

Search for the Standard Model Higgs boson in $Z + \gamma$ final
states with the ATLAS detector at the LHC

Feng Tian

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

CERN-THESIS-2014-001
// 2013



COLUMBIA UNIVERSITY

2013

©2013

Feng Tian

All Rights Reserved

ABSTRACT

Search for the Standard Model Higgs boson in $Z + \gamma$ final states with the ATLAS detector at the LHC

Feng Tian

This dissertation describes a search for the Standard Model Higgs boson in $Z\gamma$ channel, $Z \rightarrow ee (\mu\mu)$, with the ATLAS detector at the LHC using 4.6 fb^{-1} of proton-proton collisions at $\sqrt{s} = 7 \text{ TeV}$ in 2011 and 20.7 fb^{-1} of proton-proton collisions at $\sqrt{s} = 8 \text{ TeV}$ in 2012. The distribution of the ΔM (the mass difference between $M_{Z\gamma}$ and M_Z) is compared to the Standard Model (SM) background expectations. No significant deviation from the SM prediction is observed and the upper limits on the signal strength $\mu = \frac{\sigma}{\sigma_{SM}}$ of a Higgs boson with a mass between 120 and 150 GeV are derived. The expected exclusion limits at 95% confidence level range between 7.3 and 22.3 times the predicted Standard Model cross section. The observed exclusion limits range between 5.3 and 38.7 times the Standard Model cross section. For a Higgs boson with mass of 125.5 GeV, the expected and observed limits are 13.2 and 17.0 times the Standard Model respectively.

Table of Contents

I	Theory	1
1	The Standard Model	3
1.1	The Standard Model	3
1.1.1	Fermions	4
1.1.2	Bosons	6
1.1.3	Interactions	7
1.1.4	Limitations of the Standard Model	8
1.2	The Higgs Mechanism	9
1.2.1	Spontaneously Symmetry Breaking and the Higgs boson	9
1.2.2	The Higgs boson properties	10
1.3	Motivation	14
II	Experiment	17
2	The Large Hadron Collider	19
2.1	LHC introduction	19
2.2	Beam acceleration and focusing	21
2.3	Luminosity	23
3	ATLAS detector	29
3.1	The ATLAS Coordinate System	29
3.2	The Inner detector	31
3.2.1	The Pixel Detector	32

3.2.2	The Silicon Microstrip Tracker	32
3.2.3	The Transition Radiation Tracker	32
3.3	Calorimeter	33
3.3.1	EM Calorimeter(ECAL)	33
3.3.2	Hadronic Calorimeter (HCAL)	36
3.3.3	Forward Calorimeter (FCAL)	38
3.4	Muon System	38
3.4.1	Monitored Drift Tubes	38
3.4.2	Cathode Strip Chambers	38
3.4.3	Resistive Plate Chambers	40
3.4.4	Thin gap chambers	40
3.4.5	Overall Muon System Layout	42
3.5	Trigger and Data acquisition system (TDAQ)	42
3.6	Luminosity detectors	45
3.7	Summary of sub-detector performance	45
III	Data analysis	47
4	Object Reconstruction and Identification	49
4.1	Track Reconstruction	49
4.2	Vertex Reconstruction	51
4.3	Electron and Photon (Egamma) Objects	53
4.3.1	Electrons	53
4.3.2	Photons	56
4.4	Calorimeter Data Quality Monitoring	58
4.5	Muons	59
4.6	Jets	60
4.7	Missing transverse energy (MET)	61
5	Data and MC samples	63
5.1	Data Model	63

5.2	Data Samples	64
5.3	Monte Carlo Samples	65
5.3.1	Generators	66
5.3.2	Signal and Background Samples	67
6	Event selection	69
6.1	Object Level Selection	69
6.1.1	Muon selection	69
6.1.2	Electron selection	70
6.1.3	Photon selection	72
6.2	Event Level Selection	72
6.2.1	Preselection	72
6.2.2	Z boson reconstruction and selection	75
6.2.3	Higgs reconstruction and selection	77
7	Signal Property Studies	81
7.1	Signal Mass reconstruction	81
7.2	Signal Mass resolution improvement	82
7.2.1	Photon vertex correction	82
7.2.2	Z mass constraint	84
7.2.3	Summary of the photon vertex correction and the Z mass constraint	89
7.3	Signal Modeling	89
7.4	Signal acceptance and expected yield	91
7.5	Kinematic distributions	98
8	Background estimation	103
8.1	Introduction	103
8.2	Standard Model $Z + \gamma$ Background	104
8.3	Z +jets Background	105
8.3.1	2-D sideband method	105
8.3.2	Control Region Definition	107

8.3.3	Systematics	112
8.4	Other backgrounds	113
8.4.1	$t\bar{t}$ and WZ Background	113
8.5	Summary for background estimation	113
8.6	The comparison between MC simulation and data	113
8.7	Background Modeling	114
9	Statistical analysis	119
9.1	Introduction	119
9.2	Test Statistics	120
9.2.1	p_0 for discovery	120
9.2.2	Upper limit	121
9.3	CLs Method for setting upper limit	122
9.4	Likelihood Model	123
9.5	Systematic uncertainties	124
9.5.1	Theoretical uncertainties	125
9.5.2	Experimental systematic uncertainties	126
9.5.3	signal mass resolution uncertainty	129
9.5.4	signal mass peak position uncertainty	130
10	Results	133
10.1	p value for a background-only hypothesis	133
10.2	Exclusion limits on $H \rightarrow Z\gamma$ signal strength μ	133
10.3	Conclusions and outlook	135
IV	Appendices	139
A	MC samples	141
B	MC and data comparison plots	145

V Bibliography	149
Reference	159

List of Figures

1.1	The elementary particles of the Standard Model. The quarks are shown in purple, the leptons in green, the gauge bosons in red and the Higgs boson in yellow. The quarks and leptons are also classified in three generations. The generations are represented by the three first columns in the figure. Not shown are the corresponding anti-particles of the quarks and leptons. . . .	4
1.2	Higgs Potential	9
1.3	Higgs Mass Exclusion limit from Tevatron by February 27, 2012. The dotted line is the expected limit on the signal strength ($\frac{\sigma}{\sigma_{SM}}$). The solid black line is the observed limit. The green and yellow band around the expected limit correspond to $\pm 1\sigma$ and $\pm 2\sigma$ respectively. The M_H region whose observed limit is below 1 can be excluded. Most of the M_H region between 100 GeV and 200 GeV were excluded by LEP, Tevatron, ATLAS and CMS experiments except for the region between 115 GeV to 127 GeV.	11
1.4	Feynman diagram for Higgs production precesses (a) Gluon-Gluon Fusion (b) Vector Boson Fusion (c) Higgs-strahlung (VH) (d) associated production with top-quark pairs ($t\bar{t}H$)	12
1.5	Higgs Production Cross-section at 7 TeV (Top plot) and 8 TeV (Bottom plot). The blue curve with the highest cross-section corresponds to the gluon-gluon fusion process. The curve with the second largest cross-section corresponds to the vector boson fusion process. The other three curves correspond to WH , ZH and $t\bar{t}H$ process respectively. The bands are the theoretical uncertainty from the QCD scales uncertainties and PDFs+ α_s uncertainties.	13

1.6	Higgs decay branching ratio. The bands represent the theoretical uncertainties.	14
1.7	Higgs signal strength $\mu = \frac{\sigma}{\sigma_{SM}}$ from ATLAS [1].	15
1.8	Leading order Feynman diagrams for the Higgs decaying to $\gamma\gamma$	16
1.9	Leading order Feynman diagrams for the Higgs decaying to $Z\gamma$	16
2.1	LHC is located 100 meters underground between Lake Geneva and the Jura mountains. The locations of the four detectors ATLAS, CMS, ALICE and LHCb can be seen on the plot. Point 1 is where the ATLAS detector is located.	20
2.2	The CERN accelerator complex from the LINAC to the LHC. Also indicated by the colors are the accelerators used to accelerate heavy ions (blue), protons (orange) and antiprotons (red).	21
2.3	LHC dipole magnets (left) and the dipole magnetic flux (right).	22
2.4	LHC Quadrupole magnets (left) and Quadrupole magnetic flux (right). . . .	22
2.5	Beams interaction at ATLAS. The beam size at the interaction point is about $16 \mu\text{m}$	24
2.6	ATLAS daily integrated luminosity in 2011 (upper plot) and in 2012 (lower plot). The green area shows the delivered luminosity to ATLAS by the LHC and the yellow area shows the luminosity recorded by ATLAS. The recorded luminosity is less than the delivered luminosity either because the ATLAS detector data acquisition chain was busy or because one or more of its sub-detector systems are temporarily unavailable.	26
2.7	The luminosity-weighted distribution of the mean number of interactions per crossing (μ) for 2011 (upper plot) and 2012 (lower plot). The integrated luminosities and the $\langle \mu \rangle$ values are given in the figure. The mean number of interactions per crossing corresponds the mean of the poisson distribution on the number of interactions per crossing for each bunch. It is calculated from the instantaneous per bunch luminosity as $\mu = \frac{L_{bunch} \times \sigma_{inel}}{f_r}$ where L_{bunch} is the per bunch instantaneous luminosity, σ_{inel} is the inelastic cross section which is 71.5 mb (2011) and 73 mb (2012) and f_r is the LHC revolution frequency.	27

3.1	The ATLAS detector.	30
3.2	An illustration of particle detection in the subsystems of the ATLAS detector.	30
3.3	The Inner detector.	31
3.4	Inner detector thickness in units of radiation length as a function of η as measured from the center of the detector.	33
3.5	Main parameters of the inner detector.	34
3.6	ATLAS Calorimeter	34
3.7	An electromagnetic shower in the calorimeter.	35
3.8	Photograph of a partly stacked barrel electromagnetic LAr module. A total of six out of seven outer support rings into which the absorbers can be seen. The backbone behind the outer support rings and the assembly bench below the stacked modules are also visible.	36
3.9	Pseudorapidity coverage, granularity and longitudinal segmentation of the calorimeters.	37
3.10	The Monitored drift tube (MDT). Under the electric field generated by the 3080 V potential, ionization electrons drift towards the anode and initiate an electron avalanche near the wire. The green circles show the starting point of this drifting in each tube. The radius of the circles is the closest distance from the wire to the muon track. The time of arrival of electrons from the starting point of closest approach determines the drift time corresponding to the drift radius.	39
3.11	The Cathode Strip Chambers (CSC).	39
3.12	The mechanical structure of an RPC chamber.	40
3.13	The structure of thin gap chambers (TGC).	41
3.14	The overall ATLAS muon system (upper plot) and a detailed quarter view cut-away cross section of the muon system (lower plot).	43
3.15	A photograph of the end view of the ATLAS detector under construction in 2005 with the eight air-core superconducting toroid magnets visible (they are the elements in the picture with the orange “stripes”).	44
3.16	Trigger and Data acquisition system (TDAQ)	44

3.17	Sub-detector performance in 2011 (upper) and 2012 (lower)	46
4.1	An illustration of the energy loss processes for the different types of particles produced in proton-proton collisions in the detector. Different detector sub-systems are optimized for the specific particle types. The interaction point is at the bottom of the figure.	50
4.2	Track reconstruction efficiencies as a function of $ \eta $ for muons (left) and pions (right) with $p_T = 1, 5$ and 100 GeV.	51
4.3	Track reconstruction efficiencies as a function of $ \eta $ for electrons (left) with $p_T = 1, 5$ and 100 GeV and comparisons between electrons, muons and pions with $p_T = 5$ GeV (right).	52
4.4	Multiple primary vertices were reconstructed in one event. The colored circles identify the primary vertices.	53
4.5	The position and size in z of the centroid of the luminous region in ATLAS over the course of pp running in 2012 at $\sqrt{s} = 8$ TeV. The data points are the result of a maximum likelihood fit to the spatial distribution of primary vertices collected over ten minutes. The large movement at the beginning of May was associated with movement of the inner detector. Errors are statistical only.	54
4.6	Definition of variables used for loose, medium and tight electron identification cuts.	57
4.7	Discriminating variables used for loose and tight photon identification cuts.	58
4.8	Calorimeter Data Quality monitoring. The plot is an hit map in $\eta - \phi$, displaying the entries for topological clusters with transverse energy about 10 GeV. The noisy cluster, shown in red, is around $\eta - \phi = (1.6, 0)$. It will also be checked at the CaloCell level and masked in the reprocessed data. .	59
4.9	Shapes for jets algorithms reconstructing the same event. The towers in the (y, ϕ) plane are colored according to which jet they belong to. The vertical axis shows the p_T of the tower. The <i>anti</i> - k_T algorithm reconstructs jets with a regular shape, which is very attractive experimentally.	61

6.1	The d_0 (left) and z_0 (right) distribution before applying cuts.	70
6.2	The distribution of the photon shower shapes variables before and after corrections. The upper two plots are for $w_{\eta 2}$ and R_η . The lower two plots are for w_{shot} and F_{side} . The definition of these variables are shown in Figure 4.7.	73
6.3	The simulated Z mass distribution at $M_H = 125$ GeV from ggH processes for electron (left) and muon (right) channels at $\sqrt{s} = 7$ TeV (up) and $\sqrt{s} = 8$ TeV (down).	75
6.4	The observed Z mass distribution at $M_H = 125$ GeV for electron (left) and muon (right) channels at $\sqrt{s} = 7$ TeV (up) and $\sqrt{s} = 8$ TeV (down). There are two peaks on the plot. The Z peak is around 91 GeV while the low mass peak is mainly due to final state radiation (FSR) of leptons.	76
6.5	The normalized track isolation and the normalized calorimeter isolation distribution of the signal events, z +jets and $t\bar{t}$ events in muon (upper plots) and electron (lower plots) channel. The total number of events of the three samples are normalized to the same number.	78
6.6	The transverse impact parameter significance distribution of the signal events, z +jets and $t\bar{t}$ events in muon (left) and electron (right) channel. The total number of events of the three samples are normalized to the same number.	78
6.7	The observed Z mass Vs $Z\gamma$ mass distribution for muon channel at $\sqrt{s} = 8$ TeV. The final state radiation events are in the region where $Z\gamma$ mass is around 91 GeV.	79
7.1	A study of the Z boson vertex estimators. The difference between the reconstructed vertex and the true Z vertex in z is shown. The leading PV (in black) is the primary vertex with the largest Σp_T^2 and the closest PV (in red) is the primary vertex closest to the combined vertex position in z . The combined vertex is calculated by using two leptons' z_0 positions weighted by their measurement uncertainties.	83
7.2	The z position of the leading primary vertex along the beam line. The mean is at about -8 mm and the width is about 47 mm.	84

7.3	Illustration of the photon vertex correction scheme. The “0” position is the detector origin. The first (second) layer refers to the first (second) layer of the calorimeter.	85
7.4	The difference, in radians, between corrected η_x and original η_{s2} in electron channel (top) and in muon channel (bottom) for the ggh MC sample at 125 GeV for 2012.	86
7.5	The difference between corrected η_x and original η_{s2} in electron channel (top) and in muon channel (bottom) for ggh MC sample at 125 GeV for 2011. . .	87
7.6	Simulated $H \rightarrow Z\gamma$ mass reconstruction for $M_H=125$ GeV in electron channel (top) and muon channel (bottom). The red line is the reconstructed $M_{Z\gamma}$ without the photon vertex correction or the Z mass constraint applied and the green line is with the both corrections applied.	90
7.7	The reconstructed $\Delta M = M_{Z\gamma} - M_{ll}$ for simulated $H \rightarrow Z\gamma$ at mass point $M_H = 125$ GeV for the electron channel (left) and the muon channel (right). The red line is the combined fit result. The blue line is the CB component and the yellow line is the Gaussian component. Due to the energy loss of electrons by Bremsstrahlung the Gaussian component of the electron channel shifts to the right of the peak position while for the muon channel it is symmetric around the peak.	91
7.8	Signal efficiency for events produced at $\sqrt{s} = 7$ TeV. The left column is for electron channel. The right column is for muon channel. The top two plots are for ggH process and the bottom two plots are for VBF process.	92
7.9	Signal efficiency for events produced at $\sqrt{s} = 8$ TeV. The left column is for electron channel. The right column is for muon channel. The top two plots are for ggH process and the bottom two plots are for VBF process.	93
7.10	Reconstructed signal photon p_T (GeV) at different mass points from 120 GeV to 150 GeV. The black, red, green, blue, yellow, pink and cyan lines correspond to Higgs masses of 120, 125, 130, 135, 140, 145 and 150 GeV respectively.	98

7.11	Reconstructed Higgs mass at mass point $M_H=125$ GeV. The m_{cb} , σ_{cb} , a_{cb} are parameters of the CB function and refer to the mean, the width and a parameter for the CB tail respectively. The m_{gs} and σ_{gs} are parameters for Gaussian function and refer to the mean and the width respectively.	99
7.12	Reconstructed signal $\Delta M = M_H - M_Z$ at mass point $M_H=125$ GeV. The m_{cb} , σ_{cb} , a_{cb} are parameters of the CB function and refer to the mean, the width and a parameter for the CB tail respectively. The m_{gs} and σ_{gs} are parameters for Gaussian function and refer to the mean and the width respectively. . .	100
7.13	The intrinsic width of Higgs as a function of M_H	101
8.1	Feynman diagrams for $Z\gamma$ production in (a) u-channel (b) t-channel and (c) final state photon radiation (FSR) from the Z boson decay process.	104
8.2	The reconstructed $Z\gamma$ mass (left) and ΔM (right) for SM $Z\gamma$ and signal ggH at 125 GeV using MC samples. The signal $\times 50$ is shown as the red histogram.	105
8.3	The kinematic variable distributions for $H \rightarrow Z\gamma$ (in red) and SM $Z\gamma$ events (in black) for: (a) photon p_T (b) Z_{pT} (c) $\Delta R(Z, \gamma)$ (d) $\Delta\Phi$. The red and black line are the SM $Z\gamma$ background and Higgs signal events respectively. Neither of the variables is observed to be able to distinguish the two processes without loosing the signal. The two processes share the same phase space but the cross section for SM $Z\gamma$ is much larger than that for $H \rightarrow Z\gamma$. . .	106
8.4	2-D sideband method (ABCD). Region A is the signal region. Regions B, C, D are control regions. The kinematic distribution of Z +jets in signal region A is extracted from control region C. N_B to N_D is assumed to equal the ratio of N_A (without any signal events) to N_C and is thus used to calculate the background in the signal region A. The signal region is defined to have photon isolation (γ_{iso}) less than 4 GeV, and the B and D regions have photon isolation greater than 6 GeV.	107
8.5	Photon isolation for electron channel (left) and muon channel (right). Both the shape and the normalization of photon isolation for Z +jets are from control regions. The bottom of the two plots shows the data to mc ratio. .	108

8.6	The photon isolation distribution for tight and different nontight definitions for electron (left) and muon (right) channel using simulated Z +jets sample. The photon isolation in tight, non-tight (fail at least 2 out of “4 bits”) and non-tight (fail at least 2 out of “5 bits”) are shown in black, red and blue respectively. The red curve is closer to the black curve than the blue curve.	109
8.7	The distributions of $M_{Z\gamma}$ (left) and photon p_T (right) for $t\bar{t}$ events for $\sqrt{s}=8$ TeV in electron channel	113
8.8	Reconstructed $Z\gamma$ mass in the electron channel (left) and the muon channel (right). The sub-plot shows the ratio $= \frac{N_{data}}{N_{MC}}$	114
8.9	Reconstructed mass difference of $\Delta M = M_{Z\gamma} - M_Z$ in electron channel (left) and muon channel (right).	115
8.10	Spurious signal as a function of the Higgs boson mass, fitting the ΔM distribution with different models and different ranges. Pseudo-data generated from a high-statistics generator-level $Z+\gamma$ Sherpa sample. The top two plots and bottom plots correspond to 7 TeV and 8 TeV study respectively. The left two plots are the spurious signal divided by the uncertainty on the fitted signal yield. The right two plots are the statistical uncertainty (from the fit) on the signal yield.	116
8.11	Background-only model fitting on data in 2011 and 2012. Green band and yellow band correspond to 1σ and 2σ error bands respectively. The red histogram shows expected $H \rightarrow Z\gamma$ signal $\times 10$	117
9.1	Higgs cross section variations from the PDF uncertainties and the QCD scale uncertainties.	127
10.1	Expected (dashed black line) and observed (solid red line) p_0 (compatibility of the data with the background-only hypothesis) as a function of the Higgs boson mass, using 4.6 fb^{-1} of pp collisions at $\sqrt{s}=7$ TeV and 20.7 fb^{-1} of pp collisions at $\sqrt{s}=8$ TeV.	134

10.2	Observed 95% C.L. limits (solid red line) on the production cross section of a SM Higgs boson decaying to $Z\gamma$, as a function of the Higgs boson mass, using 4.6 fb^{-1} of pp collisions at $\sqrt{s} = 7 \text{ TeV}$ and 20.7 fb^{-1} of pp collisions at $\sqrt{s} = 8 \text{ TeV}$. The median expected 95% C.L. exclusion limits (dashed black line) are also shown. The green and yellow bands correspond to the $\pm 1\sigma$ and $\pm 2\sigma$ intervals.	135
B.1	Reconstructed photon transverse momentum in electron channel (left) and muon channel (right). The sub-plot shows the ratio $= \frac{N_{data}}{N_{MC}}$	145
B.2	Reconstructed electron transverse momentum. Leading electron is on the left and sub leading electron is on the right. The leading electron is the one with highest p_T . The sub leading electron is one with lower p_T . The sub-plot shows the ratio $= \frac{N_{data}}{N_{MC}}$	146
B.3	Reconstructed muon transverse momentum. Leading muon is on the left and sub-leading muon is on the right. The leading muon is the one with highest p_T . The sub leading muon is one with lower p_T . The sub-plot shows the ratio $= \frac{N_{data}}{N_{MC}}$	146
B.4	Reconstructed photon eta distribution in the electron channel (left) and muon channel (right). The dips in the plots correspond to the transition region of the barrel and the end cap of the EM calorimeter. The sub-plot shows the ratio $= \frac{N_{data}}{N_{MC}}$	147

List of Tables

1.1	Properties of leptons.	5
1.2	Properties of quarks.	6
1.3	Properties of the elementary bosons.	7
1.4	The fundamental interactions.	8
2.1	The LHC beam parameters for the nominal design, 2011 and 2012 pp collisions.	24
5.1	Integrated Luminosity for 2012 data taking periods.	64
5.2	Integrated Luminosity for 2011 data taking periods.	65
5.3	A short summary of the signal and background Monte Carlo samples for 2012 at $\sqrt{s} = 8$ TeV.	68
5.4	A short summary of the signal and background Monte Carlo samples for 2011 at $\sqrt{s} = 7$ TeV.	68
6.1	Inner Detector (ID) hit requirements for the CB and ST muon tracks. The requirements are constructed such that detector conditions are taken into account for the silicon systems: if a track traverses a dead silicon module (which is the “No. of crossed dead pixel/SCT sensors”), it is added to the hit count. nHits denotes the number of TRT hits on the muon track and nOutliers denotes the number of TRT outliers on the muon track.	71

6.2	Triggers used in 2011 and 2012. The ‘medium1’ is re-optimized to medium at higher luminosity. The ‘vh’ triggers apply a cut on the hadronic core isolation (≤ 1 GeV) at L1. The ‘i’ triggers apply a cut on track isolation at EF (PTcone20/PT). The ‘T’ is for tighter L1 requirement in order to keep the L1 rates within the allowed bandwidth without raising the threshold at the HLT.	74
7.1	Main Z boson decay modes. l indicates each type of lepton (e , μ , and τ). .	82
7.2	Higgs boson production cross section in pb for a SM Higgs boson with $m_H = 120, 125, 130, 135, 140, 145, 150$ GeV for $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV , 1211, 2-1000-1000[2[1]	94
7.3	Number of expected signal events for each production process and Higgs boson masses in 5 GeV steps between 120 and 150 GeV at $\sqrt{s} = 8$ TeV and $Z \rightarrow ee$	95
7.4	Number of expected signal events for each production process and Higgs boson masses in 5 GeV steps between 120 and 150 GeV at $\sqrt{s} = 8$ TeV and $Z \rightarrow \mu\mu$	95
7.5	Number of expected signal events for each production process and Higgs boson masses in 5 GeV steps between 120 and 150 GeV at $\sqrt{s} = 7$ TeV and $Z \rightarrow ee$	96
7.6	Number of expected signal events for each production process and Higgs boson masses in 5 GeV steps between 120 and 150 GeV at $\sqrt{s} = 7$ TeV and $Z \rightarrow \mu\mu$	96
7.7	Number of expected signal events left after each selection requirement (without applying any weights) at 120 GeV, 125 GeV and 130 GeV for the 8 TeV electron channel. The cuts are described in Chapter 6 Event selection. . .	97
7.8	Number of expected signal events left after each selection requirement (without applying any weights) at 120 GeV, 125 GeV and 130 GeV for the 8 TeV muon channel. The cuts are described in Chapter 6 Event selection.	97

8.1	non-tight definition study for electron channel 2012. The numbers in the table are the number of Z +jets events in each region. The R is defined as $R=\frac{BC}{AD}$; therefore values closer to one correspond to smaller correlations and hence indicate the optimal definition of non-tight events.	108
8.2	non-tight definition study for muon channel 2012. The numbers in the table are the number of Z +jets events in each region. The R is defined as $R=\frac{BC}{AD}$; therefore values closer to one correspond to smaller correlations and hence indicate the optimal definition of non-tight events.	109
8.3	The number of Z +jets events (N_{zjets}) for different definitions of the control regions together with the fractional amount of signal leakage into the B, C, and D regions shown in the respective columns.	111
8.4	ABCD Correlation	112
8.5	A summary of the observed data and each background component. For Z +jets, the first error is the statistical error due to the size of the Monte Carlo statistical sample and the second error is the estimated systematic error as described in the text. The error for SM $Z\gamma$ and $t\bar{t}$ are statistical errors only. The systematic error for $Z\gamma$ is fully correlated with that for Z +jets.	114
8.6	The detailed description of the functions tested in the background modeling.	117
9.1	Theoretical systematic uncertainties for the SM Higgs boson production cross section and branching fraction of the $H \rightarrow Z\gamma$ decay at the boundaries of the tested Higgs boson mass range (120 and 150 GeV). In each column, the first number is the uncertainty for $m_H = 120$ GeV, the second is the uncertainty for $m_H = 150$ GeV.	126
9.2	Summary of the systematic uncertainties on the signal yield and invariant mass distribution for $m_H = 125$ GeV, at $\sqrt{s} = 7$ TeV.	131
9.3	Summary of the systematic uncertainties on the signal yield and invariant mass distribution for $m_H = 125$ GeV, at $\sqrt{s} = 8$ TeV.	132

10.1	Expected and Observed limit for signal strength μ for Higgs mass from 120 GeV to 150 GeV.	136
A.1	Simulated samples at $\sqrt{s} = 7$ TeV. The number of events listed in the second column is either the number of events in the D3PD (for unskimmed samples) or of events in the AOD before skimming (for skimmed samples). The cross sections for the signals are from Ref. , 1211, 2-1000-1000[2[1]	142
A.2	Simulated samples at $\sqrt{s} = 8$ TeV. The number of events listed in the second column is either the number of events in the D3PD (for unskimmed samples) or of events in the AOD before skimming (for skimmed samples). The cross sections for the signals are from Ref. , 1211, 2-1000-1000[2[1]	143

Acknowledgments

There are many people who have kindly provided me help during my research work. The first is my advisor Michael Tuts. I am very grateful for his patience, guidance, insights and constantly support. Without his help, I could not come all the way to where I am. I also want to thank the other faculty members in the Columbia ATLAS group, John Parsons, Gustaaf Brooijmans and Emlyn Hughes. I benefit a lot from their challenging questions, suggestions and fruitful discussions during my graduate study.

I am grateful to my current and former Columbia classmates and postdocs for being so supportive to each other. In particular, I want to thank Andrew Altheimer, Tim Andeen, Jun Guo, Diedi Hu, Nikiforos Nikiforou, Alex Penson, Max Scherzer, Francesco Spano, Emily Thompson, Dustin Urbaniec, Eric Williams, Evan Wulf, Lei Zhou and Ning Zhou.

I feel lucky to meet so many talented people at CERN, from whom I learned a lot of knowledge and experience. I want to thank all my colleagues on the $H \rightarrow Z\gamma$ search and on the Standard Model $W\gamma/Z\gamma$ cross section measurement and the new particles searches. Especially thanks go to Louis Helary, Zhijun Liang, Giovanni Marchiori, SongMing Wang for their suggestions and inspiring discussions.

Additionally, I want to thank my high school teacher Mengjin Wang who opened the door to physics for me and showed me the beauty. I am also grateful to my undergraduate advisor Liang Han who showed me how interesting the experimental particle physics is and the whole process is just like a detective story.

Finally, I want to thank my parents for their love, support and trust. Without these, I could not follow my heart to do what I love and to be what I want to be.

for my parents

Part I

Theory

Chapter 1

The Standard Model

One of the most interesting scientific and philosophical questions we can ask is how large and how small things can be. Astrophysics answers the questions of the large scale structure of the universe. Particle physics addresses the small scale by studying the smallest building blocks of matter. Over the last sixty years, a theory to explain the nature and interaction of these building blocks, named the Standard Model (SM) of particle physics, was developed. The Standard Model has successfully explained experimental results and also made predictions that were verified later by experiments. The last missing piece of the Standard Model, the Higgs Boson, which has been searched for by experimentalists for 30 years, was discovered in July 2012 [1; 2]. In this thesis, we present a search for the decay $H \rightarrow Z\gamma$ which if observed would indicate the need for physics beyond the standard model. In the following sections, the success and limitations of the Standard Model will be introduced and the Higgs Mechanism and Higgs properties will be discussed.

1.1 The Standard Model

The Standard Model (SM) [3; 4; 5; 6] is a theory that describes the elementary particles which are the building blocks of matter as well as their interactions including the electromagnetic interaction, the weak interaction and the strong interaction. The SM combines quantum field theory and gauge theory. The gravitational interaction is not included in the SM because quantum field theories of gravity generally break down at the Planck

length¹ $l_P = 1.6 \times 10^{-35}$ m. In the SM, the elementary particles shown in Figure 1.1 are the excitations of their fields and the interaction between fields are modeled by couplings in Lagrangian. Based on their spin, these particles can be classified into two categories: Fermions (with half-integer spin) and Bosons (with integer spin). Fermions (leptons and quarks) make up the known matter in the universe and Bosons (γ , gluons, Z , $W^\pm$) mediate the interactions between these fundamental particles.

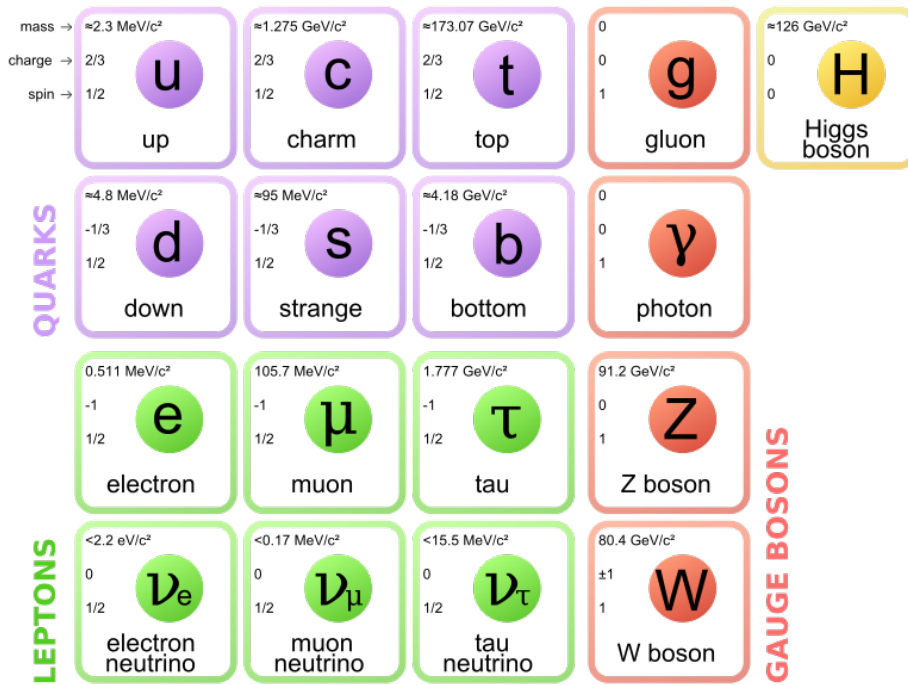


Figure 1.1: The elementary particles of the Standard Model. The quarks are shown in purple, the leptons in green, the gauge bosons in red and the Higgs boson in yellow. The quarks and leptons are also classified in three generations. The generations are represented by the three first columns in the figure. Not shown are the corresponding anti-particles of the quarks and leptons.

¹The Planck length is defined by three fundamental physical constants as $l_P = \sqrt{\frac{\hbar G}{c^3}}$ where c is the speed of light in a vacuum, G is the gravitational constant, and $\hbar$ is the reduced Planck constant.

1.1.1 Fermions

Fermions in the SM have spin one-half. They obey the Pauli exclusion principle [7], i.e. two identical fermions cannot occupy the same space at the same time. Fermions are divided into two groups, quarks and leptons. The first generation of leptons and quarks are the lightest; hence they are stable and make up ordinary matter. Protons and neutrons consist of three valence quarks, uud and udd respectively. Protons and neutrons form the nucleus, which together with electrons form atoms which in turn form molecules and hence ordinary matter.

1.1.1.1 Leptons

The electron was discovered in 1897 by J. J. Thomson [8]. The other two charged lepton flavors are called the muon, which was discovered by Carl D. Anderson in 1936 [9] and the tau, discovered around 1975 by Martin Perl and collaborators from the Stanford Linear Accelerator Center and Lawrence Berkeley National Laboratory [10]. These heavier leptons, muons and taus, have a short life time of about 10^{-6} s and 10^{-13} s respectively. They decay into electrons through a weak interaction. For example, the muon decays to an electron, an anti-electron-neutrino and a muon neutrino. Lepton number is conserved during this process. Lepton number is an additive quantum number. For leptons, each flavor has its own lepton number (denoted by L_e , L_μ , L_τ) with a value of +1, while the corresponding anti particles have a lepton number of -1 . The neutral leptons are called neutrinos, denoted by ν . There are three flavors of neutrinos: the electron neutrino, the muon neutrino and the tau neutrino. The charged leptons carry integral electric charge and weak isospin. The weak isospin (T_3) is to the weak interaction what electric charge is to electromagnetism, and what color charge is to the strong interaction. Therefore charged leptons undergo both electromagnetic interactions and weak interactions. Neutrinos are neutral (hence have zero electric charge and no electromagnetic interaction) but do carry weak isospin and therefore neutrinos only interact via the weak interaction. Neutrinos in the Standard Model are massless but experiments show that the mass is small but not zero which is still an active area of research [11; 12]. Table 1.1 lists their properties. The unit of mass can be expressed using MeV in natural unit instead of MeV/c^2 . In the following text, the natural unit is

used to express the mass and momentum.

Table 1.1: Properties of leptons.

	Particle/Antiparticle	Electric charge (e)	Spin	Mass (MeV)
1 st generation	e^-/e^+	$-1/+1$	$\frac{1}{2}$	0.511
	$\nu_e/\bar{\nu}_e$	0/0	$\frac{1}{2}$	$< 2.2 \times 10^{-6}$
2 st generation	μ^-/μ^+	$-1/+1$	$\frac{1}{2}$	105.7
	$\nu_\mu/\bar{\nu}_\mu$	0/0	$\frac{1}{2}$	< 0.17
3 st generation	τ^-/τ^+	$-1/+1$	$\frac{1}{2}$	1776.8
	$\nu_\tau/\bar{\nu}_\tau$	0/0	$\frac{1}{2}$	< 15.5

1.1.1.2 Quarks

Quarks carry fractional charges of $+\frac{2}{3}|e|$ or $-\frac{1}{3}|e|$. There are six flavors of quarks including the up quark (u), the down quark (d), the charm quark (c), the strange quark (s), the top quark (t) and the bottom quark (b). The difference between quarks and leptons is that quarks carry color charges (Red, Green and Blue) besides electric charge and weak isospin. The anti-quarks carry anti-color charges (Anti-red, Anti-green and Anti-blue). The color/anti-color charges are responsible for their strong interactions which are mediated by gluons. Gluons are massless particles carrying color charge as well. They will be introduced in the next section. The name for the up and down quarks comes from isospin symmetry ($I=\frac{1}{2}$), corresponding to the two components ($\pm\frac{1}{2}$). Quarks combine to form hadrons (composite particles). Due to the color confinement phenomenon, quarks cannot be observed directly or found in isolation. They can be found only within hadrons (except for the top quark which decays before forming a hadron). Hadrons are colorless objects. Hadrons include baryons and mesons. Baryons are composed of three quarks while mesons are composed of one quark and one antiquark. The most well known baryons are protons and neutrons. Pions ($\pi^\pm, \pi^0$) are examples of mesons.

Table 1.2: Properties of quarks.

	Particle/Antiparticle	Electric charge (e)	Spin	Mass (MeV)
1^{st} generation	$u/\bar{u}$	$+\frac{2}{3}/-\frac{2}{3}$	$\frac{1}{2}$	1.7 to 3.1
	$d/\bar{d}$	$+\frac{1}{3}/-\frac{1}{3}$	$\frac{1}{2}$	4.1 to 5.7
2^{st} generation	$c/\bar{c}$	$+\frac{2}{3}/-\frac{2}{3}$	$\frac{1}{2}$	1290^{+50}_{-110}
	$s/\bar{s}$	$+\frac{1}{3}/-\frac{1}{3}$	$\frac{1}{2}$	100^{+30}_{-20}
3^{st} generation	$t/\bar{t}$	$+\frac{2}{3}/-\frac{2}{3}$	$\frac{1}{2}$	$172900 \pm 600 \pm 900$
	$b/\bar{b}$	$+\frac{1}{3}/-\frac{1}{3}$	$\frac{1}{2}$	4190^{+180}_{-60}

1.1.2 Bosons

The elementary bosons in the SM have integer spin and therefore obey the rules of Bose-Einstein statistics [13]. They can have identical quantum numbers, being in the same place at the same time. There are three kinds of interactions in the SM and each has a set of gauge bosons responsible for the corresponding interaction. The photon (γ) is the gauge boson for the electromagnetic interaction, the $W^\pm$ and Z are the gauge bosons for the weak interaction and gluons (g) are the gauge bosons for the strong interaction. According to gauge theory, the gauge bosons should be massless² and long-ranged. However, this is not supported by experiments in which the weak interactions have been observed to have a very short range. Therefore a scheme named the Higgs mechanism [14; 15; 16; 17; 18; 19] was introduced to break $SU(2) \times U(1)$ electroweak symmetry and give mass to the $W^\pm$ and Z bosons while the photon remains massless. The Higgs mechanism will be introduced in detail in the next section.

1.1.3 Interactions

Three fundamental interactions (the electromagnetic interaction, the weak interaction and the strong interaction) are included in the SM. The electromagnetic interaction has been

²The mass term for gauge field is not allowed in the Lagrangian because the mass term is not gauge invariant under gauge transformation.

Table 1.3: Properties of the elementary bosons.

Gauge Bosons	Electric charge (e)	Spin	Mass (GeV)
Photon (γ)	0	1	0
$W^\pm$	± 1	1	80.385 ± 0.015
Z	0	1	91.1876 ± 0.0021
Gluon (g)	0	1	0

known for a long time. This force acts between electrically charged particles. Electromagnetism is a long-ranged force like gravity, but stronger. It is described by Quantum electrodynamics (QED) which is a theory of how light and matter interact, incorporating both quantum mechanics and special relativity.

Weak interactions are not commonly seen in everyday life. The interaction happens via the exchange of $W^\pm$ or Z bosons. The weak force and electromagnetic force look different at low energy, but at energy scales of 100 GeV, the two forces behave in a similar way and hence a unified theory called the electroweak theory (EWT) can be formulated to describe the interactions. The Gauge group for EWT is $SU(2) \times U(1)$. The three generators of $SU(2)$ correspond to the gauge bosons $W^\pm$ and Z and the one generator of $U(1)$ corresponds to the photon. The EWT successfully predicted the mass of the W and Z bosons as well as the existence of a Higgs field and the associated Higgs boson. Recently, this Higgs boson, which is an excitation of the Higgs field, was discovered by the ATLAS and CMS experiments [1; 2].

Strong interactions are responsible for holding the quarks together in the proton and the neutron and as well as allowing protons and neutrons to form a nucleus of an atom. The Quantum chromodynamics (QCD) is used to describe this interaction. Its gauge group is $SU(3)$. There are 8 generators of the group, and therefore 8 types of gluons which mediate this interaction. Both quarks and gluons carry “color charge” which is analogous to the electric charge for the electromagnetic interaction. A list of the fundamental interactions is shown in Table 1.4.

Table 1.4: The fundamental interactions.

Interaction	Mediators	Relative strength	Range (m)
Gravitational	gravitons (hypothetical)	1	infinity
Weak	W and Z bosons	10^{25}	10^{-18}
Electromagnetic	Photons	10^{36}	infinity
Strong	Gluons	10^{38}	10^{-15}

1.1.4 Limitations of the Standard Model

The Standard Model has been widely tested by experiments in the last few decades and so far almost all the experimental results agree with the Standard Model predictions. Despite the success of the SM, there are still some questions left which are not explained satisfactorily by the SM theory. For example, the SM theory is incompatible with general relativity which is the most successful theory of gravity by far. Therefore gravity is not explained by the SM. Moreover, neutrinos are massless according to the SM. However, neutrino oscillation experiments have shown that neutrinos have a small but nonzero mass. In addition, the SM does not explain what particles make up the dark matter in the universe or what dark energy is. Theoretically, there are a few imperfections such as the fact that the SM has nineteen numerical parameters which can only be determined by experiment. The SM does not provide a dynamical reason for the values of the parameters and/or a relation between those parameters. Moreover a “fine-tuning” problem also persists. The fine-tuning problem is that the parameters of a model must be adjusted very precisely in order to agree with observations. The fundamental parameters (couplings or masses) of some Lagrangian are very different from the parameters measured by experiment. The quantum corrections to the Higgs mass are vastly larger than the measured Higgs mass, so there is a fine-tuned cancellation between the bare mass and quantum corrections in order to agree with the measured mass. Such an accidental amount of fine-tuning is also referred to as the “naturalness problem”.

1.2 The Higgs Mechanism

1.2.1 Spontaneously Symmetry Breaking and the Higgs boson

The SM theory is based on gauge symmetry; however there is one difficulty in the electroweak theory, namely that the experimentally observed weak gauge bosons are massive and short-ranged. Therefore, on the one hand, one wants to preserve the symmetry of the equations of motion or the Lagrangian of the physical system under gauge transformations but, on the other hand, one wants the lowest energy solutions not to exhibit that symmetry. The solution to this problem is known as “spontaneous symmetry breaking” and the mechanism that realizes this is referred to as the Higgs mechanism [14; 15; 16; 17; 18; 19]. The Higgs mechanism introduced a complex doublet scalar field Φ with the potential shown in Figure 1.2

$$V(\Phi) = -\mu^2 \Phi^\dagger \Phi + \frac{\lambda}{2} (\Phi^\dagger \Phi)^2 \quad (1.1)$$

where $\mu^2 > 0$ and $\lambda > 0$. The vacuum expectation value $\nu = \sqrt{\frac{\mu^2}{\lambda}}$ is nonzero. It is a spontaneously broken symmetry because a symmetrical potential ends up in an asymmetrical ground state.

The Higgs field Φ has two complex components (or four real components, two neutral ones and two charged). Three (one neutral two charged) of the four components, together with the three weak gauge bosons will form the massive $W^\pm$ and Z . The fourth neutral component of the Higgs field corresponds to the massive scalar Higgs. One can see the quadratic mass term by expanding the potential $V(\Phi)$ near Φ_0 where the minimum is achieved. However, the mass term $\sqrt{2}\mu$ is a free parameter, so it is not specified by the SM theory and has to be determined by experiment. The experimental search for and final discovery of the Higgs boson and the measurement of its mass has taken decades. A 2012 compilation of older experimental results on indirect limits on the mass of the Higgs boson led to an upper limit of $m_H < 158$ GeV at 95% confidence level (CL) by using global fits to precision electroweak results [36]. As shown in Figure 1.3 [37], the allowable range for the Higgs mass has been narrowing based on results from LEP [38], the Tevatron [39; 40; 41] and the LHC [42; 43], leaving only a small allowed window for a Higgs of mass between 115 GeV to 127 GeV. On July 4th 2012, one Higgs-like boson was discovered at 125 GeV

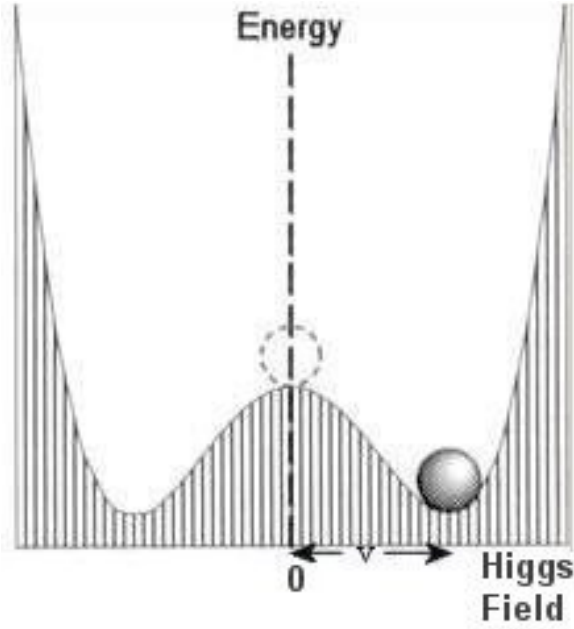


Figure 1.2: Higgs Potential

to 126 GeV by the ATLAS and CMS experiments [1; 2]. Based on this experimental discovery by ATLAS and CMS and subsequent measurements of this Higgs-like boson, Higgs and Englert were awarded the 2013 Nobel prize in physics for their theoretical work in proposing the Higgs mechanism. The discovery of the Higgs boson has started a new era of Higgs physics. Precision measurements of the Higgs properties, eventually including the measurement of the $H \rightarrow Z\gamma$ decay mode searched for in this thesis, will confirm whether this is the SM Higgs or whether more precise measurements will point to physics beyond the SM.

1.2.2 The Higgs boson properties

In the Standard Model, the Higgs field is a scalar field and the Higgs boson is spin 0 particle with zero electric charge and color charge. It is its own antiparticle. The Higgs mass is not predicted by the Standard Model but the coupling of Higgs to fermions and to gauge bosons is determined by the Standard Model as a function of the Higgs mass. The strength of the coupling is proportional to the fermion mass and to the gauge bosons mass (W or Z). In the

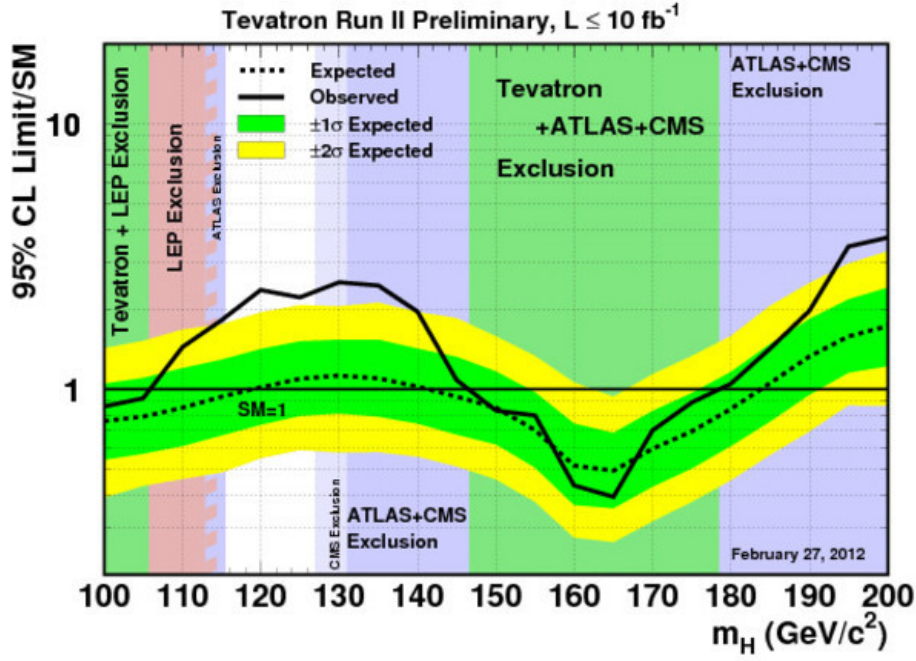


Figure 1.3: Higgs Mass Exclusion limit from Tevatron by February 27, 2012. The dotted line is the expected limit on the signal strength ($\frac{\sigma}{\sigma_{SM}}$). The solid black line is the observed limit. The green and yellow band around the expected limit correspond to $\pm 1\sigma$ and $\pm 2\sigma$ respectively. The M_H region whose observed limit is below 1 can be excluded. Most of the M_H region between 100 GeV and 200 GeV were excluded by LEP, Tevatron, ATLAS and CMS experiments except for the region between 115 GeV to 127 GeV.

following sections, the Higgs production process and decay processes will be introduced.

