Search for the Standard Model Higgs boson in the H $\rightarrow$ WW $\rightarrow$ lνqq channel

Joseph David Goodell

Charlottesville, VA

B.S., The University of Pennsylvania, 2008

A Dissertation presented to the Graduate Faculty
of the University of Virginia in Candidacy for the Degree of
Doctor of Philosophy

Department of Physics

University of Virginia

August, 2015

Abstract

One of the biggest recent successes of the standard model (SM) was the 2012 discovery of a new scalar particle consistent with an SM-like Higgs boson by the CMS and ATLAS experiments at the Large Hadron Collider (LHC). The production of Higgs particles and their subsequent decay allows many distinct final states to be observed. Presented here is a search for an SM Higgs boson with mass ≃125 GeV that decays through two W bosons, where one W decays hadronically and the other leptonically. While $H \to WW$ has been observed at the LHC in the fully-leptonic final state, analyses in the lvqq channel have not yet achieved sufficient sensitivity to a low-mass Higgs. This analysis was optimized directly for a low-mass Higgs boson and aims to complement the observations of the Higgs in this regime. The decay chain $H \to WW \to l\nu qq$ requires one W boson to have an off-shell mass; further, the presence of a neutrino in the final state makes Higgs mass reconstruction difficult. Finally, this decay channel suffers from a large irreducible background from W+jets production. This dissertation presents a search for the semi-leptonic W decay via a multivariate analysis of the 2012 8 TeV proton-proton collision data, a total luminosity of ∼19 fb^{-1}, collected at CMS.

We approve the dissertation of Joseph David Goodell.

Date of Signature

Supervisor: Prof. Christopher Neu

Committee Member: Prof. Hank Thacker

Committee Member: Prof. Roger Chevalier

PhD Committee Chair: Prof. Brad Cox

Contents

Contents **iv**

List of Figures **vii**

List of Tables **xv**

Acknowledgements **xviii**

1 Introduction **1**

2 Theoretical Background: The Standard Model **4**

2.1 The Standard Model 4

2.2 Quantum Electrodynamics 6

2.3 Electro-Weak Interaction 8

2.3.1 Yang-Mills Theory 8

2.3.2 Glashow-Weinberg-Salam model for EW Interactions 9

2.4 Strong Interaction 12

2.5 Higgs Mechanism 14

2.6 The Higgs Boson 18

2.7 Success of the Standard Model 19

2.8 Higgs Production in a p-p Collider 23

2.9 $H \to WW \to lvjj$ Production at the LHC 24

2.10 $H \to WW \to lvjj$ Backgrounds 27

3 The Large Hadron Collider **30**

3.1 Accelerator System 33

3.2 Magnets and Cryogenic System 39

3.3 Radiofrequency Cavities 42

4 Compact Muon Solenoid **44**

4.1 The Superconducting Solenoid . . . 47

4.2 The Tracker . . . 48

4.2.1 The Silicon Pixel Detector . . . 48

4.2.2 The Silicon Strip Detector . . . 49

4.3 The Electromagnetic Calorimeter . . . 50

4.4 The Hadronic Calorimeter . . . 54

4.5 The Muon Chambers . . . 57

5 Event Reconstruction **61**

5.1 Data Acquisition . . . 61

5.1.1 L1 Trigger and HLT . . . 61

5.1.2 T1 sites and data storage . . . 62

5.2 Particle Flow Event Reconstruction . . . 63

5.2.1 Iterative Tracking . . . 63

5.2.2 Calorimeter Clustering . . . 64

5.2.3 Linking Tracks and Clusters . . . 65

5.2.4 Cluster Calibration . . . 66

5.3 Physics Object Reconstruction . . . 68

5.3.1 Muons . . . 69

5.3.2 Electrons . . . 69

5.3.3 Charged Hadron . . . 70

5.3.4 Photons and Neutral Hadrons . . . 71

5.3.5 Jets . . . 71

5.3.6 Missing Transverse Energy ($\not{E}_{\mathrm{T}}$) . . . 74

6 $H \to WW \to l\nu qq$ Analysis Part 1: Selection **76**

6.1 Data and MC Samples . . . 77

6.1.1 Data Samples . . . 78

6.1.2 Signal Samples . . . 78

6.1.3 Background Samples . . . 79

6.1.4 MC Pileup Reweighting . . . 80

6.2 Multijet-QCD Background . . . 80

6.3 Event Selection . . . 82

6.3.1 Trigger . . . 83

6.3.2 Vertex Selection . . . 83

6.3.3 Electron Selection . . . 83
6.3.4 Muon Selection . . . 85
6.3.5 Jet Selection . . . 85
6.3.6 Analysis Cuts . . . 86
6.4 MC Corrections . . . 89
6.4.1 Jet Energy Corrections . . . 89
6.4.2 Jet Energy Resolution . . . 90
6.4.3 $\not{E}_{T\phi}$ Corrections . . . 91
6.4.4 CSV Reweighting . . . 93
6.4.5 TTbar Reweighting . . . 94
6.4.6 $cos(\theta_l)$ Weights . . . 94
6.4.7 QCD Reweighting . . . 96
6.5 MC Yields . . . 103

7 $H \to WW \to l\nu qq$ Analysis Part 2: MVA **106**
7.1 Multivariate Analysis: Boosted Decision Tree . . . 106
7.2 Kinematic Variable Selection and Definition . . . 110
7.3 BDT Input Optimization . . . 114
7.3.1 BDT Parameter Optimization . . . 118
7.4 BDT Input Variables: Data to Monte Carlo Comparisons . . . 118
7.5 BDT Output: Data to Monte Carlo Comparisons . . . 120
7.6 Systematic Uncertainties . . . 120

8 Results **132**
8.1 Statistical Methods: Limit Setting . . . 132
8.2 Counting Experiment Results . . . 135
8.3 Shape-based Analysis . . . 136
8.3.1 Yields and Limits Using Statistical Uncertainties . . . 136
8.3.2 Limits Using Full Systematic Uncertainties . . . 138
8.4 Summary of Results . . . 141

9 Conclusion / Summary **142**

Bibliography **144**

List of Acronyms **153**

List of Figures

1.1 The CMS experiment has observed a new boson at m$\sim$125 GeV/c^2 2

2.1 The Standard Model of elementary particles [1]. Six quarks (shown in purple) and six leptons (shown in green) comprise the fermionic component of the SM, while the force mediating gauge bosons (shown in red) comprise the bosonic component. Additionally, the Higgs Boson (the result of electroweak symmetry breaking) is shown in yellow. 5

2.2 Feynman diagram for a neutron decaying into a proton via the weak interaction. In this case, a down type quark radiates a W boson and becomes an up type quark, while the W decays leptonically. 10

2.3 Example of a gluon exchange between quarks resulting in a change of color charge in the quarks. Quarks carry either a positive or negative color charge while gluons carry one component of both positive and negative color charge. 12

2.4 The 'sombrero' potential with an unstable state at $\phi = 0$ an a non-zero minimum 15

2.5 Tree level Feynman diagrams for Higgs coupling vertices to vector bosons(a,b), fermions(c), and to itself(d). 20

2.6 Milestones in particle physics showing discovery and measurements of SM particles. 22

2.7 Higgs production cross-sections at the LHC for 8 TeV pp collisions 23

2.8 Feynman diagrams for Higgs production modes at the LHC 25

2.9 Higgs decays modes at the LHC. Red line denotes a mass of 125 GeV. 26

2.10 Feynman diagram for the SM process a Higgs boson is created through the gluon-gluon fusion process and decays semi-leptonically to two quarks, one lepton, and one neutrino. 26

2.11 Representative Feynman diagrams of SM V+jets processes 29

2.12 Representative Feynman diagrams of SM diboson processes 29

2.13 Representative Feynman diagrams of the $t\bar{t}$ process [2] 29

2.14 Representative Feynman diagrams of SM single top processes 29

3.1 Artistic representation of the LHC accelerator complex with both surface and subsurface views [3]. 30

3.2 Aerial view of the LHC complex, spanning the French-Swiss border [4]. 31

3.3 Integrated Luminosity delivered to the CMS experiment from 2010-12 [5] 33

3.4 Layout of facilities along the LHC ring at CERN [6] 34

3.5 Overview of the LHC injection chain at CERN [7] 36

3.6 Features of the Linac2, Proton Synchrotron Booster, and Proton Synchrotron . . 37

3.7 Features of Super Proton Synchrotron and Large Hadron Collider 38

3.8 Schematic of the bunch structure for filling the LHC ring. Initially, 6 groups of protons are provided to the PS which splits them into 72 bunches. The bunch trains then travel to the SPS and into the LHC with beam gaps arising due to the rise time of injection kickers. This leads to a total of 3564 possible buckets with 2808 filled (assuming 25ns bunch spacing).Reprinted from [8] 39

3.9 The LHC ring is divided into eight octants with 4 interaction points marked with stars[8] 40

3.10 Cross section view of an LHC dipole magnet and cryostat. Reprinted from [8]. . 41

3.11 Features of the LHC dipole magnets 41

3.12 Depiction of the single turn beam injection for the LHC. Reprinted from [9]. . . . 42

3.13 Schematic of a cryostat used at the LHC to accelerate the proton beams through 4 RF cavities with 2MV potentials each. 2 of these cryostats are used in succession on each beamline [10]. 43

4.1 A cutaway diagram of the CMS detector, with subsystem statistics [11]. 44

4.2 The CMS coordinate system, reprinted from [12]. 46

4.3 A slice of the CMS detector showing how particles interact and deposit energy with the various sub-systems. The tracker measures the trajectory of charged particles; electrons and photons are measured/absorbed in the ECAL, while hadrons deposit most of their energy in the HCAL. Muon chambers measure the trajectory of muons, the distinct double curve is due to the opposite magnetic field they feel outside of the solenoid. Reprinted from [13]. 47

4.4 A side view of the tracker subsystem. The pixel detector forms the innermost layers, with three concentric rings of detectors in the barrel and two in the endcap. The tracker inner barrel (TIB), tracker inner disks (TID), tracker outer barrel (TOB) and tracker end caps (TEC) are composed of concentric layers of silicon strip detectors [14]. 48

4.5 Geometric layout of the pixel detector (a) and hit coverage as a function of η(b) [14]. 49

4.6 Schematic drawing of a TIB/TID assembly. This structure is mounted inside the TOB (one on each end), and shows the 'margherita', a service distribution disk used to route out signals and control the cooling supply lines [14]. 50

4.7 Signal to noise measurement for the TOB (a) and TIB (b) sections of the silicon tracker, operated in deconvolution mode (optimal conditions) [15]. 51

4.8 Layout of the ECAL sub-detector [14] 51

4.9 Example of lead tungstate crystals with photodetectors attached. (a)A barrel crystal with APD attached. (B)An endcap crystal with VPT attached. [14]. . . 52

4.10 Relative response to laser light (440 nm) measured by the ECAL laser monitoring system, averaged over all crystals in bins of pseudorapidity, for the 2011 and 2012 data taking periods Layout of the ECAL sub-detector [16]. 53

4.11 One half of an ECAL end-cap, called a Dee. Two Dees form a disk with an inner bore for the beam-line to pass through. 5x5 supercrystal modules are mounted in preparation for installation at CMS [14]. 54

4.12 Longitudinal cross-section of the HCAL with the four sub-systems labeled [14]. . 54

4.13 Isometric view of HCAL segment [14]. 55

4.14 Russian navy shells re-used in the CMS Hadron Calorimeter [17]. 56

4.15 Cross-sectional view of the HF calorimeter. The beam-line runs parallel to the diagram, with the right side pointing toward the interaction point. Fibers run parallel to beam-line in the absorber, are bundled together (gray region) and routed to the read-out boxes [14]. 57

4.16 The layout of muon DT chambers in one of the 5 wheels that make up the muon system [14]. 58

4.17 (a)Sketch of a DT cell showing drift lines. The top and bottom plates are held at ground potential. (b) Equipotential lines for half of a drift cell. The anode wire is located on the right side of this plot[14]. 59

4.18 Muon CSC wedge, showing cathode panels in orange and anode wires in gold [14]. 59

4.19 Muon transverse momentum resolution as a function of p_{T} using the muon system only (black), the inner tracker only (blue), or both (red). Shown for $|\eta| < 0.8$ (a) and $1.2 < |\eta| < 2.4$ (b)[14]. 60

5.1 Diagram showing the data-flow through the CMS tiered computing system [14]. . 62

5.2 Track reconstruction efficiency (a) for electrons passing high purity requirements as a function of p_T for the barrel, endcap, and transition regions, and (b) the tracking fake rate for those electrons [18]. 65

5.3 (a) The distribution of differences of inter-calibration coefficients for a supermodule exposed to high energy electrons on two occasions a month apart. C_A represents the coefficient measured during the August exposure, and C_S the September exposure. The reproducibility (RMS/$\sqrt{2}$) is measured to be 0.2%. (b) Distribution of inter-calibration coefficient differences for cosmic ray data (C_{cosm}) compared to high energy electron data (C_{beam})for nine SMs [14]. 67

5.4 The effects of calibration for ECAL output due to laser monitoring. Black points show response before calibration to laser light, red points show response after [14]. 67

5.5 (a) The energy resolution as a function of true hadron energy (b) Calibration coefficients as a function of energy to estimate the neutral hadron energy fraction in the HCAL [14]. Results shown are from a χ^2 minimization on simulated events. 68

5.6 Superposition of various dimuon trigger paths on $1.1 fb^{-1}$ of data taken in early 2011 [19]. 69

5.7 Distribution of the output of the multivariate electron estimator used to define GSF electrons. To be identified as an electron by the MVA, $\xi > -0.1$. Data is shown as dots while MC simulation is shown as the solid histogram [20]. 70

5.8 The anti-kt jet clustering algorithm with distance parameter R=1.0 [21] 72

5.9 A b-quark will travel a distance L_{xy} before decaying and creating a secondary vertex. L_{xy} is measured in the plane orthogonal to the beam direction, and the impact parameter, d_0, measures the displacement from the beam line [22]. . . . 73

5.10 $\not\!\!E_T$ distribution in 900 GeV data before (blue histogram) and after (black dots) cleaning [20]. 75

6.1 Comparison of the number of primary vertices (nPV) in data to MC generated samples for the 2012 full dataset of 19fb^{-1} before pileup reweighting (a), and after weights are applied (b). 81

6.2 Particle Flow Isolation values for the electron D-D QCD multi-jet sample (a) and the muon D-D sample (b). 82

6.3 Normalized histograms of gluon-gluon fusion signal(green) W+jets background(blue). Red line shows cut level where 5% of signal is lost 87

6.4 (a) Distribution of $\not\!\!E_{T\phi}$ for both data and MC. Shown are W+jets MC in red and Single Electron data in black (b) The $\not\!\!E_{x,y}$ distributions as a function of nPV. Black and Red distributions show the MC $\not\!\!E_T$ X and Y distributions respectively, while blue and green show the data $\not\!\!E_T$ X and Y. 92

6.5 $\not\!\!E_{T\phi}$ distributions with corrections applied. (a) Shown are W+jets MC in red and Single Electron data in black (b) The $\not\!\!E_{x,y}$ distributions as a function of nPV. Black and Red distributions show the MC $\not\!\!E_T$ X and Y distributions respectively, while blue and green show the data $\not\!\!E_T$ X and Y. 93

6.6 Distribution of Top p_{T} weight for electron events (a) and muon events (b) 95

6.7 Comparison of the Data to MC agreement in the 2-jet bin for $cos(\theta_l)$ in our signal region (a). Note the overestimation of data compared to MC in the low values, and the ovestimation in the high values, leading to a linear trend in the residual plot. (b) Comparison of $cos(\theta_l)$ for events with 1 b-tagged jet. 95

6.8 (a) Corrective weights generated by comparing data and W+jets MC in a 1 b-tag control region. (b) Data to MC agreement for $cos(\theta_l)$ in the signal region after weights have been applied. 96

6.9 Ratio of number of events in the signal region to the events in the antiIso region for each of the six samples. The numbers in parentheses on the Y-axis reflect the number of entries in both regions. The first plot is empty due to a lack of MC statistics that passed our selection for low values of p_{T}. 97

6.10 $\not\!\!E_T$ distribution for all 13 η bins. The distribution of QCD (solid red) and W+jets (solid green) has been fit to the data (dots) while keeping all other processes (blue) fixed to their SM expectation. The last plot shows the χ^2/NDF of the fits. . . . 98

6.11 Left: Scale factors for QCD as a function of absolute lepton η. Right: Ratio of measured yield of W+jets events found from fitting to data to the SM expected yield of W+jets. The green band indicates the error on the expected SM W+jets cross section. 99

6.12 $\not\!\!E_T$ distributions for all η bins in the 1 jet control region electron QCD fits. . . . 99

6.13 Electron QCD scale factors (a) as a function of absolute lepton η in the 1 jet control region. (b) Ratio of the measured number of W+jets events to the SM expected number in the 1 jet control region. 100

6.14 Muon QCD scale factors (a) as a function of absolute lepton η in the 1 jet control region. (b) Ratio of the measured number of W+jets events to the SM expected number in the 1 jet control region. 101

6.15 Number of Primary Vertex distribution comparisons (a,c) and associated weights (b,c) for data-driven QCD sample for electrons(a,b) and muons (c,d). 102

7.1 Example of a decision tree found in the kinematic BDT analysis. Starting with the root node (very top green box), a sequence of binary splits using the discriminating variables provided as input is applied to the data. Each split uses the variable that at this node gives the best separation between signal and background when being cut on. The same variable may thus be used at several nodes (as seen here with the variable jet2dRLep), while others might not be used at all. The leaf nodes at the bottom end of the tree are shown in blue for signal and red for background, depending on the majority of events that end up in the respective nodes. 107

7.2 Normalized histograms of ggH M125 signal (black) and W+jets (green). Also shown are the two FOMs calculated from the distribution. The variables shown are lepton η (a), ΔR(lep,jet2) (b) , $M_{l\nu jj}$ (c), and Cos(Θ_{lep}) (d). 111

7.3 Nominal input histogram (a) for lepton p_{T} showing signal (black) and background (green), and the corresponding CDF (b) for lepton p_{T} 112

7.4 Defining the angular variables in a $H \rightarrow WW \rightarrow l\nu jj$ decay process. Reprinted from [23] 114

7.5 Example output plots from a BDT Training. A comparison of the training and test samples with K-S results (a), correlation matrix of background (b) and signal (c) input variables. 116

7.6 ROC output from example BDT training (a) and comparison of ROCs from multiple BDT trainings of the same jet bin (b). As shown in (b), the samples with 11 variables provided the greatest discrimination power. 117

7.7 Overlay of ROC curves from a BDT trained with 5 different MaxDepth Values (a). BDTs trained with a MaxDepth of 3(b) and 9 (c). Note the K-S test results showing good agreement (values $>>$ 1) for MaxDepth = 3 and severe overtraining (values $\sim$ 0) for MaxDepth = 9. 119

7.8 Overlay of ROC curves from a BDT trained with different boost factors β, and different numbers of trees. In both cases the 'nominal' value shown in dark blue performed best. 119

7.9 Data-MC comparison plots for the input variables used for the 2-jet BDT training. Shown here are ΔR(lep,jj) (a), $\Delta\phi$(met,jet) (b), $\Delta\phi$(jet,jet) (c), Cos(Θ_{WH}) (d), and Cos(Θ_{lep}) (e). 120

7.10 Data-MC comparison plots for the input variables used for the 2-jet BDT training. Shown here are (a), m_T (b), $P_{T_{lnujj}}$ (c), ht (d), ΔR(lep,jet1) (e), and ΔR(lep,jet2) (f). 121

7.11 Data-MC comparison plots for the input variables used for the 3-jet BDT training. Shown here are $p_{\mathrm{T}\ lep}$ (a), Charge$\times\eta_{lep}$ (b), M_{lvjj} (c), ht (d), and ΔR(lep,jet2) (e).121

7.12 Data-MC comparison plots for the input variables used for the 3-jet BDT training. Shown here are ΔR(lep,jj) (a), min$\Delta\phi$(lep,jet) (b), $\Delta\eta$(jet,jet) (c), $\Delta\phi$(met,jet) (d), Cos(Θ_{lep}) (e), and ΔR(lep,jet3) (f). 122

7.13 Data-MC comparison plots for the input variables used for the 3-jet BDT training. Shown here are Cos(Θ_{jet}) (a), and Cos(Θ_{WH}) (b). 122

7.14 Data-MC comparison plots for the input variables used for the $\geq$ 4-jet BDT training. Shown here are Charge$\times\eta_{lep}$ (a), M_{lvjj} (b), ht (c), ΔR(lep,jet2) (d), and ΔR(lep,jet3) (e). 123

7.15 Data-MC comparison plots for the input variables used for the $\geq$ 4-jet BDT training. Shown here are $\Delta\phi$(met,jet) (a), and $\Delta\phi$(met,lep) (b). 123

7.16 (a) Output discriminant from the training of the 2-jet BDT. (b) Data-MC comparison plot for 2-jet bin BDT discriminant. 124

7.17 (a) Output discriminant from the training of the 3-jet BDT. (b) Data-MC comparison plot for 3-jet bin BDT discriminant. 124

7.18 (a) Output discriminant from the training of the $\geq$ 4-jet BDT. (b) Data-MC comparison plot for $\geq$ 4-jet bin BDT discriminant. 125

7.19 Output discriminant from the training of the 2-jet BDT and the output for samples with JES scaled up and down for ggH signal sample (a) and W+jets sample (b). 127

7.20 Output discriminant from the training of the 2-jet BDT and the output for samples with QCD η weights scaled up and down for electron sample (a) and muon sample (b). 130

7.21 Shape uncertainty histograms for variations of the Cos(θ_{lep}) weights for (a) 2-jets, (b) 3-jets, and (c) $\geq$ 4-jets. 131

8.1 BDT Discriminant value distribution for the 2-jet (a) and 3-jet(b) categories. Each plot shows the full BDT distribution (top) with red hatched lines showing total uncertainty on backgrounds, and a ratio of data / simulation (bottom) with uncertainty shown as a gray band. 139

8.2 BDT Discriminant value distribution for the 2-jet (a) and 3-jet(b) categories in the b-tag control region. Each plot shows the full BDT distribution (top) with red hatched lines showing total uncertainty on backgrounds, and a ratio of data / simulation (bottom) with uncertainty shown as a gray band. 140

List of Tables

2.1 The quantum representation of fermions in the standard model and their associated electric charge and Hypercharge(Y). All fermions except neutrinos interact with the electromagnetic force. All left-handed doublets interact with the weak force. 11

3.1 Relationship between kinetic energy and speed of a proton in the CERN accelerator complex, reproduced from [24]. 35

6.1 The datasets analyzed for this analysis. 79

6.2 List of Signal datasets and cross sections 79

6.3 List of background MC datasets and cross sections used for normalization. 79

6.4 List of triggers 83

6.5 The requirements for a primary vertex 83

6.6 Tight and loose electron definitions 84

6.7 Tight and loose muon definitions 86

6.8 Jet definition 87

6.9 Jet Energy Resolution (JER) correction scale factors by η 91

6.10 List of parameter values for the $\not{E}_{T\phi}$ corrections. 92

6.11 Jet Energy Resolution (JER) correction scale factors by η 97

6.12 List of W+jets Scale Factor values and QCD xSec values from $\not{E}_T$ fit. 103

6.13 Expected event yield normalized to cross sections and luminosity. Top section shows background processes with all diboson processes combined as well as all single top processes combined. The middle section shows contributions from all $H \to WW$ processes that are considered as signal. Bottom section shows other Higgs processes that are not part of our signal that could contaminate our final state ('Volunteer Signal'). 104

6.14 Expected event yield normalized to cross sections and luminosity. Top section shows background processes with all diboson processes combined as well as all single top processes combined. The middle section shows contributions from all $H \to WW$ processes that are considered as signal. Bottom section shows other Higgs processes that are not part of our signal that could contaminate our final state ('Volunteer Signal'). 104

6.15 Expected event yield normalized to total yield. Background samples are normalized to total background, while Higgs samples are normalized to total $H \to WW$ contribution. Dominant background highlighted for each jet bin, here W+jets is dominant for the 2 and 3 jet bin but $t\bar{t}$ is dominant for ≥ 4 jets. 105

6.16 Expected event yield normalized to total yield. Background samples are normalized to total background, while Higgs samples are normalized to total $H \to WW$ contribution. Dominant background highlighted for each jet bin, here W+jets is dominant for all jet bins. 105

7.1 List of Variables Considered for MVA 113

7.2 Global scale factors for BDT signal inputs, by jet bin. 115

7.3 BDT Variable choices optimized by jet bin 117

7.4 Summary of the systematic uncertainties considered on the inputs to the limit calculation. 126

7.5 Systematic Uncertainty Due to JES shift 127

7.6 List of samples used for W+jets systematic uncertainty shape. 128

7.7 Systematic Uncertainty Due to Pileup Weights 128

7.8 Systematic Uncertainty Due to CSV Weights 129

7.9 Cross section uncertainties used for the limit settings 129

8.1 Shows expected event yield normalized to cross sections and luminosity for all signal and background processes after the optimized BDT discriminant cut of > 0.24. Uncertainties shown are statistical only. Data yields are also shown for direct comparison. 136

8.2 Results of the expected and observed a priori limit using the Asymptotic CL_S method for a counting experiment using a cut of BDT > 0.24. 136

8.3 Shows expected event yield normalized to cross sections and luminosity for all signal and background processes. Uncertainties shown are statistical only. Data yields are also shown for direct comparison. 137

8.4 Results of the expected a priori limit using the Asymptotic CL_S method for the kinematic BDT with all systematic uncertainties included. Expected results use the Asimov dataset hypothesis for calculation. 139

Acknowledgements

This thesis would have been impossible without the massive effort put forth by everyone working to run and maintain the Large Hadron Collider (LHC). I would like to acknowledge the amazing performace of the LHC, and the men and women who worked to build it and keep it running. I would also like to thank all the members of the Compact Muon Solenoid (CMS) collaboration, as an enormous amount of time and effort has gone into making CMS such a high precision machine. This thesis would not have been possible without the data measured and recorded by CMS, and is but the final chain in a collaborative effort.

I would like to thank the UVa HEP group for being a great place to work and learn, both in Virginia and while at CERN. Thanks to John Wood, Brian Francis, and Thomas Anderson for discussions, debates, and help navigating these years of school. A special thanks to Dr. Sarah Boutle for being such an amazing mentor throughout the years of my analysis. I can't begin to count the number of things you taught me about physics, the LHC, and how it's possible to be good scientist and still balance that with a life outside of physics.

Working on the H→WW analysis, we had a small group, but having others to collaborate with was immeasurably helpful. My thanks goes out to Alexx Perloff and Professor Ricardo Eusebi of Texas A&M University (TAMU) for their help and collaboration. Of course, no acknowledgement would be complete without the mention of my advisor Professor Chris Neu. Thank you for teaching and supporting me over these long years, giving me the opportunity to travel around the country and to CERN to participate in one of the greatest scientific collaborations there has ever been.

Lastly, I would like to thank my family for their ever present love and support. Thank you Mom for believing in and supporting me through this endeavor. Thank you to Bob and Judy Kaplan for all the visits, holidays, and meals, you have made me feel like part of the family from the beginning. To my lovely fiancée Jen. Thank you for being there with and for me these last years. There's too much to put into words here, but I'm so glad that you were there with me, and I love you.

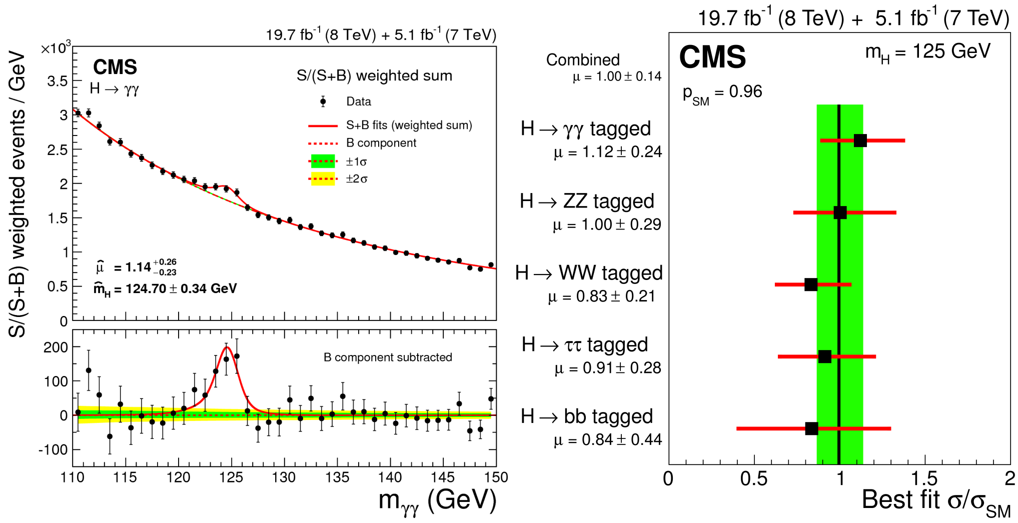
Chapter 1

Introduction

One of the biggest moments in particle physics came on July 4th, 2012 where Compact Muon Solenoid (CMS) and A Toroidal LHC Apparatus (ATLAS) announced the discovery of a new boson of mass $\simeq$125 GeV [25] [26]. All measurements so far have shown this boson to be consistent with expectations for the spin zero mediator of the Higgs field, the Higgs boson. The Higgs boson couples to all particles with mass, meaning it has a probability to decay into many of the SM particles that we can measure. This gives us the opportunity to observe the Higgs in multiple decay channels and improve on the measurement of any individual channel.

The mass of the Higgs boson was measured by CMS to be 125.7 ± 0.3 (stat) ± 0.3 (syst)[27] GeV by six decay modes: $H \to \gamma\gamma$, $H \to \tau\tau$, $H \to bb$, $H \to WW$, $H \to ZZ$, and $H \to \mu\mu$. Figure 1.1 shows the mass peak seen in the $\gamma\gamma$ channel, as well as the combined measurements of the five channels mentioned above.

The search for the Higgs boson in the semi-leptonic decay channel, $H \to WW \to$lνqq, is performed with the CMS detector, a modern general purpose particle detector located at European Center for Nuclear Research (CERN). This detector is capable of identifying photons, electrons, muons, τ leptons, and quark jets. In addition, its hermetic design and high efficiency in identifying and reconstructing all of the particles produced in the collisions makes it good at identifying a momentum imbalance in an event. Such an imbalance arises when a particle escapes detection (usually signifying the presence of a neutrino in the event which CMS cannot track) and can be measured with good precision in the direction transverse to the beam line (Σ momentum should be zero here as the proton beams collide head on). One of the central features of CMS is its namesake solenoid which provides a 3.8 Tesla magnetic field uniformly across the detector. This field bends the charged particles that are produced during collisions, allowing the particles momentum to be accurately measured. Combining these tracks with the energy information gathered by the Electromagnetic Calorimeter (ECAL) and Hadronic Calorimeter



(a) $H \rightarrow \gamma\gamma$ channel [28] (b) σ/σ_{SM} for 5 CMS decay channels [27]

Figure 1.1: The CMS experiment has observed a new boson at m$\sim$125 GeV/c^2

(HCAL) we are able to fully reconstruct the particles generated in each event.

Beyond directly detecting the decay products of the Higgs (as in $H \rightarrow \gamma\gamma$), the Higgs can decay to have particles that have short lifetimes (ex. tau, top quark, weak bosons). The decay of these particles leads to a plethora of final states that can be observed by reconstructing all of the final state particles in an event. This thesis presents the search for one of the final states, $H \rightarrow WW \rightarrow \ell\nu q\bar{q}$, in which the Higgs decays into two W bosons where one W decays leptonically and the other decays hadronically. This final state signature of one lepton, two quark jets, and a neutrino (observed as a missing energy) is a valuable addition to the CMS Higgs measurement, as it has been searched for $M_h > 2\ M_W$ [29] but not at a mass $M_h \simeq 125$ GeV.

