. The $\epsilon^{data}_{\text{top-tag, per-event}}$ is the per-event top-tagging efficiency. The $f^{MC}_{t\bar{t}}$ is the fraction of $\mathrm{t\bar{t}}$ events with respect to the $\mathrm{t\bar{t}}$+tW events measured using MC at WW level requiring no jets. The x is the fraction of events that contain two taggable legs in tW events. It corresponds to the $\epsilon_{per-leg}$ measured in tW MC.

Finally, the control region in data is defined by inverting the top-veto requirement, *i.e.*, requiring soft-muon or b-tagged jets with $10 < p_{\mathrm{T}} < 30$ GeV. The contributions from other backgrounds are subtracted so that the measured top-tagging efficiency is applied to only $t\bar{t}/tW$ events. The left plot in Figure 7.7 shows the level of agreement between data and MC in the top control region for 0-jet category. They show good agreement, and we conclude that the data subtracted by other backgrounds will give $t\bar{t}/tW$ events with high purity. The prediction of $t\bar{t}/tW$ events in the signal region is calculated by

$$N^{top}_{\text{WW region}} = (N^{\text{data}}_{top-tag} - N^{\text{other bkg}}_{top-tag}) \times \frac{1-\epsilon^{\text{data}}_{\text{top-tag, per-event}}}{\epsilon^{\text{data}}_{\text{top-tag, per-event}}}. \tag{7.6}$$

where $N^{\text{data}}_{top-tag}$ is the data count in the top CR and $N^{\text{other bkg}}_{top-tag}$ is the contribution from other background process in the top CR.

The systematic uncertainty to the top estimation in the 0-jet category is dominated by the uncertainty to the parameter x. The uncertainties of $t\bar{t}$ and tW cross sections are about 7 % and 15 % [69], and the resulting uncertainty of x is 17 %.

1-jet category method

The method for measuring top-tagging efficiency for 1-jet category is based on the fact that the b-tagging efficiency of the most energetic jet in 1-jet and 2-jet categories is approximately same. The efficiencies measured in $t\bar{t}$ MC after WW selection is 66 % and 67 % in 1-jet and 2-jet categories, respectively. Therefore, we measure the b-tagging efficiency($\epsilon^{\text{data}}_{\text{top-tag, per-event}}$) of the most energetic jet in events containing 2-jets, and apply this to the control region in the 1-jet category. The definition of control region is different from an inversion of the top-veto requirement in this case because the b-tagging efficiency is correlated with existence of soft-muons in the jet. So, the control region is defined without soft-muon requirement, and this definition is consistently applied for the top-tagged region, top-vetoed region and the top-tagging efficiency measurement. The right plot in Figure 7.7 shows the level of agreement between data and MC in the top control region in the 1-jet category. They show a good agreement, and we conclude that the data subtracted by other backgrounds gives $t\bar{t}/tW$ events with high purity.

Then, the top contribution in WW region is calculated using the measured top-tagging efficiency($\epsilon^{\text{data}}_{\text{top-tag, per-event}}$) and the number of events in the top CR($N^{\text{data}}_{top-tag} - N^{\text{other bkg}}_{top-tag}$) as shown in the equation 7.6.

The systematic uncertainty to the top estimation in the 1-jet category is dominated by data statistics in the control region which accounts for about 2 %.

Result

Table 7.5 shows the result of the data-driven estimation for $t\bar{t}/tW$. The data/simulation scale factors are consistent to unity in both 0-jet and 1-jet categories. The uncertainty is large in the 0-jet category due to the large cross section uncertainty of the tW process. The calculated scale factors are applied to the MC

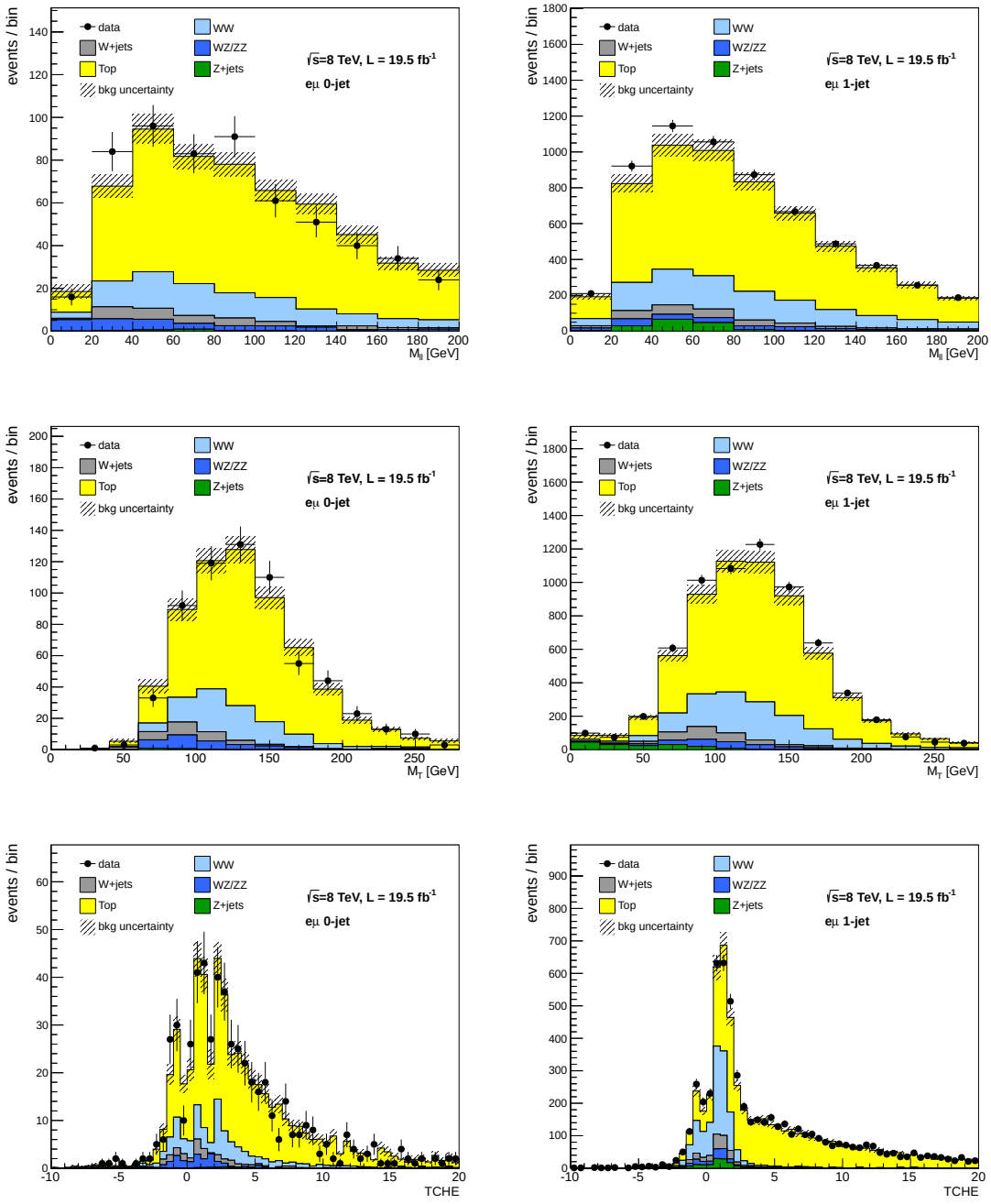


Figure 7.7: Distribution of b-tagging discriminator for most energetic jet with $p_{\mathrm{T}} > 10$ GeV in the event in the 0-jet(left) and 1-jet(right) in top control regions where top-tagging efficiency is measured.

counts in the signal region.

Table 7.5: Monte Carlo to data scale factor for the top background contribution for 19.5 fb^{-1}.

Sample	0-jet	1-jet
Estimated $t\bar{t}/tW$events in simulation	720.5 ± 4.3	2151.3 ± 13.9
Tagging efficiency (%)	49.3 ± 4.3	64.8 ± 0.5
Data events in control region	1034	4847
Background events in control region	292.6 ± 43.9	255.5 ± 51.1
$t\bar{t}/tW$estimation in data	761.4 ± 146.5	2307.9 ± 59.4
Data/simulation scale factor	1.06 ± 0.20	1.07 ± 0.03

7.4 *WW*

The non-resonant W^+W^- background is the main background in the most sensitive category(0-jet $e\mu/\mu e$)and estimated for 0-jet and 1-jet categories separately using the $m_{\ell\ell}$ distribution. Figure 7.8 shows the $m_{\ell\ell}$ distributions in 0-jet $e\mu/\mu e$ category for various $m_{\rm H}$ points, 125, 200 and 400 GeV, and WW after WW selection. For $m_{\rm H}$ hypotheses below 200 GeV, there is a negligible contribution from signal in the $m_{\ell\ell} > 100$ GeV region. Therefore, this region can be used as a WW control region. But, $m_{\rm H}$ hypotheses above 200 GeV give a significant contamination in that region, and it can not be used as a control region. So, we rely on simulation to estimate WW contribution in the signal region for the high $m_{\rm H}$ hypotheses($m_{\rm H} > 200$ GeV). For the estimation of WW background in the signal region for low $m_{\rm H}$ hypotheses($m_{\rm H} \leq 200$ GeV), we have different approaches for the shape-based and the cut-based methods.

For the shape-based method, we measure the data/simulation scale factor applying the template selection. In this method we expect that fit can constrain the WW component using the WW sideband region, *i.e.,* high $m_{\rm T}$ and/or high $m_{\ell\ell}$ region. Therefore, the measured scale factor is used to set the initial point for the

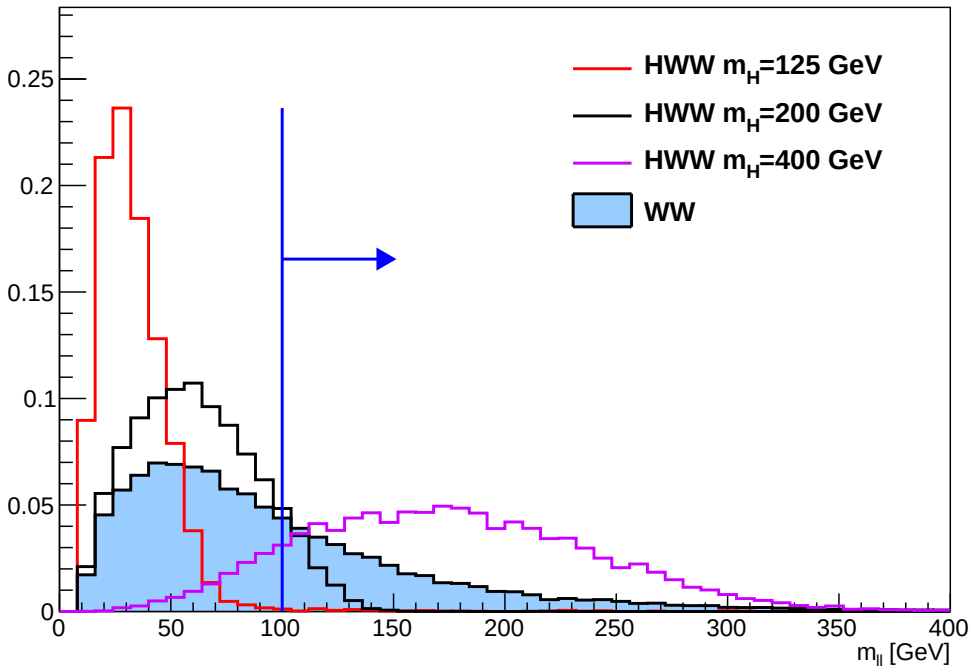


Figure 7.8: $m_{\ell\ell}$ distributions in 0-jet $e\mu/\mu e$ channel for signal events, $m_{\mathrm{H}} = 125$, 200 and 400 GeV and WW after WW selection. The control region, $m_{\ell\ell} > 100$ GeV, is marked with a blue line.

fit to start. The *pdf* for the WW normalization in the fit model is a flat function, so its value and the uncertainty are entirely determined by fit.

For the cut-based method, we define a control region in the $m_{\ell\ell}$ space where negligible signal events are expected, and extrapolate the control region yield to the signal region using the ratio of signal region to control region calculated in simulation($R_{\mathrm{SR/CR}}$). At each m_{H} hypothesis, the control region is defined with the p_{T} requirements on the leading and the trailing leptons on top of $m_{\ell\ell} > 100$ GeV.

The following equation summarizes how the estimation is done,

$$N_{\mathrm{WW}}^{\mathrm{SR}} = \left(N_{\mathrm{CR}}^{\mathrm{data}} - N_{\mathrm{CR}}^{\text{other bkg}}\right) \times R_{\mathrm{SR/CR}}^{\mathrm{MC}}. \tag{7.7}$$

We first measure the ratio($R_{\mathrm{SR/CR}}^{\mathrm{MC}}$) of events in the signal region to the control region using simulation. Then, we obtain the yield($N_{\mathrm{CR}}^{\mathrm{data}}$) in the control region($m_{\ell\ell} >$ 100 GeV) in data, and subtract the contribution from other backgrounds($N_{\mathrm{CR}}^{\text{other bkg}}$) such as $t\bar{t}/tW$, $\mathrm{W}+\mathrm{jets}$ and $Z/\gamma^* \to \ell\ell$. The data-driven estimation described in the previous sections are used for these backgrounds. Other small backgrounds are taken from simulation. We multiply the extrapolation factor to the data yield in the CR to get the prediction of WW in the signal region($N_{\mathrm{WW}}^{\mathrm{SR}}$). The same data/simulation scale factor is used for $qq \to W^+W^-$ and $gg \to W^+W^-$ to scale the prediction by simulation.

The uncertainty on the scale factor comes from the statistics in data control region and the systematic uncertainty of the other background in the control region. These two sources contribute almost equally to the final uncertainty. Table 7.6 shows the result of the data-driven estimation of WW background.

7.5 $W\gamma^*$

In the $W\gamma^*$ events, the leptons decayed from γ^* tend to have a small opening angle and a small invariant mass, and at least one of them has average p_{T} of 5 GeV. If the conversion happens early in the detector, *i.e.*, close to the interaction point, and most of the momentum of γ^* is carried by one lepton, the other lepton may not have enough energy to reach ECAL or muon station. In this case that soft

Table 7.6: The result WW background estimation for cut-based and shape-based analyses.

mass [GeV]	0-jet scale factor	1-jet scale factor
Cut-based		
115	1.10 ± 0.06	0.93 ± 0.10
120	1.10 ± 0.06	0.93 ± 0.10
125	1.10 ± 0.06	0.93 ± 0.10
130	1.10 ± 0.06	0.93 ± 0.10
135	1.10 ± 0.06	0.93 ± 0.09
140	1.10 ± 0.06	0.93 ± 0.09
150	1.08 ± 0.06	0.93 ± 0.10
160	1.08 ± 0.06	0.93 ± 0.10
170	1.07 ± 0.06	0.92 ± 0.10
180	1.07 ± 0.06	0.92 ± 0.10
190	1.07 ± 0.06	0.92 ± 0.10
200	1.07 ± 0.06	0.91 ± 0.10
Shape-based		
All masses	1.19 ± 0.06	1.09 ± 0.10

lepton can not be reconstructed, and the event contains two leptons(one from W and the other from γ^*) from interaction point and MET from leptonic W decay.

The WZ simulation is supposed to cover the $W\gamma^*$ as a part of it, but the low m_{γ^*} region is not covered appropriately due to the generator level cut $m_{\gamma^*} > 12$ GeV. The problem is that there is a significant contribution from that region of phase space all the way down to the production threshold, $m_{\gamma^*} = 2m_{e/\mu}$, to the signal region. So, we have a dedicated MC sample that was generated in LO to cover this region. Because it was generated in LO, it is important to measure the k-factor to obtain a correct normalization. The k-factor also would tell about the validation of the sample. For example, if k-factor is 10, this would indicate

that the sample is not correctly generated.

When measuring the k-factor, we consider separately the cases where the γ^* decays to a pair of electrons or muons. We use only $l^{\pm}\mu^+\mu^-$ final state because of the uncontrollable QCD backgrounds in the $l^{\pm}e^+e^-$ final state. We first define a control region dominated by the $W\gamma^*$ events. The following requirements are imposed to define the control region :

- The muon pair should have opposite charges. In case of $\mu^{\pm}\mu^+\mu^-$ final states, the pair with lowest $m_{\mu\mu}$ is considered coming from γ^*.
- The muon isolation is redefined such that it does not include muons in the isolation energy calculation in order to reconstruct γ^* from a muon pair very close to each other($\Delta R < 0.3$)
- To suppress $t\bar{t}/tW$, the number of jets should be less than 2, and events containing at least one b-tagged jets with $p_{\mathrm{T}} > 10$ GeV are excluded
- To suppress QCD background, MET > 25 GeV and $m_{\mathrm{T}}(W) > 45$ GeV are applied
- To reject $\mu\mu$ pairs from J/ψ, $\left|m_{\mu\mu} - m_{\mathrm{J}/\psi}\right| > 0.1$ GeV is applied
- To avoid interference with WZ^*, $m_{\mu\mu} < 12$ GeV is applied

The residual contributions of other background processes mostly come from W + jets, and it is estimated by the data-driven method described in section 7.2.

The measured k-factor is 1.5 which is consistent with the k-factors of other EWK processes. Figure 7.9 shows the di-muon mass distribution in the $W\gamma^*$ control region. The $W\gamma^*$ component is scaled up using the measured k-factor. There is a disagreement in the $m_{\mu\mu}$ shape, and this is due to the mis-modelling of reconstruction efficiency of the close-by muons at very low p_{T}. In order to take this into account, the k-factor is measured in two $m_{\mu\mu}$ regions, $m_{\mu\mu} < 2$ GeV and $2 < m_{\mu\mu} < 12$ GeV in $\mu^{\pm}\mu^+\mu^-$ and $e^{\pm}\mu^+\mu^-$ final states. The average spread of the four measurements is taken as a systematic uncertainty, and the resultant k-factor is 1.5 ± 0.5.

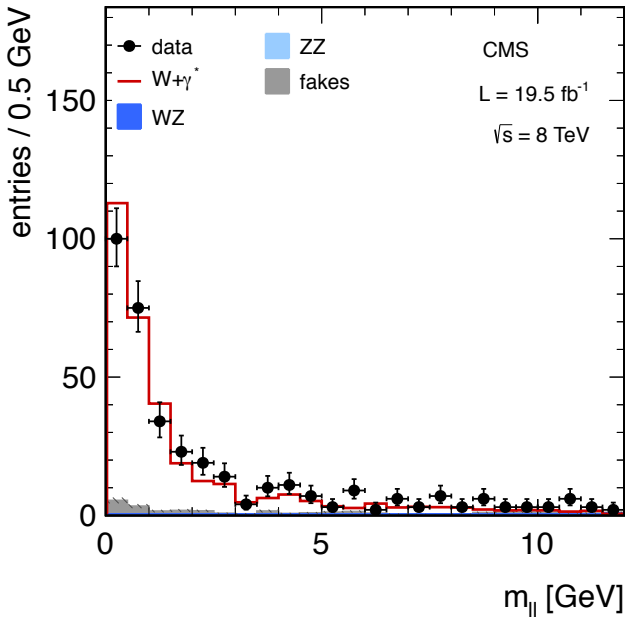


Figure 7.9: $m_{\ell\ell}$ distribution for opposite-sign muon pairs after $W\gamma^*$ selection. The $W\gamma^*$ is normalized to match the data.

7.6 Other backgrounds

7.6.1 $W\gamma$

$W\gamma$ can fake signal when the photon converts to a pair of electrons asymmetrically, giving a large fraction of its momentum to one of the electrons. This is difficult to be measured in data because it is hard to select events with asymmetric photon conversion due to fake electrons from W + jets. So, we rely on simulation after data corrections such as trigger/lepton selection efficiencies and pileup.

For the shape-based method, the templates for $W\gamma$ suffer from low statistics if the full lepton selection is applied to both leptons. So, we take the shape of the templates from $W\gamma$ events before the photon converts to leptons, and apply photon-to-electron conversion probability($P(\gamma \to e)$) as a function of η. The conversion factor is measured in simulation and shown in Figure 7.10. The function is parametrized with η because conversion probability has strong correlation with the amount of material that the photon goes through.

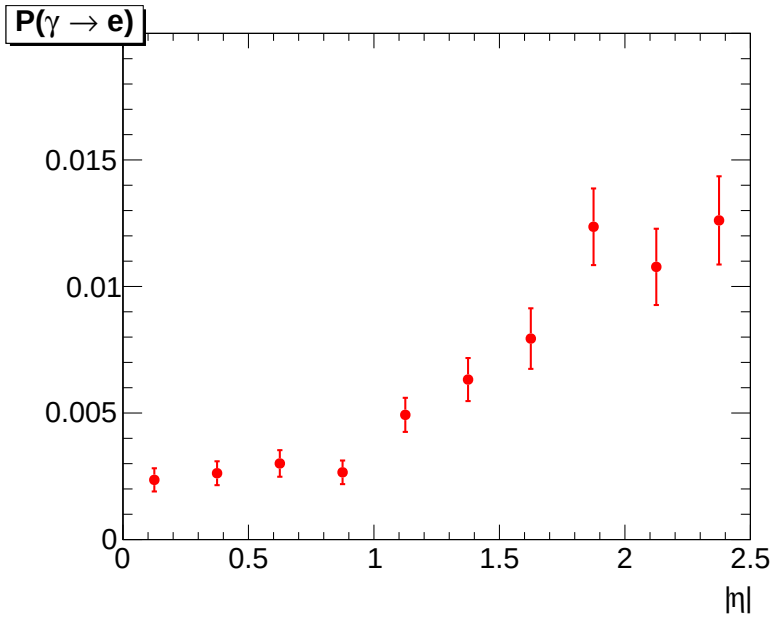


Figure 7.10: The probability for a photon to convert to a pair of electrons, and one of the electrons is selected as a good electron.

Figure 7.11 shows the normalized $m_{\ell\ell}$ and m_{T} distributions of $W\gamma$ samples, the red is the sample with 2 leptons in the final state and black is the sample with 1 lepton + γ where the conversion factor is applied. The two shapes look consistent. By using the 1 lepton + γ sample, we get an order of $\mathcal{O}(10^2)$ more statistics in the template.

7.6.2 $Z \to \tau\tau$

The $Z \to \tau\tau$ background contributes in the $e\mu/\mu e$ category in case where the taus decay to e or μ and neutrinos. In the $ee/\mu\mu$ category, it is negligible thanks to $\mathrm{proj} - \mathrm{MET}$ which is designed to reduce $Z \to \tau\tau$ background. Typically, the natural source of MET in $Z \to \tau\tau$ gives soft MET, and a modest MET selection can reduce this background. But, in the environment of large number of pileup, the instrumental contribution to MET becomes larger, and it leads to large MET values. The cross section of the $Z \to \tau\tau$ process is large, so we need to make sure that this background is under control. The simulation does not reproduce the instrumental MET properly, and we need a data-driven method to accurately

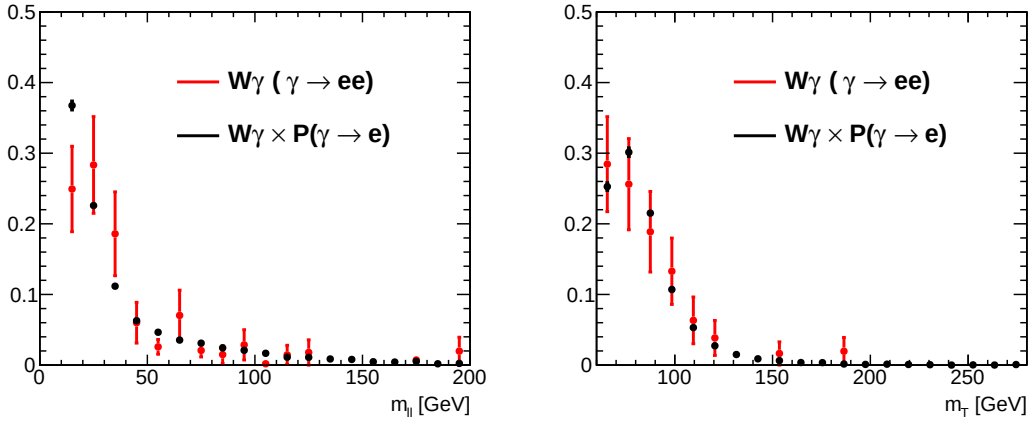


Figure 7.11: $m_{\ell\ell}$ and m_{T} distributions of $W\gamma$ samples. Red is after photon coverts to electrons and black is before the photon conversion with the conversion probability applied.

estimate the $Z \to \tau\tau$ background.

The data-driven method for the estimation of $Z \to \tau\tau$ background is done by the "tau embedding" technique [70]. In this method, we select $Z/\gamma^* \to \mu\mu$ events in data from the whole run range, and replace each muon with a simulated tau decay, $\tau \to l\bar{\nu}_l\nu_\tau$ [71]. Then, the MET is re-calculated adding the neutrinos from the tau decays. Advantage of this method is that the event environment is taken from the real data, and the modeling of pileup, UE, jets and thus MET can be done consistently with real data. This sample is normalized to the inclusive $Z \to \tau\tau$ MC at the level of requiring two leptons with lepton efficiency correction and the branching fraction.

A 10 % of systematic uncertainty of the method is assigned based on the MC closure test.

7.6.3 WZ/ZZ

WZ/ZZ backgrounds where the bosons decay leptonically are well-modelled by simulation. In addition, the contribution of these backgrounds is small. Therefore, to estimate these background, we rely on simulation after data corrections such as trigger/lepton selection efficiencies and pileup.

7.7 The result of background estimation

Table 7.8 summarizes the result of background estimation. It shows the yield of each process in the four final states, 0-jet $ee/\mu\mu$, 0-jet $e\mu/\mu e$, 1-jet $ee/\mu\mu$ and 1-jet $e\mu/\mu e$. Data and the total background yields are in good agreement. Actually, they should agree by design because backgrounds are normalized to data by data-driven estimation.

For comparison, Table 5.1 shows the yields with $m_{\mathrm{H}} = 125$ GeV Higgs selection shown in Table 5.1. The equivalent tables for other m_{H} and 7 TeV are Table 11.2 and 11.1, respectively. After Higgs selection we see that data is in better agreement with prediction when Higgs signal at m_{H}=125 GeV is added.

Table 7.7: Expected number of signal and background events from the data-driven methods for at 8 TeV after applying $\mathrm{W^+W^-}$ selection. Only statistical uncertainties are reported.

	0-jet		1-jet	
	$ee/\mu\mu$	$e\mu/\mu e$	$ee/\mu\mu$	$e\mu/\mu e$
$qq \to W^+W^-$	1782.4 ± 11.6	4402.1 ± 18.0	533.7 ± 6.0	1414.3 ± 9.7
$gg \to W^+W^-$	124.0 ± 2.2	223.7 ± 3.0	34.1 ± 1.1	76.3 ± 1.6
$t\bar{t}/tW$	273.3 ± 7.8	563.7 ± 11.1	677.2 ± 10.4	1644.2 ± 16.8
$W\gamma$	23.5 ± 5.7	140.0 ± 16.9	11.9 ± 4.8	51.8 ± 9.2
$W\gamma^*$	13.7 ± 2.0	147.9 ± 6.5	2.4 ± 0.8	23.9 ± 2.7
WZ/ZZ	56.6 ± 1.2	143.8 ± 1.7	39.7 ± 1.6	134.1 ± 2.0
W + jets	107.5 ± 4.6	694.8 ± 10.3	47.4 ± 3.6	356.2 ± 7.9
Z + jets	294.6 ± 60.3	82.5 ± 1.4	118.4 ± 30.3	248.7 ± 2.8
Total Background	2676.3 ± 62.5	6399.1 ± 30.1	1465.3 ± 33.3	3950.0 ± 23.7
$gg \to H$	83.87 ± 1.63	237.24 ± 2.67	25.68 ± 0.91	94.97 ± 1.71
$qq \to qqH$	0.87 ± 0.05	3.08 ± 0.09	3.78 ± 0.09	12.4 ± 0.17
Data	2728	6361	1477	3944

Table 7.8: Expected number of signal and background events from the data-driven methods for at 8 TeV after applying $m_{\mathrm{H}} = 125$ GeV Higgs selection shown in Table 5.1. Total uncertainties including both statistical and systematic uncertainties are reported.

	0-jet		1-jet	
	$ee/\mu\mu$	$e\mu/\mu e$	$ee/\mu\mu$	$e\mu/\mu e$
$qq \to W^+W^-$	197.2 ± 19.2	294.1 ± 28.4	37.5 ± 5.3	74.5 ± 10.4
$gg \to W^+W^-$	9.7 ± 3.1	16.0 ± 5.0	2.2 ± 0.8	4.8 ± 1.6
$t\bar{t}/tW$	9.3 ± 2.2	20.0 ± 4.3	40.4 ± 3.1	78.8 ± 4.5
$W\gamma$	3.2 ± 2.5	21.2 ± 9.8	2.5 ± 1.5	6.8 ± 4.0
$W\gamma^*$	6.1 ± 2.9	18.3 ± 8.1	0.8 ± 0.7	4.5 ± 2.3
WZ/ZZ	13.3 ± 1.3	10.2 ± 1.0	6.5 ± 1.0	10.2 ± 1.1
W + jets	28.6 ± 10.7	47.8 ± 17.6	6.5 ± 2.8	25.6 ± 9.7
Z + jets	92.2 ± 31.0	1.2 ± 0.2	14.7 ± 5.3	2.7 ± 0.4
Total Background	359.6 ± 37.6	428.8 ± 34.2	111.0 ± 8.6	208.0 ± 14.1
$gg \to H$	55.8 ± 12.2	88.9 ± 19.3	15.9 ± 5.2	37.5 ± 12.2
$qq \to qqH$	0.5 ± 0.1	1.0 ± 0.1	2.1 ± 0.2	4.6 ± 0.5
Data	421	505	140	228

This chapter is a reprint of the material as it appears in CMS Collaboration, "Measurement of Higgs boson production and properties in the WW decay channel with leptonic final states", J. High Energy Phys. 01 (2014) 096. The dissertation author was the primary investigator and author of this paper.

Chapter 8

Statistical Interpretation

In order to test a compatibility of observed data with a hypothesis, one performs statistical hypothsis tests. In the Higgs search or other searches in general, there are two hypotheses to be tested: a hypothsis that Higgs exists and a hypothesis that Higgs does not exist. For a quantitative analysis, we first construct a test statistic. It is a function of expected signal(s) and background yields(b), and the observed data. Then, probability density functions(*pdf*) for each hypothesis are constructed by pseudo-data or analytic functions. After that, the compatibility of data with each hypothesis is estimated in terms of p-value, the probability for the *pdf* to have the measurement at observed data or greater.

One complexity to consider is that the expected yields, s and b, are subject to change by many sources discussed in chapter 9. These systematic uncertainties are incorporated to the likelihood by adding *pdf* for each systematic source. The *pdf* is constructed by re-interpretating the *pdf*, $\rho(\theta|\tilde{\theta})$, using the Bayes' theorem

$$\rho(\theta|\tilde{\theta}) \sim p(\tilde{\theta}|\theta) \cdot \pi_\theta(\theta) \tag{8.1}$$

where $\pi_\theta(\theta)$ is the prior for which we use a flat function. This re-interpretation allows one to represent the systematic errors in a frequentist context [72].

This chapter discusses in detail the statistical procedure of the exclusion and the discovery of the SM Higgs boson.

8.1 Exclusion of SM Higgs boson

The procedure of constructing the likelihood and the probability density function which is agreed between ALTAS and CMS collaborations is described in detail in the reference [72].

The first step for the *pdf* construction is to define a likelihood function,

$$\mathcal{L}(X|\mu,\theta) = \prod_i^{N_{bin}} \frac{(\mu s_i(\theta) + b_i(\theta))^{X_i}}{X_i!} e^{-\mu s_i(\theta) - b_i(\theta)} \times \prod_j^{N_{nuisance}} p\left(\tilde{\theta}_j|\theta_j\right) \tag{8.2}$$

where $X = \{X_i\}$ is a set of measurements which can be the real data from measurement or a pseudo-data generated to construct the *pdf*s on a given hypothesis, N_{bin} is the number of measurements which corresponds to the number of bins or

categories, μ is the signal strength (signal strength modifier), θ is the nuisance parameter, $s_i(\theta)$ and $b_i(\theta)$ are the expected signal and background yields, respectively, in the i^{th} bin or category, $p\left(\tilde{\theta}_j|\theta_j\right)$ is the *pdf* for θ_j which is constructed from auxiliary measurements or some theoretical assumptions, and $\tilde{\theta}_j$ is the measured or assumed value of θ.

According to the Neyman-Pearson lemma [73], when performing a hypothesis test between two hypotheses, the ratio of two likelihoods constructed by the two hypotheses is the most powerful discriminator. Instead of the ratio itself, the log of the ratio is taken because of a number of advantages [1]. At the LHC, due to its asymptotic properties which is described in detail in the reference [74], the profile log-likelihood ratio (LLR) is used as a test statistic. The profile log-likelihood is constructed using values of nuisance parameters that maximize the likelihood function. By varying a nuisance parameter(θ_i) according to the constraint function($p\left(\tilde{\theta}_i|\theta_i\right)$) a new central value along with new uncertainty is determined and used for statistical interpretation. However, there can be a case where some nuisance parameters are over-constrained, *i.e.* becoming too small compared to the estimated value, because of high statistics in data. Therefore, it is important to examine the post-fit nuisance parameters and make sure that over-constraining some of them does not affect the final results. This issue is discussed in section **??**.

The functional form of the statistic is

$$\tilde{q}_\mu = \begin{cases} -2\ln\dfrac{\mathcal{L}(X|\mu,\hat{\theta}_\mu)}{\mathcal{L}(X|\hat{\mu},\hat{\theta})} & \text{if } 0 \le \hat{\mu} \le \mu \\ 0 & \text{otherwise} \end{cases} \tag{8.3}$$

where $\hat{\theta}_\mu$ is the best-fit value of θ for a given μ, $\hat{\mu}$ and $\hat{\theta}$ are the best-fit values of μ and θ, respectively, from the global fit on data. The requirement $0 \le \hat{\mu}$ is imposed because the signal rate must be positive. The requirement $\hat{\mu} \le \mu$ constrains μ to one-sided confidence level. This also means that the region, $\mu < \hat{\mu}$, is not considered more incompatible than the data observed, $\hat{\mu}$. This region is not tested for setting upper limits.

[1] Log converts multiplication of likelihoods into linear summation. Terms in exponent becomes a multiplication factor.