1.2.2.1 The Higgs boson Production

The Higgs boson can be produced in several processes including the gluon-gluon fusion process, the vector boson fusion process, a Higgs-strahlung process and associated production with top quark pairs. The gluon-gluon fusion process, whose Feynman diagram is shown in Figure 1.4(a), has the largest cross-section, as shown in Figure 1.5. In this process, the two gluons do not couple directly to the Higgs but through a fermion loop. Since the coupling of fermions to the Higgs is proportional to the fermion mass, the top quark (which has the heaviest mass) loop will therefore have the largest contribution. If there are fourth generation quarks, which is an active research area in the ATLAS Exotic group, we would see an enhancement in the Higgs production. So precision Higgs physics can further the understanding of new physics beyond Standard Model.

Another important Higgs production process is the vector boson fusion process shown in Figure 1.4(b), where two gauge bosons (W or Z) radiated from incoming quarks collide with each other to produce a Higgs and the two incoming quarks fragment into two jets in the forward region. The two forward jets plus a Higgs form a unique signature which can be used to reduce background events significantly.

Another process is Higgs-strahlung (VH) shown in Figure 1.4(c). A quark and antiquark collide with each other to produce a gauge boson (W or Z) which radiates a Higgs boson. The final state is one real W or Z and a Higgs boson.

The last process we consider is associated Higgs production with top-quark pairs ($t\bar{t}H$) as shown in Figure 1.4(d). Each of the two gluons produce a pair of top and anti top quarks, followed by one anti-top quark annihilating with a top quark from the other pair to produce a Higgs boson.

1.2.2.2 The Higgs Decay

The SM Higgs boson can decay to fermion pairs (e.g. $b\bar{b}$, $\mu^+\mu^-$, $\tau^+\tau^-$) and gauge boson pairs (W^+W^- , ZZ) directly and could also decay to two photons ($\gamma\gamma$) and a Z boson plus a photon ($Z\gamma$) mainly through top quark loop and W boson loop. The high mass region

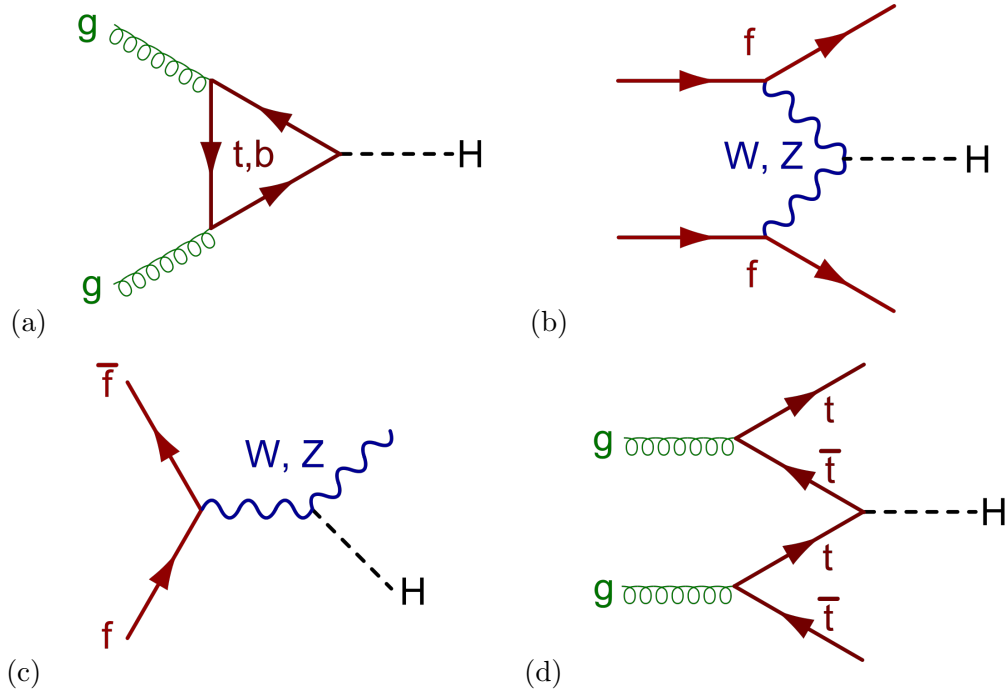


Figure 1.4: Feynman diagram for Higgs production processes (a) Gluon-Gluon Fusion (b) Vector Boson Fusion (c) Higgs-strahlung (VH) (d) associated production with top-quark pairs ($t\bar{t}H$)

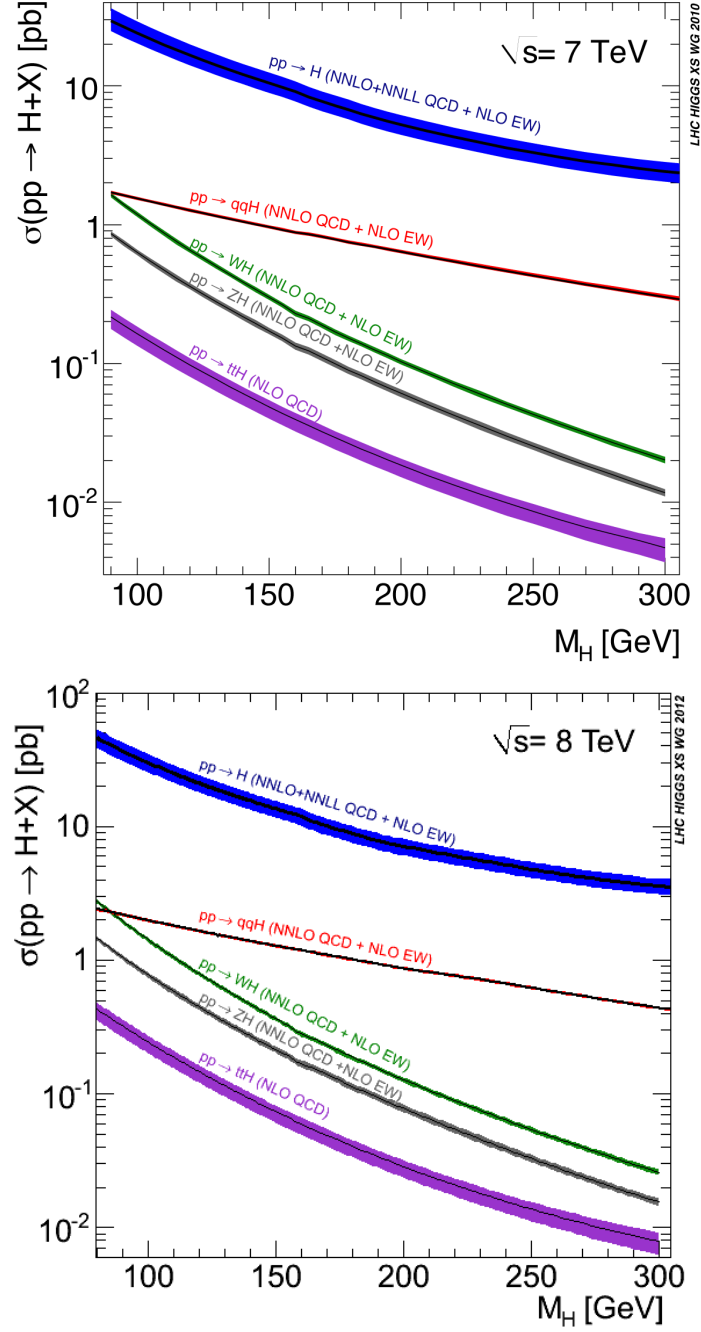


Figure 1.5: Higgs Production Cross-section at 7 TeV (Top plot) and 8 TeV (Bottom plot). The blue curve with the highest cross-section corresponds to the gluon-gluon fusion process. The curve with the second largest cross-section corresponds to the vector boson fusion process. The other three curves correspond to WH , ZH and $t\bar{t}H$ process respectively. The bands are the theoretical uncertainty from the QCD scales uncertainties and PDFs+ α_s uncertainties.

$m_H > 135$ GeV is dominated by $H \rightarrow WW$ and the low mass region, $m_H < 135$ GeV, is dominated by $H \rightarrow b\bar{b}$. However, considering the huge QCD background for $b\bar{b}$, the most sensitive channels are $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ$ which are also good for Higgs mass measurements because of the excellent electromagnetic energy resolution of the ATLAS detector. In Figure 1.6, the branching ratio of Higgs decay processes are shown.

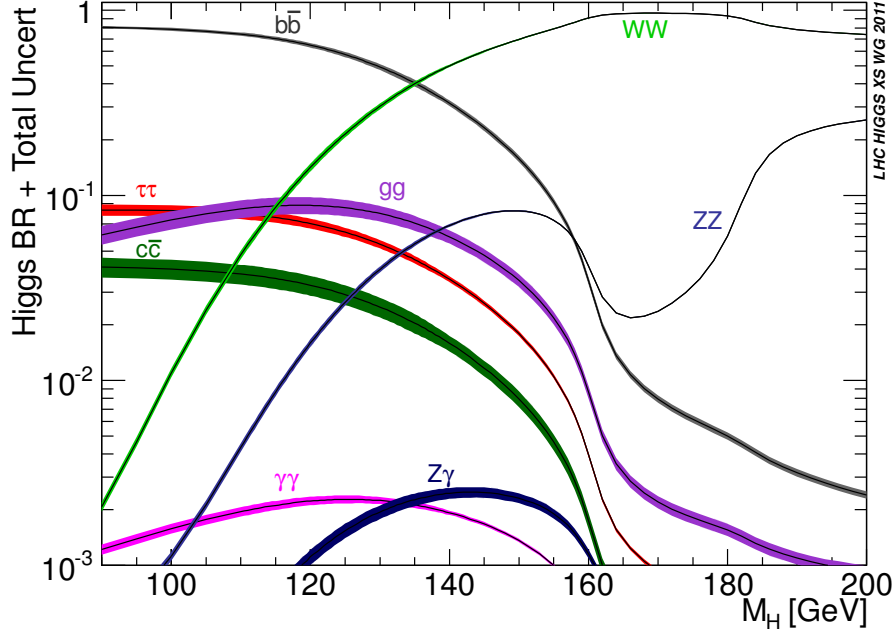


Figure 1.6: Higgs decay branching ratio. The bands represent the theoretical uncertainties.

1.3 Motivation

The Standard Model (SM) has shown its predictive power, as evidenced by the discovery of W and Z bosons and many other electroweak precision measurement results. One last missing piece is the evidence of the electroweak spontaneous symmetry breaking mechanism; that is the Higgs boson. On July 2012, a Higgs-like boson was first discovered in the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ$ channels at 125-126 GeV [1; 2]. A natural question to ask is : is this the SM Higgs boson? We will be able to answer this question by measuring this particle's spin [20; 21] and its other decay channels [22; 23; 24; 25; 27] as predicted by the SM, such as $H \rightarrow Z\gamma$

which can also be used to precisely reconstruct the Higgs mass and by measuring the coupling of the Higgs to gauge bosons and fermions since in the framework of the SM these couplings are fixed. Any discrepancy of the coupling from the SM prediction could be a signal of new physics beyond the SM [28; 29; 30; 31]. In Figure 1.7, the ATLAS result for $H \rightarrow \gamma\gamma$ shows a possible higher value than the SM prediction [1]. One possibility is that new heavy particles exist in the decay loop of $H \rightarrow \gamma\gamma$ shown in Figure 1.8. It is worth noting that $H \rightarrow Z\gamma$ in Figure 1.9 has a similar loop production process to $H \rightarrow \gamma\gamma$. Therefore by searching for Higgs through the $Z\gamma$ channel, on the one hand we can check SM Higgs mechanism for another process, and on the other hand by measuring the ratio of the signal strength of $H \rightarrow \gamma\gamma$ to $H \rightarrow Z\gamma$, we can get a deeper understanding of the electroweak symmetry breaking mechanism [28; 31].

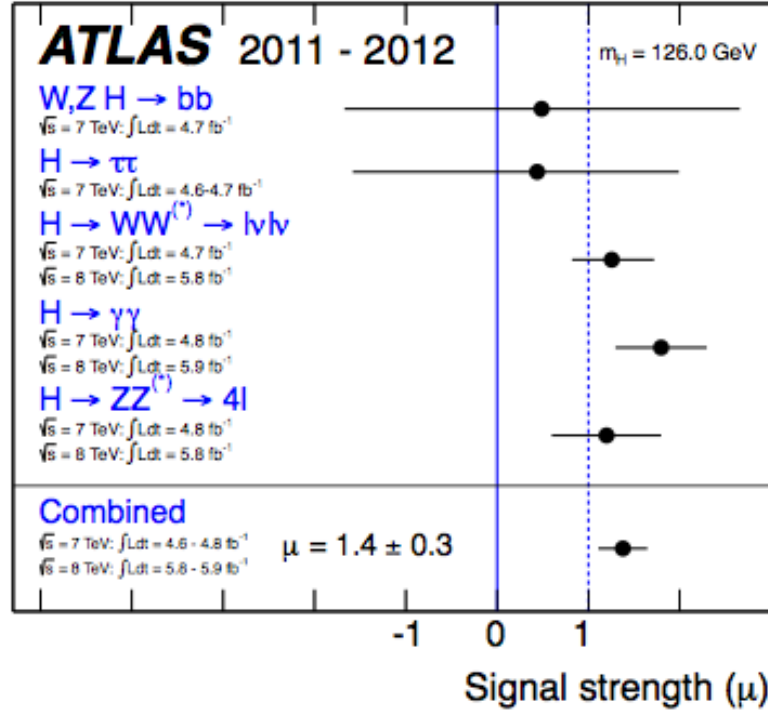


Figure 1.7: Higgs signal strength $\mu = \frac{\sigma}{\sigma_{SM}}$ from ATLAS [1].

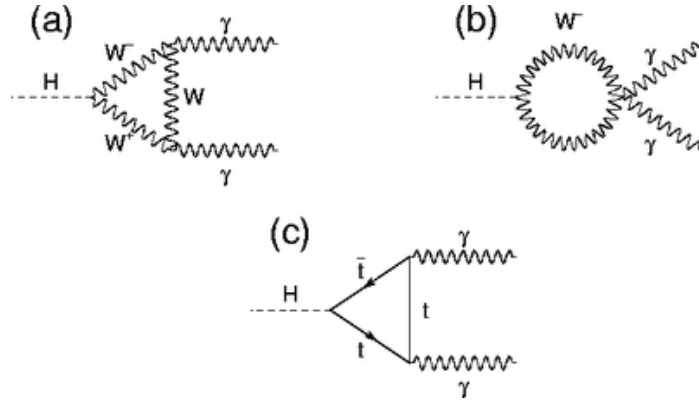


Figure 1.8: Leading order Feynman diagrams for the Higgs decaying to $\gamma\gamma$.

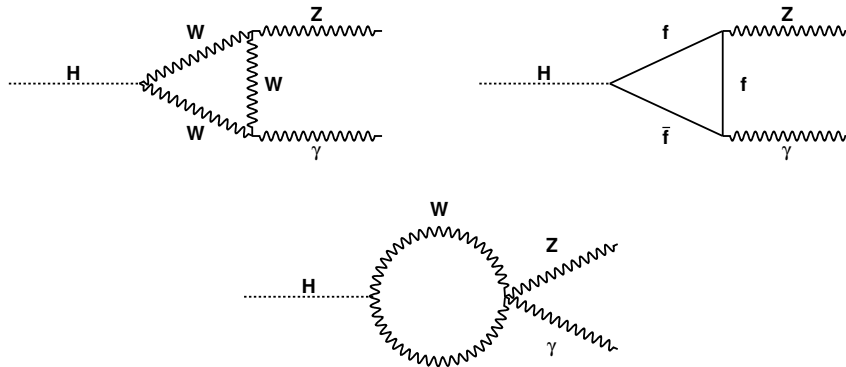


Figure 1.9: Leading order Feynman diagrams for the Higgs decaying to $Z\gamma$.

Part II

Experiment

Chapter 2

The Large Hadron Collider

Colliders are an important research tool of studying sub-atomic particles. With respect to accelerating particles, colliders can be either linear accelerator or circular accelerator. With respect to beams, colliders may collide a single beam of particles against a stationary target or two beams head-on. The two key parameters for a colliding beam accelerator are the center of mass energy and the number collisions. In the Large Hadron Collider (LHC) [26], two beams of protons are accelerated in a ring and collide with each other head-on.

The advantages of a colliding beam accelerator are

- the ability to produce higher mass particles than fixed-target experiment because the beam energies are maximumly converted to make new particles in the center of momentum frame.
- can achieve more collisions at the highest energy than fixed-target experiment

The LHC and the experiments were designed to explore the energy frontier and discover the source of electroweak symmetry breaking as well potential physics beyond the standard model such as Supersymmetry, extra-dimension, dark matter, etc.

2.1 LHC introduction

The Large Hadron Collider (LHC) shown in Figure 2.1, is currently the largest and highest-energy particle accelerator in the world. The LHC is 27 kilometers in circumference and

is located 100 meters under the French-Swiss border (the depth varies between 175 m and 50 m), near Geneva, Switzerland. The LHC collides protons on protons in contrast to the Fermilab Tevatron that collides protons on anti-protons. The design center of mass energy of the LHC is 14 TeV (corresponding to 7 TeV per beam) at a peak instantaneous luminosity of $10^{34} \text{ cm}^{-2}\text{s}^{-1}$. The luminosity is a measure of the number of collisions and is defined in section 2.3. From 2010 to 2012 the LHC initially ran at a center of mass energy that was about half of design (7 TeV), increasing to 8 TeV with the peak instantaneous luminosity increasing from 10^{32} to $10^{33} \text{ cm}^{-2}\text{s}^{-1}$ over that period. LHC p-p data taking ended on December 17th 2012. Starting in 2013, the LHC was shut down in order to do maintenance and upgrade the accelerator. In 2015, the LHC will start up at a center of mass energy of ~ 13 TeV and peak instantaneous luminosity of $\sim 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. There are four experiments located at the four interaction point of LHC, including two general purpose detectors ATLAS and CMS and two special purpose experiments ALICE (quark-gluon plasma) and LHCb (b-physics).

2.2 Beam acceleration and focusing

In order to collide two 4 TeV proton beams, the acceleration process begins with a bottle of hydrogen gas. Figure 2.2 shows the accelerator chain after the electrons have been stripped off the hydrogen to become protons. The protons are initially accelerated to 50 MeV in a linear accelerator (LINAC). The proton beams are then injected into the Proton Synchrotron Booster (PSB), which accelerates the protons to 1.4 GeV, followed by the Proton Synchrotron (PS), which raises the beam energy to 25 GeV. Protons are then sent to the Super Proton Synchrotron (SPS) where they are accelerated to 450 GeV. After this step, the proton beams are ready to enter the 27 km circumference beam pipe where they will be accelerated to 4 TeV per beam. One of the reasons that the protons have to go through several increasingly larger accelerators is that the radius of the 27 km beam pipe is fixed and the range of the magnetic field is limited¹.

In order to bend the proton beam around the 27 km circumference LHC ring, a vertical

¹For example, the magnetic field ranges from about 0.54 T to 8.3 T in 7 TeV.

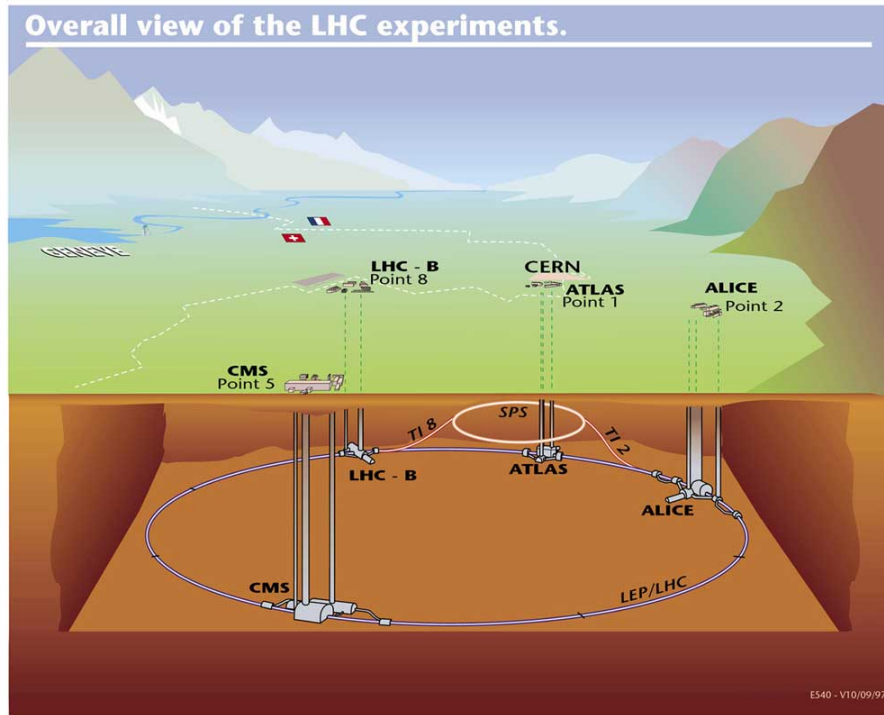


Figure 2.1: LHC is located 100 meters underground between Lake Geneva and the Jura mountains. The locations of the four detectors ATLAS, CMS, ALICE and LHCb can be seen on the plot. Point 1 is where the ATLAS detector is located.

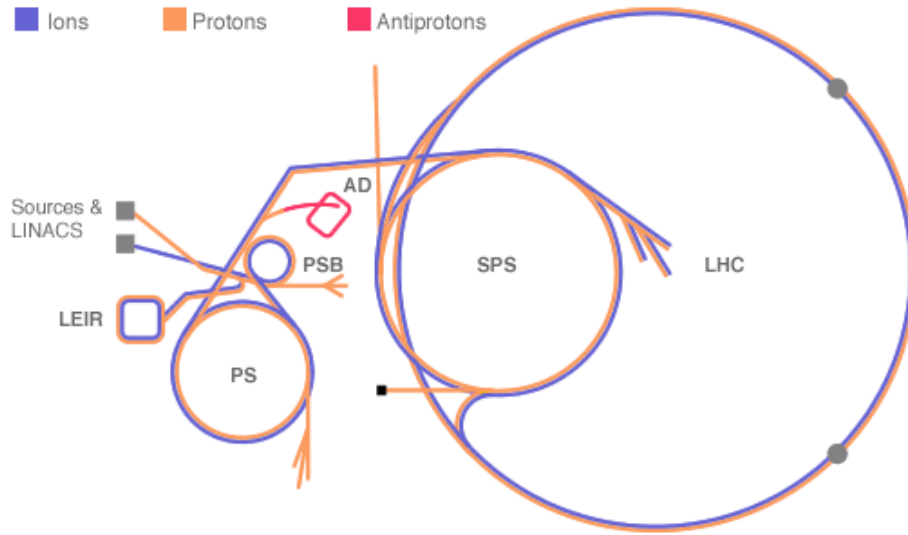


Figure 2.2: The CERN accelerator complex from the LINAC to the LHC. Also indicated by the colors are the accelerators used to accelerate heavy ions (blue), protons (orange) and antiprotons (red).

magnetic field is required to bend the beams. It is produced by the superconducting dipole magnets shown in Figure 2.3. The proton beams in the two apertures, which have to travel in opposite directions, pass through a vertical magnetic field produced by the dipole magnets. At 7 TeV these magnets have to produce a field of around 8.4 Tesla at a current of around 11,700 A. Each one is 14.3 meters long with weight of about 35 tons. A total of 1232 are needed.

In order to constrain the proton beams' width and height within the vacuum pipe, quadrupole magnets are used, shown in Figure 2.4. Quadrupole magnets are used as alternating gradient pairs. One constrains the width and the other one constrains the height. With the pairs of quadrupole magnets, the beam will oscillate around the center of the beam pipe. In total, about 392 quadrupole magnets are needed, each one 5-7 meters long.

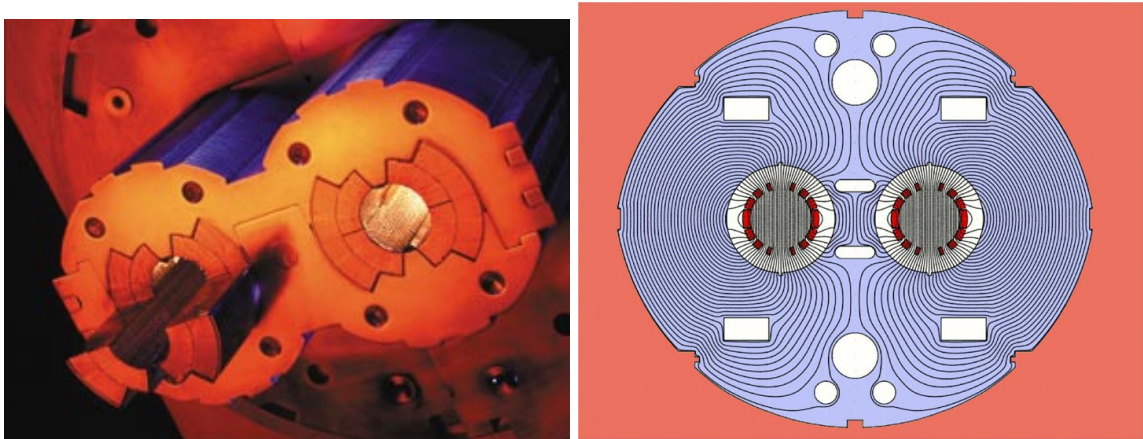


Figure 2.3: LHC dipole magnets (left) and the dipole magnetic flux (right).

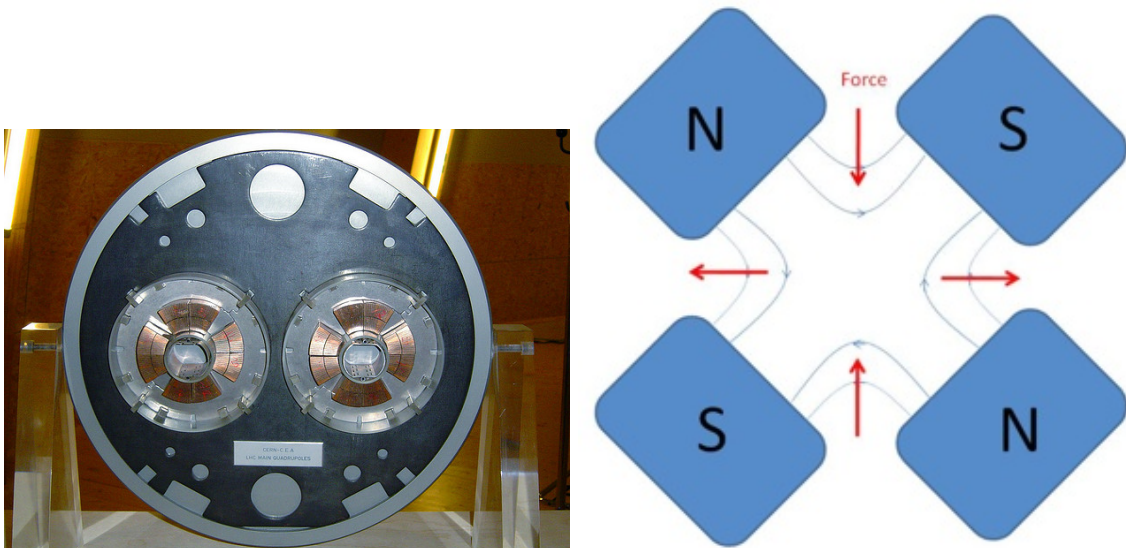


Figure 2.4: LHC Quadrupole magnets (left) and Quadrupole magnetic flux (right).

2.3 Luminosity

The Luminosity is a measure of the number of collisions that can be produced per cm^2 per second. It is an important parameter for a colliding beam accelerator with a direct impact on the expected number of collisions produced since the reaction rate is given by $R = L \times \sigma_{inel}$. The luminosity is defined as

$$L = f \times n \times \frac{N_1 \times N_2}{4\pi\sigma_x\sigma_y} = f \times n \times \frac{N_1 \times N_2}{4\epsilon\beta} \quad (2.1)$$

where:

- f is the revolution frequency;
- n is the number of bunches in one beam;
- N_{12} is the product of the particles in the two colliding bunches;
- $\sigma_x \times \sigma_y$ is the product of the transverse beam profiles in the horizontal and vertical directions;
- ϵ is the transverse emittance of a beam with units of a length. It is a measure for the average spread of particle coordinates in position-and-momentum phase space. It depends on the initial beam condition. The low emittance indicates that particles are confined to a small distance and have nearly the same momentum;
- β is the amplitude function with units of a length. It is determined by the accelerator magnet configuration. Low β indicates a narrower beam while high β indicates that the beam is wide and straight.

The amplitude function β , at the interaction point (IP), is called β^* and is shown in Figure 2.5. From Equation 2.1 one can see that luminosity L increases with n and $N_{1/2}$ and inversely with ϵ and β . The revolution frequency is determined by the accelerator design which is not easy to tune. From 2010 to 2012, the $N_{1/2}$ was increased from 1.2×10^{11} /bunch to 1.6×10^{11} /bunch and n was increased from 368 to 1380 and β^* was decreased from 3.5 m to 0.6 m. The designed LHC beam parameters and also the one for 2011 and 2012 runs are shown in Table 2.1.

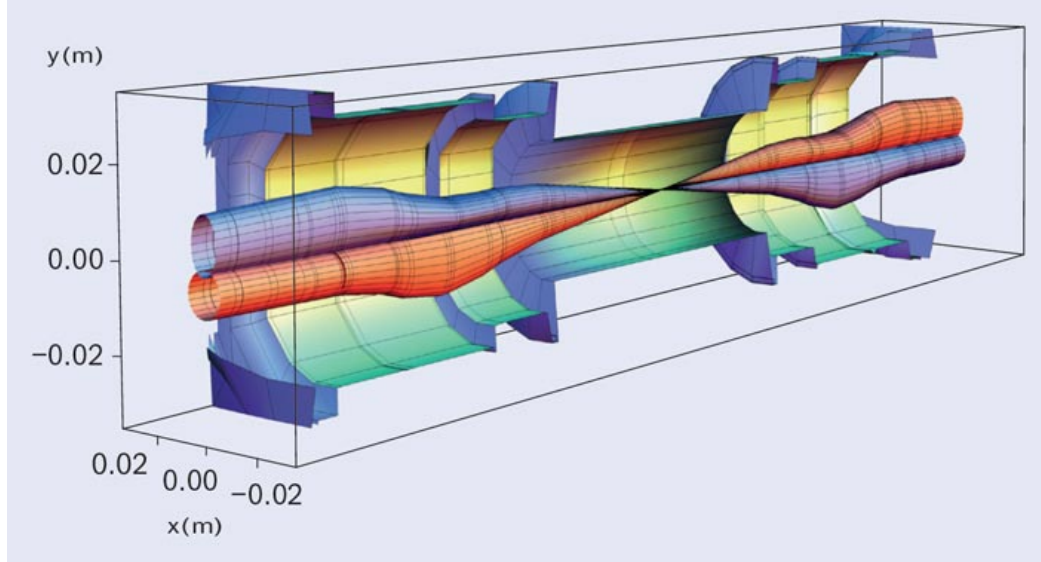


Figure 2.5: Beams interaction at ATLAS. The beam size at the interaction point is about $16 \mu\text{m}$.

Parameters	2011 runs	2012 runs	Design
Beam energy [TeV]	3.5	4	7
Number of bunches per beam	1380	1380	2808
Number of protons (10^{11} p/bunch)	1.5	1.6-1.7	1.5
β^* [m]	$1.5 \rightarrow 1$	0.6	0.55
Transverse Emittance ϵ [μm]	2.5	2.5	3.75
Bunch spacing [ns]	50	50	25
Peak instantaneous luminosity [$\text{cm}^{-2}\text{s}^{-1}$]	3.5×10^{33}	7.6×10^{33}	10^{34}
Average interactions per bunch crossing	~ 10	~ 20	23

Table 2.1: The LHC beam parameters for the nominal design, 2011 and 2012 pp collisions.

In order to have higher luminosity, the beam is squeezed just before the interaction point by using the so called “inner triplet” magnets which are a set of quadrupole magnets. They squeeze the beam size to about $16 \mu\text{m}$. The ATLAS daily integrated luminosity is shown in Figure 2.6.

The higher the luminosity, the larger the pileup. The term pileup refers to the situation where either there are multiple proton-proton interactions in a single bunch crossing (known as in-time pileup) or there are multiple proton-proton interactions in neighboring bunch crossings (known as out-of-time pileup). The dominant pileup contribution that can affect detector efficiencies and resolutions is the in-time pileup. In this analysis, we want to select isolated electrons and photons by requiring the energy deposited around the object less than a threshold. The in-time pileup tends to increase the measured energy and out-of-time pileup tends to decrease this energy [32]. This effect has been studied by a dedicated performance group and a tool was developed to correct the isolation energy in order to mitigate the pileup effect. The mean number of interactions per crossing, $\langle \mu \rangle$, which is a measure of the pileup, is shown in Figure 2.7 for the 7 TeV and 8 TeV running periods.

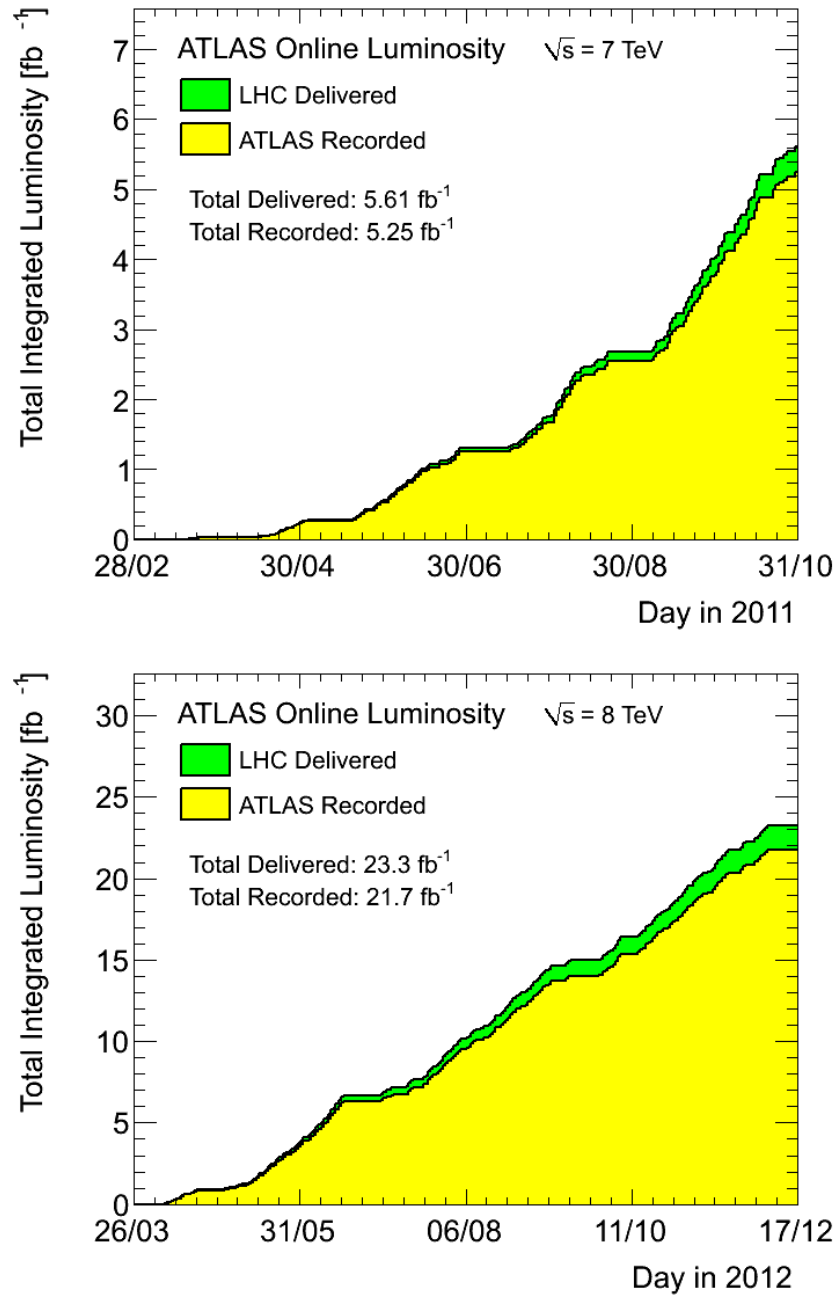


Figure 2.6: ATLAS daily integrated luminosity in 2011 (upper plot) and in 2012 (lower plot). The green area shows the delivered luminosity to ATLAS by the LHC and the yellow area shows the luminosity recorded by ATLAS. The recorded luminosity is less than the delivered luminosity either because the ATLAS detector data acquisition chain was busy or because one or more of its sub-detector systems are temporarily unavailable.

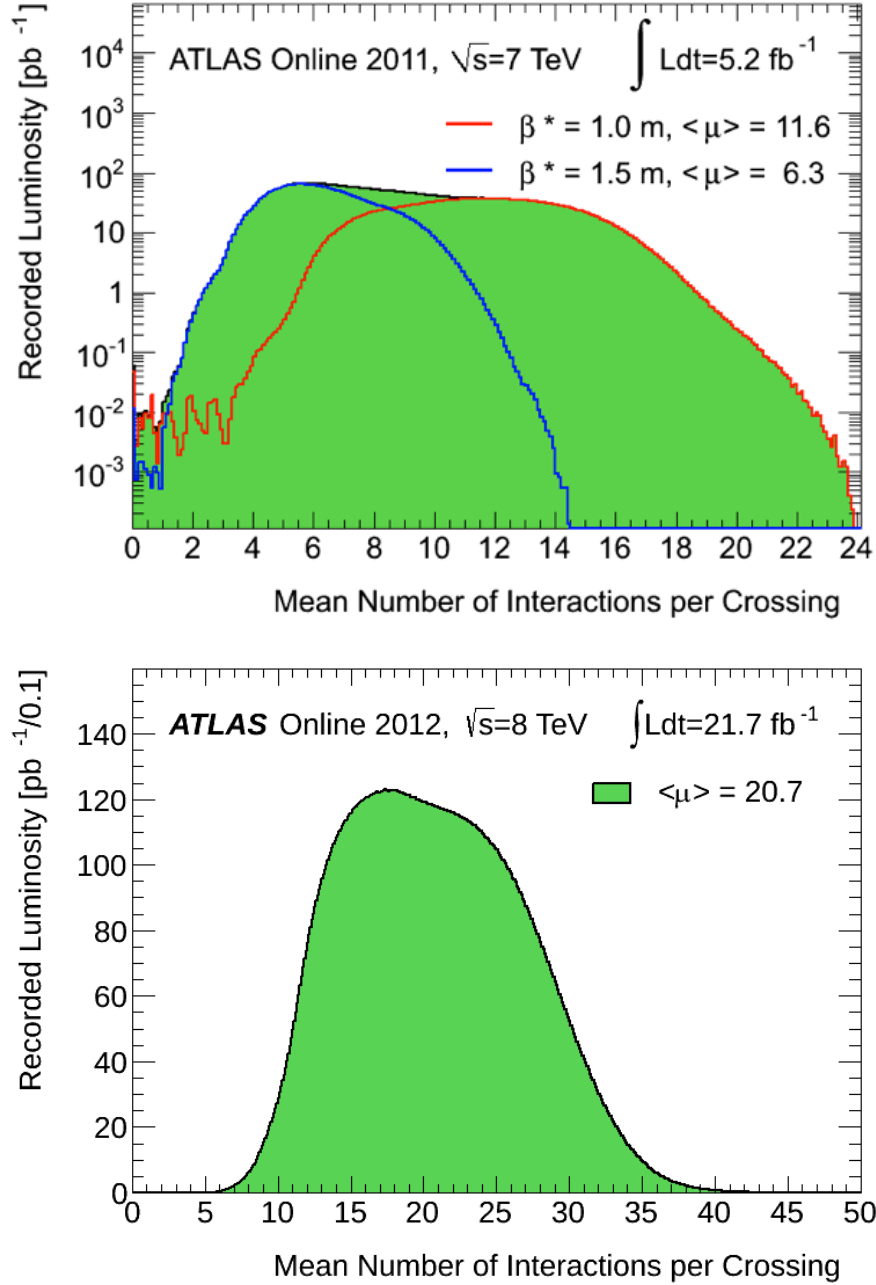


Figure 2.7: The luminosity-weighted distribution of the mean number of interactions per crossing (μ) for 2011 (upper plot) and 2012 (lower plot). The integrated luminosities and the $\langle \mu \rangle$ values are given in the figure. The mean number of interactions per crossing corresponds the mean of the poisson distribution on the number of interactions per crossing for each bunch. It is calculated from the instantaneous per bunch luminosity as $\mu = \frac{L_{bunch} \times \sigma_{inel}}{f_r}$ where L_{bunch} is the per bunch instantaneous luminosity, σ_{inel} is the inelastic cross section which is 71.5 mb (2011) and 73 mb (2012) and f_r is the LHC revolution frequency.

Chapter 3

ATLAS detector

The ATLAS detector [33; 34; 35] shown in Figure 3.1 is a multi-purpose detector located at Point 1 of the Large Hadron Collider (LHC) at CERN. It has an overall cylindrical shape with a width of 44 meters and a diameter of 22 meters. ATLAS weighs 7000 tons. The main sub-detector systems are the Inner detector, the Calorimeter and Muon systems. The Inner detector includes the pixel detector, the silicon microstrip tracker and the transition radiation tracker, and measures the trajectories and momenta of charged particle tracks. The Calorimeter includes the Electromagnetic (EM) Calorimeter and the Hadronic Calorimeter. The EM Calorimeter measures the energy deposited by electrons, positrons and photons (“egamma” objects). The Hadronic Calorimeter measures the energy deposited by hadrons such as protons, neutrons, pions, etc. The Muon systems measure the trajectories and momenta of muon tracks. An illustration of how different particles interact in the different sub-detector systems is shown in Figure 3.2. The analysis undertaken in this thesis makes use of all sub-detector systems since the Inner detector, the EM Calorimeter and the Muon systems provide direct measurements of all the particles required for this analysis and the Hadronic Calorimeter helps discriminate between egamma objects and hadrons. More details of these detectors are given in the following sections.

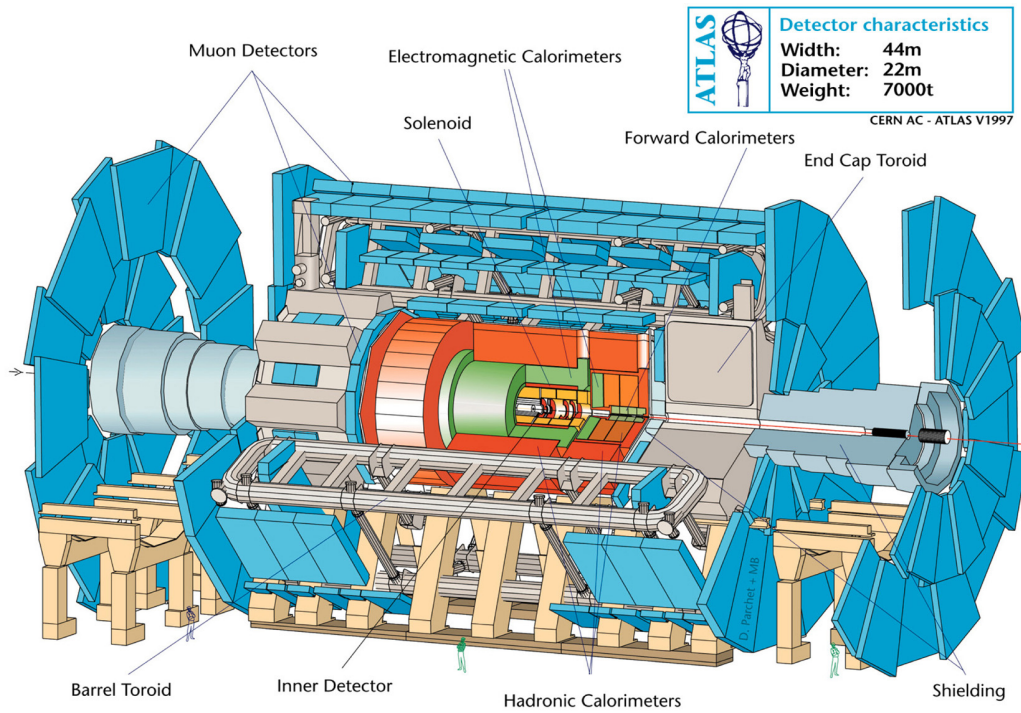


Figure 3.1: The ATLAS detector.

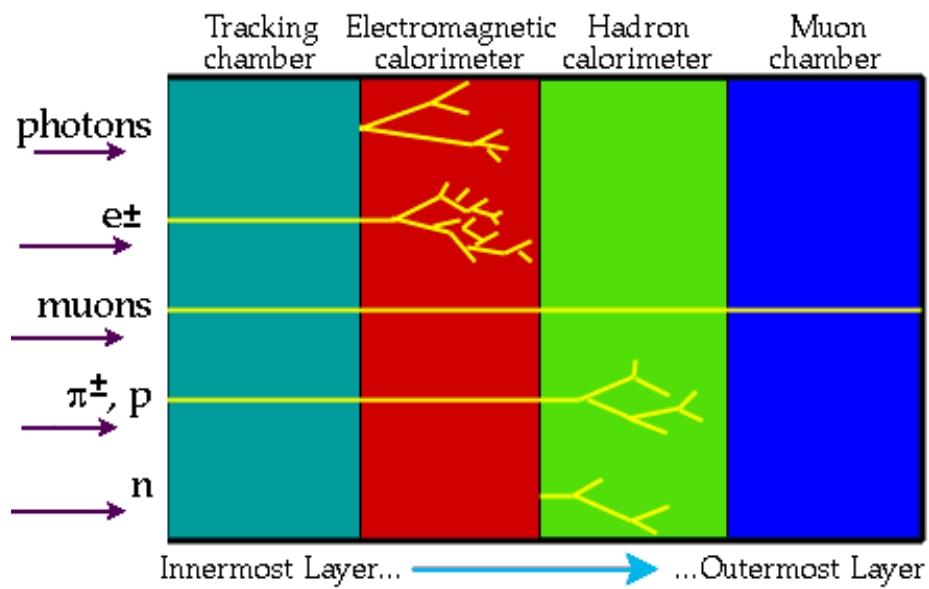


Figure 3.2: An illustration of particle detection in the subsystems of the ATLAS detector.

3.1 The ATLAS Coordinate System

In order to describe the ATLAS detector, it is necessary to define the coordinate system. The ATLAS coordinate system is defined such that the positive x -axis points to the center of the LHC ring, the y -axis points upwards and the z -axis is defined by the usual right hand rule and lies along the beam direction. The azimuthal angle ϕ is measured from the positive x -axis and the polar angle θ is measured from the positive z -axis. In experimental particle physics, pseudorapidity $\eta = -\ln[\tan(\frac{\theta}{2})]$ is a commonly used spatial variable instead of θ . Roughly speaking, this is because the hard inelastic interactions which are of interest favor the transverse direction, so the particle density is more or less flat with η . The transverse momentum p_T and the transverse energy E_T are defined in the $x - y$ plane. The p_T is the projection of the momentum onto the $x - y$ plane. The E_T is defined as $E_T = \sqrt{m^2 + p_T^2}$. For massless particles such as photon, E_T equals p_T .

3.2 The Inner detector

The Inner detector, shown in Figure 3.3, provides precise charged particle track reconstruction in the range $|\eta| < 2.5$. It consists of three sub-detectors including three layers of silicon pixel detector, four layers of silicon microstrip tracker (SCT) and a transition radiation tracker (TRT), all contained in the Central Solenoid, which provides a nominal magnetic field of 2 T.