Searching for a Higgs of mass $M_h \simeq 125$ GeV is not easy though, as it requires that at least one of the W bosons to be virtual. This means that the boson is created with an 'off shell' energy where $M_W \neq$80 GeV. Thus, reconstructing the W mass correctly is not always possible, making our signal harder to distinguish. For this analysis, the largest background is SM process of W+jets, which will directly mimic our signal when the W in that event decays leptonically. To look for our signal a Multi-Variate Analysis (MVA) technique is used to attempt to separate our $H \rightarrow\ WW$ signal from the W+jets background. The MVA technique used is a Boosted Decision Tree (BDT), which when provided with input about the event generates a single discriminant that describes the event as signal-like or background-like. This discriminant is then used for signal extraction and ultimately to place an upper limit on the production cross section of $H \rightarrow\ WW$.

This thesis is organized as follows. Chapter 2 describes the physics motivation for this SM

Higgs search as well as the theoretical framework that the SM is built on. Following this, a description of the LHC is given in chapter 3 and CMS in chapter 4. Chapter 5 describes the reconstruction of physics opbjects from signals in the detector. In Chapter 6 I describe the method for selection events and modeling the background using Monte-Carlo (MC) simulation. Chapter 7 shows the analysis of this data along with the description of the MVA method used. Finally, an interpretation of the results is shown and the thesis is concluded in chapter 9.

Chapter 2

Theoretical Background: The Standard Model

The Standard Model (SM) of particle physics concisely describes a unified representation of our knowledge of particle interactions. The SM framework was formed in the 1960s, combining the work of many different physicists. Electromagnetism was described in 1930 by Herman Weyl [30] as that of a local symmetry represented by the Lie group U(1). Then, in 1954, Yang and Mills constructed a gauge theory based on a three dimensional group SU(2) [31] that was used to describe the electroweak interaction. By the 1970s a model of the strong interaction had been added, represented by an SU(3) group describing the color interactions. Together, these forces were unified to create a representation of particle physics described by the gauge group

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \tag{2.1}$$

where $SU(3)_C$ describes the quark QCD interactions, $SU(2)_L$ describes the weak interactions among quarks and leptons, and $U(1)_Y$ describes the electromagnetic interaction. The SM is a Quantum Field Theory (QFT) that describes all of the known particles, though there are still a handful of experimental and theoretical shortcomings. Thee will be described in more detail below.

2.1 The Standard Model

The SM is comprised of twelve types of fermions and 5 types of bosons shown in Figure 2.1. The fermions are further broken down into six leptons and six quarks. The bosons are separated into the force-carrying particles ($W^{\pm}$,Z, γ, gluons) and the recently discovered Higgs Boson (H).

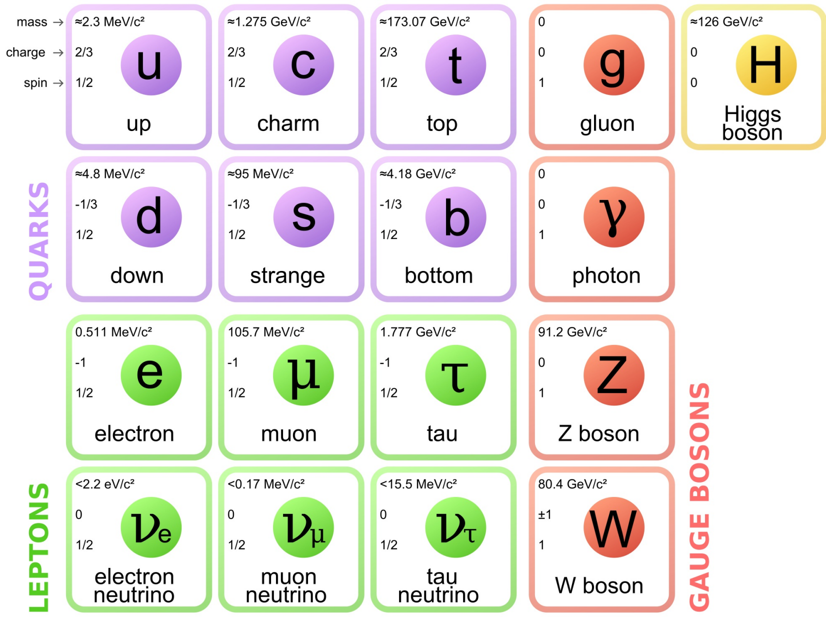


Figure 2.1: The Standard Model of elementary particles [1]. Six quarks (shown in purple) and six leptons (shown in green) comprise the fermionic component of the SM, while the force mediating gauge bosons (shown in red) comprise the bosonic component. Additionally, the Higgs Boson (the result of electroweak symmetry breaking) is shown in yellow.

All fermions are spin 1/2 particles, meaning they have an intrinsic angular momentum of $\hbar/2$. They follow Fermi-Dirac statistics and obey the Pauli exclusion principle. Under these rules no two fermions may simultaneously exist in the identical quantum state as one another. Additionally, for every fermion there is an anti-fermion with identical mass but opposite quantum numbers.

Leptons can be broken down into three generations, beginning with the first generation comprised of the electron and its associated neutrino, e and ν_e. Although the standard model predicts neutrinos to be massless, it has been shown experimentally that this is not true. Though we have yet to measure them fully, upper bounds have been placed on neutrino mass and are shown in figure 2.1. The second and third generations of leptons are composed of heavier versions of the electron, the particles known as the muon μ and tau τ leptons as well as their associated neutrinos. Each lepton generation has an associated quantum number, known as lepton number, which must be conserved in SM interactions. The first generation leptons have $L_e = +1$, while having $L_\mu = L_\tau = 0$. The second and third generation similarly have their associated lepton number =+1, while the others are zero. Antiparticles for each lepton are assigned lepton numbers of opposite sign. Lepton number conservation has been shown to be violated by neutrino oscillations [32], but in this case total lepton number ΣL_l is conserved.

Quarks are also separated into generations of hierarchical mass. The first generation of quarks is composed of the up u and down d quarks. The second and third generation of quarks are made up of the strange s and charm c quarks, and the top t and bottom b quarks respectively. Quarks have fractional charge, with the up, charm, and top quarks having charge $+2/3e$, while the down, strange, and bottom quarks have charge -1/3e. Bare quarks have never been observed in nature, so quarks are observed in bound states known as baryons or mesons. Baryons are bound states of three quarks (or anti-quarks), and mesons are bound states of two quarks. Together, baryons and mesons are collectively referred to as hadrons. Like leptons, baryons have an associated quantum number called Baryon Number B, with all quarks assigned a baryon number of +1/3. Conservation of baryon number means that the SM only permits the creation and destruction of quark-antiquark pairs. In addition to electric charge, quarks possess an additional charge known as color charge. This is commonly described as being red, green, or blue (as well as the associated anti-color charges of anti-red, anti-green, and anti-blue). Only colorless bound states have been observed, which can be seen in baryons as the combination of a red, blue, and green quark, or in mesons as the combination as a color anti-color quark pair. The additional component of color charge allows the quarks to interact via the strong interaction as well as via the weak and electromagnetic interactions.

The bosons described in the SM are known as force carriers, meaning that they mediate interactions involving the different forces. The electromagnetic interactions are mediated by photons, represented as γ. Weak interactions can involve charged interactions (mediated by $W^{\pm}$ bosons), or neutral interactions (mediated by the Z boson). Gluons, represented by g mediate strong interactions.

Additionally, the Higgs boson (h) forms the last piece of the Standard Model pantheon. Its inclusion in the SM is a result of electroweak symmetry breaking and will be addressed in section 2.5.

2.2 Quantum Electrodynamics

Quantum Electrodynamics (QED) is a quantum field theory (QFT) that describes the electromagnetic interactions. In a QFT, particles such as leptons and quarks, are represented by fields. Fields are described by a Lagrangian density, denoted by $\mathcal{L}$, but as I will be describing field theories I will hereafter simply refer to $\mathcal{L}$ as the Lagrangian. QED describes particles with spin 1/2, which are represented by a Dirac Lagrangian given by

$$\mathcal{L} = i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\bar{\psi}\psi \tag{2.2}$$

where ψ is a four-component field knows as a Dirac spinor, γ^{μ} are the four Dirac gamma matrices, and $\bar{\psi} \equiv \psi^{\dagger}\gamma^{0}$. In order for our theory to correctly describe our real world particles, it must be invariant under global and local gauge transformations. Let us first look at a global U(1) transformation where

$$\psi \rightarrow \psi' = e^{-i\alpha}\psi \tag{2.3}$$

By replacing this into equation 2.2, we see that $\mathcal{L} \rightarrow \mathcal{L}' = \mathcal{L}$, and thus our Dirac Lagrangian is invariant under global transformations. For local transformations we let $\alpha \rightarrow \alpha(\mathrm{x})$. Under this transformation

$$\mathcal{L} \rightarrow \mathcal{L} - (\partial_\mu \alpha)\bar{\psi}\gamma^\mu \psi \tag{2.4}$$

where it is evident that the Lagrangian is not invariant under such a transformation. In order to restore this invariance we replace the partial derivative in the Lagrangian with a newly defined covariant derivative:

$$D_\mu = \partial_\mu + ieA_\mu \tag{2.5}$$

where A_μ is a new gauge field representing the photon that transforms as

$$A_\mu(x) \rightarrow A_\mu(x) - \frac{1}{e}\partial_\mu \alpha(x) \tag{2.6}$$

Now, by replacing the partial derivative with this covariant derivative we can see that the covariant derivative transforms in the same way that $\psi(\mathrm{x})$ transforms which will preserve the local gauge invariance. When transforming this field we see that $D_\mu\psi \rightarrow (D_\mu)' = e^{-i\alpha}D_\mu\psi$. This results in equation 2.2 taking the locally gauge invariant form:

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} \tag{2.7}$$

where

$$F^{\mu\nu} = (\partial^\mu A^\nu - \partial^\nu A^\mu) \tag{2.8}$$

is the electromagnetic field strength tensor. It is important to note that in this Lagrangian there is no $m^2 A_\mu A^\mu$ term, which would represent the mass of the gauge field. This is good because we have identified the gauge field here as the photon which we know to be massless. It is evident though that this process of introducing a gauge field that transforms like the wave-function will only work with massless bosons, a problem we will discuss later. Lagrangian 2.7 now describes lepton interaction, and can be generalized to include all leptons by letting $\psi \rightarrow \psi_i$ and summing over all leptons.

2.3 Electro-Weak Interaction

In the standard model the electromagnetic and weak interactions are unified into a single electroweak theory [33]. The work of extending the symmetry described in section 2.2 to higher order models was accomplished by Yang and Mills in 1954 [31]. With more dimensions, instead of varying a local function α(x) you instead need a matrix (or matrices) to describe the dynamics. This generalization is known as Non-abelian gauge theory and to understand it we will start first with a fermionic doublet representing an SU(2) symmetry, then show how it combines into the description of the electroweak interactions represented by a local SU(2)$_L\times$U(1)$_Y$ gauge symmetry.

2.3.1 Yang-Mills Theory

First let's start with a doublet of Dirac fields,

$$\psi = \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \end{pmatrix} \tag{2.9}$$

where this doublet will transform under an arbitrary three dimensional rotation via the transformation [34]:

$$\psi \rightarrow \ \exp\langle \ i\alpha^i \frac{\sigma_i}{2} \rangle \psi \tag{2.10}$$

where σ_i are the Pauli sigma matrices

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \ , \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \ , \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.11}$$

whose products satisfy the identity

$$\sigma^i \sigma^j = \delta^{ij} + i\epsilon^{ijk}\sigma^k \text{ , where } \epsilon^{0123} = +1 \tag{2.12}$$

and ϵ is a totally antisymmetric tensor. Generalizing equation 2.10 from a global symmetry to a locally symmetric transformation, we impose the condition that the Lagrangian be invariant under any arbitrary transformation α^i(x). Now,

$$\psi(x) \rightarrow V(x)\psi(x) \text{ , where V(x) = } \exp(i\alpha^i(x)\frac{\sigma^i}{2}) \tag{2.13}$$

so in order to preserve local gauge invariance we must introduce three vector fields A^i_μ(x), where i = 1,2,3. As in the EM theory transformation, we similarly transform the covariant derivative

$$D_\mu = \partial_\mu - igA^i_\mu \frac{\sigma^i}{2} \tag{2.14}$$

This means that the fields A^i_μ(x) must transform

$$A^i_\mu(x)\frac{\sigma^i}{2} \to V(x)\left(A^i_\mu(x)\frac{\sigma^i}{2} + \frac{i}{g}\partial_\mu\right)V^\dagger(x) \tag{2.15}$$

Since the Pauli matrices do not commute this is not a simple calculation. For infinitesimal transformations, we can expand V(x) to first order in α and obtain a (slightly) easier relationship:

$$A^i_\mu \frac{\sigma^i}{2} \to A^i_\mu \frac{\sigma^i}{2} + \frac{1}{g}(\partial_\mu \alpha^i)\frac{\sigma^i}{2} + i\left[\alpha^i \frac{\sigma i}{2}, A^i_\mu \frac{\sigma^i}{2}\right] + ... \tag{2.16}$$

By combining the transformation in 2.16 and the infinitesimal fermion transformation, we find that the covariant derivative transforms as

$$D_\mu \psi \to \left(1 + i\alpha^i \frac{\sigma^i}{2}\right) D_\mu \psi \tag{2.17}$$

which leads to a new form for the field strength tensor $F^i_{\mu\nu}$:

$$F^i_{\mu\nu} = \partial_\mu A^i_\nu - \partial_\nu A^i_\mu + g\epsilon^{ijk} A^j_\mu A^k_\nu \tag{2.18}$$

Finally, we can put this all together to form the Yang-Mills Lagrangian

$$\mathcal{L} = -\frac{1}{4}(F^i_{\mu\nu})^2 + \bar{\psi}(i\gamma^\mu \partial_\mu - igA^i_\mu \frac{\sigma^i}{2})\psi \tag{2.19}$$

2.3.2 Glashow-Weinberg-Salam model for EW Interactions

Now that we have the mathematical framework of Yang-Mills theory, we can use it to obtain the Weinberg-Salam model of Electro-Weak (EW) interactions [33], a SU$(2)_L \times$U$(1)_Y$ gauge theory. Figure 2.2 shows a classic example of weak interaction, neutron decay. By using some important information taken from experiments [35], we know that the particles in the standard model can be represented as either left handed (spin of the particle is aligned with the direction of motion) or right handed (anti-aligned). In this model the left handed components are doublets which participate in the weak interaction, and the right handed components are singlets which only interact via the electromagnetic interaction. There are no right handed neutrinos observed in the SM, so we get

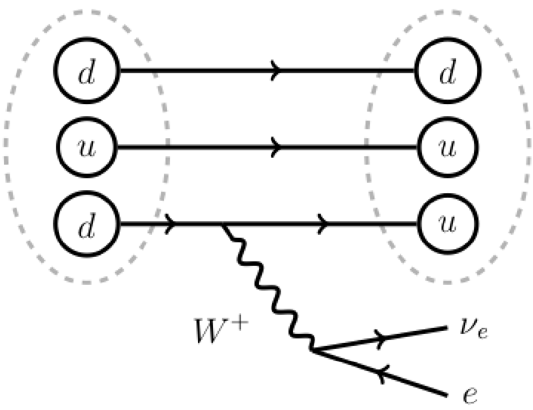


Figure 2.2: Feynman diagram for a neutron decaying into a proton via the weak interaction. In this case, a down type quark radiates a W boson and becomes an up type quark, while the W decays leptonically.

$$L_e = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}, R = e_R \tag{2.20}$$

First, looking only at the kinetic energy term of the Lagrangian we get

$$\mathcal{L}_{KE} = L_e^\dagger \sigma^\mu i\partial_\mu L_e + R^\dagger \sigma^\mu i\partial_\mu R \tag{2.21}$$

In order for this to remain invariant under global SU(2)$_L$×U(1)$_Y$ transformation we have

$$L \to L' = e^{i\theta} U L \tag{2.22}$$

$$R \to R' = e^{2i\theta} R \tag{2.23}$$

where U = $e^{-i\alpha^k \sigma^k}$, and θ and α^k are real numbers. Like before, this is not invariant under local transformations on its own, so we will introduce a U(1) gauge field B_μ(x) and three SU(2) gauge fields W_μ(x) = W_μ^k(x)σ_k. These fields transform as

$$B_\mu(x) \to B'_\mu(x) = B_\mu(x) + \frac{2}{g_1}\partial_\mu \theta(x) \tag{2.24}$$

$$W_\mu(x) \to W'_\mu(x) = U(x) W_\mu(x) U^\dagger(x) + \frac{2i}{g_2}\left(\partial_\mu U(x)\right) U^\dagger(x), \tag{2.25}$$

where g_1 and g_2 are dimensional parameters of the theory. Transforming the covariant derivative appropriately then gives

$$D_\mu L_e = \left(\partial_\mu + i\frac{g_1}{2} Y B_\mu + i\frac{g_2}{2} Y W_\mu\right) L_e \tag{2.26}$$

$$D_\mu R = \left(\partial_\mu + i\frac{g_1}{2} Y B_\mu\right) R \tag{2.27}$$

where Y is the hypercharge operator. Hypercharge is defined to be Y_L = -1 for the left-handed doublet and Y_R = -2 for the right handed singlet. Table 2.1 shows the representation of the Standard Model gauge fields with their associated electric charge and hypercharge.

The full Lagrangian can then be defined by combining the kinetic terms with gauge interaction terms

$$\begin{aligned} \mathcal{L} &= \mathcal{L}_{KE} + \mathcal{L}_{gauge} \\ &= L_e^\dagger \sigma^\mu i\partial_\mu L_e + R^\dagger \sigma^\mu i\partial_\mu R - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \sum_{i=1}^{3} \frac{1}{4} W^i_{\mu\nu} W^{i\mu\nu} \end{aligned} \tag{2.28}$$

where $B_{\mu\nu} = \partial_\mu B_\nu$ - $\partial_\nu B_\mu$ and $W_{\mu\nu} = \left[\partial_\mu + (i\frac{g_2}{2}) W_\mu\right] W_\nu$ - $\left[\partial_\nu + (i\frac{g_2}{2}) W_\nu\right] W_\mu$ representing the field strength tensors. The Lagrangian is now invariant as we have shown it, but it is still lacking any mass terms (as they would break this invariance). The mass terms of the Lagrangian will be addressed later in section 2.5.

Field Left-handed quark doublet	Notation $Q_L = \binom{u_L}{d_L}$	Hypercharge $\frac{1}{3}$	Electric Charge $\binom{2/3}{-1/3}$
Right-handed up-type quark singlet	u_R	$\frac{4}{3}$	$\frac{2}{3}$
Right-handed down-type quark singlet	d_R	-$\frac{2}{3}$	-$\frac{1}{3}$
Left-handed lepton doublet	$L_L = \binom{\nu_L}{e_L}$	-1	$\binom{0}{-1}$
Right-handed charged lepton singlet	e_R	-2	-1

Table 2.1: The quantum representation of fermions in the standard model and their associated electric charge and Hypercharge(Y). All fermions except neutrinos interact with the electromagnetic force. All left-handed doublets interact with the weak force.

The physical gauge bosons can be associated with combinations of these B and W fields. The W_1 and W_2 fields are electrically charged, while the W_3 and B gauge fields are electrically neutral. They combine linearly to become the physical bosons we observe:

$$\begin{aligned} W^\pm_\mu &= \frac{W^1_\mu \mp iW^2_\mu}{\sqrt{2}} \\ Z_\mu &= \frac{g_1 W^3_\mu - g_2 B_\mu}{\sqrt{g_1^2 + g_2^2}} &&= W^3_\mu cos(\theta_W) - B_\mu sin(\theta_W) \\ A_\mu &= \frac{g_1 W^3_\mu + g_2 B_\mu}{\sqrt{g_1^2 + g_2^2}} &&= W^3_\mu sin(\theta_W) - B_\mu cos(\theta_W) \end{aligned} \tag{2.29}$$

where θ_W is the Weinberg angle defined by $\sin(\theta_W) = g_1/\sqrt{g_1^2 + g_2^2}$. The interactions shown in equation 2.28 only couple the $W^\pm$ to the left handed doublets, but allows coupling of the Z and photon (A) to both the left and right handed components.

Quarks are added to the Lagrangian in a similar manner by placing the left hand components of the up and down quarks into an SU(2) doublet, and the right handed components in separate singlets,

$$Q_u = \begin{pmatrix} u_L \\ d_L \end{pmatrix}, u_R, d_R \tag{2.30}$$

By analogy to the leptons we can construct the Lagrangian

$$\mathcal{L}_{KE}^{quark} = Q_u^\dagger \sigma^\mu i D_\mu Q_u + u_R^\dagger \sigma^\mu i D_\mu u_R + d_R^\dagger \sigma^\mu i D_\mu d_R \tag{2.31}$$

and by adding a term like equation 2.31 for each set of quarks to the Lagrangian for the leptons

$$\mathcal{L}^{EW} = \mathcal{L}_{gauge} + \mathcal{L}_{KE}^{quark} + \mathcal{L}_{KE}^{lep} \tag{2.32}$$

This leads to a form which again only couples the W bosons to the left-handed quarks while the Z and photon couple to both left and right-handed components.

2.4 Strong Interaction

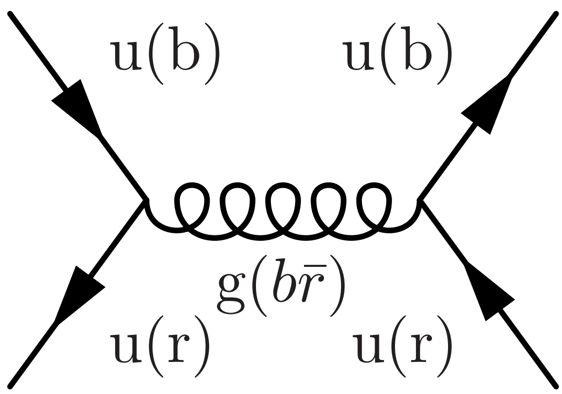


Figure 2.3: Example of a gluon exchange between quarks resulting in a change of color charge in the quarks. Quarks carry either a positive or negative color charge while gluons carry one component of both positive and negative color charge.

Quantum Chromodynamics (QCD) is the theory of the strong interactions that takes place between quarks and is represented by a local SU(3)$_C$ gauge symmetry. The C stands for color, as quarks possess an additional property known as color charge that can come in three varieties commonly called red, green, and blue. To model this, each quark is represented in a color triplet

$$q_u = \begin{pmatrix} u_r \\ u_g \\ u_b \end{pmatrix} \quad . \tag{2.33}$$

Under this representation we can define the invariant QCD quark Lagrangian to be

$$\mathcal{L}_{quark} = \sum_{i=1}^{6} \bar{q}_i i\gamma^\mu \partial_\mu q_i, \tag{2.34}$$

where q_i represents any of the 6 quark flavors. As before, we then check the invariance of $\mathcal{L}$ under the transformation $q_i \rightarrow q_i' = \mathrm{U}q_i$ where $\mathrm{U} = e^{i\alpha^a\lambda^a}$. In the case of SU(3) we now need a 3x3 matrix λ^a to describe the transformation of the quark triplet. For SU(3) these are known as the Gell-Mann matrices, of which there are 8. Under this transformation, to preserve invariance, we must introduce eight gauge fields (G_μ) and an appropriately transforming covariant derivative:

$$\begin{aligned} G^a_\mu &\rightarrow G^a_\mu + \frac{i}{g}\partial_\mu \alpha^a + f^{abc} G^b_\mu \alpha^c \\ D_\mu &= \partial_\mu + igG^a_\mu \lambda^a \end{aligned} \tag{2.35}$$

where f^{abc} is the structure constant for SU(3) that obeys the commutation relationship $[\lambda^a,\lambda^b]$ $=\mathrm{i}f^{abc}\lambda^c$. Then, the field strength tensor for QCD is defined as

$$F^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + f^{abc} G^b_\mu G^c_\nu \tag{2.36}$$

which finally leads us to the QCD Lagrangian

$$\mathcal{L}_{quark} = -\frac{1}{4} F^a_{\mu\nu} F^{a\mu\nu} + \bar{q_u}\left(i\gamma^\mu \partial_\mu - igG^a_\mu \lambda^a\right) q_u. \tag{2.37}$$

We have shown that the mathematical framework of QCD is sound, but some insight into the history of its theoretical development will help ground this understanding. The theory of hadron interactions was developed in 1964 by Gell-Mann [36] and Zweig [37], in which Gell-Mann named the fundamental particles which make up baryons and mesons to be quarks. This model included only three quarks: the up, down and strange quarks. Additionally, the existence of only two quark (hadrons) or three quark (baryons) particles appeared to violate the Pauli exclusion principle requiring that no two fermions can occupy the same quantum state. This problem was solved by the introduction of color charge by Greenberg [38], giving quarks an additional quantum number and allowing all stable hadrons to be color neutral.

This requirement can be fulfilled in two ways: combining equal parts of each color in a qqq combination, or combining a color anti-color pair in a $q\bar{q}$ combination. The combinations were introduced before as baryons (three quark particles) and mesons (two quark particles). Each quark has an associated color charge, so in order to conserve color each gluon (represented by the eight gauge fields introduced above as G_μ) must contain two color charges. With three colors and three anti-colors we would expect 9 combinations and thus 9 gluons, but a ninth state would represent a gluon singlet state which has not been observed [39].

2.5 Higgs Mechanism

Now that we have shown that gauge theories can describe the interactions of the particles in the standard model, we need to address the issue of mass. The Lagrangians for the GWS Electro-Weak (EW) theory and the strong interaction both skirt the problem of mass, and in fact require the gauge particles (W/Z bosons, gluons) to be massless. We know from experiments at European Center for Nuclear Research (CERN) that the W [40] and Z [41] bosons have a large mass, so we need to find a way to correctly describe the mass of these particles in our theory. The Higgs Mechanism allows us to generate mass terms for these particles while maintaining gauge invariance [33] [35] [42].

In order to generate these masses six physicists (in three separate groups) developed what we now call the Higgs Mechanism, but should more correctly be referred to as the Englert-Brout-Higgs-Guralnik-Hagen-Kibble mechanism. They postulated that a complex scalar field ϕ existed, which is represented by the complex scalar doublet

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{2.38}$$

This scalar field will have a Lagrangian of the form

$$\mathcal{L}_{higgs} = (D_\mu \phi)^\dagger (D_\mu \phi) - V(\phi), \tag{2.39}$$

where

$$V(\phi) = \mu^2 |\phi^\dagger \phi| + \lambda \left(|\phi^\dagger \phi|\right)^2 \tag{2.40}$$

and

$$D_\mu = \partial_\mu + i\frac{g_1}{2}\tau \cdot W_\mu + i\frac{g_2}{2} B_\mu Y \tag{2.41}$$

with $\tau = \sigma$ / 2. If we restrict this potential to values where $\lambda > 0$, then for values where $\mu^2 < 0$ we will have a ground state that is not equal to zero as seen in figure 2.4.

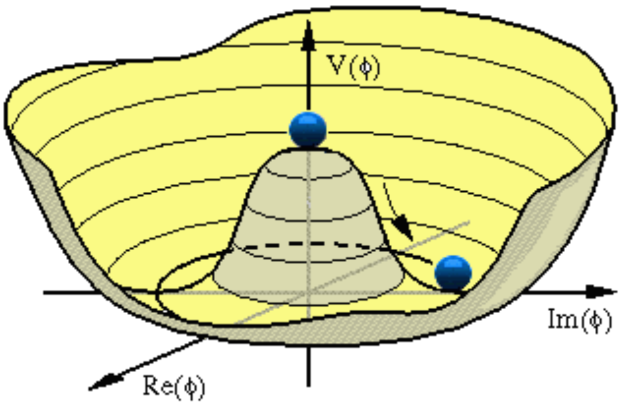


Figure 2.4: The 'sombrero' potential with an unstable state at $\phi = 0$ an a non-zero minimum

This non-zero minimum is often referred to as a Vacuum Expectation Value (VEV). By choosing the VEV judiciously we can observe the effect of this scalar field on our gauge fields, letting

$$\langle\phi\rangle_0 = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v\end{pmatrix} \tag{2.42}$$

where v = $\sqrt{\mu^2/\lambda}$. In this state ϕ is not invariant to any individual generators (τ_a or Y) but if we assign a hypercharge (Y_ϕ) to this scalar = 1 then we can define the electromagnetic charge as

$$Q = \frac{\tau^3 + Y}{2} \tag{2.43}$$

This method preserves electromagnetic symmetry while providing the desired symmetry breaking.

$$Q\langle\phi\rangle = \frac{1}{2}(\tau_3 + Y)\langle\phi\rangle = \frac{1}{2}\begin{pmatrix}Y_\phi + 1 & 0\\ 0 & Y_\phi - 1\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}0\\ v/\sqrt(2)\end{pmatrix} = \begin{pmatrix}0\\ 0\end{pmatrix} \tag{2.44}$$

thus giving us

$$SU(2)_L \times U(1)_Y \rightarrow U(1)_{EM} \tag{2.45}$$

Now that we see that we can recover the EM symmetry that we desire, let's look explicitly at the other terms of the Lagrangian to see how they gain mass. We take the covariant derivative term of the Lagrangian and act upon it with the higgs VEV $\langle\phi\rangle$:

$$\begin{aligned}\Delta\mathcal{L} &= (D_\mu\phi)^\dagger(D_\mu\phi)\\ &= \frac{1}{2}(0 \quad v)\left(g_1 W^a_\mu\tau^a + \frac{1}{2}g_2 B_\mu\right)\left(g_1 W^{\mu b}\tau^b + \frac{1}{2}g_2 B^\mu\right)\begin{pmatrix}0\\ v\end{pmatrix}.\end{aligned} \tag{2.46}$$

To see the effect on the vector bosons we can evaluate this matrix product using the values of τ_a, and in doing so we find

$$\Delta\mathcal{L} = \frac{1}{2}\frac{v^2}{4}\left[g_1^2(W^1_\mu)^2 + g_1^2(W^2_\mu)^2 + (-gW^3_\mu + g_2 B_\mu)^2\right]. \tag{2.47}$$

Using the results of equation 2.47 we can identify the three massive vector bosons and a fourth massless boson which we will define as

$$\begin{aligned} W^\pm_\mu &= \frac{1}{\sqrt{2}}(W^1_\mu \mp iW^2_\mu) && \text{with mass } m_W = g\frac{v}{2};\\ Z_\mu &= \frac{1}{\sqrt{g^2+g'^2}}(gW^3_\mu - g'B_\mu) && \text{with mass } m_Z = \frac{v}{2}\sqrt{g^2+g'^2};\\ A_\mu &= \frac{1}{\sqrt{g^2+g'^2}}(gW^3_\mu + g'B_\mu) && \text{with mass } m_A = 0.\end{aligned} \tag{2.48}$$

The last field A_μ is not present in equation 2.47 as it is massless, but was previously identified as the photon due to the gauge invariance under the τ_3+Y phase rotation. Using this information, and defining the operator $\mathrm{T}^\pm = \frac{1}{2}(\sigma_1 + i\sigma_2) = \sigma^\pm$, we can rewrite the covariant derivative in terms of the mass eigenstate fields, the charge Y, and spinor representation T. Doing this we see

$$D_\mu = \partial_\mu - \frac{ig_1}{\sqrt{2}}(W^+_\mu T^+ + W^-_\mu T^-) - \frac{i}{\sqrt{g_1^2+g_2^2}}Z_\mu(g_1^2T^3 - g_2^2Y) - \frac{g_1g_2}{\sqrt{g_1^2+g_2^2}}A_\mu(T^3+Y). \tag{2.49}$$