Then, we construct the *pdf*s based on the two hypotheses we want to test, the *signal + background* and the *background-only* hypotheses. We first find the observed(profiled) values of the nuisance parameters($\hat{\theta}_{\mu}^{obs}$ and $\hat{\theta}_{0}^{obs}$), and calculate the values of the test statistic($\tilde{q}_{\mu}^{obs}$ and $\tilde{q}_{0}^{obs}$) assuming the two hypotheses. Using the profiled nuisance parameters, $\hat{\theta}_{\mu}^{obs}$ and $\hat{\theta}_{0}^{obs}$, Monte-Carlo toys are generated to construct *pdf*s for *signal + background* and *background-only* hypotheses, $f\left(\hat{q}_{\mu}|\mu,\hat{\theta}_{\mu}^{obs}\right)$ and $f\left(\hat{q}_{\mu}|0,\hat{\theta}_{0}^{obs}\right)$, respectively. Since generating toys requires a large consumption of CPU power, we can take an advantage of the fact that $\tilde{q}_{\mu}$ follows a well-defined formula in the asymptotic limit [74]. The $\tilde{q}_{\mu}$ can be calculated analytically

$$f\left(\tilde{q}_{\mu}|\mu\right)=\frac{1}{2}\delta\left(\tilde{q}_{\mu}\right)+\begin{cases}\frac{1}{2\sqrt{2\pi}}\frac{1}{\sqrt{\tilde{q}_{\mu}}}e^{-\tilde{q}_{\mu}/2} & 0<\tilde{q}_{\mu}\leq\mu^{2}/\sigma^{2}\\ \frac{\sigma}{2\sqrt{2\pi}\mu}\exp\left[-\frac{1}{2}\frac{(\tilde{q}_{\mu}^{2}+\mu^{2}/\sigma^{2})^{2}}{(2\mu/\sigma)^{2}}\right] & \tilde{q}_{\mu}>\mu^{2}/\sigma^{2}\end{cases} \tag{8.4}$$

where $\delta(\tilde{q}_{\mu})$ is the Dirac delta function and $\sigma^{2}=\dfrac{\mu^{2}}{q_{\mu,A}}$ is the uncertainty on the test statistic evaluated using the Asimov dataset [74]. The Asimov dataset is a representative dataset made with the expected nuisance parameters. In this analysis the expected sensitivity(exclusion and significance) is calculated in the asymptotic limit. Similarly, the *pdf* for the *background-only* hypothesis, $f\left(\tilde{q}_{\mu}|0\right)$, can be obtained using the same technique.

For the measurement X, the test static($\tilde{q}_{\mu}^{obs}$) can be used to test the significance of the observation. The LHC uses the CL_{s} method [75, 76] that was developed by LEP to mitigate the problem of excluding a model one is not sensitive to. The p-value of the testing parameter is penalized by the insensitivity to distinguish the two hypotheses. The $\mathrm{CL_s}$ is defined by two p-values, p_{μ} and $1-p_{b}$. The p_{μ} is the p-value with *signal + background* hypothesis, and defined by

$$p_{\mu} = P\left(\tilde{q}_{\mu}\geq\tilde{q}_{\mu}^{obs}|signal+background\right) \tag{8.5}$$

$$= \int_{\tilde{q}_{\mu}^{obs}}^{\infty}f\left(\tilde{q}_{\mu}|\mu\right)d\tilde{q}_{\mu}. \tag{8.6}$$

A large value of p_{μ} represents a high chance that observation is compatible with the hypothesized signal strength, μ. The $1-p_{b}$ is the p-value with *background-only*

hypothesis, and defined by

$$1 - p_b = P\left(\tilde{q}_\mu \geq \tilde{q}_\mu^{obs} | background - only\right) \quad (8.7)$$

$$= \int_{\tilde{q}_\mu^{obs}}^{\infty} f\left(\tilde{q}_\mu | 0\right) d\tilde{q}_\mu. \quad (8.8)$$

A large value of $1-p_b$ represents a high chance that observation is compatible with the *background-only* hypothesis, $\mu = 0$. Thus, if data is *signal + background*-like the p_b is small. The $\mathrm{CL_s}$ is defined as a ratio of the two p-values

$$\mathrm{CL_s}\left(\mu\right) = \frac{p_\mu}{1 - p_b}. \quad (8.9)$$

Now we set the limit on the signal strength to test the compatibility of data with the SM Higgs hypothesis. The upper limit on the μ at $\alpha\%$ confidence level is the value of μ which gives $\mathrm{CL_s} = 1 - \alpha\%$. Writing for μ, we have

$$\mu^{\alpha\ \%} = \mathrm{CL_s}^{-1}(1 - \alpha\%). \quad (8.10)$$

In this case, the *signal+background* hypothesis with $\mu > \mu^{\alpha\%}$ is regarded as incompatible with data, and excluded at $\alpha\%$ $\mathrm{CL_s}$ confidence level.

When the expected limit is evaluated, it is useful to know how $\mu^{\alpha\%}$ varies because even though the true μ (call it μ') is correct, the data we actually observe can have a different value by statistical fluctuation. In the asymptotic limit, the CL_s is given by [72]

$$CL_\mathrm{s} = 1 - \alpha\ \% = \frac{1 - \Phi\left(\sqrt{q_\mu}\right)}{\Phi\left(\sqrt{q_{\mu,A}} - \sqrt{q_\mu}\right)} \quad (8.11)$$

where Φ is a cumulative distribution of standard normal distribution. Using the relations, $\sqrt{q_\mu} = \frac{\mu - \hat{\mu}}{\sigma}$ and $\sqrt{q_{\mu,A}} = \frac{\mu}{\sigma}$ [74], we have

$$1 - \alpha\ \% = \frac{1 - \Phi\left(\frac{\mu - \hat{\mu}}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma} - \frac{\mu - \hat{\mu}}{\sigma}\right)} \quad (8.12)$$

$$= \frac{1 - \Phi\left(\frac{\mu}{\sigma} - \frac{\hat{\mu}}{\sigma}\right)}{\Phi\left(\frac{\hat{\mu}}{\sigma}\right)} \quad (8.13)$$

$$= \frac{1 - \Phi\left(\frac{\mu}{\sigma} - N\right)}{\Phi\left(N\right)}. \quad (8.14)$$

where N indicates the size of the error band in the unit of σ. Solving this equation for μ, the median and the expected error bands are given by

$$\mu^{\alpha\ \%} = \sigma\left[\Phi^{-1}\left(1-\Phi\left(N\right)\left(1-\alpha\right)\right)+N\right]. \tag{8.15}$$

Note that putting $N = 0$ gives the median upper limit of μ in the asymptotic limit.

8.2 Discovery of a New Boson

In order to claim a discovery of a new particle, we test the compatibility of the *background-only* hypothesis against data. A large deviation in terms of p-value will indicate that the observed data is not compatible with the *background-only* hypothesis. Thus, something "new" needs to be taken into account.

The test can be performed by generating many pseudo-datasets assuming the *background-only* hypothesis, and construct a *pdf* of the chosen test statistic. The choice of test statistic for this test is

$$q_0 = \begin{cases} -2\ln\dfrac{\mathcal{L}(X|0,\hat{\theta}_0)}{\mathcal{L}(X|\hat{\mu},\hat{\theta}_\mu)} & \text{with } \hat{\mu}\geq 0 \\ 0 & \text{with } \hat{\mu}<0 \end{cases} \tag{8.16}$$

We don't consider downward fluctuations in data($\hat{\mu}<0$) as an compatibility with data. The downward fluctuations is more likely due to systematic uncertainties such as over-estimation of backgrounds. So, in case of $\hat{\mu}<0$ the test statistic, q_0, is set 0.

Given the test statistic, we construct the *pdf*, $f(q_0|0,\hat{\theta}_0)$, with many pseudo-datasets generated assuming *background-only* hypothesis. The p-value, q_0^{obs}, is given by

$$p_0 = P\left(q_0 \geq q_0^{obs}|background-only\right) \tag{8.17}$$

$$= \int_{q_0^{obs}}^{\infty} f\left(q_0|0,\hat{\theta}_0\right)dq_0. \tag{8.18}$$

The calculated p-value can be converted into an one-sided significance Z by finding Z that satisfies

$$p-value = \int_Z^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx. \tag{8.19}$$

This is the convention in the particle physics community when expressing an evidence or discovery of a new particle [35].

In the asymptotic limit, the *pdf* is given by a mixture of a Dirac delta function and a chi-square distribution for one degree of freedom [72]

$$f\left(q_0|0,\hat{\theta}_0\right) = \frac{1}{2}\delta\left(q_0\right) + \frac{1}{2\sqrt{2\pi}\sqrt{q_0}}e^{-\frac{q_0}{2}}. \tag{8.20}$$

In this case p-value can be obtained by

$$p_0 = 1 - \Phi\left(\sqrt{q_0}\right). \tag{8.21}$$

Using this equation, the significance is given by

$$Z_0 = \Phi^{-1}\left(1 - p_0\right) = \sqrt{q_0}. \tag{8.22}$$

In this analysis(and many other analyses as well), the sensitivity to the discovery of a new particle is measured by the expected significance. In this case, the q_0 is calculated using $\mu = 1$.

Chapter 9

Systematic Uncertainty

The expected signal and background yields, and the shape of the templates can be affected by a number of sources. The sources can be categorized as follows.

- Theoretical systematic uncertainties
- Instrumental systematic uncertainties
- Background estimation systematic uncertainties

This chapter discusses the list of systematic sources in each catetory, how they are estimated, and the result of the estimations.

9.1 Treatment of systematic uncertainties

As discussed in chapter 9, systematic uncertainties on the signal and the background yields are treated by nuisance parameters. There are two kinds of nuisance parameters; that affect normalization and that affect shape of templates. Because the normalization uncertainty can not account for variation of shape, the nuisance parameters for shape variation are treated differently.

Normalization uncertainties

For most of the normalization nuisance parameters we use the log normal function as a *pdf* [72]. The functional form of the log normal is

$$\rho(\theta) = \frac{1}{\sqrt{2\pi}\ln\kappa}\exp\left(-\frac{\left(\ln(\theta/\tilde{\theta})\right)^2}{2(\ln\kappa)^2}\right)\frac{1}{\theta} \tag{9.1}$$

where θ is a nuisance parameter, $\tilde{\theta}$ is the best measure (mean or median) of the nuisance parameter, and κ is the characteristic parameter that determines the width of the distribution. For a small $\ln\kappa(\kappa \approx 1)$, we can approximate $\ln\kappa \approx \kappa - 1$. In this case the numerator in the exponent can be effectively treated small as well, *i.e.*, large values will be suppress by the characteristics of exponential functions, and can be approximated in the same way, $\ln(\theta/\tilde{\theta}) \approx \theta/\tilde{\theta} - 1$. In this

approximation, the *pdf* becomes proportional to

$$\rho(\theta) \sim \exp\left(-\frac{\left(\theta/\tilde{\theta}-1\right)^2}{2(\kappa-1)^2}\right) = \exp\left(-\frac{\left(\theta-\tilde{\theta}\right)^2}{2\tilde{\theta}^2(\kappa-1)^2}\right). \quad (9.2)$$

This equation shows that the exponential function can be approximated by a Gaussian in case of $\kappa \approx 1$, and the width of the nuisance parameter θ can be parametrized by $\tilde{\theta}(\kappa-1)$. Therefore, $\kappa-1$ is the relative width with respect to the best estimate of the nuisance parameter. In this analysis, we express nuisance parameters in term of κ.

One feature of the log normal function is that the function dies at 0 and we can avoid negative values. This is a big advantage that we can avoid the problems such as truncation of *pdf* at 0 as it happens with Gaussian *pdf*.

There are two nuisance parameters we do not use log-normal as a pdf, the signal strength and the normalization of $qq \to W^+W^-$ in the shape-based method. Both nuisance parameters use a flat(=constant) function as a *pdf*. The rationale behind this is that there is no a priori knowledge on those nuisances. The nuisance parameter for $qq \to W^+W^-$ normalization is chosen such that fit can determine the best value of the nuisance using the signal-free region dominated by $qq \to W^+W^-$ events, without any preference of a priori knowledge.

Shape uncertainties

In the shape-based method, there are systematic sources that can change shapes by bin-to-bin migrations in the 2-dimensional templates. The normalization uncertainty described by the log-normal or the flat *pdf* does not account for this because it changes the overall normalization keeping the shape of the distribution unchanged. So, for the sources that can cause bin-by-bin migrations, we use alternate shapes. The alternate shapes are constructed by changing the source of uncertainty by $\pm1\sigma$. Then, two alternate shapes, up($+1\sigma$) and down(-1σ) shapes, are used in the statistical machinery following the vertical morphing technique [77]. This technique uses one additional parameter which follows Gaussian *pdf*, and morphs the alternate shapes such that when the value of the parameter

is +1(-1), the corresponding variation is $+1\sigma(-1\sigma)$, *i.e.*, up(down) shape.

When the alternate shapes are constructed, the correlation between m_{T} and $m_{\ell\ell}$ is also taken into account naturally. Given that there is only one morphing parameter that moves all bins by the same amount, no matter how the bins are arranged, the correlation is still conserved. This is important because we unroll the 2-dimensional template to 1-dimensional histograms because of a limitation of the available statistical tools.

In the following sections, if there is a related shape uncertainty caused by a given source, the following plots will be shown in the region where signal is populated($60 < m_{\mathrm{T}} < 120$ GeV and $12 < m_{\ell\ell} < 100$ GeV); the 2-dimensional up/down shapes relative to the central shape and m_{T} and $m_{\ell\ell}$ projections of the up/down/central shapes.

9.2 Theoretical systematic uncertainties

When data-driven estimation is not applicable, we should rely on the theoretical calculation, and it depends on various sources : The PDF and the α_s, missing higher order corrections, parton shower and underlying events and jet bin fraction. The theory uncertainties contribute to the 70 %[1] of the expected total systematic uncertainty of the signal strength. The uncertainty on the jet bin fraction of signal is the dominant theory systematic uncertainty which accounts for about half of the total theory uncertainty.

9.2.1 PDF+α_s

The parton distribution function(PDF) uncertainty together with α_s uncertainty that affects lepton acceptance and the efficiency of all cuts is estimated following the prescription recommended by PDF4LHC group [78]. We take the three PDF sets, MSTW2008, CT10 and NNPDF, and propagate the uncertainty of each prescription. The envelop of the uncertainties is taken as total uncertainty.

[1] 70 % means a half of the total uncertainty because it is added in quadrature.

The PDF+α_s uncertainty is divided into two groups depending on the production source($q\bar{q}$ or gg). The processes in the same group are assumed to be 100 % correlated while the processes in the different groups are assumed to be 100 % uncorrelated. The uncertainty for $gg \to H$ ranges from 7 % to 12 %, and the uncertainty for other signal processes is 5 %. All other processes, $gg \to W^+W^-$, $qq \to W^+W^-$, WZ/ZZ, $W\gamma$ and $W\gamma^*$ have uncertainty of 4 %.

In the shape-based method, the PDF+α_s shape uncertainty is considered for $qq \to W^+W^-$ and $gg \to W^+W^-$. Following the method described in [79], we take an envelop of different PDF sets(NNPDF2.0, CT10(CTEQ) and MSTW2008), and measure variation with respect to the default PDF set(LO CTEQ6L1 PDF) used in MC generation. Figure 9.1 and 9.2 show the up/down shapes in 2-dimensional template and its projection to m_{T} and $m_{\ell\ell}$ axes in the signal region. The variation is almost flat on the 2D plane with a size less than 5 %. The impact of changing from normalization uncertainty to shape systematics on PDF is less than 1 % to the expected significance.

9.2.2 Missing higher order corrections

The cross section of a particular process is calculated by a perturbation expansion, where the first few orders are considered due to complication of calculation with higher order terms. In case of SM Higgs production, the cross sections is calculated up to NNLO in QCD as shown in Table 1.2. Missing higher orders in the calculation should be accounted for in some way, and this is done by varying the renormalization scale(μ_{R}) and the factorization scale(μ_{F}) by a factor of 2 or 1/2 [32].

For the $gg \to H$ and $qq \to W^+W^-$ processes for $m_{\mathrm{H}} > 200$ GeV, this effect is expressed in terms of uncertainty to the exclusive jet bins and will be discussed in detail in section 9.2.4. For other process, the uncertainty due to QCD scale variation ranges from 1 - 4 %.

In the shape-based method, the effect of the μ_{R} and μ_{F} variations to shapes is considered for $qq \to W^+W^-$. The MC@NLO 4.0 [80] is used for the matrix element calculation with different choice of scales. The up variation corresponds

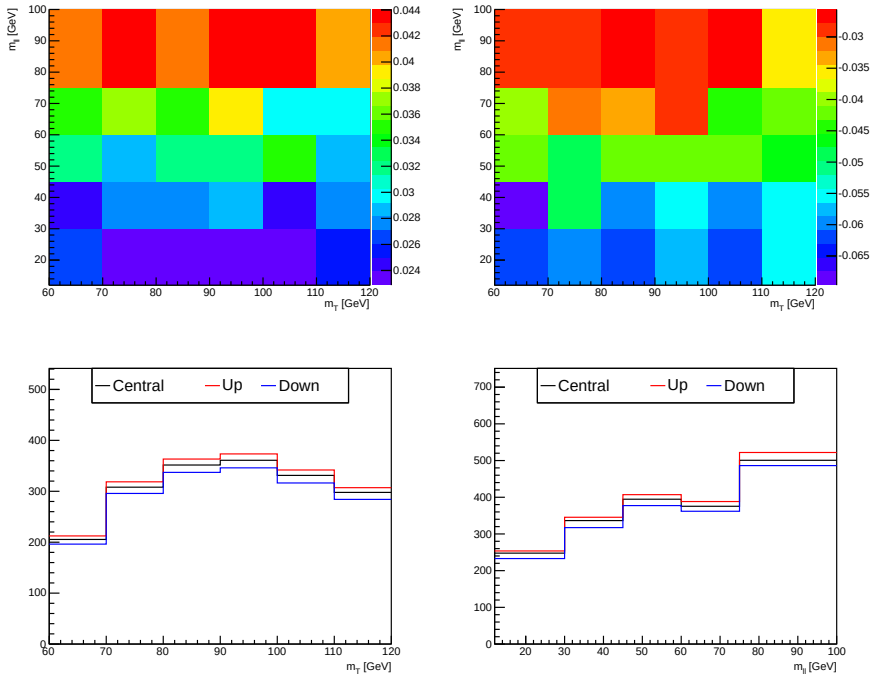


Figure 9.1: $qq \to W^+W^-$ PDF+α_s in the 0-jet category. Top are up and doen 2-dimensional shapes expressed in terms of relative difference from the nominal shape.

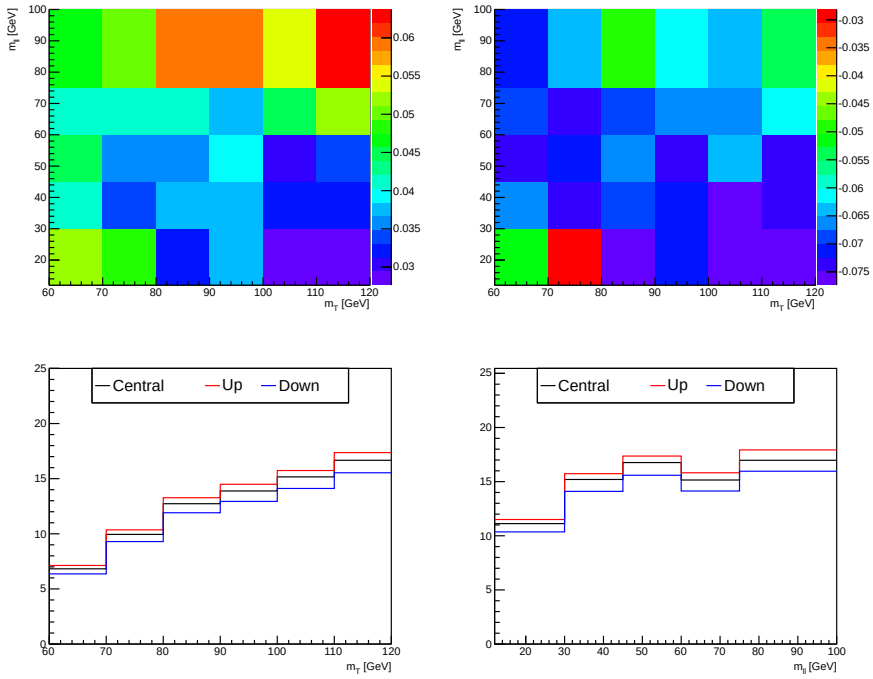


Figure 9.2: $gg \to W^+W^-$ PDF+α_s in the 0-jet category. Top are up and down 2-dimensional shapes expressed in terms of relative difference from the nominal shape.

to $\mu_{\mathrm{R}} = 0.5\mu$ and $\mu_{\mathrm{R}} = 2.0\mu$, and the down variation corresponds to $\mu_{\mathrm{R}} = 2.0\mu$ and $\mu_{\mathrm{R}} = 0.5\mu$ where μ is the nominal scale value. Figure 9.3 shows the up and down shapes in the 2-dimensional template and its projection to m_{T} and $m_{\ell\ell}$ axes in the signal region.

Another way of including the effect of missing higher order corrections is to compare results calculated at different orders. For $qq \to W^+W^-$ and $t\bar{t}/tW$ backgrounds we have the shape systematics that compare two different simulations. For $qq \to W^+W^-$ an alternate sample generated by MC@NLO 4.0 which is calculated up to NLO is used to define a up variation, and the mirror with respect to the central shape(Madgraph 5.1 [65]) which is calculated up to LO is used as a down variation. Figure 9.4 shows the up/down alternate shapes. The MC@NLO sample uses different parton shower model, Herwig++ [81], as opposed to Pythia 6.4 [82] used by the default sample, so it also accounts for the effect of uncertainty to the modeling of parton shower. For $t\bar{t}/tW$ an alternate sample generated by Madgraph 5.1 which is calculated up to LO is used to define a up variation, and the mirror with respect to the central shape(Powheg [66]) which is calculated up to NLO is used as a down variation. Figure 9.4 shows the up/down alternate shapes.

9.2.3 Parton shower and underlying events

The systematic uncertainty on modeling of parton shower(PS) and underlying events(UE) [83, 84] is evaluated by comparing different generators using different PS and UE models. A simulation chain that uses Powheg for ME calculation interfaced with Pythia 6.4 for PS is compared with a simulation chain that uses MC@NLO 4.0 for ME calculation interfaced with Herwig++ for PS. In order to exclude the effect of using different ME calculators, the Higgs p_{T} is normalized to the reference distribution [32]. The Table 9.1 shows the κ values of the PS/UE uncertainty for a given m_{H}.

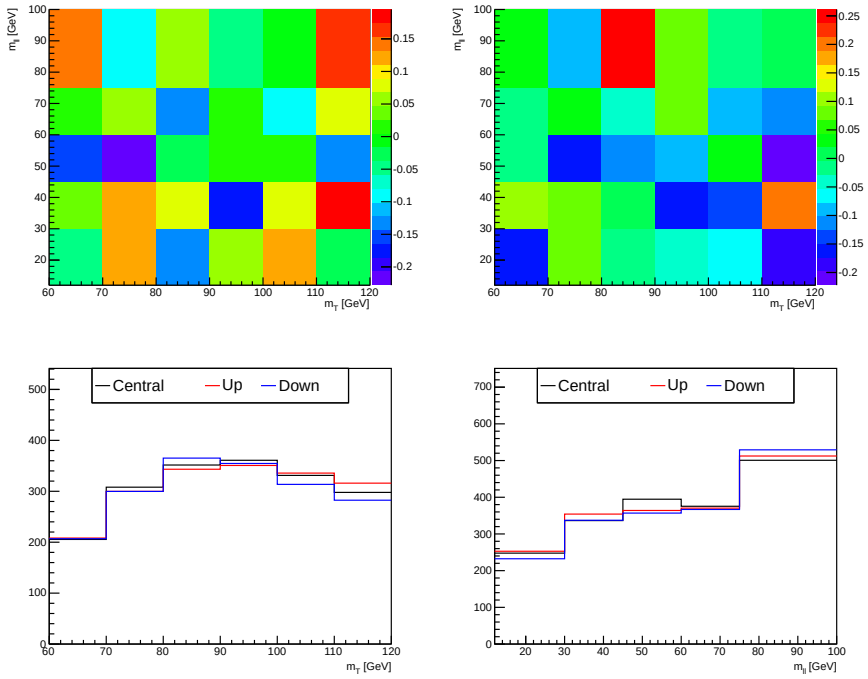


Figure 9.3: $qq \to W^+W^-$ QCD scale variation in the 0-jet category. Top are up and down 2-dimensional shapes expressed in terms of relative difference from the nominal shape.

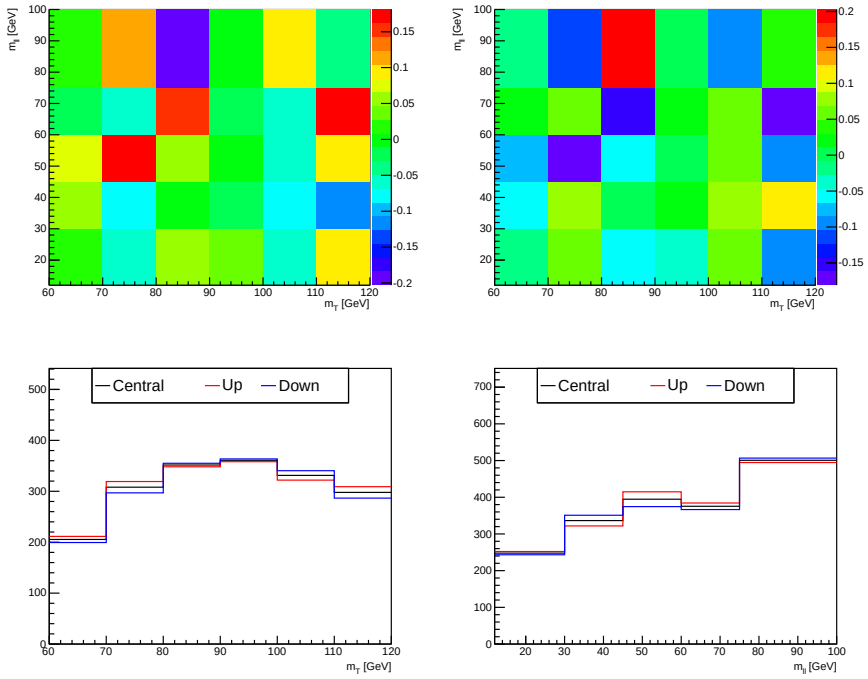


Figure 9.4: $qq \to W^+W^-$ Madgraph vs. MC@NLO in the 0-jet category. Top are up and down 2-dimensional shapes expressed in terms of relative difference from the nominal shape.

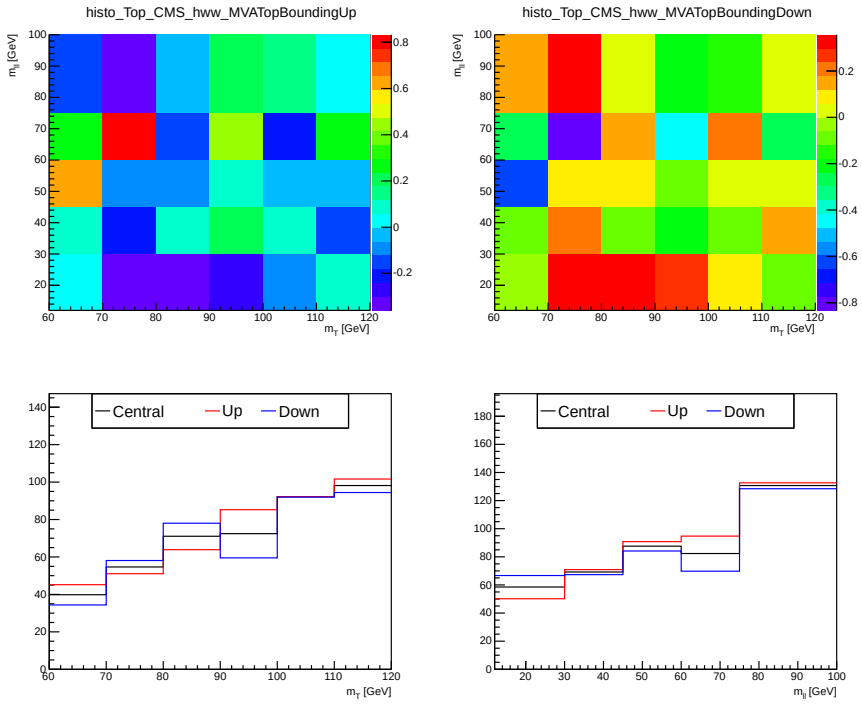


Figure 9.5: $t\bar{t}/tW$ Powheg vs. Madgraph in 1-jet category. Top are up and down 2-dimensional shapes expressed in terms of relative difference from the nominal shape.

Table 9.1: κ values of systematic uncertainty due to modeling of parton showering and underlying events.

m_{H} [GeV]	0-jet	1-jet	2-jet
115	0.941	1.128	1.212
120	0.940	1.110	1.293
130	0.937	1.113	1.237
140	0.941	1.104	1.168
150	0.942	1.093	1.156
160	0.943	1.084	1.138
170	0.946	1.075	1.108
180	0.947	1.067	1.092
190	0.948	1.068	1.083
200	0.952	1.055	1.059
250	0.955	1.058	0.990
300	0.958	1.061	0.942
350	0.964	1.068	0.889
400	0.966	1.078	0.856
450	0.954	1.092	0.864
500	0.946	1.102	0.868
550	0.931	1.117	0.861
600	0.920	1.121	0.872

9.2.4 Jet Bin Fractions

The analysis is optimized in the different jet categories, and the jets are counted with the requirement that its transverse momentum is greater than 30 GeV. For the processes that the yields in the different jet categories are estimated by data-driven methods, there is no related uncertainty because the fraction of events in the different jet categories comes from data. On the other hand, for the processes taken from simulation we need to account for this effect. The fraction of events that falls into a particular jet category is affected by the kinematics of events, and the kinematics of events is affected by missing higher order terms. In order to account for this effect, we evaluate the relevant uncertainty by varying QCD scales without any additional selection cuts.

We first get the fraction of events in the difference jet categories, $f_{\text{0-jet}}$ for 0-jet, $f_{\text{1-jet}}$ for 1-jet and $f_{\text{2-jet}}$ for $\geq$ 2-jet where the fractions are defined as

$$f_{\text{0-jet}} = \frac{\sigma_{\geq\text{0-jet}} - \sigma_{\geq\text{1-jet}}}{\sigma_{\geq\text{0-jet}}} \tag{9.3}$$

$$f_{\text{1-jet}} = \frac{\sigma_{\geq\text{1-jet}} - \sigma_{\geq\text{2-jet}}}{\sigma_{\geq\text{0-jet}}} \tag{9.4}$$

$$f_{\text{2-jet}} = \frac{\sigma_{\geq\text{2-jet}}}{\sigma_{\geq\text{0-jet}}}. \tag{9.5}$$

where $\sigma_{\geq\text{0/1/2-jet}}$ is the inclusive cross section with number of jets $\geq 0/1/2$.

The systematic uncertainties on the inclusive cross sections, $\sigma_{\geq\text{0-jet}}$, $\sigma_{\geq\text{1-jet}}$ and $\sigma_{\geq\text{2-jet}}$ are manifested by $\kappa_{\geq\text{0-jet}}$, $\kappa_{\geq\text{1-jet}}$ and $\kappa_{\geq\text{2-jet}}$, respectively. The central values of the inclusive cross sections are calculated using MSTW2008 NLO PDF at the QCD scales, $\mu_{\text{R}} = \mu_{\text{F}} = m_{\text{H}}/2$. The uncertainty on the $\sigma_{\geq\text{0-jet}}$, $\kappa_{\geq\text{0-jet}}$, is taken from the CERN Yellow Report [32], and the uncertainties on $\sigma_{\geq\text{1-jet}}$ and $\sigma_{\geq\text{2-jet}}$, $\kappa_{\geq\text{1-jet}}$ and $\kappa_{\geq\text{2-jet}}$, are calculated by the following QCD scale variations using MCFM [85] :

- $\mu_{\text{F}} = m_{\text{H}}$, $\mu_{\text{R}} = m_{\text{H}}$
- $\mu_{\text{F}} = m_{\text{H}}/4$, $\mu_{\text{R}} = m_{\text{H}}/4$
- $\mu_{\text{F}} = m_{\text{H}}$, $\mu_{\text{R}} = m_{\text{H}}/2$

- $\mu_{\rm F} = m_{\rm H}/2$, $\mu_{\rm R} = m_{\rm H}$.

following the recommendation by Higgs Cross Section WG [28]. The largest positive(Δ_+) and negative(Δ_-) uncertainties to $\sigma_{\geq 1\text{-jet}}$ and $\sigma_{\geq 2\text{-jet}}$ are taken, and they are symmetrized to be expressed in terms of κ. The symmetrization is done by

$$\kappa_{\rm symmetrized} = \sqrt{e^{\Delta_+} \times e^{\Delta_-}}, \tag{9.6}$$

where the $\kappa_{\rm symmetrized}$ is the resultant κ for the asymmetric uncertainties, Δ_+ and Δ_-. Using the values in Table 9.3, we can convert the inclusive jet bin uncertainties to the exclusive jet bin uncertainties by the formulae shown in Table 9.4 [28].

Table 9.2: Table of formulae expressing the κ values for the systematic uncertainties on the jet bin fractions due to missing higher order corrections on $\sigma_{\geq 0/1/2\text{-jet}}$ in terms of the κ values for these cross sections, $\kappa_{\geq 0/1/2\text{-jet}}$, and the jet bin fractions, $f_{0/1/2\text{-jet}}$.

Nuisance Parameter	κ's for 0-jet bin	κ's for 1-jet bin	κ's for 2-jet bin
$\theta_{\text{from } \geq 0\text{-jet}}$	$(\kappa_{\geq 0})^{\frac{1}{f_{0\text{-jet}}}}$	1.0	1.0
$\theta_{\text{from } \geq 1\text{-jet}}$	$(\kappa_{\geq 1})^{-\frac{f_{1\text{-jet}}+f_{2\text{-jet}}}{f_{0\text{-jet}}}}$	$(\kappa_{\geq 1})^{\frac{f_{1\text{-jet}}+f_{2\text{-jet}}}{f_{1\text{-jet}}}}$	1.0
$\theta_{\text{from } \geq 2\text{-jet}}$	1.0	$(\kappa_{\geq 2})^{-\frac{f_{2\text{-jet}}}{f_{1\text{-jet}}}}$	$\kappa_{\geq 2}$

For $gg \to H$ process, the numerical values of $\kappa_{\geq 0\text{-jet}}$, $\kappa_{\geq 1\text{-jet}}$ and $\kappa_{\geq 2\text{-jet}}$ are summarized in Table 9.3. Combining information in Table 9.3 and Table 9.2 we obtain the final κ's for $gg \to H$ process as shown in Table 9.4.