3.2.1 The Pixel Detector

The pixel detector ranges from 5 cm to 12 cm in radius, consisting of three barrels at average radii of ~ 5 cm, 9 cm, and 12 cm, and three disks on each side, between radii of 9 to 15 cm. A typical track will cross three pixel layers. The innermost layer of the pixel detector is called the B-layer, located just outside the beam-pipe, which is vital for good vertexing. Each pixel is $50 \mu\text{m}$ wide in R_ϕ and $300 \mu\text{m}$ long. The resolution is about $14 \times 115 \mu\text{m}^2$. In total, the pixel detector has 1744 modules. Each module consists of 46080 readout channels. So there are about 80 million channels.

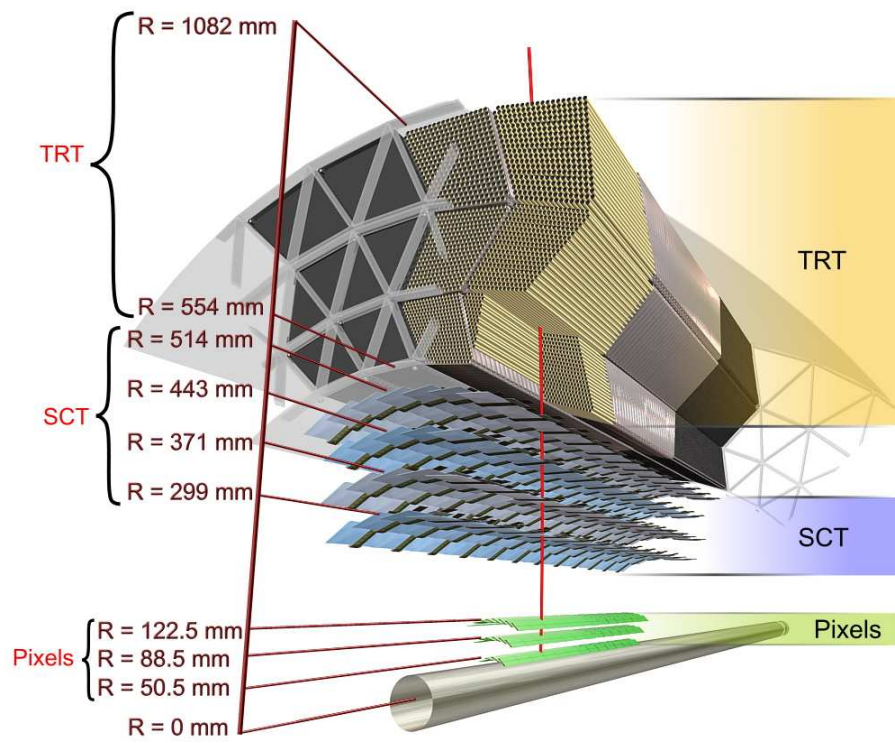


Figure 3.3: The Inner detector.

3.2.2 The Silicon Microstrip Tracker

Outside the pixel detector is the SCT occupying radii from 30 cm to 50 cm. It has four double-sided barrel layers and nine end-cap disks on each side. It has about 6.3 million readout channels and a total area of 61 square meters. The strips have an 80 μm pitch and are 12 cm long. It allows the positions of charged particles to be recorded to an accuracy of 17 μm per layer (in the direction transverse to the strips). It is designed to provide eight precision measurements per track in the intermediate radial range, contributing to the measurement of momentum, impact parameter and vertex position.

3.2.3 The Transition Radiation Tracker

The outermost detector is the TRT detector ranging in radius from 55 cm to 108 cm. It has about 350 000 readout channels. The TRT detector makes use of cylindrical drift tubes (also called straws) and provides a continuous track measurement. The straws are of 4 mm diameter and 144 cm long and are parallel to the beam. The resolution is about 200 μm . Typically the TRT provides 36 measurements per track. However, the detector has to operate at high occupancy and will be affected by the high pileup environment since its readout window is 75 ns, which is greater than the 50 ns bunch spacing. Besides track measurements, the TRT also has the capability of distinguishing electrons from charged hadrons by employing xenon gas to detect transition-radiation photons. The gas is a mixture based on 70% xenon for good X-ray absorption and 27% CO_2 and 3% O_2 to increase the electron drift velocity and for photon-quenching.

Figure 3.4 shows the radiation length (X_0) of the inner detector as a function of η measured from the center of the detector. The radiation length is the mean distance over which a high-energy electron loses all but $\frac{1}{e}$ of its energy by bremsstrahlung and $\frac{7}{9}$ of the mean free path for pair production by a high-energy photon. With the concept of radiation length, we can compare the thickness of different materials effectively.

In summary, the main parameters of the inner detector are shown in Figure 3.5.

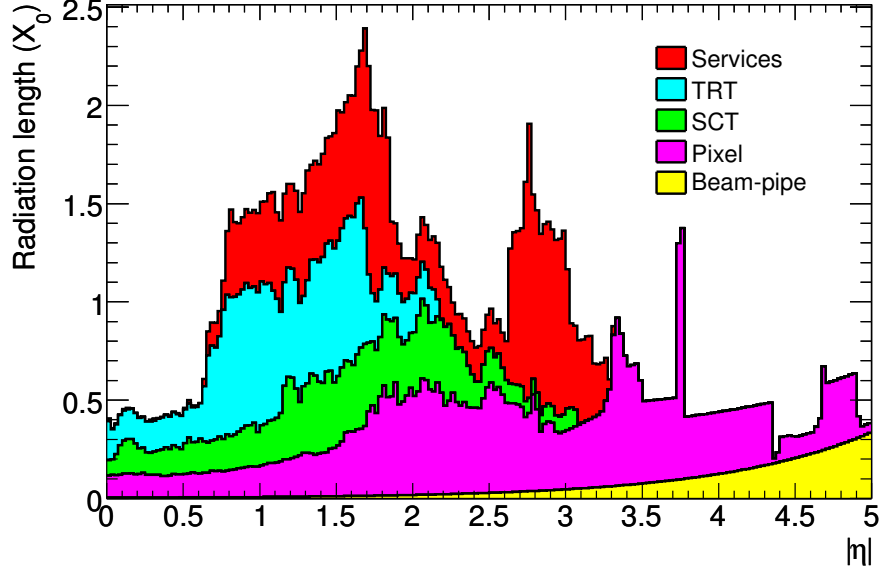


Figure 3.4: Inner detector thickness in units of radiation length as a function of η as measured from the center of the detector.

Item		Radial extension (mm)	Length (mm)
Pixel 3 cylindrical layers 2 $\times$ 3 disks	Overall envelope	$45.5 < R < 242$	$0 < z < 3092$
	Sensitive barrel	$50.5 < R < 122.5$	$0 < z < 400.5$
	Sensitive end-cap	$88.8 < R < 149.6$	$495 < z < 650$
SCT 4 cylindrical layers 2 $\times$ 9 disks	Overall envelope	$255 < R < 549$ (barrel)	$0 < z < 805$
		$251 < R < 610$ (end-cap)	$810 < z < 2797$
	Sensitive barrel	$299 < R < 514$	$0 < z < 749$
	Sensitive end-cap	$275 < R < 560$	$839 < z < 2735$
TRT 73 straw planes 160 straw planes	Overall envelope	$554 < R < 1082$ (barrel)	$0 < z < 780$
		$617 < R < 1106$ (end-cap)	$827 < z < 2744$
	Sensitive barrel	$563 < R < 1066$	$0 < z < 712$
	Sensitive end-cap	$644 < R < 1004$	$848 < z < 2710$

Figure 3.5: Main parameters of the inner detector.

3.3 Calorimeter

The ATLAS Calorimeter is designed to measure a particle's energy. It consists of the Electromagnetic (EM) Calorimeter, the Hadronic Calorimeter and the Forward Calorimeter, shown in Figure 3.6. The EM barrel is 6.4 m long, 53 cm thick with 110,000 channels. The EM endcaps each have a thickness of 0.632 m and a radius of 2.077 m. The Hadronic barrel is 5.6 m long and ranges in radii from 2.28 m to 3.87 m. The Hadronic endcaps consist of two wheels of thickness 0.8 m and 1.0 m with radii of 2.09 m. Each Forward calorimeter has three modules of radius 0.455 m and thickness 0.45 m.

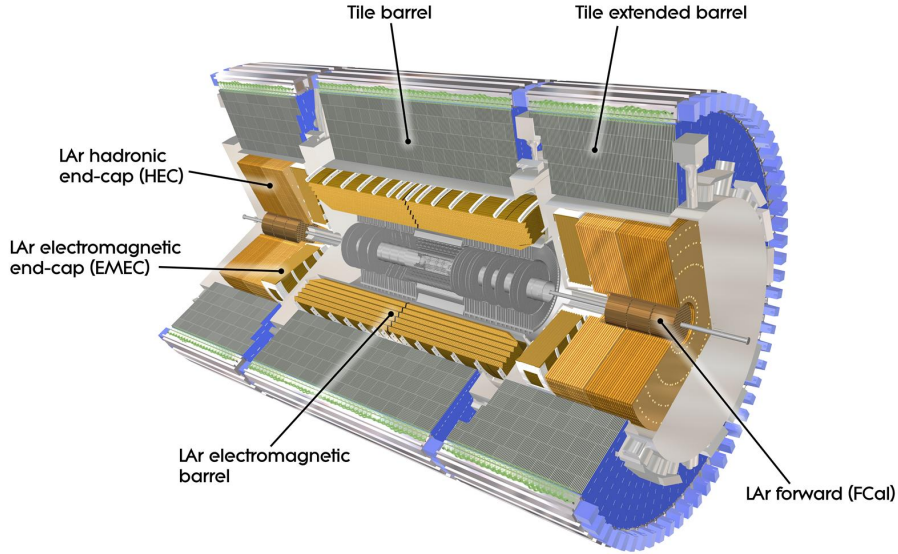


Figure 3.6: ATLAS Calorimeter

3.3.1 EM Calorimeter(ECAL)

The EM Calorimeter measures the energy of electrons, positrons and photons (called egamma objects). It is a sampling calorimeter in which the material that produces the electromagnetic shower (the absorber) is distinct from the material that measures the energy (the active material). The advantage of a sampling calorimeter is that a very dense material can be used to initiate electromagnetic showers within a more compact radius, hence keeping

detector costs down because the detector as a whole is more compact. The disadvantage is that the particle energy is not fully measured but estimated by using the sampled energy. (In contrast, if the two materials are the same, it is called a Homogeneous Calorimeter which is used, for example, by CMS.)

When an high energy electron enters the Calorimeter, a bremsstrahlung photon is then produced, and the photon decays into an electron-positron pair. The process repeats itself until the energy of newly produced photons is less than the threshold below which the Compton cross section and Photoelectric effect cross section dominate over pair production. At that point, the shower shape reaches its maximum. Similar process happens to photon as well. This cascade process leads to an electromagnetic shower as illustrated in Figure 3.7. Since the conversions occur with a mean free path of $\frac{9}{7}X_0$, then on average the photon showers happen deeper than those of electrons.

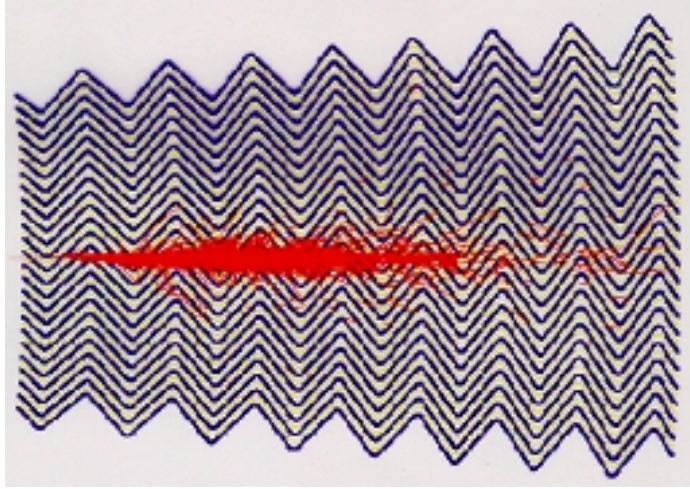


Figure 3.7: An electromagnetic shower in the calorimeter.

The EM calorimeter is divided into a barrel part ($|\eta| < 1.475$) and two end-caps ($1.375 < |\eta| < 3.2$). The barrel calorimeter consists of two identical half-barrels, separated by a small gap (6 mm) at $\eta=0$. The absorber used in ATLAS is Pb and the active material is liquid argon (LAr). When particles pass through liquid argon, electrons are liberated, collected and recorded. The choice of the materials is a tradeoff between performance and cost. The performance refers to signal linearity, energy and position resolution, radiation tolerance,

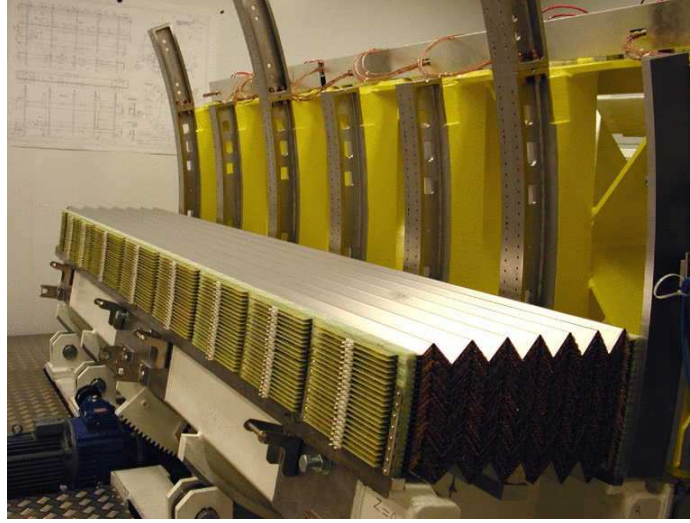


Figure 3.8: Photograph of a partly stacked barrel electromagnetic LAr module. A total of six out of seven outer support rings into which the absorbers can be seen. The backbone behind the outer support rings and the assembly bench below the stacked modules are also visible.

speed of response, etc. The accordion geometry shown in Figure 3.8 provides complete ϕ symmetry without azimuthal cracks. The Pb plate thickness has been optimized as a function of η in terms of EM calorimeter energy resolution performance. The LAr gap has a constant thickness of 2.1 mm in the barrel. In the end-cap, since the amplitude of the accordion waves increases with radius and the absorbers have constant thickness, the LAr gap increases with radius. The total thickness of the EM calorimeter is $> 24 X_0$ in the barrel and $> 26 X_0$ in the end-caps.

The EM calorimeter is divided into three sampling layers and one pre-sampler. The pre-sampler covers only $|\eta| < 1.8$. The first sampling and the pre-sampler layer with their fine granularity enhance the particle identification (γ/π^0 , e/π separation, etc.) and provide a precise η measurement. The parameters of the ATLAS Calorimeter are listed in Figure 3.9. EM objects will deposit most of their energy in the first two sampling layers since the total radiation length thickness is about $24 X_0$. The third sampling layer, with radiation length thicknesses from $2 X_0$ to $12 X_0$ measures the small remaining electromagnetic shower

EM CALORIMETER	Barrel	End-cap	
Coverage	$ \eta < 1.475$	$1.375 < \eta < 3.2$	
Longitudinal segmentation	3 samplings	3 samplings 2 samplings	$1.5 < \eta < 2.5$ $1.375 < \eta < 1.5$ $2.5 < \eta < 3.2$
Granularity ($\Delta\eta \times \Delta\phi$)			
Sampling 1	0.003×0.1	0.025×0.1 0.003×0.1 0.004×0.1 0.006×0.1 0.1×0.1	$1.375 < \eta < 1.5$ $1.5 < \eta < 1.8$ $1.8 < \eta < 2.0$ $2.0 < \eta < 2.5$ $2.5 < \eta < 3.2$
Sampling 2	0.025×0.025	0.025×0.025	$1.375 < \eta < 2.5$
Sampling 3	0.05×0.025	0.1×0.1 0.05×0.025	$2.5 < \eta < 3.2$ $1.5 < \eta < 2.5$
PRESAMPLER	Barrel	End-cap	
Coverage	$ \eta < 1.52$	$1.5 < \eta < 1.8$	
Longitudinal segmentation	1 sampling	1 sampling	
Granularity ($\Delta\eta \times \Delta\phi$)	0.025×0.1	0.025×0.1	
HADRONIC TILE	Barrel	Extended barrel	
Coverage	$ \eta < 1.0$	$0.8 < \eta < 1.7$	
Longitudinal segmentation	3 samplings	3 samplings	
Granularity ($\Delta\eta \times \Delta\phi$)			
Samplings 1 and 2	0.1×0.1	0.1×0.1	
Sampling 3	0.2×0.1	0.2×0.1	
HADRONIC LAr	End-cap		
Coverage	$1.5 < \eta < 3.2$		
Longitudinal segmentation	4 samplings		
Granularity ($\Delta\eta \times \Delta\phi$)	0.1×0.1 0.2×0.2 $1.5 < \eta < 2.5$ $2.5 < \eta < 3.2$		
FORWARD CALORIMETER	Forward		
Coverage	$3.1 < \eta < 4.9$		
Longitudinal segmentation	3 samplings		
Granularity ($\Delta\eta \times \Delta\phi$)	$\sim 0.2 \times 0.2$		

Figure 3.9: Pseudorapidity coverage, granularity and longitudinal segmentation of the calorimeters.

energy. The technical parameters are shown in Figure 3.9.

3.3.2 Hadronic Calorimeter (HCAL)

The Hadronic Calorimeter measures the energy of hadronic particles such as pions, protons, etc. The HCAL uses the same sampling technique, but is composed of different absorber and sampling materials in order to optimize the performance, cost and radiation hardness. For the region $|\eta| < 1$ (barrel) and $0.8 < |\eta| < 1.7$ (extended barrel), plastic scintillating plates are used as the active material and steel is used as absorber. For the region $1.5 < |\eta| < 3.2$ (end-cap) liquid argon is used as the active material and copper is used as the absorber. The technical parameters are shown in Figure 3.9.

3.3.3 Forward Calorimeter (FCAL)

The Forward Calorimeter is used to measure the energy of both EM objects and hadronic particles. It covers the forward region from $3.2 < |\eta| < 4.9$. It makes use of liquid argon as the active material and copper absorber for the electromagnetic part of the FCAL and tungsten absorber for the hadronic part of the FCAL.

3.4 Muon System

The outermost sub-detector is the Muon spectrometer (MS) which is used to measure muon tracks. It consists of four types of chambers. Monitored drift tube (MDT) chambers and Cathode Strip Chambers (CSCs) are used for precision position measurements. Resistive Plate Chambers (RPCs) and Thin Gap Chambers (TGCs) are used for triggering. The chambers are arranged such that particles from the interaction point traverse three stations of chambers. The positions of these stations are optimized for good hermeticity and optimum momentum resolution.

3.4.1 Monitored Drift Tubes

The Monitored drift tube (MDT) chambers are the principal detectors for track and momentum measurements. They cover an area of about 5500 m^2 in total. The MDT chambers

are made with aluminum tubes of 30 mm diameter and 400 μm wall thickness. In total, there are about 354240 tubes. Inside the tube, there is a 50 μm diameter central W-Re (97% : 3%) wire, operating at about 3080 V. The gas inside the tube is a non-flammable $\text{Ar}-\text{CH}_4-\text{N}_2$ mixture at 3 bar absolute pressure. The MDT's have a single-wire resolution of about 80 μm . Figure 3.10 shows the MDT tubes.

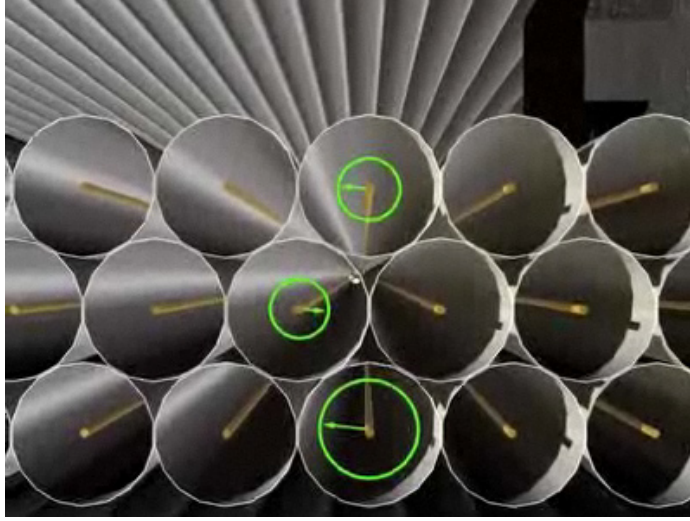


Figure 3.10: The Monitored drift tube (MDT). Under the electric field generated by the 3080 V potential, ionization electrons drift towards the anode and initiate an electron avalanche near the wire. The green circles show the starting point of this drifting in each tube. The radius of the circles is the closest distance from the wire to the muon track. The time of arrival of electrons from the starting point of closest approach determines the drift time corresponding to the drift radius.

3.4.2 Cathode Strip Chambers

The Cathode Strip Chambers (CSC) are installed nearest the interaction region in the high η region ($2 < |\eta| < 2.7$) of the first end-cap wheel where the background is highest. They are multi wire proportional chambers with cathode strip readout. When a muon passes through the chamber, it ionizes the gas on its track. Then the ionized electrons are accelerated by the strong electric field between the anode and cathode strip, drifting towards the anode. These

strong fields near the wire amplify the signal via a charge avalanche. The charge avalanche induces charge on the cathode strips shown in Figure 3.11. Based on the induced charge on the adjacent strips and the position of those strips, the centroid of the charge cloud is the position of the muon's track.

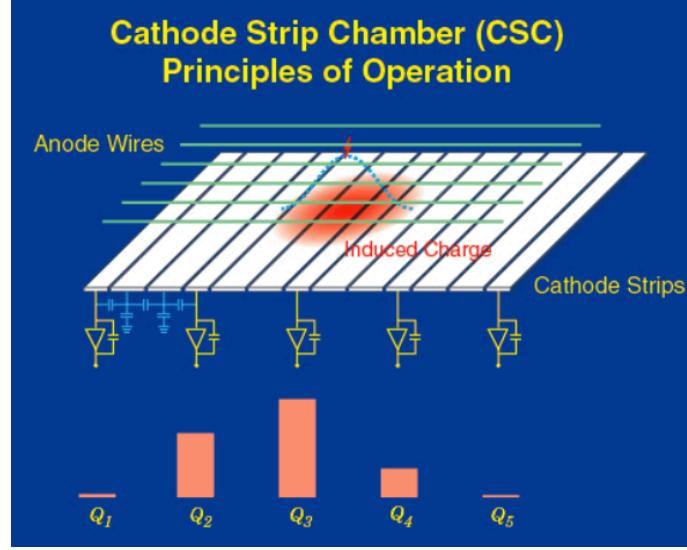


Figure 3.11: The Cathode Strip Chambers (CSC).

3.4.3 Resistive Plate Chambers

The Resistive Plate Chambers (RPC) are gaseous detectors used to trigger on muons and installed in the barrel region ($|\eta| < 1.05$). They provide a typical space-time resolution of $1 \text{ cm} \times 1 \text{ ns}$. The gas mixture is tetrafluoroethane $C_2H_2F_4$ which is non-flammable and environmentally safe. In Figure 3.12 the mechanical structure of RPC chamber is shown. The gas occupies the 2 mm gap region between the 2 mm thick resistive plates. Outside the resistive plate is a graphite layer connected to high voltage on one side and to the ground on the other side. The pick-up strips are separated by $200 \mu\text{m}$ thick insulating films which are glued on both graphite layers. The two sets of pick-up strips are orthogonally arranged. When a muon passes through the chamber, the charge avalanche around the wire induces a charge on the pick-up strips which are then read out.

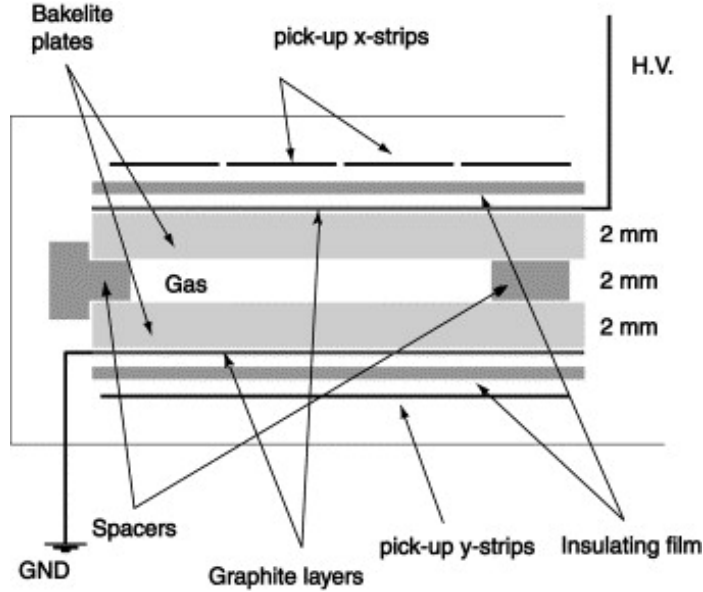


Figure 3.12: The mechanical structure of an RPC chamber.

3.4.4 Thin gap chambers

The Thin gap chambers (TGC) are similar in design to the CSC's, with the difference being that the anode wire pitch is larger than the cathode-anode distance. The TGC's cover the end-cap region ($1.05 < |\eta| < 2.4$). The TGC is shown in Figure 3.13. The pick-up strip is orthogonal to the anode wire, which can provide a second coordinate measurement.

3.4.5 Overall Muon System Layout

The Muon spectrometer is shown in Figure 3.14. In the barrel, the muon stations are positioned in three concentric layers around the beam axis, at radii of approximately 5, 7.5 and 10 meters. The innermost layer consists of MDT chambers only, situated between the Hadronic calorimeter and Toroidal magnets. The middle layer consists of one MDT chamber and two RPC chambers, situated in the Toroidal magnets. The Toroidal magnet system shown in Figure 3.15. The outermost layer consists of one MDT chamber and one RPC chamber, situated outside the Toroidal magnets. So only the middle layer is inside the Toroidal magnets. In the end-cap, CSC's are used in the innermost layer instead of MDT's because of the high particle fluxes. The second and third layers are outside the end-cap

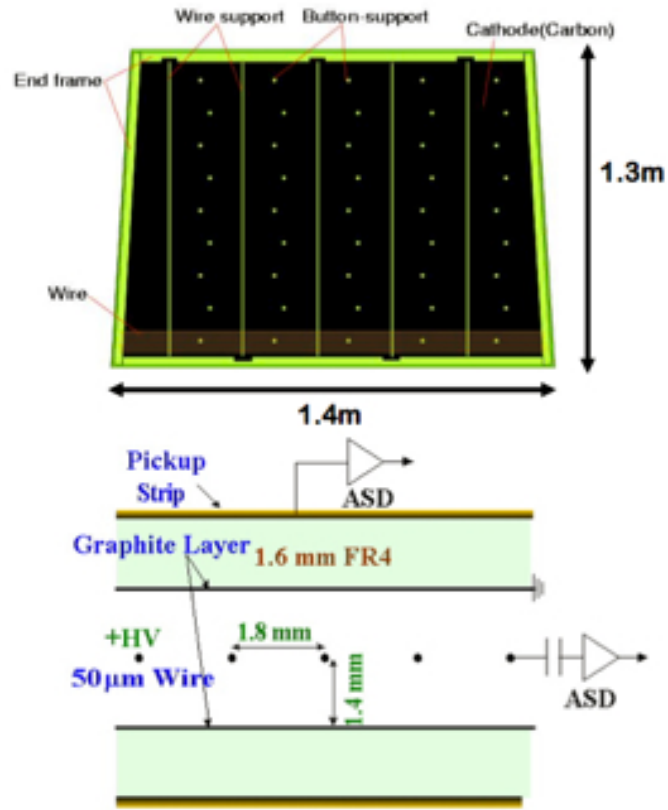


Figure 3.13: The structure of thin gap chambers (TGC).

toroid, so the muons are not bent in that region. In the end-cap, TGC's are used for the trigger.

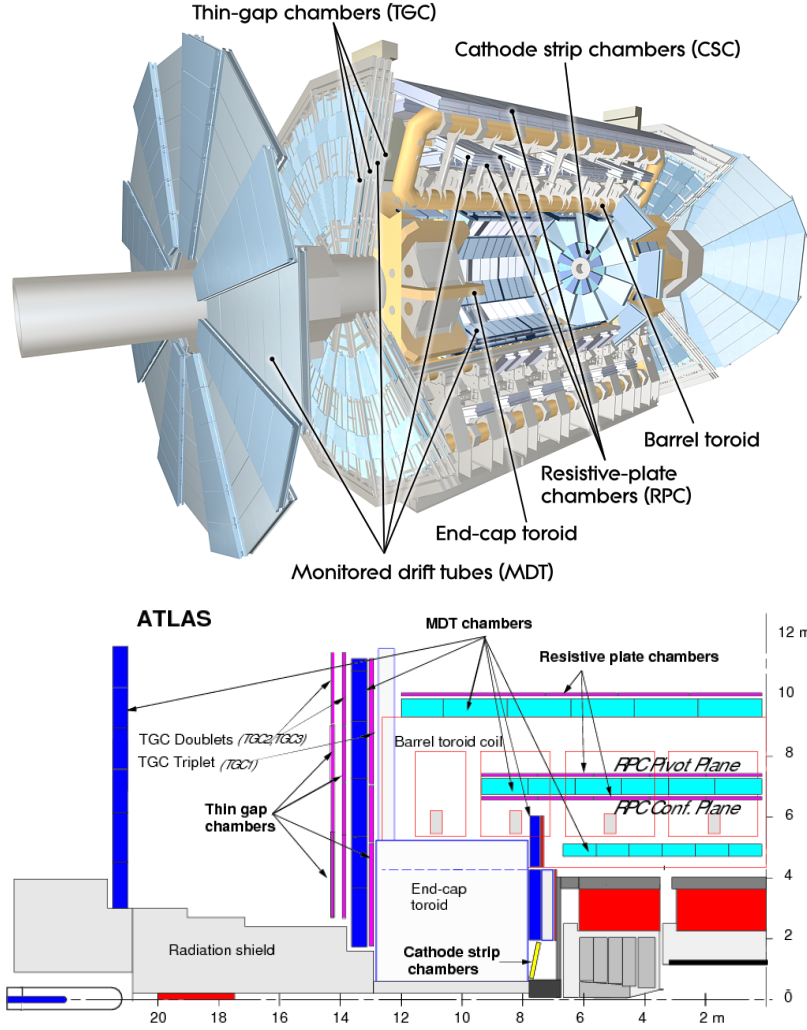


Figure 3.14: The overall ATLAS muon system (upper plot) and a detailed quarter view cut-away cross section of the muon system (lower plot).

3.5 Trigger and Data acquisition system (TDAQ)

The ATLAS trigger system is designed to record events at approximately 200 Hz from the LHCs 20 MHz bunch crossing rate, with about up to 45 interactions per bunch crossing

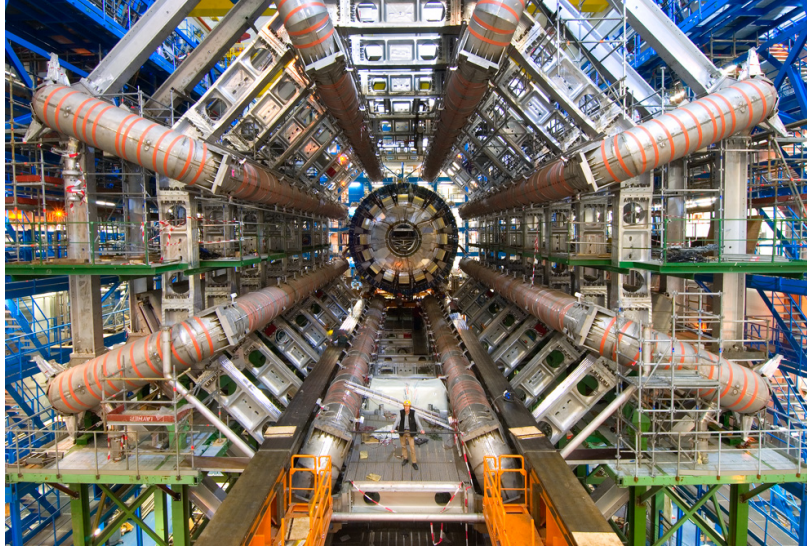


Figure 3.15: A photograph of the end view of the ATLAS detector under construction in 2005 with the eight air-core superconducting toroid magnets visible (they are the elements in the picture with the orange “stripes”).

depending on the 2011/2012 beam configuration. There are three trigger levels including L1, L2 and Event Filter (EF) in order to reduce the input rate from 20 MHz to 200 Hz step by step. The L1 trigger is a hardware-based system using information from the calorimeter and muon sub-detectors, after which the event rate will be reduced to 75 KHz. The L2 and EF level of triggers are software-based using information from all sub-detectors. After L2, the event rate is reduced to 3 KHz and after the EF it is further reduced to 200 Hz. The L2 and EF are collectively called the High Level Trigger (HLT). The TDAQ system is shown schematically in Figure 3.16.

The events selected by the trigger system are then written out to data streams. There are four physics streams including Egamma, Muons, JetTauEtmis, MinBias. Events in each of the four physics streams must satisfy a set of predefined triggers (electron or photon triggers, muon triggers, jet/tau/missing energy triggers, minimum bias triggers). The minimum bias events account for the vast majority of interactions which result from beam collisions. They do not contain hard-scattering processes. At high luminosities, multiple collisions within one beam-crossing are inevitable, causing signal events to have several minimum bias events

superimposed on them. In addition to physics streams, about 10% of events are written to an express stream which will be used as the input for Data Quality Monitoring. This will be introduced in next chapter. Calibration streams that use data from one or more sub-detectors are also recorded for further processing.

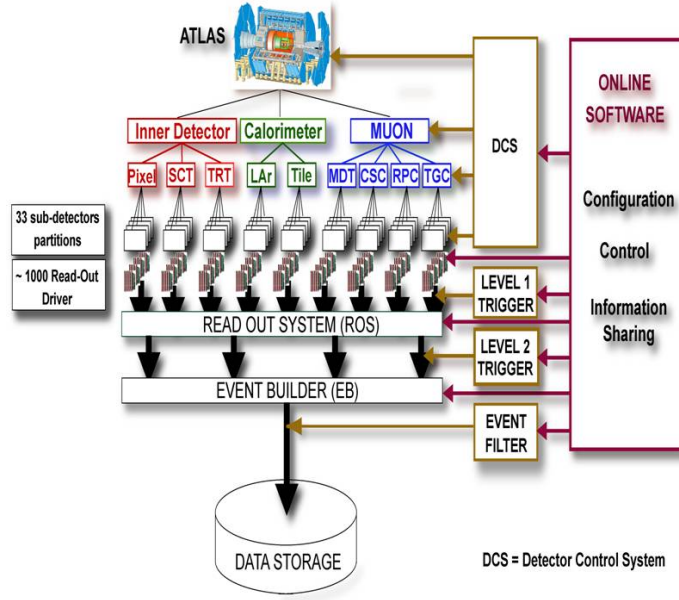


Figure 3.16: Trigger and Data acquisition system (TDAQ)

3.6 Luminosity detectors

There are two sets of luminosity detectors. The first set of detectors measure luminosity on a bunch-by-bunch basis. The second set of detectors infer the total luminosity (summed over all bunches) by monitoring detector currents which are sensitive to average particle rates over longer time scales. In 2010, when the luminosity was low, the Minimum Bias Trigger Scintillators (MBTS) were used. They are located at $z = \pm 365$ cm from the nominal interaction point, covering an eta range $2.09 < |\eta| < 3.84$. The MBTS were no longer used after 2010 due to the high luminosities. Another detector used to measure the luminosity is the Beam Conditions Monitor (BCM) located at $z = \pm 184$ cm. It has four 1 cm^2 sensors which form one horizontal pair and one vertical pair. The two pairs provide

two independent measurements. Another luminosity detector is called LUCID which is a Cherenkov detector, located at $z = \pm 17$ m. The Cherenkov photons are amplified by photomultiplier tubes (PMT) and if a preset threshold is exceeded, a ‘hit’ will be recorded for that tube in that bunch crossing. Both BCM and LUCID belong to the first set of detectors. The Tile Calorimeter (TileCal) and Forward Calorimeter (FCal) can also be used to measure luminosity since the particles passing through them will produce photoelectric current which is proportional to the luminosity. These two belong to the second set of detectors.

3.7 Summary of sub-detector performance

As a summary of the sub-detector data-taking performance, Figure 3.17 shows the luminosity weighted relative fraction of good quality data delivery by the various ATLAS subsystems during LHC fills with stable beams.

ATLAS 2011 p–p run												
Inner Tracking			Calorimeters				Muon Detectors				Magnets	
Pixel	SCT	TRT	LAr EM	LAr HAD	LAr FWD	Tile	MDT	RPC	CSC	TGC	Solenoid	Toroid
99.8	99.6	99.2	97.5	99.2	99.5	99.2	99.4	98.8	99.4	99.1	99.8	99.3
Luminosity weighted relative detector uptime and good quality data delivery during 2011 stable beams in pp collisions at $\sqrt{s}$ =7 TeV between March 13 th and October 30 th (in %), after the summer 2011 reprocessing campaign												
ATLAS p-p run: April-December 2012												
Inner Tracker			Calorimeters				Muon Spectrometer				Magnets	
Pixel	SCT	TRT	LAr	Tile	MDT	RPC	CSC	TGC	Solenoid	Toroid		
99.9	99.1	99.8	99.1	99.6	99.6	99.8	100.	99.6	99.8	99.5		
All good for physics: 95.5%												
Luminosity weighted relative detector uptime and good quality data delivery during 2012 stable beams in pp collisions at $\sqrt{s}$ =8 TeV between April 4 th and December 6 th (in %) – corresponding to 21.3 fb ⁻¹ of recorded data.												

Figure 3.17: Sub-detector performance in 2011 (upper) and 2012 (lower)

Part III

Data analysis

Chapter 4

Object Reconstruction and Identification

The physical objects that are of interest and can be reconstructed by the ATLAS detector are electrons, photons, muons, jets and Transverse Missing Energy (MET). One of main sources of MET are neutrinos, which are undetectable in the detector. Figure 4.1 illustrates how the different types of particles interact and deposit energy in the detector. One can see that the ATLAS detector sub-systems are designed to identify the various types of particles. The detector nearest the interaction point is the inner tracker which measures the trajectories of charged particles produced in the proton-proton collisions. This track information can be combined with the electromagnetic shower information measured in the calorimeter to distinguish photons (no track pointing to the shower) from electrons (with at least one track pointing to the shower). Hadronic particles, such as pions, protons and neutrons, can “shower” after interacting in the rear section of the calorimeter called the hadronic calorimeter. Muons are minimum ionizing particles that can be measured in the tracking detectors near the collision point (the inner tracker) and in the outermost muon tracking detectors. If the information from both sub-detectors is available, then the information is statistically combined to form a track.

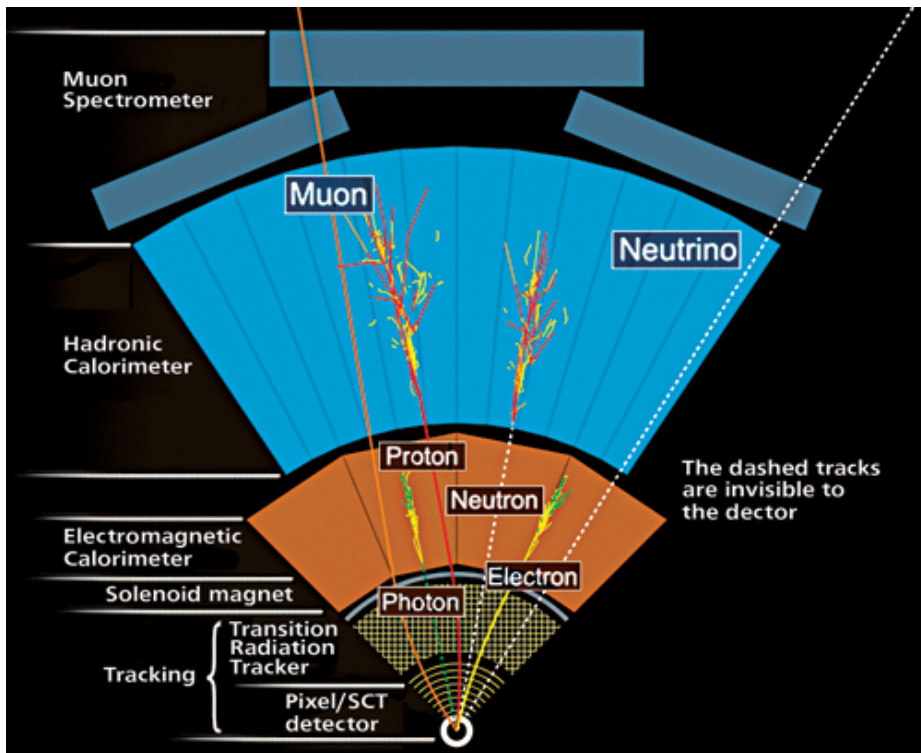


Figure 4.1: An illustration of the energy loss processes for the different types of particles produced in proton-proton collisions in the detector. Different detector subsystems are optimized for the specific particle types. The interaction point is at the bottom of the figure.

4.1 Track Reconstruction

The inner tracker provides the information about individual hits. These hits can be combined by track reconstruction algorithms to measure the trajectories of charged particles, which in turn provides the momentum of the particles by measuring their curvature in the detector's magnetic field. The measured tracks can then be combined to identify the interaction vertices for the event. Two principal algorithms are used. One is an “inside-out algorithm” which starts from 3-point seeds in the silicon detectors and adds hits moving away from the interaction point using a Kalman filter [44]. The inside-out algorithm is the baseline algorithm to reconstruct primary charged particles. The other algorithm is an “outside-in” algorithm (also referred to as a back-tracking algorithm) which is used to resolve the inefficiency of the inside-out algorithm since secondary particles may not leave enough hits in the silicon detector but could leave hits in the TRT. Photons which begin to shower in the inner tracker (“converted photons”) are one instance of such secondary particles. The outside-in algorithm starts from segments reconstructed in the TRT and then proceeds by adding silicon hits. The efficiency of the track reconstruction [45] is shown in Figures 4.2 and 4.3 for muons, pions and electrons with $p_T = 1, 5$ and 100 GeV.

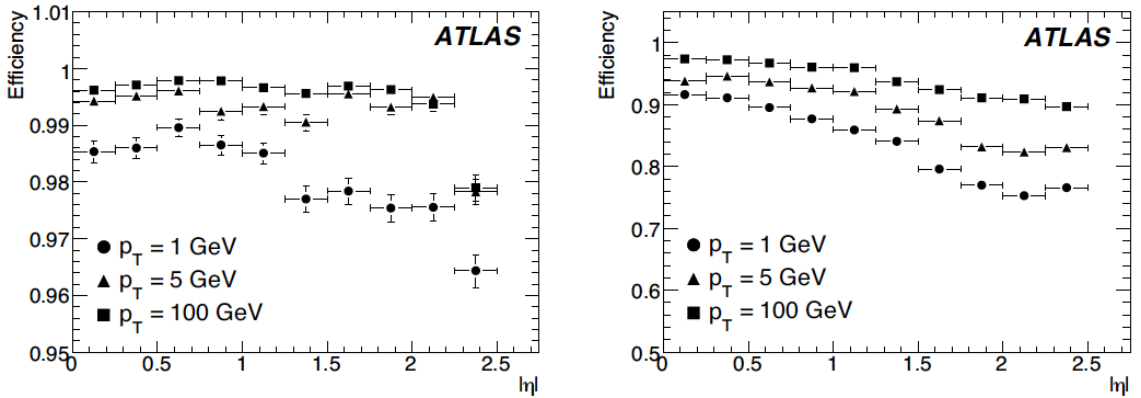


Figure 4.2: Track reconstruction efficiencies as a function of $|\eta|$ for muons (left) and pions (right) with $p_T = 1, 5$ and 100 GeV.

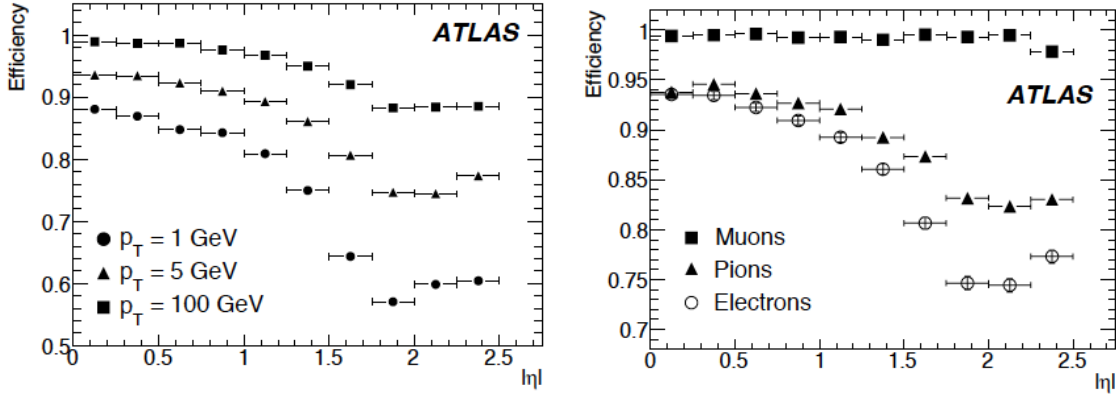


Figure 4.3: Track reconstruction efficiencies as a function of $|\eta|$ for electrons (left) with $p_T = 1, 5$ and 100 GeV and comparisons between electrons, muons and pions with $p_T = 5$ GeV (right).

4.2 Vertex Reconstruction

The reconstructed tracks are used to determine the vertices in the event. An iterative χ^2 fit is made, where the χ^2 function is defined by:

$$\chi^2 = \sum_{i=1}^N (q_i - T(\vec{V}, \vec{p}_i))^T W_i (q_i - T(\vec{V}, \vec{p}_i)) \quad (4.1)$$

where $\vec{V}$ is the assumed vertex. q_i is a vector of the five trajectory parameters $(d_0, z_0, \phi_0, \cot\theta, \frac{q}{p})^1$ of the i^{th} track. The transverse impact parameter d_0 is defined as the distance of closest approach to the beam-line. The longitudinal impact parameter z_0 is defined as the value of z of the point on the track that determines d_0 . The parameter ϕ_0 is the azimuth (the angle between the direction of motion of the track in the (x, y) plane and the x -axis) at the point of closest approach. The parameter $\cot\theta$, with θ the polar angle, gives the inverse slope of the track in the (r, z) plane. $T(\vec{V}, \vec{p}_i)$ is a vector of the predicted track parameters given the initial vertex and the initial momentum. W_i is the 5×5 weight matrix from the i^{th} track fit. An iterative fitting which tries to fit all the reconstructed tracks to a common vertex, removing at each step those tracks incompatible² with the hypothesis that they

¹A track has 5 degree of freedom (dof) including 3 spatial dof and 2 angular dof.

²Tracks incompatible with the vertex by more than approximately 7σ are used to seed a new vertex.

come from the primary vertex (secondaries from interactions, particles with a lifetime and mismeasured tracks) is performed.

At high luminosity, there are on average 20 interactions per bunch crossing. In order to identify those vertices, an algorithm [46] briefly described below is used. In a first step the algorithm assumes that all tracks are coming from the beam-line in the transverse plane (this is a good approximation due to the small transverse beam-spot size) and determines the z-coordinate of each track. Then plot the z-coordinate in a histogram with a $500 \mu m$ bin size. This histogram is then scanned to look for locations where the tracks cluster as they should do if they are coming from the same vertex. A minimum of four tracks is required to define a cluster. For each cluster, the corresponding tracks are the input to the iterative fitting algorithm described above. If the fit is successful, the vertex is retained. The procedure is repeated until no additional vertices can be found. The vertex with the maximum Σp_T^2 is the leading primary vertex. Figure 4.4 shows the reconstructed primary vertices in one event.