Equation 2.49 gives us many useful terms to look at. The last term explicitly couples the massless gauge boson A_μ with the gauge generator (T^3+Y) which we previously identified as the electric charge quantum number in equation 2.43. From this we can also identify the electron charge e as

$$e = \frac{g_1g_2}{\sqrt{g_1^2+g_2^2}}. \tag{2.50}$$

Furthermore, we can use the definitions of Z_μ and A_μ in relation to the weak mixing angle derived in equation 2.29 to rewrite the Lagrangian as

$$D_\mu = (\partial_\mu - \frac{ig_1}{\sqrt{2}}(W^+_\mu + W^-_\mu T^-) - \frac{ig_1}{\cos\theta_W}Z_\mu(T^3 - sin^2\theta_W Q) - ieA_\mu Q. \tag{2.51}$$

Now that we see how the Higgs mechanism applies to the gauge bosons, we will briefly explore its effect on the fermionic components of the Lagrangian. Combining our results from earlier (equations **??** and **??**) we get the full fermionic Lagrangian:

$$\begin{aligned}\mathcal{L}_{Fermion} = \bar{E}_L(i\gamma^u D_\mu)E_L + \bar{e}_R(i\gamma^u D_\mu)e_R \\ \bar{Q}_L(i\gamma^u D_\mu)Q_L + \bar{u}_R(i\gamma^u D_\mu)u_R + \bar{d}_R(i\gamma^u D_\mu)d_R\end{aligned} \tag{2.52}$$

This is a rather large and unwieldy equation to evaluate, so we'll look at just the first term in order to see the explicit coupling of the left handed electron to the gauge boson fields. Using our results from equation 2.51 for covariant derivative we see

$$\begin{aligned}\mathcal{L}_{E_L} =& \begin{pmatrix}\bar{\nu_L} & \bar{e_L}\end{pmatrix}\left((i\gamma^\mu(\partial_\mu - \frac{ig_1}{\sqrt{2}}(W^+_\mu T^+ + W^-_\mu T^-) - \frac{ig_1}{\cos\theta_W}Z^0_\mu(T^3 - \sin^2\theta_W Q) - ieA_\mu Q)\ )\right)\begin{pmatrix}\nu_L \\ e_L\end{pmatrix} \\ =& \bar{\nu_L} i\gamma^\mu\partial_\mu\nu_L + \bar{e_L} i\gamma^\mu\partial_\mu e_L + \frac{ig_1}{\sqrt{2}}W^+_\mu\bar{\nu_L}\gamma^\mu e + \frac{ig_1}{\sqrt{2}}W^-_\mu\bar{e_L}\gamma^\mu\nu_L \\ &+ \frac{ig_1}{\cos\theta_W}\bar{\nu_L}(1/2)\gamma^\mu\nu_L + \frac{ig_1}{\cos\theta_W}\bar{e_L}\gamma^\mu(-1/2 + \sin^2\theta_W(+1))e_L + (ie)\bar{e_L}\gamma^\mu A_\mu(-1)\end{aligned} \tag{2.53}$$

Similar terms link the rest of the components of equation 2.52 to the gauge bosons, as well as additional terms that were not shown for the higher generation of quarks and leptons. With the fermionic components liked to the gauge bosons, we look at the effect of the higgs potential ϕ on the Lagrangian. Again, for simplicity I will just look at the component related to the electron.

$$\begin{aligned}\mathcal{L}_{E_L,Yukawa} =& -\lambda_e\bar{E_L}\cdot\phi\ e_R - \lambda_e E_L\cdot\phi\ \bar{e_R} \\ =& -\frac{\lambda_e}{\sqrt{2}}(v)(\bar{e_L}e_R + e_L\bar{e_R})\end{aligned} \tag{2.54}$$

From this we can identify the mass of the electron as $m_e = \frac{\lambda_e v}{\sqrt{2}}$. In order to generate the masses of the fermions, each particle has its own λ value. This means that while the Higgs mechanism does indeed generate mass for the particles while preserving the underlying $SU(2)\otimes\ U(1)$ symmetry, it does not explain the mass hierarchy that we observe. In addition, we need to add in extra terms into the Yukawa coupling to account for the higher generation of quarks and leptons, adding in coupling terms that account for the mixing of generations. Starting by looking at the mass terms of the quark, we examine the Yukawa coupling of the quarks and inserting the Higgs VEV

$$\begin{aligned}\mathcal{L}_{q,Yukawa} =& -\lambda_d\bar{Q_L}\cdot\phi\ d_R - \lambda_u\epsilon^{ab}\bar{Q}La\phi^\dagger_b u_R + \text{hermitian conjugate terms} \\ =& -\frac{\lambda_d}{\sqrt{2}}(v)\bar{d_L}d_R + -\frac{\lambda_u}{\sqrt{2}}(v)\bar{u_L}u_R + \text{h.c. } + ...\end{aligned} \tag{2.55}$$

Like before, from equation 2.55 we can identify the mass terms for the d and u quarks to be

$$m_d = \frac{\lambda_d v}{\sqrt{2}}, \quad m_u = \frac{\lambda_u v}{\sqrt{2}} \tag{2.56}$$

Just as we found with the electron, the theory parametrizes the quark mass but does not explain the values we observe experimentally. The next step is to add mixing terms for the quark generation. By grouping the quarks into up type and down type vectors we can relate them from their original weak interaction basis to a diagonalized Higgs basis. Let's let u_L^i represent the original basis, and $u_L^{i\prime}$ represent the new basis. If

$$u_L^i = U_u^{ij} u_L^{j\prime}, \quad d_L^i = U_d^{ij} d_L^{j\prime} \tag{2.57}$$

then the two bases are related by a unitary transformations

$$u_L^i = U_u^{ij} u_L^{j\prime}, \quad d_L^i = U_d^{ij} d_L^{j\prime} \tag{2.58}$$

The interaction terms (the W boson current) with the charged gauge boson currents must then be rewritten as

$$J_W^{\mu+} = \frac{1}{\sqrt{2}} \bar{u_L^i} \gamma^\mu d_L^i = \frac{1}{\sqrt{2}} \bar{u_L^{i\prime}} \gamma^\mu (U_u^\dagger U_d) d_L^{j\prime} = \frac{1}{\sqrt{2}} \bar{u_L^{i\prime}} \gamma^\mu V_{ij} d_L^{j\prime} \tag{2.59}$$

where V_{ij} is the 3x3 Cabibbo-Kobayashi-Maskawa (CKM) matrix describing the mixing among six quarks [43] [44]. The off-diagonal terms of the CKM matrix describe the flavor mixing terms between generations, for example charm and strange mixing are related by a unitary transformations

$$V_{1j} d_L^{\prime j} = cos\theta_c d_L^{\prime j} + sin\theta_c s_L^{\prime j}, \tag{2.60}$$

with the term proportional to $\sin\theta_c$ allowing an s quark to decay weakly to a u quark.

2.6 The Higgs Boson

The investigation of fermion mass generation has focused on the scalar field that causes spontaneous symmetry breaking of our gauge theory. We've seen how its interaction has created mass terms for fermions and bosons, but there is another manifestation that we have not looked at yet: the Higgs boson itself. To see this impact we take our scalar field from equation 2.42 and parametrize it by expanding the field in terms of deviations from the ground state:

$$\phi(x) = U(x) \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \tag{2.61}$$

This spinor now contains an arbitrary real component which is given by the VEV of ϕ plus our parametrized real field h(x) with $\langle h(x) \rangle$=0. We are free to make a gauge transformation to eliminate U(x), so we will use the unitary gauge to do this. Just as before we have

$$\mathcal{L}_{higgs} = (D_\mu \phi)^\dagger (D_\mu \phi) + \mu^2 |\phi^\dagger \phi| - \lambda \left(|\phi^\dagger \phi|\right)^2, \tag{2.62}$$

where the minimum potential energy occurs at

$$v = \sqrt{\frac{\mu^2}{\lambda}}. \tag{2.63}$$

Starting by looking at the potential energy term and plugging in the values of ϕ we get

$$\begin{aligned} \mathcal{L} &= -\mu^2 h^2 - \lambda v h^3 - \frac{1}{4}\lambda h^4 \\ &= -\frac{1}{2} m_h^2 h^2 - \sqrt{\frac{\lambda}{2}} m_h h^3 - \frac{1}{4}\lambda h^4 \end{aligned} \tag{2.64}$$

where we have identified that the field h(x) is a scalar particle with mass $m_h = \sqrt{2}\mu = \sqrt{2\lambda}v$. Expanding the kinetic energy term from equation 2.62 gives us the terms we saw earlier in 2.47 plus the Higgs interaction term

$$\mathcal{L} = \frac{1}{2}(\partial_\mu h)^2 + \left[m_W^2 W^{\mu+} W_\mu^- + \frac{1}{2} m_Z^2 Z^\mu Z_\mu \right] \cdot \left(1 + \frac{h}{v}\right). \tag{2.65}$$

Additionally, we can follow that same logic looking at the fermion mass terms from before in 2.54 and 2.55 and identify the Higgs coupling to fermions as

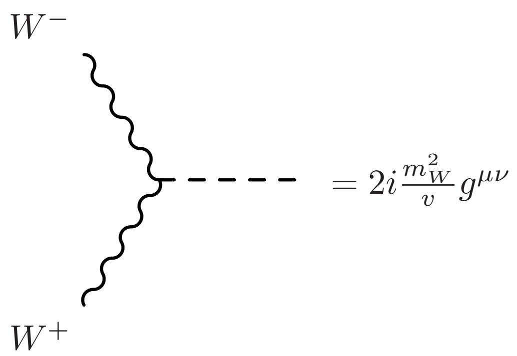
$$\mathcal{L}_f = -m_f \bar{f} f \left(1 + \frac{h}{v}\right). \tag{2.66}$$

Combining these results we can see that the Higgs couples to vector bosons, fermions, as well as itself. Figure 2.5 shows the Feynman rules for these couplings explicitly.

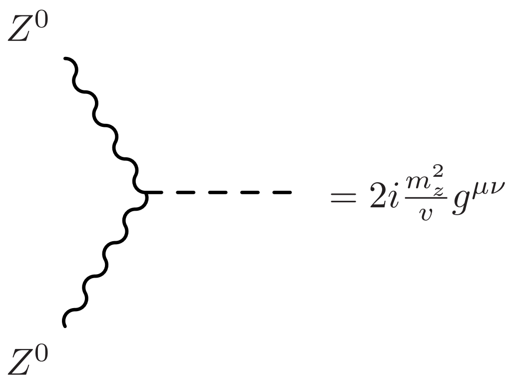
2.7 Success of the Standard Model

With the theoretical framework in place it is useful to look at how successful the standard model has been at predicting and describing the world of particle physics that we observe. As I showed in figure 2.1, we have observed and measured all of the particles shown there. These discoveries have only occurred over about the last 50 years, with many of those particles not even theorized until the 1950s and 1960s. The original quark model proposed by Gell-Mann and Zweig only included the three lightest quarks: up, down, and strange.

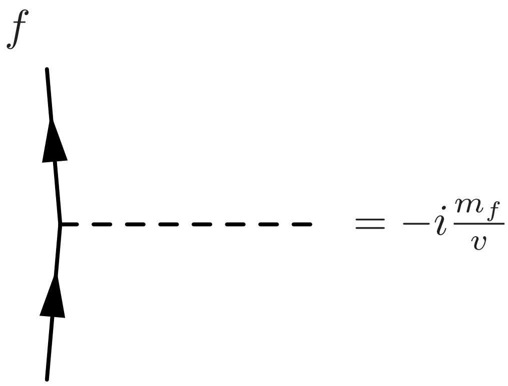
Inclusion of the charm quark was proposed by Bjørken and Glashow in 1964 [45] and its full



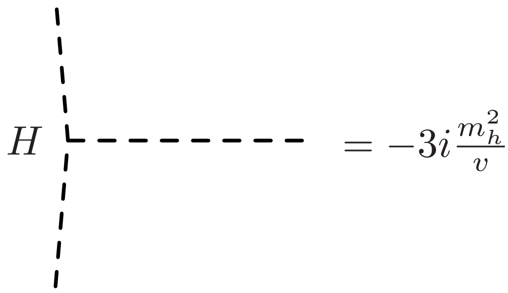
(a) Higgs coupling to W bosons



(b) Higgs coupling to Z bosons



(c) Higgs coupling to fermions



(d) Higgs self coupling

Figure 2.5: Tree level Feynman diagrams for Higgs coupling vertices to vector bosons(a,b), fermions(c), and to itself(d).

inclusion through the GIM mechanism was described in 1970 [46]. Soon after that, the charm quark was first observed in the J/ψ meson by the Stanford Linear Accelerator (SLAC) [47] and Brookhaven National Laboratory (BNL) [48]. Next came the theorization of the bottom (or beauty) quark by Kobayashi and Maskawa in 1973, as a method for describing CP violation in the weak interaction [49]. This would later lead to a Nobel prize for their theory of CP violation in 2008. Not long after that, Fermilab National Laboratory discovered the bottom quark in 1977 [50].

Following this, the W and Z bosons were discovered at CERN in 1983. In proton-antiproton collisions at $\sqrt{s} = 540$, GeV Carlo Rubbia led a team using the Super Proton Synchrotron (which is still in use today) on the experiment UA1 and with team led by Pierre Darriulat on UA2 they jointly announced discovery of the weak bosons [51]. Very few particles were observed in these first experiments, but later under the Large Electron-Positron Collider (LEP) experiment at CERN, precision measurements were made on the W and Z masses [52] [53]:

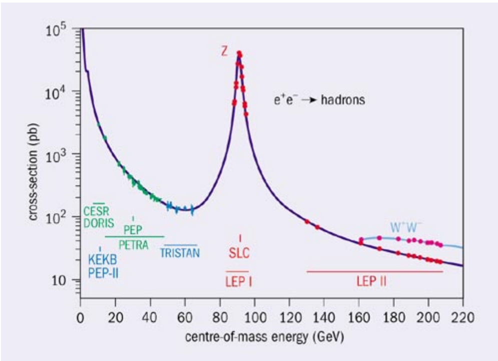
$$\begin{aligned} m_Z &= 91.1875 \pm 0.0021\,\mathrm{GeV} \\ m_W &= 80.376 \pm 0.0033\,\mathrm{GeV}. \end{aligned} \tag{2.67}$$

One more milestone for the standard model came in 1995 when the CDF and D0 experiments at the Tevatron (located at Fermilab National Laboratory) announced the observation of the 6th and final quark, the top quark. The Tevatron used proton-antiproton collisions at $\sqrt{s} = 1.4$TeV to discover the top quark with mass $m_t \sim 176$GeV [54] [55]. The top quark completed the 3rd generation of quarks predicted by Kobayashi and Maskawa, leaving the Standard Model nearly complete.

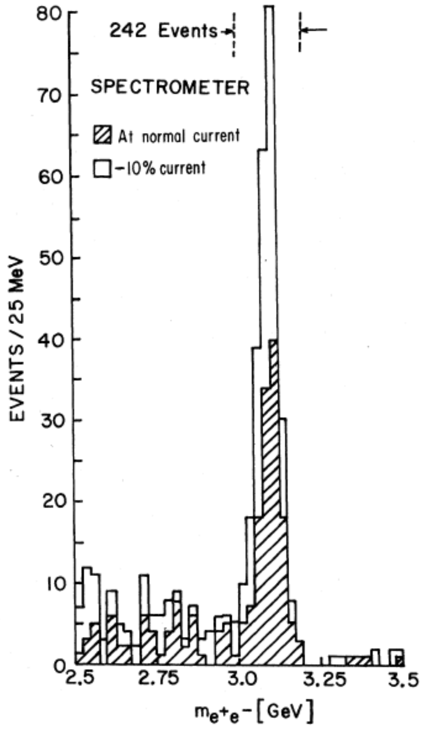
The final particle remaining to be discovered was the Higgs Boson. Both the LEP and Tevatron experiments searched for the Higgs, and though they did not observe it, they were able to exclude a large range of possible masses. Combining their results CDF and D0 were able to exclude the Higgs except for masses of 115 $< m_{Higgs} <$155 GeV, and $m_{Higgs} >$ 176 GeV, as shown in figure 2.6c.

When the LHC first started collisions in 2010, hopes were high that this would lead to the first real look at the Higgs. Then, in July of 2012, the CMS and ATLAS collaborations at CERN announced the observation of a new boson with mass $\sim$ 125GeV that is consistent with expectations for the Higgs boson [25] [26]. Measurements in the $H \to \gamma\gamma$ and $H \to ZZ$ channels at CMS report $m_H = 125.3^{+0.26}_{-0.27}(\mathrm{stat})^{+0.14}_{-0.15}(\mathrm{syst})$ GeV as shown in figure 2.6d.

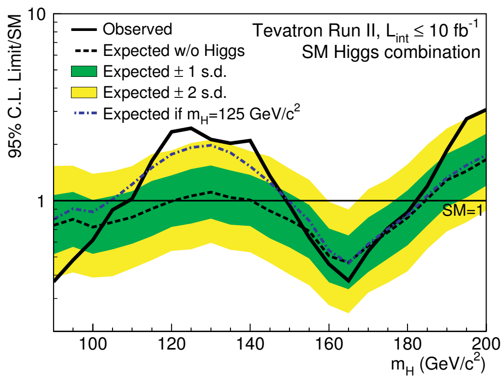
So far, the Higgs Boson has been observed in a number of different decay modes, but no direct observation has been seen in the semi-leptonic WW decay mode. A search here will add a valuable piece to the understanding of the Higgs and its coupling, and add yet another piece



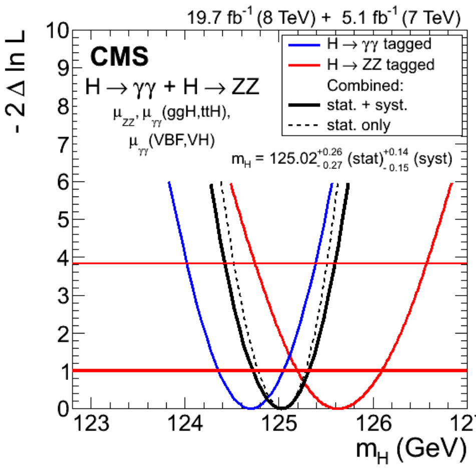
(a) Plot of the e^+e^- annihilation cross section to hadrons, showing the Z peak [56].



(b) Mass spectrum showing the existence of the J/ψ particle from BNL experiments [48]



(c) Combined exclusion plot for Higgs mass from the D0 and CDF experiments [57]



(d) Best fit mass results from the $\gamma\gamma$ and ZZ decay channels at CMS [27]

Figure 2.6: Milestones in particle physics showing discovery and measurements of SM particles.

to the combination of decay modes in which the Higgs can be observed.

2.8 Higgs Production in a p-p Collider

In order to search for the Higgs, we need to model its production and decay. The LHC is a proton-proton collider that can produce the Higgs through a number of different processes. Figure 2.7 shows the production cross sections (at 8 TeV) for the five different production channels that occur at the LHC.

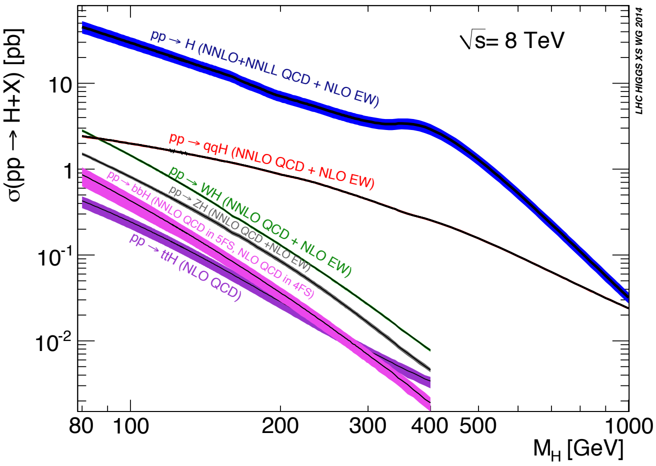


Figure 2.7: Higgs production cross-sections at the LHC for 8 TeV *pp* collisions

Notice that figure 2.7 is log scale, so the top process of Gluon-Gluon Fusion (gg-F) (in blue) is much more likely than any of the others. Also important to keep in mind is that this figure was generated before the Higgs discovery, hence the large range on the x-axis for potential Higgs mass. As this thesis focuses on a search for the $\simeq$ 125 GeV Higgs, that is the area of the figure to focus on. Since gluons are massless they can't couple directly to the Higgs, so the gluon-gluon fusion production mechanism proceeds through a fermion loop interaction (shown in figure 2.8a). This loop is dominated by the top quark because, as we saw previously, the higgs coupling to fermions is dependent on fermion mass(eq 2.66) and the top quark is by far the heaviest fermion. The cross section for Higgs production at $m_H = 125\,\text{GeV}$ and $\sqrt{s} = 8\,\text{TeV}$ is given as:

$$\sigma_{ggF} = 19.27 \quad \pm^{+7.2\%}_{-7.8\%} \text{ (QCD Scale Unc.) } \pm^{+7.4\%}_{-6.9\%} \text{ (PDF+}\alpha_S \text{ Unc.) } \text{ pb}^{-1} \tag{2.68}$$

where the QCD Scale uncertainty refers to the Next to Next to Leading Order (NNLO) radiative corrections, and PDF+α_S refers to uncertainty on the Parton Distribution Function (PDF) and strong coupling parameters.

The next leading production mechanism is though the Vector Boson Fusion (VBF) (fig 2.8b) where W^+ and W^- (or two Z^0) combine to produce a Higgs. Higgs production through VBF is also called qqH production due to the two outgoing quarks that are present in the production mechanism. Since this process involves weak vector bosons, an additional uncertainty on the EW scale is included, which has been calculated to Next to Leading Order (NLO). Production via VBF for a Higgs at $m_H = 125\,\text{GeV}$ and $\sqrt{s} = 8\,\text{TeV}$ is:

$$\sigma_{VBF} = 1.653 \;\; \pm^{+4.5\%}_{-4.5\%} (\text{ EW Unc.}) \;\; \pm^{+0.2\%}_{-0.2\%} (\text{ QCD Scale Unc.}) \;\; \pm^{+2.6\%}_{-2.8\%} (\text{ PDF}+\alpha_S \text{ Unc.}) \;\; \text{pb}^{-1} \tag{2.69}$$

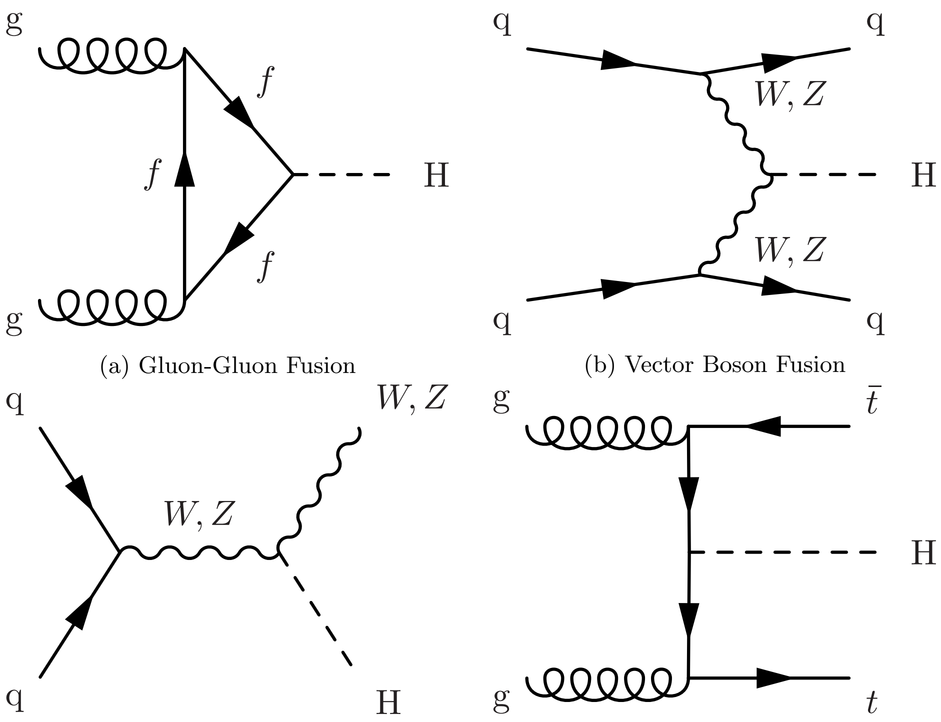
The third (as well as the fourth and fifth) leading processes for Higgs production at the LHC are collectively called associated production mechanisms. This is when the Higgs is produced along with a $W^\pm$ or Z^0 boson, or with a $t\bar{t}$ pair. WH or ZH processes are also referred to as "Higgsstralung" production as the Higgs is radiated from a vector boson in the same way a photon is radiated from an electron in traditional bremsstrahlung radiation (fig 2.8c). ttH production proceeds as shown in figure 2.8d. In total, the associated production cross sections for $m_H = 125\,\text{GeV}$ and $\sqrt{s} = 8\,\text{TeV}$ are:

$$\begin{aligned}
\sigma_{WH} &= 0.7046 \;\; \pm^{+1.0\%}_{-1.0\%} (\text{ QCD Scale Unc.}) \;\; \pm^{+2.3\%}_{-2.3\%} (\text{ PDF}+\alpha_S \text{ Unc.}) \;\; \text{pb}^{-1} \\
\sigma_{ZH} &= 0.4153 \;\; \pm^{+3.1\%}_{-3.1\%} (\text{ QCD Scale Unc.}) \;\; \pm^{+2.5\%}_{-2.5\%} (\text{ PDF}+\alpha_S \text{ Unc.}) \;\; \text{pb}^{-1} \\
\sigma_{ttH} &= 0.1293 \;\; \pm^{+3.8\%}_{-9.3\%} (\text{ QCD Scale Unc.}) \;\; \pm^{+8.1\%}_{-8.1\%} (\text{ PDF}+\alpha_S \text{ Unc.}) \;\; \text{pb}^{-1}
\end{aligned} \tag{2.70}$$

2.9 $H \to WW \to l\upsilon jj$ Production at the LHC

In this thesis we are interested in only one of many decay modes for a Higgs boson. Now that we have covered the ways to produce a Higgs, it is useful to examine the different ways in which a Higgs can decay. Figure 2.9 shows Higgs branching ratios as well at $\sigma \times BR$ for the triggerable final states. The phrase 'triggerable final state' refers to the fact that the final state incluses a physics object that can be identified to classify the event, such as a the presence of one or more leptons. As shown in figure 2.9a Higgs decay to WW has one of the highest cross sections, while figure 2.9b shows that the $l\nu jj$ final state has the highest $\sigma \times BR$ of the 4 fermion final states. Figure 2.10 shows the semi-leptonic W decay mode that we are searching for in this thesis.

In order to calculate the the total $\sigma \times BR$ for the $WW \to l\nu jj$ final state, we need to use the production cross sections from section 2.8 as well as the Branching Ratio (BR)'s for a number of SM process. The BR's considered for this final state are



(a) Gluon-Gluon Fusion

(b) Vector Boson Fusion

(c) Associated Production with W, Z

(d) Associated Production with $t\bar{t}$ pair

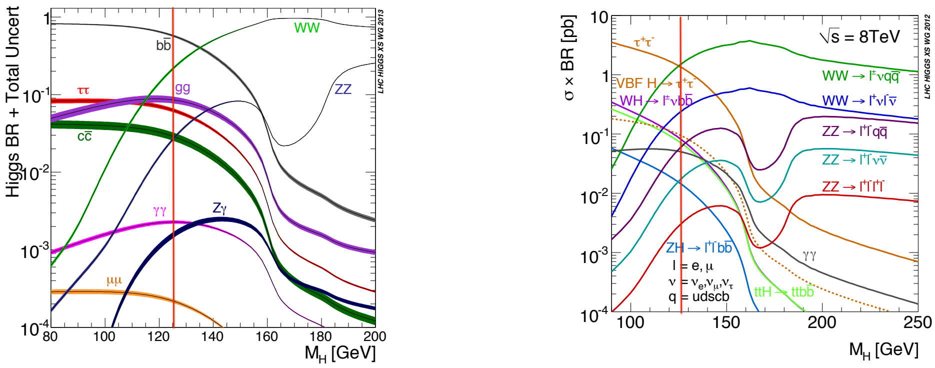
Figure 2.8: Feynman diagrams for Higgs production modes at the LHC

$$
\begin{aligned}
H \to WW &= 0.215^{+4.26\%}_{-4.20\%} \\
W \to l\nu &= 0.3257 \\
W \to qq &= 0.676 \\
WW \to l\nu qq &= 0.2203
\end{aligned}
\tag{2.71}
$$

Additionally, it will be necessary to know the BR's for various final states that are similar to our own final state, as well as states that could appear as an $l\nu jj$ final state due to various detector mis-identification or misreconstructions.

$$
\begin{aligned}
H \to ZZ &= 0.0264^{+4.28\%}_{-4.21\%} \\
H \to bb &= 0.577^{+3.21\%}_{-3.27\%}
\end{aligned}
\tag{2.72}
$$

In this analysis, we simulate production of the Higgs via Monte-Carlo (MC) generators using all of the production mechanisms. In addition, we simulate samples for background processes that are likely to appear in our final state, and using our analysis cuts we can then try to minimize their presence in our final state cuts. Using the values in the last two sections we can calculate the full $\sigma \times BR$ for all of these processes and use them to scale our MC samples appropriately. The signal samples considered are:



(a) Higgs Branching Ratio (BR)

(b) Higgs $\sigma\times$BR for 4 fermion final states

Figure 2.9: Higgs decays modes at the LHC. Red line denotes a mass of 125 GeV.

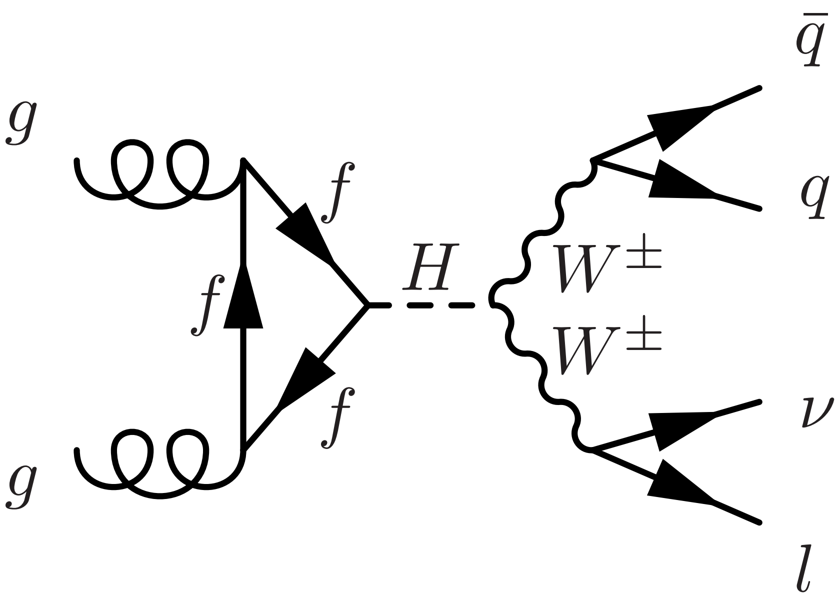


Figure 2.10: Feynman diagram for the SM process a Higgs boson is created through the gluon-gluon fusion process and decays semi-leptonically to two quarks, one lepton, and one neutrino.

$$\begin{aligned}
ggH&,\ \text{where } H \to WW \to l\nu jj = 1.823\ \text{pb}^{-1}\\
qqH&,\ \text{where } H \to WW \to l\nu jj = 0.1493\ \text{pb}^{-1}\\
WH&,\ \text{where } H \to WW = 0.1515\ \text{pb}^{-1}\\
ZH&,\ \text{where } H \to WW = 0.08929\ \text{pb}^{-1}\\
ttH&,\ \text{where } H \to WW = 0.0278\ \text{pb}^{-1},
\end{aligned} \tag{2.73}$$

and Higgs samples that could produce the same final state we are looking for (‘volunteer’ signals)

$$
\begin{aligned}
WH, \text{ where } H \to bb \to l\nu jj &= 0.1324 \text{ pb}^{-1} \\
TTH, \text{ where } H \to bb \to l\nu jj &= 0.0746 \text{ pb}^{-1} \\
WH, \text{ where } H \to ZZ &= 0.01860 \text{ pb}^{-1} \\
ZH, \text{ where } H \to ZZ &= 0.01096 \text{ pb}^{-1} \\
ttH, \text{ where } H \to ZZ &= 0.00341 \text{ pb}^{-1}.
\end{aligned}
\tag{2.74}
$$

Using the information above we can quickly see that our signal will be dominated by the ggH sample, which is what we want. It is also notable that an associated production mode of WH where $H \to bb$ has a non-negligible contribution to the number of events we expect. In fact, it is very comparable to that of the expected signal for qqH events. Later in the analysis we show how we can make cuts on certain event criteria (in this case to detect the presence of b-quark jets) in order to remove this signal 'contamination.'