For $qq \to W^+W^-$ process, the inclusive cross sections, $\sigma_{\geq 0\text{-jet}}$, $\sigma_{\geq 1\text{-jet}}$ and $\sigma_{\geq 2\text{-jet}}$ are evaluated using MC@NLO 4.0 [80] and the corresponding uncertainties are estimated by combination of QCD scale variation and the comparison with ALPGEN [86]. The uncertainties to the above inclusive cross sections are 3 %, 6 % and 42 % for $\sigma_{\geq 0\text{-jet}}$, $\sigma_{\geq 1\text{-jet}}$ and $\sigma_{\geq 2\text{-jet}}$, respectively [28]. The jet bin fractions calculated using the inclusive cross sections are $f_{0\text{-jet}} = 0.70$, $f_{1\text{-jet}} = 0.22$ and $f_{2\text{-jet}} = 0.08$. Inserting these numbers into the formulae in Table 9.2, we obtain

Table 9.3: κ values for the systematic uncertainties due to missing higher order corrections for the inclusive $gg \to H$ production cross section, $\sigma_{\geq 0\text{-jet}}$, $\sigma_{\geq 1\text{-jet}}$ and $\sigma_{\geq 2\text{-jet}}$. The corresponding κs are $\kappa_{\geq 0\text{-jet}}$, $\kappa_{\geq 1\text{-jet}}$ and $\kappa_{\geq 2\text{-jet}}$, respectively.

m_{H}[GeV]	$\kappa_{\geq 0\text{-jet}}$	$\kappa_{\geq 1\text{-jet}}$	$\kappa_{\geq 2\text{-jet}}$
115	1.106	1.226	1.149
120	1.104	1.224	1.120
130	1.100	1.230	1.117
140	1.096	1.220	1.129
150	1.095	1.220	1.124
160	1.095	1.221	1.199
170	1.090	1.222	1.175
180	1.089	1.218	1.171
190	1.087	1.217	1.171
200	1.087	1.213	1.197
250	1.083	1.208	1.230
300	1.082	1.208	1.205
350	1.090	1.207	1.209
400	1.075	1.195	1.195
450	1.078	1.194	1.196
500	1.087	1.188	1.174
550	1.089	1.191	1.194
600	1.090	1.187	1.192

Table 9.4: Table of κ values for the systematic uncertainties for the jet bin fractions due to missing higher order corrections for the total inclusive Higgs cross section, the inclusive Higgs+1jet cross section, and the inclusive Higgs+2jet cross section.

m_{H}[GeV]	0-jet bin		1-jet bin		2-jet bin
	$\kappa_{\mathrm{from}\geq\text{0-jet}}$	$\kappa_{\mathrm{from}\geq\text{1-jet}}$	$\kappa_{\mathrm{from}\geq\text{1-jet}}$	$\kappa_{\mathrm{from}\geq\text{2-jet}}$	$\kappa_{\mathrm{from}\geq\text{2-jet}}$
115	1.16	0.92	1.28	0.97	1.15
120	1.16	0.92	1.28	0.97	1.12
130	1.15	0.91	1.29	0.98	1.12
140	1.15	0.91	1.28	0.97	1.13
150	1.15	0.90	1.28	0.97	1.12
160	1.15	0.90	1.28	0.96	1.20
170	1.15	0.89	1.28	0.96	1.18
180	1.15	0.89	1.28	0.96	1.17
190	1.15	0.88	1.28	0.96	1.17
200	1.15	0.88	1.27	0.96	1.20
250	1.16	0.86	1.27	0.96	1.17
300	1.17	0.84	1.27	0.95	1.20
350	1.20	0.83	1.27	0.95	1.21
400	1.17	0.82	1.26	0.95	1.20
450	1.19	0.81	1.26	0.95	1.20
500	1.22	0.80	1.25	0.95	1.17
550	1.24	0.78	1.26	0.95	1.19
600	1.25	0.78	1.26	0.94	1.19

the κ's for the jet bin fraction uncertainties for $qq \to W^+W^-$ process as shown in Table 9.5.

Table 9.5: Table of κ values for the systematic uncertainties for the jet bin fractions due to missing higher order corrections for the total inclusive $qq \to W^+W^-$ cross section, the inclusive $qq \to W^+W^-$+1jet and $qq \to W^+W^-$+2jet cross sections.

0-jet bin		1-jet bin		2-jet bin
$\kappa_{\text{from}\geq\text{0-jet}}$	$\kappa_{\text{from}\geq\text{1-jet}}$	$\kappa_{\text{from}\geq\text{1-jet}}$	$\kappa_{\text{from}\geq\text{2-jet}}$	$\kappa_{\text{from}\geq\text{2-jet}}$
1.042	0.978	1.076	0.914	1.42

9.3 Instrumental Systematic Uncertainties

There are systematic sources induced by uncertainty on the measurement of basic quantities such as luminosity, lepton momentum and resolution, lepton selection efficiency, MET resolution and jet energy scale. Another source is the statistics of the available MC samples. These sources are classified as instrumental systematic uncertainty.

These uncertainties contribute to about 45 % of the expected total systematic uncertainty of signal strength. This means that without these uncertainties, the total uncertainty is reduced by 10 %. Thus, instrumental uncertainty makes small effect to the measurement of signal strength.

Luminosity

The luminosity at CMS is measured by the hadronic forward calorimeter(HF) and the silicon pixel detector. Thanks to small dependence on experimental conditions such as pileup, the counting of the pixel clusters is chosen for the precision luminosity measurement [87]. The measured luminosity is calibrated by Van Der Meer Scans at the ISR [87]. The dominant source of uncertainty is the choice of fit function to model the bunch shapes.

The total estimated uncertainties are 2.6 % [87] at 8 TeV and 2.2 % at 7 TeV [88].

Lepton momentum scale and resolution

The lepton momentum scale and resolution can affect the selection efficiency for the cuts applied on lepton momentum or variables constructed using lepton momenta. The size of the uncertainties of both sources is estimated by comparing Z invariant mass shape between simulation and data in the $ee/\mu\mu$ final state. The momentum scale is responsible for the location of the invariant mass peak, and the momentum resolution is responsible for the width of the distribution.

The measured uncertainties are 1.5 %(3 %) for electrons in barrel(endcap) and 1.0 %(1.7 %) for muons in barrel(endcap). For the cut-based method we use the average over barrel and endcap, 2 % for electron and 1.5 % for muon. For the shape-based method, new m_{T} and $m_{\ell\ell}$ are calculated using the lepton momenta scaled by the measured resolutions. The up shape is made by adding the smeared momentum resolution to both leptons, and the down shape is made by subtracting the smeared momentum resolution to both leptons. Figure 9.6 shows the corresponding up/down alternate shapes for $qq \to W^+W^-$ in the 0-jet $e\mu/\mu e$ category.

Lepton efficiency

The lepton selection and trigger efficiencies are measured by the Tag-and-Probe method as described in section 6.2. In this method, uncertainties come from the determination of the background contribution and the modeling of the signal and background shapes in the likelihood fit. An additional uncertainty which is prominent in the low p_{T} region in the electron case comes from the possible bias by using the N-1 technique. This is described in section 6.2.1 as well as the corresponding uncertainties in Table 6.3.

The estimated uncertainties are 1.5 % for muons and 2 % for electrons, and we use 3 % and 4 % for $\mu\mu$ and ee events, respectively, in the cut-based method. For the shape-based method, alternate shapes are constructed by scaling up/down

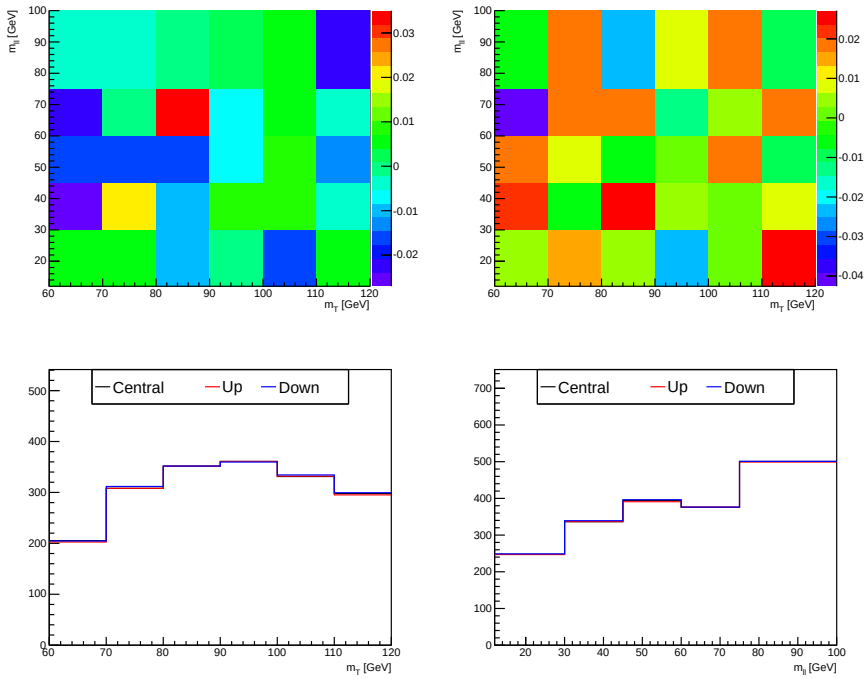


Figure 9.6: Alternate shapes for lepton momentum scale and resolution for $qq \to W^+W^-$ in the 0-jet $e\mu/\mu e$ category. Top are up and down 2-dimensional shapes expressed in terms of relative difference from the nominal shape.

the lepton efficiencies using the systematics sources as a function of lepton p_{T} and η. Figure 9.7 shows the corresponding up/down alternate shapes for $qq \to W^+W^-$ in the 0-jet $e\mu/\mu e$ category.

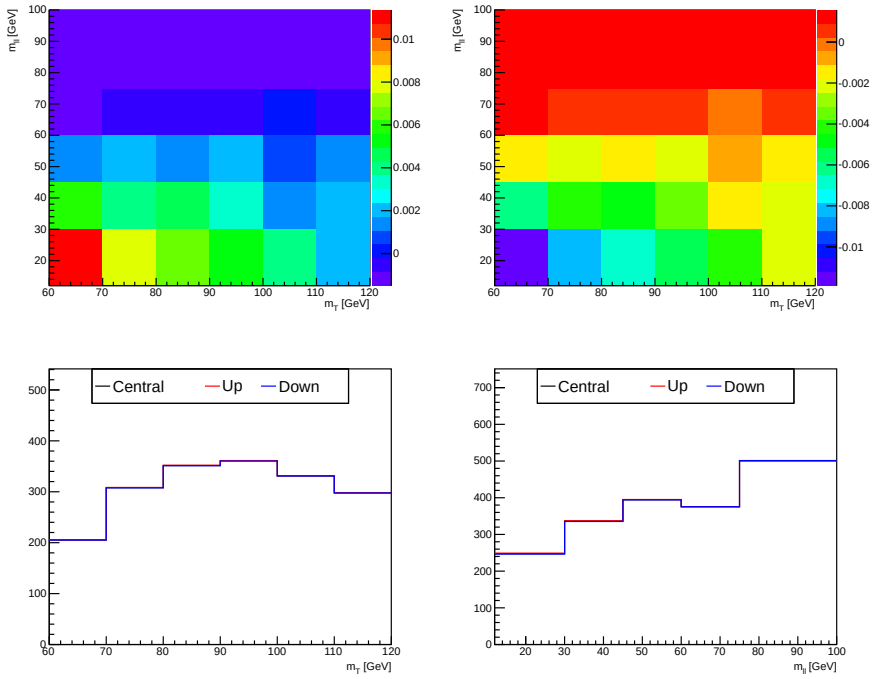


Figure 9.7: Alternate shapes for lepton efficiency for $qq \to W^+W^-$ in the 0-jet $e\mu/\mu e$ category. Top are up and down 2-dimensional shapes expressed in terms of relative difference from nominal shape.

MET resolution

The mis-modeling of MET by simulation can introduce a systematic uncertainty as we select events that pass MET and MET-related selections. The uncertainty is measured by comparing data and simulation using $Z/\gamma^* \to \ell\ell$ events. To account for the difference between data and simulation, an additional Gaussian smearing for the individual x and y components of MET are needed. For PF MET, the size of Gaussian smearing is 3.2 and 3.6 GeV for 0-jet and 1-jet categories, respectively. For trkMET, the size of Gaussian smearing is 1.0 and 4.5 GeV for 0-jet and 1-jet categories, respectively.

For the cut-based method, the resultant effect of smearing to yields is around 2 %. For the shape-based method, the up/down alternate shapes are constructed with the new m_{T} calculated using the smeared x and y components of MET. Figure 9.7 shows the corresponding up/down alternate shapes for $qq \to W^+W^-$ in the 0-jet $e\mu/\mu e$ category.

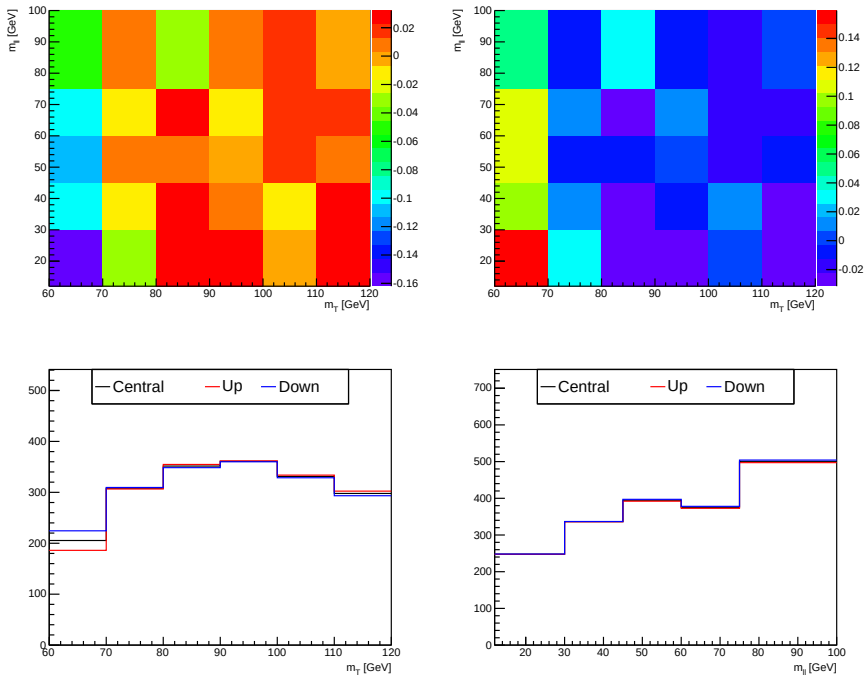


Figure 9.8: Alternate shapes for MET resolution for $qq \to W^+W^-$ in the 0-jet $e\mu/\mu e$ category. Top are up and down 2-dimensional shapes expressed in terms of relative difference from nominal shape.

Jet energy scale

The analysis is optimized in jet categories, and the jet selection is done by requiring p_{T} to be greater than 30 GeV. Thus, uncertainty on the energy of jets can lead to migration between different jet categories. From the studies on the PF jet energy resolution done in [89], the jet energy resolution ranges from 3 % to 5 % as the η of jets increase. We take 5 % for whole η range as a conservative choice.

For the cut-based method, the resultant effect of 5 % of jet energy resolution to yields is around 2 %. For the shape-based method, the up/down alternate shapes

are constructed by scaling the jet transverse momentum by ± 5 %. Figure 9.9 shows the corresponding up/down alternate shapes for $qq \to W^+W^-$ in the 0-jet $e\mu/\mu e$ category.

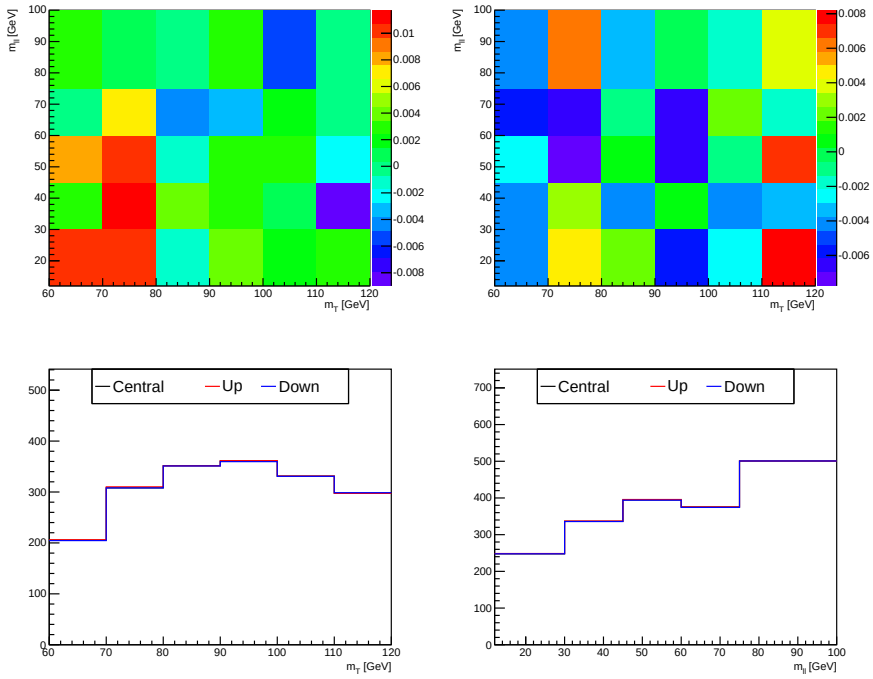


Figure 9.9: Alternate shapes for jet energy resolution for $qq \to W^+W^-$ in the 0-jet $e\mu/\mu e$ category. Top are up and down 2-dimensional shapes expressed in terms of relative difference from the nominal shape.

Statistics of simulated samples

The limited statistics of the available simulation samples should be taken into account as a source of systematic uncertainty. For the cut-based method, overall statistical uncertainty of the sample after the final selection is taken into account. For the shape-based method, the up/down alternate shapes are constructed by adding/subtracting the size of statistical uncertainty in each bin. This does not allow bin-by-bin fluctuation in the statistical machinery because all bins are either up or down by the statistical uncertainties in each bin. The ideal method should be considering each bin independently, but this requires an extensive CPU

consumption. We checked the expected and the observed significances using both approaches, and found that the results are compatible within a few %. Thus, we use the former approach. Figure 9.10 shows the corresponding up/down alternate shapes for $gg \to W^+W^-$ in the 0-jet category.

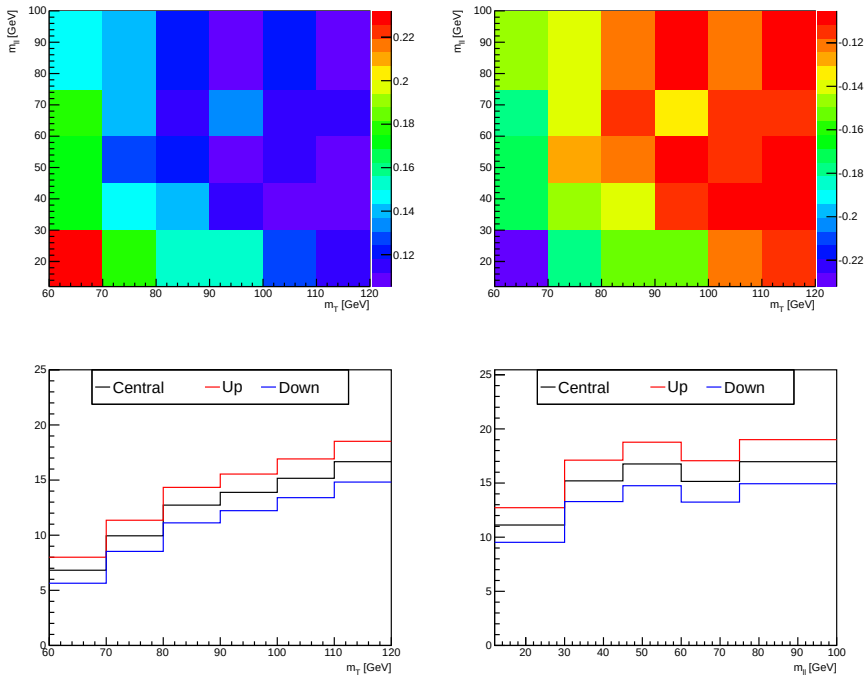


Figure 9.10: Alternate shapes for MC statistics for $gg \to W^+W^-$ in the 0-jet $e\mu/\mu e$ category. Top are up and down 2-dimensional shapes expressed in terms of relative difference from nominal shape.

9.4 Background Estimation Uncertainty

The expected background contributions in the signal region are estimated by data-driven methods or taken from simulation. All procedures and the source of systematic uncertainties are discussed in chapter 7. These uncertainties are related to the normalization of each background. But, there is a shape systematic uncertainty that can cause a variation of shapes.

These uncertainties contribute to about 55 % of the expected total systematic uncertainty of signal strength.

W + jets **alternate shapes**

As described in section 7.2, the W + jets background is estimated by a data-driven method which measures the fake rate in the QCD di-jet sample. One of the systematic sources is the variation of the away jet p_{T}, and the uncertainty of 30 % is assigned to take this account. In the shape-based method, the affect of the varying the away jet p_{T} to the shape is considered as well. The alternate up shape is constructed using alternate jet p_{T} thresholds, 15 GeV for muons and 20 GeV for electrons, and the relative difference in shape is taken. The alternate down shape is taken as a mirror of the up shape with respect to the central shape. Figure 9.11 shows the corresponding up/down alternate shapes for W + jets when muon is an FO.

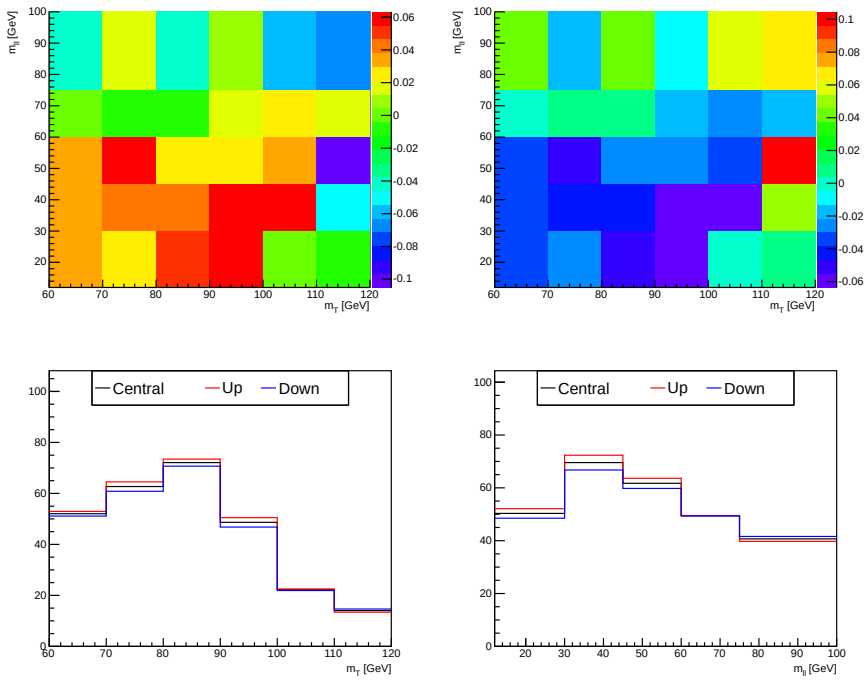


Figure 9.11: Alternate shapes for W($\mu\nu_{\mu}$) + jets. Top are up and down 2-dimensional down shapes expressed in terms of relative difference from the nominal shape.

9.5 Summary

Table 9.6: Summary of systematic uncertainties and their relative contributions to the expected total systematic uncertainty of signal strength.

Systematic sources	sub-division	contribution to total systematic uncertainty of signal strength	Total
Theory uncertainty	signal(jet bin fraction)	50 %	70 %
	signal(except jet bin fraction)	35 %	
	background	30 %	
Instrumental uncertainty	MC statistics	25 %	45 %
	except MC statistics	40 %	
Background Estimation	W + jets	40 %	55 %
	except W + jets	40 %	

Table 9.6 shows a summary of systematic uncertainties and their contributions to the expected total systematic uncertainty of signal strength. Each contribution in the table is estimated by removing relevant nuisance parameters in the fit. Therefore, these numbers are not the exact values in the full fit because removing some nuisance parameters can alter correlations between them. But, these numbers give a sense of which component is dominant and which is not.

The dominant systematic source is the theory uncertainty, particularly, the jet bin fraction uncertainty of signal. The instrumental uncertainty contributes little compared to other sources. The background estimation uncertainties is dominated by W + jets.

In chapter 11.5 it will be shown that the systematic and statistical uncertainties of signal strength are at the same level with current data. The analysis will be limited by systematic uncertainties as more data is added. Therefore, to improve the analysis, one needs to reduce the theoretical uncertainty on the signal.

Chapter 10

Fit Validation

10.1 Validation of the Fit Model Using Real Data

In the shape-based method, we rely on simulation for the background shapes except $\mathrm{W}+\mathrm{jets}$. Therefore, it is important to confirm that our fit model, both shape and normalization, correctly describes what is observed in data.

The idea is to choose a control region which is dominated by (a) process(es) under study, and to fit the region using our fit model. If the fit model is not correct then the post-fit shape and normalization would not describe the data correctly, *i.e.* difference between data and the post-fit shape and normalization would be large.

This section describes the validation of the fit model for the dominant backgrounds, $qq \to W^+W^-$, $t\bar{t}/tW$, $\mathrm{W}+\mathrm{jets}$ and $W\gamma/W\gamma^*$, using dedicated control regions.

10.1.1 Validation of $qq \to W^+W^-$ fit model

The most sensitive channel, 0-jet $e\mu/\mu e$, is dominated by the $qq \to \mathrm{W^+W^-}$ background. Therefore validating the $qq \to \mathrm{W^+W^-}$ fit model is crucial. This section describes a validation study using two control regions dominated by $qq \to \mathrm{W^+W^-}$ events. The basic idea is that if the fit model is not correct then the prediction using one region may not be compatible with that using another region. To test this, we use the post-fit result in one control region, and predict the shape and the normalization in another region. Following are the procedure and the result of this test.

We first select two control regions with similar statistics. It is important that both regions have a similar number of events so that they have a similar statistical power to constrain backgrounds. Figure 10.1 shows signal region(SR) which is drawn in black and two control regions, CR1 and CR2, which are drawn in red and in blue, respectively. The SR is defined as $60 < m_{\mathrm{T}} < 120$ GeV and $12 < m_{\ell\ell} < 60$ GeV. The CR1 is defined as $120 < m_{\mathrm{T}} < 280$ GeV and $12 < m_{\ell\ell} < 200$ GeV, and the CR2 is defined as $60 < m_{\mathrm{T}} < 120$ GeV and $60 < m_{\ell\ell} < 200$ GeV. The composition of signal and backgrounds in the two

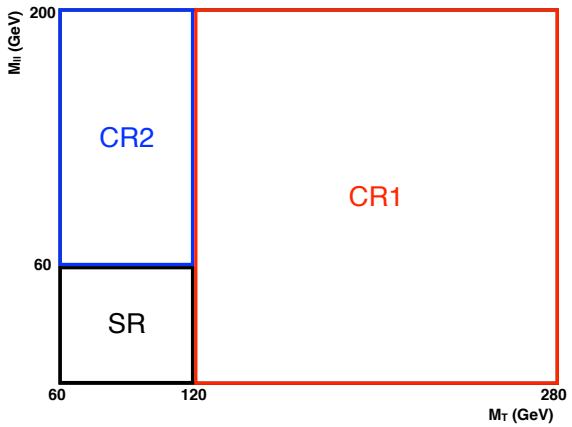


Figure 10.1: Definition of signal region(SR) and two control regions(CR1 and CR2).

Table 10.1: Expected signal and background yields in the two control regions(CR1 and CR2) and the full template region(Full range).

Region	Signal	$qq \to W^+W^-$	$gg \to W^+W^-$	WZ/ZZ	$t\bar{t}/tW$	
CR1	27.0	1321.0	113.2	46.2	289.6	
CR2	13.5	1672.7	54.5	51.5	146.1	
Full range	238.4	3969.6	210.6	132.6	498.7	
Region	$W(e\nu_e)$ + jets	$W(\mu\nu_\mu)$ + jets	$W\gamma$	$W\gamma^*$	$Z \to \tau\tau$	Data
CR1	54.9	22.4	6.0	19.3	2.8	1892
CR2	108.1	128.2	19.3	21.4	19.9	2155
Full range	282.8	331.8	115.6	167.8	46.0	5729

control regions is summarized in Table 10.1. Both regions are dominated by $qq \rightarrow W^+W^-$with purity of around 70(75)% in CR1(2). The signal contamination is negligible(less than 1.5%) in both regions.

Because this test is only for $qq \rightarrow W^+W^-$, other processes have to be fixed in the fit. Therefore, all other processes are fixed to the post-fit normalization and shape of the nominal fit except $qq \rightarrow \mathrm{W^+W^-}$ process. Hereafter, this configuration will be called "full range" fit. In the full range fit, all nuisances for $qq \rightarrow \mathrm{W^+W^-}$ are included, but all nuisances for other processes are dropped because shape and normalization for those processes are already post-fit results. As a sanity check

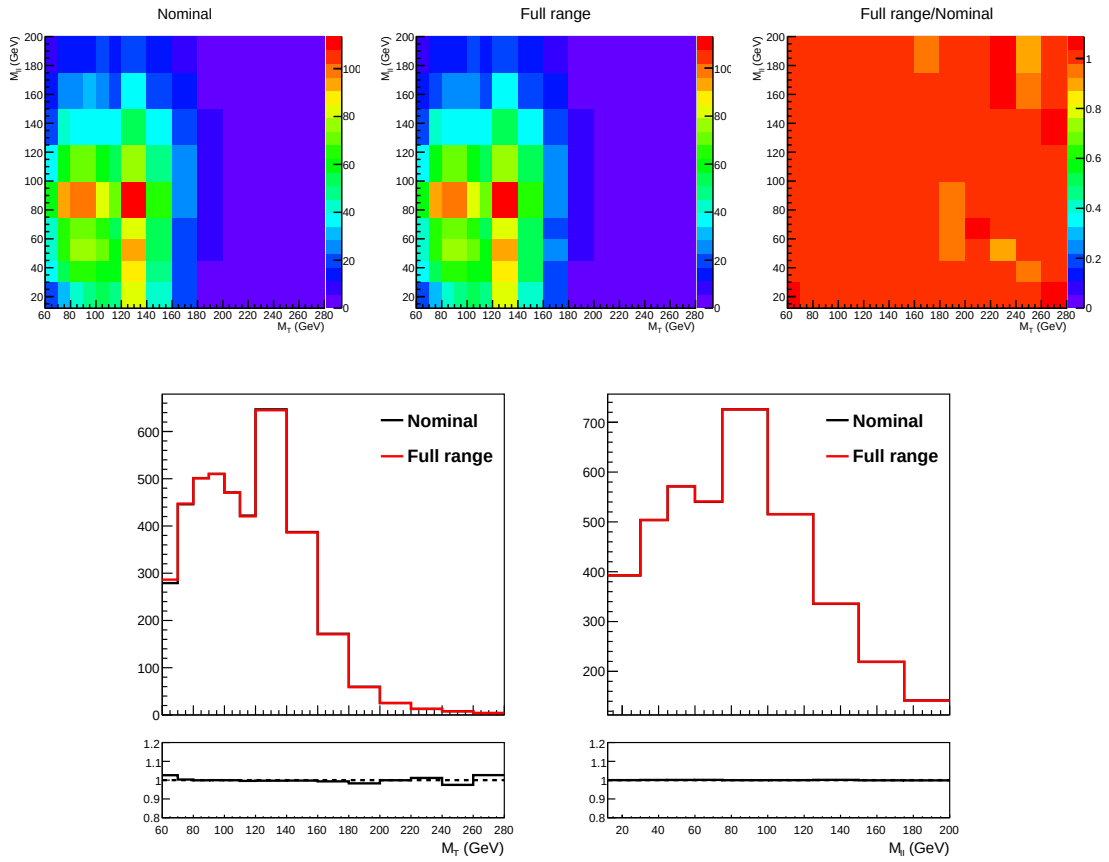


Figure 10.2: Comparison of post-fit shapes from the nominal and the full range fits.

we compare the post-fit shape and normalization of $qq \rightarrow W^+W^-$ from the full range fit with those from the nominal fit in Figure 10.2 and Table 10.2. Both shapes and normalizations from the full range fit are consistent with the nominal

Table 10.2: Comparison of normalization from nominal and full range fits.

Fit	$qq \to VH$	$qq \to qqH$	$gg \to H$	$qq \to W^+W^-$
Nominal	4.8	1.9	144.3	3945.4
Full range	4.8	1.9	143.6	3947.0

fit results.

Since the full range fit has been validated, we perform two fits using only one of the control regions. The fit using CR1(2) will be denoted as CR1(2) fit hereafter. In a fit using one control region, the other region is removed in the fit, and the normalization of each process is fixed accordingly.