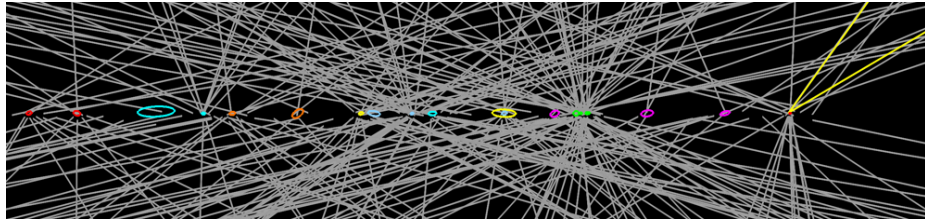


Figure 4.4: Multiple primary vertices were reconstructed in one event. The colored circles identify the primary vertices.

The centroid position of the reconstructed primary vertices collected about every 10 minutes over a three month period is shown in Figure 4.5. The size of the luminous region in z is between 40-50 mm in 2012.

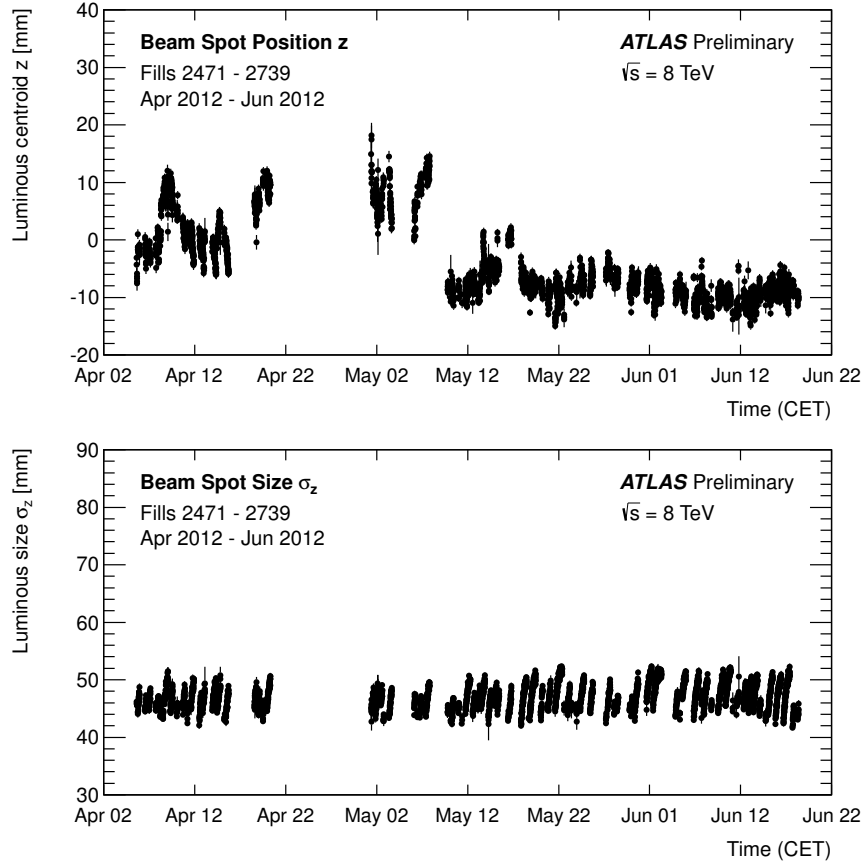


Figure 4.5: The position and size in z of the centroid of the luminous region in ATLAS over the course of pp running in 2012 at $\sqrt{s} = 8$ TeV. The data points are the result of a maximum likelihood fit to the spatial distribution of primary vertices collected over ten minutes. The large movement at the beginning of May was associated with movement of the inner detector. Errors are statistical only.

4.3 Electron and Photon (Egamma) Objects

4.3.1 Electrons

Electrons are reconstructed by using the Calorimeter system and the Tracking system. The Calorimeter is divided into a grid of $N_\eta \times N_\phi = 200 \times 256$ elements (each of size $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$). A sliding window algorithm [47] is used to build the EM clusters. The “sliding window” refers to the search strategy of changing the position of the window so that the sum of the transverse energy contained in the window is a local maximum and is above a threshold E_{thresh} . The window size is 3×5 in units of 0.025×0.025 in $\eta - \phi$ space, corresponding to the granularity of the calorimeter sampling layer 2 (refer to Table 3.9). If the sum of all the calorimeter cell energies within the window size is above 2.5 GeV, the cluster is reconstructed (seed cluster). After an energy comparison, duplicate clusters are removed from nearby seed clusters. The position of the cluster is computed as the energy-weighted η and ϕ barycenters of all cells within a fixed-size window. In MC simulations, the efficiency of the algorithm is 100% for electrons with $E_T > 15$ GeV from W and Z decays [49].

The reconstructed tracks are extrapolated to the middle layer of the calorimeter. If the distance between the track impact point and the cluster position satisfy $\Delta\eta < 0.05$ and $\Delta\phi < 0.1$, an electron candidate is reconstructed. If there is more than one track, the one closest to the seed cluster is chosen based on distance defined by $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

Then the electron cluster is rebuilt using $3(\Delta\eta) \times 7(\Delta\phi)$ longitudinal towers of cells in the barrel EM calorimeter and using 5×5 towers of cells in the end-cap. The energy deposited in each layer as well as shower shape variables (in Figure 4.7) are calculated for further use in the identification process.

Since the inner detector does not cover the forward region ($2.5 < |\eta| < 4.9$), only the calorimeter information is used to build electron clusters in that region.

In the 2012 analysis, the bremsstrahlung losses have been taken into account by using the Gaussian Sum Filter (GSF) [50] while in 2010 and 2011, the standard electron track reconstruction algorithm is the same as that for other charged particles without any special treatment of the bremsstrahlung energy loss due to the amount of material in the inner de-

tector. Actually the GSF electron algorithm was available as an alternative to the standard electron reconstruction algorithm in 2011 and became the standard in 2012.

The distribution of relative energy losses due to bremsstrahlung (the BetheHeitler model) is sharply peaked at zero with a long tail, so the standard tracking algorithm might miss part of an electron track which has a large energy loss. In order to reduce this inefficiency and keep using the Kalman filter technique³, the Bethe-Heitler model of electron energy loss by bremsstrahlung is approximated by a weighted sum of a few Gaussian distributions and then run a number of Kalman filters in parallel. The approximation is performed by minimizing some kind of distance between the cumulative distribution function (CDF) or the Kullback–Leibler distance (a loss function) [50]. Although the quality of the approximation gets better by increasing the number of Gaussian components in the mixture, there is not much improvement beyond six components. The refitting step is only applied to the tracks with a transverse momentum $p_T > 400$ MeV and $|\eta| < 2.5$ that are assigned to electrons. The resulting collection of bremsstrahlung corrected tracks are then subsequently used by the standard electron reconstruction algorithm.

An electron cluster is rejected in some cases which may result in a bad energy measurement. The electron cluster is rejected if [49]:

- it contains a region in the EM calorimeter strip layer or middle layer where the readout has failed.
- it contains a region where the high voltage has failed.
- it contains a region with isolated noisy cells which are in the core region (defined as 3×3 cells) of the electron cluster in the middle layer.

In order to evaluate the quality of the electron candidates, we define three quality levels: loose++, medium++ and tight++. The ‘++’ sign means re-optimized selection criteria for high pileup environment. The requirements for the three quality levels are shown in Figure 4.6. **Loose++** electrons have requirements on electron hadronic leakage and shower shape variables in the middle layer of the EM calorimeter. The hadronic leakage

³The standard Kalman filter model assumes Gaussian noise.

is defined as the ratio of the E_T in the 1st sampling of the hadronic calorimeter to the E_T of the EM cluster and the ratio of E_T in the hadronic calorimeter to E_T of the EM cluster. **Medium**++ electrons have additional requirements, beyond the loose++ ones, on the shower shape variables in the strip layer of the EM calorimeter and on the Track quality and matching cut. **Tight**++ electrons have further requirements, beyond the medium++ ones, that make use of B-layer hits, Track matching variables and TRT hits.

4.3.2 Photons

Photons are reconstructed by using the Tracking system together with the calorimeter system. If there is no reconstructed track matched to an electromagnetic cluster, this egamma object will be classified as an **unconverted photon** and if there is a reconstructed conversion vertex matched to the cluster, it will be classified as a **converted photon**. A conversion can occur when a photon traverses the material of the detector. A converted photon will decay into a electron and positron pair. The differential cross section for photon conversions of energies of 1 GeV and above in terms of the quantity $x = \frac{E_{electron}}{E_\gamma}$ is [51]:

$$\frac{d\sigma}{dx} = \frac{A}{X_0 N_A} \left(1 - \frac{4}{3}x(1-x)\right) \quad (4.2)$$

where A is the atomic mass of the target given in g/mol, and $N_A = 6.022 \times 10^{23}$ is Avogadro's number. X_0 is known as the radiation length of the material through which the photon passes. Due to the term $x(1-x)$, the electron and positron do not simply share the photon energy equally. Especially when the photon energy is low, one of either the electron or the positron may have too low an energy to be reconstructed by the detector. Therefore, some converted photons only have one track, which causes difficulties in resolving between a single-track converted photon and an electron. These ambiguous objects often correspond to early showers in the tracker. In order to maximize both the electron and photon reconstruction efficiency, the ambiguous objects are often classified as both an electron and a converted photon. At the selection stage, the B-layer (the first layer of the pixel detector) hit is an useful variable to distinguish between electrons and converted photons. Photons are identified by using **Loose/Tight** criteria, which are a cut-based requirement. A Loose photon has a requirement on the Hadronic which is defined as the ratio of E_T in

Type	Description	Variable name
Loose cuts		
Acceptance of the detector	* $ \eta < 2.47$	
Hadronic leakage	* Ratio of E_T in the first layer of the hadronic calorimeter to E_T of the EM cluster (used over the range $ \eta < 0.8$ and $ \eta > 1.37$)	R_{had1}
	* Ratio of E_T in the hadronic calorimeter to E_T of the EM cluster (used over the range $ \eta > 0.8$ and $ \eta < 1.37$)	R_{had}
Second layer of EM calorimeter	* Ratio in η of cell energies in 3×7 versus 7×7 cells. * Lateral width of the shower.	R_η $w_{\eta 2}$
Medium cuts (includes Loose)		
First layer of EM calorimeter.	* Total shower width. * Ratio of the energy difference associated with the largest and second largest energy deposit over the sum of these energies	w_{stot} E_{ratio}
Track quality	* Number of hits in the pixel detector (≥ 1). * Number of hits in the pixels and SCT (≥ 7). * Transverse impact parameter (< 5 mm).	d_0
Track matching	* $\Delta\eta$ between the cluster and the track (< 0.01).	$\Delta\eta_1$
Tight cuts (includes Medium)		
b-layer	* Number of hits in the b-layer (≥ 1).	
Track matching	* $\Delta\phi$ between the cluster and the track (< 0.02). * Ratio of the cluster energy to the track momentum * Tighter $\Delta\eta$ cut (< 0.005)	$\Delta\phi_2$ E/p $\Delta\eta_1$
Track quality	* Tighter transverse impact parameter cut (< 1 mm).	d_0
TRT	* Total number of hits in the TRT. * Ratio of the number of high-threshold hits to the total number of hits in the TRT.	
Conversions	* Electron candidates matching to reconstructed photon conversions are rejected	

Figure 4.6: Definition of variables used for loose, medium and tight electron identification cuts.

the first sampling layer of the hadronic calorimeter to the E_T of the EM cluster in the range $|\eta| < 0.8$ and $|\eta| > 1.37$; in the range $0.8 < |\eta| < 1.37$ it is the ratio of E_T in the full hadronic calorimeter to E_T of the EM cluster). In addition information on the EM middle layer is used from the R_η and $w_{\eta 2}$ variables. A Tight photon has a requirement, in addition to the Loose cuts, on R_ϕ from the middle EM layer and the EM Strip layer variables (w_{s3} , w_{stot} , F_{side} , ΔE , E_{ratio}). For a detailed description of these variables please refer to Figure 4.7

Category	Description	DV	Loose	Tight
Acceptance	$ \eta < 2.37$, $1.37 < \eta < 1.52$ excluded	–		✓
Hadronic leakage	Ratio of E_T in the first sampling of the hadronic calorimeter to E_T of the EM cluster (used over the ranges $ \eta < 0.8$ and $ \eta > 1.37$)	R_{had1}	✓	✓
	Ratio of E_T in all the hadronic calorimeter to E_T of the EM cluster (used over the range $0.8 < \eta < 1.37$)	R_{had}	✓	✓
EM Middle layer	Ratio between the sum $E_{3 \times 7}^{S2}$ of the energies of the cells contained in a 3×7 $\eta \times \phi$ rectangle (measured in cell units), and the sum $E_{7 \times 7}^{S2}$ of the cell energies in a 7×7 rectangle, both centered around the cluster seed	R_η	✓	✓
	Lateral width of the shower in the η direction	$w_{\eta 2}$	✓	✓
	Ratio between the sum $E_{3 \times 3}^{S2}$ of the energies of the cells contained in a 3×3 $\eta \times \phi$ rectangle (measured in cell units), and the sum $E_{3 \times 7}^{S2}$ of the cell energies in a 3×7 rectangle, both centered around the cluster seed	R_ϕ		✓
EM Strip layer	Lateral shower width for three strips around maximum strip	w_{s3}		✓
	Total lateral shower width	w_{stot}		✓
	Fraction of energy outside core of three central strips but within seven strips	F_{side}		✓
	Difference between the energy associated with the second maximum in the strip layer, and the energy reconstructed in the strip with the minimal value found between the first and second maxima	ΔE		✓
	Ratio of the energy difference associated with the largest and second largest energy deposits over the sum of these energies	E_{ratio}		✓

Figure 4.7: Discriminating variables used for loose and tight photon identification cuts.

4.4 Calorimeter Data Quality Monitoring

Noisy or sporadically noisy channels sometimes can produce fake energy deposits in the Calorimeter which may give rise to a sliding window cluster. This could lead to fake electrons or photons. In order to identify the noisy channels, data quality monitoring has been performed through 2-dimensional $\eta - \phi$ hit maps and energy deposit maps. The hit map records the entries of cells or clusters with energy above a specific threshold in the $\eta - \phi$ map. The noisy channels are then recorded in the database and these channels are masked at the reconstruction level for all kinds of objects. An example of a noisy cluster is shown as the red pixel in Figure 4.8.

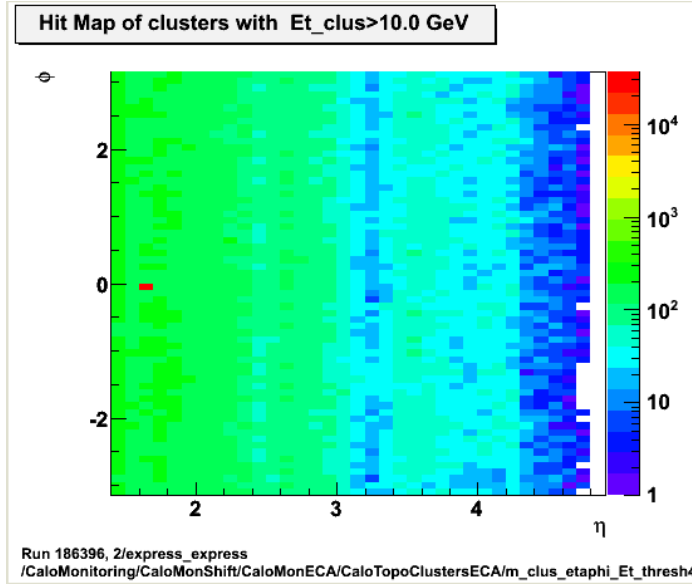


Figure 4.8: Calorimeter Data Quality monitoring. The plot is an hit map in $\eta - \phi$, displaying the entries for topological clusters with transverse energy about 10 GeV. The noisy cluster, shown in red, is around $\eta - \phi = (1.6, 0)$. It will also be checked at the CaloCell level and masked in the reprocessed data.

4.5 Muons

Muons are reconstructed by using the Inner Detector (ID) and the Muon spectrometer (MS). There are four classes of Muons:

Combined Muon (CB) Track reconstruction is performed in the ID and the MS independently and then a combined track is formed.

Stand-alone muon (SA) The muon track is only reconstructed in the MS. The impact parameter of the muon at the interaction point is determined by extrapolating the track back to the beam line and also taking into account the energy loss in the calorimeter.

Segment-tagged muon (ST) The muon track is measured in the ID and that track, when extrapolated to the Muon spectrometer, is required to be associated with straight track segments in the precision muon chambers. The measured track parameters of the associated ID track are used.

Calorimeter-tagged muon (CT) A track in the ID is identified as a muon if the associated energy deposit in the calorimeters is compatible with the hypothesis of a minimum ionizing particle. These muons are used in the analysis to cover the region around $\eta = 0$ which is not equipped with muon chambers.

4.6 Jets

A Jet is the signature of quarks or gluons. It is a cone of hadrons and other particles which are produced by the fragmentation of quarks and gluons due to the color confinement. Because only colorless objects can be observed independently, when quarks or gluons (that carry color charge) are produced, they combine with other colored particles produced from the vacuum to become colorless hadrons immediately. These hadrons are observed by the detector and reconstructed as a jet. Since jets are not used in this analysis, I will only briefly introduce the Jet reconstruction algorithm.

We define a distance parameter $d_{iB} = k_{ti}^{2p}$ for each cluster, and d_{ij} for each pair of clusters :

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \times \frac{\Delta R_{ij}^2}{R^2} \quad (4.3)$$

where k_{ti} and k_{tj} are the momentum of the i^{th} and the j^{th} cluster. p is a parameter with value 0, +1 and -1 corresponding to the Cambridge/Aachen algorithm, k_T and $anti - k_T$ algorithm respectively. ΔR_{ij} is the distance between i and j in $\eta - \phi$ plane. R is the distance parameter or cone size of the jet.

The procedure for reconstructing a jet is:

- find the minimum of d_{iB} and d_{ij} as d_{min}
- if d_{min} equals d_{ij} , merge clusters i and j
- if d_{min} equals d_{iB} , reconstruct cluster i as a jet and remove it from the cluster collection
- repeat until no clusters remain

As for the $anti - k_T$ algorithm, soft particles tend to merge with hard particles first. If a hard particle has no hard neighbors within a distance $2R$, then it will merge all the soft particles within a circle of radius R , resulting in a conical jet. When the ΔR between a hard particle and a soft particle is larger than R , the soft particle does not merge with the hard particle, therefore the key feature of the $anti - k_T$ algorithm is that the soft particles do not modify the shape of the jet as can be seen from Figure 4.9.

4.7 Missing transverse energy (MET)

Missing transverse energy (MET) is not used in this analysis, so I will only briefly discuss it here. MET is the energy imbalance in the transverse plane. Since the two proton beams collide with each other along the z direction, then the sum of the transverse momentum is expected to be zero. If there are no invisible particles, a non-zero MET may result from noisy or dead cells, inefficiency in reconstructing muons and fake muons. But this MET value should be small. If there are high p_T invisible particles produced in a collision, the MET is expected to be large. MET is important for finding new invisible particles. In ATLAS, “refined-final” Missing ET is recommended. It uses calorimeter cell energies

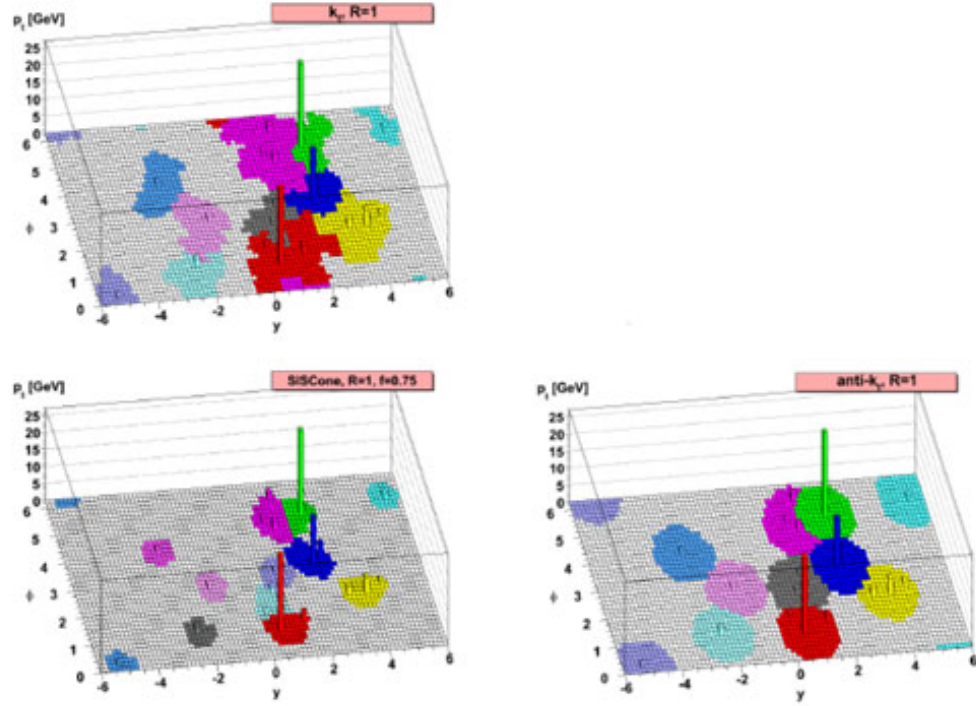


Figure 4.9: Shapes for jets algorithms reconstructing the same event. The towers in the (y, ϕ) plane are colored according to which jet they belong to. The vertical axis shows the p_T of the tower. The *anti* - k_T algorithm reconstructs jets with a regular shape, which is very attractive experimentally.

within $|\eta| < 4.9$ which are calibrated according to the object (electron, jet etc.) that they are attached to and muons reconstructed in the muon spectrometer.

As MET is not used in this analysis, a brief description of how the MET is calculated is shown below.

The component in x and y direction is calculated by $E_{x(y)}^{miss} = E_{x(y)}^{miss,calo} + E_{x(y)}^{miss,\mu}$ where the $E_{x(y)}^{miss,calo}$ term is from the calorimeter and the $E_{x(y)}^{miss,\mu}$ term is from the muon spectrometer.

The calorimeter term $E_{x(y)}^{miss,calo}$ is calculated by:

$$E_{x(y)}^{miss,Calo} = E_{x(y)}^{miss,e} + E_{x(y)}^{miss,\gamma} + E_{x(y)}^{miss,\tau} + E_{x(y)}^{miss,jets} + E_{x(y)}^{miss,softjets} + E_{x(y)}^{miss,Calo,\mu} + E_{x(y)}^{miss,CellOut} \quad (4.4)$$

where $E_{x(y)}^{miss,CellOut}$ is for the cells that are not associated with any reconstructed objects.

The muon term $E_{x(y)}^{miss,\mu}$ is calculated from the muon momentum reconstructed with $|\eta| < 2.7$.

$$E_{x(y)}^{miss,\mu} = -\sum_{muons} E_{x(y)}^{\mu} \quad (4.5)$$

Chapter 5

Data and MC samples

5.1 Data Model

The ATLAS detector records a few petabytes of data per year, which is too large to be distributed and analyzed. Therefore a few additional reduced datasets are produced for analysis.

- Raw data is the detector output data after being triggered. The raw data size is about 1.6 Mb/Event.
- Event Summary Data (ESD) is a summary of the raw data which contains sufficient information to do track re-fitting, calibration and particle identification etc. The ESD data size is about 500 Kb/Event.
- Analysis Object Data (AOD) is a further summary of the ESD but contains sufficient information for most physics analysis. The AOD is widely used by researchers. The AOD data size is about 100 Kb/Event.
- Derived ESD (dESD) is derived from the ESD and used for detector and performance studies.
- Derived AOD (dAOD) is derived from the AOD and used for physics analyses with the exact definition defined by the needs of group analyses.

- Ntuple (D3PD) is a ROOT format data type produced from AOD/ESD/dAOD/dESD. The Ntuple is compressed and can be processed quickly. So most people choose to produce Ntuples for further use. An Ntuple is usually a few Kb/Event or even smaller depending on individual needs.

5.2 Data Samples

The data sample used in this analysis corresponds to an integrated luminosity of 4.6 fb^{-1} with a relative uncertainty of 1.8% at collision energy of 7 TeV in 2011 and of 20.7 fb^{-1} with a relative uncertainty of 3.6% at collision energy of 8 TeV in 2012 [70] [71]. The data are separated into data periods. During each period, the configuration of the detector and the beam are stable. The integrated luminosity for each period of 2012 and 2011 are shown in Table 5.1 and Table 5.2 respectively.

The data for the electron channel and for the muon analyses are from the Egamma stream and Muons stream respectively. Each High-level trigger output is organized into streams. The events in the Egamma stream must pass at least one of a set of predefined egamma triggers. Likewise, the events in Muons stream must pass at least one of a set of predefined muon triggers.

The data quality requirement is implemented through Good run list (GRL). The GRL is centrally produced. It requires all the sub-detectors are in a good condition when taking data. Also the data was checked by Data Quality Monitoring experts run by run (or lumiblock by lumiblock, where lumiblock is an unit of data-taking period, about 2 minutes). If a run/lumiblock is good, then it will be included in the GRL for physics analysis.

5.3 Monte Carlo Samples

Monte Carlo (MC) samples are simulations generated by computer programs. MC simulation consists of several models of how collisions happen and evolve including parton distribution function (pdf), leading-order (LO) and next-to-leading-order (NLO) matrix calculation, parton shower, hadronization, detector response, etc. MC samples are important to the studies of detector performance, signal efficiency, systematic uncertainties, etc.

Date	Period	Run number	Integrated Luminosity (pb^{-1})
Apr 04 - Apr 19, 2012	A	200804 $\rightarrow$ 201556	809.24
May 01- Jun 18, 2012	B	202660 $\rightarrow$ 205113	5197.27
Jul 01 - Jul 23, 2012	C	206248 $\rightarrow$ 207397	1433.91
Jul 24 - Aug 22, 2012	D	207447 $\rightarrow$ 209025	3355.74
Aug 23 - Sep 16, 2012	E	209074 $\rightarrow$ 210308	2577.49
Sep 26 - Oct 07, 2012	G	211522 $\rightarrow$ 212272	1299.57
Oct 13 - Oct 26, 2012	H	212619 $\rightarrow$ 213359	1473.68
Oct 27 - Nov 02, 2012	I	213431 $\rightarrow$ 213819	1032.90
Nov 02 - Nov 25, 2012	J	213900 $\rightarrow$ 215091	2636.88
Nov 30 - Dec 05, 2012	L	215414 $\rightarrow$ 215643	857.44
Total	A-L	200804 $\rightarrow$ 215643	20674.12

Table 5.1: Integrated Luminosity for 2012 data taking periods.

Date	Period	Run number	Integrated Luminosity (pb^{-1})
Mar 21 - Jul 29, 2011	B $\rightarrow$ I	177986 $\rightarrow$ 186493	1428.37
Jul 30 - Aug 04, 2011	J	186516 $\rightarrow$ 186755	223.49
Aug 04 - Aug 22, 2011	K	186873 $\rightarrow$ 187815	572.78
Sep 07 - Oct 30, 2011	L, M	188902 $\rightarrow$ 191933	2362.58
Total	B-M	177986 $\rightarrow$ 191933	4587.23

Table 5.2: Integrated Luminosity for 2011 data taking periods.

The more MC events used, the smaller the statistical uncertainty of the quantity estimated using MC samples. We generated sufficient signal Higgs events with equivalent luminosity of $\sim 4 \times 10^4 \text{ fb}^{-1}$ (2000 and 10000 times more than the data luminosity of 2012 and 2011 data respectively). However, due to the limitation of computing power, for the background events we only generated the events of a few times more than the data luminosity which is shown in Table 5.3. The first step of MC production is to generate parton level events by calculating the matrix elements (ME) from LO/NLO Feynman diagrams. The generated four vectors are the input to the shower program which implements the parton showering and then hadronization in order to keep the final state particles color neutral. The next step is to propagate these particles through matter (the ATLAS detector). When particles hit the detector, the interaction between particles and materials and detector response are modelled by GEANT4 [52]. The MC produces energy deposits that are used to simulate the detector signals seen in the detector; from that point on the all the same software used to analyze real data is used to analyze the MC data.

There are several MC generators used in this analysis which will be described later. They include Pythia [53] and Sherpa [54] (two general purpose generators), Alpgen [55] (Matrix element+Parton shower generator), Powheg [56] and MC@NLO [57] (two NLO generator+Parton shower generators).

5.3.1 Generators

The simulated events are made using a variety of Monte Carlo generators. In this section, a few generators are introduced including Pythia, Sherpa, Alpgen, POWHEG BOX and MC@NLO. Pythia is a full chain generator which provides showering and hadronization for some other MC generators in addition to event generation. Sherpa and Alpgen are most notably for the production of electroweak bosons in association with jets: these profited from the implemented CKKW and MLM parton-showering and matrix-element matching technique. POWHEG BOX and MC@NLO are two next-to-leading-order generators. They both can simulate the Higgs production. The POWHEG BOX is used in this analysis because it produces events with positive (constant) weight and it can be easily interfaced to any modern shower generator.

5.3.1.1 Pythia

Pythia [53] is a general purpose high-energy physics Monte Carlo sample generator. It comprises a coherent set of physics models for the evolution from a few-body hard processes to a complex multi-hadronic final state. While previous versions of Pythia6 were written in Fortran, the current release Pythia8 is in C++. It currently only works with proton-proton, proton-antiproton, e^-e^+ and $\mu^-\mu^+$ incoming beams. Besides the event generation, the parton showering and hadronization are also implemented in Pythia. This makes Pythia easy to use and also provides the interface of showering and hadronization for other generators such as POWHEG.

5.3.1.2 Sherpa

Sherpa [54] is a general purpose Monte Carlo generator that provides complete hadronic final states in simulations of high-energy particle collisions. Sherpa simulation can be achieved for the following types of collisions: lepton-lepton, lepton-photon, photon-photon, lepton-hadron and hadron-hadron collisions. Sherpa covers particle production processes at tree level in the Standard Model and in models beyond the Standard Model. Its algorithm of merging matrix elements and parton showers is described at [58]. This Matrix Elements (ME) + Parton shower matching technique describes well the production of electroweak bosons in association with jets which are wide separated with high p_T .

5.3.1.3 Alpgen

ALPGEN [55] is a tree-level matrix element (ME) calculator for a fixed number of partons in the final-state. ALPGEN describes multi-partonic final states based on the exact evaluation of the relevant Feynman diagrams in QCD and EW interactions at leading order without virtual corrections (loops). ALPGEN only generates the hard processes without any form of hadronization. In order to carry out the showering and hadronization for the output quarks and gluons, it has to interface with other general purpose generators like Pythia mentioned above. However, a parton produced during the shower evolution may be theoretically equivalent to the one already generated in the matrix element calculation. This problem is solved by using a matching procedure for matrix elements and parton showers in order to

remove events which occur twice. In the case of ALPGEN, the so called MLM matching scheme is implemented [55].

5.3.1.4 POWHEG BOX

The POWHEG BOX [56] is a general computer framework for implementing next-to-leading-order (NLO) calculations in shower Monte Carlo programs according to the POWHEG method [56]. It can be interfaced with most modern shower Monte Carlo programs for parton showering and hadronization. The current available processes include Higgs boson production in gluon fusion [59], Higgs boson production in vector boson fusion [60], Single vector-boson production with decay [61], Vector boson plus one jet production with decay [62], etc.

5.3.1.5 MC@NLO

MC@NLO [57] is the first proposed method to combine NLO matrix element contributions with shower Monte Carlo algorithms. Together with POWHEG, they are the main two NLO generators.

5.3.2 Signal and Background Samples

A short summary of signal and background samples are shown in Table 5.3 for 2012 and in Table 5.4 for 2011. A detailed list is shown in Appendix A.

In order to estimate signal selection efficiency and to model signal shape, about 100,000 signal events ($H \rightarrow Z\gamma$) were generated. It is equivalent to a dataset of luminosity of about 30,000 fb⁻¹ to 700,000 fb⁻¹ for the ggH and the VBF processes in range M_H from 120 GeV to 150 GeV. It is calculated by $Lumi = \frac{N_{event}}{\sigma_{gen}\epsilon_{filter}}$ where the N_{event} is the number of generated events; the σ_{gen} is the process cross-section; the ϵ_{filter} is the generator level filter efficiency if filters are applied.

Process	Generator	Nevents	Equivalent lumi (fb^{-1})
ggH ($H \rightarrow Z\gamma$)	Powheg	$\sim 100,000$	$\sim 4 \times 10^4$
VBF ($H \rightarrow Z\gamma$)	Powheg	$\sim 100,000$	$\sim 4 \times 10^5$
SM $Z\gamma$	Sherpa	$\sim 1,200,000$	~ 40
Z+jets	Alpgen	$\sim 8,500,000$	~ 10
$t\bar{t}$	MC@NLO	$\sim 8,000,000$	~ 400

Table 5.3: A short summary of the signal and background Monte Carlo samples for 2012 at $\sqrt{s} = 8$ TeV.

Process	Generator	Nevents	Equivalent lumi (fb^{-1})
ggH ($H \rightarrow Z\gamma$)	Powheg	$\sim 100,000$	$\sim 4 \times 10^4$
VBF ($H \rightarrow Z\gamma$)	Powheg	$\sim 100,000$	$\sim 4 \times 10^5$
SM $Z\gamma$	Sherpa	$\sim 400,000$	~ 15
Z+jets	Alpgen	$\sim 10,000,000$	~ 10
$t\bar{t}$	MC@NLO	$\sim 500,000$	~ 40

Table 5.4: A short summary of the signal and background Monte Carlo samples for 2011 at $\sqrt{s} = 7$ TeV.

Chapter 6

Event selection

This chapter presents the event selection strategy behind the search for $H \rightarrow Z\gamma$. The object level selection (muon, electron and photon) is presented first in section 6.1 and followed by event level selection requirement in section 6.2.

6.1 Object Level Selection

6.1.1 Muon selection

In this analysis, four kinds of muon candidates are used: combined muons (CB), stand-alone muons (SA), segment tagged muons (ST) and calorimeter tagged muons (CT). The four reconstruction methods are described in chapter 4.

CB, SA and ST muons are required to have transverse momentum $p_T > 10$ GeV, while CT muons must have $p_T > 15$ GeV. CB and ST muons are required to have $|\eta| < 2.7$, SA muons are required to have $2.5 < |\eta| < 2.7$, and CT muons must have $|\eta| < 0.1$. Also the quality of all kinds of muons are required to be tight. The track associated with CB, ST and CT muons is required to have a transverse impact parameter $|d_0| < 1$ mm and longitudinal impact parameter $|z_0| < 10$ mm. The distributions of d_0 and z_0 are shown in Figure 6.1.

Combined and segment/calorimeter tagged muons have associated inner detector tracks. The inner detector hit quality requirements of the CB and ST muon tracks is given in Table 6.1 for years 2011 and 2012. In track reconstruction, a hit refers to a measurement that has been associated with a track. A hole is an expected measurement given the track trajectory that

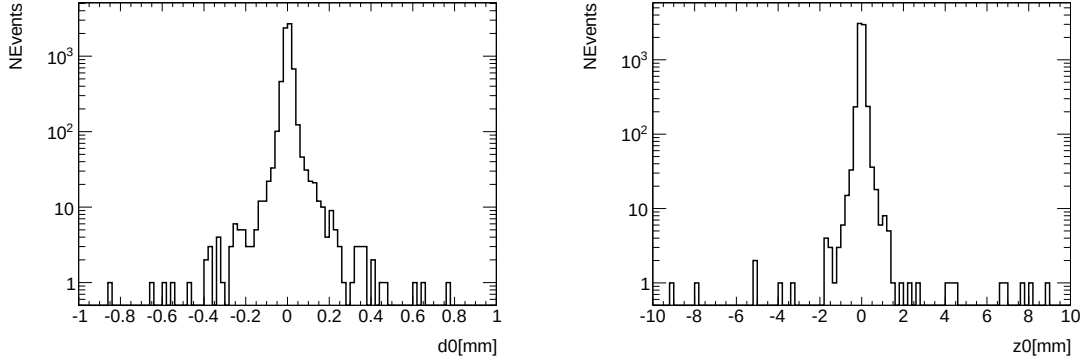


Figure 6.1: The d_0 (left) and z_0 (right) distribution before applying cuts.

has not been assigned to the track. In other words, we expect a hit if a track passes through a module and the module measurement is assigned to the track while a hole is that either the module measurement is not assigned to the track or the module does not response to the track at all. Holes result from silicon inefficiency or problems within the pattern recognition. Inactive modules are excluded from the definition of holes. For CT muons, the TRT hit requirements are replaced by $n\text{Hits} + n\text{Outliers} < 6$ OR $\frac{n\text{Outliers}}{n\text{Hits} + n\text{Outliers}} < 0.9$ where $n\text{Hits}$ denotes the number of TRT hits on the muon track and $n\text{Outliers}$ denotes the number of TRT hits on the track associated as outliers¹.

The Muon Spectrometer hit requirements that only apply to SA muons is $n\text{CSCEtaHits} + n\text{CSCPhiHits} > 0$ and $n\text{MDTEMHits} > 0$ and $n\text{MDTEOHits} > 0$, where $n\text{CSCEtaHits}$ and $n\text{CSCPhiHits}$ are the number of hits in CSC η and ϕ strips respectively. $n\text{MDTEMHits}$ and $n\text{MDTEOHits}$ are the number of hits in the MDT endcap middle station and outer station respectively.

The overlap between muon candidates reconstructed by different algorithms, for the same particle, is removed by requiring that SA muons are not within $\Delta R = 0.2$ of a ST muon and CT muons are not within $\Delta R = 0.2$ of a CB or SA or ST muons.

¹A outlier is a hit associated to the track but the hit is produced by a module which is not passed by the track.

2012	
ID Si hit requirement	$\text{expectBLayerHit}=\text{false}$ or $\text{numberOfBLayerHits} \geq 1$ No. of pixel hits + No. of crossed dead pixel sensors > 0 No. of SCT hits + No. of crossed dead SCT sensors > 4 No. of pixel holes + No. of SCT holes < 3
TRT hit requirements:	
$0.1 < \eta \leq 1.9$	$\text{nHits} + \text{nOutliers} > 5 \ \& \ \frac{\text{nOutliers}}{\text{nHits}+\text{nOutliers}} < 0.9$
$ \eta < 0.1$ or $ \eta \geq 1.9$	if $(\text{nHits} + \text{nOutliers} > 5)$: $\frac{\text{nOutliers}}{\text{nHits}+\text{nOutliers}} < 0.9$
2011	
ID Si hit requirement	$\text{expectBLayerHit}=\text{false}$ or $\text{numberOfBLayerHits} \geq 1$ No. of pixel hits + No. of crossed dead pixel sensors > 1 No. of SCT hits + No. of crossed dead SCT sensors > 5 No. of pixel holes + No. of SCT holes < 3
TRT hit requirements:	
$ \eta < 1.9$	$\text{nHits} + \text{nOutliers} > 5 \ \& \ \frac{\text{nOutliers}}{\text{nHits}+\text{nOutliers}} < 0.9$
$ \eta \geq 1.9$	if $(\text{nHits} + \text{nOutliers} > 5)$: $\frac{\text{nOutliers}}{\text{nHits}+\text{nOutliers}} < 0.9$

Table 6.1: Inner Detector (ID) hit requirements for the CB and ST muon tracks. The requirements are constructed such that detector conditions are taken into account for the silicon systems: if a track traverses a dead silicon module (which is the “No. of crossed dead pixel/SCT sensors”), it is added to the hit count. nHits denotes the number of TRT hits on the muon track and nOutliers denotes the number of TRT outliers on the muon track.

6.1.2 Electron selection

Electron candidates are required to have transverse energy $E_T > 10$ GeV and pseudorapidity $|\eta| < 2.47^2$. An electron candidate is rejected if its reconstructed energy is affected by a dead front-end-board (FEB) or high-voltage (HV) failures or masked cells in the core (defined as the 3×3 cells in the middle layer), namely $el_OQ \ \& \ 1446 = 0$ (all checks are good). The electron candidates must have a well-reconstructed ID track pointing to an

²The Inner detector η coverage is $|\eta| < 2.5$

electromagnetic calorimeter cluster and the shower shapes of the cluster must satisfy a set of identification criteria (the requirement of shower shapes and hadronic leakage are called loose++ in ATLAS). The four-momentum of the electron is formed by using the energy measured by the calorimeter and the pseudo rapidity and azimuth measured in the inner detector. If the electron passes through an active B-layer (the first layer of the pixel detector, closest to the beam), the track is required to have a hit in the B-layer in order to reduce fake electron candidates from photon conversions. The longitudinal impact parameter of the track with respect to the primary vertex is required to be less than 10 mm.

If two selected electrons share the same ID track, the electron with the higher E_T is kept. In 2012 data, because of higher pileup, if two electron candidates are close to each other with $|\Delta\eta| < 0.075$ and $|\Delta\phi| < 0.125$, the electron with higher E_T is kept. If an electron candidate shares an ID track with a selected muon within $\Delta R = 0.02$, the electron candidate is rejected.

6.1.3 Photon selection

The photon is critical to this analysis since it is one of the final state particles we are looking for. The photon candidates are required to have transverse energy $E_T > 15$ GeV and pseudorapidity $|\eta| < 1.37$ or $1.52 < |\eta| < 2.37$. You might notice that the difference of the η range between the photon and electron selection. The reason is that for photon, one important background is $\pi^0 \rightarrow \gamma\gamma$. In order to have a good $\gamma - \pi^0$ separation, we want to use the fine granularity of the strip layer. Therefore the Barrel-End-cap transition region ($1.37 < |\eta| < 1.52$) is excluded. For $|\eta| > 2.37$, the strips width changes from $\frac{0.025}{4}$ to 0.025, therefore this region is excluded as well. A photon candidate is rejected if its reconstructed energy is affected by a dead front-end-board (FEB) or high-voltage (HV) failures or masked cells in the core, namely $ph_OQ \& 34214 = 0$. A tight identification requirement is applied. Since there is a difference of shower shapes between Monte Carlo simulation and data, the Monte Carlo shower shapes are corrected first by using a tool developed by the ATLAS egamma working group and then a tight identification cut is applied [73]. The shower shapes of photon clusters before and after the correction are shown in Figure 6.2.

In addition, an isolation requirement is applied in order to reduce jet-faked photons.

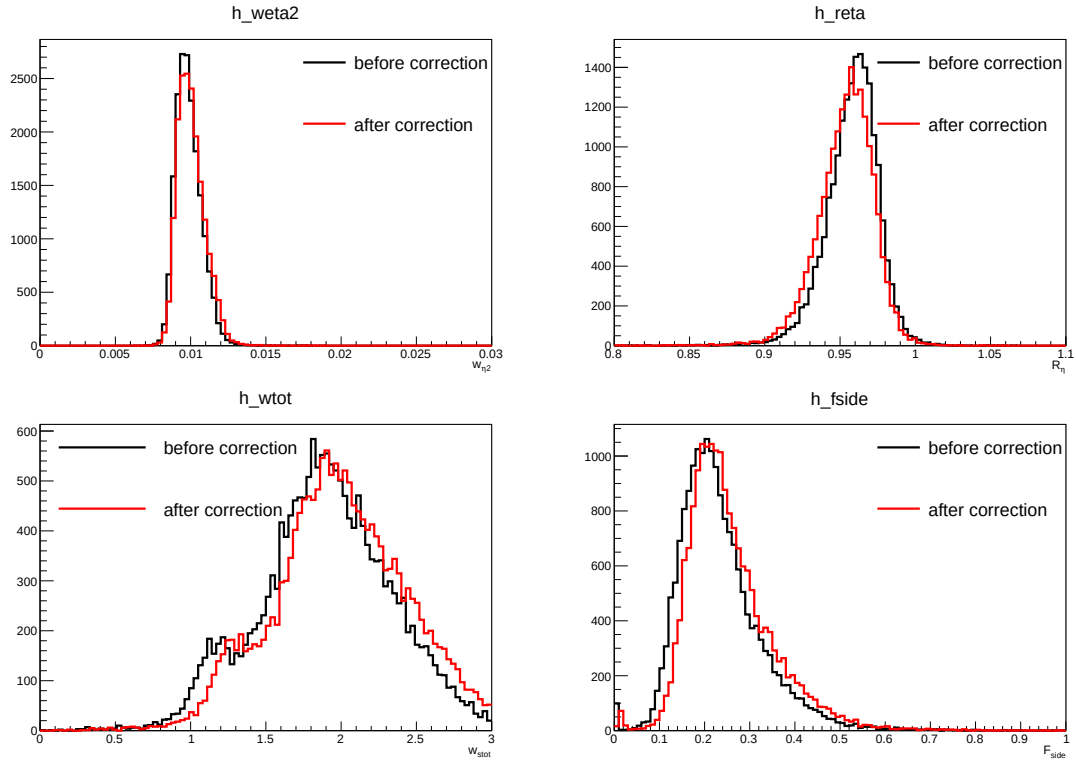


Figure 6.2: The distribution of the photon shower shapes variables before and after corrections. The upper two plots are for $w_{\eta 2}$ and R_{η} . The lower two plots are for w_{shot} and F_{side} . The definition of these variables are shown in Figure 4.7.

The calorimeter isolation, E_T^{iso} , based on topological clusters [47] in a cone of $\Delta R = 0.4$ around the photon candidate is required to be smaller than 4 GeV. The photon isolation distribution is shown in Figure 8.5. The Z +jets background dominates the region with $E_T^{iso} > 4$ GeV. Topological clusters are based on a clustering algorithm which clusters together neighboring cells that have energies that are significant compared to the expected noise (The threshold ranges from 3 to 6 corresponding to the low to high threshold.). This results in clusters that have a variable number of cells, in contrast to the fixed-sized clusters produced by the sliding-window algorithm.

If photon candidates are within $\Delta R < 0.3$ of a selected electron or muon candidate, then they are rejected.

6.2 Event Level Selection

6.2.1 Preselection

In order to remove problematic data runs, a Good Run List (GRL) requirement [48] is applied on data. Bad data events with LAr or Tile errors are also removed. For 2012, corrupted data events (a fraction of about 6×10^{-6} of total events) are removed as well. The Trigger requirement for both data and MC for all the run periods is given in Table 6.2.

6.2.2 Z boson reconstruction and selection

Z boson is important to this analysis since it is one of the final state particles we are looking for. Z boson candidates are reconstructed from a pair of opposite charge, same flavor leptons (e/μ). In the case of multiple Z candidates in the same event³, the candidate with the invariant mass closest to the Z pole is chosen. Also, the leptons used to reconstruct Z bosons are required to match the trigger object(s) of the (single or double) lepton trigger used to select the event. An invariant mass requirement $M_{ll} > 65$ GeV is applied in order to remove biases in the simulated Z +jet samples which have a generator-level dilepton

³The multiple Z events only account for less than 0.8% of the total selected signal events.

Year/period	Electron triggers	Muon triggers
2011 B-I	e20_medium	mu18_MG
	2e12_medium	2mu10_loose
2011 J	e20_medium	mu18_MG_medium
	2e12_medium	2mu10_loose
2011 K	e22_medium	mu18_MG_medium
	2e12T_medium	2mu10_loose
2011 L-M	e22vh_medium1	mu18_MG_medium
	2e12Tvh_medium	2mu10_loose
2012 A-L	e24vhi_medium1	mu24i_tight
	e60_medium1	mu36_tight
	2e12Tvh_loose1	mu18_tight_mu8_EFFS
		2mu13

Table 6.2: Triggers used in 2011 and 2012. The ‘medium1’ is re-optimized to medium at higher luminosity. The ‘vh’ triggers apply a cut on the hadronic core isolation (≤ 1 GeV) at L1. The ‘i’ triggers apply a cut on track isolation at EF (PTcone20/PT). The ‘T’ is for tighter L1 requirement in order to keep the L1 rates within the allowed bandwidth without raising the threshold at the HLT.

invariant mass cut with a lower threshold of 60 GeV. The simulated and observed Z mass distribution are shown in Figure 6.3 and Figure 6.4 respectively.