2.10 $H \to WW \to lvjj$ Backgrounds

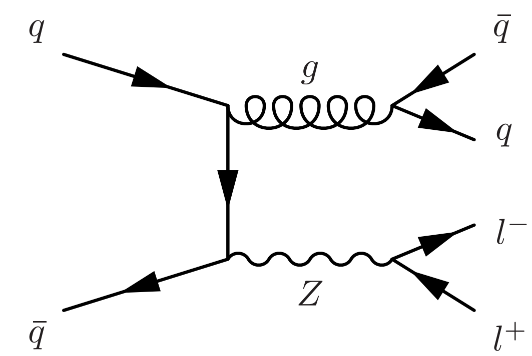
As I have described above in section 2.9 we are only interested in events that have a final state of one lepton, two quarks, and one neutrino. When identifying or reconstructing events in our detector, there are three categories of events that can make it into our selection by mimicking the event signature of $\ell^{\pm}\nu q\bar{q}$ we are looking for. To identify this signature we select for a final state of one isolated lepton (electron or muon), two high p_{T} jets, and at least 25 GeV of $\not{E}_T$. In order to identify our signal we need to consider all SM processes that could also result in that final state.

1. True signal events that are from $H \to WW \to l\nu jj$ events (most important if not the most numerous)
2. 'Volunteer signal' events: events that are Higgs decays where the Higgs does not decay through the semi-leptonic W channel. An example of this would be a $H \to bb$ event where an extra lepton was identified.
3. Background events: events from Standard Model processes that have final states which look like $l\nu jj$
   (a) Irreducible: processes that produce the $l\nu jj$ final state naturally, such as SM WW production where the decay is semi-leptonic
   (b) Reducible: processes that only partially reproduce the $l\nu jj$ final state, such as $t\bar{t}$ which will have extra jets and b-jets associated with it.

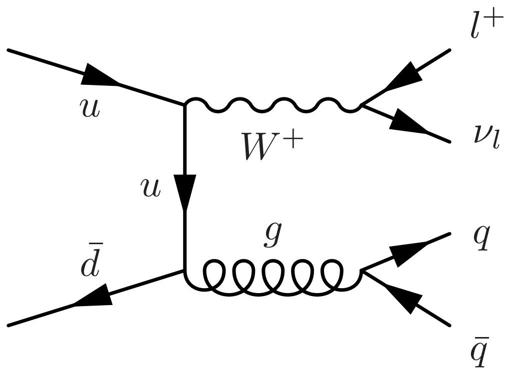
The backgrounds considered in this analysis are as follows:

- W+jets: the production of a single W vector boson in association with quarks or gluons can mimic our final state when the W decays leptonically. The large cross section makes this by far the dominant background to contend with, so accurate modeling is imperative.
- Drell-Yan Z/γ^* + jets: production of single Z/γ^* bosons in association with quarks or gluons, where one lepton goes undetected because of acceptance or inefficiency effects, and the hadronic activity mimics the final state signature of the hadronic W decay products.
- WW: non resonant WW production is an irreducible background for our analysis
- WZ: mimics our final state if the Z decays hadronically or if the W decays hadronically and the Z decays leptonically where one of the leptons is not identified.
- ZZ: if one Z decays hadronically, and only one lepton is identified in the event.
- $t\bar{t}$: top quarks decay primarily to a b quark and a W boson via the weak interaction. The presence of two Ws in the final state can clearly reproduce our signal signature, though the presence of extra b quarks is useful in cuts to limit this background. Due to acceptance and inefficiencies in reconstruction we can still get contamination from $t\bar{t}$ in our selection.
- Single Top: production proceeds via three distinct channels [58].
    1. t-channel: a top is produced via the exchange of a virtual W boson between a b quark and another quark.
    2. s-channel: a top quark is produced with a $\bar{b}$ quark after the annihilation of a pair of quarks.
    3. tW-channel: a top quark is produced in association with a W boson via gluon-b quark interaction.
- QCD Multi-jet: events with multiple jets contribute to the background due to the non-negligible probability of a jet being mistakenly reconstructed as a lepton. This background is difficult to model via MC so we use a data-driven approach described in section 6.2.

Representative Feynman diagrams for these processes are shown in figures 2.11, 2.12, 2.13, and 2.14. All of these SM processes will play a part in the analysis, with more comprehensive descriptions in chapter 6. Before we get to that, we need to understand the machine that makes these collisions happen, the LHC described in chapter 3, and the detector that collects and reconstructs our events CMS, which is described in chapter 4.

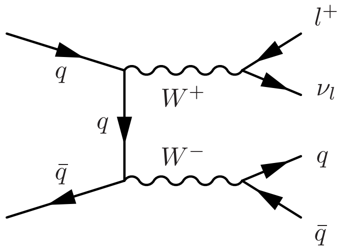


(a) Production of a Z^0 in association with jets

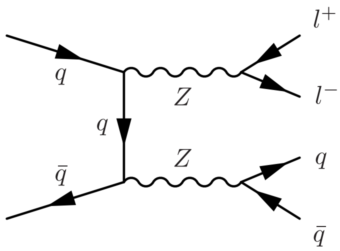


(b) Production of a W^+ in association with jets

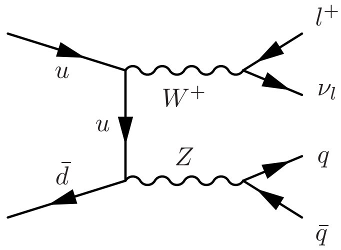
Figure 2.11: Representative Feynman diagrams of SM V+jets processes



(a) Production of a W^+W^- pair



(b) Production of two Z bosons



(c) Production of a W^+ and Z boson

Figure 2.12: Representative Feynman diagrams of SM diboson processes

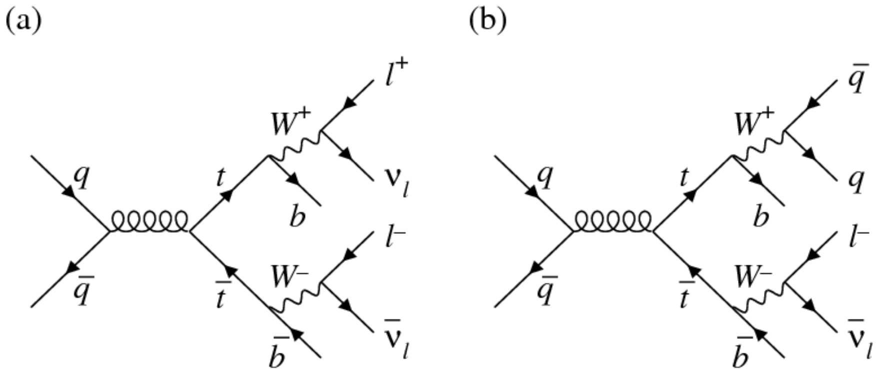
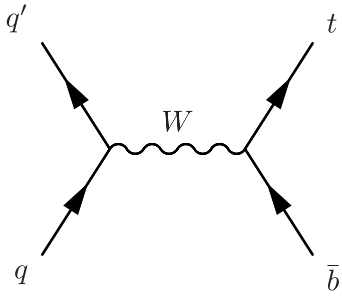
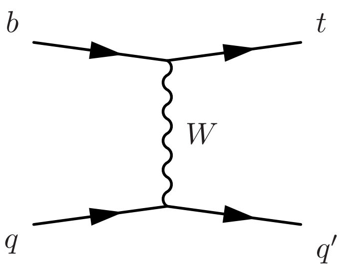


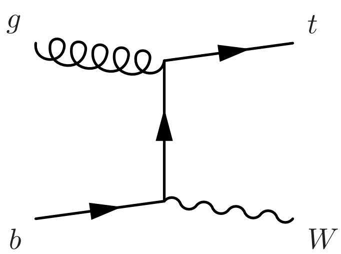
Figure 2.13: Representative Feynman diagrams of the $t\bar{t}$ process [2]



(a) Production of a single top quark via the s-channel



(b) Production of a single top quark via the t-channel



(c) Production of a single top quark via the tW-channel

Figure 2.14: Representative Feynman diagrams of SM single top processes

Chapter 3

The Large Hadron Collider

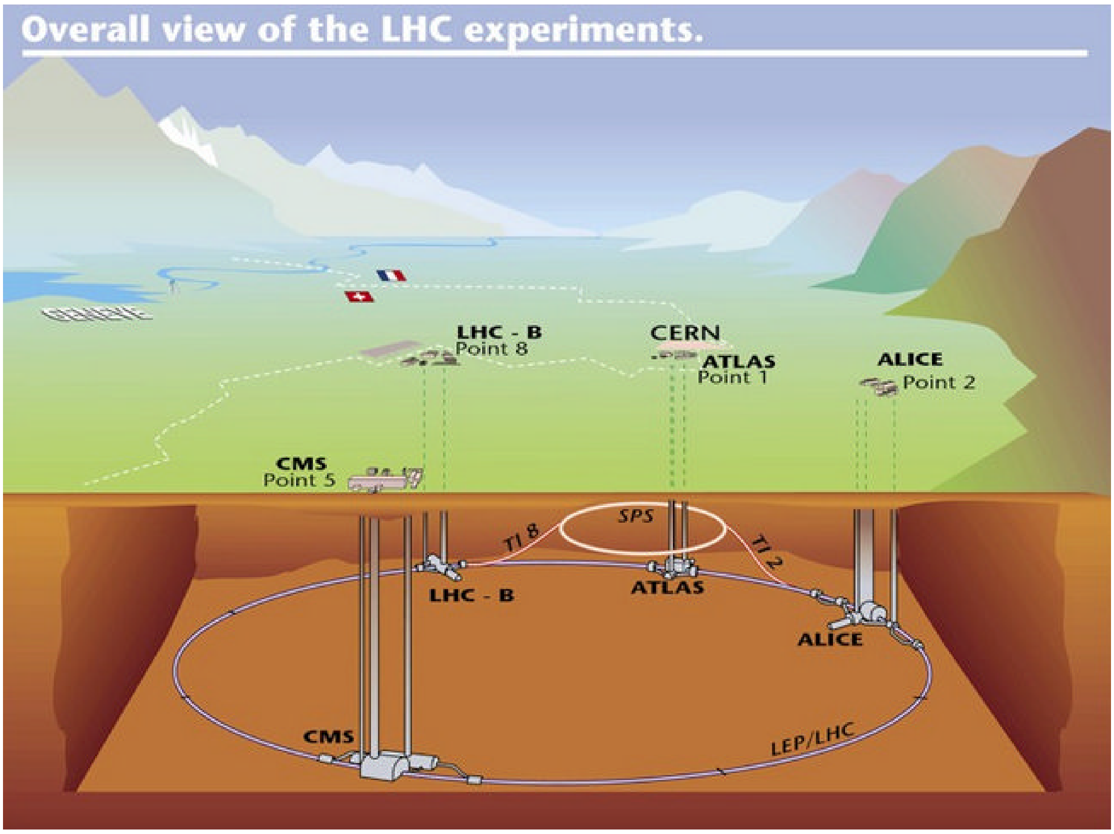


Figure 3.1: Artistic representation of the LHC accelerator complex with both surface and subsurface views [3].

The Large Hadron Collider (LHC) is the world's largest particle accelerator. Located on the border of Switzerland and France just outside of the city of Geneva, it is run by European Center for Nuclear Research (CERN). The LHC is primarily a proton-proton collider designed to collide anti-circulating proton beams at a center of mass energy of $8TeV$, but it also can accommodate collisions of fully stripped lead ions ($^{208}Pb^{82+}$) with a total center of mass energy of a staggering $1.15PeV$ [8].

As shown in figures 3.1 and 3.2, the main campus of labs is located at the point marked CERN, while the experiments are located at various spots around the ring. The two large multipurpose physics detectors Compact Muon Solenoid (CMS) and A Toroidal LHC Apparatus (ATLAS) are shown, along with the more specialized detectors of A Large Ion Collider Experiment (ALICE) and Large Hadron Collider beauty (LHCb). There are three more experiments (Large Hadron Collider forward (LHCf), Total Cross Section, Elastic Scattering and Diffraction Dissociation (TOTEM), Monopole and Exotics Detector At the LHC (MoEDAL)) that also use the LHC ring but are not pictured.

Figure 3.2: Aerial view of the LHC complex, spanning the French-Swiss border [4].

The LHC itself is a two ring superconducting proton accelerator built in existing tunnels that were used for Large Electron-Positron Collider (LEP)(LEP collided electrons and positrons from 1989-2000). Its goal is to reveal physics beyond the Standard Model by generating large numbers of particle collisions at higher energies than ever before, up to $14TeV$. The high energy in the collisions allows for heavy particles to be created, while the high collision rate makes it more likely to see rare physics processes. We can measure the number of events per second generated at the LHC to be a product of the machine luminosity and the cross section of the events we are looking for:

$$N_{events} = L\sigma_{event} \tag{3.1}$$

The machine luminosity depends on beam parameters and can be described explicitly for a Gaussian beam as:

$$L = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi\epsilon_n \beta^*} F. \tag{3.2}$$

- N_b - Number of of particles per bunch. Designed for high luminosity, the LHC seeks to maximize the density of particles in each bunch. This density is limited by the linear tune shift of beam-beam interaction given by

$$\xi = \frac{N_b r_p}{4\pi\epsilon_n} \tag{3.3}$$

  where r_p is the classical proton radius, and ϵ_n is the transverse beam emittance ϵ_n=3.75mum. When combined, this gives a maximum bunch intensity for the LHC to be $N_b = 1.15 \times 10^{11}$.

- n_b - the number of bunches per beam. This is limited by the spacing of bunches, designed for a nominal 25ns spacing. This spacing allows for a maximum of 2808 proton bunches per beam.

- f_{rev} - the revolution frequency of the beams. This is set by the size of the LHC giving f_{rev} = 11.2kHz

- γ_r - the relativistic gamma factor of the protons. This is determined by the energy used in collisions. For the 2012 run this was 4TeV.

- ϵ_n - the transverse normalized beam emittance. This is determined by measuring the spread of the beam in the tranverse direction, and is ϵ_n=3.75mum for the LHC.

- β^*- the beta function at the collision point. This describes the size of the beam, and is minimized at interaction points to maximize the probability of collisions during beam crossing. For the LHC $\beta^* = 0.55$

- F - the geometric luminosity reduction factor due to crossing angle and an interaction point(IP) is defined as

$$F = \left(1 + \left(\frac{\theta_c \sigma_z}{2\sigma_*}\right)^2\right)^{-1/2} \tag{3.4}$$

  where θ_c is the full crossing angle at the IP, σz is the RMS bunch length, and σ_* is the transverse RMS beam size at the IP.

The LHC was designed to deliver a luminosity of L = $10^{34} cm^{-2} s^{-1}$ for proton operation. This luminosity is delivered to CMS and ATLAS, with lower luminosity delivered to the other experiments. In 2010 and 2011 the LHC ran at center of mass energy $\sqrt{s} = 7\,\mathrm{TeV}$, and delivered a combined $\sim 6\,\mathrm{fb}^{-1}$ of data. In 2012 the energy was increased to $\sqrt{s} = 8\,\mathrm{TeV}$, and the LHC delivered $\sim 23\,\mathrm{fb}^{-1}$ of data to CMS. Figure 3.3 shows the integrated luminosity delivered to CMS in 2010-12.

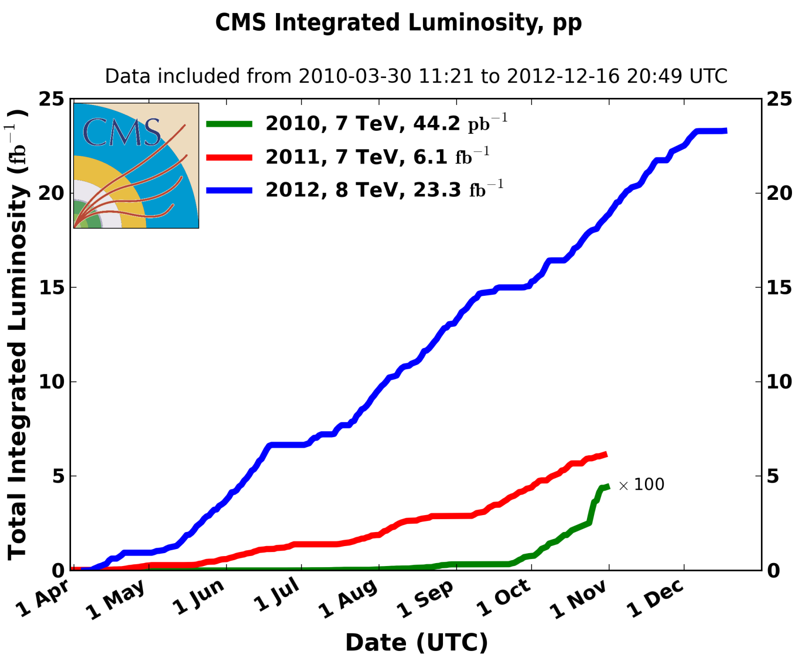


Figure 3.3: Integrated Luminosity delivered to the CMS experiment from 2010-12 [5]

The rest of chapter 3 will describe the injection scheme for the LHC, the different types of magnets and how they are used, and finally the radio-frequency cavities that accelerate the protons to the design energies.

3.1 Accelerator System

The LHC is comprised of a number of interconnected accelerator rings. The main LHC ring was built in the existing tunnel that was bored for the Large Electron-Positron Collider (LEP) experiment between 1984 and 1989. It is comprised of 8 straight sections and 8 arcs, lying between 45 and 170 m below the surface with a full circumference of 26.7km. Figure 3.4 shows the location of the structures around the LHC ring, highlighting the CMS and ATLAS experiments at points 5 and 1 respectively.

Before the protons make it into the LHC ring, they must first undergo numerous acceleration and bunching procedures. Ultimately, the protons that are collided come from a bottle of hydrogen gas attached to CERN's Linac2 linear accelerator [59]. In Linac2 the hydrogen passes

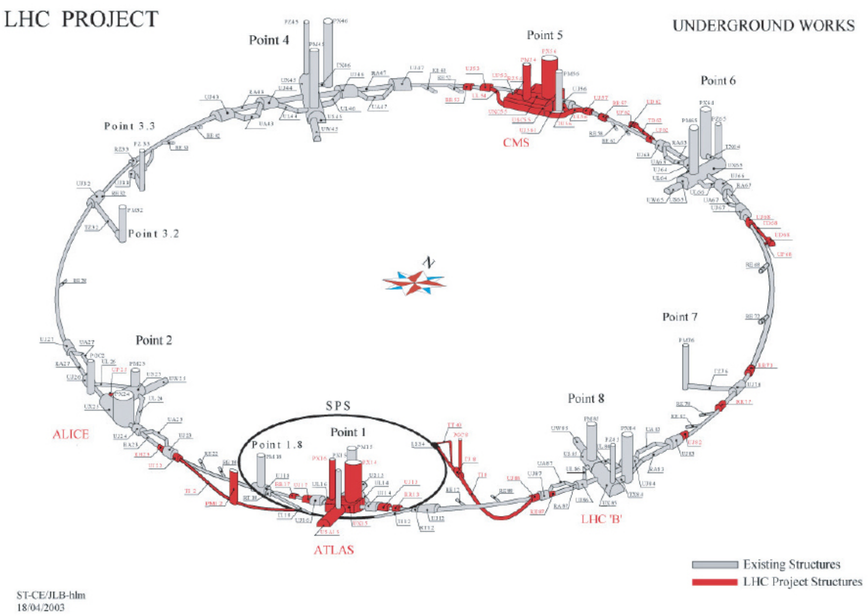


Figure 3.4: Layout of facilities along the LHC ring at CERN [6]

through an electric field which strips the electrons off, leaving just the protons to enter into the accelerator. Linear accelerators work by using Radio-Frequency (RF) cavities to produce a series of electromagnetic fields that exert a force on the particles inside pushing them in one direction down the beamline. A more detailed description of the RF cavities will be provided in section 3.3, but there are some things that will be important to note now.

An RF cavity is a specially shaped, hollow conductor, that the beam passes through. By applying an oscillating electric field to this specially shaped cavity, you can determine the resonant frequency of the RF cavity (as well as its harmonics which are the integer multiples of the fundamental resonant frequency). By using a resonant frequency that matches the revolution frequency of the proton, you can ensure that the proton receives an accelerating force from the RF field [60]. This resonant field generates a number of useful results:

1. Protons feel an accelerating force each time they pass through the RF cavity. Once the revolution frequency of the proton reaches the fundamental frequency of the RF cavity, $f_{RF} = n \times f_{rev}$, the proton will be entering the RF cavity just as the field is alternating through its point of zero field. Once they reach this speed they will feel no acceleration from the cavity.

2. Protons moving too fast or two slow in relation to this equilibrium will either feel an acceleration or deceleration from the RF cavity. This results in diffuse groups or protons being bunched into a group going the same speed.

3. Driving an RF cavity at a harmonic frequency n will result in n bunches of protons being formed due to this splitting.

4. In order to increase the energy of the protons over a large range, you must increase the frequency of the cavity to maintain synchronization with the revolution frequency.

By using different RF cavities and running them at the various harmonics, each part of the accelerator chain is able to increase the energy of the protons, as well as split them into the specified number of bunches.

Digressing a little bit, we need to address the issue of proton energy. We have been describing the center of mass energy of proton collisions, but what exactly does that mean? At rest, a proton has a mass of $\sim 938 MeV/c^2$. Usually, in particle physics, the c^2 term is dropped and we would describe the proton of having an energy of 938MeV. No particle can move with speeds faster than the speed of light in a vacuum, but there is no limit to the energy a particle can attain. In high-energy accelerators like the LHC, particles are accelerated to very close to the speed of light. When the speed of a particle nears the speed of light, the classical Newtonian kinetic energy term ($\frac{1}{2}mv^2$) no longer correctly describes the energy. Instead, we must use the relativistic kinetic energy (KE = (1-γ)mc^2), where c is the speed of light and $\gamma = 1/\sqrt{1-(v/c)^2}$. In these conditions, as the energy increases,the increase in speed is minimal. Table 3.1 shows the relationship between kinetic energy of a proton at each stage of acceleration at the LHC and its speed.

Kinetic Energy of Proton	Speed (%c)	Accelerator
50 MeV	31.4	Linac 2
1.4 GeV	91.6	PS Booster
25 GeV	99.93	PS
450 GeV	99.9998	SPS
7 TeV	99.9999991	LHC

Table 3.1: Relationship between kinetic energy and speed of a proton in the CERN accelerator complex, reproduced from [24].

The protons are accelerated in a series of steps shown in figure 3.5, essentially from rest in the form of hydrogen gas, up to their final energy in the LHC ring. As mentioned above, the protons are first stripped and accelerated through Linac 2, reaching an energy of 50MeV. From there, they enter the Proton Synchrotron Booster (PSB), which accelerates them up to 1.4 GeV before delivering them to the Proton Synchrotron to be brought to 25 GeV. The next stage of acceleration is the Super Proton Synchrotron (SPS), which brings the protons up to 450 GeV before finally delivering them into the LHC. Once in the LHC, they are accelerated to the specified beam energy (4TeV in 2012, but designed to go up to 7TeV) before they are collided. Images of the various accelerators that are described here can be seen in figure 3.6 and 3.7.

In addition to accelerating the protons up to the necessary energy level, the injection chain is where the protons get separated and grouped into bunches. As with the energy of the protons, we'll begin with Linac2 where the protons start. It is here where a group of protons is formed

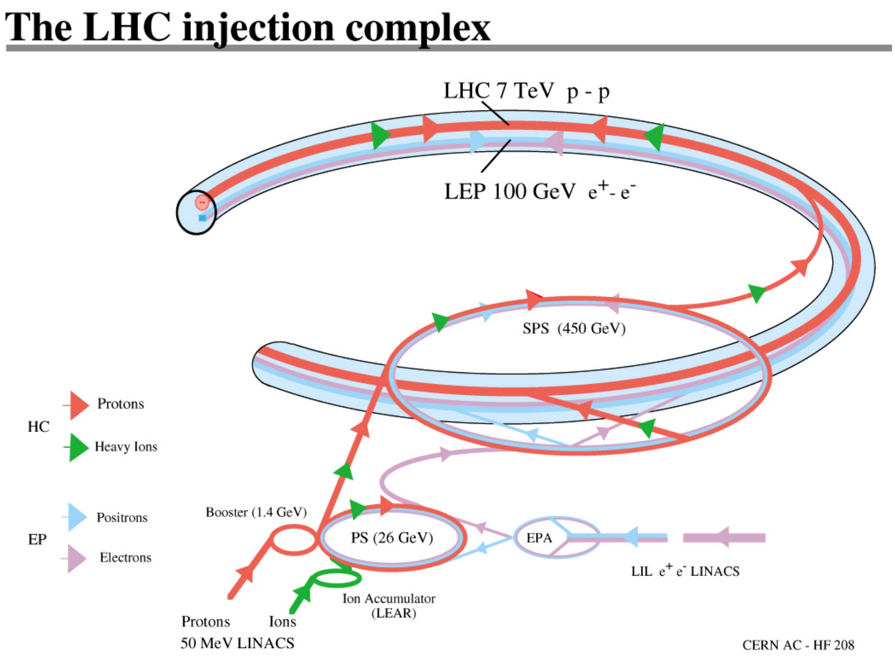
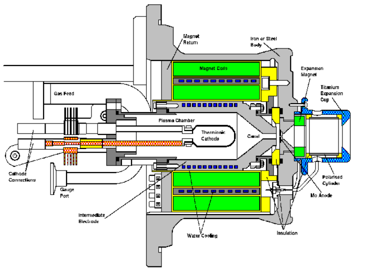


Figure 3.5: Overview of the LHC injection chain at CERN [7]

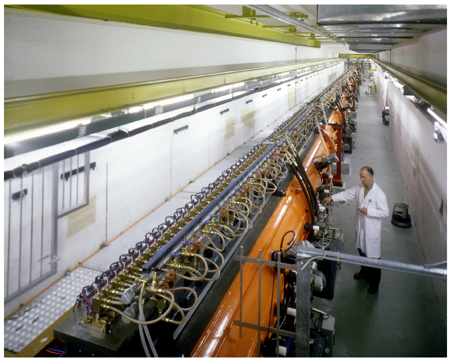
by controlling the input of hydrogen gas. It is not until the next stage of the PSB where the protons begin to be divided. By the time the protons reach the end of Linac2, they have been accelerated by use of RF cavities and collimated by using quadrupole magnets. Quadrupole (and higher order) magnets work by producing a field that squeezes the protons in a particular direction, the specifics on how they control the proton beams will be explored in section 3.2. For now we will just assume that each of the synchrotrons uses many magnets to steer and focus the protons.

Before reaching the PSB, the beam can be split into 4 separate groups in order to take advantage of the 4 separate stacked synchrotrons that make up the PSB. This works to limit the transverse emittance of beam by reducing the number of protons that need to be accelerated in each group. The PSB takes only 1.2s to accelerate a bunch of protons from the 50MeV of energy they have on arrival to the 1.4GeV it delivers to the Proton Synchrotron (PS). For LHC fills, the full splitting into 4 groups is not always used. The PS is engineered to accept 6 packets from the PSB, which is done in either a 3+3 or 4+2 configuration in sequential 1.2s batches.

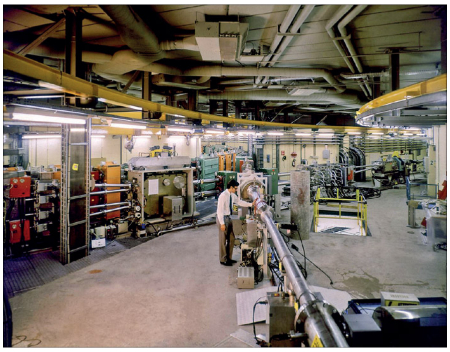
Once the protons are in the PS the process of splitting is begun. The PS ring is 628m in diameter and operated at RF harmonic h = 7 on arrival of the packets from the PSB. Each harmonic of the field provides a minimum, or a ‘bucket’, that allows the PS to capture one bunch in each of the ‘buckets’ produced by the harmonics. One bucket is left empty here, so it starts witha total of six bunches. The bunches are then each split into three smaller bunches while at 1.4GeV using RF cavities operating on harmonics h =7, 14, and 21. While bunched on harmonic h = 21 the protons are accelerated up to 25GeV before being split again. Here they



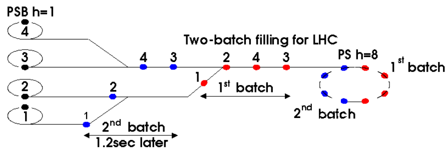
(a) Schematic of Duoplasmatron which takes hydrogen gas and strips the electrons to generate a proton beam for Linac2 [61]



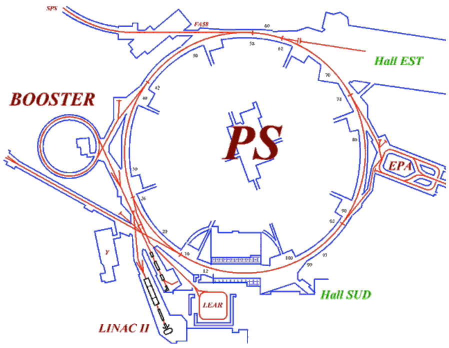
(b) Linac2 in it's cavern[62].



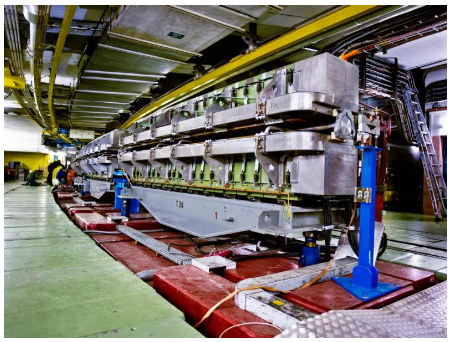
(c) Proton Synchrotron Booster input line[63].



(d) Schematic of batch filling of Proton Synchrotron from PSB[64].

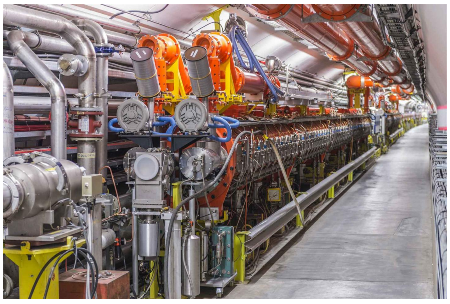


(e) Diagram of the PS complex layout showing Linac2, PSB and PS[65].



(f) Proton Synchrotron dipole magnets used for beam steering.B[66].

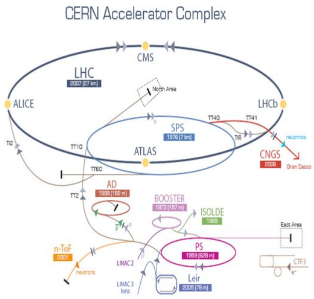
Figure 3.6: Features of the Linac2, Proton Synchrotron Booster, and Proton Synchrotron



(a) Part of the SPS accelerator[67].

(b) A section of dipole magnets along the SPS[67].

(c) Riding along the ∼27km LHC tunnel you can see the curve of the dipoles [68].



(d) Graphic of the entire CERN accelerator Complex[69].