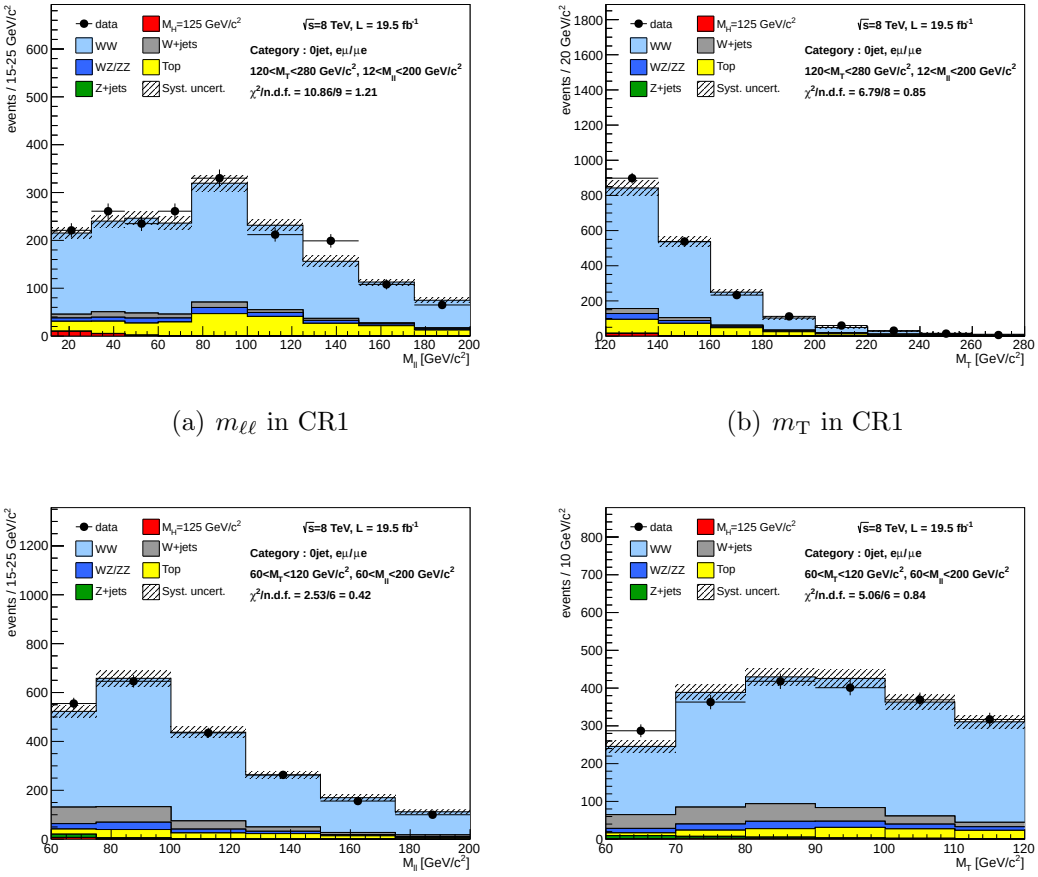
Figure 10.3 shows the m_{T} and $m_{\ell\ell}$ distributions in CR1 and CR2 using the post-fit result from the other control regions. The uncertainty band is the post-fit uncertainty obtained by pseudo-data sets generated using the post-fit uncertainties. The variances in each bin of $qq \to W^+W^-$ and all other backgrounds are taken from the full range and the nominal fits, respectively, and are added in quadrature. The agreement between data and prediction is measured by

$$\chi^2/n.d.f. = \frac{1}{N_{bin}} \sum_{i=1}^{N_{bin}} \frac{\left(N_i^{data} - N_i^{simulation}\right)^2}{\left(\sigma_i^{data}\right)^2 + \left(\sigma_i^{simulation}\right)^2} \tag{10.1}$$

where N_{bin} is the number of bins, $N_i^{data(simulation)}$ is the yield of data(simulation) in the i^{th} bin and $\sigma_i^{data(simulation)}$ is the uncertainty of data(simulation) in the i^{th} bin. As shown on each plot, the $\chi^2/n.d.f.$ is close to 1, which means that all distributions show a good agreement with data. This indicates that our $qq \to W^+W^-$ fit model is not biased.

Table 10.3 shows the best-fit μ(signal strength) values from full range, CR1 and CR2 fits. Using different control regions results in consistent best-fit μ values. This is another evidence that our fit model is correct.

The fit model is further examined by using an alternate shape(MC@NLO or Powheg) as the central shape and the central shape(Madgraph) as an alternate shape. Ideally, if the fit model is consistent with data, switching central



(a) $m_{\ell\ell}$ in CR1

(b) m_{T} in CR1

(c) $m_{\ell\ell}$ in CR2

(d) m_{T} in CR2

Figure 10.3: $m_{\ell\ell}$(a,c) and m_{T}(b,d) distributions in CR1 at the top and CR2 at the bottom using the fit results of CR2 and CR1.

Table 10.3: Comparison of the best-fit μ values from the full range, CR1 and CR2 fits, where μ is the signal strength.

	full range fit	CR1 fit	CR2 fit
Best-fit μ	0.63	0.63	0.62

and alternate shapes should not affect the final result because any difference between data and prediction should be covered by the shape uncertainty. Table 10.4

Table 10.4: Exclusion limit on the production rate of $H \to W^+W^- \to 2l2\nu$ at CL_s=95 %, significance and the observed(best-fit) signal strength using different central shapes for $qq \to W^+W^-$ process in the shape-based analysis.

Central shape	Exclusion limit observed / expected	Significance observed / expected	Signal strength observed
Madgraph(default)	1.2 / 0.4	4.0σ / 5.2σ	0.76 ± 0.21
MC@NLO	1.2 / 0.4	4.2σ / 5.3σ	0.82 ± 0.24
Powheg	1.2 / 0.4	3.9σ / 5.1σ	0.74 ± 0.21

shows exclusion limit normalized by SM prediction, significance in the unit of standard deviation(σ) and the observed(best-fit) signal strength using different central shapes for $qq \to W^+W^-$ process in the shape-based analysis. It shows that the results do not depend on the choice of central shape. As mentioned, this is an indication that our fit model is correct.

Based on above observations, we conclude that our fit model for $qq \to W^+W^-$ process fits data correctly.

10.1.2 Validation of $t\bar{t}/tW$ fit model

By inverting the top-tagging requirement, we can select an event sample dominated by $t\bar{t}/tW$ events. We perform the same shape-based analysis with templates constructed using the top-tagged events. The Figure 10.4 shows the $m_{\ell\ell}$ and m_{T} distributions after normalizing the shape to the post-fit result. The agreement between the post-fit fit result and data is good, and we conclude that the $t\bar{t}/tW$ fit model is correct.

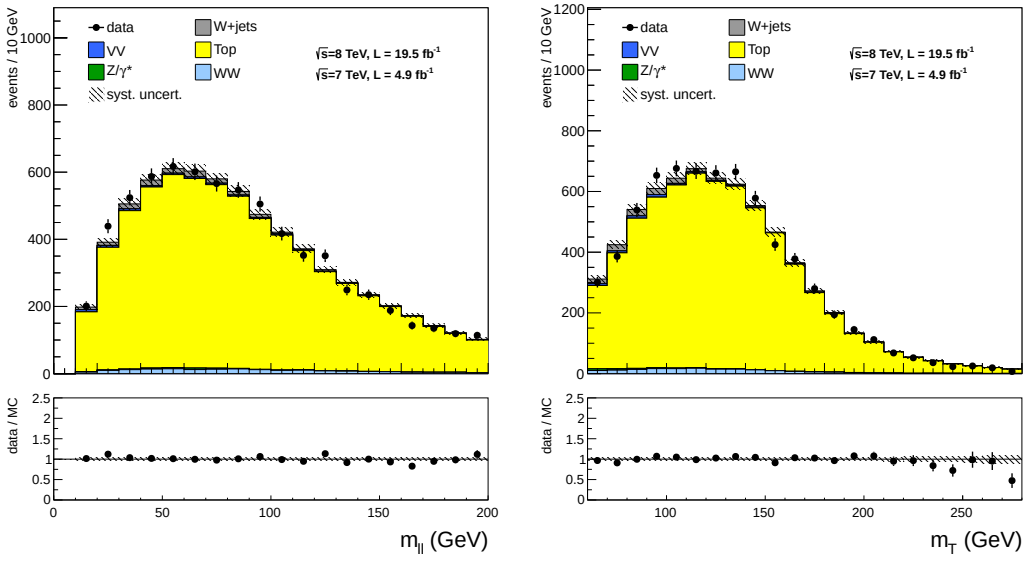


Figure 10.4: The post-fit $m_{\ell\ell}$ and m_{T} distributions in top-tagged region.

10.1.3 Validation of $\mathrm{W}+\mathrm{jets}$ and $W\gamma/W\gamma^*$ fit models

By inverting the opposite-sign requirement on the leptons, we can select an event sample dominated by $\mathrm{W}+\mathrm{jets}$ and $W\gamma/W\gamma^*$ events. We perform the same shape-based analysis with templates constructed using the same-sign events. The Figure 10.5 shows the $m_{\ell\ell}$ and m_{T} distributions after normalizing the shape to the post-fit result. The agreement between the post-fit fit result and data is good, and we conclude that the $\mathrm{W}+\mathrm{jets}$ and $W\gamma/W\gamma^*$ fit models are correct.

10.2 Post-fit Analysis

Fit is a tool that determines the normalization and shape of the signal and the backgrounds that describes the data best. We use the maximum likelihood fit that scans nuisances, and finds the point where the likelihood has its maximum. The tool can do what is best in terms of finding the maximum, but sometimes the result does not make sense, which might indicate that our fit model is not correct.

This section thus discusses the post-fit result of the fit to confirm that the

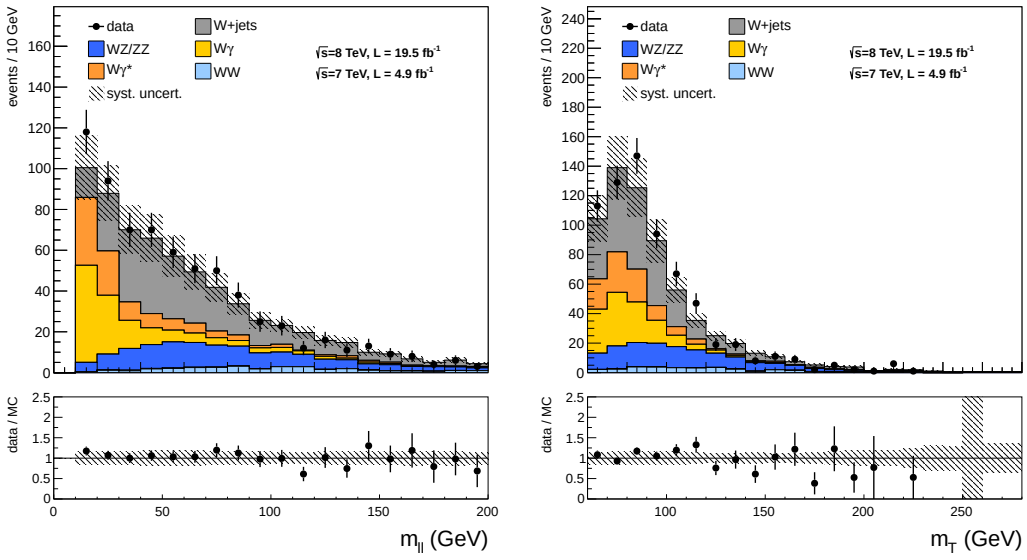


Figure 10.5: The post-fit $m_{\ell\ell}$ and m_{T} distributions in same-sign region.

fit is stable, and the result makes sense.

10.2.1 Post-fit result of nuisances

In an ideal world where prediction of the central value of a nuisances is exactly what nature gives, the nuisance should not change by fit. But, in reality prediction can be wrong, thus we assign uncertainty on the nuisances. On the other hand, if the prediction is wrong by a large margin, the post-fit value of the nuisance can be far from the prediction, even larger than the uncertainty we assign. So, we need to make sure that the post-fit nuisances are within the assigned uncertainty. A large variation of nuisances may indicate that the fit model does not describe data correctly.

The measure to check this is the "Pull" which is defined as

$$\mathrm{Pull} = \frac{\theta_{\text{post-fit}} - \theta_{\text{pre-fit}}}{\sigma_{\text{pre-fit}}} \tag{10.2}$$

where $\theta_{\text{pre-fit(post-fit)}}$ is the central value of a nuisance before(after) the fit, and $\sigma_{\text{pre-fit}}$ is the assigned uncertainty to that nuisance. If the prediction is perfect, and there

is no variation of the nuisance by fit, then the Pull should be 0. The Pull with larger than 1, *i.e.* the variation is larger than the assigned error, may indicate that the fit model is not correct or a lack of understanding of that nuisance.

One of the reasons why we rely on fit is that the data can constrain nuisances more than the assigned uncertainties. If there is a region that is dominated by a certain background process then the fit can use that region to constrain that background. Also, as discussed in section 8.1, profiling nuisances can sometimes over-constrains them. Therefore, we should verify that the nuisances deserve to be over-constrained, and if they do not, what the impact of these nuisance parameters to the final result is, for example, to the measurement of signal strength.

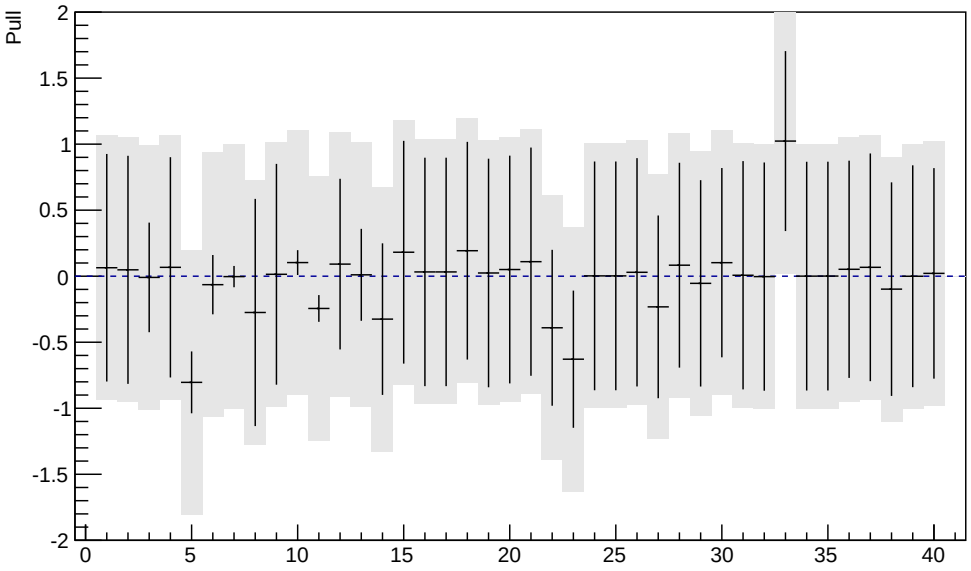
In order to assess any effects of this kind, we define the normalized uncertainty, σ_{norm}, which is defined as a ratio of the post-fit uncertainty($\sigma_{\text{post-fit}}$) to the pre-fit uncertainty($\sigma_{\text{pre-fit}}$) of a nuisance,

$$\sigma_{\text{norm}} = \frac{\sigma_{\text{post-fit}}}{\sigma_{\text{pre-fit}}}. \tag{10.3}$$

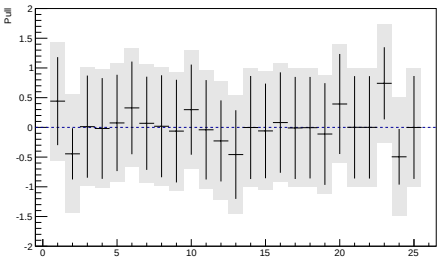
Nuisances with σ_{norm} much smaller than 1 should be examined to avoid the issue of over-constraining.

Figure 10.6 shows the post-fit result of nuisance parameters. The x axis is the indices of the nuisance parameters, and the y axis is the Pull and the σ_{norm}. The black line shows the Pull as the central value and σ_{norm} as an uncertainty bar. The grey area is drawn with the Pull as central value and the error bar being 1. This is drawn to visualize the size of the post-fit uncertainty with respect to the assigned uncertainty. The nuisances are divided into 5 groups; nuisances that are used (a) in all final states, (b) only in the $e\mu/\mu e$ 0-jet category, (c) only in the $ee/\mu\mu$ 0-jet category, (d) only in the $e\mu/\mu e$ 1-jet category, and (e) only in the $ee/\mu\mu$ 1-jet category.

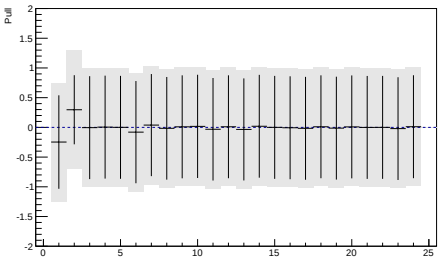
There is one nuisance parameter that has Pull greater than 1. It is one of the theory uncertainties on the $gg \to W^+W^-$process. Given that this process is not the dominant background process, it does not affect the result of the statistical interpretation. This nuisance causes a large variation of $gg \to W^+W^-$ normalization by fit, which is discussed in section 10.2.2.



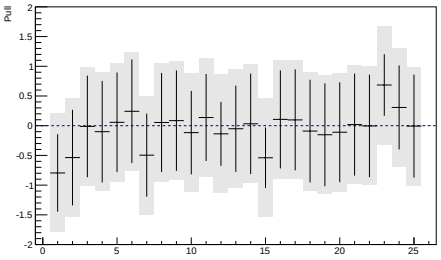
(a)



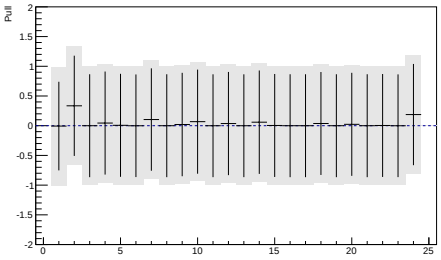
(b)



(c)



(d)



(e)

Figure 10.6: Post-fit results on the nuisance parameters.

There are 5 shape nuisance parameters that are constrained by more than 70 %, *i.e.* σ_{norm} is less than 0.3. They are instrumental nuisances for lepton momentum resolution, MET resolution and alternate shape nuisances for $qq \to W^+W^-$ and $t\bar{t}/tW$. The analysis is not designed to constrain the instrumental nuisances as opposed to, for example, $Z \to \ell\ell$ where lepton momentum resolution can be constrained using the width of Z peak. Thus, some nuisances of this kind can be over-constrained as an artifact of the fit allowing others nuisance not being constrained enough. In section 9.3, it is mentioned that the instrumental uncertainties does not affect the final result. This was re-confirmed by that the final result does not change even without these nuisance parameters or replacing them by normalization nuisance parameters used by cut-based method.

Apart from the instrumental nuisance parameters, the $qq \to W^+W^-$ and $t\bar{t}/tW$ shape nuisance parameters can be constrained by the information in the signal-free region, *i.e.*, high m_{T} and/or high $m_{\ell\ell}$. But, the degree by which they are constrained is still questionable. In section 10.1.1 it was confirmed that switching central and alternate shape of $qq \to W^+W^-$ does not affect the final result. In the fit's point of view switching the central and an alternate shape is effectively forcing 1σ variation with respect to the default configuration. Table 10.4 shows that the result is not affected by switching the $qq \to W^+W^-$ central shape to two alternate shapes, MC@NLO and Powheg. This means that even though these nuisance parameters were not constrained, the final result would not change. They have small impact on the final quantities we measure. So, we conclude that the over-constraint of the these nuisance parameters is acceptable.

10.2.2 Post-fit Result of Normalization

In this section we check if there is any process of which post-fit normalization changes dramatically. A large variation which touches the boundary of the allowed uncertainty range might indicate that the fit model is not correct.

Tables 10.5 - 10.12 show the pre-fit and post-fit normalizations of each process, and the difference in absolute yield and relative fraction in %, in each categories. In general most of the processes are stable, but $gg \to W^+W^-$has a

Table 10.5: The pre-fit and post-fit normalizations in $e\mu/\mu e$ 0-jet category in 8 TeV.

Process	N(pre-fit)	N(post-fit)	Difference(raw)	Difference(%)
$qq \to ZH$	1.7	1.3	-0.4	-23.9
$qq \to WH$	5.7	4.5	-1.2	-20.4
$qq \to qqH$	2.9	2.2	-0.7	-23.9
$gg \to H$	228.1	169.6	-58.5	-25.6
$qq \to W^+W^-$	3981.7	3935.7	-46.0	-1.2
$gg \to W^+W^-$	211.3	292.1	80.8	38.3
WZ/ZZ	132.6	132.2	-0.4	-0.3
$t\bar{t}/tW$	499.4	422.9	-76.4	-15.3
$\mathrm{W}(\mathrm{e}\nu_\mathrm{e}) + \mathrm{jets}$	284.7	250.6	-34.1	-12.0
$W\gamma$	115.6	107.8	-7.7	-6.7
$W\gamma^*$	130.7	119.0	-11.7	-8.9
$Z \to \tau\tau$	46.0	47.2	1.3	2.7
$\mathrm{W}(\mu\nu_\mu) + \mathrm{jets}$	332.2	249.0	-83.2	-25.0

large change. This is driven by the one theory nuisance which was pulled by 1σ. This nuisance is anti-correlated with those of $qq \to W^+W^-$ and $t\bar{t}/tW$, which are moved down by the fit. In $ee/\mu\mu$ channel the shift is less than 10 % of signal, for example, Table 10.7 shows that the shift is 3.2 which is about 4 % of the signal. In $e\mu/\mu e$ channel the shift is as large as 50 % of signal, but its shape is different from that of signal as shown in Figure 5.7 and 5.8 and thus the shift in the region where signal is populated is negligible. Therefore, it does not affect the final result of statistical interpretation.

This chapter is a reprint of the material as it appears in CMS Collaboration, "Measurement of Higgs boson production and properties in the WW decay channel

Table 10.6: The pre-fit and post-fit normalizations in $e\mu/\mu e$ 1-jet category in 8 TeV.

Process	N(pre-fit)	N(post-fit)	Difference(raw)	Difference(%)
$qq \to ZH$	1.9	1.5	-0.4	-21.7
$qq \to WH$	6.9	5.4	-1.5	-21.8
$qq \to qqH$	11.1	8.5	-2.7	-24.0
$gg \to H$	88.5	68.8	-19.8	-22.3
$qq \to W^+W^-$	1203.1	1279.0	75.9	6.3
$gg \to W^+W^-$	69.0	88.4	19.4	28.1
WZ/ZZ	116.6	114.8	-1.8	-1.6
$t\bar{t}/tW$	1436.8	1348.7	-88.1	-6.1
$\mathrm{W}(\mathrm{e}\nu_\mathrm{e}) + \mathrm{jets}$	130.4	120.0	-10.4	-8.0
$W\gamma$	29.1	27.7	-1.3	-4.6
$W\gamma^*$	20.0	11.1	-8.9	-44.5
$Z \to \tau\tau$	76.8	78.9	2.2	2.9
$\mathrm{W}(\mu\nu_\mu) + \mathrm{jets}$	153.0	124.0	-29.0	-19.0

with leptonic final states", J. High Energy Phys. 01 (2014) 096. The dissertation author was the primary investigator and author of this paper.

Table 10.7: The pre-fit and post-fit normalizations in $ee/\mu\mu$ 0-jet category in 8 TeV.

Process	N(pre-fit)	N(post-fit)	Difference(raw)	Difference(%)
$qq \to ZH$	0.1	0.1	-0.0	-23.4
$qq \to WH$	0.5	0.4	-0.1	-23.4
$qq \to qqH$	0.5	0.4	-0.1	-23.4
$gg \to H$	55.8	41.9	-13.9	-25.0
$qq \to W^+W^-$	197.6	201.1	3.5	1.8
$gg \to W^+W^-$	9.8	12.9	3.2	32.3
WZ/ZZ	13.3	13.4	0.1	0.7
$t\bar{t}/tW$	9.3	8.8	-0.5	-5.8
$\mathrm{Z} + \mathrm{jets}$	92.2	100.5	8.3	9.0
$\mathrm{W}(\mathrm{e}\nu_\mathrm{e}) + \mathrm{jets}$	12.4	12.6	0.2	1.6
$W\gamma$	3.2	3.1	-0.1	-4.5
$W\gamma^*$	5.1	4.6	-0.5	-9.7
$\mathrm{W}(\mu\nu_\mu) + \mathrm{jets}$	16.5	17.1	0.6	3.7

Table 10.8: The pre-fit and post-fit normalizations in $ee/\mu\mu$ 1-jet category in 8 TeV.

Process	N(pre-fit)	N(post-fit)	Difference(raw)	Difference(%)
$qq \to ZH$	0.3	0.2	-0.1	-22.9
$qq \to WH$	0.7	0.5	-0.2	-22.9
$qq \to qqH$	2.1	1.6	-0.5	-23.0
$gg \to H$	15.9	12.6	-3.3	-20.6
$qq \to W^+W^-$	37.8	39.6	1.8	4.8
$gg \to W^+W^-$	2.2	3.0	0.8	36.1
WZ/ZZ	6.5	6.6	0.1	1.7
$t\bar{t}/tW$	40.4	40.9	0.5	1.2
$\mathrm{Z} + \mathrm{jets}$	14.7	16.3	1.6	10.8
$\mathrm{W}(\mathrm{e}\nu_\mathrm{e}) + \mathrm{jets}$	2.8	2.9	0.1	2.3
$W\gamma$	2.5	2.4	-0.1	-2.1
$W\gamma^*$	0.7	0.6	-0.1	-8.4
$\mathrm{W}(\mu\nu_\mu) + \mathrm{jets}$	3.7	3.9	0.2	4.8

Table 10.9: The pre-fit and post-fit normalizations in $e\mu/\mu e$ 0-jet category in 7 TeV.

Process	N(pre-fit)	N(post-fit)	Difference(raw)	Difference(%)
$qq \to qqH$	0.4	0.3	-0.1	-23.6
$gg \to H$	50.3	37.6	-12.8	-25.4
$qq \to W^+W^-$	828.8	810.5	-18.3	-2.2
$gg \to W^+W^-$	40.8	52.7	11.9	29.3
WZ/ZZ	17.7	17.8	0.1	0.6
$t\bar{t}/tW$	91.2	99.1	7.9	8.7
$\mathrm{Z} + \mathrm{jets}$	4.9	4.6	-0.4	-7.2
$\mathrm{W}(\mathrm{e}\nu_\mathrm{e}) + \mathrm{jets}$	88.3	84.4	-3.9	-4.4
$W\gamma$	19.7	18.6	-1.1	-5.5
$W\gamma^*$	36.4	35.9	-0.5	-1.5
$\mathrm{W}(\mu\nu_\mu) + \mathrm{jets}$	62.7	45.8	-17.0	-27.0

Table 10.10: The pre-fit and post-fit normalizations in $e\mu/\mu e$ 1-jet category in 7 TeV.

Process	N(pre-fit)	N(post-fit)	Difference(raw)	Difference(%)
$qq \to qqH$	2.1	1.6	-0.5	-23.5
$gg \to H$	17.1	13.6	-3.5	-20.6
$qq \to W^+W^-$	246.3	273.6	27.3	11.1
$gg \to W^+W^-$	14.0	17.7	3.7	26.5
WZ/ZZ	18.1	18.1	-0.0	-0.0
$t\bar{t}/tW$	226.7	203.9	-22.8	-10.0
$\mathrm{Z} + \mathrm{jets}$	8.3	3.2	-5.1	-61.3
$\mathrm{W}(\mathrm{e}\nu_\mathrm{e}) + \mathrm{jets}$	34.4	28.5	-5.9	-17.1
$W\gamma$	3.6	3.4	-0.2	-4.3
$W\gamma^*$	4.8	5.0	0.2	3.5
$\mathrm{W}(\mu\nu_\mu) + \mathrm{jets}$	26.2	19.2	-7.0	-26.7

Table 10.11: The pre-fit and post-fit normalizations in $ee/\mu\mu$ 0-jet category in 7 TeV.

Process	N(pre-fit)	N(post-fit)	Difference(raw)	Difference(%)
$qq \to qqH$	0.1	0.1	-0.0	-22.9
$gg \to H$	10.0	7.5	-2.5	-24.7
$qq \to W^+W^-$	45.0	44.5	-0.4	-1.0
$gg \to W^+W^-$	2.0	2.6	0.6	28.9
WZ/ZZ	0.9	0.9	0.0	1.3
$t\bar{t}/tW$	2.0	2.0	-0.0	-0.3
$\mathrm{Z} + \mathrm{jets}$	10.6	9.7	-0.9	-8.2
$\mathrm{W}(\mathrm{e}\nu_\mathrm{e}) + \mathrm{jets}$	2.1	2.2	0.0	0.9
$W\gamma^*$	0.7	0.7	-0.1	-9.8
$\mathrm{W}(\mu\nu_\mu) + \mathrm{jets}$	1.0	1.1	0.0	1.9

Table 10.12: The pre-fit and post-fit normalizations in $ee/\mu\mu$ 1-jet category in 7 TeV.

Process	N(pre-fit)	N(post-fit)	Difference(raw)	Difference(%)
$qq \to qqH$	0.4	0.3	-0.1	-22.5
$gg \to H$	2.6	2.1	-0.5	-20.1
$qq \to W^+W^-$	9.7	9.9	0.1	1.5
$gg \to W^+W^-$	0.7	0.9	0.2	32.0
WZ/ZZ	0.7	0.7	0.0	1.9
$t\bar{t}/tW$	6.4	6.3	-0.1	-1.4
$\mathrm{Z} + \mathrm{jets}$	9.8	9.8	-0.0	-0.2
$\mathrm{W}(\mathrm{e}\nu_\mathrm{e}) + \mathrm{jets}$	0.7	0.7	0.0	1.6
$W\gamma^*$	0.2	0.2	-0.0	-8.8
$\mathrm{W}(\mu\nu_\mu) + \mathrm{jets}$	0.2	0.2	0.0	3.4

Chapter 11

Results

The question we have is whether the SM Higgs boson exists or not. This chapter discusses the interpretation of data to answer this question.

For the cut-based method, the table of yields and the distributions of m_{T} and $m_{\ell\ell}$ after Higgs selection are shown. For the shape-based method, the 2-dimensional and the $\frac{S}{S+B}$-weighted 1-dimensional data - backgrounds plots are shown to see by eye the compatibility of data with the m_{H}=125 GeV hypothesis. Then, the exclusion limit and the significance are shown combining shape-based and cut-based results in $e\mu/\mu e$ and $ee/\mu\mu$ categories, respectivley, using 7 and 8 TeV data. Finally, the measurement on the production rate in terms of the signal strength is discussed.

11.1 Cut-based Method results

Figure 11.1 and 11.2 show the m_{T} and $m_{\ell\ell}$ distributions with the m_{H}=125 GeV selection in the 0-jet and 1-jet categories, respectively. All lepton final states, $e\mu/\mu e$, $ee/\mu\mu$ and inclusive category from the top, are shown. Table 11.1 and 11.2 show the yields of each process and the corresponding uncertaintiesin 7 and 8 TeV, respectively. The data shows a good agreement with the assumption of m_{H}=125 GeV Higgs boson.

11.2 Shape-based method results

Figure 11.3 shows the 2-dimensional templates of the post-fit signal and the data subtracted by the post-fit backgrounds in the 0-jet and 1-jet categories. The plots show only the region where signal is populated($60 < m_{\mathrm{T}} < 120$ GeV and $12 < m_{\ell\ell} < 100$ GeV). The data - background plots in 0-jet show good agreement with the m_{H}=125 GeV plot. The 1-jet plot do not have enough data to draw any conclusions.

Figure 11.4 shows the stacked and data - backgrounds m_{T} and $m_{\ell\ell}$ distributions using the post-fit results of shape-based method in $e\mu/\mu e$ final states combining 7 and 8 TeV. Before projected to $m_{\ell\ell}$ and m_{T} each bin of 2-dimensional

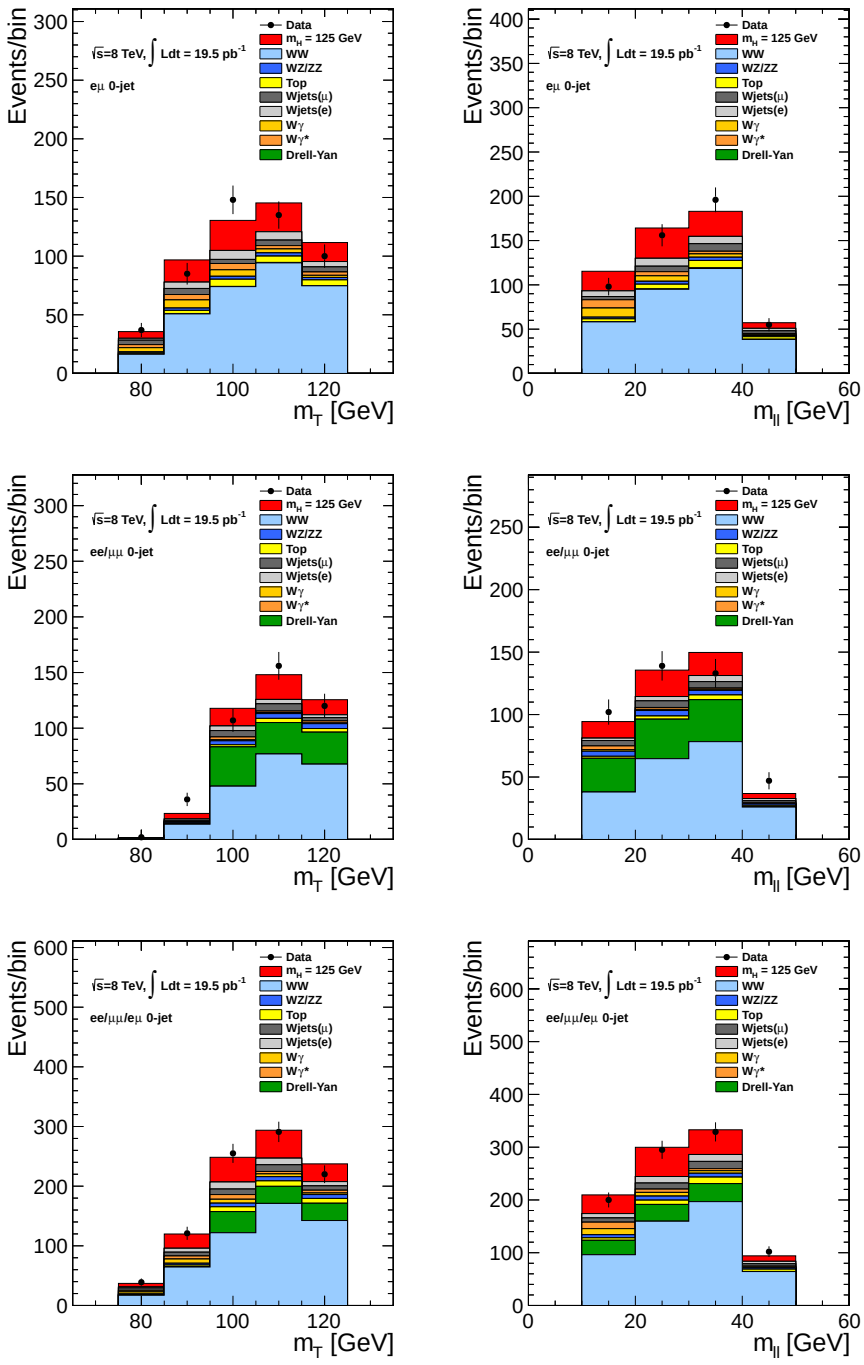


Figure 11.1: m_{T}(left) and $m_{\ell\ell}$(right) distribution for m_{H}=125 GeV analysis in 0-jet category. Top is for $e\mu/\mu e$, middle is for $ee/\mu\mu$, and the bottom is for the inclusive channel.