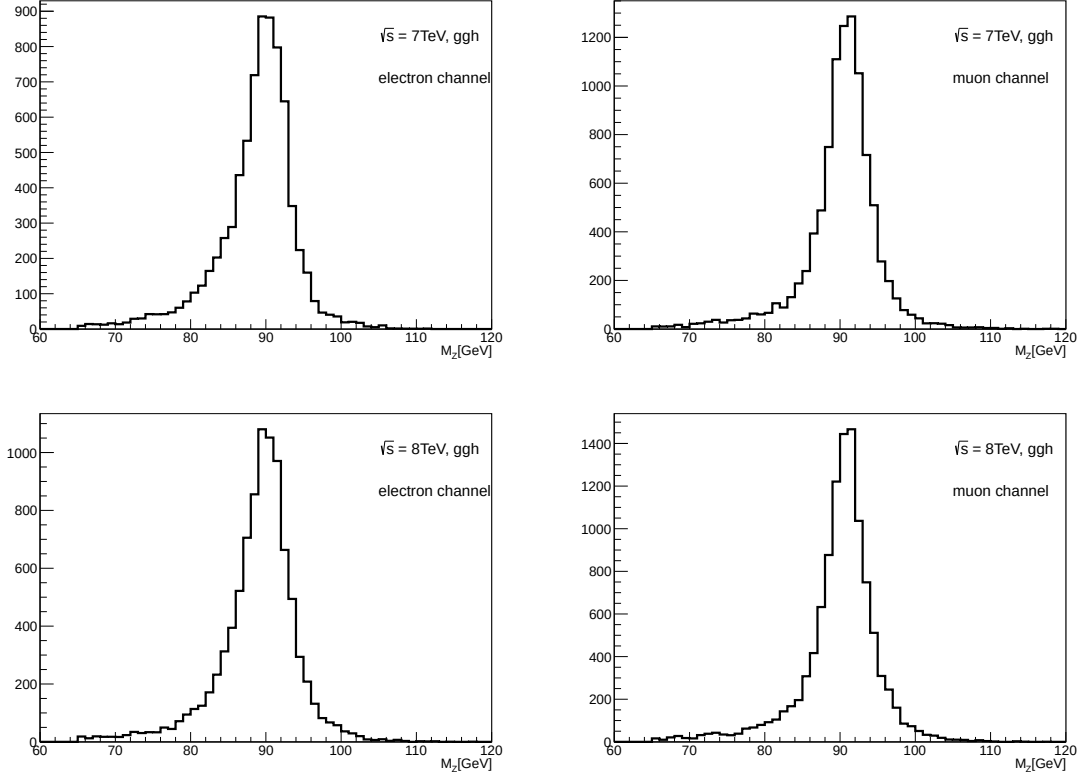


Figure 6.3: The simulated Z mass distribution at $M_H = 125$ GeV from ggH processes for electron (left) and muon (right) channels at $\sqrt{s} = 7$ TeV (up) and $\sqrt{s} = 8$ TeV (down).

6.2.3 Higgs reconstruction and selection

Higgs boson candidates are reconstructed from Z boson and photon with the highest transverse energy⁴. Normalized track isolation and calorimeter isolation requirements are applied to the leptons used to reconstruct the Higgs boson. The normalized track isolation and calorimeter isolation distribution are shown in Figure 6.5. The normalized track isolation is defined as the sum of the transverse momenta of tracks, $\sum p_T$, inside a cone of $\Delta R = 0.2$

⁴The signal events with more than one selected photon is about 0.2% of the total selected events.

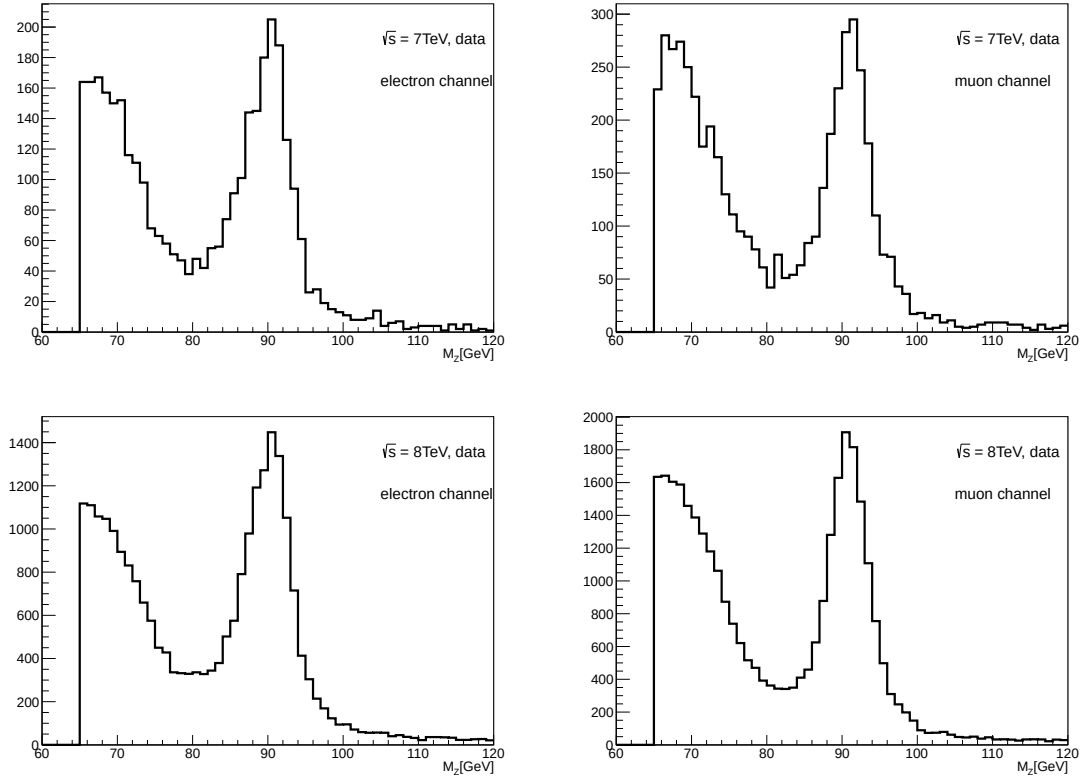


Figure 6.4: The observed Z mass distribution at $M_H = 125$ GeV for electron (left) and muon (right) channels at $\sqrt{s} = 7$ TeV (up) and $\sqrt{s} = 8$ TeV (down). There are two peaks on the plot. The Z peak is around 91 GeV while the low mass peak is mainly due to final state radiation (FSR) of leptons.

around the lepton, excluding the lepton track, divided by the lepton p_T . The normalized track isolation is required to be smaller than 0.15. The normalized calorimeter isolation is defined as the sum of the calorimeter cell energies in 2011 (positive-energy topological clusters were used in 2012) within a cone of $\Delta R = 0.2$ around the candidate lepton divided by the lepton E_T . It is required to be less than 0.3 in 2011 and 0.2 in 2012 for electrons ⁵, and less than 0.3 (0.15 for SA muons) for muons in both years.

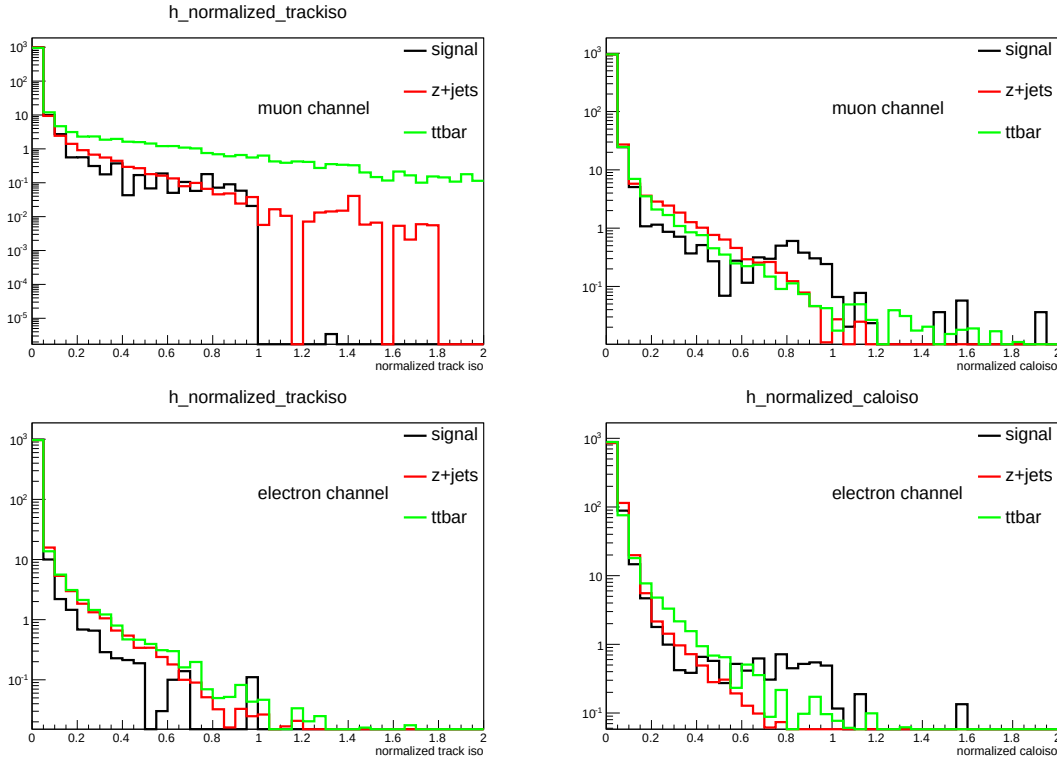


Figure 6.5: The normalized track isolation and the normalized calorimeter isolation distribution of the signal events, z +jets and $t\bar{t}$ events in muon (upper plots) and electron (lower plots) channel. The total number of events of the three samples are normalized to the same number.

In addition, a track impact parameter significance requirement is applied. Its distribution is shown in Figure 6.6. For electrons, the Impact parameter significance is required to

⁵The calorimeter isolation based on topoclusters is less than that based on cells. In order to have similar background rejection power, the normalized calorimeter isolation requirement is 0.2 in 2012 and 0.3 in 2011.

be smaller than 6.5 and for muons it is required to be smaller than 3.5.

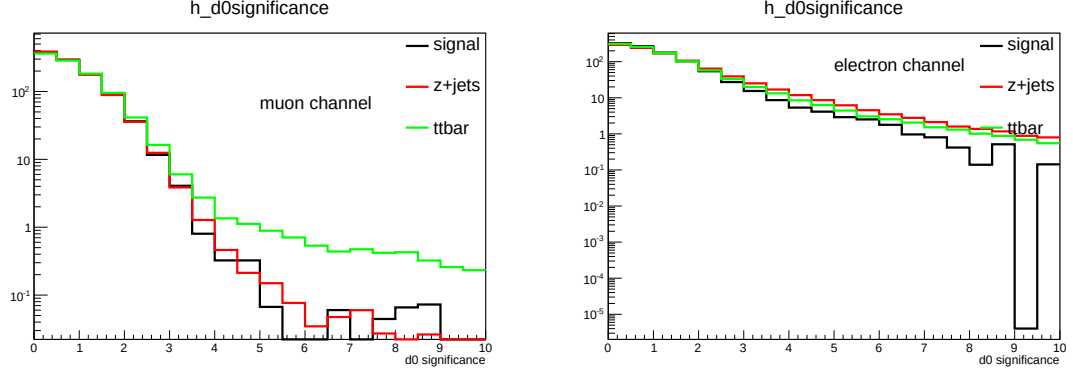


Figure 6.6: The transverse impact parameter significance distribution of the signal events, z +jets and $t\bar{t}$ events in muon (left) and electron (right) channel. The total number of events of the three samples are normalized to the same number.

Finally, to suppress final state radiation (FSR) events with $Z + \gamma$ signature, the mass of Z candidates is required to be above 81.18 GeV. This FSR effect is shown in Figure 6.7.

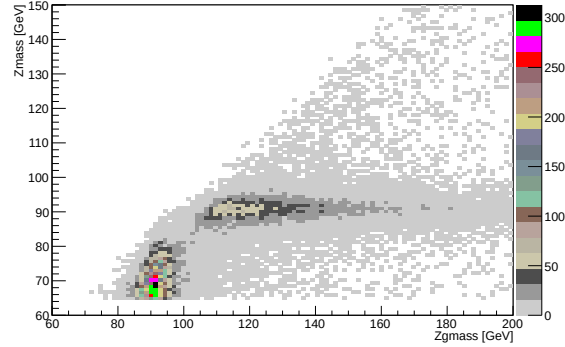


Figure 6.7: The observed Z mass Vs $Z\gamma$ mass distribution for muon channel at $\sqrt{s} = 8$ TeV. The final state radiation events are in the region where $Z\gamma$ mass is around 91 GeV.

Chapter 7

Signal Property Studies

When we are looking for a new particle, there are in general two ways to do it. One is the direct search by measuring the invariant mass of the particle, which is the mass in the rest frame of the particle. The invariant mass is one of the key characteristics as it remains the same in any reference frame. The other one is the indirect search by measuring the cross-section or coupling and trying to find difference between the measured quantity and the expected one. In this analysis we use the direct search and hence the mass reconstruction is our focus here. The following chapters will introduce the signal ($H \rightarrow Z\gamma$) mass reconstruction, techniques for improving the reconstructed mass resolution, the signal shape model and its expected yield.

7.1 Signal Mass reconstruction

The reconstructed $Z\gamma$ mass is of interest and it is calculated by

$$m^2 = p_\mu p^\mu = E^2 - \vec{P}^2 = (E_Z + E_\gamma)^2 - (\vec{P}_Z + \vec{P}_\gamma)^2 \quad (7.1)$$

The Z boson is a real Z (also referred to as on shell). It means that the equation

$$m^2 = E^2 - \vec{P}^2 \quad (7.2)$$

is satisfied and that the calculated m is the rest mass of the particle. (If a particle is a virtual one, its calculated m is not close to its rest mass and the m^2 could even be negative.)

<i>Z</i> decay modes	
mode	Fraction
e^+e^-	$(3.363 \pm 0.004)\%$
$\mu^+\mu^-$	$(3.366 \pm 0.007)\%$
$\tau^+\tau^-$	$(3.367 \pm 0.008)\%$
l^+l^-	$(3.3658 \pm 0.0023)\%$
invisible	$(20.00 \pm 0.06)\%$
hadrons	$(69.91 \pm 0.06)\%$

Table 7.1: Main *Z* boson decay modes. *l* indicates each type of lepton (*e*, μ , and τ).

The mean lifetime of the Higgs with a mass of 126 GeV is about 1.56×10^{-22} second, so that it will decay immediately after being created. The *Z* boson also has a short lifetime of about 3×10^{-25} second. It decays immediately after being produced. The *Z* boson decay modes are shown in Table 7.1. Hence we cannot measure *Z* boson directly, instead we measure the lepton pair from *Z* boson, from which we reconstruct the *Z* kinematics.

In the experiment, we need to measure the photon and leptons' energy and momentum in order to calculate the *Z* mass. The two high quality tracks of the lepton pair make the leading primary vertex the best estimator of the Higgs decay vertex. Photons that are measured by the calorimeter, using calorimeter coordinates, need to be corrected to the primary vertex. In this way, all three objects are constrained to come from the same vertex. The photon vertex correction is introduced in the next section. Any improvements to the mass peak resolution lead to an improved signal to background ratio which directly improves the signal significance. Additional improvements to the significance can be achieved via a method called the *Z* mass constraint; it has been used to improve the *Z* γ mass resolution. The *Z* mass constraint uses the standard Breit-Wigner *Z* shape model and the assumption that the detector resolution can be modeled by a Gaussian function.

7.2 Signal Mass resolution improvement

7.2.1 Photon vertex correction

The photon η is measured using the calorimeter (the $\eta-\phi$ coordinates are with respect to the detector origin). When a collision event occurs, the position of the collision is referred to a vertex. However, due to pileup effects shown in Figure 2.7, there may be many reconstructed vertices. The one with the highest sum of squared track p_T associated to that vertex, namely Σp_T^2 , is called the leading primary vertex. The photon vertex correction assigns the leading primary vertex (PV) as the photon vertex. Based on the presence of two leptons' tracks of high quality, Figure 7.1 shows that the leading primary vertex is a better estimator of the true Z vertex than the primary vertex closest to the weighted-averaged z position of the two leptons.

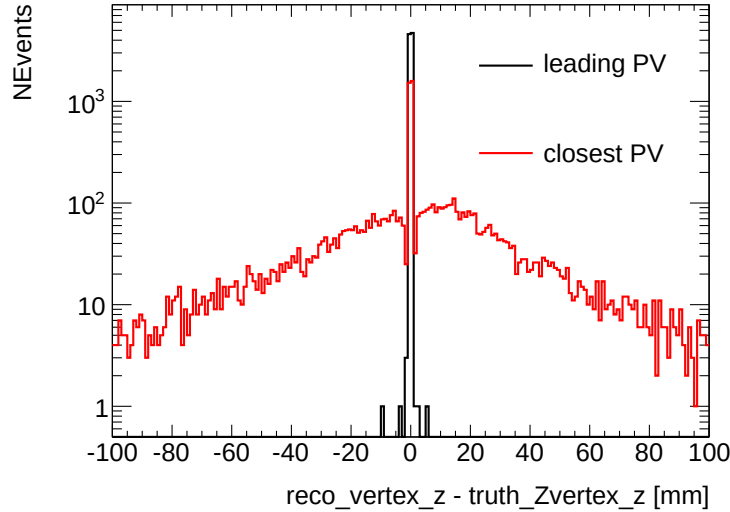


Figure 7.1: A study of the Z boson vertex estimators. The difference between the reconstructed vertex and the true Z vertex in z is shown. The leading PV (in black) is the primary vertex with the largest Σp_T^2 and the closest PV (in red) is the primary vertex closest to the combined vertex position in z . The combined vertex is calculated by using two leptons' z_0 positions weighted by their measurement uncertainties.

In Figure 7.3, we illustrate how the correction is determined, where:

- R as the perpendicular distance from the barycenter of photon in the second layer of Electromagnetic Calorimeter (ECAL) to the beam line
- L is the distance along the beam line between detector origin and the barycenter of the photon in the second layer of the ECAL. The barycenter of the photon is the energy weighted average position
- θ_{s2} is the angle corresponding to η_{s2} , where η_{s2} is measured using two points: the detector origin and the photon energy barycenter in the second layer of the ECAL
- θ_p is the angle corresponding to η_p , where η_p is measured using the photon energy barycenter in the first and second layers of the ECAL
- θ_x is the angle corresponding to the corrected η_x , where η_x is measured using two points: the primary vertex and the photon energy barycenter in the second layer
- Z_{vertex} is the z position along the beam line corresponding to η_p
- Z_{PV} shown in Figure 7.2 is the z position along the beam line corresponding to η_x
- 0 is the detector origin

The geometric relations of these variables are:

$$\begin{cases} \tan(\theta_p) = \frac{R}{L - Z_{vertex}} \\ \tan(\theta_{s2}) = \frac{R}{L} \\ \tan(\theta_x) = \frac{R}{L - Z_{PV}} \end{cases} \quad (7.3)$$

By solving these three equations, we find $\tan\theta_x = \frac{\tan\theta_{s2} * Z_{vertex}}{Z_{vertex} - Z_{PV} * (1 - \frac{\tan\theta_{s2}}{\tan\theta_p})}$; by using the definition $\eta = -\ln(\tan(\frac{\theta}{2}))$, we determine the corrected η . Note: if the corrected $\theta < 0$, θ should be transformed to $[0, \pi]$ by adding π . A typical correction distribution is shown in Figure 7.4 and 7.5.

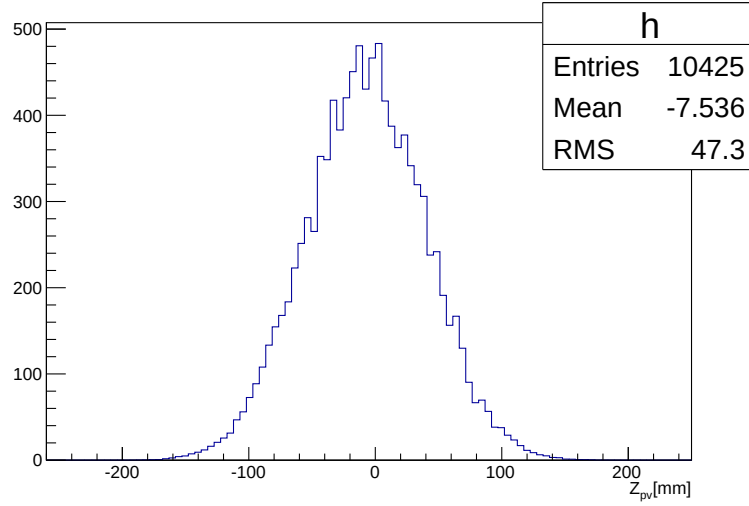


Figure 7.2: The z position of the leading primary vertex along the beam line. The mean is at about -8 mm and the width is about 47 mm.

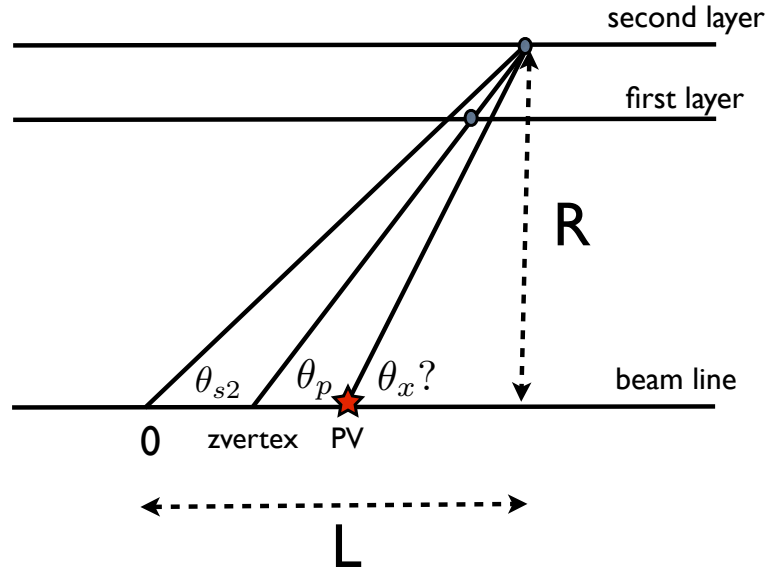


Figure 7.3: Illustration of the photon vertex correction scheme. The “0” position is the detector origin. The first (second) layer refers to the first (second) layer of the calorimeter.

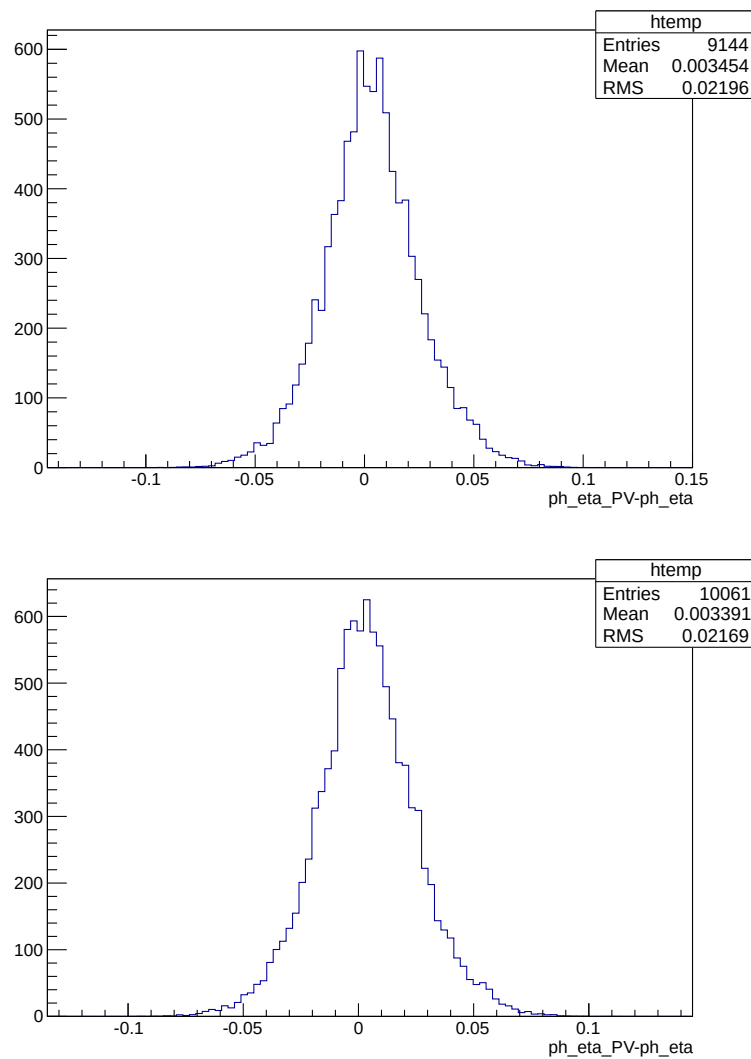


Figure 7.4: The difference, in radians, between corrected η_x and original η_{s2} in electron channel (top) and in muon channel (bottom) for the ggh MC sample at 125 GeV for 2012.

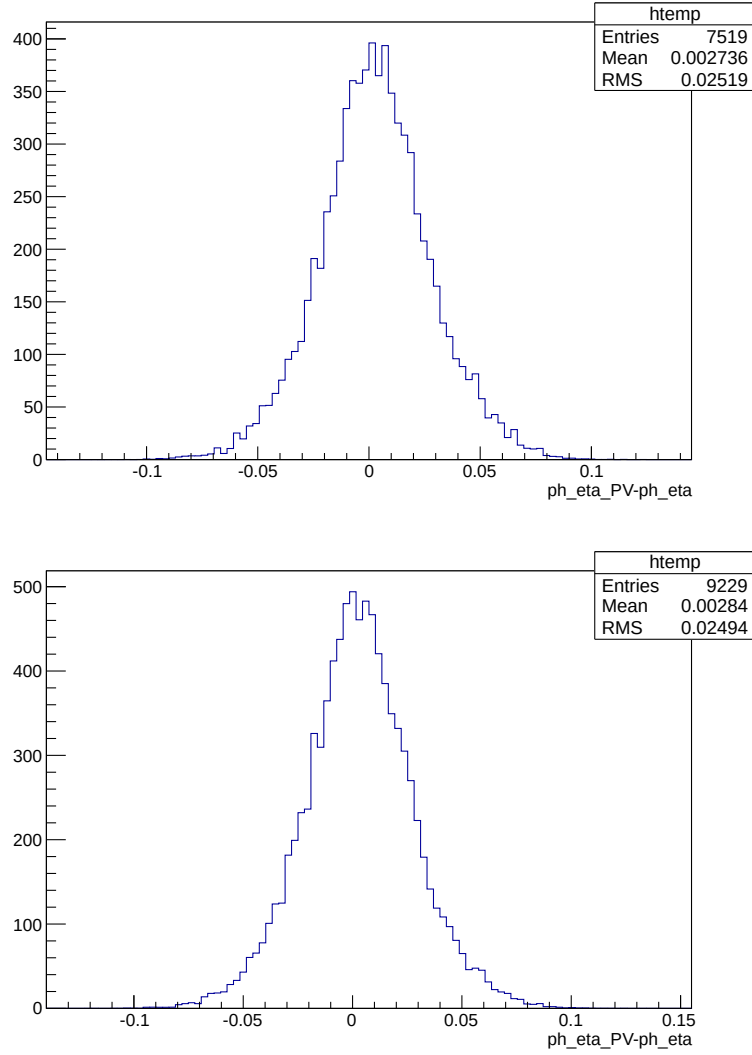


Figure 7.5: The difference between corrected η_x and original η_{s2} in electron channel (top) and in muon channel (bottom) for ggH MC sample at 125 GeV for 2011.

7.2.2 Z mass constraint

The Z mass constraint refers to the procedure of correcting the reconstructed Z boson mass by noting that the Z boson is real and an assumption that the measured mass uncertainty can be modeled by a Gaussian function. This algorithm was first used by the Higgs $\rightarrow ZZ$ group [72]. We define:

- m_{ll} as the reconstructed Z mass. $m_{ll} = \sqrt{(E_{l1} + E_{l2})^2 - (\vec{P}_{l1} + \vec{P}_{l2})^2}$
- m_{true} as the true Z mass, sampled from a relativistic Breit-Wigner distribution.
- σ_m as the detector mass resolution of the reconstructed Z mass. It is an event-by-event value whose typical value is 1.5 GeV, ranging from a few hundred MeV to a few GeV.
- Γ as the natural width of Z boson, which is 2.45 GeV.

The Likelihood function

$$L(m_{true}) = P(m_{ll}|m_{true})P(m_{true}), \quad (7.4)$$

and

$$P(m_{true}) \sim \frac{1}{(m_{true}^2 - M_z^2)^2 + \Gamma^2 M_z^2} \quad (7.5)$$

where $M_z = 91.187$ GeV.

It models the relativistic Breit-Wigner distribution. The detector mass resolution is modeled by a Gaussian function.

$$P(m_{ll}|m_{true}) = \frac{1}{\sqrt{2\pi}\sigma_m} e^{-\frac{(m_{ll}-m_{true})^2}{2\sigma_m^2}} \quad (7.6)$$

σ_m is the Z mass resolution on an event-by-event basis and is computed from the six components of two lepton momenta $\vec{x} \equiv \{p_x^1, p_y^1, p_z^1, p_x^2, p_y^2, p_z^2\}$ and from their covariance matrix Σ . The component $\Sigma_{ij} = \text{cov}(\vec{x}_i, \vec{x}_j)$. That is $\sigma_m = \frac{\sigma_{m^2}}{2m}$, which comes from $\sigma_{m^2} = \sqrt{(\frac{\partial m^2}{\partial m})^2 \sigma_m^2}$. Also $\sigma_{m^2} = \sqrt{J^T \Sigma J}$, where $J = \{\frac{\partial m^2}{\partial x_1}, \frac{\partial m^2}{\partial x_2}, \dots, \frac{\partial m^2}{\partial x_6}\}^T$. The covariance matrix Σ of $\vec{x}$ is calculated using the measured variables $\{d_0, z_0, \phi, \theta, E\}$ for electrons and $\{d_0, z_0, \phi, \theta, \frac{q}{p}\}$ for muons and also their respective covariance matrix V . The transverse impact parameter d_0 is defined as the distance of closest approach of a track to the beam

line. The longitudinal impact parameter z_0 is defined as the value of z of the point on the track that determines d_0 . The parameter ϕ is the azimuth (the angle between the direction of motion of the track in the (x, y) plane and the x -axis) at the point of closest approach. The parameter $\tan(\theta)$, with θ the polar angle, gives the slope of the track in the $r - z$ plane where $r = \sqrt{x^2 + y^2}$. E is the energy of the electron measured by the Calorimeter. $\frac{q}{p}$ is the charge over momentum for the muons. Assume Σ can be modeled by

$$\begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{pmatrix}$$

So we need to change variables from $\{d_0, z_0, \phi, \theta, E\}$ for electrons and from $\{d_0, z_0, \phi, \theta, \frac{q}{p}\}$ for muons to $\{d_0, z_0, p_x, p_y, p_z\}$ and hence get Σ by a Jacobian transformation. The Jacobian matrix for electron is

$$J' = \frac{\partial\{d_0, z_0, p_x, p_y, p_z\}}{\partial\{d_0, z_0, \phi, \theta, E\}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -P \sin(\theta) \sin(\phi) & P \cos(\theta) \cos(\phi) & \sin(\theta) \cos(\phi) \\ 0 & 0 & P \sin(\theta) \cos(\phi) & P \cos(\theta) \sin(\phi) & \sin(\theta) \sin(\phi) \\ 0 & 0 & 0 & 0 & \cos(\theta) \end{pmatrix}$$

The Jacobian matrix for the muon is transformed first from $\{d_0, z_0, \phi, \theta, \frac{q}{p}\}$ to $\{d_0, z_0, \phi, \theta, p\}$ and then to $\{d_0, z_0, p_x, p_y, p_z\}$.

$$J' = \frac{\partial\{d_0, z_0, p_x, p_y, p_z\}}{\partial\{d_0, z_0, \phi, \theta, \frac{q}{p}\}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -P \sin(\theta) \sin(\phi) & P \cos(\theta) \cos(\phi) & \sin(\theta) \cos(\phi) \\ 0 & 0 & P \sin(\theta) \cos(\phi) & P \cos(\theta) \sin(\phi) & \sin(\theta) \sin(\phi) \\ 0 & 0 & 0 & 0 & \cos(\theta) \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -p^2 \end{pmatrix}$$

Then $\Sigma_{e/\mu} = J'_{e/\mu} V_{e/\mu} J'^T_{e/\mu}$. Σ_i corresponds to the bottom right 3×3 block of $\Sigma_{e/\mu}$.

7.2.3 Summary of the photon vertex correction and the Z mass constraint

With both photon vertex correction and Z mass constraint applied, and fitting with a Crystal Ball + Gaussian function, an effective $\sigma_{eff} \equiv f * \sigma_{cb} + (1 - f)\sigma_{gs}$, the effective

signal resolution at 125 GeV is improved by 23%, going from 2.35 GeV to 1.8 GeV, in the electron channel, and it is improved by 26%, going from 2.12 GeV to 1.57 GeV, in the muon channel. The reconstructed $M_{Z\gamma}$ with and without photon vertex correction and Z mass constraint is shown in Figure 7.6.

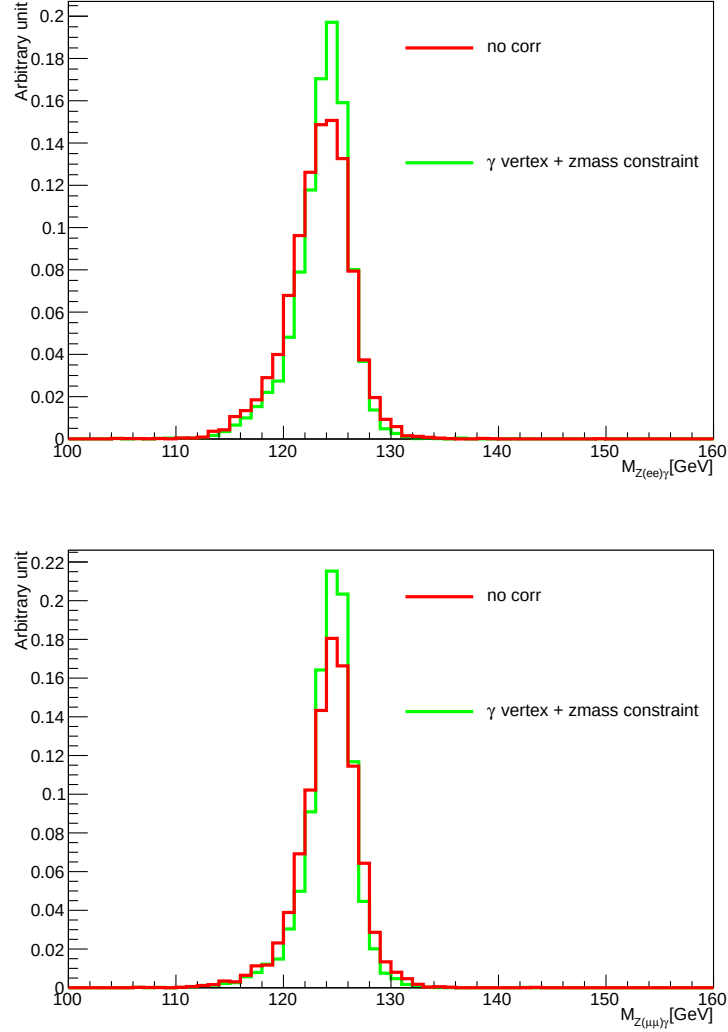


Figure 7.6: Simulated $H \rightarrow Z\gamma$ mass reconstruction for $M_H=125$ GeV in electron channel (top) and muon channel (bottom). The red line is the reconstructed $M_{Z\gamma}$ without the photon vertex correction or the Z mass constraint applied and the green line is with the both corrections applied.

7.3 Signal Modeling

A Crystal Ball function (CB) [74] plus a Gaussian function is used to describe the reconstructed signal $M_{Z\gamma}$ and $\Delta M = M_{Z\gamma} - M_{ll}$ shape. An example plot with the CB and Gaussian components are shown in Figure 7.7. The CB function characterizes the peak of the signal and Gaussian function describes the wide tail on the two sides. A simultaneous fit is performed across all available signal samples in each of four categories (for each category there are 7 samples for ggH production and 7 samples for VBF production ranging from $M_H = 120$ GeV to 150 GeV in steps of 5 GeV). There are four parameters for a CB function including μ_{cb} , σ_{cb} , α_{cb} and n_{cb} . μ_{cb} models mean of a peak. σ_{cb} models the width of a peak. The Crystal Ball function is a Gaussian that is “connected” to an exponential tail at $\alpha_{cb} \sigma_{cb}$ of the Gaussian. The sign of α_{cb} determines if the tail happens on the left or on the right. The n_{cb} models the slope of the tail. There are two parameters for a Gaussian function including μ_{gs} and σ_{gs} , modeling the mean and width respectively. μ_{cb} , μ_{gs} , σ_{cb} are fit with a linear model of Higgs mass M_H and α_{cb} , n_{cb} and $\frac{\sigma_{gs}}{\sigma_{cb}}$ are global variables, independent of M_H .

The Crystal Ball function is defined as :

$$f(x; \mu, n, \alpha, \sigma) = N \begin{cases} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, & \frac{x-\mu}{\sigma} > -|\alpha| \\ A(B - \frac{x-\mu}{\sigma})^{-n}, & \frac{x-\mu}{\sigma} \leq -|\alpha| \end{cases} \quad (7.7)$$

where

$$A = \left(\frac{n}{|\alpha|}\right)^n e^{-\frac{|\alpha|^2}{2}}$$

$$B = \frac{n}{|\alpha|} - |\alpha|$$

$$N = \frac{1}{\sigma(C+D)}$$

$$C = \frac{n}{|\alpha|} \frac{1}{n-1} e^{-\frac{|\alpha|^2}{2}}$$

$$D = \sqrt{\frac{\pi}{2}} (1 + \text{erf}(\frac{|\alpha|}{\sqrt{2}}))$$

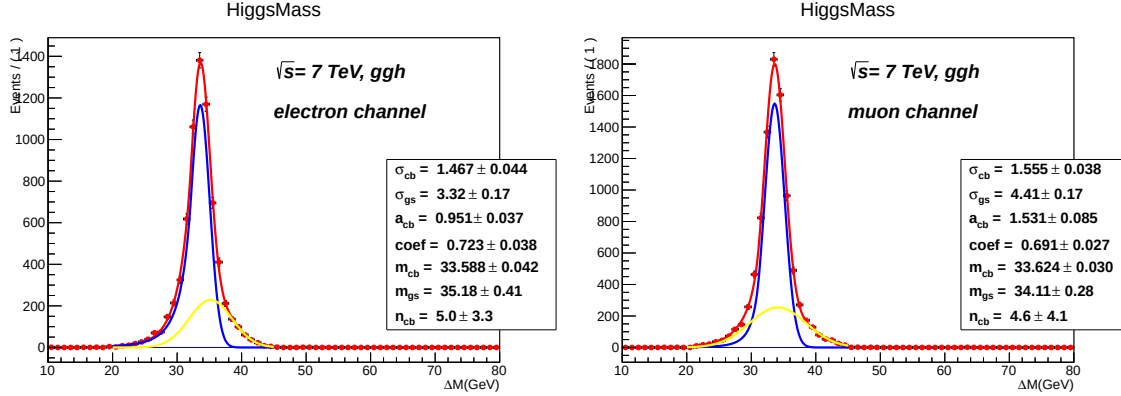


Figure 7.7: The reconstructed $\Delta M = M_{Z\gamma} - M_{ll}$ for simulated $H \rightarrow Z\gamma$ at mass point $M_H = 125$ GeV for the electron channel (left) and the muon channel (right). The red line is the combined fit result. The blue line is the CB component and the yellow line is the Gaussian component. Due to the energy loss of electrons by Bremsstrahlung the Gaussian component of the electron channel shifts to the right of the peak position while for the muon channel it is symmetric around the peak.

7.4 Signal acceptance and expected yield

ATLAS full detector simulation samples were run for Higgs mass points ranging from 120 GeV to 150 GeV in steps of 5 GeV. Each sample has about 100 K events, which will be used to estimate the signal selection efficiency and hence the expected signal yield.

$$N_{i,l}^{expected} = Luminosity \times \sigma_i(H) \times BR(H \rightarrow Z\gamma) \times BR(Z \rightarrow ll) \times \varepsilon_{i,l} \quad (7.8)$$

where i indicates the Higgs production process including gluon-gluon fusion (ggH), vector boson fusion (VBF), WH, ZH, $t\bar{t}H$ and l indicates the lepton flavor: electron and muon. $\sigma_i(H)$ and $BR(H \rightarrow Z\gamma)$ and $BR(Z \rightarrow ll)$ are listed in Table 7.2. $\varepsilon_{i,l}$ and $N_{i,l}^{expected}$ are listed in Table 7.3, 7.4, 7.5, 7.6.

For the ggH process, the signal cross section is computed up to next-to-next-to-leading order (NNLO) in QCD [75; 76; 77; 78; 79; 80; 81] and up to Next-to-leading order (NLO) electroweak (EW) corrections [82; 83]. For the VBF process, full QCD and EW corrections up to NLO [84; 85; 86; 87] and approximate NNLO QCD corrections [88] are used to

calculate the cross section. For the associated W/ZH processes (VH), the cross section is calculated up to NNLO [89; 90; 91] in QCD correction and NLO [92] in EW corrections. The cross sections for the $t\bar{t}H$ process are estimated up to NLO QCD [93; 94; 95; 96; 97].

A typical cutflow without applying correction weights is shown in Table 7.7 for the 2012 electron channel and Table 7.8 for the 2012 muon channel.

The signal selection efficiency for each mass point is modeled by a second order polynomial function of M_H , in order to interpolate the efficiency where full detector simulation samples are not available. This is shown in Figure 7.8 for 7 TeV samples and in Figure 7.9 for 8 TeV samples.

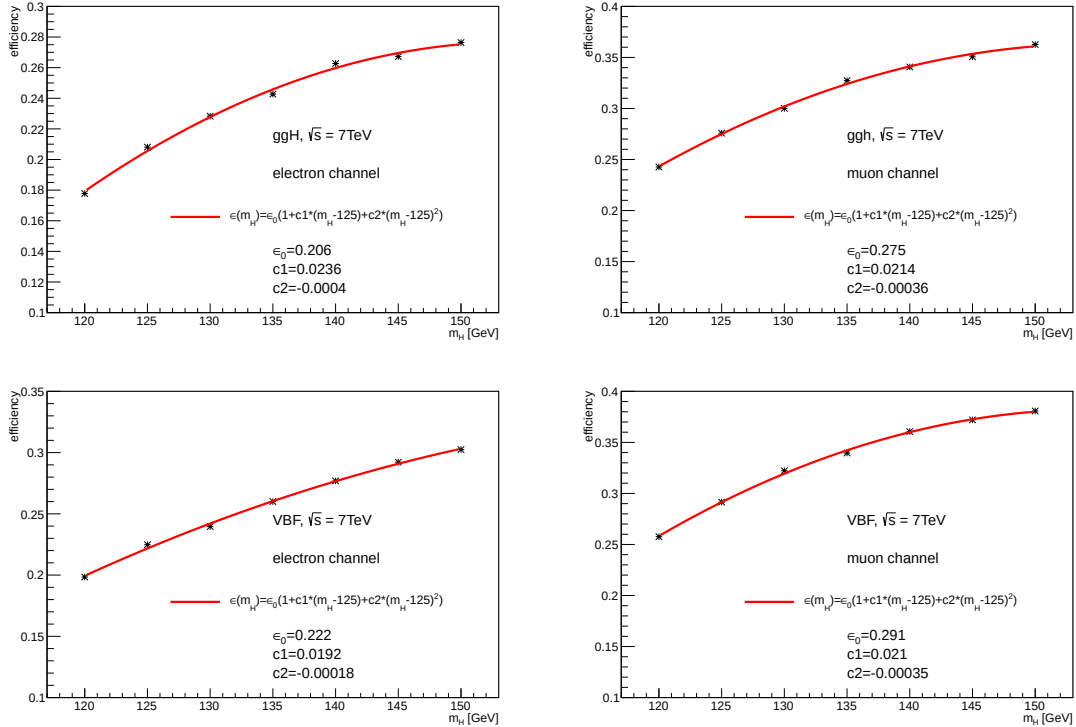


Figure 7.8: Signal efficiency for events produced at $\sqrt{s} = 7$ TeV. The left column is for electron channel. The right column is for muon channel. The top two plots are for ggH process and the bottom two plots are for VBF process.

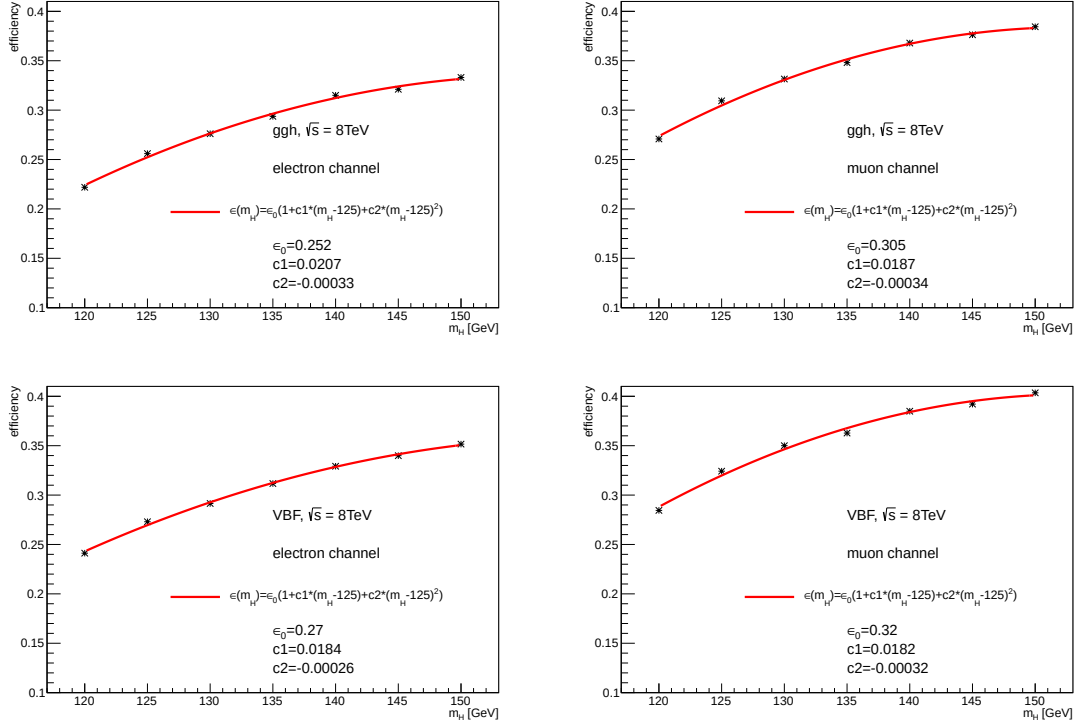


Figure 7.9: Signal efficiency for events produced at $\sqrt{s} = 8$ TeV. The left column is for electron channel. The right column is for muon channel. The top two plots are for ggH process and the bottom two plots are for VBF process.

Table 7.2: Higgs boson production cross section in pb for a SM Higgs boson with $m_H = 120, 125, 130, 135, 140, 145, 150$ GeV for $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV [63], and the branching ratio $\mathcal{B}_{H \rightarrow Z\gamma}$ of a Higgs boson decaying to $Z\gamma$ [63]. In the last column, the product of the total production cross section and the Higgs and Z boson decay branching fractions (in fb) is given.