Figure 3.7: Features of Super Proton Synchrotron and Large Hadron Collider

are each split twice more using 20MHz and 40 MHz RF systems. This results in the original 6 bunches being split into a total of 72 bunches in the PS on harmonic h=84, with 12 consecutive buckets remaining empty. These empty buckets provide a gap of $\sim 320ns$ which allows for the rise-time of the ejection kicker.

Before the packets are moved from the PS to the SPS they are first shortened from $\sim 11ns$ to $\sim 4ns$ in length via a rotation in phase space from an 80MHz h = 168 mode. In the SPS, the protons enter a nearly 7km in diameter ring that uses more than a thousand electromagnets to focus and steer the beam [70]. The SPS can store up to 4 bunch trains delivered from the PS at a time and accelerate them from 25GeV up to 450GeV. Due to rise-time of the injection kicker into the SPS there is a 220ns gap at the end of each bunch train. The large acceleration produced in the SPS necessitates the use of tunable RF cavities.

Finally, the bunches are injected into the LHC ring. The SPS injects bunch trains in groups of 3 or 4 at a time into the LHC. At the end of each train is another gap due to the LHC injection kicker rise-time. Finally, once the LHC ring is filled it has an orbit of $88.924\mu s$, leaving a 3 μs abort gap at the end of the orbit. The entire LHC injection scheme is summarized in Figure 3.8. LHC injection occurs near points 2 and 8 (one injection location for each of the beam directions) through use septum and kicker magnets. These magnets precisely time the bunch injection as well as deflecting the incoming beam into the correct orbit of the LHC. Once in the LHC the bunches are accelerated up to their final energy (4TeV in 2012, but up to 6.5 TeV beams have been generated in 2015) before collisions.

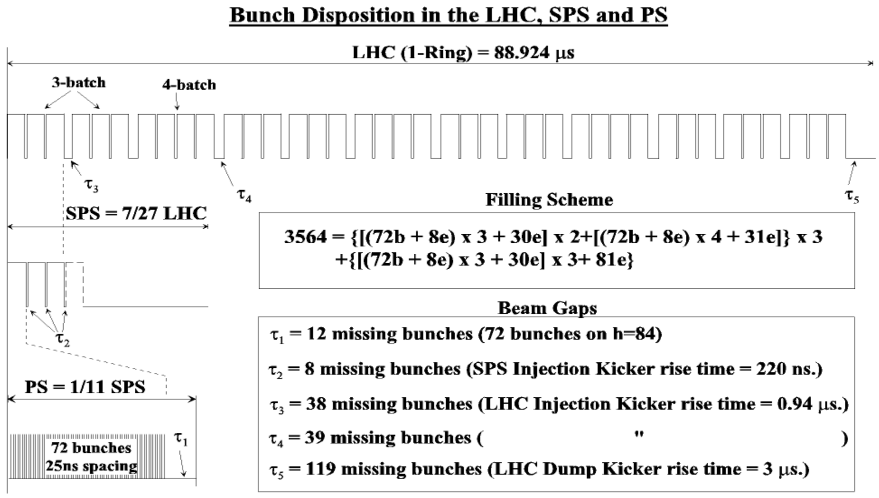


Figure 3.8: Schematic of the bunch structure for filling the LHC ring. Initially, 6 groups of protons are provided to the PS which splits them into 72 bunches. The bunch trains then travel to the SPS and into the LHC with beam gaps arising due to the rise time of injection kickers. This leads to a total of 3564 possible buckets with 2808 filled (assuming 25ns bunch spacing).Reprinted from [8]

.

The LHC ring is divided into eight octants with eight straight sections (one on either side of the 4 interaction points (IP)) and 4 curved regions. Figure 3.9 shows the distribution of these octants in the LHC, showing what each region is used for. The low β description for each IP refers to size of the beam, a reminder that in addition to crossing at each IP, the beams are squeezed to maximize interactions.

3.2 Magnets and Cryogenic System

The LHC uses many different kinds of magnets, from beam injection using septum and kicker magnets to the main dipole magnets used for steering to the higher order (quadrupole, sextupole, octupole) magnets used to focus the beam. There are 1232 main dipole magnets in the LHC that are each 15 meters long and weigh $\sim$35 tons. These magnets are superconducting due to the large magnetic field required of them (8.33 Tesla, 100k times Earth's magnetic field) and thus have to be kept very cold.

The superconducting coils are kept at 1.9 K, a temperature that is achieved through use of more than 120 tons of superfluid helium. In order to keep the coils this cold, a complex cooling system that comprises the largest cyrogentic system in the world is needed [71]. Figure 3.10 shows a cross section of the dipole / cryostat. From the inside out, we have the two beam pipes which are each surrounded by the superconducting magnet coils made from niobium-titanium

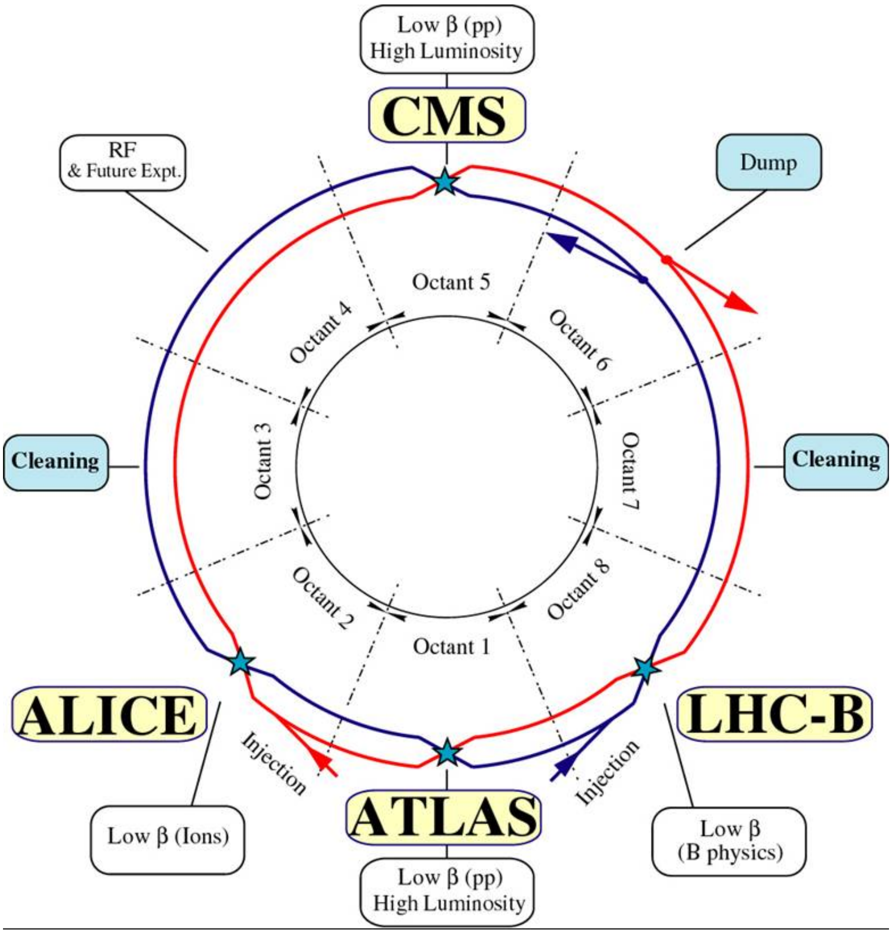


Figure 3.9: The LHC ring is divided into eight octants with 4 interaction points marked with stars[8]

.

(NbTi). Around this lies the iron yoke which serves as a large cold mass (held at 1.9 K) as well as path for the magnetic field to loop through. Around this is a vacuum vessel that serves as an insulator.

The dipole magnets contain two separate vacuum systems, one for the beam pipes and a second to insulate the magnet [72]. In order to provide the centripetal Lorentz force needed on the beam, the dipoles are set with the fields pointing vertically up or down (depending on the direction of the beam). The field lines of the dipole are shown in figure 3.11b. Though the magnets do have a curvature, it is hard to notice when looking at any magnet individually (3.11a).

In addition to the dipoles there are many quadrupole and higher order magnets that are used to correct the fields and focus the beams. Quadrupole magnets provide a squeezing force on the beam in one plane, so by providing two quadrupoles in succession that are rotated 90 degrees in relation to one another, you can squeeze the beam in the x and y plane successively to keep it centered in the beam pipe. In the straight sections of the LHC, there are special magnets that perform the final squeezing and bending of the beams for collisions. These are called low-β

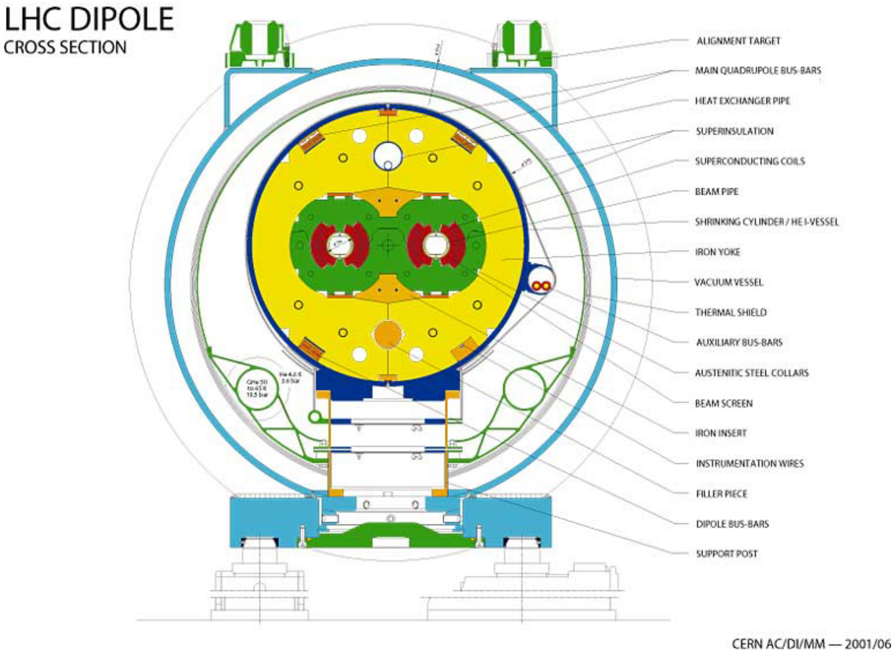
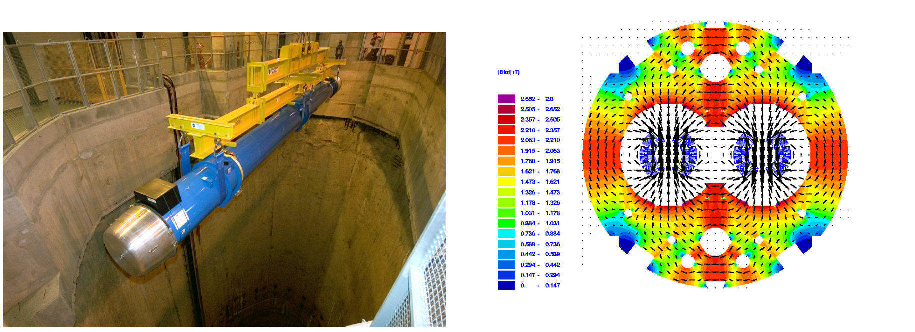


Figure 3.10: Cross section view of an LHC dipole magnet and cryostat. Reprinted from [8].

inner triplets which must provide a very high field gradient of 215 T/m as well as be able to withstand high radiation doses.



(a) Part of the SPS accelerator[73].

(b) Drawing of a magnetic field lines for a dipole[74].

Figure 3.11: Features of the LHC dipole magnets

Finally, there are special magnets used for injection and extraction of the proton beams. The injection scheme the LHC uses is a single-turn injection characterized by two types of magnets: a septum magnet for bending and a kicker magnet for alignment. A drawing of this beam injection is shown in figure 3.12. The septum magnet has two regions, one where it produces a homogeneous field to deflect the incoming beam horizontally into alignment with the target beam, and a second that has no field where the circulating beam passes through without being deflected [9] [75]. Then, a kicker magnet produces a very short pulsed magnetic field that deflects the bunch vertically into the final orbit of the LHC. Since the kicker magnet is in the beamline, the length of the kick must be short enough to not interrupt the circulating beam and timed precisely to insert the new bunches into the beam. For extraction the system works in a similar way, with the kicker providing a short pulse to displace part of the circulating beam and the septum then steering the beam away to be deposited in a beam dump.

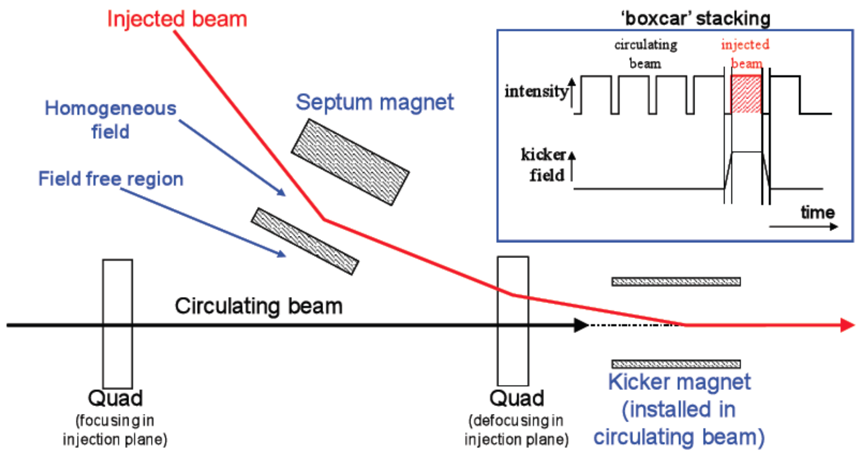


Figure 3.12: Depiction of the single turn beam injection for the LHC. Reprinted from [9].

3.3 Radiofrequency Cavities

The radiofrequency cavities at the LHC, much like the RF cavities in the synchrotrons, capture and accelerate the injected beam using a 400MHz superconducting cavity system [60]. Two independent RF systems of 8 cavities are required (one for each beam) to provide the 16MV needed when the beam is at full energy (7 TeV per beam), though only half of that is needed when the beams enter the LHC. Each cavity supplies a potential difference of 2MV to the beam, accelerating the proton bunches through the potential difference every time they pass through. In this way, the beam is accelerated from 450GeV on entry to the LHC up to 7TeV, getting an accelerating 'kick' from this electric field on every revolution around the ring.

The main RF system is housed in Point 4 along the ring, and consists of two 4-cavity cyromodules per beam. A diagram of one of these cryomodules is shown in figure 3.13. Each of the cavities is made out of niobium sputtered copper to allow the cavities to become superconducting (from the niobium), while the copper helps to dissipate any heat build-up and reduce the possibility of a quench. The superconducting cavities are needed as they dissipate much less power than a normal conducting cavity and allow for a narrow resonance width. Operating at 400MHz limits the bunch length to $< 2ns$ and each cavity is powered by a 300 kW Klystron [8].

A klystron is a source of RF power that works by weak RF source to accelerate and bunch electrons. By using a system of chambers it can build up a resonance of electrons from the relatively weak RF input, and then this much stronger RF resonance can be delivered via waveguide cables to the superconducting cavities of the main RF cryomodules. Having a klystron for each

RF cavity allows for complex feedback loops which allow precise control of the field in each cavity. Tight control of the field avoids any problems with coupling between cavities that could occur if one klystron was feeding more than one RF cavity.

In total, there are 16 400MV klystrons delivering a total of 4800 kW of power to the superconducting cavities. These cavities operate at 4.5 K, and use a similar liquid helium cooling system to that of the magnets. Though the RF cavities achieve a maximum voltage of 16MV, some of that energy is used to control the beam. In reality, each time the proton beam passes through the RF cavities, an energy of ∼485KeV is imparted to the beams. At a revolution frequency of over 11,000 times a second, this means it takes the beam about 20 minutes to ramp up from 450 GeV to the full 7TeV.

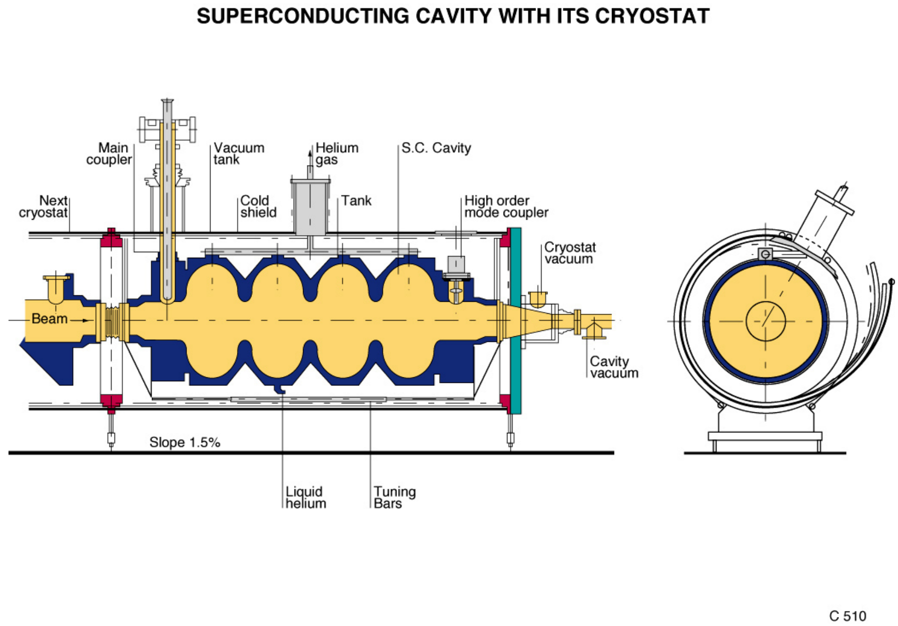


Figure 3.13: Schematic of a cryostat used at the LHC to accelerate the proton beams through 4 RF cavities with 2MV potentials each. 2 of these cryostats are used in succession on each beamline [10].

Chapter 4

Compact Muon Solenoid

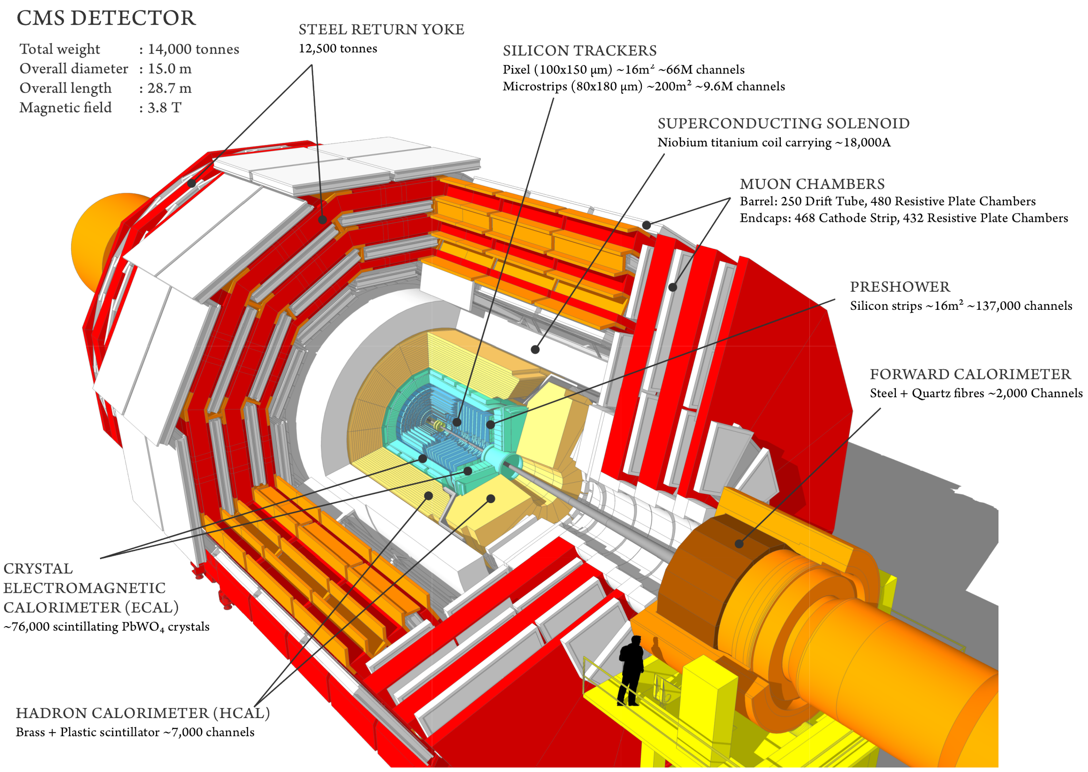


Figure 4.1: A cutaway diagram of the CMS detector, with subsystem statistics [11].

The Compact Muon Solenoid (CMS) experiment is one of two general purpose particle detector operating at Point 5 on the LHC ring, sitting opposite A Toroidal LHC Apparatus (ATLAS) at Point 1. It is capable of a wide range of physics measurements, including the identification and reconstruction of charged and neutral hadrons, photons, electrons, muons, and taus. Its 4π hermetic coverage surrounding the Interaction Point (IP) also allows for measurements of neutrinos through identification of a momentum imbalance in the measured collision. CMS was built in 15 sections on the surface, which were then lowered into place and assembled in a large underground cavern near Cessy France [76].

CMS was built with 4 primary design goals:

1. Good muon identification and momentum resolution over a wide range of momenta and angles, as well as good dimuon mass resolution;

2. Good charged-particle momentum resolution and reconstruction efficiency in the inner tracker. Efficient triggering and offline tagging of τ's and b-jets, requiring pixel detectors close to the interaction region;

3. Good electromagnetic energy resolution, good diphoton and dielectron mass resolution ($\sim$1% at 100 GeV);

4. Good missing transverse energy ($\not\!\!E_T$) and dijet mass resolution.

CMS takes its name from the design and detector characteristics that comprise it. The 'Compact' part of its name refers to relatively small size of CMS compared to other modern particle detectors, with a total length of 28.7 m and a diameter of 15 m. Although CMS is the size of 4 story building, it still qualifies as 'compact' when comparing it to its counterpart at CERN, ATLAS, which is 44m long and 15 m tall. The word 'Muon' in the name refers to CMS's ability to detect and reconstruct muons by using three muon detection systems which provide superior p_{T} and time resolution for muons. Lastly, 'Solenoid' refers to the enormous solenoidal magnet that makes up the heart of CMS. This solenoid is 13m long with an inner diameter of 6m, and provides a uniform field of 3.8 Tesla across it's interior. This strong field provides tremendous bending power allowing CMS to precisely measure charged particles. Unless otherwise stated, all technical information in this chapter is taken from [14].

The coordinate system of CMS is centered on the nominal IP. From here, $\hat{y}$ points directly up to the sky, $\hat{x}$ points toward the center of the LHC ring, and $\hat{z}$ points counter-clockwise along the LHC ring. In polar coordinates $\hat{r}$ is defined as the direction radially outward from the IP, $\hat{\phi}$ is the azimuthal angle measured relative to the positive x-axis, and $\hat{\theta}$ is the polar angle measured with respect to the positive z-axis. Figure 4.2

The pseudorapidity, η, is defined as -ln($tan(\theta/2)$), which is a good approximation of the rapidity of relativistic particles (y):

$$y = \frac{1}{2}\ln\left(\frac{E + p_z c}{E - p_z c}\right) \tag{4.1}$$

The components of momentum and energy transverse to the beam line, p_{T} and E_{T}, are defined as p_{T} = —p—cos(ϕ) and E_T = Ecos(ϕ).

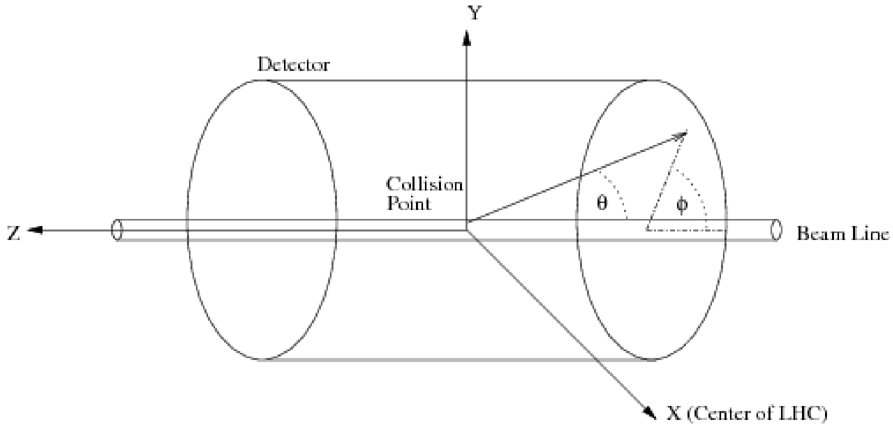


Figure 4.2: The CMS coordinate system, reprinted from [12].

CMS is composed of multiple sub-detectors that are arranged in concentric cylindrical layers surrounding the IP of *pp* collisions. The sub-detector closest to the beam-line is the tracker, which is made up of 3 layers of silicon pixel detectors followed by 10 layers of silicon strip detectors. The pixel dector portion has an inner radius of 4.4cm, and has a total η coverage up to η=2.5. It also plays an important role in determining the location of the IP, the impact parameters of charged particles, and measurement of displaced vertices (the location of a particle decay some distance from the IP) which are critical for the identification of b-quarks.

Beyond the silicon tracker lies the Electromagnetic Calorimeter (ECAL), which absorbs energy from electromagnetically interacting particles. The ECAL is made up of a single layer of lead-tungstate ($PbWO_4$) crystals which act as both an absorption and scintillation medium. The ECAL is split into 2 pieces, the barrel and the endcap, which together cover a region up to $\eta = 3$. The last layer of detector inside the solenoid is the Hadronic Calorimeter (HCAL). The HCAL uses brass absorber plates combined with a plastic scintillator to sample the energy of hadrons, while steel plates form the inner and outermost plates for structural strength. Four sub-detectors combine to create the HCAL, the hadron barrel (HB), the endcap (HE), the outer (HO), and the forward (HF) calorimeters. All of these except for the HO are located inside the solenoid.

The solenoid itself provides a uniform 3.8T field that bends particles as they traverse the detector, allowing for accurate measurements of a particle's momentum. Outside the solenoid is a large iron return yoke weighting 10,000 tons which serves to increase the field homogeneity in the tracker volume and to reduce the stray field by returning the magnetic flux of the solenoid [77]. The yoke is interleaved with the muon system, which used three types of detectors: drift tubes in the barrel region and cathode strip chambers in the endcap to read out muon tracks, while resistive strip chambers are used throughout to provide independent trigger and timing measurements. Figure 4.3 shows a representation of a wedge of the CMS detector and how different types of particles deposit their energy in the sub-systems.

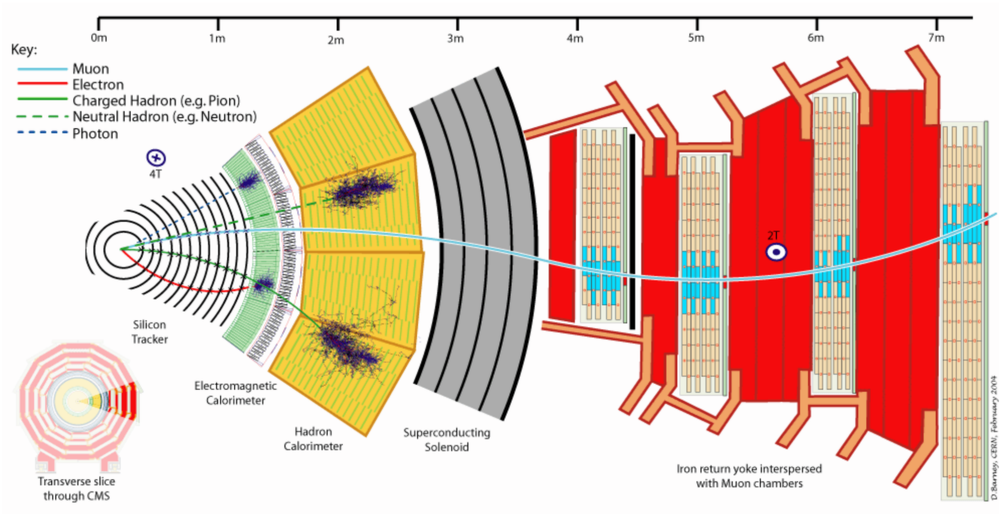


Figure 4.3: A slice of the CMS detector showing how particles interact and deposit energy with the various sub-systems. The tracker measures the trajectory of charged particles; electrons and photons are measured/absorbed in the ECAL, while hadrons deposit most of their energy in the HCAL. Muon chambers measure the trajectory of muons, the distinct double curve is due to the opposite magnetic field they feel outside of the solenoid. Reprinted from [13].

At the design center of mass energy ($\sqrt{s}$= 14 TeV), the total *pp* cross section is expected to be roughly 100mb, resulting in approximately 10^9 collisions per second. This is a staggering large number of events, so many that they can not all be analyzed. An online selection process knows as 'triggering' reduces this number to about 100 events/second for storage and later analysis. The rest of this chapter will discuss the main sub-systems in further detail.

4.1 The Superconducting Solenoid

The free bore solenoid magnet at the heart of the CMS detector is 6m in diameter and 12.5 meters in length, storing an energy of 2.6 gigajoules at full current. This magnet is distinctive because it constructed of a 4-layer winding of stabilized reinforces NbTi superconductor. The strain on the magnet material from magnet pressure is much larger than previous detector magnets, and required that a large fraction of the CMS coil have structural function. These coils, and their reinforcements, make up the 200 ton cold mass of the solenoid. The cold mass operates at a temperature of 4.6K, requiring it's own vacuum and cryogenics systems. As mentioned above, the iron return yoke weighs 10,000 tons, accounting for the majority of the weight of the entire CMS detector.

4.2 The Tracker

The inner tracking system of CMS was designed to provide precise measurements of the trajectories of charged particles, as well as precise reconstruction of secondary vertices required for τ and b-jet reconstruction. The intense flux of particles at LHC design luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$ results in an average of $\sim$1000 particles from 20 overlapping pp collisions for each bunch crossing. This corresponds to a hit rate density of 1 MHz/mm^2 at a radius of 4 cm, 60 kHz/mm^2 at 22 cm, and 3 kHz/mm^2 at 115 cm from the beam-line. The intense flux of particles will also cause severe radiation damage to the tracking system, necessitating a radiation hard detector.

In order to keep the occupancy of the detector low, below 1%, a pixelated detector is needed for radii <10cm. Beyond 10cm a silicon strip dectector is used up to a radius of 1.1m. Additionally, in order to reduce the signal to noise ratio the detector is operated at a temperate of -10°C. Operating at this temperature the signal to noise ratio of 10:1 is achieved. Together, the pixel and strip detector have an acceptance of $|\eta| < 2.5$, and with $200m^2$ of active silicon they make up the largest silicon tracker ever built. Figure 4.4 shows a side view of the tracker with sub-systems labeled.