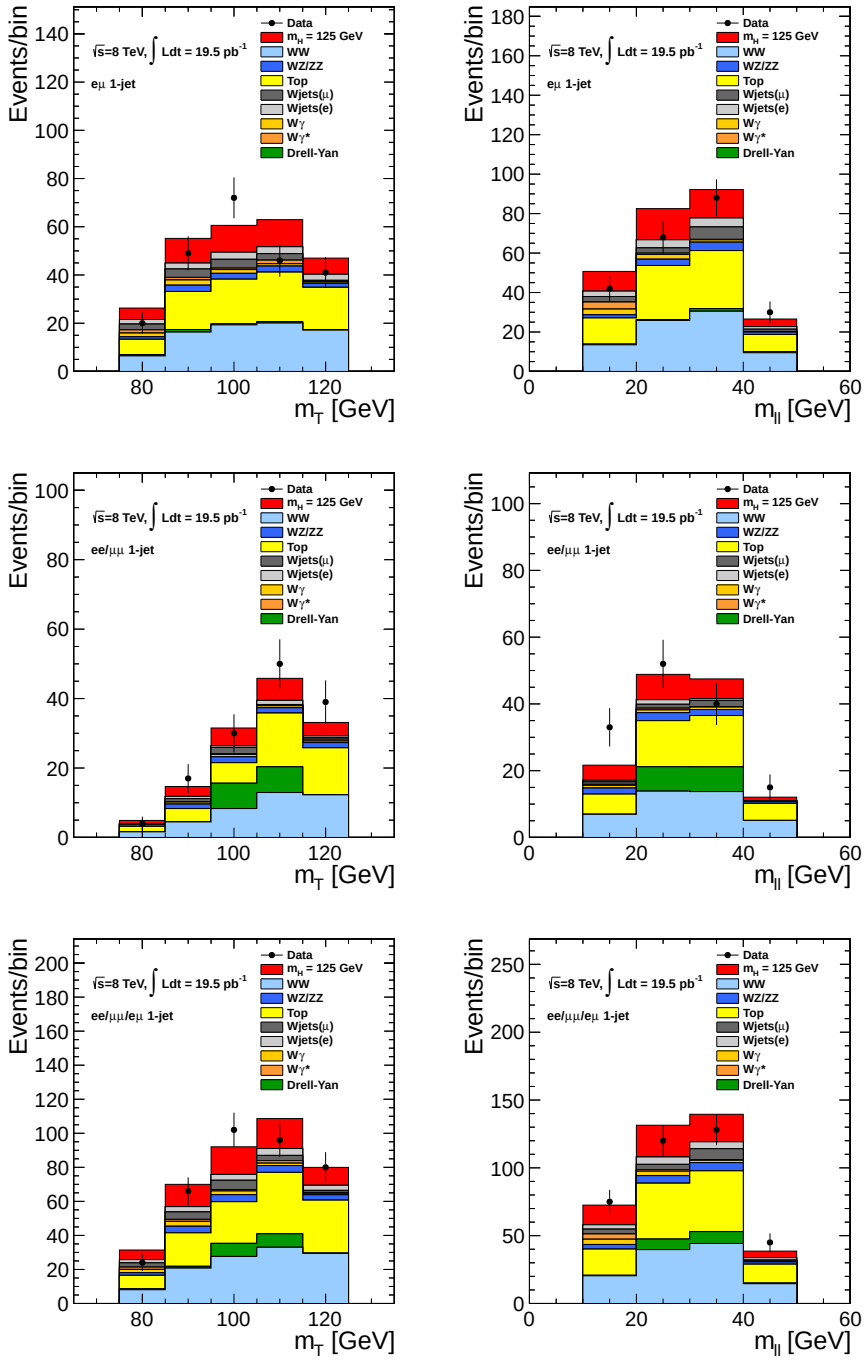


Figure 11.2: m_{T}(left) and $m_{\ell\ell}$(right) distribution for m_{H}=125 GeV analysis in 1-jet category. Top is for $e\mu/\mu e$, middle is for $ee/\mu\mu$, and the bottom is for the inclusive channel.

Table 11.1: Table of yields in cut-based analysis at 7 TeV in 4.9 fb^{-1}. All final states are shown separately. Both yields and uncertainties are shown. Z+jets includes $Z \to \tau\tau$.

process	qqH	ggH	qqWW	ggWW	VV	Top	Z+jets	W(e)+jets	$W\gamma$	$W\gamma^*$	W(μ)+jets	$\sum$Bkg	Data
						m_{H}= 124 GeV							
$e\mu/\mu e$ 0-jet	0.2 ± 0.0	17.8 ± 3.9	68.0 ± 7.3	3.3 ± 1.1	1.5 ± 0.2	3.9 ± 0.9	0.5 ± 0.4	6.5 ± 2.5	3.5 ± 2.7	2.9 ± 1.3	4.2 ± 1.8	94.3 ± 8.6	99
$e\mu/\mu e$ 1-jet	0.7 ± 0.1	6.3 ± 2.1	18.1 ± 3.2	1.0 ± 0.4	1.7 ± 0.2	13.1 ± 1.0	0.6 ± 0.4	3.1 ± 1.3	1.0 ± 1.0	1.0 ± 0.6	3.0 ± 1.4	42.7 ± 4.1	48
$ee/\mu\mu$ 0-jet	0.1 ± 0.0	8.1 ± 1.8	40.9 ± 4.5	1.7 ± 0.5	0.8 ± 0.1	1.8 ± 0.5	10.5 ± 4.1	2.0 ± 0.8	0.0 ± 0.0	0.9 ± 0.5	1.1 ± 0.7	59.7 ± 6.3	60
$ee/\mu\mu$ 1-jet	0.3 ± 0.0	2.3 ± 0.8	8.7 ± 1.6	0.6 ± 0.2	0.6 ± 0.1	5.7 ± 0.5	9.7 ± 4.0	0.6 ± 0.3	0.0 ± 0.0	0.3 ± 0.3	0.1 ± 0.4	26.4 ± 4.4	29
						m_{H}= 160 GeV							
$e\mu/\mu e$ 0-jet	0.9 ± 0.1	73.4 ± 16.0	40.5 ± 4.4	4.4 ± 1.4	0.7 ± 0.1	4.1 ± 1.0	0.0 ± 0.0	1.3 ± 0.6	0.9 ± 1.0	0.3 ± 0.2	1.2 ± 0.8	53.4 ± 5.0	59
$e\mu/\mu e$ 1-jet	4.0 ± 0.4	33.1 ± 10.5	16.7 ± 3.0	1.5 ± 0.5	1.0 ± 0.1	14.1 ± 1.1	0.1 ± 0.0	1.2 ± 0.6	0.0 ± 0.0	0.3 ± 0.2	0.4 ± 0.5	35.3 ± 3.3	32
$ee/\mu\mu$ 0-jet	0.6 ± 0.1	59.3 ± 12.9	34.4 ± 3.8	3.2 ± 1.0	0.5 ± 0.1	3.3 ± 0.8	3.4 ± 3.5	0.4 ± 0.3	0.0 ± 0.0	0.6 ± 0.4	0.1 ± 0.4	45.9 ± 5.4	50
$ee/\mu\mu$ 1-jet	2.8 ± 0.3	24.3 ± 7.7	12.0 ± 2.2	1.2 ± 0.4	0.5 ± 0.1	10.2 ± 0.9	9.0 ± 3.8	0.2 ± 0.2	0.0 ± 0.0	0.1 ± 0.2	1.4 ± 0.9	34.7 ± 4.6	47
						m_{H}= 200 GeV							
$e\mu/\mu e$ 0-jet	0.4 ± 0.0	28.3 ± 6.4	64.3 ± 7.1	7.1 ± 2.2	1.0 ± 0.1	11.1 ± 2.5	0.1 ± 0.0	2.5 ± 1.1	0.0 ± 0.0	0.1 ± 0.2	0.3 ± 0.5	86.7 ± 7.9	85
$e\mu/\mu e$ 1-jet	2.3 ± 0.2	13.7 ± 4.1	28.7 ± 5.2	2.7 ± 0.9	1.1 ± 0.1	31.1 ± 2.2	0.2 ± 0.0	1.8 ± 0.8	0.0 ± 0.0	0.0 ± 0.0	0.4 ± 0.6	66.0 ± 5.8	49
$ee/\mu\mu$ 0-jet	0.4 ± 0.0	24.0 ± 5.4	49.8 ± 5.5	5.8 ± 1.8	0.7 ± 0.1	6.9 ± 1.6	3.8 ± 2.6	0.7 ± 0.2	0.0 ± 0.0	0.0 ± 0.0	0.6 ± 0.7	68.3 ± 6.6	70
$ee/\mu\mu$ 1-jet	1.5 ± 0.2	9.8 ± 3.0	20.2 ± 3.7	2.0 ± 0.7	0.7 ± 0.1	20.6 ± 1.5	15.2 ± 5.1	0.4 ± 0.3	0.0 ± 0.0	0.0 ± 0.0	1.7 ± 1.0	60.9 ± 6.6	56
						m_{H}= 400 GeV							
$e\mu/\mu e$ 0-jet	0.2 ± 0.0	11.0 ± 3.0	34.6 ± 3.8	2.6 ± 0.8	0.9 ± 0.1	17.4 ± 3.9	0.1 ± 0.0	3.1 ± 1.2	0.0 ± 0.0	1.4 ± 0.7	0.2 ± 0.5	60.2 ± 5.7	58
$e\mu/\mu e$ 1-jet	0.7 ± 0.1	7.6 ± 2.3	28.8 ± 4.6	1.5 ± 0.5	1.1 ± 0.1	34.2 ± 2.4	1.0 ± 0.7	3.5 ± 1.4	0.0 ± 0.0	0.3 ± 0.3	0.8 ± 0.7	71.2 ± 5.5	60
$ee/\mu\mu$ 0-jet	0.1 ± 0.0	8.8 ± 2.4	25.6 ± 2.9	2.0 ± 0.6	0.5 ± 0.1	11.1 ± 2.5	3.0 ± 0.3	2.2 ± 0.9	0.0 ± 0.0	6.1 ± 2.7	0.2 ± 0.4	50.8 ± 4.8	45
$ee/\mu\mu$ 1-jet	0.5 ± 0.1	5.3 ± 1.6	16.3 ± 2.6	0.9 ± 0.3	0.6 ± 0.1	19.8 ± 1.4	6.6 ± 2.7	1.6 ± 0.7	0.0 ± 0.0	1.6 ± 0.8	0.1 ± 0.4	47.5 ± 4.2	65
						m_{H}= 600 GeV							
$e\mu/\mu e$ 0-jet	0.1 ± 0.0	1.7 ± 0.6	10.1 ± 1.2	0.8 ± 0.3	0.2 ± 0.0	5.3 ± 1.2	0.0 ± 0.0	1.1 ± 0.5	0.0 ± 0.0	0.3 ± 0.1	0.0 ± 0.0	17.9 ± 1.8	16
$e\mu/\mu e$ 1-jet	0.3 ± 0.0	1.5 ± 0.5	10.1 ± 1.7	0.5 ± 0.2	0.4 ± 0.1	9.8 ± 0.8	1.0 ± 0.7	1.7 ± 0.7	0.0 ± 0.0	0.2 ± 0.2	0.6 ± 0.5	24.1 ± 2.2	19
$ee/\mu\mu$ 0-jet	0.1 ± 0.0	1.2 ± 0.5	5.5 ± 0.7	0.6 ± 0.2	0.2 ± 0.0	3.3 ± 0.8	1.0 ± 0.1	0.9 ± 0.4	0.0 ± 0.0	0.0 ± 0.0	0.2 ± 0.3	11.6 ± 1.2	13
$ee/\mu\mu$ 1-jet	0.2 ± 0.0	1.0 ± 0.3	5.2 ± 0.9	0.3 ± 0.1	0.2 ± 0.0	4.9 ± 0.5	0.6 ± 0.1	0.9 ± 0.4	0.0 ± 0.0	0.3 ± 0.3	0.1 ± 0.3	12.6 ± 1.2	16

Table 11.2: Table of yields in cut-based analysis at 8 TeV in 19.5 fb^{-1}. All final states are shown separately. Both yields and uncertainties are shown. Z+jets includes $Z \to \tau\tau$. W+jets includes both W($e\nu_e$) + jets and W($\mu\nu_\mu$) + jets.

process	qqH	ggH	qqWW	ggWW	VV	Top	Z+jets	$W\gamma$	$W\gamma^*$	W+jets	$\sum$Bkg	Data
						m_H= 125 GeV						
$e\mu/\mu e$ 0-jet	1.0 ± 0.1	88.9 ± 19.3	294.1 ± 28.4	16.0 ± 5.0	10.2 ± 1.0	20.0 ± 4.3	1.2 ± 0.2	21.2 ± 9.8	18.3 ± 8.1	47.8 ± 17.6	428.8 ± 34.2	505
$e\mu/\mu e$ 1-jet	4.6 ± 0.5	37.5 ± 12.2	74.5 ± 10.4	4.8 ± 1.6	10.2 ± 1.1	78.8 ± 4.5	2.7 ± 0.4	6.8 ± 4.0	4.5 ± 2.3	25.6 ± 9.7	208.0 ± 14.1	228
$ee/\mu\mu$ 0-jet	0.5 ± 0.1	55.8 ± 12.2	197.2 ± 19.2	9.7 ± 3.1	13.3 ± 1.3	9.3 ± 2.2	92.2 ± 31.0	3.2 ± 2.5	6.1 ± 2.9	28.6 ± 10.7	359.6 ± 37.6	421
$ee/\mu\mu$ 1-jet	2.1 ± 0.2	15.9 ± 5.2	37.5 ± 5.3	2.2 ± 0.8	6.5 ± 1.0	40.4 ± 3.1	14.7 ± 5.3	2.5 ± 1.5	0.8 ± 0.7	6.5 ± 2.8	111.0 ± 8.6	140
						m_H= 160 GeV						
$e\mu/\mu e$ 0-jet	6.2 ± 0.6	371.6 ± 81.1	175.6 ± 17.1	20.7 ± 6.5	5.8 ± 0.6	24.9 ± 5.5	0.1 ± 0.1	4.6 ± 3.3	1.7 ± 1.1	5.8 ± 2.7	239.3 ± 19.5	285
$e\mu/\mu e$ 1-jet	24.4 ± 2.6	181.2 ± 57.6	66.1 ± 9.3	6.6 ± 2.2	7.5 ± 0.9	82.9 ± 4.7	0.5 ± 0.2	0.2 ± 0.2	0.9 ± 0.7	8.7 ± 3.8	173.3 ± 11.0	226
$ee/\mu\mu$ 0-jet	4.7 ± 0.5	321.2 ± 70.1	148.4 ± 14.5	15.9 ± 5.0	9.7 ± 1.0	14.0 ± 3.2	18.8 ± 9.7	1.0 ± 0.7	0.8 ± 0.5	5.7 ± 2.6	214.3 ± 18.6	256
$ee/\mu\mu$ 1-jet	13.7 ± 1.5	103.6 ± 33.0	42.6 ± 6.1	3.7 ± 1.3	5.4 ± 0.8	47.6 ± 3.2	8.8 ± 4.0	1.3 ± 1.0	0.0 ± 0.0	4.1 ± 2.1	113.6 ± 8.3	134
						m_H= 200 GeV						
$e\mu/\mu e$ 0-jet	3.1 ± 0.3	150.5 ± 34.1	286.2 ± 27.8	32.3 ± 10.0	10.1 ± 1.1	54.6 ± 11.2	0.6 ± 0.2	2.1 ± 2.2	1.7 ± 1.3	7.0 ± 3.2	394.6 ± 31.8	471
$e\mu/\mu e$ 1-jet	13.1 ± 1.4	78.1 ± 23.7	117.5 ± 16.5	11.1 ± 3.7	10.4 ± 1.2	187.0 ± 8.3	1.4 ± 0.3	2.6 ± 2.4	0.3 ± 0.3	13.7 ± 5.5	344.0 ± 19.4	421
$ee/\mu\mu$ 0-jet	2.6 ± 0.3	120.5 ± 27.3	230.9 ± 22.5	29.0 ± 9.0	19.5 ± 1.9	41.8 ± 8.7	20.1 ± 7.7	1.8 ± 1.1	1.1 ± 0.8	7.0 ± 3.2	351.1 ± 27.1	390
$ee/\mu\mu$ 1-jet	7.4 ± 0.8	47.9 ± 14.6	77.6 ± 11.0	7.4 ± 2.5	11.1 ± 1.5	120.1 ± 6.2	16.6 ± 5.7	0.0 ± 0.0	0.0 ± 0.0	6.2 ± 2.8	239.0 ± 0.0	261
						m_H= 400 GeV						
$e\mu/\mu e$ 0-jet	1.2 ± 0.1	63.0 ± 17.3	195.6 ± 23.6	14.0 ± 4.5	9.7 ± 1.1	91.9 ± 18.3	0.2 ± 0.1	2.5 ± 2.6	2.7 ± 1.7	9.4 ± 3.6	326.2 ± 30.6	306
$e\mu/\mu e$ 1-jet	4.4 ± 0.5	42.5 ± 12.9	125.1 ± 19.7	7.1 ± 2.4	10.9 ± 1.2	212.6 ± 9.0	2.0 ± 0.4	0.4 ± 0.4	1.0 ± 0.8	17.6 ± 6.9	376.7 ± 22.5	361
$ee/\mu\mu$ 0-jet	1.0 ± 0.1	53.9 ± 14.8	171.5 ± 20.8	10.8 ± 3.5	21.3 ± 2.1	71.8 ± 14.4	30.2 ± 23.4	0.5 ± 0.5	0.8 ± 0.7	7.4 ± 2.8	314.2 ± 34.8	290
$ee/\mu\mu$ 1-jet	3.0 ± 0.3	29.7 ± 9.0	69.8 ± 11.1	4.1 ± 1.4	11.4 ± 1.3	122.3 ± 6.0	21.2 ± 16.5	0.0 ± 0.0	1.1 ± 0.9	5.4 ± 2.6	235.3 ± 0.0	215
						m_H= 600 GeV						
$e\mu/\mu e$ 0-jet	0.6 ± 0.1	10.1 ± 3.7	61.2 ± 7.6	5.1 ± 1.7	4.0 ± 0.5	30.0 ± 6.4	0.1 ± 0.1	2.6 ± 2.6	1.3 ± 0.9	3.4 ± 1.4	107.7 ± 10.6	95
$e\mu/\mu e$ 1-jet	2.1 ± 0.2	8.9 ± 2.8	47.1 ± 7.5	2.6 ± 0.9	4.7 ± 0.6	65.0 ± 4.1	0.8 ± 0.2	0.0 ± 0.0	0.4 ± 0.5	9.1 ± 3.8	129.7 ± 0.0	110
$ee/\mu\mu$ 0-jet	0.5 ± 0.1	8.8 ± 3.2	55.7 ± 6.9	4.4 ± 1.5	7.5 ± 0.8	21.9 ± 4.8	0.0 ± 0.0	0.5 ± 0.5	0.0 ± 0.0	2.7 ± 1.2	92.8 ± 8.7	94
$ee/\mu\mu$ 1-jet	1.3 ± 0.1	5.7 ± 1.8	24.6 ± 4.0	1.7 ± 0.6	4.2 ± 0.6	30.9 ± 2.2	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	1.8 ± 1.1	63.2 ± 0.0	63

templates is weighted by $\frac{S}{S+B}$ using the variable which is not plotted. Mathematically, when projecting to x axis, the weight applied to each 2-dimensional bin is defined as

$$w_{ij} = \frac{\sum_{i=1}^{N_{bin}} s_{ij}}{\sum_{i=1}^{N_{bin}} s_{ij} + b_{ij}}. \tag{11.1}$$

Because the bins are summed in the x direction, the x information is not used when constructing the weight. So, if x and y variables are independent, the projected distribution is not biased by the weight. This is the case of $m_{\ell\ell}$ and m_{T}, which are weakly correlated. Therefore, the projected $m_{\ell\ell}$ and m_{T} plots are unbiased by the weight. After applying the weight, the plots are normalized such that the signal yield is same with the one before applying the weight. The hatched area is the post-fit background uncertainty estimated by toys generated using the post-fit nuisance parameters. The data - backgrounds plots show a clear excess on top of backgrounds and a good agreement with m_{H}=125 GeV hypothesis in terms of normalization and shape.

11.3 Exclusion limit of SM Higgs boson

Following the procedure described in section 8.1, we calculate the 95 % CL_{s} limit on the signal strength, the ratio of observed signal yield to the expected signal yield at a given Higgs mass. The expected median limit and its $1\sigma/2\sigma$ uncertainty bands are shown in yellow and green, respectively, along with the observed limit.

Figure 11.5 shows the exclusion limits of SM Higgs boson combining all categories in 7 and 8 TeV. The top is the result using only cut-based results in all categories. The observed and expected exclusions of SM boson at CL_{s}= 95 % are 132 - 212 and 310 - 550 GeV, and 120 - 480 GeV, respectively. The bottom is the result of the cut-based method in the $ee/\mu\mu$ category and the shape-based method in the $e\mu/\mu e$ category. The observed and expected exclusions of SM boson at CL_{s}= 95 % are 128-600 GeV and 115-575 GeV, respectively.

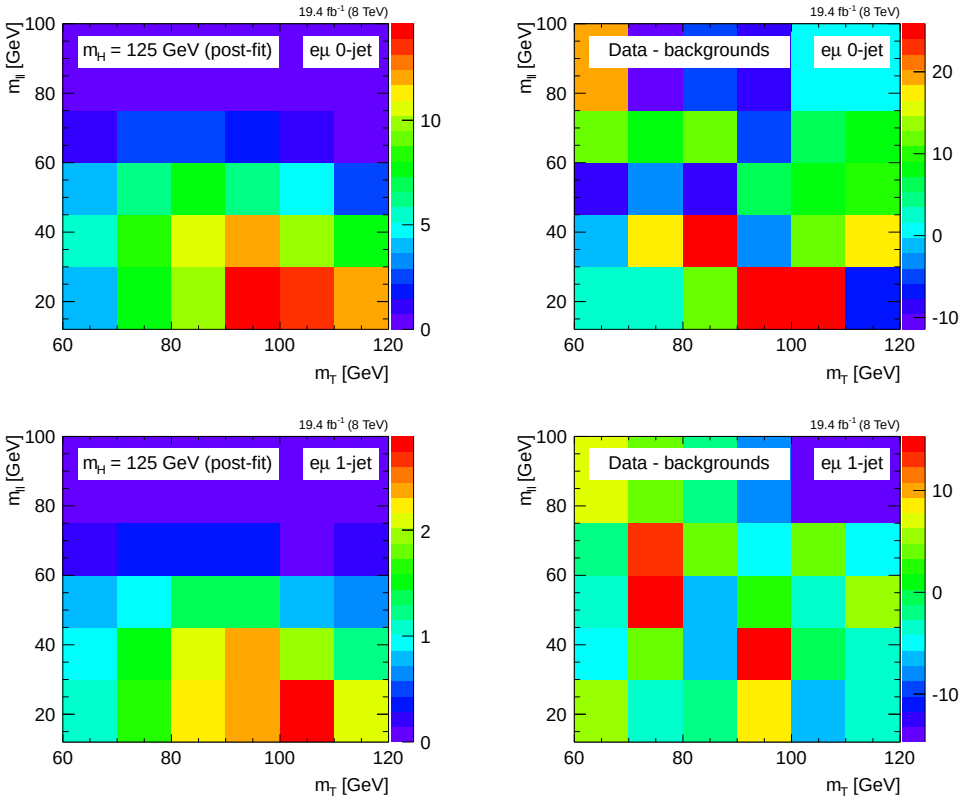


Figure 11.3: 2-dimensional templates of post-fit signal on the left and data subtracted by post backgrounds on the right. The plots show only signal region defined by $60 < m_{\mathrm{T}} < 120$ GeV and $12 < m_{\ell\ell} < 100$ GeV.

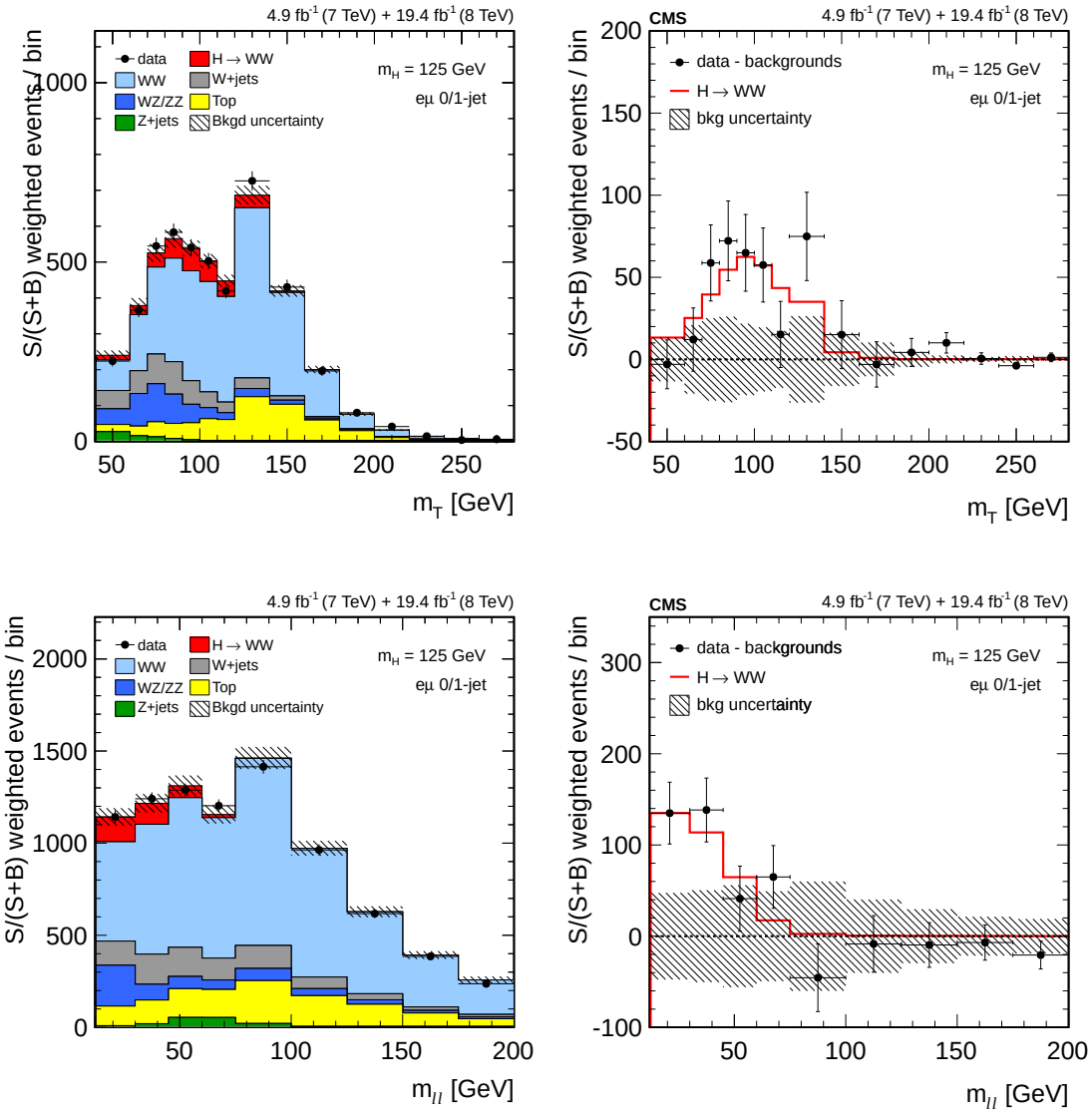


Figure 11.4: Stacked and data - backgrounds m_{T}(top) and $m_{\ell\ell}$(bottom) distributions using post-fit results of shape-based analysis in $e\mu/\mu e$ final states combining 7 and 8 TeV.

Figure 11.6 shows the exclusion limit of the second SM-Higgs-like boson considering a SM Higgs at m_{H}=125 GeV as a background. The cut-based result is used in $ee/\mu\mu$ category and the shape-based result is used in $e\mu/\mu e$ category. The observed and expected exclusions of the second SM-Higgs-like boson at CL_{s}= 95 % are 118-600 GeV and 115-600 GeV, respectively.

11.4 Discovery of a new boson

Following the procedure described in section 8.2, we study the compatibility of data with the background-only hypothesis. The measure is expressed as significance. A Large deviation indicates that there is additional contribution on top of backgrounds. In this section, the expected and observed significances are shown for selective mass points, $m_{\mathrm{H}} = 125, 160, 200, 400$ and 600 GeV.

Tables 11.3 - 11.5 show observed and expected significances at the selected m_{H}, for 7 TeV, 8 TeV and combination of 7 and 8 TeV. The cut-based method is used in $ee/\mu\mu$ category and the shape-based method is used in $e\mu/\mu e$ category. At m_{H}=125 GeV, the observed and expected significances are 4.0σ and 5.2σ, respectively, when all categories are combined.

Figure 11.7 shows the observed and expected significance for low Higgs mass hypotheses($m_{\mathrm{H}} \leq$ 200 GeV). The solid black line represents the expected significance assuming m_{H}= 125 GeV signal. The green and yellow bands represent the $1\sigma/2\sigma$ uncertainty band of the expected significance estimated by pseudo data. The dotted line represents the significance assuming existence of SM Higgs at the given mass. The blue line is the observed significance. The observed data is within 1σ of the expected significance assuming the existence of SM Higgs boson at m_{H}= 125 GeV.

Table 11.6 shows the observed and expected significances using different generators for the $qq \to W^+W^-$ process. Alternative generators, MC@NLO and Powheg, were used replacing the default generator, Madgraph. The result shows that the significance is insensitive to the choice of the default generator, *i.e.*, central shape of the $qq \to W^+W^-$ background.

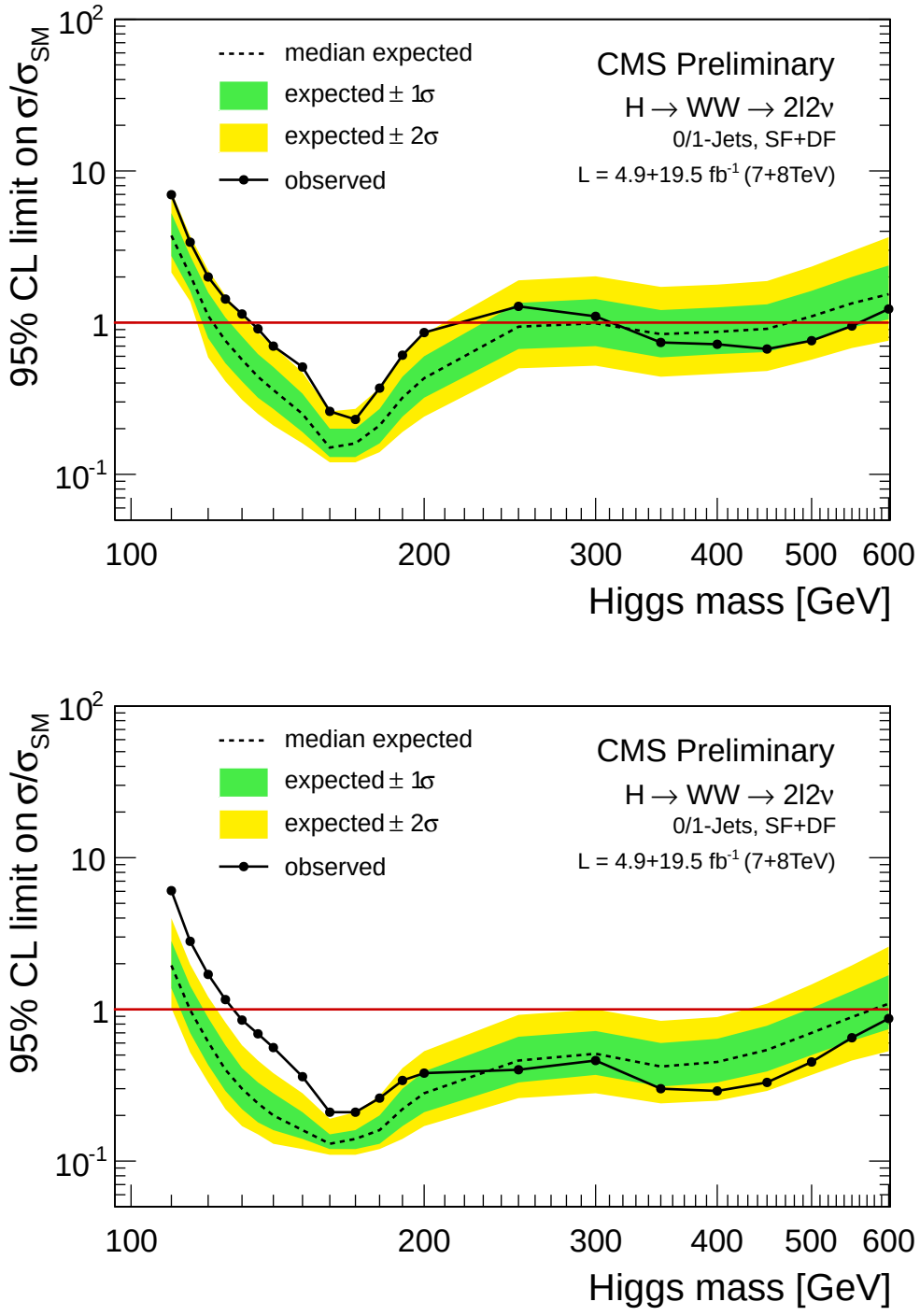


Figure 11.5: Exclusion limits of SM Higgs boson combining all categories in 7 and 8 TeV. The top is the result of the cut-based method in all categories.