$\sqrt{s}$ [TeV]	m_H [GeV]	$\mathcal{B}_{H \rightarrow Z\gamma}$ [10^{-3}]	$\mathcal{B}_{Z \rightarrow \ell\ell}$ [%]	σ_{pp} [pb]	σ_{gg} [pb]	σ_{VBF} [pb]	σ_{WH} [pb]	σ_{ZH} [pb]	$\sigma_{t\bar{t}H}$ [pb]	$\sigma(pp \rightarrow H \rightarrow Z(\ell\ell)\gamma)$ [fb]
7	120	1.11	3.366	19.05	16.65	1.28	0.66	0.36	0.10	0.71
	125	1.54	3.366	17.52	15.32	1.22	0.57	0.32	0.09	0.91
	130	1.95	3.366	16.19	14.16	1.17	0.50	0.28	0.08	1.06
	135	2.27	3.366	14.98	13.11	1.12	0.44	0.25	0.07	1.14
	140	2.46	3.366	13.91	12.18	1.07	0.39	0.22	0.06	1.15
	145	2.48	3.366	12.94	11.33	1.02	0.34	0.19	0.05	1.08
	150	2.31	3.366	12.08	10.58	0.98	0.30	0.17	0.05	0.94
8	120	1.11	3.366	24.18	21.13	1.65	0.80	0.45	0.15	0.90
	125	1.54	3.366	22.32	19.52	1.58	0.70	0.39	0.13	1.16
	130	1.95	3.366	20.66	18.07	1.51	0.61	0.35	0.12	1.36
	135	2.27	3.366	19.19	16.79	1.45	0.54	0.31	0.10	1.47
	140	2.46	3.366	17.86	15.63	1.39	0.47	0.27	0.09	1.48
	145	2.48	3.366	16.67	14.59	1.33	0.42	0.24	0.08	1.39
	150	2.31	3.366	15.59	13.65	1.28	0.37	0.22	0.07	1.21

Table 7.3: Number of expected signal events for each production process and Higgs boson masses in 5 GeV steps between 120 and 150 GeV at $\sqrt{s} = 8$ TeV and $Z \rightarrow ee$.

mass [GeV]	$gg \rightarrow H$		VBF		WH		ZH		$t\bar{t}H$		Total
	$\varepsilon[\%]$	N_{evt}	$\varepsilon[\%]$	N_{evt}	$\varepsilon[\%]$	N_{evt}	$\varepsilon[\%]$	N_{evt}	$\varepsilon[\%]$	N_{evt}	N_{evt}
120	22.19	3.63	24.12	0.31	23.15	0.14	23.15	0.08	23.15	0.03	4.18
125	25.60	5.36	27.30	0.46	26.45	0.12	26.45	0.11	26.45	0.04	6.17
130	27.61	6.78	29.15	0.60	28.38	0.24	28.38	0.13	28.38	0.05	7.79
135	29.37	7.80	31.17	0.71	30.27	0.26	30.27	0.15	30.27	0.05	8.96
140	31.50	8.44	32.92	0.78	32.21	0.26	32.21	0.15	32.21	0.05	9.96
145	32.11	8.10	33.99	0.78	33.05	0.24	33.05	0.14	33.05	0.05	9.30
150	33.31	7.32	35.16	0.72	34.24	0.20	34.24	0.12	34.24	0.04	8.41

Table 7.4: Number of expected signal events for each production process and Higgs boson masses in 5 GeV steps between 120 and 150 GeV at $\sqrt{s} = 8$ TeV and $Z \rightarrow \mu\mu$.

mass [GeV]	$gg \rightarrow H$		VBF		WH		ZH		$t\bar{t}H$		Total
	$\varepsilon[\%]$	N_{evt}	$\varepsilon[\%]$	N_{evt}	$\varepsilon[\%]$	N_{evt}	$\varepsilon[\%]$	N_{evt}	$\varepsilon[\%]$	N_{evt}	N_{evt}
120	27.08	4.43	28.45	0.36	27.76	0.17	27.76	0.10	27.76	0.03	5.09
125	30.94	6.48	32.42	0.55	31.68	0.24	31.68	0.13	31.68	0.04	7.45
130	33.15	8.14	35.00	0.72	34.07	0.28	34.07	0.16	34.07	0.05	9.35
135	34.81	9.24	36.28	0.83	35.54	0.30	35.54	0.17	35.54	0.06	10.61
140	36.79	9.86	38.49	0.92	37.64	0.30	37.64	0.18	37.64	0.06	11.31
145	37.63	9.49	39.21	0.90	38.42	0.28	38.42	0.16	38.42	0.06	10.88
150	38.44	8.44	40.35	0.83	39.39	0.23	39.39	0.14	39.39	0.05	9.69

Table 7.5: Number of expected signal events for each production process and Higgs boson masses in 5 GeV steps between 120 and 150 GeV at $\sqrt{s} = 7$ TeV and $Z \rightarrow ee$.

mass [GeV]	$gg \rightarrow H$		VBF		WH		ZH		$t\bar{t}H$		Total
	$\varepsilon[\%]$	N_{evt}	$\varepsilon[\%]$	N_{evt}	$\varepsilon[\%]$	N_{evt}	$\varepsilon[\%]$	N_{evt}	$\varepsilon[\%]$	N_{evt}	N_{evt}
120	17.78	0.51	19.84	0.04	18.81	0.02	18.81	0.01	18.81	0.003	0.59
125	20.81	0.76	22.48	0.07	21.64	0.03	21.64	0.02	21.64	0.004	0.88
130	22.83	0.98	23.96	0.09	23.40	0.04	23.40	0.02	23.40	0.005	1.12
135	24.27	1.12	26.00	0.10	25.14	0.04	25.14	0.02	25.14	0.006	1.29
140	26.27	1.22	27.69	0.11	26.98	0.04	26.98	0.02	26.98	0.006	1.40
145	26.72	1.16	29.22	0.12	27.97	0.04	27.97	0.02	27.97	0.006	1.34
150	27.64	1.05	30.25	0.11	28.95	0.03	28.95	0.02	28.95	0.005	1.21

Table 7.6: Number of expected signal events for each production process and Higgs boson masses in 5 GeV steps between 120 and 150 GeV at $\sqrt{s} = 7$ TeV and $Z \rightarrow \mu\mu$.

mass [GeV]	$gg \rightarrow H$		VBF		WH		ZH		$t\bar{t}H$		Total
	$\varepsilon[\%]$	N_{evt}	$\varepsilon[\%]$	N_{evt}	$\varepsilon[\%]$	N_{evt}	$\varepsilon[\%]$	N_{evt}	$\varepsilon[\%]$	N_{evt}	N_{evt}
120	24.27	0.70	25.76	0.06	25.02	0.03	25.02	0.02	25.02	0.004	0.80
125	27.58	1.00	29.15	0.09	28.36	0.04	28.36	0.02	28.36	0.006	1.16
130	30.01	1.28	32.22	0.11	31.12	0.05	31.12	0.03	31.12	0.007	1.48
135	32.73	1.51	33.96	0.13	33.34	0.05	33.34	0.03	33.34	0.008	1.73
140	34.07	1.58	36.06	0.15	35.06	0.05	35.06	0.03	35.06	0.008	1.82
145	35.06	1.53	37.21	0.15	36.14	0.05	36.14	0.03	36.14	0.008	1.75
150	36.26	1.37	38.08	0.13	37.17	0.04	37.17	0.02	37.17	0.006	1.57

cut	120GeV	125GeV	130GeV
Total	99,998	99,999	99,999
Trigger	30,749	31,470	31,843
Good Photon	12,809	14,717	15,905
$M_Z > 65\text{GeV}$	8,403	9,667	10,401
Trigger Match	8,394	9,655	10,392
Track isolation	8,306	9,524	10,270
Calo isolation	8,229	9,407	10,153
Impact parameter	8,127	9,302	10,011
$M_Z > 81.18\text{GeV}$	7,534	8,646	9,372

Table 7.7: Number of expected signal events left after each selection requirement (without applying any weights) at 120 GeV, 125 GeV and 130 GeV for the 8 TeV electron channel. The cuts are described in Chapter 6 Event selection.

cut	120GeV	125GeV	130GeV
Total	99,968	9,9961	99,966
Trigger	29,267	29,785	29,786
Good Photon	12,274	13,978	15,158
$M_Z > 65\text{GeV}$	9,706	11,150	11,946
Trigger Match	9,704	11,148	11,946
Track isolation	9,589	11,007	11,802
Calo isolation	9,536	10,941	11,734
Impact parameter	9,494	10,880	11,667
$M_Z > 81.18\text{GeV}$	8,960	10,250	11,007

Table 7.8: Number of expected signal events left after each selection requirement (without applying any weights) at 120 GeV, 125 GeV and 130 GeV for the 8 TeV muon channel. The cuts are described in Chapter 6 Event selection.

7.5 Kinematic distributions

The reconstructed photon p_T of a 125 GeV Higgs peaks at 25 GeV (Figure 7.10), so that a 15 GeV p_T cut is applied in order to have a high signal selection efficiency while suppressing jet-faked photon background. The photon p_T distribution for other M_H are shown in Figure 7.10 as well.

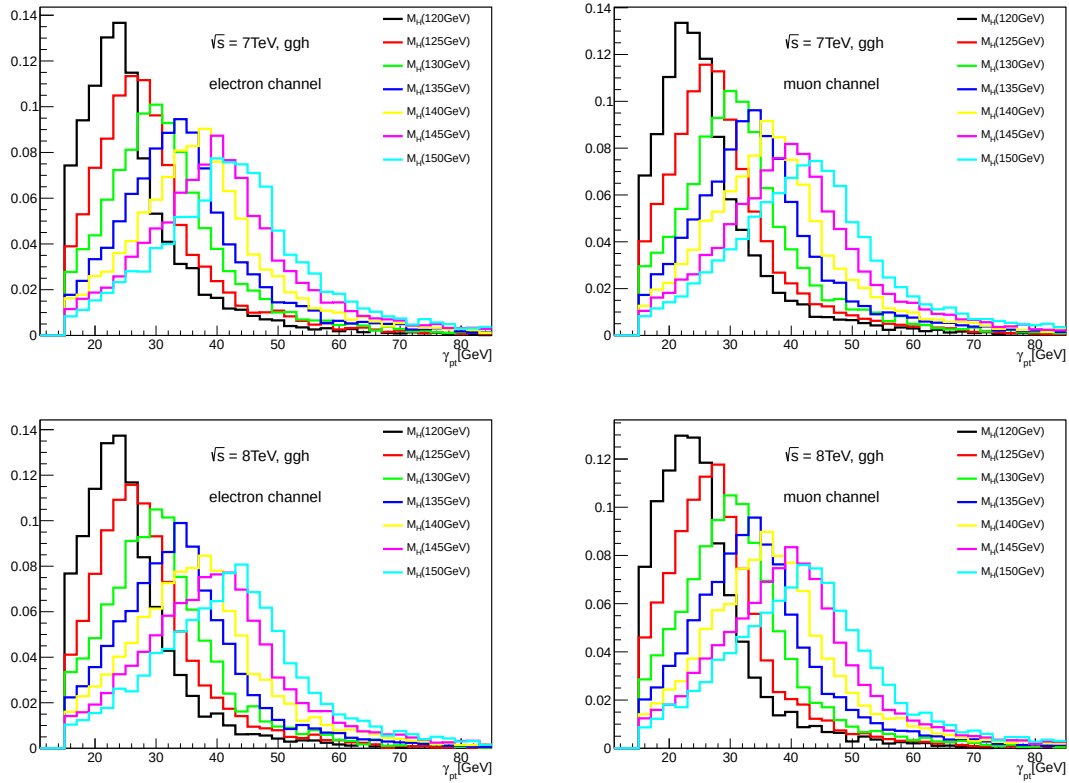


Figure 7.10: Reconstructed signal photon p_T (GeV) at different mass points from 120 GeV to 150 GeV. The black, red, green, blue, yellow, pink and cyan lines correspond to Higgs masses of 120, 125, 130, 135, 140, 145 and 150 GeV respectively.

The reconstructed Higgs mass and $\Delta M = M_{l\gamma} - M_l$ for $M_H = 125 \text{ GeV}$ are shown in Figure 7.11 and Figure 7.12 respectively. The ΔM variable is eventually used as the discriminating variable because it is to a large extent insensitive to the contribution from the FSR of $H \rightarrow \mu\mu$ (The FSR of $H \rightarrow ee$ in the SM is expected to be negligible due to the

small electron mass). This could not be distinguished from $H \rightarrow Z\gamma$ if $M_{ll\gamma}$ were chosen as the discriminating variable, as $M_{ll\gamma}$ peaks at the Higgs boson mass M_H for both $H \rightarrow Z\gamma$ and $H \rightarrow \mu\mu^* \rightarrow \mu\mu\gamma$. On the other hand, the ΔM distribution does not produce a peak of $H \rightarrow \mu\mu^* \rightarrow \mu\mu\gamma$ and therefore its contamination does not bias the fitted $H \rightarrow Z\gamma$ yield.

The width of the Higgs mass is dominated by detector resolution as the intrinsic width of the Higgs at $M_H=125$ GeV is about 0.03 GeV. The intrinsic width of Higgs as a function of M_H is shown Figure 7.13. The width of the ΔM is dominated by the Z boson intrinsic width.

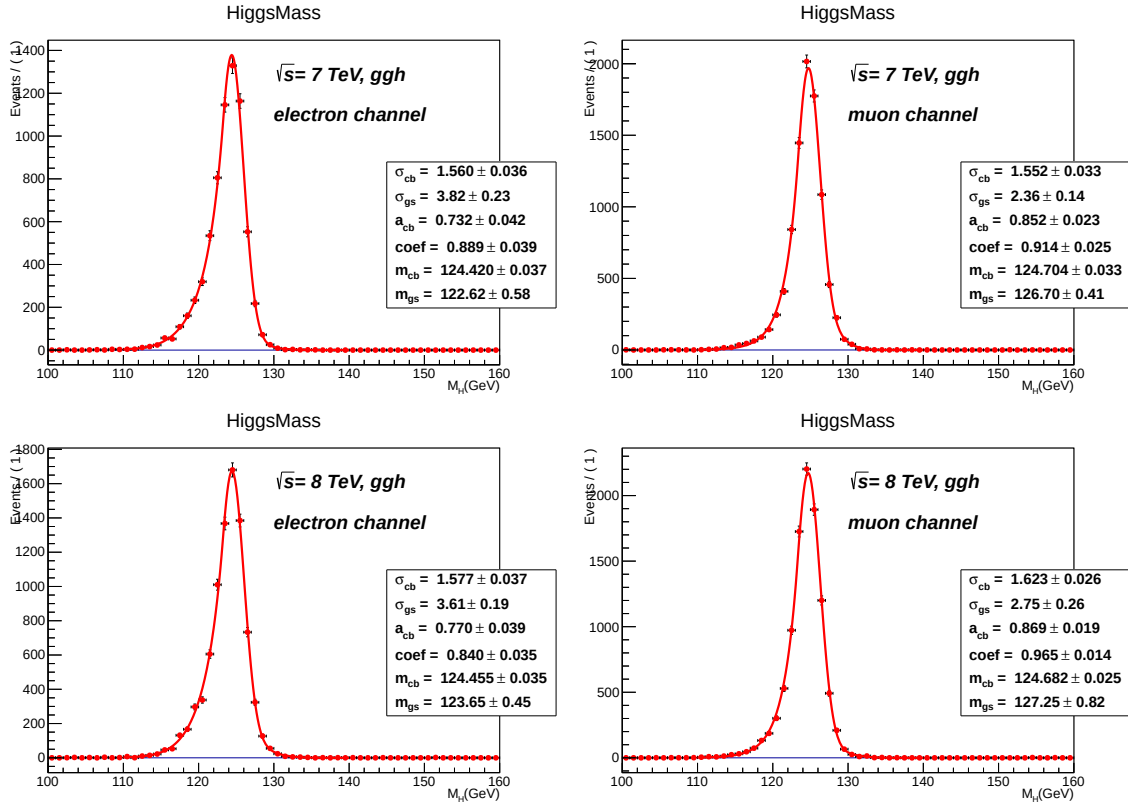


Figure 7.11: Reconstructed Higgs mass at mass point $M_H=125$ GeV. The m_{cb} , σ_{cb} , a_{cb} are parameters of the CB function and refer to the mean, the width and a parameter for the CB tail respectively. The m_{gs} and σ_{gs} are parameters for Gaussian function and refer to the mean and the width respectively.

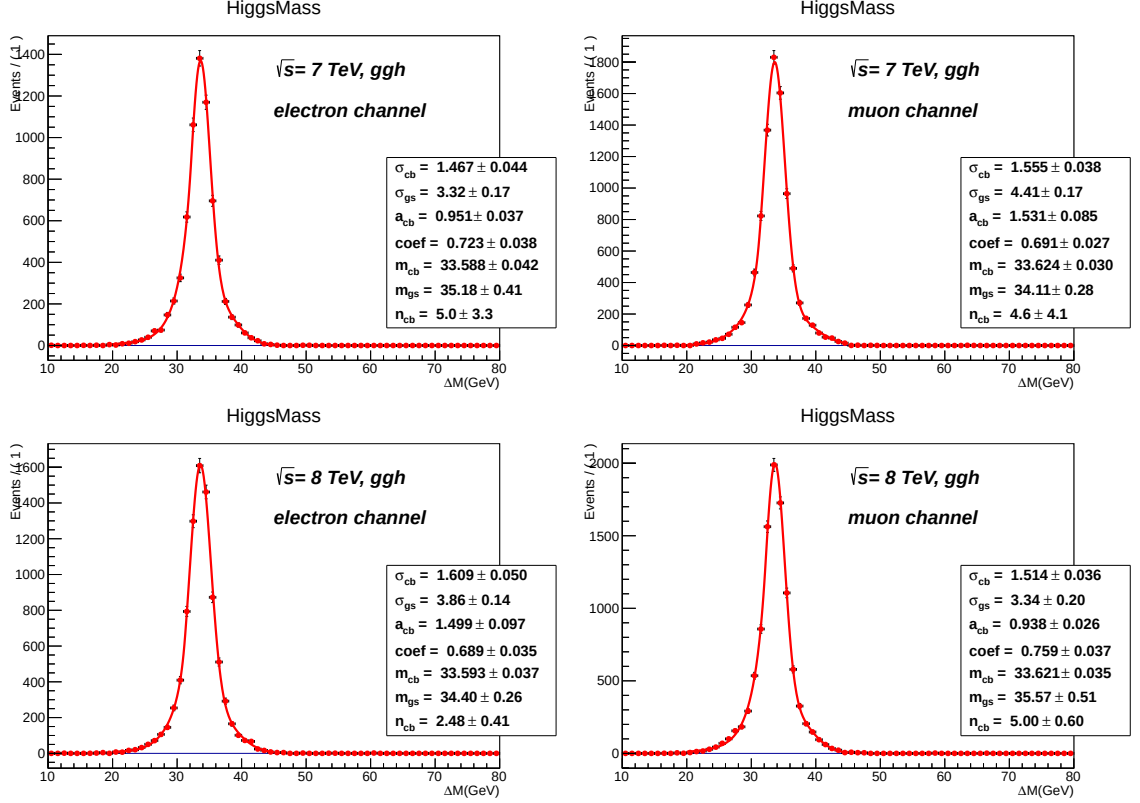


Figure 7.12: Reconstructed signal $\Delta M = M_H - M_Z$ at mass point $M_H = 125$ GeV. The m_{cb} , σ_{cb} , a_{cb} are parameters of the CB function and refer to the mean, the width and a parameter for the CB tail respectively. The m_{gs} and σ_{gs} are parameters for Gaussian function and refer to the mean and the width respectively.

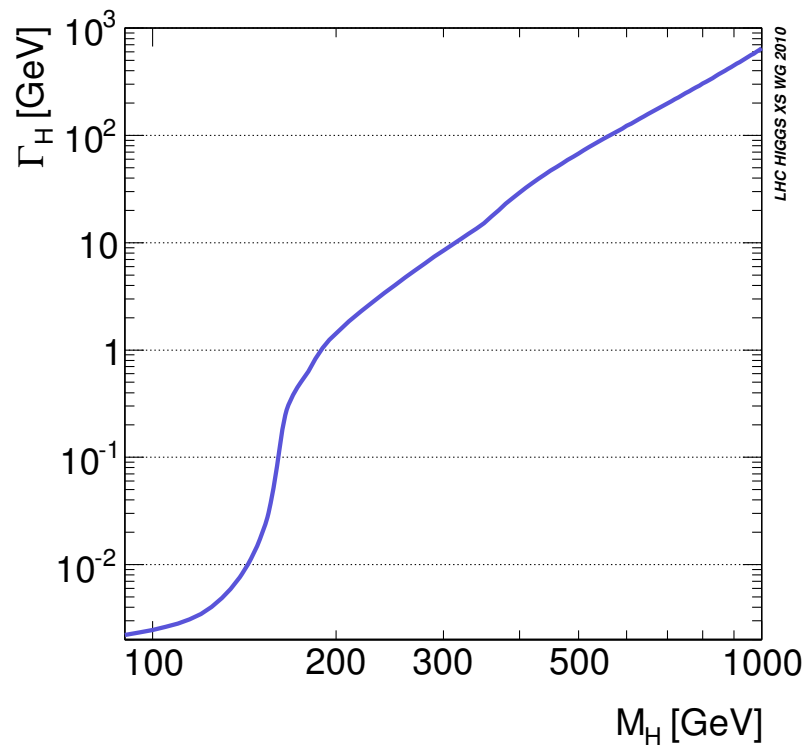


Figure 7.13: The intrinsic width of Higgs as a function of M_H .

Chapter 8

Background estimation

This chapter describes the background processes that contribute to the final state $l\bar{l}\gamma$. A general introduction is given in section 8.1. SM $Z\gamma$ process is described in section 8.2. The Z +jets background and the estimation methodology is presented in section 8.3. The remaining backgrounds are introduced in section 8.4. A summary of the components of the background and a comparison between MC prediction and data are shown in section 8.5 and section 8.6 respectively. Finally, a study of the background modeling is presented in section 8.7.

8.1 Introduction

Since our signal consists of two same-flavor, opposite-charge leptons plus a photon, the main background processes that contribute to this final state are Standard Model (SM) $Z\gamma$, Z +jets and $t\bar{t}$. The dominant background is due to SM $Z\gamma$ events which are indistinguishable from the signal events. The shape of this background is estimated using Monte Carlo events with the normalization calculated from a data driven method. The Z +jets events form the second largest background, where some jets are misidentified as photons. The remaining backgrounds are $t\bar{t} \rightarrow l^+l^- + jets + X$ where some jets are misidentified as photons and WZ leptonic-decay events where one lepton was either misidentified as a photon or was not in the phase-space we were looking at. $t\bar{t}$ has a small contribution and WZ events turned out to be negligible. In this chapter, the method of estimating the reducible

backgrounds is explained.

8.2 Standard Model $Z + \gamma$ Background

The Standard Model (SM) $Z\gamma$ process is the main background. The Feynman diagram for the SM $Z\gamma$ production processes are shown in Figure 8.1. The signal $H \rightarrow Z\gamma$ is small

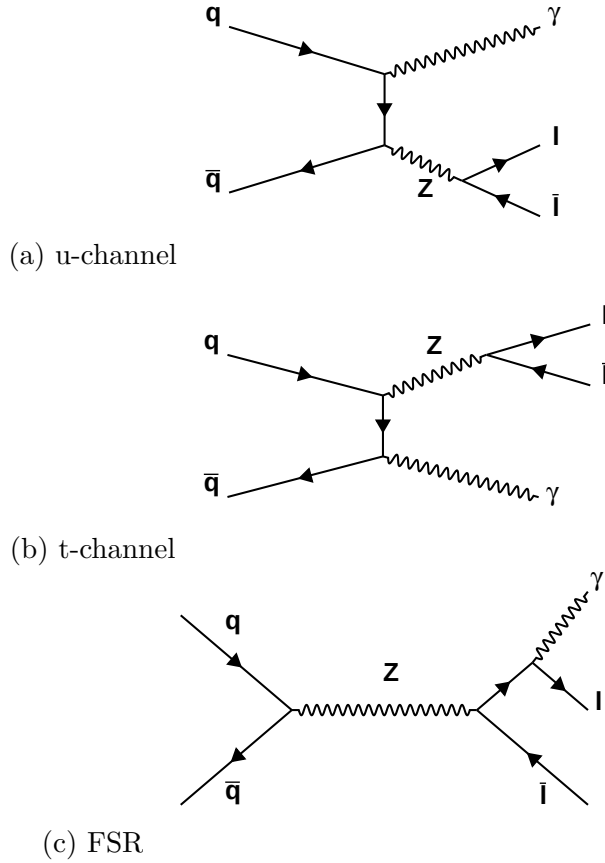


Figure 8.1: Feynman diagrams for $Z\gamma$ production in (a) u-channel (b) t-channel and (c) final state photon radiation (FSR) from the Z boson decay process.

compared to the SM $Z\gamma$ background, which is shown in Figure 8.2.

There is no kinematic variable to effectively distinguish $H \rightarrow Z\gamma$ from SM $Z\gamma$. The distribution of a few kinematic variables (photon p_T , Z p_T , the distance between Z and photon in $\eta - \phi$ plane and the Φ difference between Z and photon) are shown in Figure 8.3.

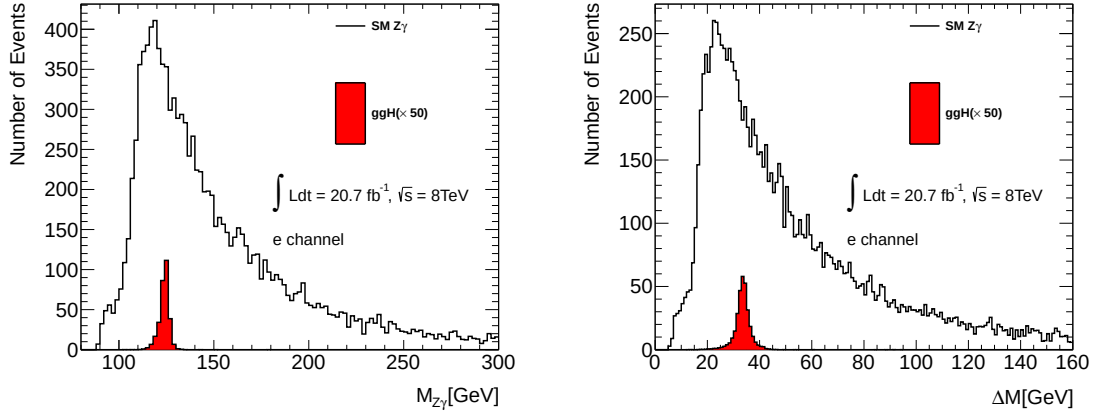


Figure 8.2: The reconstructed $Z\gamma$ mass (left) and ΔM (right) for SM $Z\gamma$ and signal ggH at 125 GeV using MC samples. The signal $\times 50$ is shown as the red histogram.

The Monte Carlo sample of SM $Z\gamma$ is used to get these distributions because there is no control region to use since the final states are the same. We did not use higher photon p_T cut because it moves the SM $Z\gamma$ turn-on peak towards the signal peak, which makes the background shape modeling difficult. No cut on $\Delta\Phi$ is used because that will cause a significant lost in signal as well. But this is a potential variable to use to divide the samples into more categories in order to enhance the signal over background ratio in a future analysis.

8.3 Z +jets Background

The Z +jet events form the second largest background, accounting for about 15% of the observed background near the signal region.

In this analysis, a data driven method has been used to study the yield and shape of Z +jets background.

8.3.1 2-D sideband method

To study the Z +jets background, 2-D sideband method (or ABCD method) is used, shown in Figure 8.4. The idea of this method is to use 2 independent observables to define 3 control

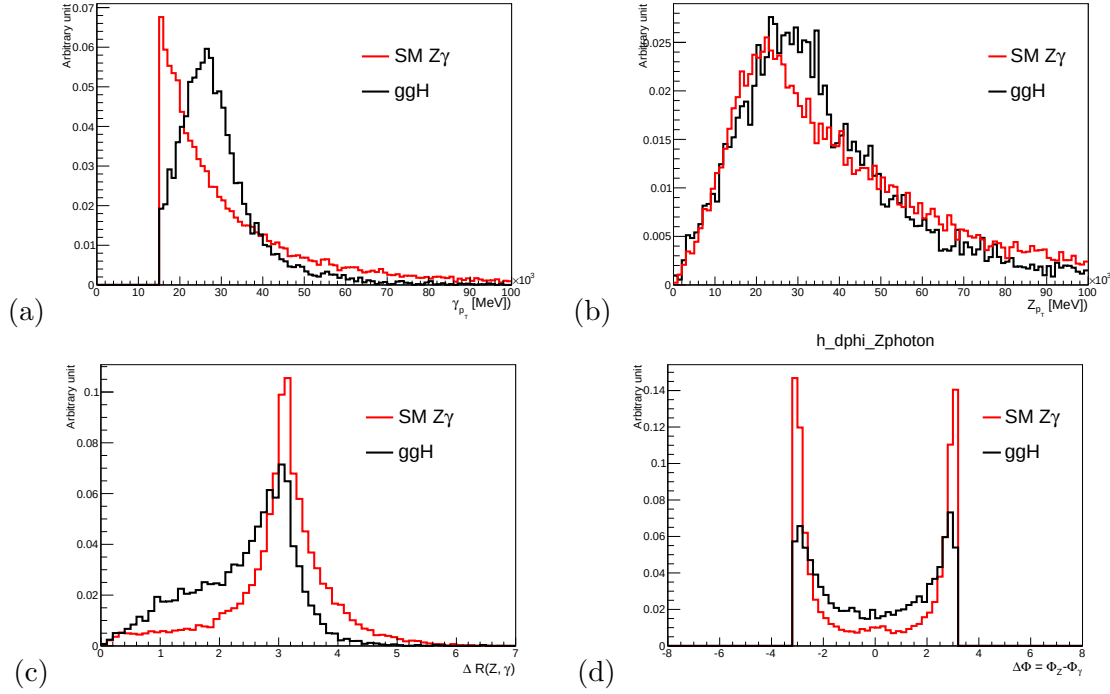


Figure 8.3: The kinematic variable distributions for $H \rightarrow Z\gamma$ (in red) and SM $Z\gamma$ events (in black) for: (a) photon p_T (b) Z_{pT} (c) $\Delta R(Z, \gamma)$ (d) $\Delta\Phi$. The red and black line are the SM $Z\gamma$ background and Higgs signal events respectively. Neither of the variables is observed to be able to distinguish the two processes without losing the signal. The two processes share the same phase space but the cross section for SM $Z\gamma$ is much larger than that for $H \rightarrow Z\gamma$.

regions (B, C, D) and one signal region (A). Because of the independence, $\frac{N_C}{N_A} = \frac{N_D}{N_B}$. So we have $N_A = \frac{N_B N_C}{N_D}$, where N_A, N_B, N_C, N_D represents the number of Z +jets events in signal region A and three control regions B, C, D respectively. However, in reality, the two variables may not be completely independent. A correlation factor $R = \frac{N_B N_C}{N_A N_D}$ has been calculated based on Z +jets Monte Carlo samples. The difference between 1 and R will be used as one source of systematics of Z +jets background estimation.

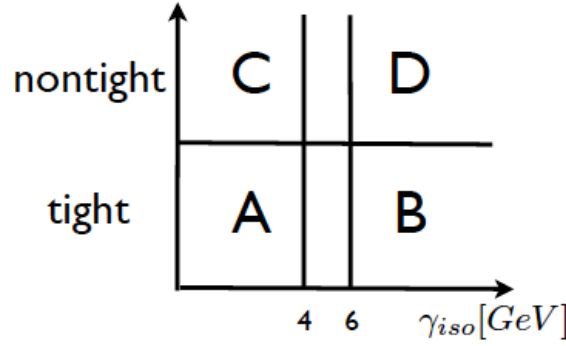


Figure 8.4: 2-D sideband method (ABCD). Region A is the signal region. Regions B, C, D are control regions. The kinematic distribution of Z +jets in signal region A is extracted from control region C. N_B to N_D is assumed to equal the ratio of N_A (without any signal events) to N_C and is thus used to calculate the background in the signal region A. The signal region is defined to have photon isolation (γ_{iso}) less than 4 GeV, and the B and D regions have photon isolation greater than 6 GeV.

8.3.2 Control Region Definition

Photon isolation and identification tightness are the two observables used for the 2-D sideband method. Figure 8.5 shows the distribution of the photon isolation variable for the electron and muon channels. The data points are well fit (both in normalization and shape) by the backgrounds calculated using the 2-D sideband method.

Recall that there are 5 shower shape variables used to characterize electromagnetic showers in the first layer of the EM Calorimeter, which is shown in Figure 4.7, namely $w_{s3}, w_{stot}, F_{side}, \Delta E, E_{ratio}$ (refer to Figure 4.7). These are referred to as “5 bits” in the

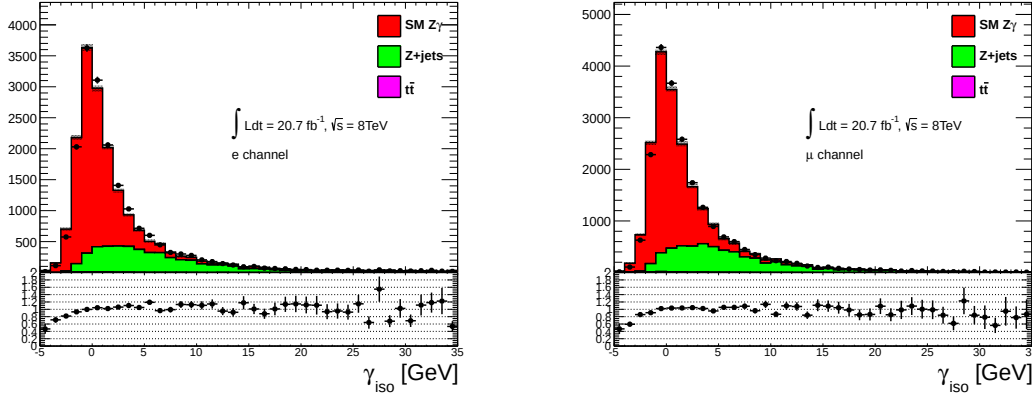


Figure 8.5: Photon isolation for electron channel (left) and muon channel (right). Both the shape and the normalization of photon isolation for Z +jets are from control regions. The bottom of the two plots shows the data to mc ratio.

following description. Similarly “4 bits” refers to all but the w_{stot} variable which is found to be correlated with the isolation variable.

The non-tight control regions (C and D) are defined as requiring that at least 2 of 4 EM Calorimeter first layer shower shape requirements fail (i.e. 2 of the 4 bits fail), but that all the calorimeter middle layer shower shape cuts pass and that the hadronic leakage cut also passes. In order to arrive at that optimal definition of “non-tight”, four cases were considered: 1. fail at least 1 out of “5 bits” ; 2. fail at least 2 out of “5 bits” ; 3. fail at least 1 out of “4 bits” ; 4. fail at least 2 out of “4 bits”. The best two choices (with ratio $R = \frac{N_B N_C}{N_A N_D}$ closer to 1) require failing at least 1 or 2 of “4 bits”. The other two cases show a stronger correlation with the ABCD regions, therefore they are not used. This is shown in Figure 8.6. The reason for defining the non-tight control region to have failed at least two of “4 bits” instead of just one of the “4 bits” is to reduce the signal leakage into the control regions. Since we want to use the Z +jets dominated control region to estimate the Z +jets events in signal region, failing at least two of the “4 bits” requirement guarantees a nearly pure Z +jets sample in the control region with a $Z + \gamma$ contamination of less than 10% of the total number of events.

The non-tight definition study using Z +jets Monte Carlo samples is shown in Table 8.1

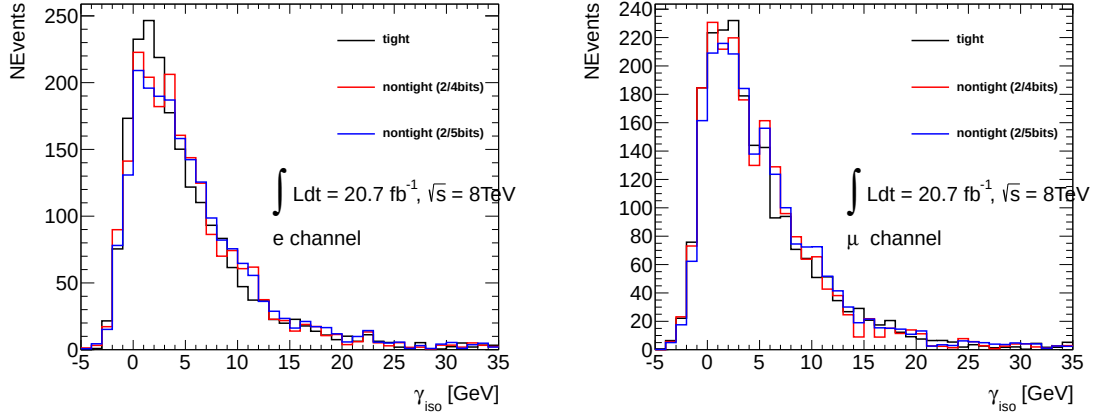


Figure 8.6: The photon isolation distribution for tight and different nontight definitions for electron (left) and muon (right) channel using simulated Z +jets sample. The photon isolation in tight, non-tight (fail at least 2 out of “4 bits”) and non-tight (fail at least 2 out of “5 bits”) are shown in black, red and blue respectively. The red curve is closer to the black curve than the blue curve.

for electron channel and in Table 8.2 for muon channel.

non-tight def	A	B	C	D	R
fail at least 1/5	3074.9 ± 86.5	1821.3 ± 66.9	5713.6 ± 117.9	4211 ± 101.7	0.80 ± 0.05
fail at least 2/5	3074.9 ± 86.5	1821.3 ± 66.9	4044.5 ± 99.2	3150.0 ± 87.9	0.76 ± 0.05
fail at least 1/4	3074.9 ± 86.5	1821.3 ± 66.9	4034.0 ± 99.1	2651.33 ± 80.7	0.90 ± 0.05
fail at least 2/4	3074.9 ± 86.5	1821.3 ± 66.9	2418.0 ± 76.7	1629.04 ± 63.2	0.88 ± 0.06

Table 8.1: non-tight definition study for electron channel 2012. The numbers in the table are the number of Z +jets events in each region. The R is defined as $R = \frac{BC}{AD}$; therefore values closer to one correspond to smaller correlations and hence indicate the optimal definition of non-tight events.

The non-isolation control region is defined as requiring the photon isolation variable $topoEtcone40 > 6$ GeV. In summary, four regions are:

1. Tight and Isolated region (signal region A) : candidates with photon isolation < 4

non-tight def	A	B	C	D	R
fail at least 1/5	3562.9 ± 93.1	2121.5 ± 72.2	6863.8 ± 129.1	4866.4 ± 109.6	0.84 ± 0.04
fail at least 2/5	3562.9 ± 93.1	2121.5 ± 72.2	5069.7 ± 110.9	3690.9 ± 95.4	0.82 ± 0.04
fail at least 1/4	3562.9 ± 93.1	2121.5 ± 72.2	4879.9 ± 108.9	3012.4 ± 86.2	0.96 ± 0.05
fail at least 2/4	3562.9 ± 93.1	2121.5 ± 72.2	3150.4 ± 87.5	1911.0 ± 68.6	0.98 ± 0.06

Table 8.2: non-tight definition study for muon channel 2012. The numbers in the table are the number of Z +jets events in each region. The R is defined as $R = \frac{BC}{AD}$; therefore values closer to one correspond to smaller correlations and hence indicate the optimal definition of non-tight events.

GeV and photons that pass the tight id cut since we are looking for well isolated photon with good quality.

2. Tight and Non-isolated region (control region B) : candidates with photon isolation > 6 GeV and photons that pass the tight id cut.
3. Non-tight and Isolated region (control region C) : candidates with photon isolation < 4 GeV and photons that pass the non-tight id cut.
4. Non-tight and Non-isolated region (control region D) : candidates with photon isolation > 6 GeV and photons that pass the non-tight id cut.

The ABCD method is based on two assumptions:

1. The two variables used (γ isolation and tightness) are independent. So the ratio N_C/N_A is equal to N_D/N_B .
2. The signal leakage to the control regions is negligible. So we consider all the events in the B, C, D control regions are pure background events (Z +jets).

In fact, the two assumptions are not quite correct since a 10% correlation is found between control regions and therefore a non-negligible fraction of signal events are found in the control regions. In order to correct for this contamination of the control region, the

formula has been extended from $N_A^{zjet} = \frac{N_B^{zjet} N_C^{zjet}}{N_D^{zjet}}$ to

$$N_A - N_A^{sig} - N_A^{EW} = \frac{1}{R} \frac{(N_B - C_b N_A^{sig} - N_B^{EW})(N_C - C_c N_A^{sig} - N_C^{EW})}{N_D - C_d N_A^{sig} - N_D^{EW}} \quad (8.1)$$

where

- N_A, N_B, N_C, N_D are the observed data events in A, B, C, D region.
- N_A^{sig} is the number of signal events in signal region A.
- $N_A^{EW}, N_B^{EW}, N_C^{EW}, N_D^{EW}$ are the contamination of $t\bar{t}$ events.
- C_b, C_c, C_d are the signal leakage ratio, defined as the number of SM $Z\gamma$ events in regions B, C, D region to the number of SM $Z\gamma$ events in region A respectively.
- $R = \frac{BC}{AD}$ is the correlation factor between A, B, C, D regions.

By solving this equation, we get

$$N_A^{sig} = \frac{E[-1 + \sqrt{1 + F}]}{G} \quad (8.2)$$

Hence,

$$N_A^{Zjets} = N_A - N_A^{EW} - N_A^{sig} \quad (8.3)$$

where

- $E = ((N_C - N_C^{EW} C_b + (N_B - N_B^{EW}) C_c - R(N_A - N_A^{EW}) C_d - R(N_D - N_D^{EW}))$
- $F = \frac{4(C_b C_c - C_d)[(N_A - N_A^{EW})(N_D - N_D^{EW})R - (N_B - N_B^{EW})(N_C - N_C^{EW})]}{(N_C - N_C^{EW}) C_b + (N_B - N_B^{EW}) C_c - R(N_A - N_A^{EW}) C_d - R(N_D - N_D^{EW})}$
- $G = 2(RC_d - C_b C_c)$
- $R = \frac{N_B^{Zjets} N_C^{Zjets}}{N_A^{Zjets} N_D^{Zjets}}$

8.3.3 Systematics

8.3.3.1 Due to Non-isolation region selection

Since γ isolation is a continuous variable, different non-isolated definitions used will result in slightly different Z +jets estimation. The difference partially comes from limited statistics. A systematic uncertainty is assigned by ± 1 GeV variation of γ isolation (1 GeV is about the one standard deviation of the photon isolation distribution.).

		B	C	D	N_{zjets}
2012	Electron				
	nominal	0.1179 ± 0.0026	0.0151 ± 0.0009	0.0035 ± 0.0004	2099 ± 117
	noniso>5GeV	0.1333 ± 0.0028	0.0151 ± 0.0009	0.0041 ± 0.0005	2203 ± 114
	noniso>7GeV	0.1058 ± 0.0025	0.0151 ± 0.0009	0.0033 ± 0.0004	2121 ± 129
	nontight(fail $\geq$ 1bit)	0.1179 ± 0.0026	0.0802 ± 0.0021	0.0125 ± 0.0008	2199 ± 120
	Muon				
	nominal	0.0725 ± 0.0018	0.0165 ± 0.0008	0.0026 ± 0.0003	2564 ± 116
	noniso>5GeV	0.0885 ± 0.0020	0.0165 ± 0.0008	0.0032 ± 0.0004	2618 ± 116
2011	noniso>7GeV	0.0605 ± 0.0016	0.0165 ± 0.0008	0.0023 ± 0.0003	2536 ± 123
	nontight(fail $\geq$ 1bit)	0.0725 ± 0.0018	0.0845 ± 0.0019	0.0081 ± 0.0006	2579 ± 112
	Electron				
	nominal	0.1179 ± 0.0026	0.0151 ± 0.0009	0.0035 ± 0.0004	282 ± 48
	noniso>5GeV	0.1050 ± 0.0049	0.0188 ± 0.0020	0.0030 ± 0.0008	288 ± 45
	noniso>7GeV	0.0884 ± 0.0044	0.0188 ± 0.0020	0.0027 ± 0.0007	287 ± 55
	nontight(fail $\geq$ 1bit)	0.0965 ± 0.0047	0.0905 ± 0.0045	0.0110 ± 0.0015	309 ± 49
	Muon				
	nominal	0.0725 ± 0.0018	0.0165 ± 0.0008	0.0026 ± 0.0003	349 ± 46
	noniso>5GeV	0.0648 ± 0.0031	0.0205 ± 0.0017	0.0027 ± 0.0006	329 ± 41
	noniso>7GeV	0.0422 ± 0.0024	0.0205 ± 0.0017	0.0019 ± 0.0005	352 ± 49
	nontight(fail $\geq$ 1bit)	0.0504 ± 0.0027	0.0899 ± 0.0036	0.0076 ± 0.0010	366 ± 44

Table 8.3: The number of Z +jets events (N_{zjets}) for different definitions of the control regions together with the fractional amount of signal leakage into the B, C, and D regions shown in the respective columns.

8.3.3.2 Due to Non-tight region definition

There are several ways to define non-tight region. The definition used in this analysis is based on a study of the correlation of between control regions (R-ratio) using MC Z +jets sample, provides the lowest correlation. However, the simulations of shower shape variables are not perfect. It is worth looking at the Z +jets estimation using a different definition of non-tight (fail at least one of the “4 bits”) and assigning the difference as one source of systematic uncertainty.

The Z +jets is evaluated under different non-isolation and non-tight definition and shown in Table 8.3.

8.3.3.3 Due to Correlation between control regions

One of the main assumptions of this 2D-sideband method is that the two variables used to define control regions are independent. However, in reality, a non-negligible correlation is found by using Z +jets MC samples. It may due to a correlation between the two variables from the experimental measurement and/or due to Z +jets being composed of Z +1jet/2jets/3jets ... Even though for each component the two variables may be independent when summed over a number of jets, a correlation may be introduced since the ratio of the probability of being non-tight and tight are sensitive to the effects of higher order terms. Overall, this uncertainty is taken into account by using a correlation factor in Table 8.4.

Table 8.4: ABCD Correlation

	2012 e	2012 μ	2011 e	2011 μ
R	0.88 ± 0.06	0.98 ± 0.06	0.97 ± 0.07	1.05 ± 0.07

8.4 Other backgrounds

8.4.1 $t\bar{t}$ and WZ Background

The $t\bar{t}$ background contributes only 0.5% to the total observation. Since this is a small contribution, its normalization and shape both come from MC simulation rather than from

data driven estimates. The distributions of $M_{Z\gamma}$ and photon p_T for $t\bar{t}$ events in electron channel are shown in Figure 8.7. The WZ background is much smaller than $t\bar{t}$ background (only 10% of $t\bar{t}$ events, estimated from Monte Carlo sample), so it was neglected in this analysis.

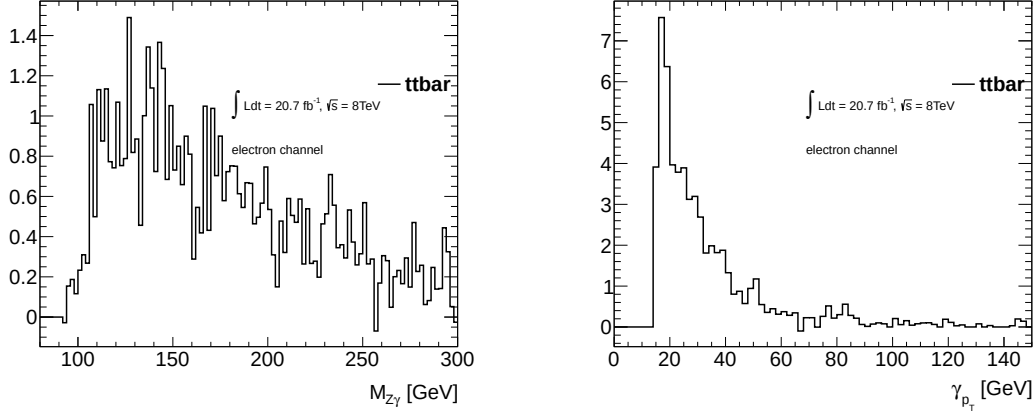


Figure 8.7: The distributions of $M_{Z\gamma}$ (left) and photon p_T (right) for $t\bar{t}$ events for $\sqrt{s}=8$ TeV in electron channel .

8.5 Summary for background estimation

The observed data and the background components ($Z\gamma$, Z +jets, $t\bar{t}$) are listed in Table 8.5. The systematic error for Z +jets is calculated by adding the systematic errors in quadrature. For the case where there are two different systematic error estimates ($\gamma_{iso} > 5$ GeV and $\gamma_{iso} > 7$ GeV) the largest estimate is used.

8.6 The comparison between MC simulation and data

The good agreement between the Monte Carlo simulation and the data can be seen in the plots of the $Z\gamma$ reconstructed mass (Figure 8.8) and $\Delta M = M_{l\bar{l}\gamma} - M_{l\bar{l}}$ (Figure 8.9), where the lower plots show the ratio of the data to the Monte Carlo simulation results. Additional comparisons of the data with the Monte Carlo are shown in Appendix B. Overall the data is well modeled by the Monte Carlo.