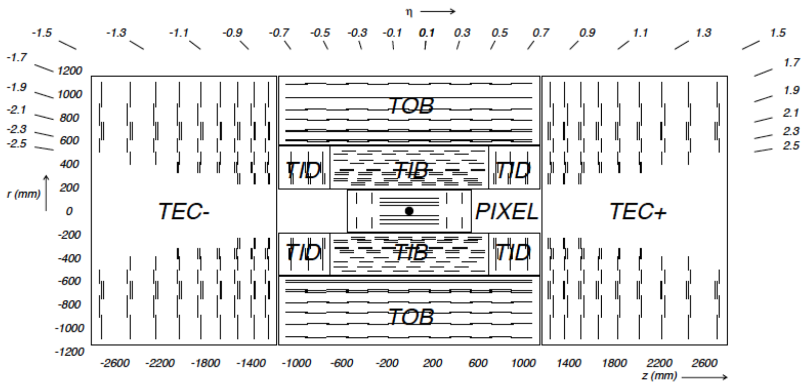
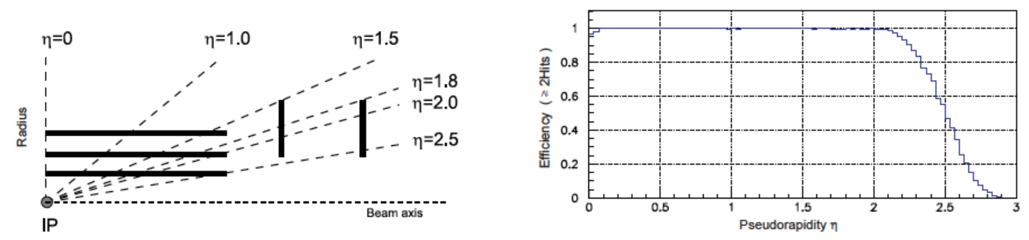


Figure 4.4: A side view of the tracker subsystem. The pixel detector forms the innermost layers, with three concentric rings of detectors in the barrel and two in the endcap. The tracker inner barrel (TIB), tracker inner disks (TID), tracker outer barrel (TOB) and tracker end caps (TEC) are composed of concentric layers of silicon strip detectors [14].

4.2.1 The Silicon Pixel Detector

The pixel detector is composed of three barrel layers at radii between 4.4cm and 10.2cm. It is completed by an endcap section consisting of 2 disks. It is the closest system to the interaction region, and is responsible for the small impact parameter resolution and secondary vertex reconstruction. Each pixel cell is $100{\times}150\mu m^2$. Together, the barrel layers and endcap disks contain 66 million pixels, and are arranged to provide 3 tracking points for nearly the entire $|\eta| < 2.5$ range. Figure 4.5 shows the layout of the barrel and endcap layers.



(a) (b)

Figure 4.5: Geometric layout of the pixel detector (a) and hit coverage as a function of η(b) [14].

The proximity to the interaction region required a radiation tolerant design. This lead to an n+ pixel on n- substrate detector design allowing for partial depleted operation at very high particle fluences. Additionally, the magnetic field induces a Lorentz drift on the electrons in the ϕ direction. This results in the charge from one pixel being shared among neighboring pixels. Particle hits are reconstructing by reading out the analog pulse height of a pixel and interpolating among multiple pixels, leading to a spatial resolution of 15-20 μm.

The pixels are grouped into multi-pixel sensors, a grid of 52×80 pixels. Each grid has a readout chip (ROC) attached to it, which serve to amplify and buffer the charge from the grid while providing zero suppression to the pixel sensor. A Token Bit Manager (TBM) chip controls the readout from multiple ROCs, providing Level-1 trigger and clock information to the ROCs. Finally, the signal is digitized and read out by a pixel front end digitizer (pxFED).

4.2.2 The Silicon Strip Detector

The silicon strip detector is made up of four components (shown in figure 4.4): the tracker inner barrel (TIB), the tracker inner disks (TID), the tracker outer barrel (TOB), and the tracker end caps (TEC). The TIB, TID, and inner four rings of the TECs are comprised of sensors with a thickness of 320μm, while the outer 3 rings of the TEC and the TOB are made of 500μm thick sensors. The thicker sensors compensate for the increased capacitance in the outer strips due to their increased length, and serve to maintain a signal:noise ratio of at least 10:1 everywhere. There are a total of 15,148 detector modules that are distributed as shown in the longitudinal cross section of figure 4.4.

The TIB and TOB are arranged in straight rows along $\hat{z}$, with repeating rows covering the full 2π extent of ϕ. The TIB consists of 4 concentric cylinders with radii ranging from 255mm to 498 mm, while the TOB consists of a single mechanical wheel made of 688 self contained 'rods' providing support and cooling for 6 or 12 silicon modules. The three TID disks are arranged into three concentric circular rings of increasing r. The TEC modules are are affixed to wedges in ϕ called 'petals', with nine petals needed to cover all of ϕ. Figure 4.6 shows a schematic of the TIB/TID sub-assembly. The signal to noise ratios achieves in the TIB and TOB are shown in figure 4.7

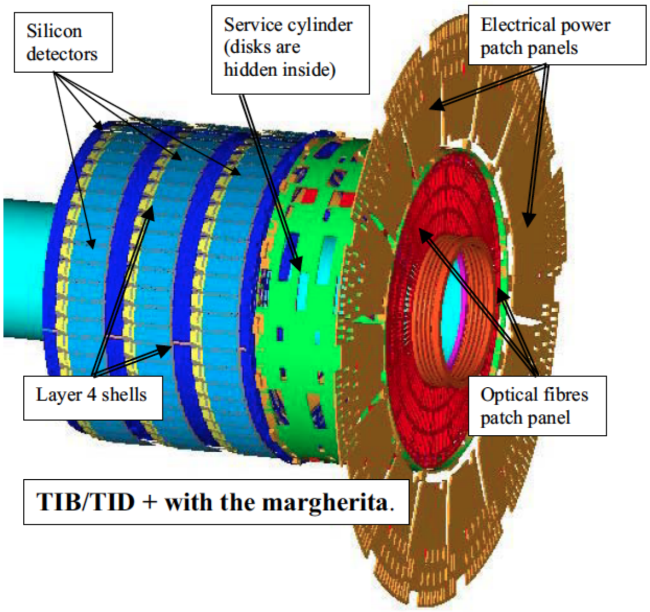


Figure 4.6: Schematic drawing of a TIB/TID assembly. This structure is mounted inside the TOB (one on each end), and shows the 'margherita', a service distribution disk used to route out signals and control the cooling supply lines [14].

The silicon strips are wire bonded to a readout chip called an APV25, which have 128 channels. Two APD25 chips are multiplexed to one read out channel, meaning that strips can only be read out in multiples of 256. The APD25 chips serve to amplify, shape, and buffer the signals before they are read out to a front end driver (FED) system. The superior performance of the tracker over the hadronic calorimeter for low energy charged hadrons has been exploited in the particle flow $\not\!\!E_T$ and jet reconstruction techniques, which is described in section 5.2.

4.3 The Electromagnetic Calorimeter

The Electromagnetic Calorimeter (ECAL) is composed of 75,848 lead tungstate ($PbWO_4$) crystals, divided into a barrel section (EB) and two endcap (EE) sections. In the barrel, 61,200 crystals are arranged in grids of 20×85 in order to cover the entire $\phi \times \eta$ section. These 1700 crystal grids are called supermodules (SM), laid out end to end there are 36 total SM needed to cover the barrel. Each endcap is composed of 7,324 crystals clustered in 5×5 crystal groups called superclusters (SC). The crystals have a fast response, provide fine granularity, and are radiation resistant, making them ideal for the LHC environment. In the endcaps, an additional detector, the preshower, provides additional spatial resolution with silicon microstrip detectors, similar to those in the tracker. Figure 4.8 shows the layout of the ECAL.

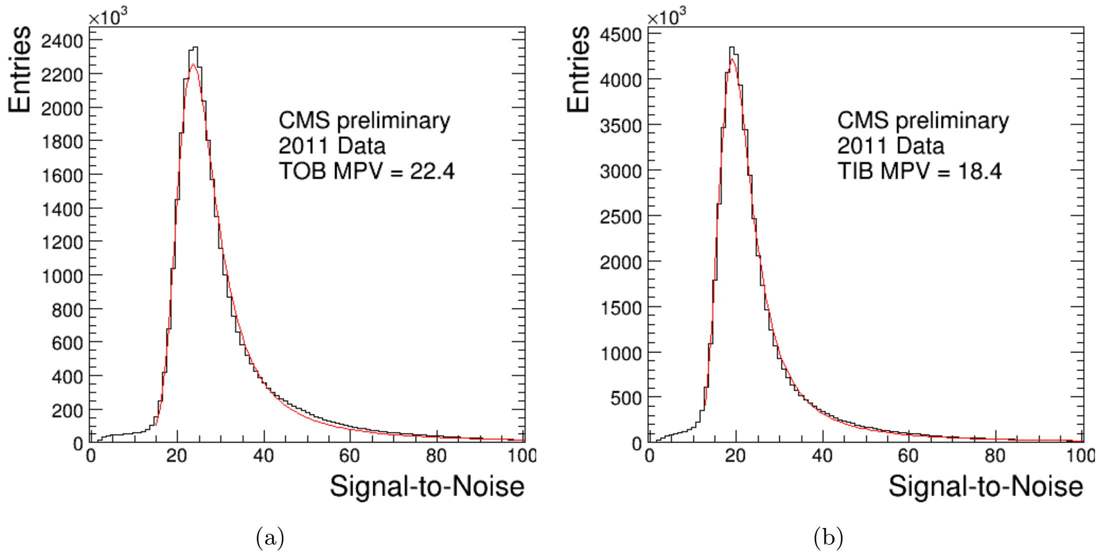


Figure 4.7: Signal to noise measurement for the TOB (a) and TIB (b) sections of the silicon tracker, operated in deconvolution mode (optimal conditions) [15].

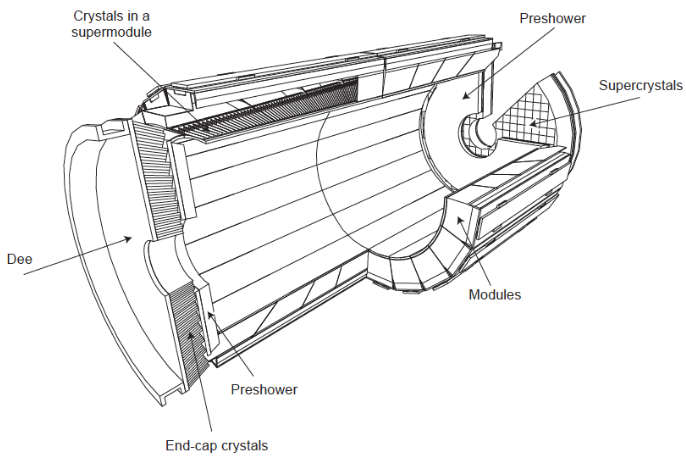
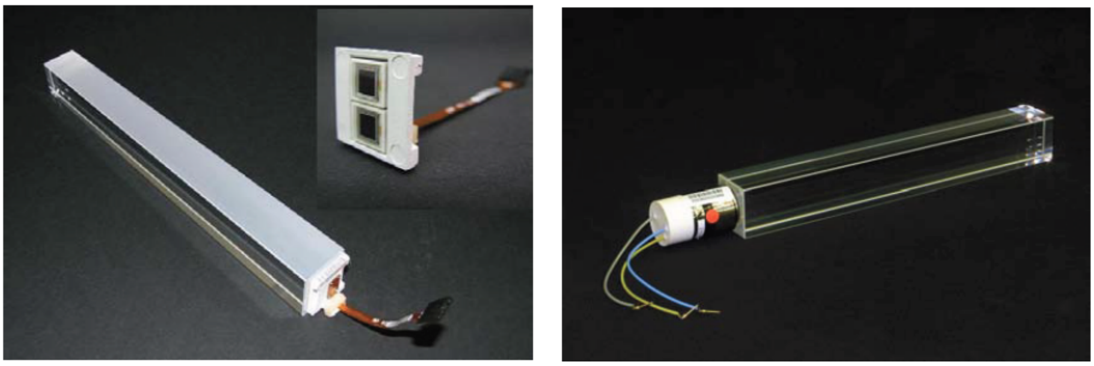


Figure 4.8: Layout of the ECAL sub-detector [14]

Lead tungstate is a great material for electromagnetic calorimetry. It has a high density, 8.28 g/cm^3, short radiation length (χ_0=0.89cm), resulting in increased likelihood of a transiting particle to interact with the crystal as well as containing many radiation lengths in a single crystal. The EB crystals (front face) sit at a distance of 1.29m from the IP and are slightly tapered, with the front face measuring $22 \times 22mm^2$ and the rear face measuring $26 \times 26mm^2$. The crystals have a length of 230mm, corresponding to 25.8 χ_0. The EB crystals are slightly different, with a front face cross section of $28.62 \times 28.62mm^2$ and a rear face of $30 \times 30mm^2$.

As a charged particle or photon begins to deposit energy in a crystal, it begins a process known as an electromagnetic shower, where it fragments into many lower energy photons and electrons. Particles (such as electrons) which are bent by the magnetic field in CMS create bremsstrahlung photon radiation. The intensity of this radiation is inversely proportional to the mass of the particle squared, so due to the short radiation length of $PbWO_4$ anything heavier than an electron will pass through the crystal without loosing much energy. Additionally, the crystals have a small Moliere radius, 2.2cm, which is the radius of a cylinder that encloses of 90% of the electromagnetic shower's energy deposition. Given the geometry of the crystals in the EB and EE this means that a small grid of crystals we receive all of the energy deposited by a high energy photon or electron.

$PbWO_4$ crystals also have very useful scintillation properties. They are optically clear, emitting a blue-green scintillation light with a broad maximum at 420-430nm. Additionally, the scintillation decay time is the same order of magnitude as the minimum bunch crossing time of the LHC, with 80% of the scintillation light emitted in 25ns. The light output of the crystals varies with temperature, requiring a precise 18°C operating temperature. The light output from the crystals is collected by photodetectors attached to the ends, avalanche photodiodes (APDs) are used in the barrel while vacuum phototriodes (VPTs) are used in the endcaps. Figure 4.9 shows examples of barrel and endcap crystals with attached photodetectors.



(a) (b)

Figure 4.9: Example of lead tungstate crystals with photodetectors attached. (a)A barrel crystal with APD attached. (B)An endcap crystal with VPT attached. [14].

Although $PbWO_4$ crystals are radiation resistant, they still suffer from transparency loss due to radiation-induced lattice damage, as shown in figure 4.10. Additionally, any unforeseen changes in gain due to changes in the amplifier or photodetectors will degrade the ECAL resolution. To account for this, a calibration system is in place in the ECAL that uses laser and LED pulses to compute corrections to the crystal gains.

The preshower detector is a two-layer sampling calorimeter that sits in front of the ECAL end-caps. Lead radiators initiate electromagnetic showers from electrons and photons, and silicon strips are placed behind them to measure trajectories and deposited energy of passing particles. The goal of the preshower detector is to identify neutral pions in the endcaps (which have a much higher multiplicity here than in the barrel due to the endcap being a higher η region near the beam-line).The total thickness is 20cm, which corresponds to a 2 radiation lengths in the first

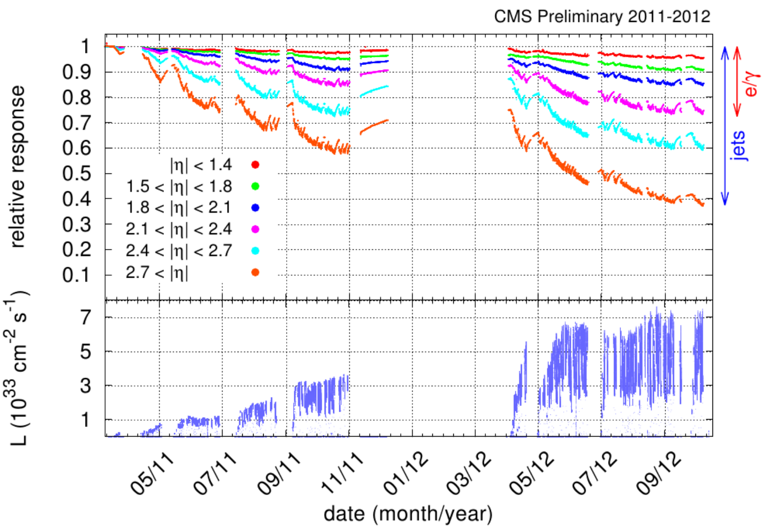


Figure 4.10: Relative response to laser light (440 nm) measured by the ECAL laser monitoring system, averaged over all crystals in bins of pseudorapidity, for the 2011 and 2012 data taking periods Layout of the ECAL sub-detector [16].

layer, and another radiation length in the second layer. The lead layer causes 95% of photons are converted to e^+e^- pairs after passing through. The preshower, like the ECAL endcaps, are formed into two Dees (one on either side of the beam-pipe), and covers the region $1.653 < |\eta| < 2.6$.

The EE is made up of two endcaps which are each separated into two halves called Dees, covering a region of $1.479 < |\eta| < 2.6$. Each Dee holds 3,662 crystals contained in 138 5×5 SCs, as well as 18 special partial supercrystals on the inner and outer circumference. The EE sits its a longitudinal distance of 315.4 cm from the nominal interaction point, with the crystal faces focused at a point 1.2m beyond the interaction point. Figure 4.11 shows an ECAL Dee with the crystals installed, grouped into SCs.

Read-out of the ECAL has to be able to acquire the small signals from the photo-detectors with high speed and precision. The on-detector electrons are designed to read a complete trigger tower (5×5 crystals in $\eta \times \phi$) or a super-crystal for EB and EE respectively. It is made up of five Very Front End (VFE)boards, one Front End (FE) board, two (EB) or six (EE) Gigabit Optical Hybrids (GOH), one Low Voltage Regulator Card (LVR) and a motherboard. Once triggered, the APD (or VPT in the EE) is sampled 10 times at a 40 MHz sampling rate, and amplified by a multi-gain amplifier (MGPA),with nominal gains of 1, 6, and 12 contained on the VFE. These digitized samples are sent to the FE board, where they are buffered (for $\approx 3\mu s$) before receiving the Level-1 trigger, where they are sent to the off-detector electronics Data Concentrator Card (DCC) via the GOHs.

Figure 4.11: One half of an ECAL end-cap, called a Dee. Two Dees form a disk with an inner bore for the beam-line to pass through. 5x5 supercrystal modules are mounted in preparation for installation at CMS [14].

4.4 The Hadronic Calorimeter

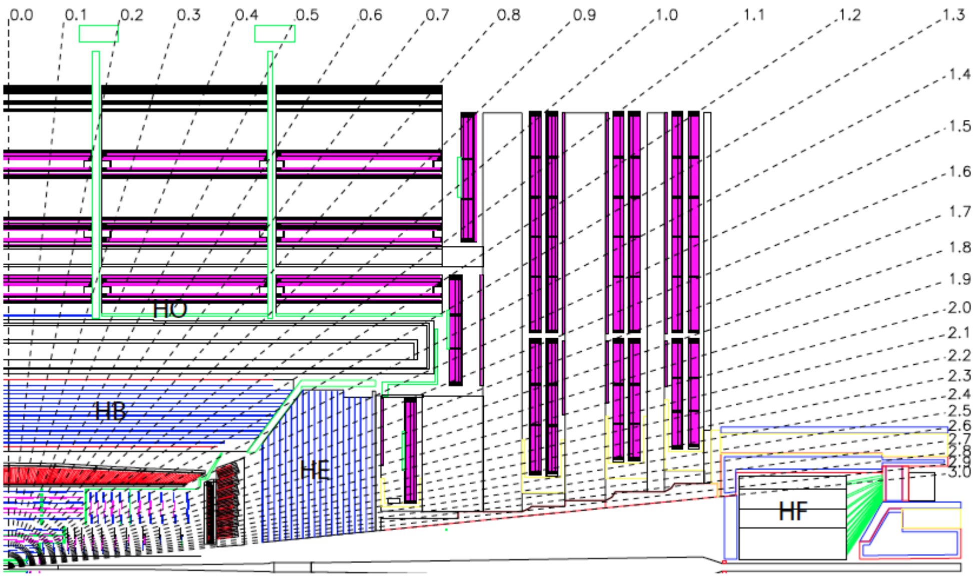


Figure 4.12: Longitudinal cross-section of the HCAL with the four sub-systems labeled [14].

The Hadronic Calorimeter (HCAL) is is divided into four sub-systems shown in figure 4.12. The barrel (HB), the endcap (HE), the outer calorimeter (HO), combine to cover a region from $|\eta| < 3$, while the forward calorimeter (HF) extends the coverage out to $|\eta| < 5.2$. The HCAL barrel and endcaps sit behind the tracker and ECAL systems, as seen from the interaction point. The hadron calorimeters are particularly important for measuring jets as well as neutrinos through the measurement of missing transverse energy ($\not\!\!E_T$).

The geometry of CMS restricts the HCAL radially, as it must fit between the outer edges of the ECAl (R=1.77m) and the inner surface of the solenoid (R=2.95m). This resulted in the HB located inside the magnet coil, while an the HO was placed outside the coil to complement the measurements from the HB. The HB itself is divided into two half-barrel sections, covering a pseudorapidity range $|\eta| < 1.3$. It is a sampling calorimeter made 14 layers of brass absorber plates alternated with plastic scintillator tiles, and uses steel plates on the front and rear layers for structural support. The brass absorber plates are C26000/Cartridge Brass, chosen to maximize the number of interaction lengths, as well as having good physical properties and reasonable cost. The HB is constructed of 36 azimuthal wedges (shown in figure 4.13 which are bolted together ins such a fashion as to minimize the crack between them to less than 2mm.

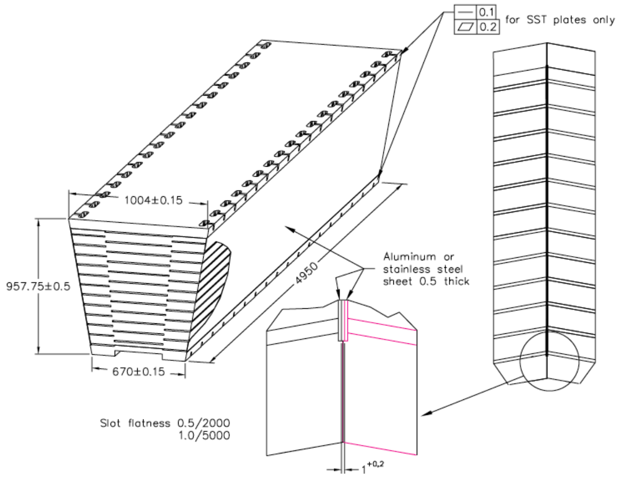


Figure 4.13: Isometric view of HCAL segment [14].

When a hadron passes through the HCAL, the brass and steel plates absorb energy and initiate the decay of the hadron into a number of lighter particles. These particles pass through the scintillator layers, which absorb energy from the interactions or collisions with the passing particles. The scintillator then emits light in the blue-violet range of the visible spectrum proportional to the amount of energy absorbed by the scintillator. These photons carried out by wavelength shifting fibers (WSFs), which absorb and re-emit the light in green part of the visible spectrum. The brass absorbers have a nuclear interaction length, or the length necessary to reduce the number of charged particles in a hadron shower by 1/e, of 16.42 cm, and a radiation length of 1.49 cm. This results in the HB containing a large part of most hadron showers produced at LHC energies, though a small portion will still pass through the entire radial distance.

The outer barrel layer, HO is designed to measure any part of the hadron shower that passes through the HB. It is located outside of the solenoid, and is composed of an absorber layer equal to $1.4/\sin\theta$ interaction lengths. The HO is separated into 5 sections along the z-axis, with all but the center section having one layer of absorber. The central section corresponds to $\eta = 0$, meaning it has the minimal amount of absorber depth under it, so two layers of absorber are used here.

The endcap calorimeter, HE, covers a substantial portion of the rapidity range, $1.3 < |\eta| < 3.0$. This region contains $\approx 34\%$ of the final state particles produced in collisions. The high luminosity of the LHC requires that the HE be very radiation tolerant, and its location inside the end of the solenoid requires the use of a non-magnetic material. Like the HB, the HE is composed of C26000 cartridge brass found, and is also a sampling calorimeter made by alternating absorber and scintillator layers. Interestingly, the construction of the HE was the responsibility of the Russian and Dubna groups, and the high quality of brass needed was difficult to find. Eventually, they found the brass they needed in World War II navy artillery shells, which were melted down and used to form the HE plates [78]. Figure 4.14 shows a pile of these shells before being melted down.

Figure 4.14: Russian navy shells re-used in the CMS Hadron Calorimeter [17].

The forward calorimeter (HF), located 11.2m from the interaction point, extends the HCAl coverage from $3.0 < |\eta| < 5.0$. It also experiences the greatest particle fluxes at the highest energies, receiving an average of 760 GeV per pp collision compared to 100 GeV the rest of the HCAL absorbs. The HF uses a Cherenkov-based, radiation hard, technology which utilizes fused-silica core quartz fibers as the scintillating medium. The HF consists of a steel absorber structure with grooved plates, with the quartz fibers inserted into these grooves. Thirty-six wedges form a cylindrical detector around the beam pipe, with the fibers transporting the light output photo-multiplier tubes housed in a read-out box. A diagram of the HF is shown in figure 4.15

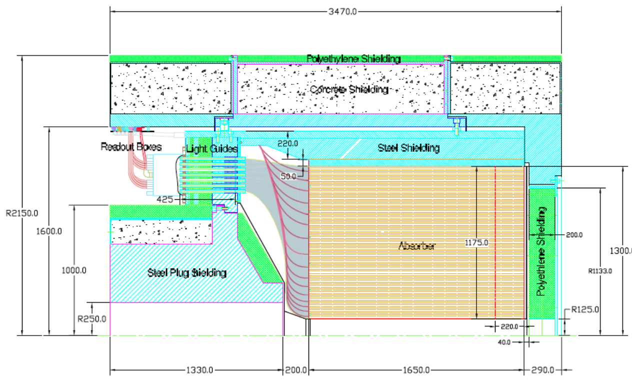


Figure 4.15: Cross-sectional view of the HF calorimeter. The beam-line runs parallel to the diagram, with the right side pointing toward the interaction point. Fibers run parallel to beam-line in the absorber, are bundled together (gray region) and routed to the read-out boxes [14].

4.5 The Muon Chambers

Muon detectors is a very useful tool for recognizing signatures of interesting processes over high background rates due the relative ease of detecting muons combined with the low rate of radiative loss muons experience in the tracker material compared to electrons. The CMS muon system has three main functions: muon identification, momentum measurement, and triggering. Momentum resolution and triggering capability are enabled by the large magnetic field, while a large material thickness that the muons travel through allows for a high likelihood of identification.

The muon system is composed to three types of gaseous detectors: drift tubes (DTs), cathode strip chambers (CSCs), and resistive plate chambers (RPCs). Like the other sub-detectors, the shape of CMS lends itself to have a cylindrical barrel section and 2 planar endcap sections. In the barrel region, the muon rate is low and the magnetic field is uniform, a topology in which drift chambers are well suited. The DTs are organized into 4 stations which are interspersed among the layers of the steel return yoke, and cover the pseudorapidity region $|\eta| < 1.2$.

The barrel region is divided into 5 longitudinal, cylindrical sections around the beam-line, known as wheels. In each wheel there there 4 concentric layers of drift tube stations, one on either side of the magnet return yoke, and two interspersed inside of it. Each wheel is divided into 12 azimuthal sections, making 48 sections in the barrel, as shown in figure 4.16. The inner three layers each have two sets of 4 chambers that measure the $r - \phi$ bending plane as well as 4 chambers that provide measurement in the $\hat{z}$ direction. The fourth layer does not contain any measurement on the z-plane.

Each DT is made up of 4 layers of rectangular drift cells combined in either 2 or 3 superlayers, with each superlayer staggered by half a cell. A single drift cell is comprised of a hollow 13×42 mm tube, with a 1.5mm wall to provide isolation between adjacent cells. This means the

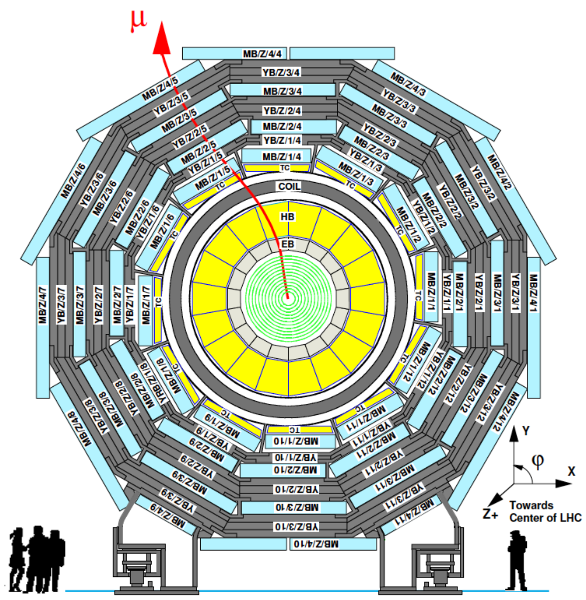
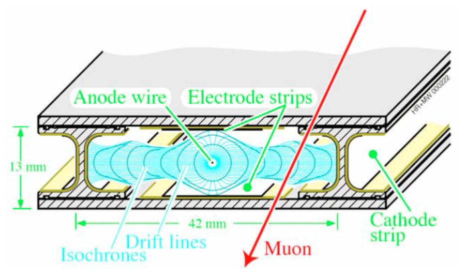


Figure 4.16: The layout of muon DT chambers in one of the 5 wheels that make up the muon system [14].

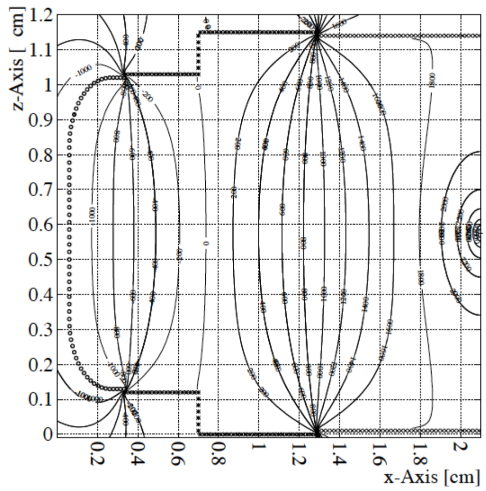
maximum drift time in each cell is 380ns, a small enough value to produce negligible occupancy in the muon system. The thickness of the walls also provides a decoupling of the several layers of cells in each tube, a function that helps with reconstructing high p_T muons. The cell is filled with a mixture of 85% argon + 15%CO_2 gas mixture, and contains a gold plated anode wire that is held at 3600V that runs down the center of the cell. The walls of the cell are held at 1800 V or -1200 V, creating a electric field across the chamber. When a muon passes through the chamber, it's charge ionizes molecules of the CO_2 gas, causing the electrons to drift towards the anode wire, and the CO_2 ions drift towards the wall. As the electrons approach the anode, they are accelerated and liberate secondary electrons from other CO_2 molecules, creating an avalanche of electrons near the wire, resulting in a drop in voltage as they are collected. The voltage drop is read out by front end electronics as a signal that a muon has passed through the chamber. Figure 4.17 shows a sketch of a single drift cell and the equipotential lines in it.

In addition to the drift tubes, resistive plate chambers (RPCs) are used in the barrel to complement the muon measurement. Unlike the long drift time associated with DTs, RPCs can tag the time of an ionizing event in much less that 25ns. Signals from the RPCs provide time and position information about a muon hit, and associate the muon hit with a specific bunch crossing. RPCs are located on the inner and outer surfaces of the first two drift stations in the barrel, and only on the inner surface of stations 3 and 4. In the endcap, RPCs are mounted on both faces of the muon endcap system.

The muon endcap is composed of cathode strip chambers (CSCs). Each endcap is composed of 4 layers of trapezoidal CSCs (shown in figure 4.18) with 468 CSV distributed in each endcap.



(a)



(b)

Figure 4.17: (a)Sketch of a DT cell showing drift lines. The top and bottom plates are held at ground potential. (b) Equipotential lines for half of a drift cell. The anode wire is located on the right side of this plot[14].

A CSC is a multiwire proportional chamber, comprised of 6 anode wire planes interleaved among 7 cathode panels. The wires run azimuthally and define the track's radial coordinate, and the space between panels is filled with a gas that is 40% Argon, 50% CO_2, and 10% CF_4. The wires are held at a positive voltage while the strips are held at negative voltage, so when a muon ionizes the gas the positive ions drift toward the anode as in the DTs. Unlike DTs, an image charge is induced in the cathode strips which, which provides a 2-d measurement of the muon (r and ϕ).