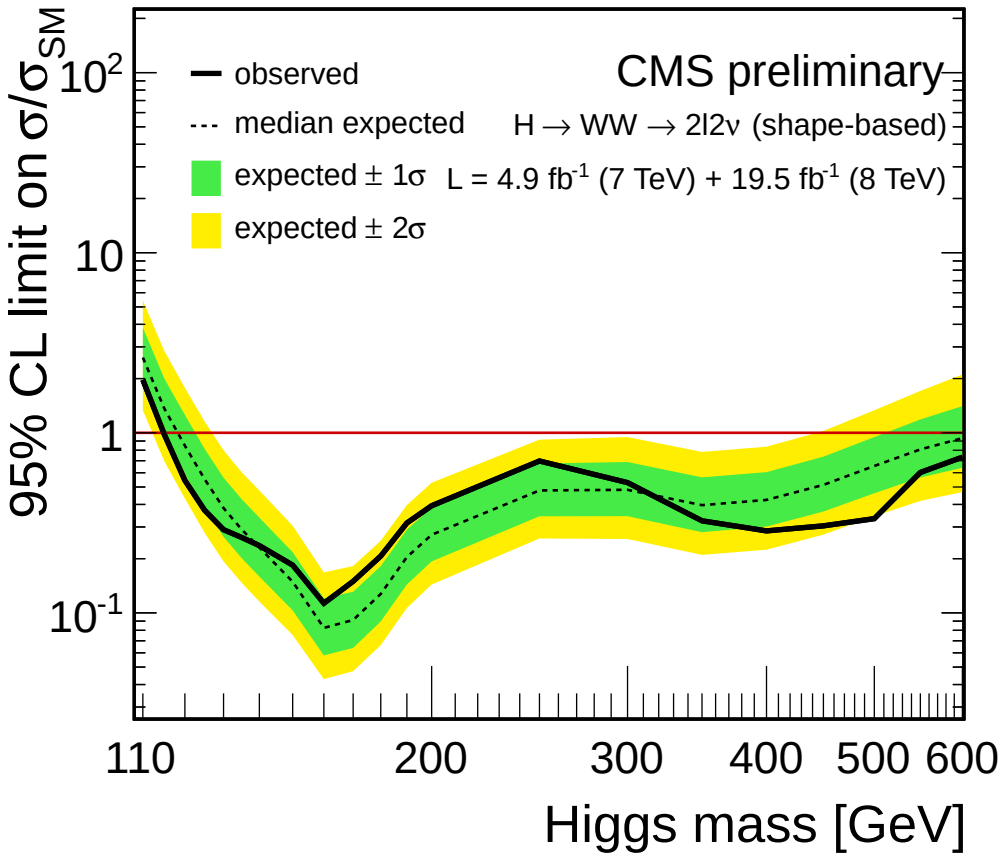


Figure 11.6: Exclusion limit of the second SM-Higgs-like boson considering the SM Higgs at m_{H}=125 GeV as a background.

Table 11.3: Observed and expected significances in 7 TeV. Cut-based analysis is used in $ee/\mu\mu$ final states and shape-based analysis is used in $e\mu/\mu e$ final states

	2D		Cut-based	
Higgs Mass(GeV)	Observed	Expected	Observed	Expected
125	2.3	2.5	0.8	1.7
160	0.9	10.4	0.0	8.2
200	0.0	3.7	0.0	2.9
400	0.2	1.9	0.0	1.5
600	0.0	0.9	0.0	0.8

Table 11.4: Observed and expected significances in 8 TeV. Cut-based analysis is used in $ee/\mu\mu$ final states and shape-based analysis is used in $e\mu/\mu e$ final states

	2D		Cut-based	
Higgs Mass(GeV)	Observed	Expected	Observed	Expected
125	3.5	4.7	2.1	2.6
160	4.1	20.5	0.0	11.5
200	1.4	6.9	2.5	4.3
400	0.0	3.8	0.0	2.1
600	0.0	1.8	0.0	1.3

Table 11.5: Observed and expected significances combining 7 TeV and 8 TeV results. Cut-based analysis is used in $ee/\mu\mu$ final states and shape-based analysis is used in $e\mu/\mu e$ final states.

	2D		Cut-based	
Higgs Mass(GeV)	Observed	Expected	Observed	Expected
125	4.0	5.2	2.1	2.7
160	4.0	22.0	2.8	11.6
200	1.3	7.2	2.5	4.3
400	0.0	4.1	0.0	2.2
600	0.0	2.0	0.0	1.4

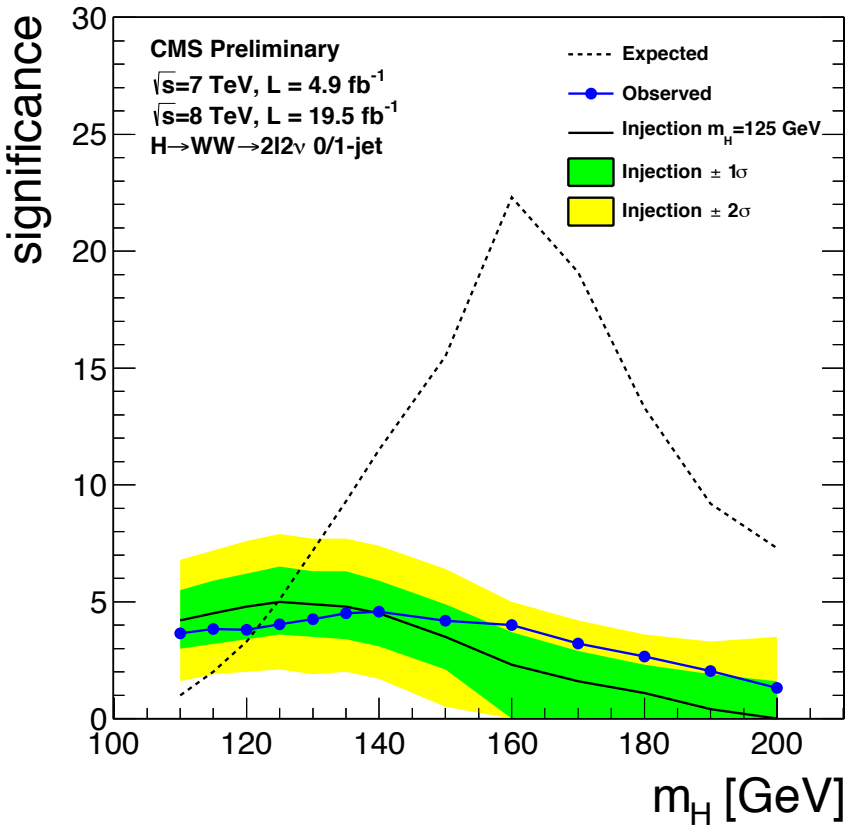


Figure 11.7: Observed and expected significance as a function of m_{H} for the low Higgs mass hypotheses($m_{\mathrm{H}} \leq 200$ GeV).

Table 11.6: The observed and expected significances using different generators for the $qq \to W^+W^-$ process. Alternative generators, MC@NLO and Powheg, were used replacing the default generator, Madgraph.

MC@NLO		Powheg		Madgraph	
Observed	Expected	Observed	Expected	Observed	Expected
4.2	5.3	3.9	5.1	4.0	5.2

11.5 Measurement of Production rate($\sigma \times BR$)

As mentioned before, we measure the signal strength at the measured Higgs mass. The mass measurement comes from $H \to ZZ \to 4l$ and $H \to \gamma\gamma$, and it turned out to be around 125 GeV [90,91]. For the results shown in this section, the cut-based analysis is used in $ee/\mu\mu$ final states and shape-based analysis is used in $e\mu/\mu e$ final states.

Figure 11.8 shows the best fit signal strength as a function of m_{H} for the low Higgs mass hypotheses($m_{\mathrm{H}} \leq 200$ GeV) using all categories in 7 TeV and 8 TeV. The green band shows 1σ error of the global fit. The signal strength is within 1σ of SM Higgs production rate in the range m_{H}= 118 - 125 GeV. Since the exact measured m_{H} is 125.6 GeV [90], we need to know how the signal strength is sensitive to the variation of m_{H} around 125 GeV. The signal strength at m_{H}=126 GeV and m_{H}=124 GeV are 0.82 and 0.72, respectively.

Figure 11.9 shows the $-2\Delta \ln \mathcal{L}$ scan of μ. The black curve represents the case where both systematic and statistical uncertainties are taken into account, while the blue curve represents the case where only statistical uncertainty is considered. To obtain the latter, all nuisances are fixed to the post-fit values, and fitted again. The uncertainty for the blue curve comes solely from statistics of data. The red lines represent the 1σ uncertainty band for each curve. The measured signal strength is 0.76 ± 0.21. One can extract the contribution of systematic uncertainties to the signal strength by subtracting uncertainty in blue from the uncertainty in black. Separating the statistical and systematic uncertainties, the measured

signal strength is $0.76 \pm 0.13(stat.) \pm 0.16(syst.)$.

Figure 11.10 shows the fitted signal strength(μ) for m_{H}=125 GeV in the individual categories. The dotted vertical line and the green band are the central value and the uncertainty band of the signal strength, respectively, obtained from the combination of all categories. The figure shows that all categories are consistent with each other within the uncertainties.

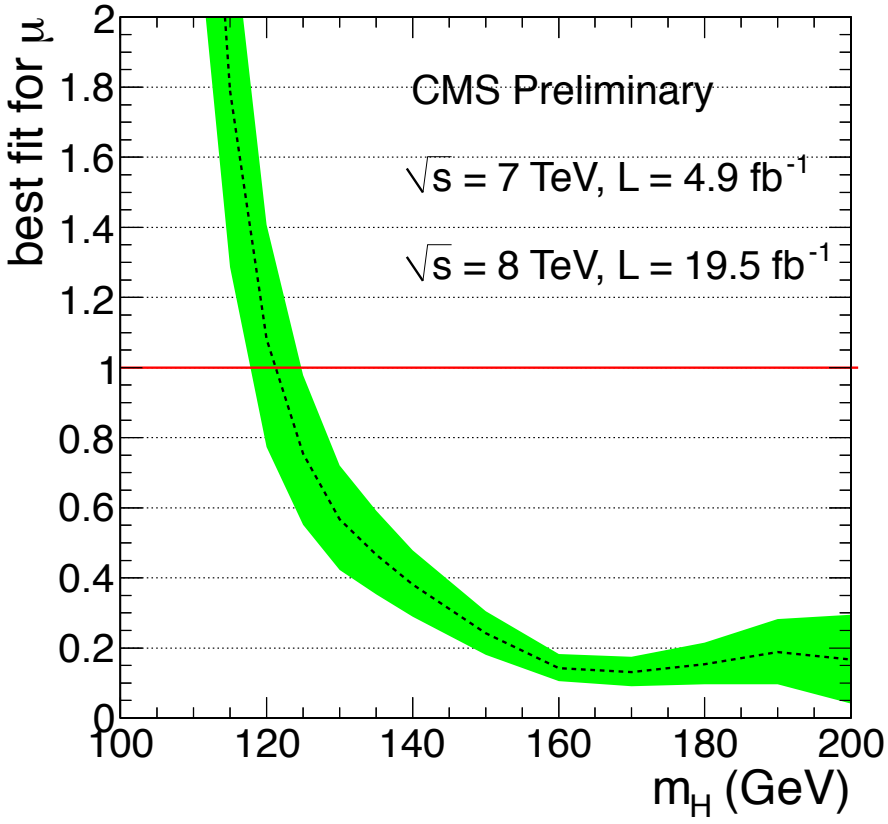


Figure 11.8: The best fit signal strength(μ) as a function of m_{H} for low Higgs mass hypotheses($m_{\mathrm{H}} \leq 200$ GeV). all categories are combined. Cut-based analysis is used in $ee/\mu\mu$ final states and shape-based analysis is used in $e\mu/\mu e$ final states

This chapter is a reprint of the material as it appears in CMS Collaboration, "Measurement of Higgs boson production and properties in the WW decay channel with leptonic final states", J. High Energy Phys. 01 (2014) 096. The dissertation

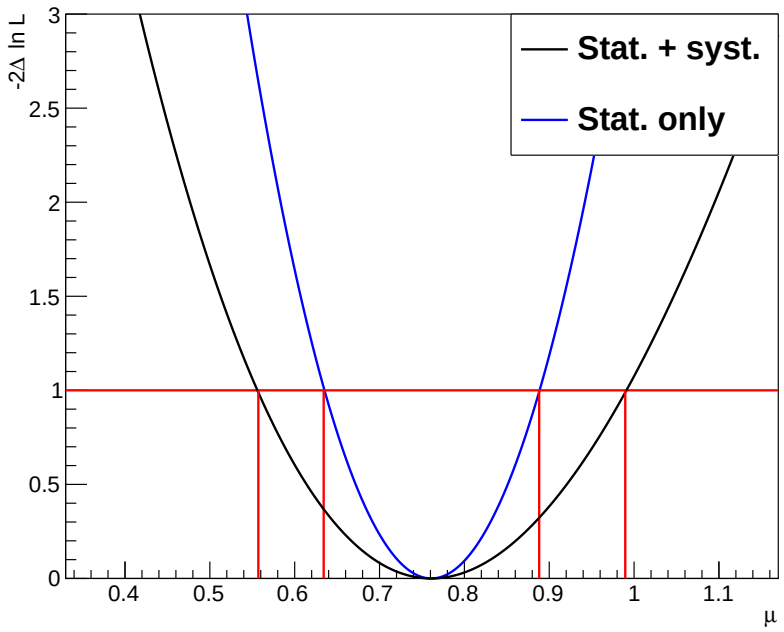


Figure 11.9: $-2\Delta \ln \mathrm{L}$ scan of μ with(black) and without(blue) systematic uncertainty.

author was the primary investigator and author of this paper.

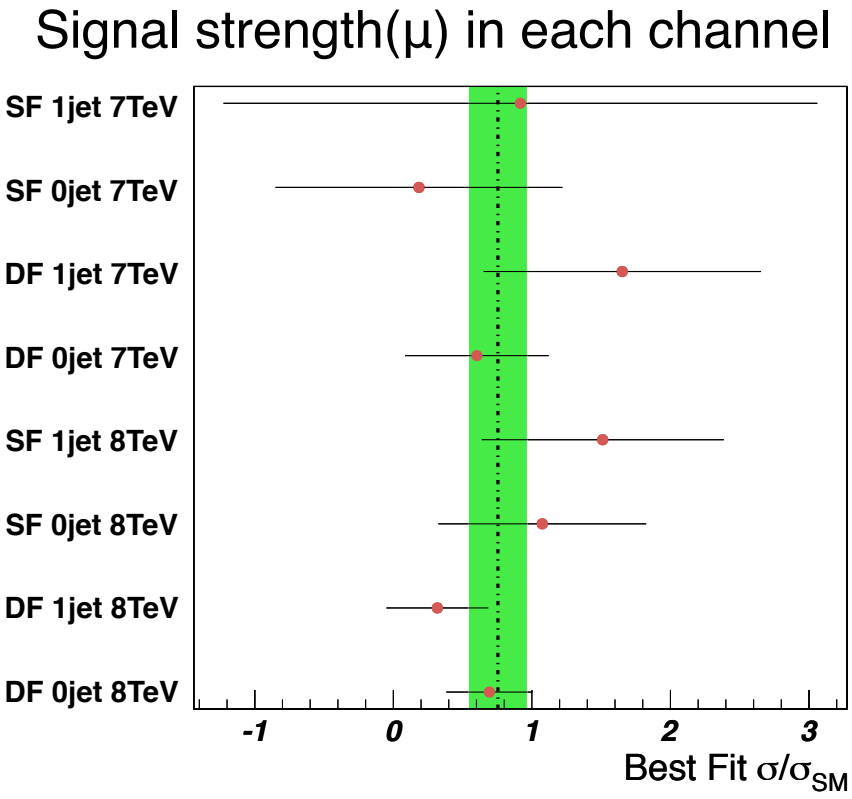


Figure 11.10: Signal strength(μ) for individual categories for m_{H}=125 GeV. Cut-based analysis is used in $ee/\mu\mu$ final states and shape-based analysis is used in $e\mu/\mu e$ final states.

Chapter 12

Study on Spin-Parity of the New Boson

After the discovery of the new boson, the next most important topic is to examine its properties. The SM Higgs is predicted to have $J^{CP} = 0^{++}$ where J is the spin, C is the charge, and P is the parity. In some BSM models such as SUSY, the Higgs sector can be expanded to CP-odd scalar or psedoscalar particles, or a mixture of them that leads to CP-violation. Spin-1 model is excluded because the new boson decays to two massless bosons, *i.e.*, photons, which is not possible for a massive spin-1 resonance due to Yang's theorem [92]. So, we intend to study alternative spin-0 and spin-2 models.

In the reference [93], phenomenological studies of the scattering amplitudes of SM Higgs or exotic boson wth different spin-parity natures are performed. The study shows that $H \to W^+W^- \to 2l2\nu$ channel has a good sensitivity to distinguish the SM Higgs from a graviton-like spin-2 resonance that couples to the WW through a minial couplings. In this chapter this model is denoted as 2^+_{min} and the SM Higgs boson is denoted as 0^+. We also test the pseudo-scalar spin-0 boson which is denoted as 0^-.

This section descirbes the spin-2 model, the helicity argument of WW decay, method of study and the result.

12.1 Models

A general form of the scattering amplitude of a spin-zero resonance that decays to two vector bosons is [93]

$$A(X \to V_1V_2) \;=\; \frac{1}{v^2}\epsilon_1^{\mu*}\epsilon_2^{\nu*}\left(a_1 g_{\mu\nu} m_{\rm H}^2 + a_2 q_\mu q_\nu + a_3 \epsilon_{\mu\nu\alpha\beta} q_1^\alpha q_2^\beta\right) \qquad (12.1)$$

In this parametrization, the SM Higgs has $a_1 \neq 0$ and $a_2 = a_3 = 0$, and the spin-0 pseudo-scalar boson has $a_3 \neq 0$ and $a_1 = a_2 = 0$. The helicity amplitudes are combination of A_{00}, A_{++} and A_{--} in case of SM Higgs boson, and A_{++} and A_{--} for the spin-0 pseudo-scalar boson where A_{00} is both Ws are longitudinally polarized, A_{++} is when both Ws are helicity +1, and A_{--} is when both Ws are helicity -1.

A general form of scattering amplitude of a spin-two resonance that decays to two vector bosons is given in [93]. The minimal coupling scenario where only

two couplings are non-zero gives the scattering amplitude

$$A(X \to V_1 V_2) \quad = \quad \frac{1}{\Lambda^2}\left[2g_2^{(2)} t_{\mu\nu} \frac{q_\alpha q_\beta}{\Lambda^2} f^{(1)\mu\alpha *} f^{(2)\nu\beta *} + 2m_V^2 g_5^{(2)} t_{\mu\nu} \epsilon_1^{\mu *} \epsilon_2^{\nu *}\right] \quad (12.2)$$

where Λ is the energy scale where new physics occur, $g_2^{(2),(5)}$ are the couplings, $t_{\mu\nu}$ is the wave function of X given by a symmetric traceless tensor, q_α is the 4-momentum of X, $f^{(i)\mu\alpha}$ is the field strength tensor given by $\epsilon_i^\mu q_i^\alpha - \epsilon_i^\alpha q_i^\mu$, m_V is the mass of gauge boson, and ϵ_i^μ is the polarization vector of the gauge boson i. This is the model similar to the spin-2 graviton with minimal couplings [93]. Apart from the spin-0 case, the spin-2 amplitude is a combination of all possible helicity amplitudes, so it is not easy to isolate the amplitude to specific helicity amplitudes.

Because of the difference in the spin structure of 0^+, 0^- and 2^+_{min} models, the angular distributions of the decaying particles in these models are different. Figure 12.1 shows the azimuthal angle difference(Φ) between two leptons at the generator level. It shows that the angular distribution is different between models.

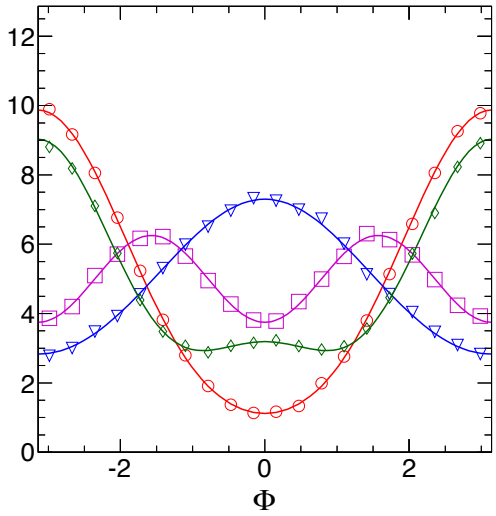


Figure 12.1: The azimuthal angle difference(Φ) between two leptons at the generator level. The red circle is the SM Higgs boson(0^+), the magenta square is the 0^- model, and the blue triangle is the 2^+_{min} model.

The spin-2 sample is generated by the JHUGen [93, 94] generator with the matrix element calculation at a leading-order(LO). The rest of the simulation for parton showering, hadronization, and underlying events is done by PYTHIA [82].

The JHUGen generator is validated by generating the SM Higgs boson events, and comparing it with the result of MCFM [85]. At the reconstruction level, JHUGen+PYTHA is validated by comparing kinematic distributions with the result of POWHEG+PYTIHA. Both comparisons show good agreement [95].

12.2 Test Method

Templates

In order to distinguish between the SM Higgs boson and an alternate model, we first construct the 2-dimensional templates with m_{T} and $m_{\ell\ell}$ for the signal and the background processes in $e\mu/\mu e$ 0-jet and 1-jet categories.. We use the same 2-dimensional templates as used for the SM Higgs search, which are described in detail in section 5.2. At the generator level, $\Delta\phi_{\ell\ell}$ is the best variable that separates the two hypotheses. But, after selections are applied to suppress backgrounds, and the boost of Higgs system is taken into account, the separation power is diluted, and $m_{\ell\ell}$ gives a better separation power.

Figure 12.2 shows the 2-dimensional templates zoomed in the signal region ($60 < m_{\mathrm{T}} < 120$ GeV and $0 < m_{\ell\ell} < 100$ GeV) for SM Higgs boson(0^+), pseudo-scalar boson(0^-) and the 2^+_{min} model in 0-jet and 1-jet categories. In both categories it is clearly seen that the 2^+_{min} and 0^+ hypotheses have different shapes, and this information can be used to discriminate one from the other. On the other hand, the difference between 0^+ and 0^- is not seen as much as the 2^+_{min} case. Thus, the separation between 0^+ and 0^- model is not expected to be as strong as the 2^+_{min} case. For the backgrounds, the exactly same templates are used as in the search analysis. They are shown in Figure 5.7 - 5.10 for 0-jet.

Test statistic

For the statistical interpretation, we construct the test statistic(q). It is defined as the difference of the log likelihood between the two hypotheses, *i.e.*, alternate model(J^P) vs. SM Higgs boson(0^+),

$$q_{J^P} = -2\ln \mathcal{L}_{J^P}/\mathcal{L}_{0^+} \tag{12.3}$$

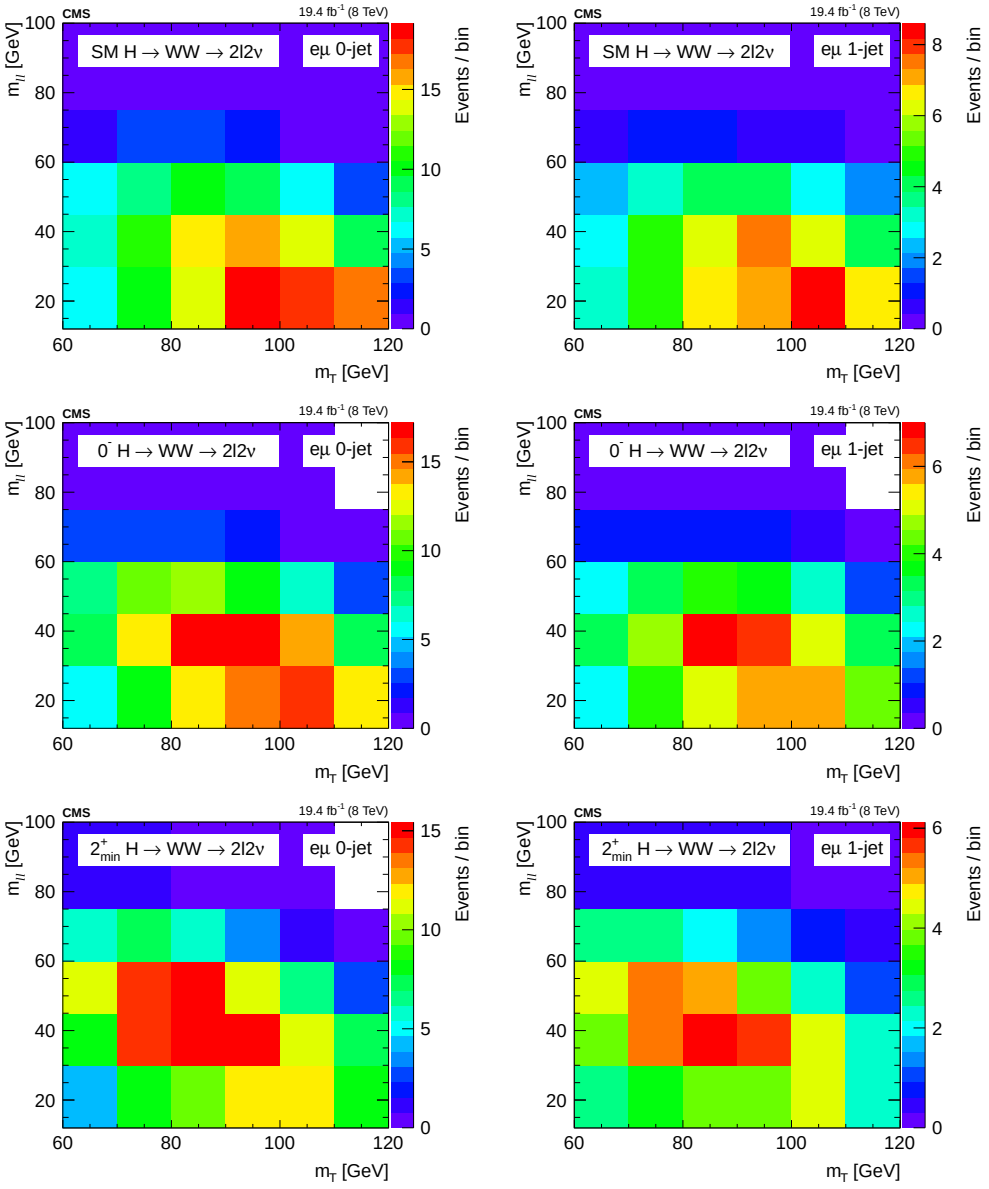


Figure 12.2: 2-dimensional templates for 0^+, 0^- and 2^+_{min} models.

The likelihood is same as the one used in the SM Higgs search :

$$\mathcal{L}(X|\mu,\theta) = \prod_{i}^{N_{bin}} \frac{(\mu s_i(\theta) + b_i(\theta))^{X_i}}{X_i!} e^{-\mu s_i(\theta) - b_i(\theta)} \times \prod_{j}^{N_{nuisance}} p\left(\tilde{\theta}_j|\theta_j\right). \quad (12.4)$$

The only difference in the two likelihoods is the signal component, s_i.

We perform a maximum likelihood fit for each hypothesis with the same dataset, and the difference in the log-likelihood is taken as a test statistic. In the

two fits, the signal strength is allowed to float independently, and the nuisance parameters in the two models are treated independently. The q_{J^P} is calculated using the best-fit values of the nuisance parameters.

Quantifying the separation

The separation is quantified by measuring compatibility of data with the hypothesis under consideration. The expected separation is defined as

$$P(q > q_{0^+}^{\text{expected}} | \text{ alternate model}) \tag{12.5}$$

where $q_{0^+}^{\text{expected}}$ is the peak position of q assuming 0^+. For the observed separation, we can use the observed q from data.

12.3 Results

Because the 2^+_{min} model can be generated via both $gg \to X$ and $q\bar{q} \to X$ modes, and the result depends on the fraction of the two modes, the results is reported as a function of the fraction of the $q\bar{q} \to X$ production mode, $f_{q\bar{q}}$. The 0^- model is generated via only $gg \to X$, so only one result is reported. The alternate models are normalized to the SM Higgs cross section. The $qq \to qqH$ and $qq \to VH$ modes are not tested in this study, *i.e.* the SM Higgs prediction is used for both hypotheses.

Figure 12.3 shows the distribution of $q_{2^+_{min}}$ with full data combining 0-jet and 1-jet categories. Top plots show the result with $f_{q\bar{q}} = 0$ % and the bottom plots show the result with $f_{q\bar{q}} = 100$ %. The left plots show the result using the expected signal signal strength, $\mu = 1$, and the right plots show the result using the best-fit value of the signal strength, $\mu \approx 0.75$. Figure 12.4 shows the expected median, $\pm 1\sigma$(green) and $\pm 2\sigma$(yellow) bands for q assuming 0^+. When the observed μ is used, the expected separation ranges from 1.8σ to 2.9σ as $f_{q\bar{q}}$ goes from 0 to 100 %. The incompatibility of data with 2^+_{min} model ranges from 1.2σ to 3.1σ as $f_{q\bar{q}}$ goes from 0 to 100 %. We can express this result in terms of CL_{s}. The CL_{s} ranges from 16.3 % to 0.2 % as $f_{q\bar{q}}$ goes from 0 to 100 %. All these result shows that data prefers 0^+ to 2^+_{min} hypothesis.

Figure 12.5 shows the distribution of q_{0^-} with full data combining 0-jet and 1-jet categories. The left plots show the result using the expected signal signal strength, $\mu = 1$, and the right plots show the result using the best-fit value of the signal strength, $\mu \approx 0.75$. When the observed μ is used, The expected separation is 0.8σ. As expected in the 2-dimensional templates in Figure 12.2, the sensitivity of separation between 0^+ and 0^- is not as good as the 2^+_{min} case. The observed separation is 1.2σ, and it corresponds to CL_s=34.7 %.

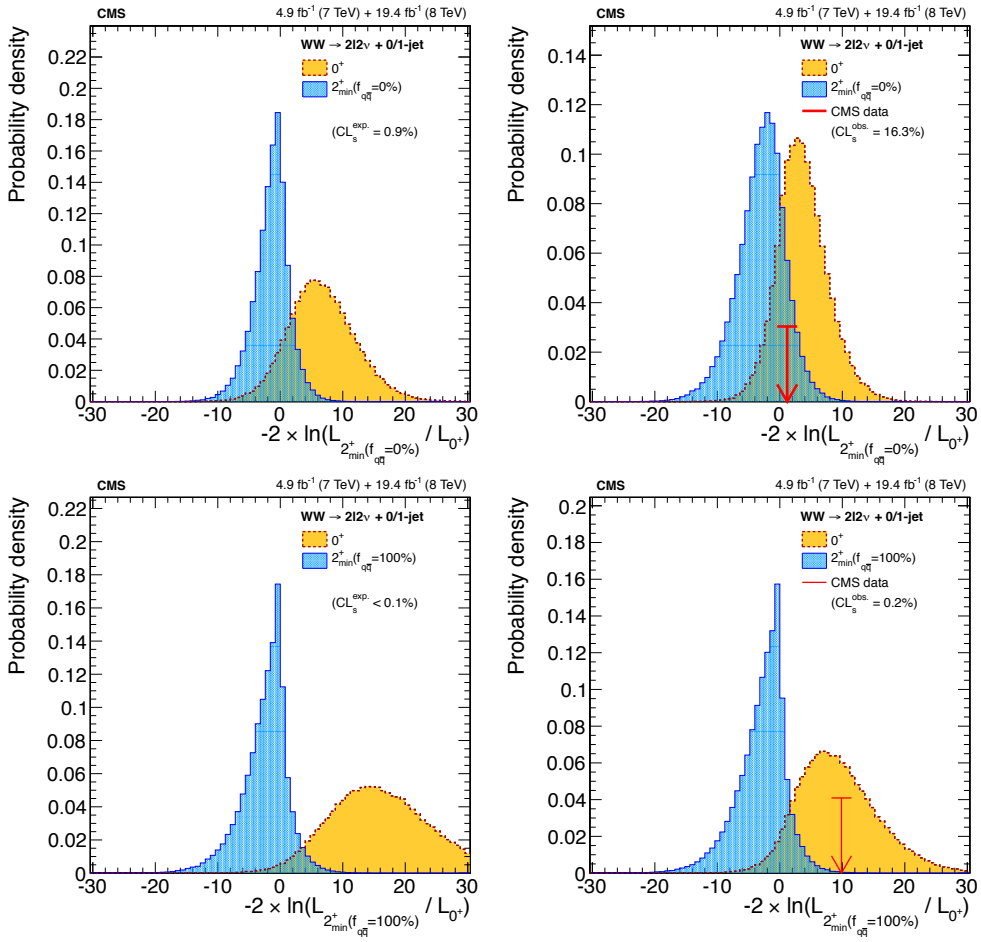


Figure 12.3: Distributions of $q_{2^+_{min}}$ assuming 0^+ and 2^+_{min} models. The blue is the expected distribution assuming 2^+_{min}, and the orange is the expected distribution assuming 0^+ hypothesis.

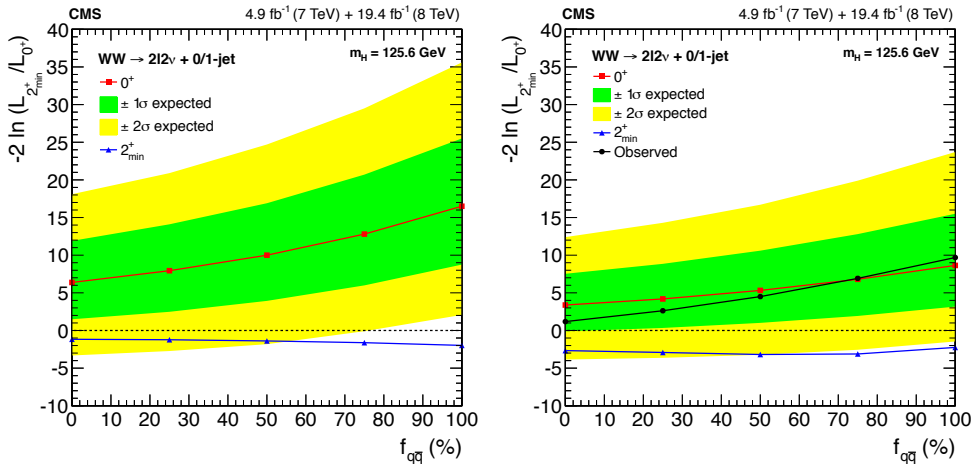


Figure 12.4: Median of $q_{2^+_{min}}$ as a function of $f_{q\bar{q}}$. The blue line is the median of q assuming 2^+_{min}. Left is with $\mu = 1$ and the right is with observed μ. Right plot also shows the q calculated using data.

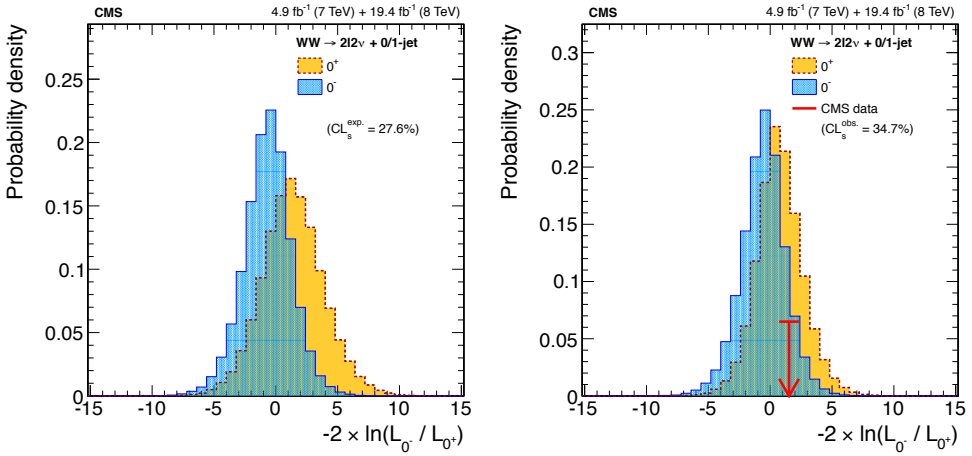


Figure 12.5: Distributions of q_{0^-} assuming 0^+ and 0^- models. Blue is the expected distribution assuming 0^-, and orange is the expected distribution assuming 0^+ hypothesis.