	2012		2011	
sources	$Z(ee)\gamma$	$Z(\mu\mu)\gamma$	$Z(ee)\gamma$	$Z(\mu\mu)\gamma$
Observed Events	13966	16460	1909	2620
SM $Z\gamma$	11807 ± 166	13822 ± 173	1628 ± 65	2264 ± 69
Z +jets	$2099 \pm 117 \pm 320$	$2564 \pm 116 \pm 77$	$282 \pm 48 \pm 29$	$349 \pm 46 \pm 31$
$t\bar{t}$	60 ± 2	74 ± 2	6 ± 1	7 ± 1
Purity($=\frac{N_{Z\gamma}}{N_{total}}$)	84.5%	84.0%	85.2%	86.3%

Table 8.5: A summary of the observed data and each background component. For Z +jets, the first error is the statistical error due to the size of the Monte Carlo statistical sample and the second error is the estimated systematic error as described in the text. The error for SM $Z\gamma$ and $t\bar{t}$ are statistical errors only. The systematic error for $Z\gamma$ is fully correlated with that for Z +jets.

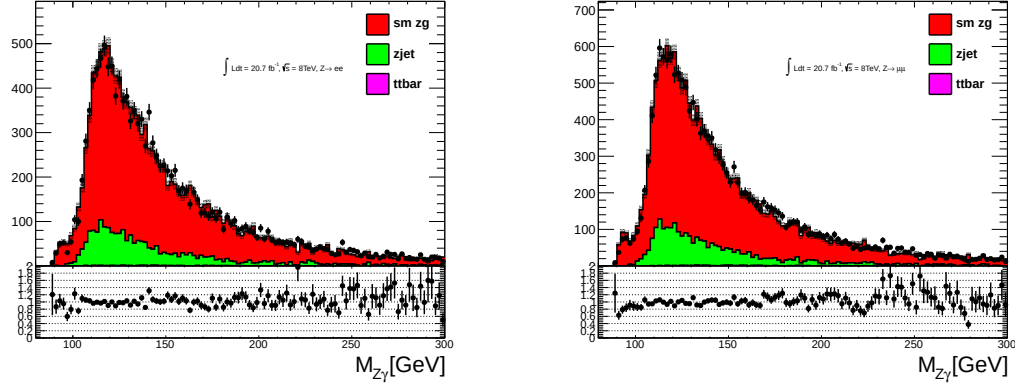


Figure 8.8: Reconstructed $Z\gamma$ mass in the electron channel (left) and the muon channel (right). The sub-plot shows the ratio $= \frac{N_{data}}{N_{MC}}$.

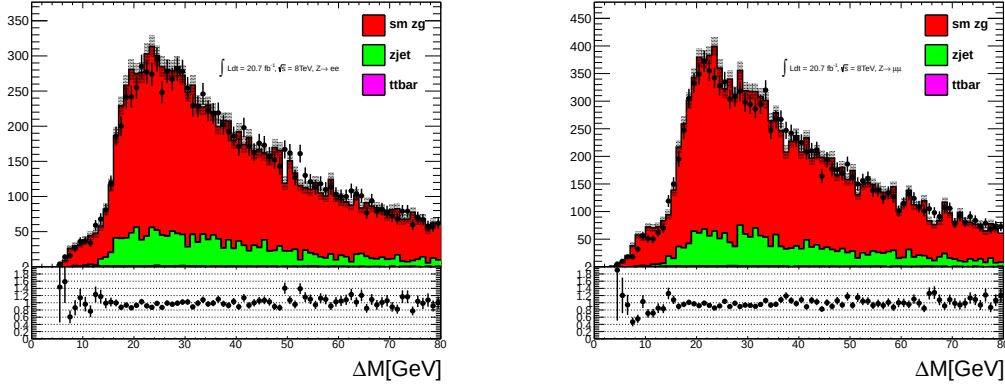


Figure 8.9: Reconstructed mass difference of $\Delta M = M_{Z\gamma} - M_Z$ in electron channel (left) and muon channel (right).

8.7 Background Modeling

Since we are searching for a small signal over a large background, in order to avoid the limited statistics of the Monte Carlo samples with the full detector simulation and both the theoretical and the experimental uncertainties associated with the background Monte Carlo samples, we use the analytical functional form to describe the background shape. Several functions have been studied, shown in Figure 8.10, to model the background shape including a Chebyshev function family [98], an exponential of a second order polynomial and a double exponential function. Initially, the template from the MC samples with the full detector simulation was used to generate data-like pseudoexperiments for testing the performance of various background models. The advantage of this method is that the model is tested on the realistic background shape. However, due to the limited statistics of the MC samples, it yields significant fluctuations that can lead to overestimate the expected bias. Therefore a high statistics (10 million events) generator-level¹ $Z + \gamma$ sample was used in this study. Due to the CPU-time constraints, huge samples of fully simulated $Z + \gamma$ and $Z + \text{jets}$ are not available. The $Z + \gamma$ is the dominant background, accounting for about 85% of the total background and the same functional form can describe both the $Z + \gamma$ and $Z + \text{jets}$

¹Monte Carlo events without detector simulation

reconstructed ΔM shape. The detector effect should only smear the generator-level shape but should not affect dramatically the overall shape of the background distribution. In the future, with more computing power, this study can be improved by the fully simulated background samples with high statistics. In the model selection, the goal is to find a function that fits the data well, minimizes the possibility of creating spurious signals, yet is not so flexible that it would subsume any potential signal within the background shape. The spurious signal is the fake signal generated when fitting background only events with signal + background model. We use a figure of merit in evaluating the potential background functions for which the size of the spurious signal is less than 20% of the statistical uncertainty on the fitted signal yield [1]. Figure 8.11 shows that the fourth and third order Chebyshev Polynomials provide the best performance except for the bins of 143-145 GeV where the fraction is between 20% and 30%. Although an 4th order Chebyshev Polynomial gives smaller fitting error, it provides a worse limit because it is too flexible to be sensitive to signal. The exponential of a second order polynomial and the double exponential functions show larger spurious signals. Therefore a 3rd order Chebyshev Polynomial is used to model the four categories of ΔM distribution within 24 - 64 GeV region. A description of the functions used in this study is given in Table 8.6.

Function	Description	Function form
Exp(pol2)	exponential of a 2^{nd} order polynomial	$e^{a_1(\Delta M-24)+a_2(\Delta M-24)^2}$
Cheb3	3^{rd} order Chebyshev polynomial	$T_0(x) = 1, T_1(x) = x, T_2(x) = 2x^2 - 1, T_3(x) = 4x^3 - 3x, T_4(x) = 8x^4 - 8x^2 + 1$
Cheb4	4^{th} order Chebyshev polynomial	

Table 8.6: The detailed description of the functions tested in the background modeling.

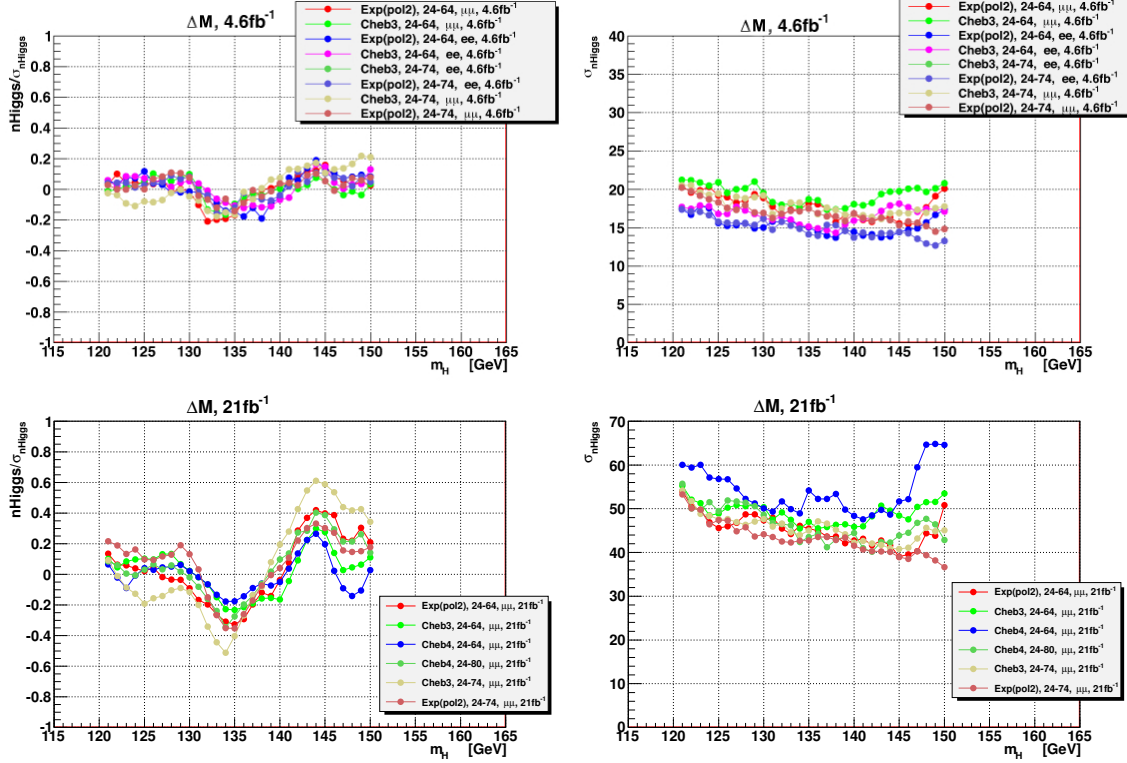


Figure 8.10: Spurious signal as a function of the Higgs boson mass, fitting the ΔM distribution with different models and different ranges. Pseudo-data generated from a high-statistics generator-level $Z + \gamma$ Sherpa sample. The top two plots and bottom plots correspond to 7 TeV and 8 TeV study respectively. The left two plots are the spurious signal divided by the uncertainty on the fitted signal yield. The right two plots are the statistical uncertainty (from the fit) on the signal yield.

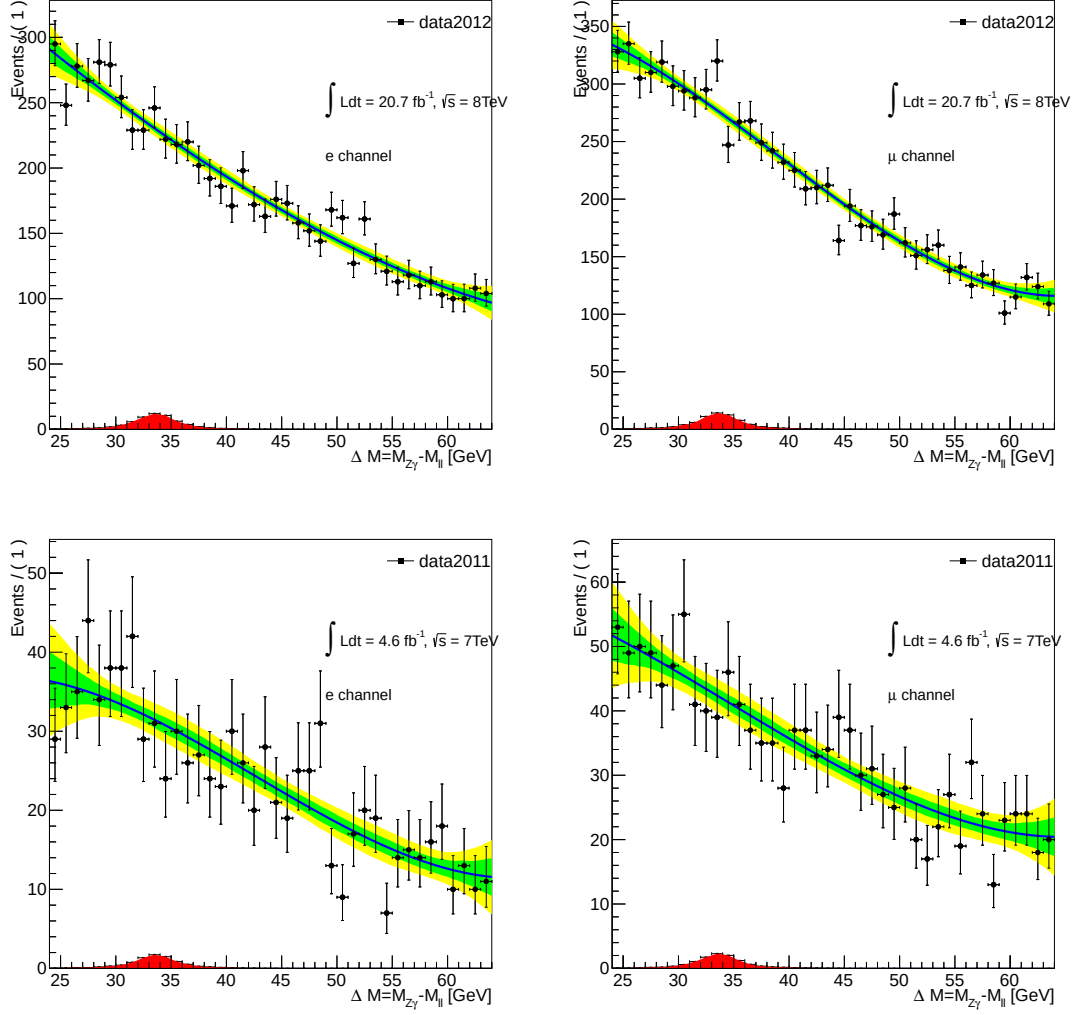


Figure 8.11: Background-only model fitting on data in 2011 and 2012. Green band and yellow band correspond to 1σ and 2σ error bands respectively. The red histogram shows expected $H \rightarrow Z\gamma$ signal $\times 10$.

Chapter 9

Statistical analysis

This chapter introduces the statistical method used to quantify the null search result. From Chapter 8, we did not see significant evidence of a signal in the data. Hence we will establish an upper limit on the signals in this chapter. In order to do the hypothesis test, we need to summarize the data-set into one value which is called the “test statistic” in the statistical language. With the test statistic and its distribution, we can calculate the probability of obtaining a test statistic at least as extreme as the one that was actually observed (refers to the p value) and hence the upper limit can be calculated at a certain confidence level. An alternative p value which is used by the ATLAS community is the CLs [104] which is able to mitigate the false exclusion. This is described in section 9.3. A detailed description of the statistical limit setting procedure is given in section 9.4 and section 9.5.

9.1 Introduction

In this section we present a discussion of the statistical methods used to analyze our null result. The first step is to characterize how consistent the data are with a background only hypothesis - this is done by calculating the p value, which measures the probability of obtaining a test statistic at least as extreme as the one that was actually observed. Once we have established that the data do not exhibit a signal we can calculate upper limits on the signal strength $\mu = \frac{\sigma}{\sigma_{SM}}$. The statistical methods used to determine these upper

limits are discussed in this section and include the test statistics¹, the Modified Frequentist method (CLs Method), the likelihood model and how systematic uncertainties were taken into account. Three concepts which recur in the following sections are :

- **Parameter of interest (POI):** The signal strength ($\mu = \frac{\sigma}{\sigma_{SM}}$) is a measure of how close the upper limits on the cross section are compared to the Standard Model expectations, thus, for example, an upper limit that falls below a value of one for a particular mass range would indicate that that mass range is ruled out (for that Standard model process). μ is the parameter that is under investigation in this analysis.
- **Nuisance parameters (θ):** A nuisance parameter is a parameter that is not under investigation in this analysis but still has an impact on the POI. For example, the luminosity uncertainty and the energy scale uncertainty are nuisance parameters.
- **The Profile Likelihood ratio:** The profile likelihood ratio is used as recommended by the ATLAS group and it can also take systematics into account [99]. The likelihood is a distribution function of the parameters (both the POI and the nuisance parameters) given the data set. The “profile” means to set the nuisance parameters to the best fit value of the likelihood function, therefore the final test statistic only depends on the POI.

9.2 Test Statistics

A test statistic is a summary of the data as mention above. For different hypothesis tests, we use different test statistics which are sensitive to what we want to test. In the following sub-sections, the test statistics for the discovery and for the upper limits are presented. In addition, the method for calculating the p0 and the upper limits are given as well.

9.2.1 p0 for discovery

To quantify the agreement between the data and a background only hypothesis ($\mu=0$), the p value (p0) is calculated. A small p0 indicates that based on a background only model, the

¹The test statistic is a function of the data. It is a numerical summary of the data-set.

data are not consistent with that model. In such a case the null hypothesis (background only model) should be rejected. In the calculation of p_0 , only upward fluctuations of data are used while downward fluctuations are ignored because we expect the new signal process has a non-negative contribution to the event rate. The test statistic is defined as:

$$q_0 = \begin{cases} -2 \ln \frac{L(0, \hat{\hat{\theta}}(0))}{L(\hat{\mu}, \hat{\theta})} & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases} \quad (9.1)$$

where L is the likelihood model; the $\hat{\mu}$ and the $\hat{\theta}$ are the maximum likelihood estimator (the values that maximize L) for the signal strength and the nuisance parameters respectively; the $\hat{\hat{\theta}}(0)$ is the conditional maximum likelihood estimator (the value that maximize L under certain condition) for the nuisance parameters given zero signal strength. We can see that the value of q_0 is greater than 0 when $\hat{\mu} \geq 0$ and equal to 0 when $\hat{\mu} < 0$. The probability density function of q_0 denoted by $f(q_0|0)$ is given as

$$f(q_0|0) = \frac{1}{2} \delta(q_0) + \frac{1}{2} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{q_0}} e^{-q_0/2} \quad (9.2)$$

where a delta function at 0 with weight $\frac{1}{2}$ describes the downward fluctuation combined with a chi square function with one degree of freedom with weight $\frac{1}{2}$ describing the $\hat{\mu} > 0$ cases. You will see a different test statistic in the upper limits section which is designed for setting upper limit where upward fluctuations are ignored.

The p value is calculated using the right side tail from q_0^{obs} to $+\infty$:

$$p_0 = \int_{q_0^{obs}}^{+\infty} f(q_0|0) dq_0 \quad (9.3)$$

There is a p value for each signal mass point M_H . Although the mass parameter does not enter in the equation explicitly, it is associated with the signal strength μ which appears in the denominator of q_0 .

9.2.2 Upper limit

The profile likelihood ratio test statistic is used to set the upper limit, which is an optimal test based on the Neyman-Pearson lemma [100] if there are no nuisance parameters. We

expect that if the nuisance parameters are well constrained by control experiments², the likelihood ratio test statistic is still the most powerful one to use. As recommended by ATLAS collaboration [99], the profile likelihood ratio $\tilde{\lambda}(\mu)$ is defined as:

$$\tilde{\lambda}(\mu) = \begin{cases} \frac{L(\mu, \hat{\hat{\theta}}(\mu))}{L(0, \hat{\hat{\theta}}(0))} & \hat{\mu} < 0 \\ \frac{L(\mu, \hat{\hat{\theta}}(\mu))}{L(\hat{\mu}, \hat{\hat{\theta}})} & \hat{\mu} \geq 0 \end{cases} \quad (9.4)$$

where μ is the signal strength which is the parameter of interest (POI) in this analysis and θ is the nuisance parameter space (e.g. luminosity uncertainty, $e\gamma$ energy scale uncertainty etc.). $\hat{\mu}$ and $\hat{\theta}$ are the maximum likelihood estimators [101], and $\hat{\hat{\theta}}(\mu)$ is the conditional maximum likelihood estimator given μ . Since in this analysis we are looking for a signal with $\mu \geq 0$, then if $\hat{\mu} < 0$, $\hat{\mu}$ is set to 0. The $\tilde{\lambda}(\mu)$ measures the compatibility between the data and the model being tested.

The test statistic for setting the upper limit on μ is computed as

$$\tilde{q}_\mu = \begin{cases} -2 \ln \tilde{\lambda}(\mu) & \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases} = \begin{cases} -2 \ln \frac{L(\mu, \hat{\hat{\theta}}(\mu))}{L(0, \hat{\hat{\theta}}(0))} & \hat{\mu} < 0 \\ -2 \ln \frac{L(\mu, \hat{\hat{\theta}}(\mu))}{L(\hat{\mu}, \hat{\hat{\theta}})} & 0 \leq \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases} \quad (9.5)$$

From the definition of $\tilde{q}_\mu$ we can see that the larger the value of $\tilde{q}_\mu$ is, the greater the difference between the data and the hypothesized value of μ . Note that when $\hat{\mu} > \mu$, $\tilde{q}_\mu$ is 0. This indicates that when data fluctuates upward, we do not consider this as a discrepancy. Whereas in the test statistics for discovery, a downward data fluctuation is not considered as a discrepancy.

Having defined the test statistic, the next question is how to test a hypothesis with the parameter μ . In hypothesis testing, we cannot prove the hypothesis but only reject the hypothesis. In order to reject a hypothesis, it is useful to define the concept of p value. Given a model, the p value is the probability of obtaining a test statistic at least as extreme as the one that was actually observed. In this analysis, we perform one sided hypothesis

²The control experiments here refer to the experiments used to measure the uncertainties of the nuisance parameters.

testing, so the formula for the p value for a background-only model is:

$$p_b = \int_{\tilde{q}_\mu^{obs}}^{\infty} f(\tilde{q}_\mu|0) d\tilde{q}_\mu. \quad (9.6)$$

And the p value for a signal+background model with signal strength μ is:

$$p_{s+b} = \int_{\tilde{q}_\mu^{obs}}^{\infty} f(\tilde{q}_\mu|\mu) d\tilde{q}_\mu, \quad (9.7)$$

where $f(\tilde{q}_\mu|\mu)$ is the distribution of $\tilde{q}_\mu$, the subscript μ indicates the model being tested and the second argument of $f(\tilde{q}_\mu|\mu)$ gives the value of μ assumed in the distribution of the data.

Instead of requiring the p value for p_{s+b} to be less than the level α used in limit setting, we use $CLs = \frac{p_{s+b}}{1-p_b}$ which will be introduced in the CLs Method section.

9.3 CLs Method for setting upper limit

As mentioned above, a modified frequentist $CLs = \frac{p_{s+b}}{1-p_b}$ is used as a p value in ATLAS and CMS analyses. The method is described in detail in [104]. The reason for this definition is to mitigate the effect of a downward data fluctuation which would lead the data to disfavor both the signal+background model and the background-only model. In the case of a downward data fluctuation, p_{s+b} and $1 - p_b$ are both small and therefore tend to cancel each other out, so this definition of CLs minimizes the possibility that a downward fluctuation of the data leads an unrealistically low upper limits.

9.4 Likelihood Model

In this section we discuss the likelihood method. The Likelihood function is different conceptually from probability, although their forms are similar. The probability function has predefined parameters and can be used to predict the odds of outcomes. While the likelihood allows one to answer questions like: given an experimental observation, what can we say about the parameter of the model? In our case, the parameter is the signal strength μ . The total likelihood function is :

$$L\left(\mu, \theta = \bigcup_{c=1}^{n_{cat}} \theta_c | x = \bigcup x_c\right) = \prod_{c=1}^{n_{cat}} L_c(\mu, \theta_c | x_c) \quad (9.8)$$

where

$$L_c(\mu, \theta_c | x_c) = e^{-N'_c N_c} \prod_{k=1}^{N_c} \mathcal{L}_c(x_k | \mu, \theta_c) \quad (9.9)$$

where $N'_c = N_{signal,c} + N_{bkg,c}$. The $N_{signal,c}$ and $N_{bkg,c}$ are the signal events and background events respectively. When maximizing the likelihood function, the poisson factor $e^{-N'_c N_c}$ forces $\hat{N}_{signal,c} + \hat{N}_{bkg,c} = N_c$.

For each $\mathcal{L}_c(x | \mu, \theta_c)$, it is a weighted sum of the signal and the background probability density function:

$$\mathcal{L}_c(x | \mu, \theta_c) = \frac{N_{signal,c}(\mu, \theta_c^{norm})}{N_{signal,c} + N_{bkg,c}} f_{signal,c}(x | \theta_c^{shape}) + \frac{N_{bkg,c}}{N_{signal,c} + N_{bkg,c}} f_{bkg,c}(x | \theta_c^{bkg}) \quad (9.10)$$

where $N_{signal,c}(\mu, \theta_c^{norm})$ is the sum of different Higgs production processes. The systematic uncertainties are incorporated in the expression for $N_{signal,c}(\mu, \theta_c^{norm})$. Usually there are two ways to consider these systematics: either through the log-normal constraint or through the normal constraint. The log-normal has the form of $e^{\sqrt{\log(1+\sigma^2)}\theta}$ where θ is a normally distributed variable. In this analysis we use the log-normal constraint for theoretical scale uncertainties and experimental uncertainties that affect the signal yield. We have gaussian constraints for the uncertainties arising from the variations of the PDF set's eigenvalues and branching ratio.

$$N_{signal,c}(\mu, \theta_c^{norm}) = [N_{ggH,c}(\theta^{ggH}) + N_{VBF,c}(\theta^{VBF}) + N_{WH,c}(\theta^{WH}) + N_{ZH,c}(\theta^{ZH}) + N_{ttH,c}(\theta^{ttH})] * f(\theta_{exp}, \theta_{BR}) \quad (9.11)$$

where

$$\begin{aligned}
f(\theta_{exp}, \theta_{BR}) = & \exp(\sqrt{\log(1 + \sigma_{lumi}^2)} \theta_{lumi}) \times \\
& \exp(\sqrt{\log(1 + \sigma_{phID}^2)} \theta_{phID}) \times \\
& \exp(\sqrt{\log(1 + \sigma_{leptonID}^2)} \theta_{leptonID}) \times \\
& \exp(\sqrt{\log(1 + \sigma_{phES}^2)} \theta_{phES}) \times \\
& \exp(\sqrt{\log(1 + \sigma_{phRES}^2)} \theta_{phRES}) \times \\
& \exp(\sqrt{\log(1 + \sigma_{e\gamma_{iso}}^2)} \theta_{e\gamma_{iso}}) \times \\
& \exp(\sqrt{\log(1 + \sigma_{trig}^2)} \theta_{trig}) \times \\
& \exp(\sqrt{\log(1 + \sigma_{generator}^2)} \theta_{generator}) \times \\
& (1 + \sigma_{BR} * \theta_{BR})
\end{aligned} \tag{9.12}$$

The meaning of the subscripts: lumi is the luminosity; phID is the photon identification; leptonID is the lepton identification; phES is the photon energy scale; phRES is the photon resolution; $e\gamma_{iso}$ is the electron and photon isolation; trig is the trigger; BR is the branching ratio.

and

$$\begin{aligned}
N_{ggH,c}(\theta^{ggH}) &= N_c^{ggH,SM} \exp(\sqrt{\log(1 + \sigma_{ggHscale}^2)} \theta_{ggHscale}) (1 + \sigma_{ggPDF,ggH} \theta_{ggPDF}) \\
N_{VBF,c}(\theta^{VBF}) &= N_c^{VBF,SM} \exp(\sqrt{\log(1 + \sigma_{VBFscale}^2)} \theta_{VBFscale}) (1 + \sigma_{q\bar{q}PDF,VBF} \theta_{q\bar{q}PDF}) \\
N_{WH,c}(\theta^{WH}) &= N_c^{WH,SM} \exp(\sqrt{\log(1 + \sigma_{WHscale}^2)} \theta_{WHscale}) (1 + \sigma_{q\bar{q}PDF,WH} \theta_{q\bar{q}PDF}) \\
N_{ZH,c}(\theta^{ZH}) &= N_c^{ZH,SM} \exp(\sqrt{\log(1 + \sigma_{ZHscale}^2)} \theta_{ZHscale}) (1 + \sigma_{q\bar{q}PDF,ZH} \theta_{q\bar{q}PDF}) \\
N_{ttH,c}(\theta^{ttH}) &= N_c^{ttH,SM} \exp(\sqrt{\log(1 + \sigma_{ttHscale}^2)} \theta_{ttHscale}) (1 + \sigma_{ggPDF,ttH} \theta_{ggPDF})
\end{aligned} \tag{9.13}$$

where $N_c^{i,SM}$ are the mean values of the SM predictions, $N_{i,c}$ includes the systematic errors from scale and pdf uncertainties ('i' is ggH, VBF, WH, ZH, ttH), the θ 's are standard normally distributed variables and the σ 's are the uncertainties for each source of systematic error (the subscript PDF stands for the parton density function uncertainties and the scale stands for the scale uncertainties).

9.5 Systematic uncertainties

The systematic uncertainties result from uncertainties of the intrinsic nature of the measurement devices, assumptions made in the measurement methods, the theoretical model used to make predictions. These uncertainties, in contrast to statistical uncertainties, are not arising from the sample size but due to the limited knowledge of the data or incomplete theoretical framework. In this section, the theoretical uncertainties and experimental uncertainties were described.

9.5.1 Theoretical uncertainties

9.5.1.1 Parton density function uncertainties

When calculating the cross section for any process in hadron-hadron or lepton-hadron collision experiments, the effect of the parton density function (PDF) needs to be included. The parton density function $f_i(x, Q^2)$ gives the probability of finding, in the proton, a parton of flavor i (quarks or gluon) carrying a fraction x of the proton momentum with Q being the energy scale of the hard interaction. These parton density functions cannot be calculated theoretically and must be measured experimentally using a global fit to extract the parameters. There are several groups in the world that work on PDF parameterizations. The different PDF's arise from the use of different input data, different parametrization forms, different treatment of errors, etc. Hence, using different PDF's will result in different cross sections. These cross section calculations have been carried out by the ATLAS Higgs cross section working group. The treatment of PDF systematic uncertainties is to use a Gaussian distribution with the σ in the Gaussian distribution defined as

$$\sigma = \max(\sigma^+, \sigma^-) \quad (9.14)$$

where the σ^+ and σ^- are the upper and lower range of the PDF uncertainty respectively.

9.5.1.2 Scale uncertainties

The factorization scale parameter, μ_F , and the renormalization scale parameter, μ_R , are also used in the cross section calculation. Usually they are set to Q ($\mu_F = \mu_R = Q$) where

Q is the scale of the process (characterized by momentum transfer). The scale uncertainty evaluation was carried out by varying the scale from $Q/2$ to $2Q$ [63]. The treatment of scale systematic uncertainties uses a log-normal distribution with the σ of the Log-normal distribution assigned as

$$\sigma = \max(\sigma^+, \sigma^-) \quad (9.15)$$

where the σ^+ and σ^- are the upper and lower range of the scale uncertainty respectively. Using a log-normal distribution function instead of a gaussian function guarantees that the cross section is always positive.

9.5.1.3 Summary of PDF and Scale Systematic Uncertainties

In determining the PDF and scale systematic uncertainties, we use the Ref. [63] values as a function of the Higgs mass. In Table 9.1 we provide those systematic uncertainties at the lower and upper boundary of the Higgs mass search region (120 GeV and 150 GeV). These systematic uncertainties vary linearly over that search range, as is shown in Figure 9.1.

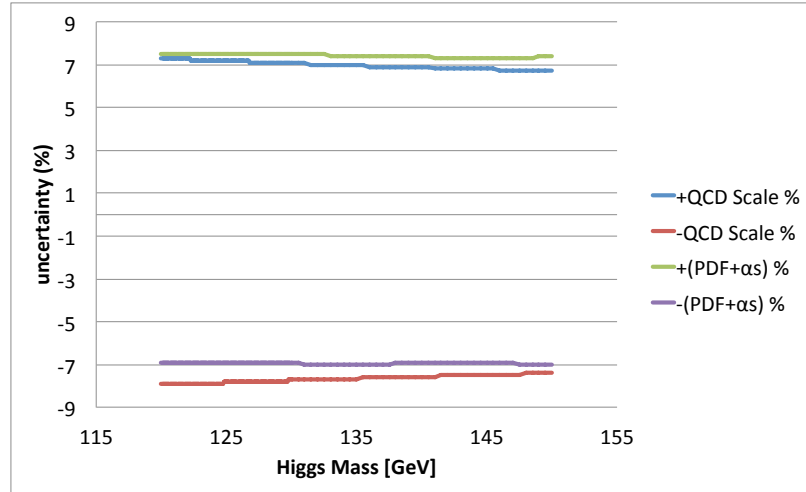


Figure 9.1: Higgs cross section variations from the PDF uncertainties and the QCD scale uncertainties.

Uncertainty source	Relative uncertainty (%)			
	7 TeV		8 TeV	
	120 GeV	150 GeV	120 GeV	150 GeV
$\sigma_{\sigma^-}^+(gg \rightarrow H)$ (scale)	+7.2 -7.9	+6.6 -7.4	+7.3 -7.9	+6.7 -7.4
$\sigma_{\sigma^-}^+(gg \rightarrow H)$ (PDF)	+7.6 -7.0	+7.6 -7.5	+7.5 -6.9	+7.4 -7.0
$\sigma_{\sigma^-}^+(\text{VBF})$ (scale)	+0.3 -0.4	+0.2 -0.1	+0.2 -0.2	+0.3 -0.2
$\sigma_{\sigma^-}^+(\text{VBF})$ (PDF)	+2.4 -2.1	+2.7 -2.1	+2.6 -2.8	+2.5 -2.7
$\sigma_{\sigma^-}^+(WH)$ (scale)	+0.4 -0.7	+0.4 -0.8	+0.1 -0.6	+0.1 -0.7
$\sigma_{\sigma^-}^+(WH)$ (PDF)	+3.4 -3.4	+3.3 -3.3	+3.4 -3.4	+3.3 -3.3
$\sigma_{\sigma^-}^+(ZH)$ (scale)	+1.5 -1.2	+1.8 -1.6	+1.5 -1.4	+2.1 -1.8
$\sigma_{\sigma^-}^+(ZH)$ (PDF)	+3.5 -3.5	+3.7 -3.7	+3.5 -3.5	+3.6 -3.6
$\sigma_{\sigma^-}^+(t\bar{t}H)$ (scale)	+3.4 -9.4	+2.9 -9.1	+3.9 -9.3	+3.5 -9.1
$\sigma_{\sigma^-}^+(t\bar{t}H)$ (PDF)	+8.4 -8.4	+8.4 -8.4	+7.8 -7.8	+8.0 -8.0
$\mathcal{B}(H \rightarrow Z\gamma)$	+9.4 -9.3	+6.0 -6.2	+9.4 -9.3	+6.0 -6.2

Table 9.1: Theoretical systematic uncertainties for the SM Higgs boson production cross section and branching fraction of the $H \rightarrow Z\gamma$ decay at the boundaries of the tested Higgs boson mass range (120 and 150 GeV). In each column, the first number is the uncertainty for $m_H = 120$ GeV, the second is the uncertainty for $m_H = 150$ GeV.

9.5.2 Experimental systematic uncertainties

The experimental uncertainties on the signal yield that are used in setting limits include: (a) the photon, electron and muon identification efficiency; (b) the electron and muon reconstruction efficiency; (c) the photon, electron and muon energy resolution; (d) the photon and electron energy scale; (e) the photon and electron isolation; (f) trigger; (g) the acceptance of the kinematic requirements; and (h) luminosity. These sources of experimental systematic error are discussed below and summarized in Tables 9.2 and 9.3 for 7 TeV and 8 TeV running respectively.

9.5.2.1 Photon identification efficiency (γ ID efficiency)

The uncertainty on the photon identification efficiency is based on the comparison of the efficiency obtained using MC and data-driven measurements. This uncertainty affects the signal yield. For 7 TeV, the signal yield systematic uncertainty is estimated by varying the photon identification efficiency scale factors by up to 8.78% (η , p_T and converted/unconverted dependent) and recomputing the signal yield. The range obtained in this way contributes $\pm 2.9\%$ to the systematic uncertainty. For 8 TeV, based on recommendations from the egamma working group [66], the preliminary systematic uncertainty on the efficiency of the cut-based tight selection for unconverted photons with $E_T^\gamma > 15$ GeV is 2.5% up to an E_T^γ of 40 GeV and 1.5% when $E_T^\gamma > 40$ GeV with $|\eta| < 1.81$ and 2.5% when $|\eta| > 1.81$. For converted photons, the systematic uncertainty is 2.5% for $E_T^\gamma < 40$ GeV and 1.5% for $E_T^\gamma > 40$ GeV.

9.5.2.2 Electron reconstruction and identification efficiency (electron reco+ID efficiency)

The electron reconstruction and identification efficiency scale factors are determined using the tag-and-probe method. The scale factors are applied to the MC to take into account the efficiency differences between data and MC. The scale factors and uncertainties are provided by a dedicated ATLAS performance working group [67]. The uncertainty of the reconstruction and identification efficiency is obtained by varying the efficiency scale factors within their $\pm 1\sigma$ uncertainties (For reco efficiency scale factor uncertainty: range from

0.6% to 5%, η and p_T dependent; for ID efficiency scale factor uncertainty: range from 0.7% to 6.9%, η and p_T dependent) and observing the change in the final acceptance. This uncertainty affects the signal yield.

9.5.2.3 Muon reconstruction and identification efficiency (μ reco+ID efficiency)

The muon reconstruction and identification efficiencies have been measured from the experimental data using a tag-and-probe method with dimuon decays on the Z . The scale factors are applied to the MC to take into account the efficiency differences between data and MC. The scale factors and uncertainties are provided by the ATLAS combined muon performance working group [68]. The reconstruction and identification efficiency uncertainties are obtained by varying the efficiency scale factors within their $\pm 1\sigma$ uncertainties (range from 0.2% to 0.4%) and observing the change in the final acceptance. This uncertainty affects the signal yield.

9.5.2.4 Electron and photon energy scale (e/γ Energy Scale)

The energy scale correction is applied to the data in order to match the data to the MC. There are four sources of systematic errors: the Z scale uncertainties (size of the data sample, the fitting method, choice of generator); the presampler scale uncertainty; the material thickness uncertainty; $low - p_T$ electrons uncertainty (active only below $p_T=20$ GeV). The energy scale factor as a function of η and p_T is obtained using $Z \rightarrow ee$ events. The scale factor is calculated by correcting the $Z \rightarrow ee$ line shape from data to the MC Z lineshape which is used as a reference. Therefore the size of the data sample of $Z \rightarrow ee$ events, the fitting method and the choice of MC generator are some of the sources of the scale factor errors. In addition, the presampler energy scale is another source because it is used to correct for energy lost upstream of the active EM calorimeter. The material thickness uncertainty arises from the imperfect knowledge of the material in front of the EM calorimeter which affects the e/γ energy measurement since the deposited energy in any additional material is not measured. However, on average the energy scale factor can recover this energy loss. The systematic uncertainty of the additional material is evaluated using a MC with additional material in front of the calorimeters. The $low - p_T$ uncertainty arises

from the energy scale factor extracted using $J/\Psi \rightarrow ee$ after applying the baseline calibration using $Z \rightarrow ee$. The four sources of energy scale uncertainties are varied independently and the total uncertainty is summed quadratically. It affects both the signal yield and the peak position.

9.5.2.5 Electron and photon isolation (e/γ isolation)

For the electron channel, the uncertainty is obtained by varying the electron and γ isolation criteria simultaneously by ± 100 MeV which is the average difference observed between data and MC for photons and electrons. For the muon channel, only the γ isolation has been varied. The relative changes of the signal efficiency is assigned as one source of the systematic error.

9.5.2.6 Photon and electron energy resolution (e/γ Energy Resolution)

The systematic error is obtained by varying the smearing correction within its uncertainties (up to 5%, η and p_T dependent) and observing the relative variation in the signal yield. The electron and photon smearing correction are varied at the same time. This affects both the signal yield and the signal shape resolution.

9.5.2.7 Muon energy resolution (μ Energy Resolution)

The uncertainty of the muon energy resolution is estimated by varying the muon momentum corrections within their uncertainties (less than 1%) and observing the relative variation in the signal yield. This affects both the signal yield and the signal shape resolution.

9.5.2.8 Trigger

The uncertainty is obtained by varying the trigger scale factor within its $\pm 1\sigma$ uncertainties (up to 4%) which is provided by a dedicated performance study group [69].

9.5.2.9 Acceptance of the kinematic requirements

The uncertainties related to the acceptance of the selection criteria are estimated by comparing the acceptance estimated with simulated signal events generated either using POWHEG

or MCFM, both interfaced to PYTHIA. This yields an uncertainty of 4%.

9.5.2.10 Luminosity

The uncertainty on the integrated luminosity is 1.8% and 3.6% for the 2011 and 2012 data respectively [70] [71].

9.5.3 signal mass resolution uncertainty

9.5.3.1 Electron and photon energy resolution (e/γ Energy resolution)

An effective signal resolution is defined as $\sigma_{eff} = f * \sigma_{cb} + (1 - f) * \sigma_{gs}$, where f is the fraction of Crystal Ball function and $(1 - f)$ is the fraction of the Gaussian function, σ_{cb} is the variance parameter for Crystal Ball and σ_{gs} is the one for a Gaussian function. σ_{cb} and σ_{gs} are obtained by fitting and then σ_{eff} s are calculated. The uncertainty of σ_{eff} is estimated by varying the electron and photon energy smearing corrections within its uncertainties.

9.5.3.2 Muon energy resolution (μ Energy Resolution)

The uncertainty of the muon energy resolution on the signal ΔM distribution is estimated by varying the muon momentum corrections within its uncertainties (less than 1%). The relative variation of the ΔM resolution is taken as a systematic uncertainty.

9.5.4 signal mass peak position uncertainty

9.5.4.1 Electron and photon energy scale (e/γ Energy Scale)

As stated above, four sources of systematic uncertainties have been considered and varied independently. The corrections for the electron and the photon are considered to be correlated. The peak position uncertainty due to the e/γ energy scale is calculated by summing the values in quadrature.

9.5.4.2 Muon momentum scale

The uncertainty on the peak position, ΔM , is obtained by varying the muon momentum scale within its uncertainties (about 0.1%) and the shift of the peak position is considered as a systematic. This uncertainty turns out to be negligible.

Systematic Uncertainty	$H \rightarrow Z(ee)\gamma(\%)$	$H \rightarrow Z(\mu\mu)\gamma(\%)$	Uncertainty Treatment
Signal Yield			
Luminosity	± 1.8	± 1.8	correlated
Trigger	± 0.2	± 0.7	uncorrelated
MC Generator kinematics	± 4.0	± 4.0	
γ ID efficiency	± 2.9	± 2.9	correlated
electron reco+ID efficiency	± 3.0	—	uncorrelated
μ reco+ID efficiency	NA	± 0.7	uncorrelated
e/γ Energy Scale	± 0.3	± 0.2	correlated
ZeeAllUp	± 0.16	± 0.04	
Material	± 0.02	± 0.11	
Presampler	± 0.06	± 0.03	
Low pt	± 0.24	± 0.12	
e vs γ differences	± 0.10	± 0.07	
e/γ isolation	± 0.3	± 0.2	correlated
e/γ energy resolution	± 0.2	< 0.1	correlated
μ momentum scale	—	± 0.2	negligible
μ momentum resolution	—	± 0.1	negligible
Signal ΔM resolution			
e/γ Energy Resolution	± 5.00	± 2.40	correlated
μ Energy Resolution	—	± 1.50	
Signal ΔM peak position			
e/γ Energy Scale	± 0.2 GeV	± 0.2 GeV	correlated
muon momentum scale	—	negligible	negligible

Table 9.2: Summary of the systematic uncertainties on the signal yield and invariant mass distribution for $m_H = 125$ GeV, at $\sqrt{s} = 7$ TeV.

Systematic Uncertainty	$H \rightarrow Z(ee)\gamma(\%)$	$H \rightarrow Z(\mu\mu)\gamma(\%)$	Uncertainty Treatment
Signal Yield			
Luminosity	± 3.6	± 3.6	correlated
Trigger	± 0.4	± 0.8	uncorrelated
MC Generator kinematics	± 4.0	± 4.0	
γ ID efficiency	± 2.4	± 2.4	correlated
electron reco+ID efficiency	± 2.7	—	uncorrelated
μ reco+ID efficiency	—	± 0.6	uncorrelated
e/γ Energy Scale	± 1.4	± 0.3	correlated
ZeeAllUp	± 1.39	± 0.18	
Material	± 0.02	± 0.15	
Presampler	± 0.04	± 0.02	
Low pt	± 0.23	± 0.20	
e vs γ differences	± 0.09	± 0.09	
e/γ isolation	± 0.3	± 0.4	correlated
e/γ energy resolution	± 0.2	< 0.1	correlated
μ momentum scale	—	± 0.1	negligible
μ momentum resolution	—	< 0.1	negligible
Signal Δm resolution			
e/γ energy resolution	± 5.00	± 2.40	correlated
μ momentum resolution	NA	0.00	negligible
Signal Δm peak position			
e/γ energy scale	± 0.2 GeV	± 0.2 GeV	correlated
muon momentum scale	—	negligible	negligible

Table 9.3: Summary of the systematic uncertainties on the signal yield and invariant mass distribution for $m_H = 125$ GeV, at $\sqrt{s} = 8$ TeV.

Chapter 10

Results

The statistical result used to interpret the data is based on the asymptotic formulae [99] implemented in RooStats [102][103].

10.1 p value for a background-only hypothesis

The statistical procedures used to test the background-only hypothesis is described in Ref [99]. The compatibility of the data with background-only ($\mu = 0$) is tested in order to quantify the significance of a possible observation. The local p_0 as a function of Higgs mass M_H is shown in Figure 10.1. Using the full 2011 and 2012 data, corresponding to 4.6 fb^{-1} of pp collisions at $\sqrt{s} = 7 \text{ TeV}$ and 20.7 fb^{-1} of pp collisions at $\sqrt{s} = 8 \text{ TeV}$, the expected p_0 ranges between 0.40 and 0.48 for M_H between 120 GeV and 150 GeV. The observed p_0 distribution is compatible with the data being composed of background only. The smallest p_0 (0.039) corresponding to a significance of 1.76σ , occurs for a mass of 141.5 GeV. This is due to the upward data fluctuation around $\Delta M = 51 \text{ GeV}$ in $\sqrt{s} = 8 \text{ TeV}$ electron channel. The expected p_0 at $M_H = 125 \text{ GeV}$ is 0.448, corresponding to a significance of 0.13σ , while the observed one is 0.268 (0.62σ). Since there is no significant signal observed, we derive the upper limits on the signal strength as a function of Higgs mass, M_H , in the next section.

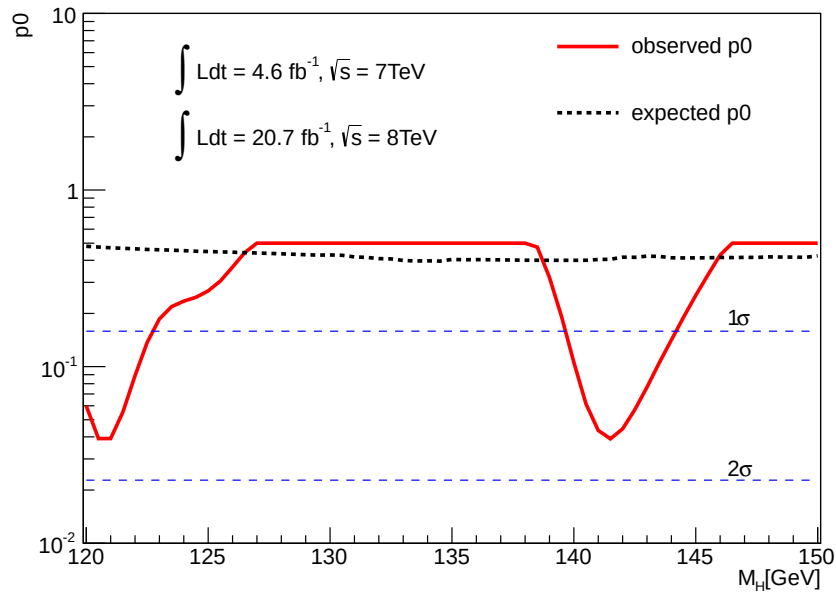


Figure 10.1: Expected (dashed black line) and observed (solid red line) p_0 (compatibility of the data with the background-only hypothesis) as a function of the Higgs boson mass, using 4.6 fb^{-1} of pp collisions at $\sqrt{s} = 7 \text{ TeV}$ and 20.7 fb^{-1} of pp collisions at $\sqrt{s} = 8 \text{ TeV}$.

10.2 Exclusion limits on $H \rightarrow Z\gamma$ signal strength μ

The upper limit on the signal strength μ is calculated for a range of Higgs masses from $M_H = 120$ GeV to $M_H = 150$ GeV, as shown in Figure 10.2. The observed limit and expected limit at 125.5 GeV, corresponding to the mass of the recently discovered Higgs-like boson, is 17.0 and 13.2 respectively. The bump of the observed limit around $M_H = 141$ GeV is due to the upward data fluctuation around $\Delta M = 51$ GeV in $\sqrt{s} = 8$ TeV electron channel. The values of the observed and expected limits is shown in Table 10.1.