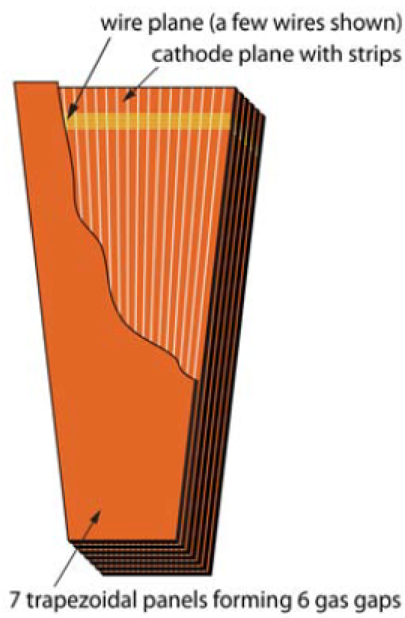
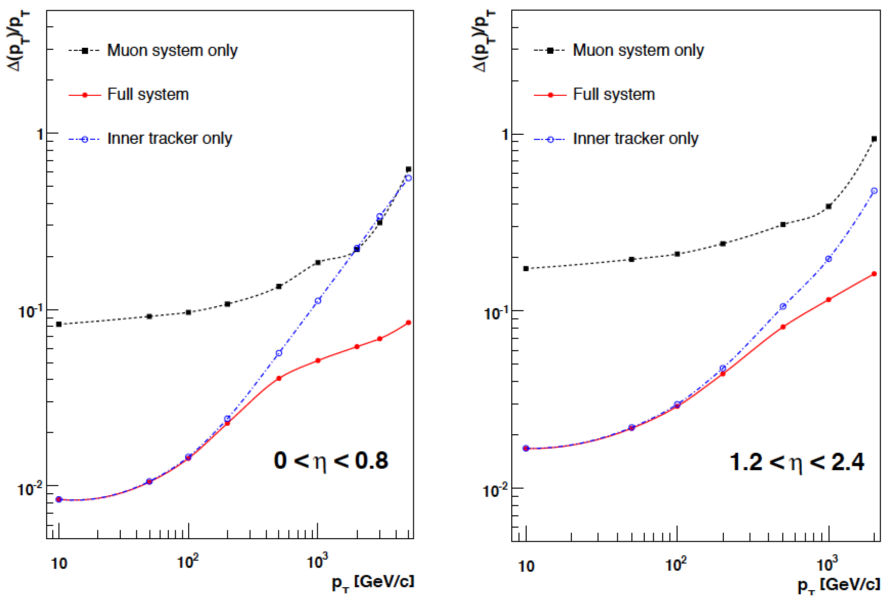


Figure 4.18: Muon CSC wedge, showing cathode panels in orange and anode wires in gold [14].

CSC track information can be combined with tracks from the inner tracking system to form more precise tracks than either system could form on it's own (figure 4.19). Each CSC has at least a 99% efficiency per chamber of finding track stubs based on it's trigger, and at least a 92% probability per chamber of identifying the correct bunch crossing in which the muon originated. Since each muon track consists of 3-4 CSC hits, the correct bunch crossing is identified more than 99% of the time. Additionally, CSCs can provide up to $75 \mu m$ off-ine spatial resolution in $r - \phi$.



(a) (b)

Figure 4.19: Muon transverse momentum resolution as a function of p_{T} using the muon system only (black), the inner tracker only (blue), or both (red). Shown for $|\eta| < 0.8$ (a) and $1.2 < |\eta| < 2.4$ (b)[14].

Chapter 5

Event Reconstruction

5.1 Data Acquisition

The CMS Data Acquisition (DAQ) and trigger system was specifically designed to collect and analyze data at a rate of 40MHz (corresponding to a 25 ns collision rate). The large design cross section of the LHC results in many overlapping collisions, with many of these collisions being the result of secondary interactions at lower energies than the main collision, and are unlikely to contain interesting information. Additionally, due to electronics limitation though, only a few 100 Hz of events can be recorded for later processing. To reduce the rate of events a trigger system is used, which aims to identify interesting events with the potential to reveal new physics. In order to achieve this reduced rate CMS makes use of a two level system employing a Level-1 (L1) trigger in the detector electronics, and a High Level Trigger (HLT) composed of an underground computer farm that performs a more sophisticated reconstruction.

5.1.1 L1 Trigger and HLT

Each sub-detector system has a piece of electronics called a front-end system (FED) whose job is to continuously store the 40MHz data in pipe-lined buffers. These buffers are sufficiently large to store information for $3.2\mu s$, which corresponds to the abort gap of the LHC proton beam. During this gap an L1 trigger is formed synchronously via a timing system, and the data is read into the DAQ system. The DAQ system itself is composed of 8 'slices', each of which is an identical autonomous system capable of handling 12.4 kHz event rate. The L1 trigger is responsible for cutting the event rate down to 50kHz before it is passed on to the HLT, which is achieved through use of a global trigger that can execute up to 128 separate trigger algorithms in parallel to analyze the event kinematics to search for minimally interesting criteria.

The HLT itself is a computer farm that uses information compiled by the detectors to build a more complete reconstruction of an event, using specific criteria to make cuts on the incoming data. The design rejection factor of the HLT is 1000, reducing further the rate of events that are stored. The HLT has many separate paths that an event can pass in order to make it into a particular dataset. They form separate datasets based on the trigger applied: single muon, single electron, diphoton, etc. Once an event has passed the HLT, it is transferred to the CERN tier-0 prompt reconstruction facility where the raw detector data is converted into a data storage file type that can be accessed by physicists.

5.1.2 T1 sites and data storage

CMS computing operates on a tiered computing structure. A tier-0 computing center is located at CERN where the data is transferred from the HLT and a first set of reconstruction occurs. From there, it is transferred to one of six Tier-1 computing centers located around the world. At the Tier-1 centers a full reconstruction of the data is performed, and the data is stored there. Physicists in the collaboration can access some of the data at Tier-1 centers for processing and storage at Tier-2 centers, of which there are 25, which typically house the final data files. Figure 5.1 shows a diagram of the tier system and the data-flow between the tiers.

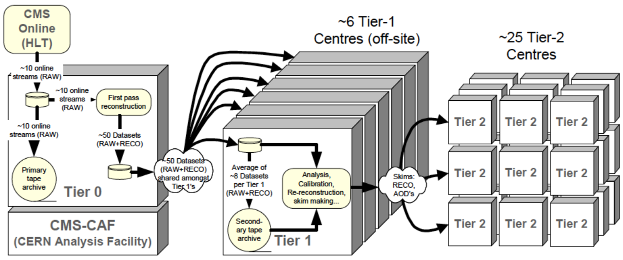


Figure 5.1: Diagram showing the data-flow through the CMS tiered computing system [14].

The data itself also is processed in three data tiers. The first layer of this is the RAW data, which is created by unpacking detector streams passed on from the L1 and HLT, typically formed of light measurements from the different calorimeters and additionally information provided by the L1 trigger. This RAW data is reconstructed (as will be described below) into physics objects that can be grouped and analyzed. This new form of the data is known as RECO, for reconstructed, and it stores the detector information as well as the physics object information. After RECO, an analysis object data (AOD) is formed from a subset of the RECO information. AOD objects are typically comprised of only high-level physics objects, making them much smaller files. These AOD datasets are shipped to the Tier-2 centers where physicists can access them and being their analyses.

5.2 Particle Flow Event Reconstruction

With all of the data measured and stored at CMS coming in the form of electronic signals, the reconstruction of these signals into real physics objects is paramount to the success of the detector. At CMS, an algorithm known as particle-flow is used to reconstruct and identify all stable particles produced in an event: electrons, muons, photons, charged hadrons and neutral hadrons [79]. Once these objects are identified, their information is used to build jets, determine the missing transverse energy ($\not\!\!E_T$), reconstruct and identify taus, identify b-jets, and many other calculations.

The CMS detector is ideally suited to identify and separate these particles. The ECAL granularity allows for excellent energy resolution of photons, and its nearly hermetic design and location inside the magnetic field allows photons to be separated from charged particle deposits. Although the HCAL is 25 times coarser than the ECAL, which on its own would not allow spatial separation of charged and neutral hadrons from high p_{T} jets, combining calorimeter information with angular and energy resolution of the tracker gives superior reconstruction. Electrons are reconstructed from a combination of tracks and energy deposits in the ECAL, both from the electron itself and from Bremsstrahlung photons it radiates while still in the tracker. Muons are reconstructed in isolation as well as in jets with a very large EM component, due to a very high efficiency by using muon chamber information in combination to that from the tracker.

Most of the stable particles produced in pp collisions have low p_{T}, with the average p_{T} of a constituent particle in a 500 GeV/c jet on the order of 10 GeV/c [79]. To identify interesting and exotic particles it is necessary to accurately reconstruct and identify as many of the final state particles in and event as possible. CMS uses a combination of information from each sub-detector to build 'elements' that can be used in the particle-flow algorithm, in the form of charged particle tracks, calorimeter clusters, and muon tracks.

5.2.1 Iterative Tracking

The tracker can measure the momentum of charged hadrons to a very high accuracy, and gives a precise measurement of the the direction at the production vertex. The tracking efficiency must have nearly 100% efficiency, while keeping the tracking rate small, as nearly two-thirds of a jet's energy is carried by charged particles. In order to achieve this CMS uses an iterative tracking strategy, known as the Combinatorial Track Finder (CTF) [18]. The CTF uses multiple iterations to reconstruct tracks, identifying those that are the easiest (high p_{T}, closest to the interaction region) first and removing the hits associated with that track from the next iteration. This reduces the combinatorial complexity over each iteration.

Each iteration proceeds through these four steps [79]:

- Use seeds to provide initial track candidates. Seeds use only a few hits (2-3) to define the initial trajectory parameters.
- Extrapolate seed trajectory over expected flight path searching for additional tracker hits that can be assigned to the track candidate.
- A filter is used to provide the best estimate of the parameters of each trajectory. CMS uses a Kalman filter, which applies a small uncertainty to the location of the seed hits and fits the initial track to this estimate. Then, it looks deeper in the detector for more hits that fall with the error of this estimate.
- Tracks are rated by their quality, with certain criteria needed to pass as a track.

A total of six iterations are used, each with a different seed later of the tracker, or different p_{T} and impact parameter requirements.

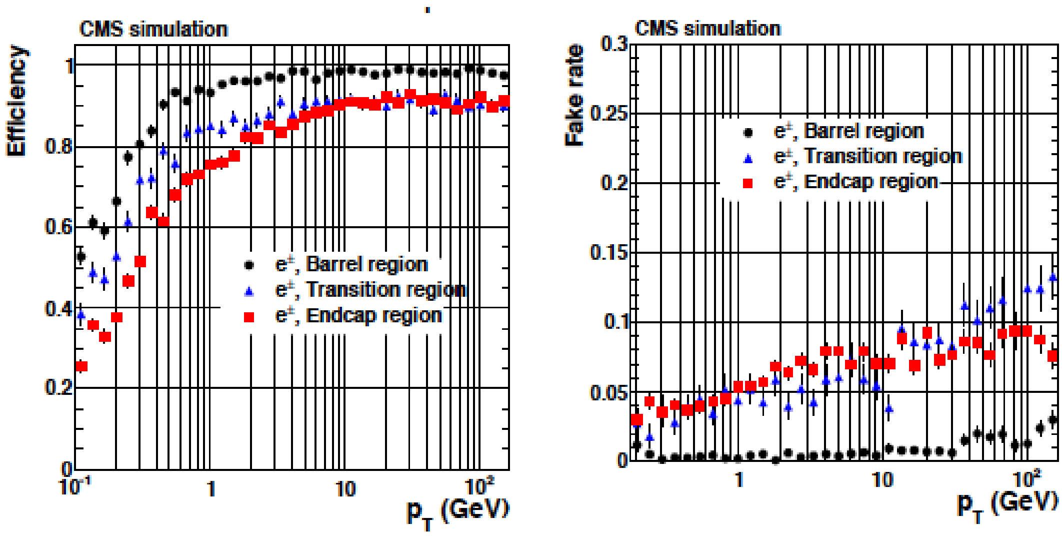
The first iterations have strict criteria in order to achieve a negligibly small fake rate. Once the hits that are associated with these tracks are removed, the seeding criteria is loosened. Loosening this criteria increases the tracking efficiency, while removing the hits associated with earlier tracks keeps the fake rate low. By the third iteration, more than 90% of jets associated with charged hadron jets are identified. For the rest of the iterations, the constraint on the track starting close to the interaction point are slowly relaxed. This allows for reconstruction of secondary charged particles created from photon conversions and nuclear interactions in the tracker volume. Figure 5.2 shows the tracking efficiency and fake rate for electron tracks as a function of η.

5.2.2 Calorimeter Clustering

Clustering in the calorimeters is the process of grouping detector cells that register hits together to measure the energy and direction of stable neutral particles. Additionally, clustering seeks to separate the neutral particles from energy deposits associated with charged hadrons, reconstruct electrons (including all associated Bremsstrahlung photons), and measure the energy of charged hadrons for which tracks were not determined accurately. The clustering algorithm is performed separately in each sub-detector: ECAL barrel and endcap, HCAL barrel and endcap, and in the preshower.

The clustering proceeds via three steps [79]:

1. Identify 'cluster seeds'. These are defined as the cell in a calorimeter with a local maximum of energy (above some set threshold).



(a) (b)

Figure 5.2: Track reconstruction efficiency (a) for electrons passing high purity requirements as a function of p_{T} for the barrel, endcap, and transition regions, and (b) the tracking fake rate for those electrons [18].

2. Expand from the seed to grow 'topological clusters'. This is done by aggregating calorimeter cells that have at least one side in common with the seed cell, and also has energy over some set threshold.

3. Repeat the process of cluster growing, now using new cells that are part of the cluster. The energy threshold limit corresponds to two standard deviations of the electronics noise in the detector ($\sim$ 80MeV in the EB and $\sim$300Mev in EE).

Each 'seed' gives rise to a 'particle-flow cluster' in the manner described above. If a cell is identified by two clusters, the energy is shared between the clusters according to the distance from the cell to the center of each cluster. The cluster energies and positions are iteratively determined as new cells are added to the cluster

5.2.3 Linking Tracks and Clusters

Each particle created in a collision is expected to give rise to multiple particle-flow elements across more than one CMS sub-detector; an example being an electron that would leave a charged particle track and several ECAL cluster energy deposits. To link these elements together an algorithm produces 'blocks' of linked elements, which serves to fully reconstruct a single particle while avoiding possible double counting from separate detectors.

The link between tracks and calorimeter clusters proceeds by extrapolating the last measured hit in the tracker to one of three detectors [79]:

i The two layers of the preshower detector, ii the ECAL, at a depth corresponding the expected maximum of the electron shower profile,

iii the HCAL, to a depth corresponding to one interaction length (typical distance for a hadron shower).

The track is linked to a cluster in these detectors if the extrapolated position is within the cluster boundaries. In an attempt to account for uncertainties on the shower position, and mechanical separations such as cracks in the detector, this position is expanded by one cell in each direction. Additionally, to link Bremsstrahlung photons to their associated electron, tangents to the track are extrapolated to the ECAL and cluster found within those boundaries is also linked.

Similarly, links between the calorimeters are formed when a cluster from the more granular calorimeter (PS or ECAL) is within the cluster envelope of the less granular calorimeter (ECAL or HCAL). The link distance is defined as the distance in $\eta - \phi$ between the two cluster positions. Finally, muon tracks are linked to charged particle tracks by a global fit between the two sets of tracks.

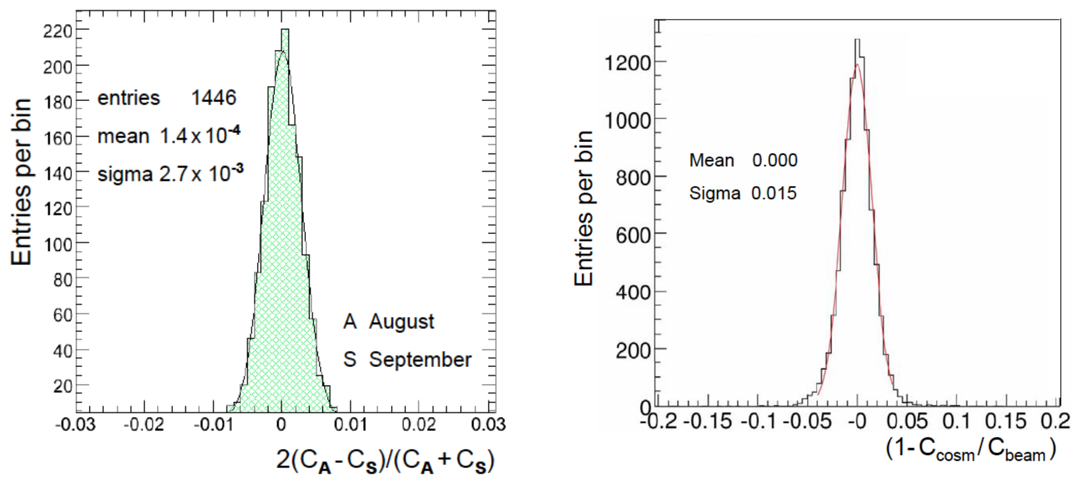
5.2.4 Cluster Calibration

A critical step in reconstructing particles is the calorimeter energy calibration, which defines the conversion of scintillation light and the subsequent photo-detector current to the energy deposited in the calorimeter by a particle. This process is done separately for the ECAL and HCAL, with calibrations of the ECAL for photon and electron reconstruction being performed before its installation in CMS.

For the ECAL, the essential issues are uniformity and stability over the entire detector, so that showers in different locations at different times are recorded accurately in relation to each other. The main source of channel to channel variation in the ECAL barrel is the variation of crystal light yield, which has an RMS $\approx 15\%$ amongst barrel crystals, though the RMS among supermodules is lower at $\approx 8\%$[14]. In the endcap, the VPT signal yields have an RMS of variation of $\approx 25\%$. Preliminary measurement in lab of crystal light yield and photodetector response reduced the variation to 5% in the EB and 10% in the EE.

Once built, each supermodule was exposed to ~ 1 week of cosmic rays. The amount of energy deposited by a muon is known to be $\sim$250MeV, allowing calibration of the crystals. Additionally, 9 supermodules were exposed to high energy electrons, with one SM exposed an additional time a month later. Figure 5.3 shows the comparison of inter-calibration coefficients, showing very good reproducibility within statistical precision.

Additionally, the ECAL performance has a strong dependence on the amount of integrated luminosity that they have been exposed to. This was shown in figure 4.10, plotting how the



(a) (b)

Figure 5.3: (a) The distribution of differences of inter-calibration coefficients for a supermodule exposed to high energy electrons on two occasions a month apart. C_A represents the coefficient measured during the August exposure, and C_S the September exposure. The reproducibility (RMS/$\sqrt{2}$) is measured to be 0.2%. (b) Distribution of inter-calibration coefficient differences for cosmic ray data (C_{cosm}) compared to high energy electron data (C_{beam})for nine SMs [14].

response of the crystals degrades over time. To correct for this, constant monitoring of the crystal response is needed in order to generate additional calibrations for the crystals. This monitoring is achieved through the use of a monitoring system that uses blue and orange LED light, as well as blue laser light, to measure the response of each crystal to a known source. The response of each crystal are averaged in rings of η (as crystals in the same η region are exposed to the same amount of radiation), and a calibration correction is calculated. Figure 5.4 shown the effects of correction from this monitoring.

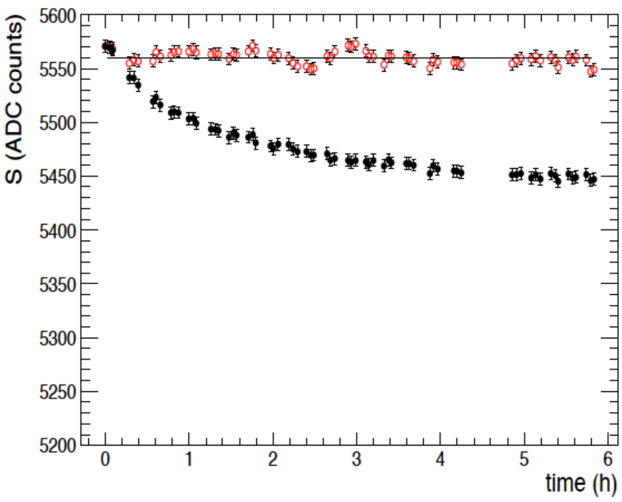


Figure 5.4: The effects of calibration for ECAL output due to laser monitoring. Black points show response before calibration to laser light, red points show response after [14].

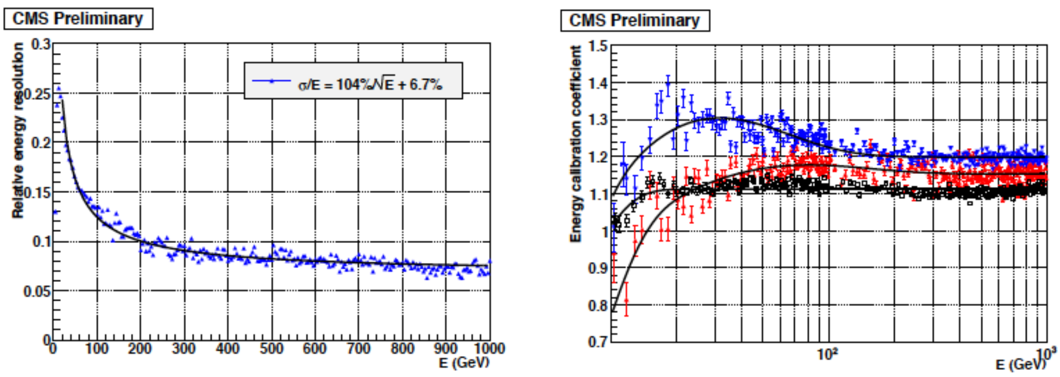
The HCAL is calibrated using 50 GeV pions that do not interact with the ECAL. In general though, hadrons deposit some energy in both the ECAL and HCAL. Even after calibrating the HCAL, substantial corrections are needed as the HCAL response to hadrons is nonlinear. It is important to note that only about 10% of the energy is affected by this, representing the contribution of neutral hadrons, but a correction procedure is still needed. To do this, we define the energy calibration in terms of contributions from the ECAL and HCAL as:

$$E_{calib} = a + b(E,\eta)E_{ECAL} + c(E,\eta)E_{HCAL} \tag{5.1}$$

where E_{ECAL} and E_{HCAL} are the energies measured in the ECAL and HCAL respectively, η is the pseudorapidity of the HCAL cluster, and E is an estimate of the true energy (the larger of the total charged particle momentum or the total calorimetric energy). For a given value of a, the values of b and c are obtained by minimizing the following χ^2 in each bin of E:

$$\chi^2 = \sum_{i=1}^{N} \frac{\left(E^i_{calib} - E^i\right)^2}{\sigma_i^2\left(E^i_{calib}\right)} \tag{5.2}$$

where E^i and σ_i are the true energy and expected calorimetric energy resolution of the i^{th} hadron. The sum extends over all events in either (a) the barrel or endcap regions of the calorimeter, or (b) solely the HCAL, solely the ECAl or in both calorimeters [79]. The coefficients of eq. 5.1 are determined via minimizing eq.5.2. Figure 5.5 shows the results of this fit using data, with the coefficient a obtained iteratively.



(a) (b)

Figure 5.5: (a) The energy resolution as a function of true hadron energy (b) Calibration coefficients as a function of energy to estimate the neutral hadron energy fraction in the HCAL [14]. Results shown are from a χ^2 minimization on simulated events.

5.3 Physics Object Reconstruction

With the tracks formed, the calorimeter clusters reconstructed, and the linking of the clusters to tracks, particles can then be reconstructed. The particle flow process begins by reconstructing muons, then electrons and photons, and finally charged and neutral hadrons. As each particle is reconstructed, the tracks and clusters associated with it are removed from the collection of blocks used to form candidate particles, which ensures that energy deposits attributed to one particle are not used a second time. The hadrons are then clustered together to form jets, and these jets can additionally be identified as coming from τ leptons or b-quarks.

5.3.1 Muons

As mentioned above, the first step in the particle flow algorithm is the identification of muon objects. To begin, and object known as a 'global muon' is identified. A 'global muon' is a muon that has tracks in the silicon pixel and strip detectors that have been matched to tracks in the muon chambers. If the combined momentum of the muon is compatible with the momentum determined by the tracker, then it is stored as a 'particle flow muon'. When the muon is removed from the candidate blocks, an estimate of the energy deposited in the associated ECAL(HCAL) clusters must also be removed, which was measured to be 0.5(3)±100%GeV in a cosmic ray study.

Muon resolution using the combined information from the muon chambers and the trackers was already shown in figure 4.19. Muon ID at CMS is very efficient, such that a dimuon spectrum can be measured with great accuracy, as shown in figure 5.6.

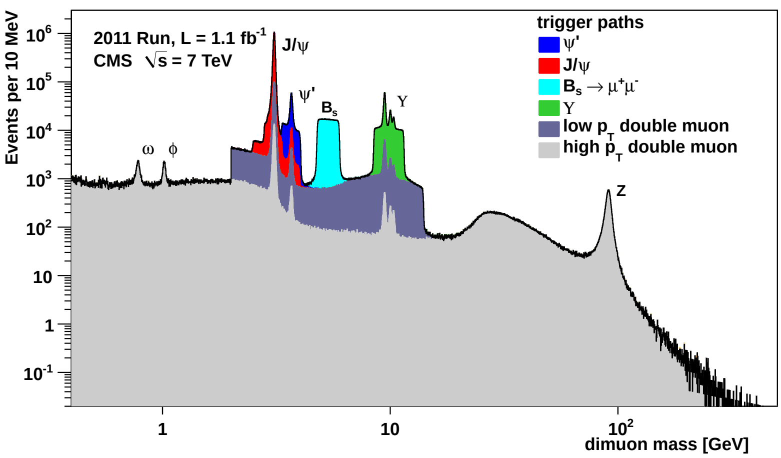


Figure 5.6: Superposition of various dimuon trigger paths on $1.1 fb^{-1}$ of data taken in early 2011 [19].

5.3.2 Electrons

Electron reconstruction follows muons as the second step in the particle-flow reconstruction process. Electrons tend to give rise to short tracks, and loose energy in the tracker layers through Bremsstrahlung radiation on their way to the calorimeter [79]. Each track is submitted to a pre-identification state which uses the tracker as a pre-shower to help identify possible electron tracks. These pre-identified tracks are then re-fit with a Gaussian-Sum Filter (GSF) in an attempt to follow their trajectories into the ECAL [80]. The GSF algorithm is used here because of the the default track recognition used employs a Kalman filter as described above. A Kalman filter approximates the energy loss of a particle using a single Gaussian method, while the GSF method approximates energy loss using a mixture of Gaussians. It has been shown by Bethe and Heitler that the energy loss of electrons is best described by a mixture of Gaussians [81].

Using the GSF method the change in direction of the electron due to Bremsstrahlung radiation is taken into account. This allows for the linking of ECAL clusters related to Bremsstrahlung photons by extrapolating tangents to these changes in direction and identifying ECAL clusters not associated with with any other track. The final step is to combine several observables built from measurement in the tracker and ECAL into a multivariate identifier for electrons. These include measurements such as the energy of the seed cluster, the momentum of the GSF track, as well as the shower width and the fraction of energy measured by the HCAL [20]. The resultant MVA estimator is used to identify electron candidates as particle-flow electrons. Figure 5.7 shows the output of this MVA in simulation compared to data taken in commissioning the detector in 2009.

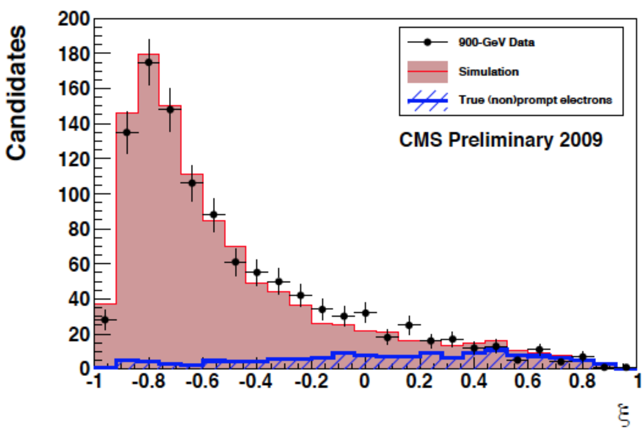


Figure 5.7: Distribution of the output of the multivariate electron estimator used to define GSF electrons. To be identified as an electron by the MVA, $\xi > -0.1$. Data is shown as dots while MC simulation is shown as the solid histogram [20].

5.3.3 Charged Hadron

After electron identification, the next step is to identify charged hadrons. Before this is done, tighter quality cuts are applied to the remaining tracks requiring that the relative uncertainty of the measure p_{T} is smaller than the energy resolution in the calorimeters expected for charged hadrons [79]. Only about 0.2% of jets are rejected by this procedure, but even then the energy is not lost as it is measured with higher precision in the ECAL. Each track can be linked to a number of ECAL and HCAL energy clusters. Comparing the total calibrated energy associated with the track, a 'particle flow charged hadron' is identified if the energy agrees with that measured in the tracker. In this case the charged hadron momenta are redefined by a fit of the measurements in the tracker and calorimeters.

5.3.4 Photons and Neutral Hadrons

The next step is the identification of clusters in the ECAL and HCAL that are linked to tracks which have a significantly larger energy than the total associated with the charged particles that have been identified. If the energy excess is large than the expected calorimeter energy resolution, a 'particle flow photon' is identified and sometimes also a 'particle flow neutral hadron'. If the energy excess is larger than the total energy excess in the ECAL, a particle flow photon is created with energy equal to that found in the ECAL, and a neutral hadron is created with the remaining energy (deposited in the HCAL). In the case that the ECAl excess is greater than the HCAL excess, only a particle flow photon is formed. This process gives precedence in the ECAL to photons, because in jets nearly 25% of the energy is carried by photons, while neutral hadrons only deposit $\sim 3\%$ of their energy in the ECAL [79].

5.3.5 Jets

Once all of the calorimeter blocks and tracks have been formed into particle-flow objects, jets can be formed by clustering groups of the charged hadrons, neutral hadrons, and photons together. The energy fraction in jets is divided amongst charged particles, photons, and neutral hadrons with a breakdown of roughly 65%, 25%, and 10% for the respective constituents. As the energy calibrations in the calorimeter only affect the 10% of energy from neutral hadrons, we expect jets formed by clustering reconstructed particles to be much more accurate than jets reconstructed with solely calorimeter information (Calo-jets). To form the particle-flow jets an iterative algorithm called the anti-k_T algorithm [21] is used.

Jet clustering algorithms works by defining a distance parameter between between two candidate particles i and j, d_{ij}, and the distance between the entity and the beam, d_{iB}. These are defined as:

$$\begin{aligned} d_{ij} &= min\left(k_{ti}^{2p}, k_{tj}^{2p}\right)\frac{\Delta_{ij}^2}{R^2} \\ d_{iB} &= k_{ti}^{2p} \end{aligned} \tag{5.3}$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, and k_{ti}, y_i , and ϕ_i are respectively the transverse momentum, rapidity and azimuth of particle i. R is a user defined radius parameter, and p governs the relative power of energy vs geometric scales. For the anti-k_T algorithm, p = -1, and eq 5.3 reduces to:

$$d_{ij} = \min(\frac{1}{p_{iT}^2}, \frac{1}{p_{jT}^2})\frac{\Delta_{ij}^2}{\mathrm{R}^2}. \tag{5.4}$$

The algorithm loops over all particle-flow candidate objects, calculating d_{ij} for each pair of objects. Once it does this, it selects the two objects with the lowest value of d_{ij} and combines them. This process is repeated until the smallest value of $d_{ij} > d_{iB}$ for all the remaining pairs. This cutoff limit of $1/p_T^2$ defines a maximum size that the algorithm will look to cluster particles inside. The construction of d_{ij} using the inverse p_{T} squared has the result of producing values of d_{ij} that are smaller for objects with a higher p_{T}, given equal separation. The result of this is that softer p_{T} particles will tend to cluster to higher p_{T} particle long before they would cluster amongst themselves. If no hard particles are present, the jet object will simply cluster soft p_{T} particles in a circle in ($\eta - \phi$ space) of radius R.