12.4 Conclusion of spin-parity study

A test on the spin-parity nature of the new boson is performed against two alternate models, 2^+_{min} and 0^-. The result shows that data favors 0^+ to 2^+_{min} with CL_s=0.2-16.3 % depending on the fraction of $q\bar{q} \to X$ mode. The CL_s with 0^-

model is 34.7 %.

This chapter is a reprint of the material as it appears in CMS Collaboration, "Measurement of Higgs boson production and properties in the WW decay channel with leptonic final states", J. High Energy Phys. 01 (2014) 096. The dissertation author was the primary investigator and author of this paper.

Chapter 13

Conclusion

Using the 4.9 fb^{-1} and 19.5 fb^{-1} of data collected by the CMS detector in 2011 and 2012 at 7 and 8 TeV center of mass energy, respectively, we have performed a search for the standard model Higgs boson in the range $m_{\rm H}$=110 - 600 GeV and a study on its spin-parity nature at $m_{\rm H}$=125 GeV in the full leptonic final states of W^+W^- decay mode. The analysis is performed in the 4 categories divided by the lepton flavor and the number of jets because the composition of background is different in the different categories.

There are two neutrinos in the final state, so $m_{\rm H}$ is not be able to be reconstructed, thus not measured. We measure an overall excess on top of backgrounds, therefore an accurate and precise estimation of backgrounds is the key in this analysis. For the reliable estimation of the major backgrounds, we apply date-driven methods to all major backgrounds. A new analysis method which uses the 2-dimensional templates of $m_{\rm T}$ and $m_{\ell\ell}$ has been developed, and enhanced the search sensitivity of $m_{\rm H}$=125 GeV SM Higgs boson by 25 %. The new method is used only in the $e\mu/\mu e$ category because of the poor modeling of $Z/\gamma^* \to \ell\ell$ template in the $ee/\mu\mu$ category. The cut-based method is used in the $ee/\mu\mu$ category. The new analysis uses templates constructed using simulation, and we validate the fit model of the major backgrounds using dedicated control regions in data.

The expected exclusion limit of the SM Higgs at $CL_{\rm s}$ = 95 % is 115 − 575 GeV, and we observe an excess in the low $m_{\rm H}$ region giving the observed exclusion limit 128 − 600 GeV. The excess of the data is quantified in terms of imcompatibility of the *background-only* hypothesis. The expected signifance at $m_{\rm H}$=125 GeV is 5.2σ, and we observe 4.0σ. The production rate is measured in the unit of expected $\sigma \times BR$ in SM, and the measured rate is $0.76 \pm 0.13(stat.) \pm 0.16(syst.)$.

The spin-parity nature of the new boson has been tested by performing hypothesis tests using a graviton-like spin-2 and a psedu-scalar spin-0 models. The data favors SM Higgs to the spin-2 and spin-0 models by $CL_{\rm s}$=0.2-16.3 % and 34.7 %, respectively.

The result of this analysis was used for the historic discovery of the new

boson and understanding of its spin-parity nature. In the July 2012, the CMS and the ATLAS collaborations announced the discovery of a new boson [61,96] around m_{H}=125 GeV, which led to the Nobel Prize in Physics in 2013. Figure 13.1 shows

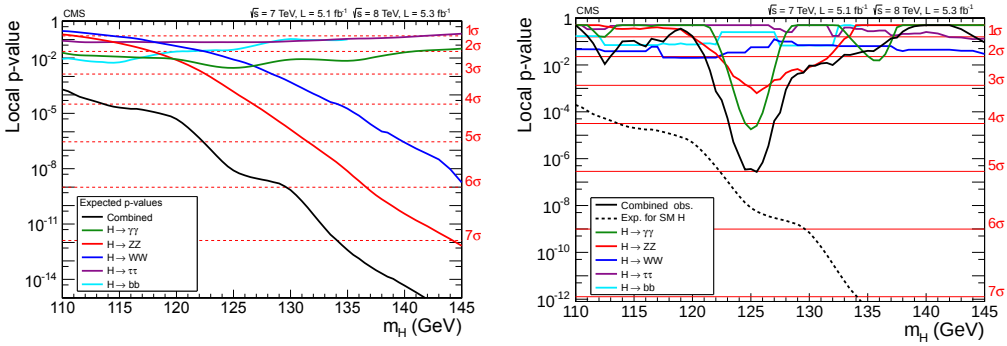


Figure 13.1: Expected(left) and observed(right) significance using data collected by July 2012.

the expected and observed significance from the all data available in July 2012. The $H \to W^+W^- \to 2l2\nu$ mode has a good sensitivity in the higher mass region, but at the mass point where the new boson was discovered, the excess is about 2σ. This analysis was used to exclude large range of high mass region, and narrowed down the search window to the low m_{H} region.

Using the full data collected in 2011 and 2012, the signal strength was measured combining all channels [97]. Figure 13.2 shows the signal strength measured at m_{H}=125.7 GeV by the five most sensitive channels. The $H \to W^+W^- \to 2l2\nu$ channel measures the signal strength at the best precision.

The spin-parity study on the spin-2 model was combined with the $H \to ZZ \to 4l$ channel. Figure 13.3 shows the $q_{2^+_{min}}$ distrbutions in case the spin-2 resonance is produced via gluon-gluon interaction. The combined result shows that the expected and the observed exclusion of the spin-2 model are CL_{s}=1.2 % and 0.6 %, respectively.

In conclusion, the Higgs boson, the last piece of the standard model, was discovered in July 2012, and $H \to W^+W^- \to 2l2\nu$ channel made a significant contribution to the discovery by excluding the SM Higgs hypothesis in the high

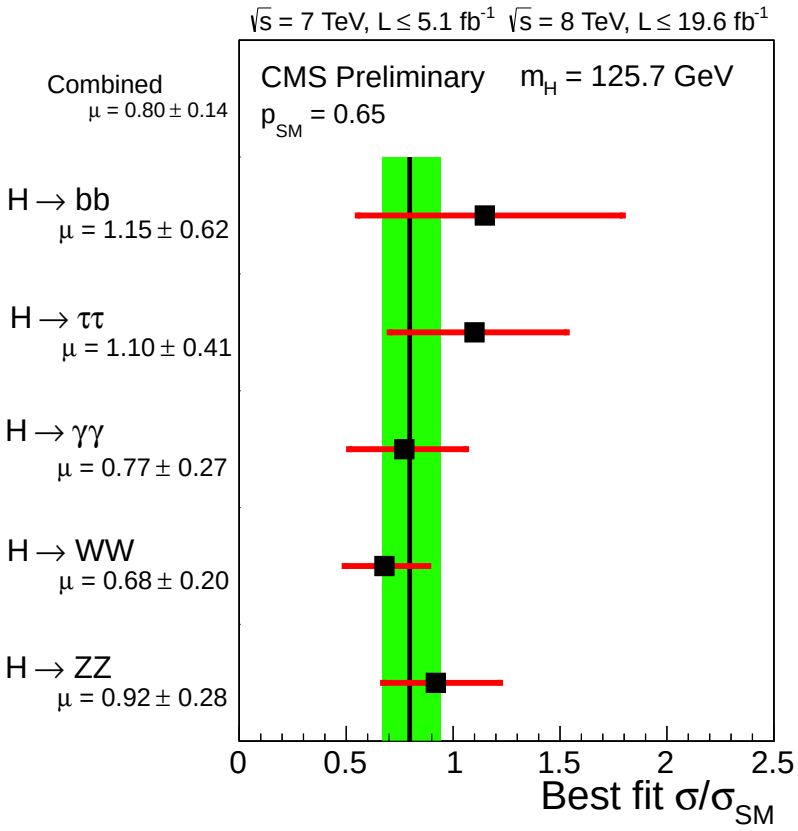


Figure 13.2: Signal strength measured by the five most sensitive channels. The black line and the green bands shows the combined measurement, $\mu = 0.80 \pm 0.14$.

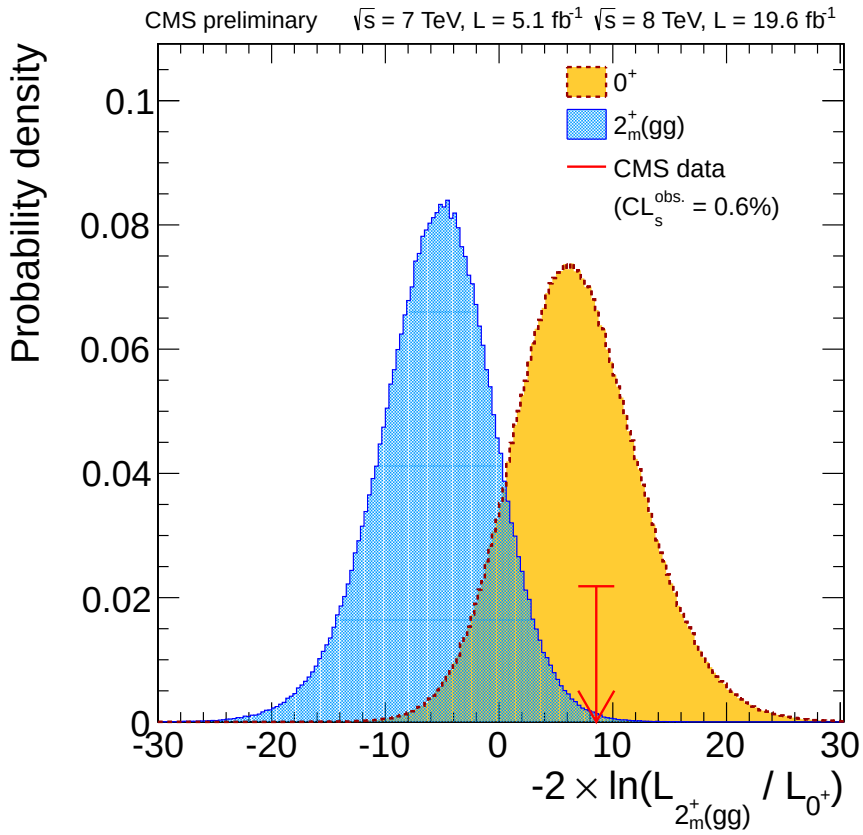


Figure 13.3: Distribution of $q_{2^+_{min}}$ assuming 0^+ and 2^+_{min} models. Blue is the expected distribution assuming 2^+_{min}, and orange is the expected distribution assuming 0^+ hypothesis. The plot shows the result using the best-fit value of the signal strength.

m_{H} region. The measured production rate and the test on the spin-parity shows that the new boson is compatible with the SM Higgs boson at m_{H}=125 GeV.

This chapter is a reprint of the materials as it appears in CMS Collaboration, "Measurement of Higgs boson production and properties in the WW decay channel with leptonic final states", J. High Energy Phys. 01 (2014) 096 and CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", Phys. Lett. B 716, 30 (2012). The dissertation author was the primary investigator and author of this paper.

Appendix A

More 2-dimensional templates

The Figure A.1 - A.4 show the templates in the $e\mu/\mu e$ 0-jet category at 8 TeV and the relative statistical uncertainty of the template with respect to the total background for each process. The same background templates are used for $m_{\mathrm{H}} < 300$ GeV hypotheses.

The Figure A.5 - A.8 show the templates in the $e\mu/\mu e$ 0-jet category at 8 TeV and the relative statistical uncertainty of the template with respect to the total background for each process. The same background templates are used for $m_{\mathrm{H}} \geq 300$ GeV hypotheses.

Figure A.9 - A.12 show the templates in the $e\mu/\mu e$ 0-jet category at 8 TeV and the relative statistical uncertainty of the template with respect to the total background for each process. The same background templates are used for $m_{\mathrm{H}} \geq 300$ GeV hypotheses.

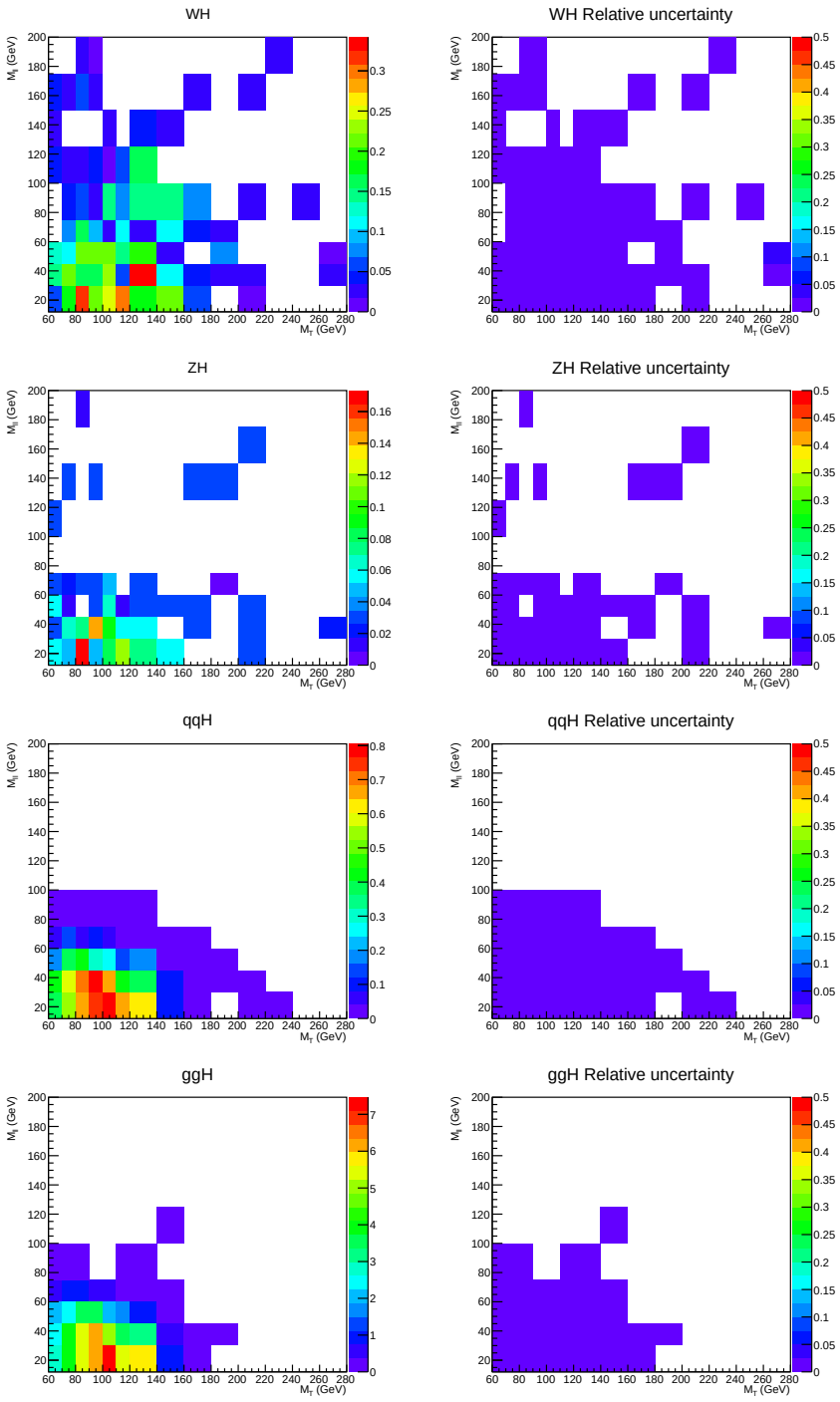


Figure A.1: Templates(left) and relative statistical uncertainty of the MC sample(right) of $qq \to WH$, $qq \to ZH$, $qq \to qqH$ and $gg \to H$. The templates are used for $m_{\mathrm{H}} = 125$ GeV analysis in the 1-jet category.

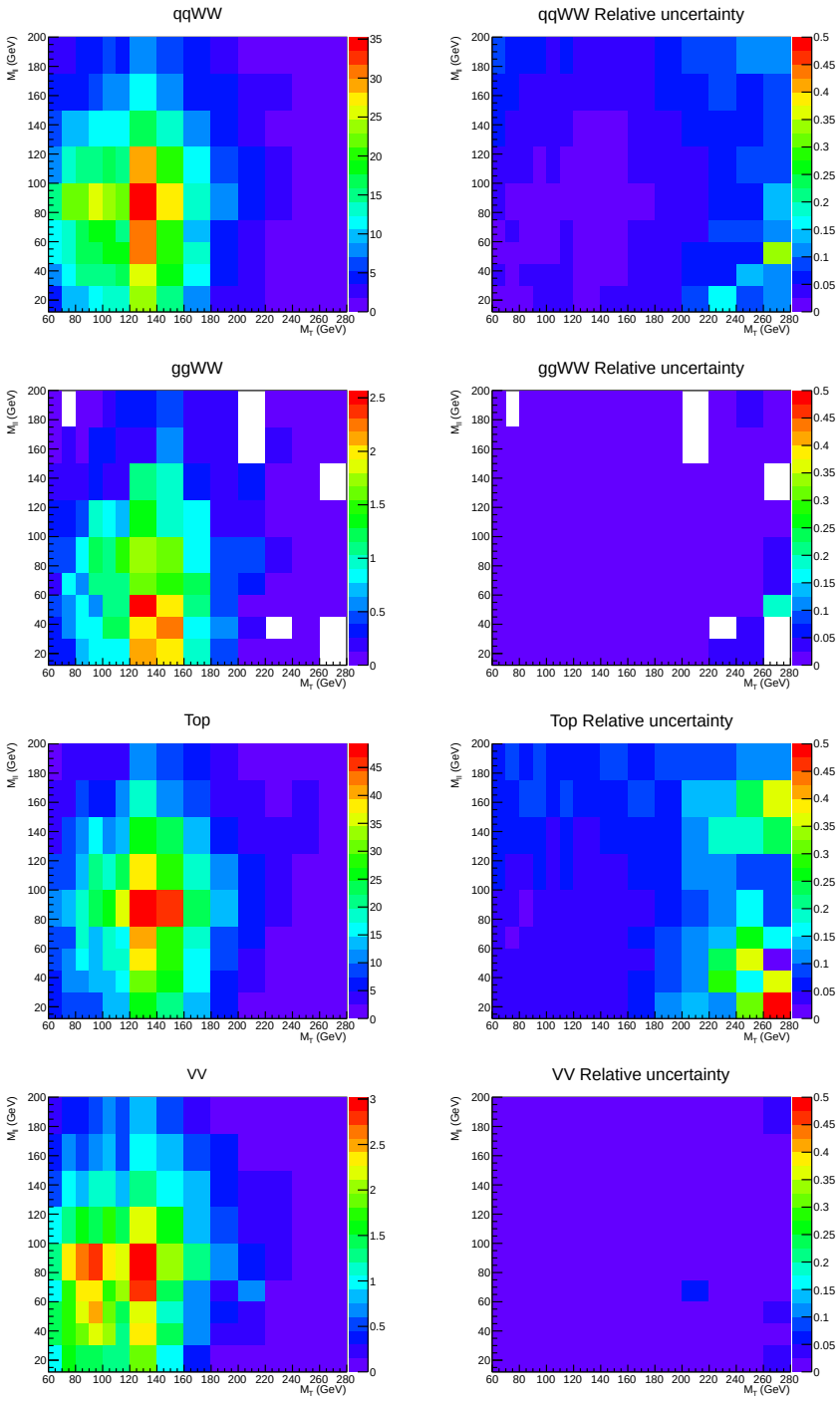


Figure A.2: Templates(left) and relative statistical uncertainty of the MC sample(right) of $qq \to W^+W^-$, $gg \to W^+W^-$, $t\bar{t}/tW$ and WZ/ZZ. The templates are used for $m_{\mathrm{H}} = 125$ GeV analysis in the 1-jet category.

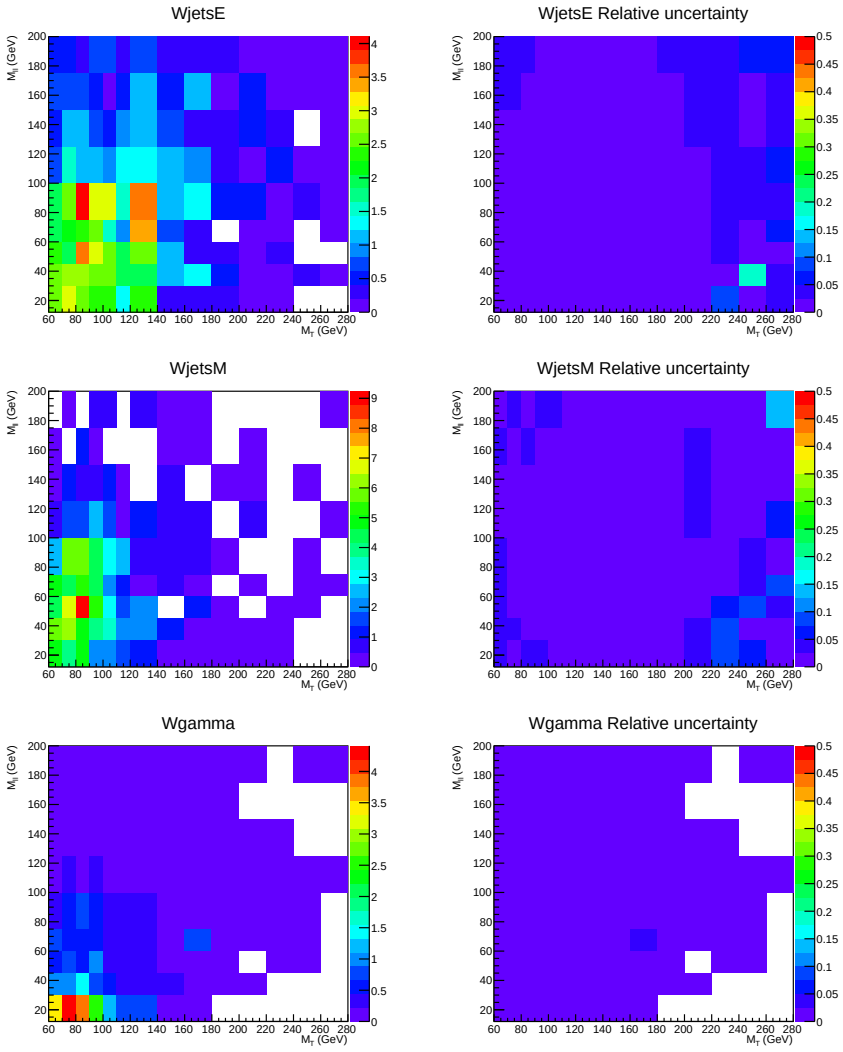


Figure A.3: Templates(left) and relative statistical uncertainty of the MC sample(right) of $\mathrm{W}(\mathrm{e}\nu_{\mathrm{e}}) + \mathrm{jets}$, $\mathrm{W}(\mu\nu_{\mu}) + \mathrm{jets}$ and $W\gamma$. The templates are used for $m_{\mathrm{H}} = 125$ GeV analysis in the 1-jet category.

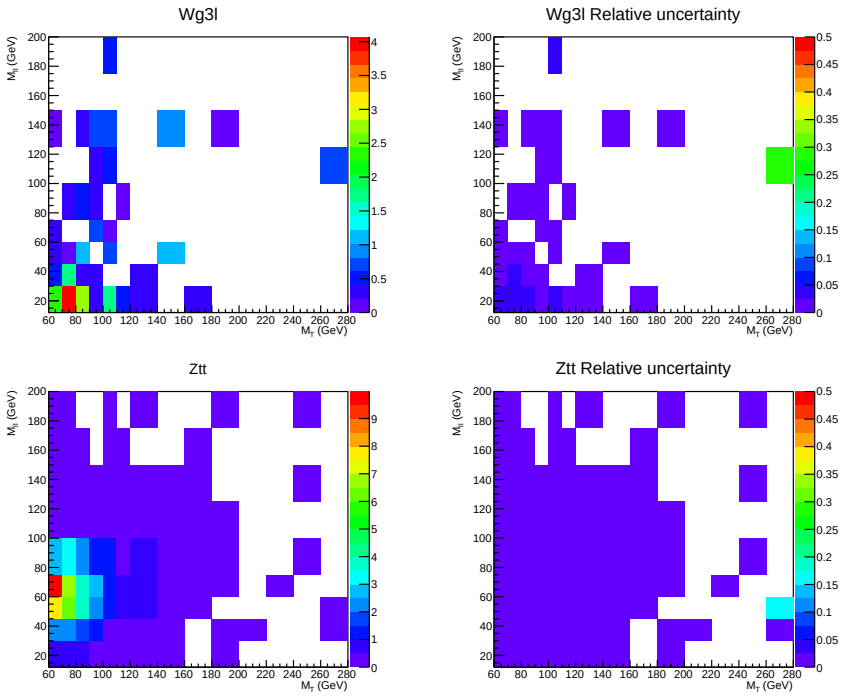


Figure A.4: Templates(left) and relative statistical uncertainty of the MC sample(right) of $W\gamma^*$ and $Z \to \tau\tau$. The templates are used for $m_{\mathrm{H}} = 125$ GeV analysis in the 1-jet category.

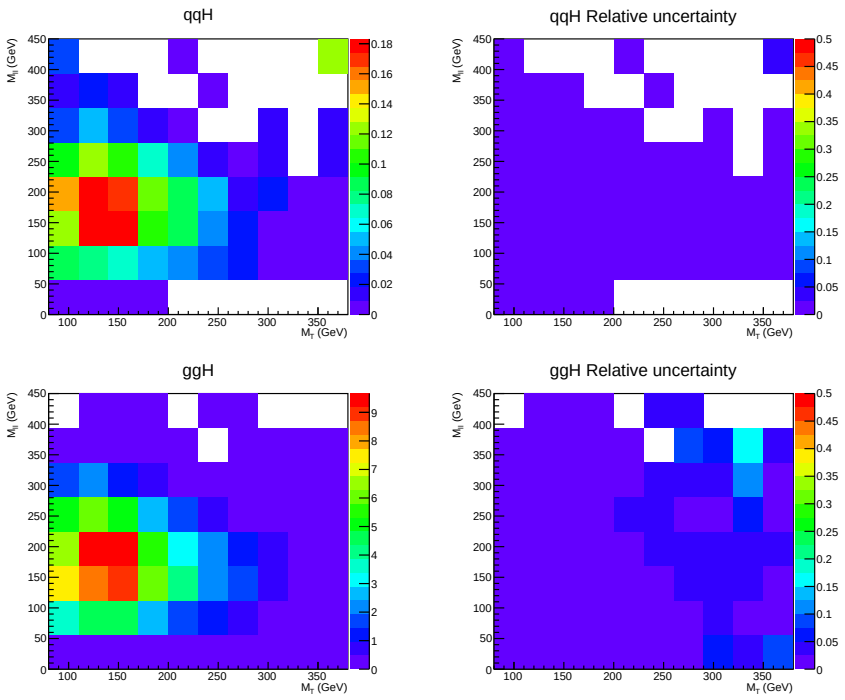


Figure A.5: Templates(left) and relative statistical uncertainty of the MC sample(right) of $qq \to qqH$ and $gg \to H$. The templates are used for $m_{\mathrm{H}} = 400$ GeV analysis in the 0-jet category.

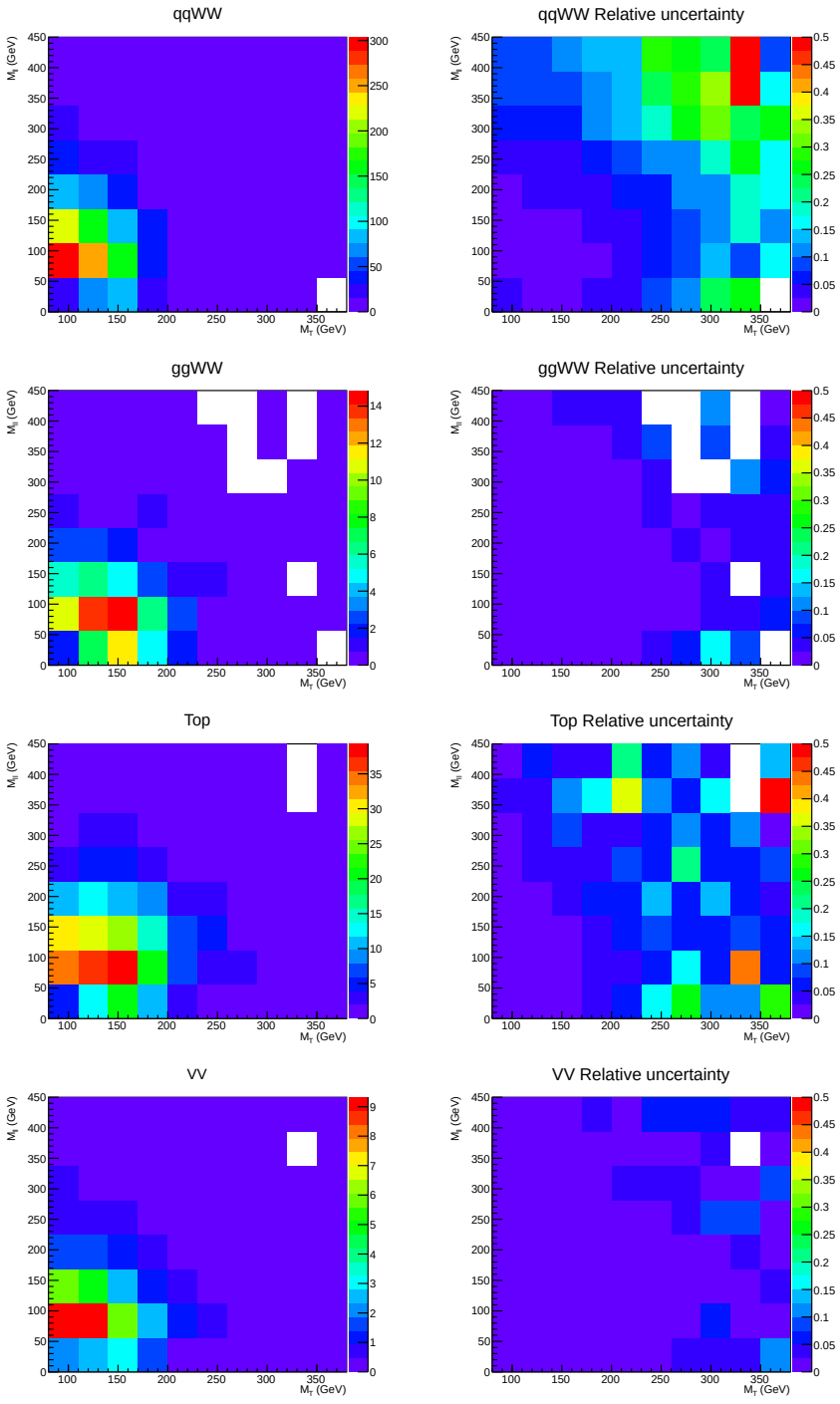


Figure A.6: Templates(left) and relative statistical uncertainty of the MC sample(right) of $qq \to W^+W^-$, $gg \to W^+W^-$, $t\bar{t}/tW$ and WZ/ZZ. The templates are used for $m_{\mathrm{H}} = 400$ GeV analysis in the 0-jet category.

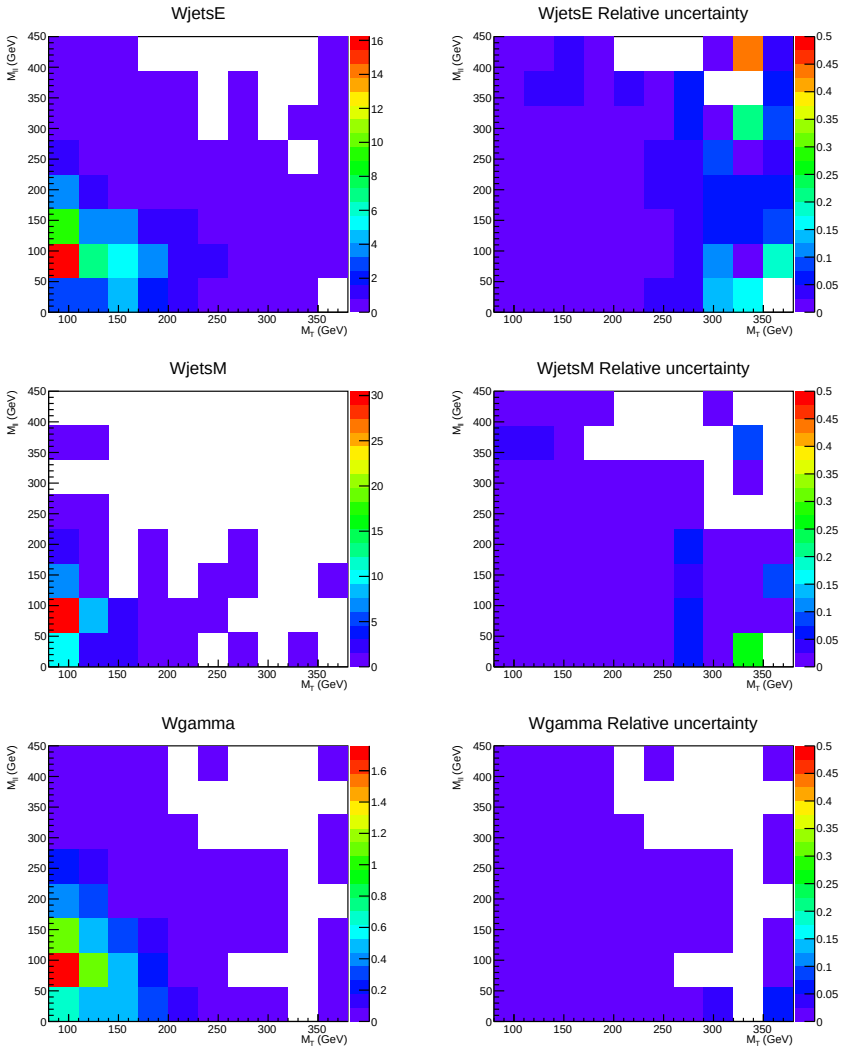


Figure A.7: Templates(left) and relative statistical uncertainty of the MC sample(right) of $\mathrm{W}(\mathrm{e}\nu_{\mathrm{e}}) + \mathrm{jets}$, $\mathrm{W}(\mu\nu_{\mu}) + \mathrm{jets}$ and $W\gamma$. The templates are used for $m_{\mathrm{H}} = 400$ GeV analysis in the 0-jet category.