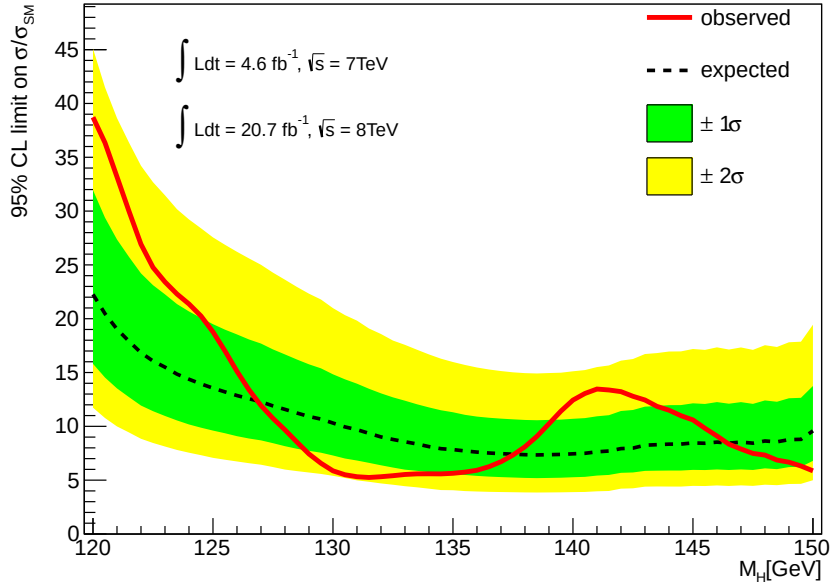


Figure 10.2: Observed 95% C.L. limits (solid red line) on the production cross section of a SM Higgs boson decaying to $Z\gamma$, as a function of the Higgs boson mass, using 4.6 fb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV and 20.7 fb^{-1} of pp collisions at $\sqrt{s} = 8$ TeV. The median expected 95% C.L. exclusion limits (dashed black line) are also shown. The green and yellow bands correspond to the $\pm 1\sigma$ and $\pm 2\sigma$ intervals.

Mass (GeV)	Observed	Expected	Mass (GeV)	Observed	Expected
120	38.71	22.26	136	5.91	7.60
121	33.18	19.02	137	6.71	7.46
122	26.91	16.84	138	8.10	7.36
123	23.41	15.49	139	10.23	7.36
124	21.36	14.37	140	12.43	7.44
125	18.77	13.55	141	13.46	7.63
126	15.11	12.86	142	13.22	7.92
127	11.93	12.28	143	12.44	8.24
128	9.66	11.57	144	11.49	8.34
129	7.42	10.91	145	10.59	8.46
130	5.86	10.30	146	9.10	8.53
131	5.31	9.69	147	7.89	8.53
132	5.31	9.01	148	7.32	8.63
133	5.52	8.56	149	6.66	8.76
134	5.58	8.11	150	5.86	9.57
135	5.63	7.83			

Table 10.1: Expected and Observed limit for signal strength μ for Higgs mass from 120 GeV to 150 GeV.

10.3 Conclusions and outlook

Searches for the Standard Model Higgs boson have been performed in the $H \rightarrow Z\gamma$ channel with the ATLAS experiment at the LHC using 4.6 fb^{-1} of pp collision data recorded during 2011 at the center-of-mass energy of 7 TeV and 20.7 fb^{-1} recorded during 2012 at the center-of-mass energy of 8 TeV.

This is the first direct search for Higgs using the $l\bar{l}\gamma$ signature via the $Z\gamma$ channel in ATLAS. We did not see a significant excess and therefore upper limits have been set on the $H \rightarrow Z\gamma$ signal strength, μ , in the Higgs mass, M_H , range from 120 GeV to 150 GeV. In particular, the upper limit at $M_H = 125.5 \text{ GeV}$ is 13.2 (expected) and 17.0 (observed).

The search is currently being updated using more kinematic categories with different signal to background ratios in order to improve sensitivity. And the analysis will also benefit from an improved understanding of the ATLAS detector and more precise calibrations.

With the data accumulated in the next data-taking period (LHC will resume running in 2015), we should have enough data to either discover or exclude a Higgs-like boson in 125 GeV - 126.5 GeV region which has been discovered in the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ$ channels [1]. Moreover, since the leptons and photons can be well reconstructed and measured, the spin analysis can be carried out in the future as well, in order to test if the new particle is a scalar boson as required by the Standard Model.

Part IV

Appendices

Appendix A

MC samples

A list of the simulated 7 TeV and 8 TeV samples are provided in Table A.1 and Table A.2 respectively.

Table A.1: Simulated samples at $\sqrt{s} = 7$ TeV. The number of events listed in the second column is either the number of events in the D3PD (for unskinned samples) or of events in the AOD before skimming (for skinned samples). The cross sections for the signals are from Ref. [64], while for background samples the cross sections are those returned by the MC generators. The equivalent luminosity is computed as $N_{\text{evts}}/(\sigma_{\text{gen}}\epsilon_{\text{fit}})$.

D3PD name	$H \rightarrow Z\gamma$ signal	Events	σ_{gen} [nb]	ϵ_{fit}	eq. lumi [fb ⁻¹]
mc11.7TeV.167480.PowhegPythia8_AU2CT10- <i>ggH</i> 120_Zllgam.merge.NTUP.PHOTON.e1749.s1570.s1571.r3108.r3109.p1373		100000	1.866e-06	1	5.359e+10
mc11.7TeV.167481.PowhegPythia8_AU2CT10- <i>ggH</i> 125_Zllgam.merge.NTUP.PHOTON.e1749.s1570.s1571.r3108.r3109.p1373		100000	2.382e-06	1	4.198e+10
mc11.7TeV.167482.PowhegPythia8_AU2CT10- <i>ggH</i> 130_Zllgam.merge.NTUP.PHOTON.e1749.s1570.s1571.r3108.r3109.p1373		99999	2.788e-06	1	3.587e+10
mc11.7TeV.167484.PowhegPythia8_AU2CT10- <i>ggH</i> 135_Zllgam.merge.NTUP.PHOTON.e1749.s1570.s1571.r3108.r3109.p1373		99999	3.005e-06	1	3.328e+10
mc11.7TeV.167485.PowhegPythia8_AU2CT10- <i>ggH</i> 140_Zllgam.merge.NTUP.PHOTON.e1749.s1570.s1571.r3108.r3109.p1373		99999	2.374e-06	1	4.212e+10
mc11.7TeV.167486.PowhegPythia8_AU2CT10- <i>ggH</i> 145_Zllgam.merge.NTUP.PHOTON.e1749.s1570.s1571.r3108.r3109.p1373		70000	1.922e-06	1	3.642e+10
mc11.7TeV.167487.PowhegPythia8_AU2CT10- <i>ggH</i> 150_Zllgam.merge.NTUP.PHOTON.e1749.s1570.s1571.r3108.r3109.p1373		99999	1.464e-06	1	6.831e+10
mc11.7TeV.167490.PowhegPythia8_AU2CT10-VBFH120_Zllgam.merge.NTUP.PHOTON.e1749.s1570.s1571.r3108.r3109.p1373		100000	1.433e-07	1	6.978e+11
mc11.7TeV.167491.PowhegPythia8_AU2CT10-VBFH125_Zllgam.merge.NTUP.PHOTON.e1749.s1570.s1571.r3108.r3109.p1373		100000	1.900e-07	1	5.265e+11
mc11.7TeV.167492.PowhegPythia8_AU2CT10-VBFH130_Zllgam.merge.NTUP.PHOTON.e1749.s1570.s1571.r3108.r3109.p1373		100000	2.300e-07	1	4.348e+11
mc11.7TeV.167494.PowhegPythia8_AU2CT10-VBFH135_Zllgam.merge.NTUP.PHOTON.e1749.s1570.s1571.r3108.r3109.p1373		100000	2.560e-07	1	3.906e+11
mc11.7TeV.167495.PowhegPythia8_AU2CT10-VBFH140_Zllgam.merge.NTUP.PHOTON.e1749.s1570.s1571.r3108.r3109.p1373		99899	2.083e-07	1	4.796e+11
mc11.7TeV.167496.PowhegPythia8_AU2CT10-VBFH145_Zllgam.merge.NTUP.PHOTON.e1749.s1570.s1571.r3108.r3109.p1373		100000	1.735e-07	1	5.764e+11
mc11.7TeV.167497.PowhegPythia8_AU2CT10-VBFH150_Zllgam.merge.NTUP.PHOTON.e1749.s1570.s1571.r3108.r3109.p1373		100000	1.356e-07	1	7.375e+11
$Z \rightarrow \ell\ell\gamma$ background					
mc11.7TeV.145161.Sherpa-Zeegamma-3jets.merge.NTUP.PHOTON.e1443.s1372.s1370.r3108.r3109.p1373		399799	0.027344	1	1.46211e+07
mc11.7TeV.145162.Sherpa-Zmumugamma-3jets.merge.NTUP.PHOTON.e1443.s1372.s1370.r3108.r3109.p1373		399800	0.027301	1	1.46442e+07
$Z(\rightarrow \ell\ell) + jets$ background					
mc11.7TeV.147771.Sherpa-CT10_Zmumu.merge.NTUP.PHOTON.e1443.s1372.s1370.r3108.r3109.p1373		9995488	1.0452	1	9.56323e+06
mc11.7TeV.107650.AlpgenJimmyZeeenp0.p120.merge.NTUP.PHOTON.e835.s1299.s1300.r3043.r2993.p1373		6618284	0.6696	1	9.88394e+06
mc11.7TeV.107651.AlpgenJimmyZeeenp1.p120.merge.NTUP.PHOTON.e835.s1299.s1300.r3043.r2993.p1373		1334897	0.13452	1	9.92341e+06
mc11.7TeV.107652.AlpgenJimmyZeeenp2.p120.merge.NTUP.PHOTON.e835.s1299.s1300.r3043.r2993.p1373		2004195	0.040706	1	4.92359e+07
mc11.7TeV.107653.AlpgenJimmyZeeenp3.p120.merge.NTUP.PHOTON.e835.s1299.s1300.r3043.r2993.p1373		549949	0.011262	1	4.88333e+07
mc11.7TeV.107654.AlpgenJimmyZeeenp4.p120.merge.NTUP.PHOTON.e835.s1299.s1300.r3043.r2993.p1373		149948	0.0028447	1	5.27114e+07
mc11.7TeV.107655.AlpgenJimmyZeeenp5.p120.merge.NTUP.PHOTON.e835.s1299.s1300.r3043.r2993.p1373		50000	0.00075691	1	6.60581e+07
mc11.7TeV.147770.Sherpa-CT10_Zee.merge.NTUP.PHOTON.e1443.s1372.s1370.r3108.r3109.p1373		9998665	1.0458	1	9.56078e+06
mc11.7TeV.107660.AlpgenJimmyZmumuNp0.p120.merge.NTUP.PHOTON.e835.s1299.s1300.r3043.r2993.p1373		6615230	0.66956	1	9.87997e+06
mc11.7TeV.107661.AlpgenJimmyZmumuNp1.p120.merge.NTUP.PHOTON.e835.s1299.s1300.r3043.r2993.p1373		1334296	0.13455	1	9.91673e+06
mc11.7TeV.107662.AlpgenJimmyZmumuNp2.p120.merge.NTUP.PHOTON.e835.s1299.s1300.r3043.r2993.p1373		1999941	0.040642	1	4.92087e+07
mc11.7TeV.107663.AlpgenJimmyZmumuNp3.p120.merge.NTUP.PHOTON.e835.s1299.s1300.r3043.r2993.p1373		549896	0.011279	1	4.8754e+07
mc11.7TeV.107664.AlpgenJimmyZmumuNp4.p120.merge.NTUP.PHOTON.e835.s1299.s1300.r3043.r2993.p1373		150000	0.002837	1	5.28728e+07
mc11.7TeV.107665.AlpgenJimmyZmumuNp5.p120.merge.NTUP.PHOTON.e835.s1299.s1300.r3043.r2993.p1373		50000	0.00076123	1	6.56832e+07
$t\bar{t}$ background					
mc11.7TeV.109345.T1_McAtNLO-Jimmy-2leptonsMll60GeV.merge.NTUP.PHOTON.e961.s1310.s1300.r3043.r2993.p1373		499948	0.14519	0.081719	4.21371e+07
mc11.7TeV.109345.T1_McAtNLO-Jimmy-2leptonsMll60GeV.merge.NTUP.PHOTON.e961.s1310.s1300.r3043.r2993.p1373		499948	0.14519	0.081719	4.21371e+07

Table A.2: Simulated samples at $\sqrt{s} = 8$ TeV. The number of events listed in the second column is either the number of events in the D3PD (for unskimmed samples) or of events in the AOD before skimming (for skimmed samples). The cross sections for the signals are from Ref. [65], while for background samples the cross sections are those returned by the MC generators. The equivalent luminosity is computed as $N_{\text{evts}}/(\sigma_{\text{gen}}\epsilon_{\text{filt}})$.

D3PD name	Events	σ_{gen} [nb]	ϵ_{filt}	eq. lumi [nb ⁻¹]
<i>H</i> $\rightarrow$ <i>Z</i> γ signal				
mc12.8TeV.167480.PowhegPythia8_AU2CT10.ggH120_Zllgam.merge.NTUP_PHOTON.el746.s1581.s1586.r3658.r3549.p1344	100000	2.368e-06	1	4.223e+10
mc12.8TeV.167481.PowhegPythia8_AU2CT10.ggH125_Zllgam.merge.NTUP_PHOTON.el746.s1581.s1586.r3658.r3549.p1344	99999	3.035e-06	1	3.294e+10
mc12.8TeV.167482.PowhegPythia8_AU2CT10.ggH130_Zllgam.merge.NTUP_PHOTON.el746.s1581.s1586.r3658.r3549.p1344	99999	3.558e-06	1	2.811e+10
mc12.8TeV.167484.PowhegPythia8_AU2CT10.ggH135_Zllgam.merge.NTUP_PHOTON.el746.s1581.s1586.r3658.r3549.p1344	99999	3.848e-06	1	2.599e+10
mc12.8TeV.167485.PowhegPythia8_AU2CT10.ggH140_Zllgam.merge.NTUP_PHOTON.el746.s1581.s1586.r3658.r3549.p1344	99998	3.046e-06	1	3.283e+10
mc12.8TeV.167486.PowhegPythia8_AU2CT10.ggH145_Zllgam.merge.NTUP_PHOTON.el746.s1581.s1586.r3658.r3549.p1344	100000	2.475e-06	1	4.040e+10
mc12.8TeV.167487.PowhegPythia8_AU2CT10.ggH150_Zllgam.merge.NTUP_PHOTON.el746.s1581.s1586.r3658.r3549.p1344	89999	1.888e-06	1	4.767e+10
mc12.8TeV.167490.PowhegPythia8_AU2CT10.VBFH120_Zllgam.merge.NTUP_PHOTON.el746.s1581.s1586.r3658.r3549.p1344	79999	1.848e-07	1	4.329e+11
mc12.8TeV.167491.PowhegPythia8_AU2CT10.VBFH125_Zllgam.merge.NTUP_PHOTON.el746.s1581.s1586.r3658.r3549.p1344	99900	2.454e-07	1	4.071e+11
mc12.8TeV.167492.PowhegPythia8_AU2CT10.VBFH130_Zllgam.merge.NTUP_PHOTON.el746.s1581.s1586.r3658.r3549.p1344	99999	2.975e-07	1	3.361e+11
mc12.8TeV.167494.PowhegPythia8_AU2CT10.VBFH135_Zllgam.merge.NTUP_PHOTON.el746.s1581.s1586.r3658.r3549.p1344	100000	3.319e-07	1	3.013e+11
mc12.8TeV.167495.PowhegPythia8_AU2CT10.VBFH140_Zllgam.merge.NTUP_PHOTON.el746.s1581.s1586.r3658.r3549.p1344	99997	2.707e-07	1	3.694e+11
mc12.8TeV.167496.PowhegPythia8_AU2CT10.VBFH145_Zllgam.merge.NTUP_PHOTON.el746.s1581.s1586.r3658.r3549.p1344	99500	2.261e-07	1	4.401e+11
mc12.8TeV.167497.PowhegPythia8_AU2CT10.VBFH150_Zllgam.merge.NTUP_PHOTON.el746.s1581.s1586.r3658.r3549.p1344	79999	1.771e-07	1	4.517e+11
<i>Z</i> $\rightarrow$ $\ell\ell\gamma$ background				
mc12.8TeV.145161.Sherpa_CT10.eegammaPt10.merge.NTUP_PHOTON.el434.s1499.s1504.r3658.r3549.p1211	1184995	0.032261	1	3.67315e+07
mc12.8TeV.145162.Sherpa_CT10.mumugammaPt10.merge.NTUP_PHOTON.el434.s1499.s1504.r3658.r3549.p1211	1199997	0.032317	1	3.71321e+07
<i>Z</i> ($\rightarrow \ell\ell$) + jets background				
mc12.8TeV.147770.Sherpa_CT10.Zee.merge.NTUP_PHOTON.el161.s1469.s1470.r3542.r3549.p1208.p1210	999568	1.2079	1	8.27847e+06
mc12.8TeV.107650.AlpgenJimmy_AUET2CTEQ6L1.ZeeNp0.merge.NTUP_PHOTON.el218.s1469.s1470.r3542.r3549.p1208.p1210	6604283	0.71177	1	9.27868e+06
mc12.8TeV.107651.AlpgenJimmy_AUET2CTEQ6L1.ZeeNp1.merge.NTUP_PHOTON.el218.s1469.s1470.r3542.r3549.p1208.p1210	1329994	0.15517	1	8.57121e+06
mc12.8TeV.107652.AlpgenJimmy_AUET2CTEQ6L1.ZeeNp2.merge.NTUP_PHOTON.el218.s1469.s1470.r3542.r3549.p1208.p1210	404798	0.048745	1	8.3044e+06
mc12.8TeV.107653.AlpgenJimmy_AUET2CTEQ6L1.ZeeNp3.merge.NTUP_PHOTON.el218.s1469.s1470.r3542.r3549.p1208.p1210	109998	0.014225	1	7.73272e+06
mc12.8TeV.107654.AlpgenJimmy_AUET2CTEQ6L1.ZeeNp4.merge.NTUP_PHOTON.el218.s1469.s1470.r3542.r3549.p1208.p1210	30000	0.0037595	1	7.97978e+06
mc12.8TeV.107655.AlpgenJimmy_AUET2CTEQ6L1.ZeeNp5.merge.NTUP_PHOTON.el218.s1469.s1470.r3542.r3549.p1208.p1210	10000	0.0010945	1	9.13659e+06
mc12.8TeV.147771.Sherpa_CT10.Zmumu.merge.NTUP_PHOTON.el161.s1469.s1470.r3542.r3549.p1208.p1256	999975	1.2078	1	8.2795e+06
mc12.8TeV.107660.AlpgenJimmy_AUET2CTEQ6L1.ZmumuNp0.merge.NTUP_PHOTON.el218.s1469.s1470.r3542.r3549.p1208.p1256	6609982	0.71211	1	9.28225e+06
mc12.8TeV.107661.AlpgenJimmy_AUET2CTEQ6L1.ZmumuNp1.merge.NTUP_PHOTON.el218.s1469.s1470.r3542.r3549.p1208.p1256	1334897	0.15477	1	8.62504e+06
mc12.8TeV.107662.AlpgenJimmy_AUET2CTEQ6L1.ZmumuNp2.merge.NTUP_PHOTON.el218.s1469.s1470.r3542.r3549.p1208.p1256	404897	0.048912	1	8.27807e+06
mc12.8TeV.107663.AlpgenJimmy_AUET2CTEQ6L1.ZmumuNp3.merge.NTUP_PHOTON.el218.s1469.s1470.r3542.r3549.p1208.p1256	110000	0.014226	1	7.73232e+06
mc12.8TeV.107664.AlpgenJimmy_AUET2CTEQ6L1.ZmumuNp4.merge.NTUP_PHOTON.el218.s1469.s1470.r3542.r3549.p1208.p1256	29999	0.0037838	1	7.92827e+06
mc12.8TeV.107665.AlpgenJimmy_AUET2CTEQ6L1.ZmumuNp5.merge.NTUP_PHOTON.el218.s1469.s1470.r3542.r3549.p1208.p1256	10000	0.0011148	1	8.97022e+06
<i>t\bar{t}</i> background				
mc12.8TeV.110001.McAtNloJimmy_CT10.ttbar_dilepton.merge.NTUP_PHOTON.el193.s1469.s1470.r3542.r3549.p1208.p1210	7694555	0.02181	1	3.528e+08
mc12.8TeV.110001.McAtNloJimmy_CT10.ttbar_dilepton.merge.NTUP_PHOTON.el193.s1469.s1470.r3542.r3549.p1208.p1256	7694555	0.02181	1	3.528e+08

Appendix B

MC and data comparison plots

photon p_T in Figure B.1, electron p_T in Figure B.2, muon p_T in Figure B.3 and photon η in Figure B.4.

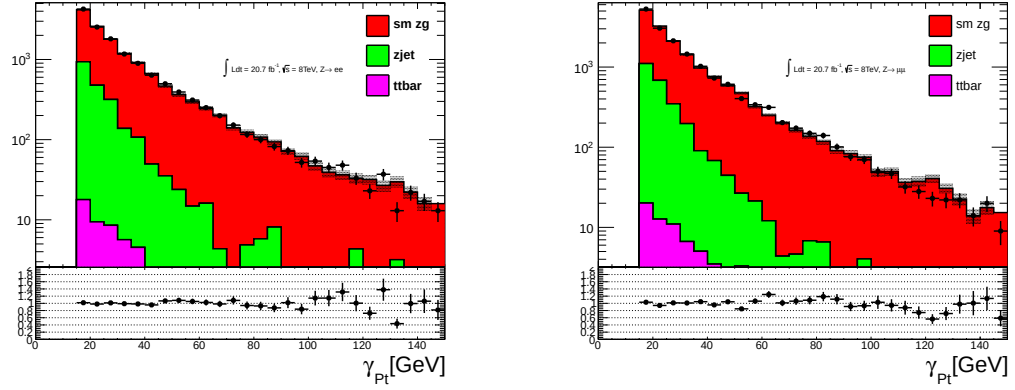


Figure B.1: Reconstructed photon transverse momentum in electron channel (left) and muon channel (right). The sub-plot shows the ratio $= \frac{N_{data}}{N_{MC}}$.

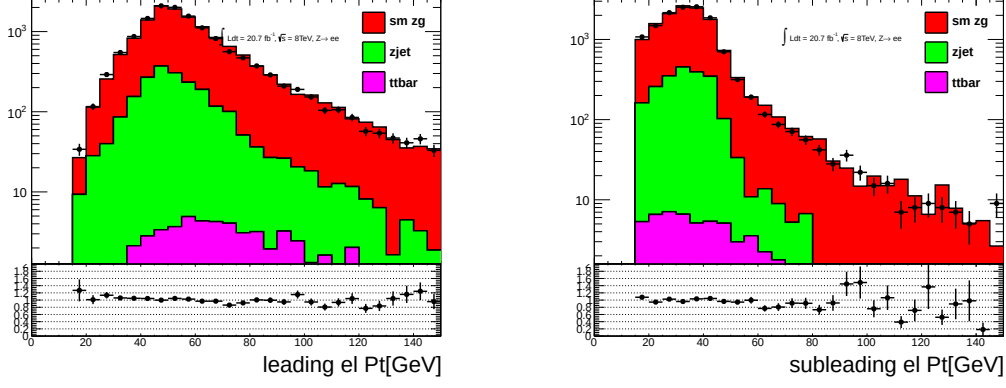


Figure B.2: Reconstructed electron transverse momentum. Leading electron is on the left and sub leading electron is on the right. The leading electron is the one with highest p_T . The sub leading electron is one with lower p_T . The sub-plot shows the ratio $= \frac{N_{data}}{N_{MC}}$.

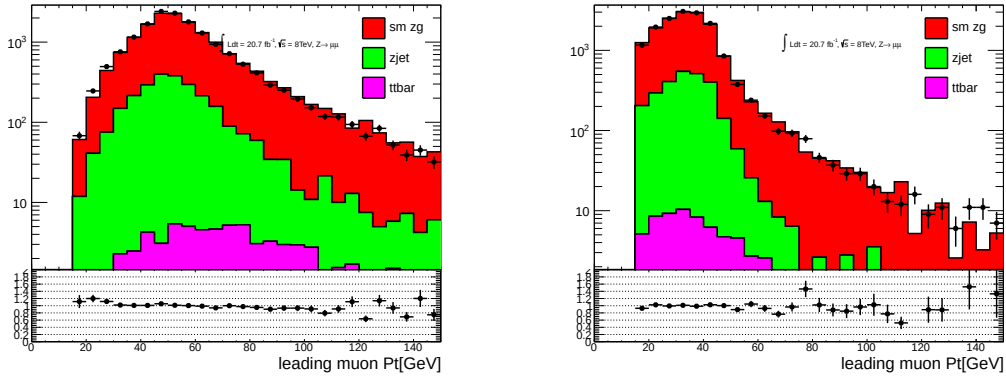


Figure B.3: Reconstructed muon transverse momentum. Leading muon is on the left and sub-leading muon is on the right. The leading muon is the one with highest p_T . The sub leading muon is one with lower p_T . The sub-plot shows the ratio $= \frac{N_{data}}{N_{MC}}$.

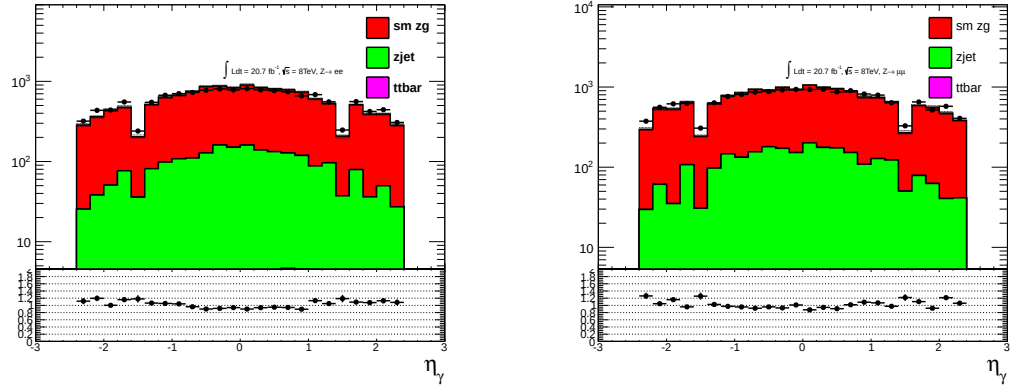


Figure B.4: Reconstructed photon η distribution in the electron channel (left) and muon channel (right). The dips in the plots correspond to the transition region of the barrel and the end cap of the EM calorimeter. The sub-plot shows the ratio $= \frac{N_{data}}{N_{MC}}$.

Part V

Bibliography

Bibliography

- [1] The 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 (2012) 1-29, arXiv:1207.7214 [hep-ex] <http://arxiv.org/abs/1207.7214>
- [2] The CMS Collaboration, “ Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”, Phys. Lett. B 716 (2012) 30, arXiv:1207.7235 [hep-ex] <http://arxiv.org/abs/1207.7235>
- [3] S.L. Glashow, Nucl. Phys. 22 (4) (1961) 579. <http://www.sciencedirect.com/science/article/pii/0029558261904692>
- [4] S. Weinberg, Phys. Rev. Lett. 19 (1967) 1264. http://prl.aps.org/abstract/PRL/v19/i21/p1264_1
- [5] A. Salam, “Weak and electromagnetic interactions”, in Elementary particle theory: relativistic groups and analyticity, N. Svartholm, ed., p. 367. Almqvist Wiksell, 1968. Proceedings of the eighth Nobel symposium.
- [6] G. 't Hooft, M. Veltman, Nucl. Phys. B 44 (1972) 189. <http://www.sciencedirect.com/science/article/pii/0550321372902799>
- [7] Griffiths, David J. (2004). “Introduction to Quantum Mechanics” (2nd ed.). Prentice Hall. ISBN 0-13-805326-X.
- [8] J. J. Thomson, “Cathode rays” Philosophical Magazine, 44 (1897) 293. http://www.tandfonline.com/doi/abs/10.1080/14786449708621070?journalCode=tphm16#.UmArnmSif_c

- [9] S.H. Neddermeyer, C.D. Anderson (1937). “Note on the Nature of Cosmic-Ray Particles”, Phys. Rev. 51 (1937) 884. http://prola.aps.org/abstract/PR/v51/i10/p884_1
- [10] M. L. Perl et al. (1975). “Evidence for Anomalous Lepton Production in e^+e^- Annihilation”, Phys. Rev. Lett. 35 (1975) 1489. http://prl.aps.org/abstract/PRL/v35/i22/p1489_1
- [11] Fukuda, Y., et al (1998). “Measurements of the Solar Neutrino Flux from Super-Kamiokande’s First 300 Days”. Physical Review Letters 81 (6): 11581162. arXiv:hep-ex/9805021. <http://arxiv.org/abs/hep-ex/9805021>
- [12] T. Araki et al. (KamLAND Collaboration) (2005). “Measurement of Neutrino Oscillation with KamLAND: Evidence of Spectral Distortion”. Physical Review Letters 94 (8): 081801. arXiv:hep-ex/0406035. <http://arxiv.org/abs/hep-ex/0406035>
- [13] McQuarrie, Donald A. (2000). Statistical Mechanics (1st ed.). Sausalito, CA 94965: University Science Books. p. 55. ISBN 1-891389-15-7.
- [14] F. Englert and R. Brout, Broken symmetry and the mass of gauge vector mesons, Phys. Rev. Lett. 13 (1964) 321. http://prl.aps.org/abstract/PRL/v13/i9/p321_1
- [15] P. W. Higgs, Broken symmetries, massless particles and gauge elds, Phys. Lett. 12 (1964) 132. <http://www.sciencedirect.com/science/article/pii/0031916364911369>
- [16] P. W. Higgs, Broken symmetries and the masses of gauge bosons, Phys. Rev. Lett. 13 (1964) 508. http://prl.aps.org/abstract/PRL/v13/i16/p508_1
- [17] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, “Global conservation laws and massless particles”, Phys. Rev. Lett. 13 (1964) 585. http://prl.aps.org/abstract/PRL/v13/i20/p585_1
- [18] P. W. Higgs, “Spontaneous symmetry breakdown without massless bosons”, Phys. Rev. 145 (1966) 1156. http://prola.aps.org/abstract/PR/v145/i4/p1156_1

- [19] T. W. B. Kibble, “Symmetry breaking in non-Abelian gauge theories”, *Phys. Rev.* 155 (1967) 1554. http://prola.aps.org/abstract/PR/v155/i5/p1554_1
- [20] The ATLAS collaboration, “Evidence for the spin-0 nature of the Higgs boson using ATLAS data”, *Phys. Lett. B* 726 (2013) 120, arXiv:1307.1432 [hep-ex]. <http://arxiv.org/abs/1307.1432>
- [21] The CMS Collaboration, “Study of the mass and spin-parity of the Higgs boson candidate via its decays to Z boson pairs”, *Phys. Rev. Lett.* 110 (2013) 081803, arXiv:1212.6639 [hep-ex]. <http://arxiv.org/abs/1212.6639>
- [22] The ATLAS collaboration, “Search for the Standard Model Higgs boson in the $H \rightarrow Z\gamma$ decay mode with pp collisions at $\sqrt{s} = 7$ and 8 TeV”, ATLAS-CONF-2013-009. <http://cds.cern.ch/record/1523683>
- [23] The ATLAS collaboration, “Search for a Standard Model Higgs boson in $H \rightarrow \mu\mu$ decays with the ATLAS detector”, ATLAS-CONF-2013-010. <http://cds.cern.ch/record/1523695>
- [24] ATLAS collaboration, “Search for invisible decays of a Higgs boson produced in association with a Z boson in ATLAS”, ATLAS-CONF-2013-011. <http://cds.cern.ch/record/1523696>
- [25] The CMS collaboration, “Search for invisible Higgs produced in association with a Z boson”, CMS-PAS-HIG-13-018. <http://cds.cern.ch/record/1561758>
- [26] L. Evans, P. Bryant (Eds.), *LHC Machine*, JINST, 3 (2008) S08001. <http://iopscience.iop.org/1748-0221/3/08/S08001>
- [27] The CMS collaboration, “Search for Higgs Boson Production in Association with a Top-Quark Pair and Decaying to Bottom Quarks or Tau Leptons”, CMS-PAS-HIG-13-019. <http://cds.cern.ch/record/1564682>
- [28] Cheng-Wei Chiang, Kei Yagyu, “Higgs boson decays to $\gamma\gamma$ and $Z\gamma$ in models with Higgs extensions”, *Phys. Rev. D* 87, 033003 (2013), arXiv:1207.1065 [hep-ph]. <http://arxiv.org/abs/1207.1065>

- [29] Ian Low, Joseph Lykken, Gabe Shaughnessy, “Singlet scalars as Higgs imposters at the Large Hadron Collider”, *Phys. Rev. D* 84, 035027 (2011), arXiv:1105.4587 [hep-ph]. <http://arxiv.org/abs/1105.4587>
- [30] Bogumila Swiezewska, Maria Krawczyk, “Diphoton rate in the Inert Doublet Model with a 125 GeV Higgs boson”, *Phys. Rev. D* 88 (2013) 035019, arXiv:1212.4100 [hep-ph]. <http://arxiv.org/abs/1212.4100>
- [31] Don Bunk, Jay Hubisz, Bithika Jain, “Higgs Decays in Gauge Extensions of the Standard Model”, arXiv:1309.7988 [hep-ph]. <http://arxiv.org/abs/1309.7988>
- [32] Hance, M, et al. “Performance Studies for e/gamma Calorimeter Isolation”, ATL-COM-PHYS-2011-1186, <https://cds.cern.ch/record/1379530>
- [33] ATLAS Collaboration, “ATLAS: Letter of intent for a general-purpose pp experiment at the large hadron collider at CERN”, CERN-LHCC-92-004 (1992). <http://cds.cern.ch/record/291061>
- [34] ATLAS Collaboration, “ATLAS Technical Proposal for a General-Purpose pp Experiment at the Large Hadron Collider at CERN”, CERN-LHCC-94-43 (1994). <http://cds.cern.ch/record/290968>
- [35] ATLAS Collaboration, “The ATLAS Experiment at the CERN Large Hadron Collider”, *JINST* 3 (2008) S08003. <http://iopscience.iop.org/1748-0221/3/08/S08003>
- [36] ALEPH, CDF, D, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the Tevatron Electroweak Working Group, and the SLD electroweak and heavy flavour groups, “Precision Electroweak Measurements and Constraints on the Standard Model”, CERN-PH-EP-2010-095 (2010). arXiv:1012.2367 [hep-ex]. <http://arxiv.org/abs/1012.2367>
- [37] The TEVNPH Working Group, for the CDF, D0 Collaborations, “Combined CDF and D0 Search for Standard Model Higgs Boson Production with up to 10.0 fb^{-1} of

- Data”, FERMILAB-CONF-12-065-E, CDF Note 10806, D0 Note 6303, arXiv:1203.3774 [hep-ex]. <http://arxiv.org/abs/1203.3774>
- [38] ALEPH, DELPHI, L3 and OPAL Collaborations, The LEP Working Group for Higgs boson searches, “Search for the standard model Higgs boson at LEP”, Phys. Lett. B 565 (2003) 61. <http://arxiv.org/abs/hepex/0306033>
- [39] CDF Collaboration, T. Aaltonen et al., “Combined search for the standard model Higgs boson decaying to a bb pair using the full CDF data set” accepted by Phys. Rev. Lett. (2012), arXiv:1207.1707 [hep-ex]. <http://arxiv.org/abs/1207.1707>
- [40] D Collaboration, V. M. Abazov et al., “Combined search for the standard model Higgs boson decaying to bb using the D Run II data set” submitted to Phys. Rev. Lett. (2012), arXiv:1207.6631 [hep-ex]. <http://arxiv.org/abs/1207.6631>
- [41] CDF Collaboration, D Collaboration, “Evidence for a particle produced in association with weak bosons and decaying to a bottom-antibottom quark pair in Higgs boson searches at the Tevatron” submitted to Phys. Rev. Lett. (2012), arXiv:1207.6436 [hep-ex]. <http://arxiv.org/abs/1207.6436>
- [42] ATLAS Collaboration, “Combined search for the Standard Model Higgs boson in pp collisions at $\sqrt{s} = 7\text{TeV}$ with the ATLAS detector” Phys. Rev. D86 (2012) 032003. <http://prd.aps.org/abstract/PRD/v86/i3/e032003>
- [43] CMS Collaboration, “Combined results of searches for the standard model Higgs boson in pp collisions at $\sqrt{s} = 7\text{TeV}$ ” Phys. Lett. B 710 (2012) 26. <http://arxiv.org/abs/1202.1488>
- [44] Kalman, R.E. (1960). “A new approach to linear filtering and prediction problems”, Journal of Basic Engineering 82 (1): 35-45. doi:10.1115/1.3662552. <http://fluidsengineering.asmedigitalcollection.asme.org/article.aspx?articleid=1430402>

- [45] ATLAS Collaboration, “Expected Performance of the ATLAS Experiment Detector, Trigger and Physics” CERN-OPEN-2008-020, 2008. <http://cds.cern.ch/record/1125884>
- [46] ATLAS DETECTOR AND PHYSICS PERFORMANCE, Technical Design Report, Volume I CERN-LHCC 99-14, 1999. <http://cds.cern.ch/record/391176>
- [47] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal, “Calorimeter Clustering Algorithms: Description and Performance”, ATL-LARG-PUB-2008-002, ATL-COM-LARG-2008-003, 2008. <http://cds.cern.ch/record/1099735>
- [48] ATLAS Collaboration, “Data-Quality Requirements and Event Cleaning for Jets and Missing Transverse Energy Reconstruction with the ATLAS Detector in Proton-Proton Collisions at a Center-of-Mass Energy of $\sqrt{s} = 7$ TeV,” Tech. Rep. ATLAS-CONF-2010-038, CERN, Geneva, 2010. <http://cds.cern.ch/record/1277678>
- [49] The ATLAS Collaboration, “Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data” Eur. Phys. J. C72 (2012) 1909, arXiv:1110.3174. <http://arxiv.org/abs/1110.3174>
- [50] R. Fruehwirth, “A Gaussian-mixture approximation of the Bethe-Heitler model of electron energy loss by bremsstrahlung” Comp. Phys. Comm. 154 (2003) 131. <http://www.sciencedirect.com/science/article/pii/S0010465503002923>
- [51] The ATLAS Collaboration. Csc note-photon conversions in atlas. August 2008. <http://www.physics.smu.edu/web/research/preprints/SMU-HEP-08-13.pdf>
- [52] Agostinelli, S., Allison, J., Amako, K., Apostolakis, J., Araujo, H., Arce, P., Asai, M., Axen, D. et al. (2003). “Geant4 -a simulation toolkit”. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 506 (2003) 250. <http://www.sciencedirect.com/science/article/pii/S0168900203013688>

- [53] Torbjorn Sjostrand, Stephen Mrenna, Peter Skands, “A Brief Introduction to PYTHIA 8.1” Comput.Phys.Commun.178:852-867,2008 arXiv:0710.3820 [hep-ph]. <http://arxiv.org/abs/0710.3820>
- [54] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, et al., “Event generation with SHERPA 1.1,” JHEP 0902 (2009) 007, arXiv:0811.4622 [hep-ph]. <http://arxiv.org/abs/0811.4622>
- [55] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, “ALPGEN, a Generator for Hard Multiparton Processes in Hadronic Collisions,” JHEP 0307 (2003) 001, arXiv:hep-ph/0206293 [hep-ph]. <http://arxiv.org/abs/hep-ph/0206293>
- [56] P. Nason, “A new method for combining NLO QCD with shower Monte Carlo” algorithms, JHEP 11 (2004) 040 hep-ph/0409146 <http://arxiv.org/abs/hep-ph/0409146>
- [57] S. Frixione and B. R. Webber, “Matching NLO QCD computations and parton shower” simulations, JHEP 06 (2002) 029 hep-ph/0204244. <http://arxiv.org/abs/hep-ph/0204244>
- [58] S. Hoche and F. Krauss and S. Schumann and F. Siegert QCD matrix elements and truncated showers, JHEP 05 (2009) 053 arXiv:0903.1219. <http://arxiv.org/abs/0903.1219>
- [59] S. Alioli, P. Nason, C. Oleari and E. Re, JHEP 0904 (2009) 002, arXiv:0812.0578. <http://arxiv.org/abs/0812.0578>
- [60] P. Nason and C. Oleari, JHEP 1002 (2010) 037 arXiv:0911.5299. <http://arxiv.org/abs/0911.5299>
- [61] S. Alioli, P. Nason, C. Oleari and E. Re, JHEP 0807 (2008) 060 arXiv:0805.4802. <http://arxiv.org/abs/0805.4802>
- [62] S. Alioli, P. Nason, C. Oleari and E. Re, JHEP 1101 (2011) 095 arXiv:1009.5594. <http://arxiv.org/abs/1009.5594>
- [63] LHC Higgs Cross Section Working Group,

<https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections>

- [64] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and R. Tanaka (Eds.), Handbook of LHC Higgs cross sections: 1. Inclusive observables, 2011. arXiv:1101.0593 [hep-ph]. CERN-2011-002. <http://arxiv.org/abs/1101.0593>
- [65] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and R. Tanaka (Eds.), Handbook of LHC Higgs Cross Sections: 2. Differential distributions, 2012. arXiv:1201.3084 [hep-ph]. <http://arxiv.org/abs/1201.3084>
- [66] ATLAS Collaboration, egamma working group, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/PhotonID2012>
- [67] ATLAS Collaboration, egamma working group, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/EfficiencyMeasurements>
- [68] ATLAS Collaboration, combined muon working group, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/MCPAnalysisGuidelinesData2012>
- [69] ATLAS Collaboration, TrigMuonEfficiency, <https://twiki.cern.ch/twiki/bin/viewauth/Atlas/TrigMuonEfficiency>
- [70] ATLAS Collaboration, “Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC”, arXiv:1302.4393 [hep-ex]. <http://arxiv.org/abs/1302.4393>
- [71] ATLAS Collaboration, “Preliminary Luminosity Determination in pp Collisions at $\sqrt{s} = 8$ TeV using the ATLAS Detector in 2012”, ATL-COM-LUM-2012-013, <https://cds.cern.ch/record/1494059>
- [72] ATLAS Collaboration, “Measurements of the properties of the Higgs-like boson in the four lepton decay channel with the ATLAS detector using 25 fb^{-1} of proton-proton collision data”, ATLAS-CONF-2013-013, <http://cds.cern.ch/record/1523699>
- [73] ATLAS Collaboration, ElectronGamma Working Group, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/PhotonIDTool>

- [74] J. E. Gaiser, Appendix-F Charmonium Spectroscopy from Radiative Decays of the J/Psi and Psi-Prime, Ph.D. Thesis, SLAC-R-255 (1982). <http://inspirehep.net/record/183554>
- [75] H. Georgi, S. Glashow, M. Machacek, D.V. Nanopoulos, Phys. Rev. Lett. 40 (1978) 692. http://prl.aps.org/abstract/PRL/v40/i11/p692_1
- [76] A. Djouadi, M. Spira, P.M. Zerwas, Phys. Lett. B 264 (1991) 440. <http://www.sciencedirect.com/science/article/pii/037026939190375Z>
- [77] S. Dawson, Nucl. Phys. B 359 (1991) 283. <http://www.sciencedirect.com/science/article/pii/0550321391900612>
- [78] M. Spira, A. Djouadi, D. Graudenz, P.M. Zerwas, Nucl. Phys. B 453 (1995) 17, arXiv:hep-ph/9504378. <http://arxiv.org/abs/hep-ph/9504378>
- [79] R. Harlander, W.B. Kilgore, Phys. Rev. Lett. 88 (2002) 201801, arXiv:hep-ph/0201206. <http://arxiv.org/abs/hep-ph/0201206>
- [80] C. Anastasiou, K. Melnikov, Nucl. Phys. B 646 (2002) 220, arXiv:hep-ph/0207004. <http://arxiv.org/abs/hep-ph/0207004>
- [81] V. Ravindran, J. Smith, W.L. van Neerven, Nucl. Phys. B 665 (2003) 325. <http://www.sciencedirect.com/science/article/pii/S0550321303004577>
- [82] U. Aglietti, R. Bonciani, G. Degrassi, A. Vicini, Phys. Lett. B 595 (2004) 432. <http://www.sciencedirect.com/science/article/pii/S0370269304009268>
- [83] S. Actis, G. Passarino, C. Sturm, S. Uccirati, Phys. Lett. B 670 (2008) 12. <http://www.sciencedirect.com/science/article/pii/S0370269308012689>
- [84] R. Cahn, S. Dawson, Phys. Lett. B 136 (1984) 196; R. Cahn, S. Dawson, Phys. Lett. B 138 (1984) 464 (Erratum) <http://www.sciencedirect.com/science/article/pii/0370269384911808>
- [85] M. Ciccolini, A. Denner, S. Dittmaier, Phys. Rev. Lett. 99 (2007) 161803. <http://prl.aps.org/abstract/PRL/v99/i16/e161803>

- [86] M. Ciccolini, A. Denner, S. Dittmaier, Phys. Rev. D 77 (2008) 013002. <http://prd.aps.org/abstract/PRD/v77/i1/e013002>
- [87] K. Arnold, M. Bahr, G. Bozzi, F. Campanario, C. Englert, et al. Comput. Phys. Commun. 180 (2009) 1661, arXiv:0811.4559 [hep-ph]. <http://arxiv.org/abs/0811.4559>
- [88] P. Bolzoni, F. Maltoni, S.-O. Moch, M. Zaro, Phys. Rev. Lett. 105 (2010) 011801. <http://prl.aps.org/abstract/PRL/v105/i1/e011801>
- [89] S. Glashow, D.V. Nanopoulos, A. Yildiz, Phys. Rev. D 18 (1978) 1724. http://prd.aps.org/abstract/PRD/v18/i5/p1724_1
- [90] T. Han, S. Willenbrock, Phys. Lett. B 273 (1991) 167. <http://www.sciencedirect.com/science/article/pii/0370269391905728>
- [91] O. Brein, A. Djouadi, R. Harlander, Phys. Lett. B 579 (2004) 149. <http://arxiv.org/abs/hep-ph/0307206>
- [92] M.L. Ciccolini, S. Dittmaier, M. Kramer, Phys. Rev. D 68 (2003) 073003. <http://prd.aps.org/abstract/PRD/v68/i7/e073003>
- [93] Z. Kunszt, Nucl. Phys. B 247 (1984) 339. <http://www.sciencedirect.com/science/article/pii/0550321384905534>
- [94] W. Beenakker, et al., Phys. Rev. Lett. 87 (2001) 201805.
- [95] W. Beenakker, et al., Nucl. Phys. B 653 (2003) 151. <http://arxiv.org/abs/hep-ph/0107081>
- [96] S. Dawson, L.H. Orr, L. Reina, D. Wackerroth, Phys. Rev. D 67 (2003) 071503. <http://arxiv.org/abs/hep-ph/0211438>
- [97] S. Dawson, C. Jackson, L.H. Orr, L. Reina, D. Wackerroth, Phys. Rev. D 68 (2003) 034022. <http://prd.aps.org/abstract/PRD/v68/i3/e034022>

- [98] Arfken, G. "Chebyshev (Tschebyscheff) Polynomials" and "Chebyshev Polynomials—Numerical Applications." section 13.3 and 13.4 in *Mathematical Methods for Physicists*, 3rd ed. Orlando, FL: Academic Press, pp. 731-748, 1985.
- [99] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J. C* 71 (2011) 1554. <http://arxiv.org/abs/1007.1727>
- [100] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses," *Phil. Trans. R. Soc. Lon.* A231 (1933) pp. 289-337.
- [101] Pratt, John W. (1976). "F. Y. Edgeworth and R. A. Fisher on the efficiency of maximum likelihood estimation". *The Annals of Statistics* 4 (3): 501-514.
- [102] L. Moneta, K. Belasco, K.S. Cranmer, S. Kreiss, A. Lazzaro, et al., "The RooStats Project", *PoS ACAT2010* (2010) 057, arXiv:1009.1003. <http://arxiv.org/abs/1009.1003>
- [103] K. W. Verkerke and D. Kirkby, "The RooFit toolkit for data modeling", *Tech. Rep. physics/0306116*, SLAC, Stanford, CA, Jun, 2003. arXiv:physics/0306116. <http://arxiv.org/abs/physics/0306116>
- [104] A. L. Read, "Presentation of search results: The CL(s) technique", *J. Phys. G: Nucl. Part. Phys* 28 (2002) 2693. <http://iopscience.iop.org/0954-3899/28/10/313/>