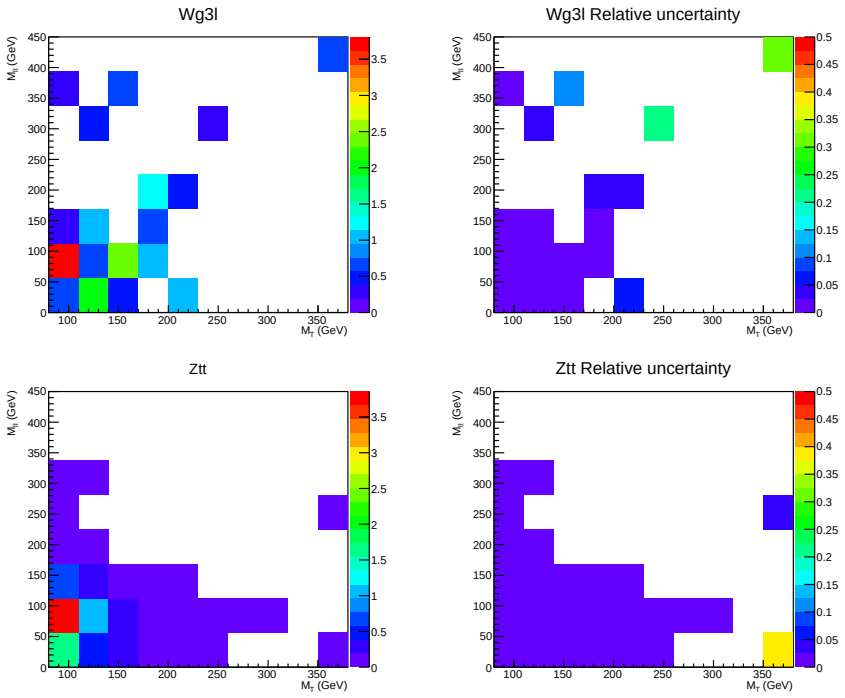


Figure A.8: Templates(left) and relative statistical uncertainty of the MC sample(right) of $W\gamma^*$ and $Z \to \tau\tau$. The templates are used for $m_{\mathrm{H}} = 400$ GeV analysis in the 0-jet category.

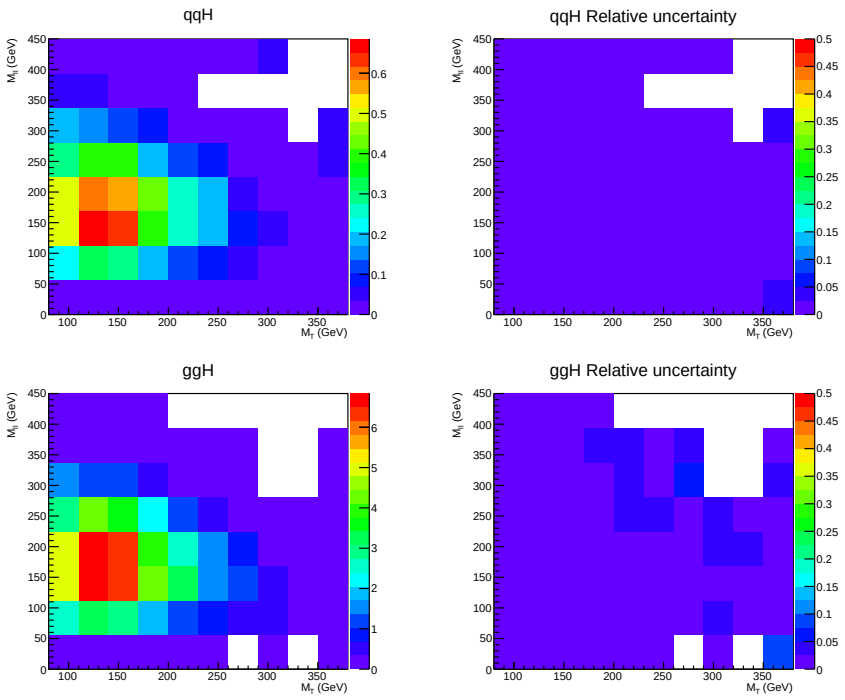


Figure A.9: Templates(left) and relative statistical uncertainty of the MC sample(right) of $qq \to qqH$ and $gg \to H$. The templates are used for $m_{\mathrm{H}} = 400$ GeV analysis in the 1-jet category.

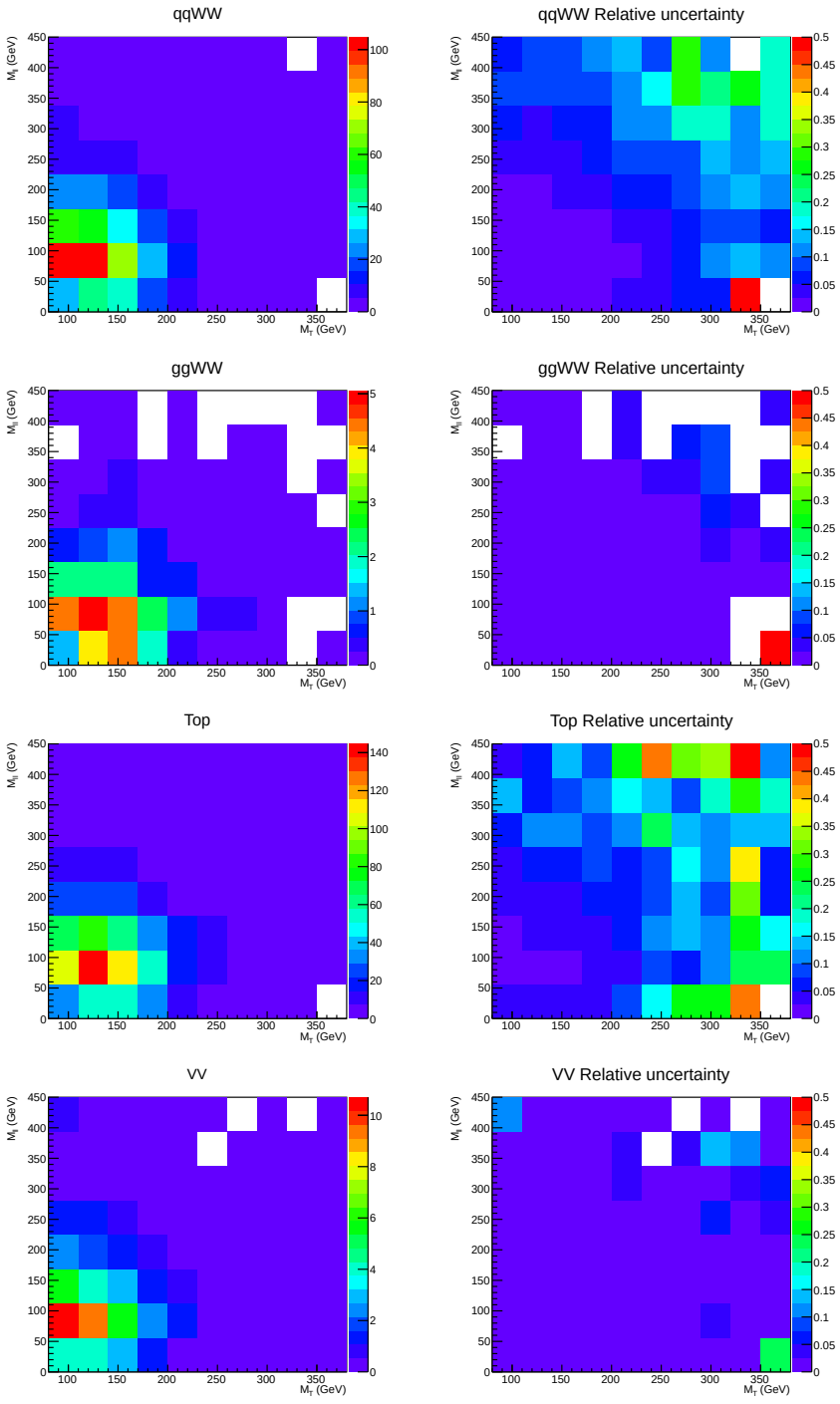


Figure A.10: Templates(left) and relative statistical uncertainty of the MC sample(right) of $qq \to W^+W^-$, $gg \to W^+W^-$, $t\bar{t}/tW$ and WZ/ZZ. The templates are used for $m_{\mathrm{H}} = 400$ GeV analysis in the 1-jet category.

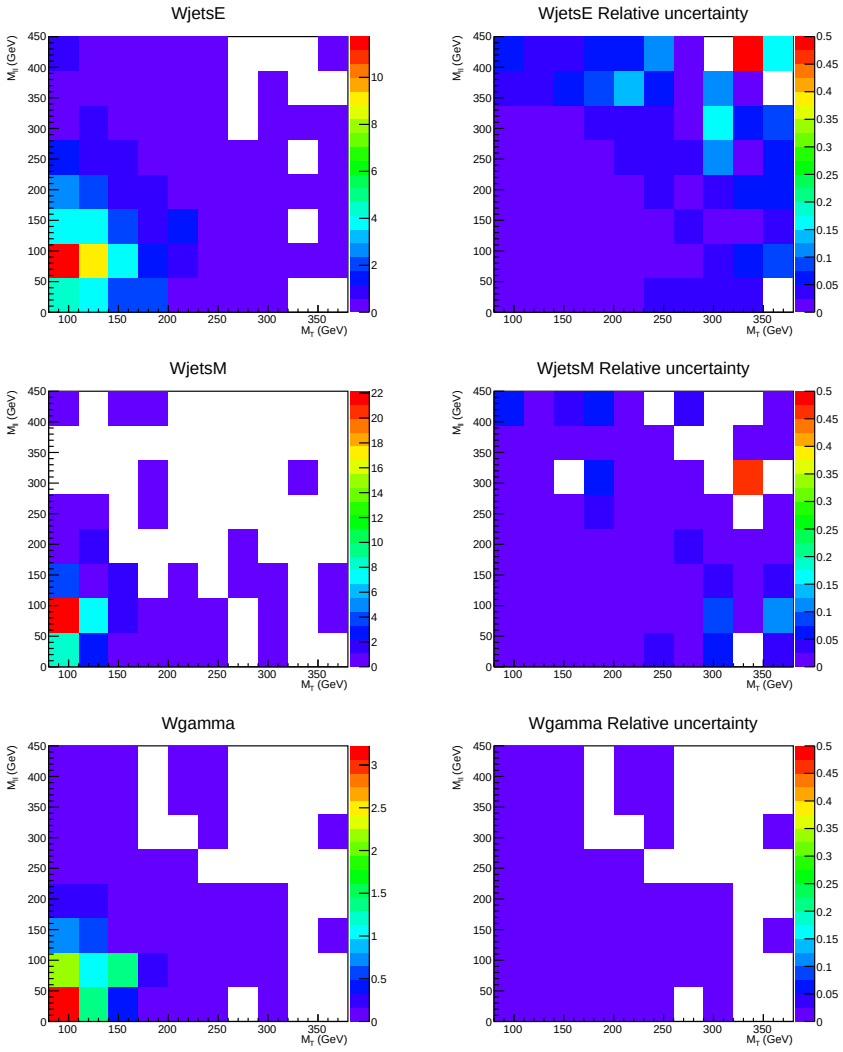


Figure A.11: Templates(left) and relative statistical uncertainty of the MC sample(right) of $\mathrm{W}(\mathrm{e}\nu_{\mathrm{e}}) + \mathrm{jets}$, $\mathrm{W}(\mu\nu_{\mu}) + \mathrm{jets}$ and $W\gamma$. The templates are used for $m_{\mathrm{H}} = 400$ GeV analysis in the 1-jet category.

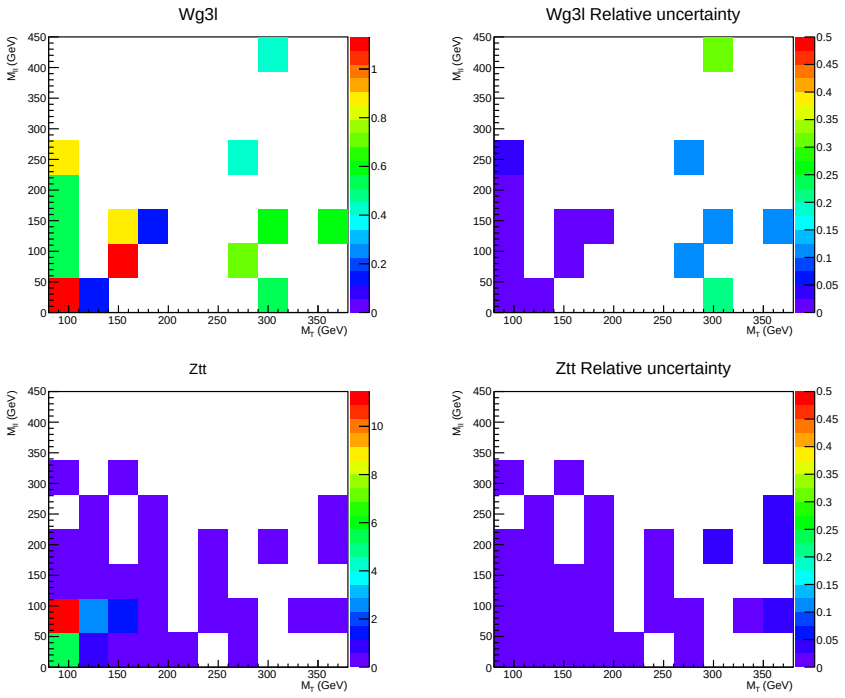


Figure A.12: Templates(left) and relative statistical uncertainty of the MC sample(right) of $W\gamma^*$ and $Z \to \tau\tau$. The templates are used for $m_{\mathrm{H}} = 400$ GeV analysis in the 1-jet category.

Bibliography

[1] V. Barger, K. Cheung, A. Djouadi, B. A. Kniehl, and P. M. Zerwas. Higgs bosons: Intermediate mass range at e^+e^- colliders. *Phys. Rev. D*, 49:79–90, Jan 1994.

[2] The ALEPH, DELPHI, L3, OPAL Collaborations, the LEP Electroweak Working Group. Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP. 2013.

[3] Ioana Anghel. Status of the CMS silicon strip tracker and commissioning results. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 604(12):277 – 280, 2009.

[4] D. Abbaneo. Layout and performance of the CMS silicon strip tracker. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 518(12):331 – 335, 2004.

[5] CMS Collaboration. *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*. Technical Design Report CMS. CERN, Geneva, 2006. There is an error on cover due to a technical problem for some items.

[6] CMS Collaboration. Performance of CMS hadron calorimeter timing and synchronization using test beam, cosmic ray, and lhc beam data. *JINST*, 5:T03013, 2010.

[7] Min Suk on behalf of the CMS Collaboration Kim. CMS reconstruction improvement for the muon tracking by the RPC chambers. *PoS*, RPC2012:045, 2012.

[8] Silvia Maselli. Calibration of the barrel muon drift tubes system in CMS. Technical Report CMS-CR-2009-100, CERN, Geneva, May 2009.

[9] Thomas Lenzi. Development and study of different muon track reconstruction algorithms for the level-1 trigger for the CMS muon upgrade with gem detectors. 2013.

[10] CMS Collaboration. Plots for ichep 2012 tracking pog. *CMS Physics Results*, 2012.

[11] CMS Collaboration. Tracking pog plots for vertex 2012. *CMS Physics Results*, 2012.

[12] CMS Collaboration. Properties of the higgs-like boson in the decay $h \to zz \to 4l$ in pp collisions at $\sqrt{s} = 7$ and 8 TeV. *Physics Analysis Summary*, HIG-13-002, 2013.

[13] CMS Collaboration. Status of the 8 TeV Jet Energy Corrections and Uncertainties based on 11 $\text{fb}^{-1}$of data in CMS. *CMS Performance Note*, DPS-2013/011, 2013.

[14] MET performance in 8 TeV data. Technical Report CMS-PAS-JME-12-002, CERN, Geneva, 2013.

[15] CMS Collaboration. Measurement of Higgs boson production and properties in the WW decay channel with leptonic final states. *J. High Energy Phys.*, 01(arXiv:1312.1129. CMS-HIG-13-023. CERN-PH-EP-2013-221):096. 83 p, Dec 2013. Comments: Submitted to JHEP.

[16] S.L. Glashow. Partial Symmetries of Weak Interactions. *Nucl.Phys.*, 22:579–588, 1961.

[17] Steven Weinberg. A Model of Leptons. *Phys.Rev.Lett.*, 19:1264–1266, 1967.

[18] A. Salam. Weak and electromagnetic interactions, in Elementary particle physics: relativistic groups and analyticity (proceedings of the 8th Nobel symposium). *Svartholm, ed., Almqvist & Wikselll.*, 1968.

[19] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys.Rev.Lett.*, 13:321–323, 1964.

[20] Peter W. Higgs. Broken symmetries, massless particles and gauge fields. *Phys.Lett.*, 12:132–133, 1964.

[21] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys.Rev.Lett.*, 13:508–509, 1964.

[22] G.S. Guralnik, C.R. Hagen, and T.W.B. Kibble. Global Conservation Laws and Massless Particles. *Phys.Rev.Lett.*, 13:585–587, 1964.

[23] Peter W. Higgs. Spontaneous Symmetry Breakdown without Massless Bosons. *Phys.Rev.*, 145:1156–1163, 1966.

[24] T.W.B. Kibble. Symmetry breaking in nonAbelian gauge theories. *Phys.Rev.*, 155:1554–1561, 1967.

[25] Guido Altarelli and G. Parisi. Asymptotic Freedom in Parton Language. *Nucl.Phys.*, B126:298, 1977.

[26] Yuri L. Dokshitzer. Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics. *Sov.Phys.JETP*, 46:641–653, 1977.

[27] V.N. Gribov and L.N. Lipatov. Deep inelastic e p scattering in perturbation theory. *Sov.J.Nucl.Phys.*, 15:438–450, 1972.

[28] G. Passarino R. Tanaka (eds.) S. Alekhin J. Alwall E. A. Bagnaschi A. Banfi J. Blumlein S. Bolognesi N. Chanon T. Cheng L. Cieri A. M. Cooper-Sarkar M. Cutajar S. Dawson G. Davies N. De Filippis G. Degrassi A. Denner D. D'Enterria S. Diglio B. Di Micco R. Di Nardo R. K. Ellis A. Farilla S. Farrington M. Felcini G. Ferrera M. Flechl D. de Florian S. Forte S. Ganjour M. V. Garzelli S. Gascon-Shotkin S. Glazov S. Goria M. Grazzini J.-Ph. Guillet C. Hackstein K. Hamilton R. Harlander M. Hauru S. Heinemeyer S. Hoche J. Huston C. Jackson P. Jimenez-Delgado M. D. Jorgensen M. Kado S. Kallweit A. Kardos N. Kauer H. Kim M. Kovac M. Kramer F. Krauss C.-M. Kuo S. Lehti Q. Li N. Lorenzo F. Maltoni B. Mellado S. O. Moch A. Muck M. Muhlleitner P. Nadolsky P. Nason C. Neu A. Nikitenko C. Oleari J. Olsen S. Palmer S. Paganis C. G. Papadopoulos T .C. Petersen F. Petriello F. Petrucci G. Piacquadio E. Pilon C. T. Potter J. Price I. Puljak W. Quayle V. Radescu D. Rebuzzi L. Reina J. Rojo D. Rosco G. P. Salam A. Sapronov J. Schaarschmidt M. Schonherr M. Schumacher F. Siegert P. Slavich M. Spira I. W. Stewart W. J. Stirling F. Stockli C. Sturm F. J. Tackmann R. S. Thorne D. Tommasini P. Torrielli F. Tramontano Z. Trocsanyi M. Ubiali S. Uccirati M. Vazquez Acosta T. Vickey A. Vicini W. J. Waalewijn D. Wackeroth M. Warsinsky M. Weber M. Wiesemann G. Weiglein J. Yu G. Zanderighi LHC Higgs Cross Section Working Group: S. Dittmaier, C. Mariotti. Handbook of LHC Higgs Cross Sections: 2. Differential Distributions. 2012.

[29] Stefano Goria, Giampiero Passarino, and Dario Rosco. The Higgs Boson Lineshape. *Nucl.Phys.*, B864:530–579, 2012.

[30] H.M. Pilkuhn. *The interactions of hadrons*. North-Holland Pub. Co., 1967.

[31] Abdelhak Djouadi. The anatomy of electroweak symmetry breaking: Tome i: The higgs boson in the standard model. *Physics Reports*, 457(14):1 – 216, 2008.

[32] S Dittmaier, C Mariotti, G Passarino, R Tanaka, J Baglio, P Bolzoni, R Boughezal, O Brein, C Collins-Tooth, S Dawson, S Dean, A Denner, S Farrington, M Felcini, M Flechl, D de Florian, S Forte, M Grazzini, C Hackstein, T Hahn, R Harlander, T Hartonen, S Heinemeyer, J Huston, A Kalinowski, M Krmer, F Krauss, J S Lee, S Lehti, F Maltoni, K Mazumdar, S-O

Moch, A Mck, M Mhlleitner, P Nason, C Neu, C Oleari, J Olsen, S Palmer, F Petriello, G Piacquadio, A Pilaftsis, C T Potter, I Puljak, J Qian, D Rebuzzi, L Reina, H Rzehak, M Schumacher, P Slavich, M Spira, F Stckli, R S Thorne, M Vazquez Acosta, T Vickey, A Vicini, D Wackeroth, M Warsinsky, M Weber, G Weiglein, C Weydert, J Yu, M Zaro, and T Zirke. *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*. CERN, Geneva, 2011. Comments: 153 pages, 43 figures, to be submitted to CERN Report. Working Group web page: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections.

[33] Hong-Jian He, Yu-Ping Kuang, and C.-P. Yuan. Equivalence theorem and probing the electroweak symmetry breaking sector. *Phys.Rev.*, D51:6463–6473, 1995.

[34] Christopher F. Kolda and Hitoshi Murayama. The Higgs mass and new physics scales in the minimal standard model. *JHEP*, 0007:035, 2000.

[35] Particle Data Group. Review of Particle Physics (RPP). *Phys.Rev.*, D86:010001, 2012.

[36] CMS Collaboration. Combined results of searches for the standard model higgs boson in pp collisions at. *Physics Letters B*, 710(1):26 – 48, 2012.

[37] ATLAS Collaboration. Combined search for the standard model higgs boson using up to 4.9 fb^{-1} of pp collision data at with the atlas detector at the lhc. *Physics Letters B*, 710(1):49 – 66, 2012.

[38] CMS Collaboration. Search for the standard model Higgs boson decaying to a W pair in the fully leptonic final state in pp collisions at $\sqrt{s} = 7$ TeV. *Phys.Lett.*, B710:91–113, 2012.

[39] LHC RUNNING. http://lhc-closer.es/1/3/5/0.

[40] Wolfgang Adam, Boris Mangano, Thomas Speer, and Teddy Todorov. Track Reconstruction in the CMS tracker. Technical Report CMS-NOTE-2006-041, CERN, Geneva, Dec 2006.

[41] R. Frhwirth, P. Kubinec, W. Mitaroff, and M. Regler. Vertex reconstruction and track bundling at the LEP collider using robust algorithms. *Computer Physics Communications*, 96(23):189 – 208, 1996.

[42] P. Azzurri and B. Mangano. Optimal filtering of fake tracks. *CMS Internal Note*, (2008/017), 2008.

[43] M. Pioppi. Iterative tracking. *CMS Internal Note*, (2007/065), 2008.

[44] W. Erdmann. Offline primary vertex reconstruction with deterministic annealing clustering. *CMS Note*, IN-2011/014, 2011.

[45] W. Waltenberger R. Fruewirth and P. Vanlaer. Adaptive vertex fitting. *CMS Note*, NOTE-2007/008, 2007.

[46] F. Ferri D. Futyan P. Meridiani I. Puljak C. Rovelli R. Salerno Y. Sirois S. Baffioni, C. Charlot. Electron reconstruction in CMS. *Eur.Phys.J.*, C49:1099–1116, 2007.

[47] H. Bethe and W. Heitler. On the stopping of fast particles and on the creation of positive electrons. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 146(856):pp. 83–112, 1934.

[48] W Adam, R Frhwirth, A Strandlie, and T Todorov. Reconstruction of electrons with the gaussian-sum filter in the cms tracker at the lhc. *Journal of Physics G: Nuclear and Particle Physics*, 31(9):N9, 2005.

[49] CMS Collaboration. Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus and E_T^{miss}. *Physics Analysis Summary*, PFT-09-001, 2009.

[50] Stephen D. Ellis V. Daniel Elvira K. Frame S. Grinstein Robert Hirosky R. Piegaia H. Schellman R. Snihur V.Sorin Dieter Zeppenfeld Gerald C. Blazey, Jay R. Dittmann. Run II jet physics. pages 47–77, 2000.

[51] *Phys. Rev. D*, 45:1448–1458, Mar 1992.

[52] *Phys. Rev. D*, 53:6000–6016, Jun 1996.

[53] Gavin P. Salam and Gregory Soyez. A Practical Seedless Infrared-Safe Cone jet algorithm. *JHEP*, 0705:086, 2007.

[54] Stephen D. Ellis and Davison E. Soper. Successive combination jet algorithm for hadron collisions. *Phys.Rev.*, D48:3160–3166, 1993.

[55] Yuri L. Dokshitzer, G.D. Leder, S. Moretti, and B.R. Webber. Better jet clustering algorithms. *JHEP*, 9708:001, 1997.

[56] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Anti-k(t) jet clustering algorithm. *JHEP*, 0804:063, 2008.

[57] Matteo Cacciari and Gavin P. Salam. Dispelling the n^3 myth for the k_t jetfinder. *Phys.Lett.*, B641:57–61, 2006.

[58] CMS Collaboration. Determination of jet energy calibration and transverse momentum resolution in CMS. *J. Instrum.*, 6(arXiv:1107.4277. CMS-JME-10-011. CERN-PH-EP-2011-102):P11002. 67 p, Jul 2011.

[59] CMS Collaboration. Identification of b-quark jets with the CMS experiment. *J. Instrum.*, 8(arXiv:1211.4462. CMS-BTV-12-001. CERN-PH-EP-2012-262):P04013. 67 p, Nov 2012. Comments: Submitted to the Journal of Instrumentation.

[60] CDF Collaboration. Measurements of inclusive W and Z cross sections in p anti-p collisions at s**(1/2) = 1.96-TeV. *J.Phys.*, G34:2457–2544, 2007.

[61] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys.Lett.*, B716:30–61, 2012.

[62] Andreas Hoecker, Peter Speckmayer, Joerg Stelzer, Jan Therhaag, Eckhard von Toerne, and Helge Voss. TMVA: Toolkit for Multivariate Data Analysis. *PoS*, ACAT:040, 2007.

[63] Pileup Jet Identification. Technical Report CMS-PAS-JME-13-005, CERN, Geneva, 2013.

[64] L. Bauerdick, K. Burkett, I. Fisk, Y. Gao, O. Gutsche, B. Hooberman, S. Jindariani, J. Linacre, V. Martinez Outschoorn, S. Tkaczyk A. Apyan, G. Bauer, J. Bendavid, E. Butz, M. Chan, V. Dutta, G. Gomez-Ceballos, M. Goncharov, K. Hahn, P. Harris, M. Klute, A. Levin, S. Nahn, C. Paus, D. Ralph, F. Stoeckli, K. Sumorok, K. Sung, R. Wolf, S. Xie, M. Yang, M. Zanetti D. Barge, C. Campagnari, D. Kovalskyi, V. Krutelyov W. Andrews, G. Cerati, D. Evans, F. Golf, I. MacNeill, S. Padhi, Y. Tu, F. Wurthwein, A. Yagil, J. Yoo I. Kravchenko. A multivariate approach for Drell-Yan rejection in $H \to W^+W^-$ analysis. *CMS Analysis Note*, AN-2012/211, 2012.

[65] Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer. Madgraph 5: going beyond. *Journal of High Energy Physics*, 2011(6):1–40, 2011.

[66] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP*, 0711:070, 2007.

[67] The CMS collaboration. Search for the standard model higgs boson decaying to a w pair in the fully leptonic final state in pp collisions at sqrt(s) = 8 tev. *Physics Analysis Summary*, (CMS-PAS-HIG-12-038).

[68] CMS Collaboration. Measurements of Inclusive W and Z Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV. *JHEP*, 1101:080, 2011.

[69] Nikolaos Kidonakis. NNLL threshold resummation for top-pair and single-to production. 2012.

[70] CMS Collaboration. Evidence for the 125 GeV Higgs boson decaying to a pair of tau leptons. 2014.

[71] Stanislaw Jadach, Johann H. Kuhn, and Zbigniew Was. TAUOLA: A Library of Monte Carlo programs to simulate decays of polarized tau leptons. *Comput.Phys.Commun.*, 64:275–299, 1990.

[72] CMS Collaborations The ATLAS and the LHC Higgs Combination Group. *Procedure for the LHC Higgs boson search combination in summer 2011. ATL-PHYS-PUB-2011-11, CMS-NOTE-2011-005*, 2011.

[73] Neyman, Jerzy; Pearson, Egon S. On the problem of the most efficient tests of statistical hypothese. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engi neering Science*, 231(694-704):289–337, 1933.

[74] E. Gross G. Cowan, K. Cranmer and O. Vitells. *Asymptotic formulae for likelihood-based tests of new physics. The European Physical Journal C*, **71**:1–19, 2011.

[75] A L Read. Modified frequentist analysis of search results (the CL_s method). (CERN-OPEN-2000-205), 2000.

[76] Thomas Junk. Confidence level computation for combining searches with small statistics. *Nucl.Instrum.Meth.*, A434:435–443, 1999.

[77] J. S. Conway. Incorporating nuisance parameters in likelihoods for multisource spectra. *ArXiv e-prints*, March 2011.

[78] Amanda Cooper-Sarkar Albert de Roeck Joel Feltesse Stefano Forte Alexander Glazov Joey Huston Ronan McNulty Torbjorn Sjostrand Robert Thorne Michiel Botje, Jon Butterworth. The PDF4LHC Working Group Interim Recommendations. 2011.

[79] E. Butz M. Chan V. Dutta P. Everaerts G. Gomez-Ceballos M. Goncharov K. Hahn P. Harris M. Klute I. Kravchenko S. Nahn C. Paus D. Ralph M. Rudolph F. Stoeckli K. Sumorok K. Sung S. Tkaczyk R. Wolf S. Xie M. Yang M. Zanetti G. Bauer, J. Bendavid. Acceptance uncertainties for the inclusive w and z cross section measurements. *CMS*, (AN-2011/055), 2011.

[80] Stefano Frixione and Bryan R. Webber. Matching NLO QCD computations and parton shower simulations. *JHEP*, 0206:029, 2002.

[81] G.Marchesini S.Moretti K.Odagiri P.Richardson M.H.Seymour B.R.Webber G.Corcella, I.G.Knowles. HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes). *JHEP*, 0101:010, 2001.

[82] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 0605:026, 2006.

[83] CMS Collaboration. Measurement of the Underlying Event Activity at the LHC with $\sqrt{s} = 7$ TeV and Comparison with $\sqrt{s} = 0.9$ TeV. *JHEP*, 1109:109, 2011.

[84] CMS Collaboration. Measurement of the underlying event in the Drell-Yan process in proton-proton collisions at $\sqrt{s} = 7$ TeV. *Eur.Phys.J.*, C72:2080, 2012.

[85] John M. Campbell and R.K. Ellis. MCFM for the Tevatron and the LHC. *Nucl.Phys.Proc.Suppl.*, 205-206:10–15, 2010.

[86] Michelangelo L. Mangano, Mauro Moretti, Fulvio Piccinini, Roberto Pittau, and Antonio D. Polosa. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 0307:001, 2003.

[87] The CMS collaboration. Cms luminosity based on pixel cluster counting - summer 2013 update. *Physics Analysis Summary*, (CMS-PAS-LUM-13-001), 2013.

[88] CMS Collaboration. Rapidity distributions in exclusive Z + jet and γ + jet events in pp collisions at $\sqrt{s}$=7 TeV. *Phys.Rev.*, D88:112009, 2013.

[89] CMS Collaboration. Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS. *JINST*, 6:P11002, 2011.

[90] CMS Collaboration. Measurement of the properties of a Higgs boson in the four-lepton final state. Technical Report arXiv:1312.5353. CMS-HIG-13-002. CERN-PH-EP-2013-220, CERN, Geneva, Dec 2013. Comments: Submitted to Phys. Rev. D.

[91] Updated measurements of the Higgs boson at 125 GeV in the two photon decay channel. Technical Report CMS-PAS-HIG-13-001, CERN, Geneva, 2013.

[92] C. N. Yang. Selection rules for the dematerialization of a particle into two photons. *Phys. Rev.*, 77:242–245, Jan 1950.

[93] Andrei V. Gritsan Kirill Melnikov Markus Schulze Nhan V. Tran-Andrew Whitbeck Sara Bolognesi, Yanyan Gao. On the spin and parity of a single-produced resonance at the LHC. *Phys.Rev.*, D86:095031, 2012.

[94] Zijin Guo Kirill Melnikov Markus Schulze Nhan V. Tran Yanyan Gao, Andrei V. Gritsan. Spin determination of single-produced resonances at hadron colliders. *Phys.Rev.*, D81:075022, 2010.

[95] I. Fisk Y. Gao O. Gutsche B. Hooberman S. Jindariani-J. Linacre V. Martinez Outschoorn S. Tkaczyk A. Apyan G. Bauer J. Bendavid E. Butz M. Chan V. Dutta G. Gomez-Ceballos M. Goncharov K. Hahn P. Harris M. Klute A. Levin S. Nahn C. Paus D. Ralph F. Stoeckli K. Sumorok K. Sung R. Wolf S. Xie M. Yang M. Zanetti D. Barge C. Campagnari D. Kovalskyi V. Krutelyov W. Andrews G. Cerati D. Evans F. Golf I. MacNeill S. Padhi-Y. Tu F. Wurthwein A. Yagil J. Yoo I. Kravchenko L. Bauerdick, K. Burkett. The Spin of a Single Produced Resonance in the Fully Leptonic W^+W^- Final State. *CMS AN*, CMS AN-2012/453, 2012.

[96] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys.Lett.*, B716:1–29, 2012.

[97] Combination of standard model Higgs boson searches and measurements of the properties of the new boson with a mass near 125 GeV. Technical Report CMS-PAS-HIG-13-005, CERN, Geneva, 2013.