The clustering of the anti-k_T algorithm leads to jets with a large p_{T} being reconstructed as perfect circles, while softer p_{T} jets can have a more ambiguous shape. Figure 5.8 shows a display of the clustering from the anti-k_T algorithm. In this figure, notice that the green jet at y = 2 and ϕ=5 has a circular shape, while it deforms the smaller jet next to it, making it a crescent moon shape.

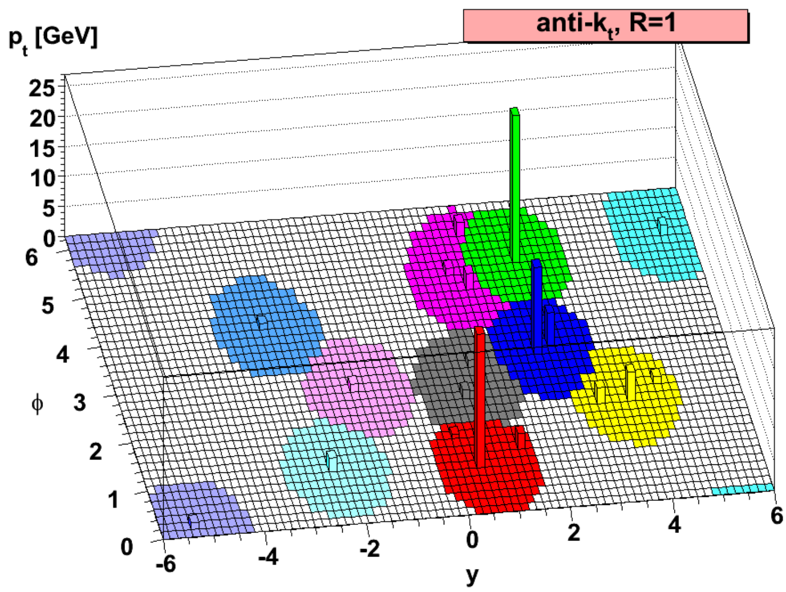


Figure 5.8: The anti-kt jet clustering algorithm with distance parameter R=1.0 [21]

B-tagged Jets

One jets are reconstructed, there is still more information that can be gathered. For instance, jets that originate from b-quarks can be distinguished from other jets because of unique properties of the b-quark. Due to the fact that the b-quark is much heavier than it's lighter relatives (∼3 times heavier than the charm quark and ∼40 times heavier than the strange quark), b-quarks will have a larger transverse momentum then the light-flavor quarks. Also, b-quarks have a longer lifetime than its lighter relatives, which means that when they are created in a collision, they tend to travel a small, but observable, distance in the detector before they decay. This results in a new vertex being formed some distance away from the primary collision vertex, dubbed a 'secondary vertex'. Figure 5.9 shows a cartoon of a b-quark jet that has traveled a distance L_{xy} from the primary vertex.

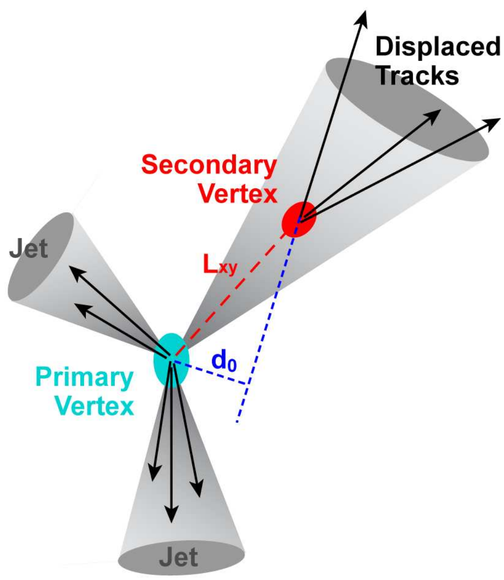


Figure 5.9: A b-quark will travel a distance L_{xy} before decaying and creating a secondary vertex. L_{xy} is measured in the plane orthogonal to the beam direction, and the impact parameter, d_0, measures the displacement from the beam line [22].

Additionally, because the b-quarks belong to the third generation of quarks they are much more likely to have a lepton in the decay products. This lepton will not be originating from the primary vertex, and in the case of muons this is very easy for CMS to track. To identify the b-quarks, many different kinematic variables related to the jet are combined in a multivariate discriminator called the Combined Secondary Vertex (CSV) algorithm [82]. Using the result of this algorithm different cut values are set on the CSV discriminant to define b-jet tagging, depending on fake rate desired.

Tau Reconstruction

Tau leptons are the 3rd generation leptons, a much larger unstable version of electrons and muons. Tau's decay via the weak interaction, generating a tau neutrino (ν_τ) and a W boson. If the W boson decays hadronically (to quarks), the tau lepton can be reconstructed by analyzing the resulting jets. Tau jets are characterized by the number of charged hadrons that make them up, which must be either one ($\sim$75%) or three ($\sim$25%) hardrons to conserve charge. To determine if a jet is from a tau, the particle flow jets are clustered a second time, using a smaller distance parameter. Tau jets are very tightly collimated, so this second distance parameter is used to determine if a jet is from a hadronically decaying tau.

5.3.6 Missing Transverse Energy ($\not\!\!E_{\mathrm{T}}$)

In order to identify Missing Transverse Energy ($\not\!\!E_{\mathrm{T}}$) CMS makes use of its hermetic design which ensures that nearly all of the particles produced in a collision pass through, and are reconstructed by, the detector. This hermeticity allows for the measurement of a momentum imbalance in an event, which can be calculated after measuring all of the constituent particles. The pp beam collides head on, so we know there is no inherent momentum transverse to the beamline in collision, meaning the p_{T} of all of the particles must balance out. Thus, we define MET ($\not\!\!E_T$) to be

$$\not\!\!E_{\mathrm{T}} = | - \sum_{i=1}^{nPF} \vec{p_{Ti}}|, \tag{5.5}$$

where nPF is the number of particle-flow candidates in the events, and $\vec{p_{Ti}}$ is the vector sum of their transverse momentum.

In the Standard Model, only long lived weakly interacting neutral particles will pass through the CMS detector without being measured. This only occurs with neutrinos in the SM, but many Beyond the Standard Model theories (such as SUSY), also predict stable neutral particles that would be observable through $\not\!\!E_T$. To calibrate the particle flow $\not\!\!E_T$ algorithm, events that produce many jets but have no intrinsic $\not\!\!E_T$ were used as a baseline. Additionally, any mis-measurement in the calorimeters due to detector noise can lead to spuriously high $\not\!\!E_T$ value, which was corrected for in calibration runs with 900GeV data from 2009 [20]. Figure 5.10 shows the resultant $\not\!\!E_T$ distribution before and after this calibration.

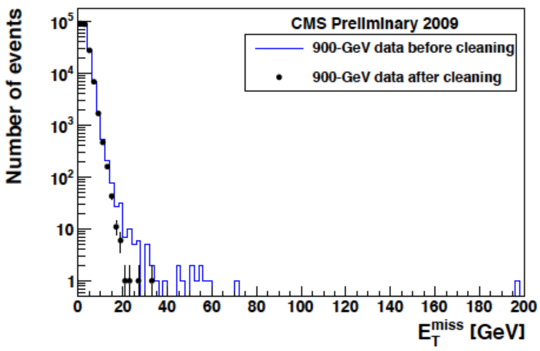


Figure 5.10: $\not\!\!E_T$ distribution in 900 GeV data before (blue histogram) and after (black dots) cleaning [20].

Chapter 6

$H \to WW \to l\nu qq$ Analysis Part 1: Selection

In order to perform our search for the Higgs boson we must start by identifying events that are consistent with the final state of $l\nu$ jj that we are searching for. The semi-leptonic decay of a Higgs via two W bosons leaves us with a specific set of criteria that we can search for. Not only are there four final state particles (one lepton l, the neutrino ν which we observe through $\not\!\!E_{\mathrm{T}}$, and two jets j), these particles have specific characteristics we can look for. For the leptonic W, we can expect a relatively high $\not\!\!E_{\mathrm{T}}$ value as the neutrino should carry a lot of energy, as well as an isolated lepton because the jets we expect in our event will be coming from the hadronic W which has recoiled in the opposite direction. For the hadronic W, we expect two jets with high p_T that are also well isolated.

Since we are searching for a low mass Higgs (∼125 GeV), at least one of the W bosons that it decays to must have an off-shell mass. This means that at least one of the W bosons will be a virtual W boson (usually denoted by W^* to imply it has mass much different from the usual 80 GeV). This virtual W changes the kinematic distributions of our final state particles, for instance limiting the usefulness of a cut on the reconstructed W mass in selecting our events.

For our analysis, the background process that provided the majority of events in our selection is that of SM W+jets production. As described earlier, this process is when a W boson is generated in addition to one or more gluons which will hadronize and produce jets in our selection. In addition to W+jets, the processes that contribute to our backgrounds are Z+jets, diboson (WW, WZ, ZZ), $t\bar{t}$, and single top processes as described in section 2.10. All of these processes were simulated via MC generators. In addition, the QCD multi-jet process was modeled using a data-driven technique and will be described below in section 6.2.

The following section will describe an analysis utilizing the full 2012 data run (19fb^{-1} for both muon and electron samples) collected by the CMS detector at a center of mass energy of 8 TeV. I will describe the data and background samples used (both Monte-Carlo (MC) simulated and data-driven) as well as the event selection and all of the correction factors applied to background samples in order to better model the events we see in data.

6.1 Data and MC Samples

Data is collected by the CMS detector for pp collisions via an HLT trigger path and stored offline for analysis. Background samples are simulated via a Monte Carlo simulation technique which I will briefly describe here. Event simulation proceeds through a number of stages in order to properly model not only the physics of the event, but how the event will be observed in our detector:

- **Stage 1**: Calculate the probability that some set of initial state particles with certain momenta will create a final state of particles with certain momenta. This involves calculating the scattering amplitude (to some order in perturbation theory) using the Feynman rules derived from the Lagrangian. The scattering amplitude is a multi-dimensional probability function, which depends on the initial and final state momenta of the particles in the process.

- **Stage 2**: Calculate the decays of any final state particles produced in stage 1, what is known as a parton shower.

- **Stage 3**: Simulate the response of the CMS detector when an interaction described in stages 1 and 2 occurs at the interaction point.

The scattering amplitude introduced above is often referred to as a Matrix Element (ME). This name arises from the fact that the initial state particles in an interaction are described in a vector, and mathematical transformation one vector into another involves a matrix. This particular matrix describes the probability of creating the final state particles we are simulating. In order to calculate this we need to know the momenta of the incoming particles, which is defined by the beam energy at the LHC. When colliding protons, however, it is really the quarks and gluons inside the protons that do the interacting. The distribution of energy inside a proton is divided amongst the valence quarks (the two up and one down quark that make up its structure), gluons, and sea quarks. This is described by a Parton Distribution Function (PDF), which comes directly from experimental measurements of the proton structure.

The generator algorithm calculates the ME from a given Lagrangian, and then uses the PDF to assign momentum values to the constituent partons given the initial momentum that we have defined. These momentum values are assigned by randomly sampling the PDF of each parton, and it is from this sampling process known as Monte Carlo that we take the name to describe our simulated samples. From here, the generators use the ME to calculate the final state partons for each interaction.

These final state particles are often quarks and gluons, which we will not observe directly in our detector. That is because these partons undergo a process called hadronization where each of the colored partons are transformed into color singlet hadrons, a process that creates many more particles and was mentioned above in Stage 2 as the parton shower. All of the particles created in hadronization will have a component of momentum in the direction of the initial particle, and can be grouped together in what is knows as a hadron jet.

Lastly, once we have the complete picture of an event through hadronization, the response by the CMS detector to such an event must be simulated. The software that does this is called Geant4 [83], which models every element of the detector including readout electronics and support structures. In addition, Geant4 describes the energy deposition of the particles as it travels through the detector and the digitization and readout of the signals we would receive from such energy deposition. After this the signals must be reconstructed into physics objects, a process which proceeds as described in chapter 5.

6.1.1 Data Samples

The results presented here are based on the full 2012 CMS dataset, which corresponds to ~ 19 fb^{-1} of 8 TeV data. Table 6.1 lists the datasets used for this analysis, which are based on High Level Trigger (HLT)s used to select events with single muons or single electrons. Luminosities are quoted from a calculation on minimum bias events with the HF detector and are reported with a 2.6% uncertainty [84].

6.1.2 Signal Samples

The $H \rightarrow WW \rightarrow l\nu jj$ signal is modeled using Pythia6 [85] Monte Carlo Generator. These events were generated in the "Summer12" MC regime. The samples, NLO cross sections, and decay modes are listed in Table 6.2 along with the branching ratio (BR) to their final state. All samples here are generated with $M_h = 125$ GeV/c^2.

Dataset	Run Range	Integrated Luminosity
/SingleMu/Run2012A-13Jul2012-v1/AOD	190645-196531	0.809 fb^{-1}
/SingleMu/Run2012A-recover-06Aug2012-v1/AOD	190782-190949	0.082 fb^{-1}
/SingleMu/Run2012B-13Jul2012-v1/AOD	193834-196531	4.383 fb^{-1}
/SingleMu/Run2012C-24Aug2012-v1/AOD	198022-198523	0.489 fb^{-1}
/SingleMu/Run2012C-PromptReco-v2/AOD	194631-203002	6.285 fb^{-1}
/SingleMu/Run2012D-PromptReco-v1/AOD	194480-208686	7.231 fb^{-1}
Total SingleMu	**190645–208686**	**19.279 fb^{-1}**
/SingleElectron/Run2012A-13Jul2012-v1/AOD	190645-196531	0.809 fb^{-1}
/SingleElectron/Run2012A-recover-06Aug2012-v1/AOD	190782-190949	0.082 fb^{-1}
/SingleElectron/Run2012B-13Jul2012-v1/AOD	193834-196531	4.336 fb^{-1}
/SingleElectron/Run2012C-24Aug2012-v1/AOD	198022-198523	0.489 fb^{-1}
/SingleElectron/Run2012C-PromptReco-v2/AOD	194631-203002	6.194 fb^{-1}
/SingleElectron/Run2012D-PromptReco-v1/AOD	194480-208686	7.238 fb^{-1}
Total SingleElectron	**190645–208686**	**19.148 fb^{-1}**

Table 6.1: The datasets analyzed for this analysis.

Signal Higgs Production			
Production Mechanism	Dataset	Cross Sect.	BR
gluon-gluon fusion	ggH; M_H=125, Decays via H→WW→lνjj	19.27pb	0.0946
vector-boson fusion (VBF)	qqH; M_H=125, Decays via H→WW→lνjj	1.578pb	0.0946
Ass. Prod. Higgs	WH, ZH, and TTH; M_H=125, Decays via H→ WW, inclusive	1.249pb	0.215
Non signal Higgs Production			
Ass. Prod. Higgs	WH, ZH, and TTH; M_H=125, Decays via H→ ZZ, inclusive	1.249pb	0.0264
Ass. Prod. Higgs	WH; M_H=125, Decays via H→bb, W→lν	0.7046pb	0.1879
Ass. Prod. Higgs	TTH; M_H=125, Decays via H→bb	0.1293pb	0.577

Table 6.2: List of Signal datasets and cross sections

6.1.3 Background Samples

The background samples were modeled using MC generated events utilizing Madgraph [86] as a tree level matrix element generator, or Powheg [87] for NLO ME generation matched to Pythia [85], in order to simulate the hard and soft hadron interactions as well as the parton shower, fragmentation, and decay. They are all generated as part of the DR53X set of samples, using the S10 pileup scenario. The samples, their parent datasets, and NLO cross section σ are listed in Table 6.3.

Sample	Dataset	Generator	Cross Sect.
W+jets	W + inclusive jets W→lν	Madgraph	37509 pb
t$\bar{t}$	t$\bar{t}$ + jets	Madgraph	225.197 pb
Z/γ^*+jets	Z/γ^* →ll, M_{ll} >50	Madgraph	3387.6 pb
WW	WW	Pythia	54.838 pb
WZ	WZ	Pythia	33.21 pb
ZZ	ZZ	Pythia	17.654 pb
Single t			
t-Channel	t, t-channel production	Powheg	56.4 pb
tW-Channel	t, tW-channel production	Powheg	11.1 pb
s-Channel	t, s-channel production	Powheg	3.79 pb
Single $\bar{t}$			
t-Channel	$\bar{t}$, t-channel production	Powheg	30.7 pb
tW-Channel	$\bar{t}$, tW-channel production	Powheg	11.1 pb
s-Channel	$\bar{t}$, s-channel production	Powheg	1.76 pb
QCD			
Electron	See table 6.1 for a list of SingleElectron datasets	N/A	N/A
Muon	See table 6.1 for a list of SingleMu datasets	N/A	N/A

Table 6.3: List of background MC datasets and cross sections used for normalization.

6.1.4 MC Pileup Reweighting

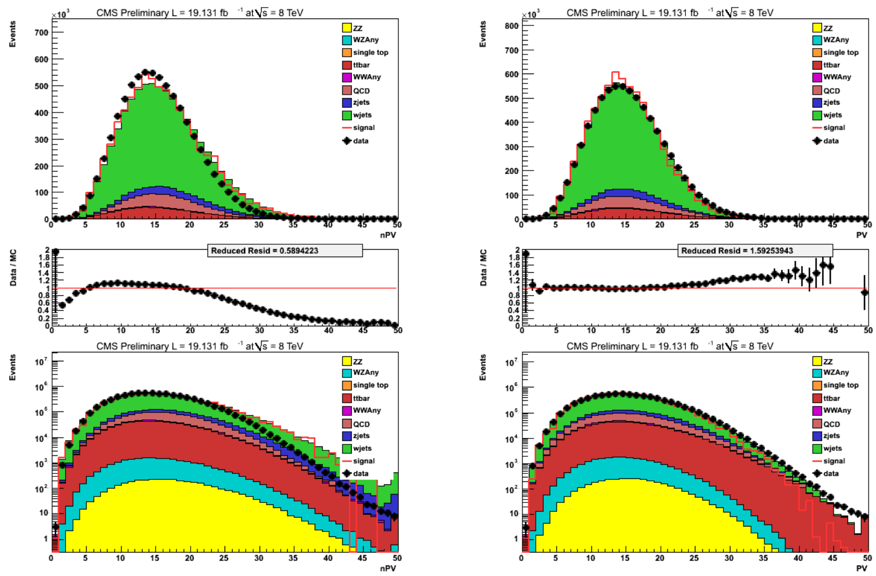
During 2012 the instantaneous luminosity delivered to CMS by the LHC increased, resulting in a large number of interactions being reconstructed in the same time window. These overlapping events, known as 'in time' pileup for interactions occurring in the same bunch crossing, makes the reconstruction and isolation of physics objects difficult. Additionally, there can be 'out of time' pileup caused by interactions in bunch crossings to either side of the primary interaction point. It is important that our MC simulations match the pileup distributions seen in the data, so in order to do this minimum bias events are added to all MC generated events. When MC events are generated, the true pileup distribution that will be seen in data is unknown, so they are generated with a large number of pileup interactions and are then reweighted to match the data later. The number of pileup events is a function of instantaneous luminosity $\mathcal{L}_{inst.}$ and total inelastic cross section $\sigma_{inelastic}$. We used a value of $\sigma_{inelastic} = 69.3$ mb following the CMS approved value [88]. In order to assess the effect of a systematic uncertainty due to choice of $\sigma_{inelastic}$, a $\pm 7\%$ variation was used.

In order to calculate pileup weights, we must know the distributions for the number of interactions in the data and in the MC samples. For data, this distribution is estimated by using a tool provided by the CMS collaboration called pileupCalc [88], that uses information about the data runs to generated a distribution of the average number of interactions in the data. For input, you provide a JSON file, which is a file listing which run numbers you want to include. We used the full 2012 'golden' JSON file for data, Cert_190456-208686_8TeV_PromptReco_Collisions12_JSON.txt, which includes all good data runs in 2012. For MC we used the Summer12 s10 MC distribution as all of our MC samples were generated using this regime.

The pileup weights are applied to all MC generated events, and we checked results by comparing the number of primary vertices distribution in the data and MC. Figure 6.1 shows how the MC looks before and after the pileup weights are applied for our combined electron + muon analysis.

6.2 Multijet-QCD Background

In order to model the QCD multi-jet background we decided to use a data-driven technique. MC based QCD samples do exist, but the QCD process is very difficult to model as it involves many orders of QCD perturbations to describe fully. While MC calculation techniques have improved in this area, our event selection also provides another difficulty. QCD processes involve many jets, and do not have a true lepton in them. Thus, by selecting on an isolated jet, we vastly reduce the number of MC events that pass our criteria, and the sample we are left with is very



(a) (b)

Figure 6.1: Comparison of the number of primary vertices (nPV) in data to MC generated samples for the 2012 full dataset of 19fb^{-1} before pileup reweighting (a), and after weights are applied (b).

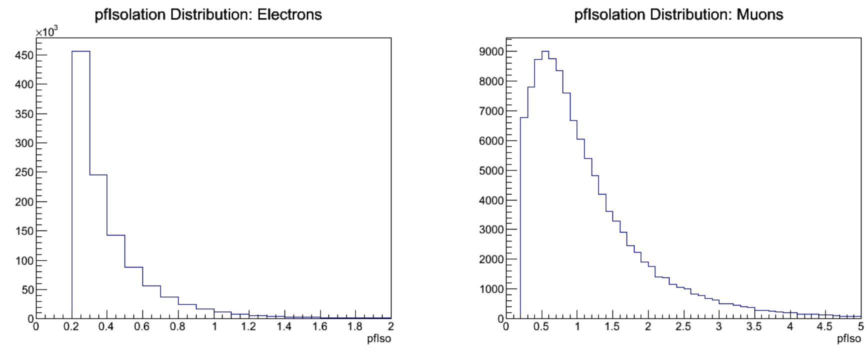
statistically limited. Together, these issues led us to pursue a data-driven sample.

The goal when selecting this sample was to extract events from the data that can be treated as true QCD multi-jet events, but are in a sample completely orthogonal to the data sample on which we will perform a signal extraction. QCD events are characterized by having many jets. When one or more of these jets deposits its energy in the electromagnetic calorimiter it can be mis-reconstructed as a lepton, and thus pass our selection. Normally we require that the lepton be very isolated to protect from this jet faking a lepton, but as we expect this from a QCD event we invert our lepton isolation. This isolation inversion also ensures orthogonality to our signal sample.

Full electron selection criteria are defined in section 6.3; here we will just address isolation cuts. All leptons in our events must pass different levels of isolation cuts, the loosest being pfIsolation < 0.2 (PF stands for particle flow, the process described in chapter 5). For QCD multi-jet event selection we require that the pfIsolation > 0.2. Both the electron and muon QCD samples were generated by running on the entire 2012 dataset. Failed jobs during processing let to a slightly smaller sample of data runs that were processed, but many more events were selected in this way than were available via MC generation. The samples are later scaled to account for the mismatch in integrated luminosity between the isolated and anti-isolated data samples.

For our electron QCD sample, in addition to the pfIsolation cut, we must turn off the electron MVA identification requirements as these help define a very strict definition for an electron. As we are looking for fake electrons, relaxing this requirement is crucial. These events must also pass an electron trigger, 'HLT_Ele27_WP80_v*', which is identical to the trigger used for signal sample. For the muon QCD sample, the trigger is changed to 'HLT_Mu24_eta2p1_v*' as the trigger used in our signal sample selection has an isolation requirement on it. There is no mvaID requirement for muons so inverting the pfIsolation requirement to > 0.2 is the only change in muon identification.

In order to give good separation from the data sample we use for signal extraction, we increased the minimum pfIsolation and also put an upper limit on the pfIsolation values in order to keep our sample from artificially skewing to higher nPV values. For electrons, this meant requiring pfIso $\{0.3,0.7\}$, and for the muon sample required pfIso $\{0.3,2\}$. In addition, we use the 'sideband' selection of the QCD (events with $0.2 <$pfIsolation< 0.3 for the low sideband, and events with pfIsolation > 0.7 (> 0.2) for the high sideband for electrons (muons)) sample in order to define our systematic uncertainty on the QCD selection. Figure 6.2 shows the electron and muon sample pfIsolation values, our selection chooses the majority of the sample while leaving enough statistics on either sideband for our systematic uncertainty calculations.



(a) (b)

Figure 6.2: Particle Flow Isolation values for the electron D-D QCD multi-jet sample (a) and the muon D-D sample (b).

6.3 Event Selection

In this section, we will define the physics object and preselection requirements used in this analysis. All events must pass one of the triggers described in section 3.1, and have at least one good primary vertex. Leptons are categorized as either 'tight' or 'loose' by selection criteria shown below. Events must also have exactly one tight lepton, and have two or more jets in them. Events with an additional 'loose' lepton are rejected.

6.3.1 Trigger

For this analysis we require all the events to pass one of the single lepton triggers shown in Table 6.4. Muon+jet events must pass the SingleMu trigger, and electron+jet events must pass the SingleEle trigger. Our electron trigger requires a minimum p_T of 27 GeV and uses a working point 80 (WP80) selection. This selection means that 80% of the electrons passing this trigger were shown to be true electrons in a MC calibration sample. The muon trigger requires that our muons be isolated and have p_T of 24 GeV or higher.

As stated in section 6.2, the data-driven electron QCD sample uses the same trigger as the electron+jet events. However, the trigger for the muon QCD sample is changed to a similar, but non-isolated trigger.

Dataset	Trigger Name
SingleMu	HLT_IsoMu24_eta2p1_v*
SingleEle	HLT_Ele27_WP80_v*
QCD_Muon	HLT_Mu24_eta2p1_v*
QCD_Electron	HLT_Ele27_WP80_v*

Table 6.4: List of triggers

6.3.2 Vertex Selection

Every event is required to have at least one good primary vertex (PV). In addition, this primary vertex must pass these requirements:

- The number of degrees of freedom used to find the PV must be larger than 4,
- The absolute value of the z-coordinate($|dZ|$) of the PV must be smaller than 24 cm,
- The absolute value of the ρ-coordinate of the PV must be smaller than 2 cm,
- The PV must not be identified as fake.

These are summarized in table 6.5.

Cut	Value
Degrees of Freedom	≥ 4
$\|dZ\|$	≤ 24 cm
$\|\rho\|$	≤ 2.0 cm

Table 6.5: The requirements for a primary vertex

6.3.3 Electron Selection

Electrons are selected using the particle flow algorithm of reconstructing the electron object. Electrons are classified as 'tight', 'loose', or passing neither set of cuts. The classification cuts are summarized in table 6.6. The cut parameters are defined as:

- p_T - the component of the momentum transverse to the beam-line.

- ID - electron ID is determined via a multivariate analysis (MVA) technique, which provides a discriminant value to separate fake from real electrons, and is trained with events that are required to pass a HLT trigger (mvaTrigV0), or not (mvaNonTrigV0). The value of MVA required to define a 'tight' or 'loose' electron is dependent on where the electron is found in the detector. "SC η" refers to the η location of the supercluster of ECAL crystals that the electron is found in.

- Conversion Veto - "passConversionVeto" ID ensures that the electron has not been reconstructed from a photon which has converted to an electron positron pair

- $|\eta|$ - the absolute value of the pseudorapidity of the electron

- $|d_0(\mathrm{PV})|$ - the absolute value of the transverse distance of the extrapolated electron track to the primary vertex, as calculated from the beam spot (BS).

- $|d_Z(\mathrm{PV})|$ - the absolute value of the longitudinal distance of the extrapolated electron track to the primary vertex position.

Electron identification uses the multivariate technique with a triggering MVA [89] for electrons with $p_\mathrm{T} \geq 20$ GeV/c. Tight electrons must pass the electron MVA cuts, as well as have a minimum p_T of 27 GeV/c and $|\eta| < 2.5$. In addition, they must have $|d_0(\mathrm{PV})| < 0.02$ cm, $|d_Z(\mathrm{PV})| < 1$ cm, and pass the conversion veto. Loose electrons must pass looser electron MVA cuts, as well as have a minimum p_T of 15 GeV/c and $|\eta| < 2.5$. In addition, they must have $|d_0(\mathrm{PV})| < 0.04$ cm, $|d_Z(\mathrm{PV})| < 2$ cm, and also pass the conversion veto. In addition electrons with $\eta > 1.4442$ and $\eta < 1.566$ are excluded as this region is the gap between the barrel and endcap section of the ECAL.

Cut	Tight	Loose
p_T	> 27 GeV/c	> 15 GeV/c
ID Cuts		
SC $\|\eta\| < 0.8$	MVA ID("mvaTrigV0") $>$ 0.977 & pfIsolation $<$0.093	MVA ID("mvaNonTrigV0") $>$ 0.877 & pfIsolation $<$0.426
SC $\|\eta\| > 0.8$ & SC $\|\eta\| < 1.479$	MVA ID("mvaTrigV0") $>$ 0.956 & pfIsolation $<$0.095	MVA ID("mvaNonTrigV0") $>$ 0.811 & pfIsolation $<$0.481
SC $\|\eta >\|$ 1.479 & SC $\|\eta\| < 2.5$	MVA ID("mvaTrigV0") $>$ 0.966 & pfIsolation $<$0.171	MVA ID("mvaNonTrigV0") $>$ 0.707 & pfIsolation $<$0.390
$\|\eta\|$	< 2.5	< 2.5
$\|d0(PV)\|$	< 0.02 cm	< 0.04 cm
$\|dZ(PV)\|$	< 1 cm	2 cm
ID	passConversionVeto	passConversionVeto

Table 6.6: Tight and loose electron definitions

6.3.4 Muon Selection

Muons are selected using the particle flow algorithm of reconstructing the muon object. Muons are categorized as 'tight' and 'loose' based on a cuts based identification. Variable definitions are identical to those defined in section 6.3.3, with selection cuts shown in table 6.7. Additional definitions are described here:

- pfIsolation - this is a ratio of the energy deposits remaining in the calorimeter to that found in the tracker after the contribution from the muon has been removed, in a cone size ΔR = 0.3 around the muon track.
- Tracker / Global / PF Muon - This refers to whether the muon was reconstructed with a χ^2 fit to the tracks from the tracker only (tracker muon), the tracker and the muon chambers (global muon), or if the particle was reconstructed from the particle-flow algorithm (PFmuon).
- N_{layers}(tracker) - the number of layers in the tracker with hits used in the muon track reconstruction.
- X^2 of track fit - the reduced χ^2 of the track fit(raw χ^2/Number of Degrees of Freedom in the fit).
- N_{layers}(pixel) - the number of layers in the inner pixel detector containing hits used in the muon track reconstruction.
- $N_{segments}(\mu)$ - the number of segments in the muon chambers used to reconstruct the muon tracks.

Tight muons must have a $p_{\mathrm{T}} > 24$ GeV/c and $|\eta| < 2.1$. They must also be reconstructed as a global muon and a PF muon, as well as having a pfIsolation < 0.12. Tight muons must have a minimum of 6 hits in the tracker, as well as at least one muon chamber and one pixel hit. These muons must also have $dZ < 0.2$ cm and $d0 < 0.5$ cm. Loose muons must have $p_{\mathrm{T}} > 10$ GeV/c, pfIsolation < 0.2, and $|\eta| < 2.5$. It must be reconstructed as a PF muon, but can be either a global or tracker muon.

6.3.5 Jet Selection

Jets are reconstructed using the anti-k$_T$ algorithm with a cone size of $\Delta R = 0.5$. We use particle flow jets (PF) with charged hadron subtraction (chs). Jet energy corrections (JEC) are applied to both MC and data at the initial n-Tupling stage. Additional Jet energy resolution (JER) corrections are applied later based on the jet η as recommended by CMS [90]. These corrections

Cut	Tight μ	Loose μ
p_T	> 24 GeV/c	> 10 GeV/c
pfIsolation	< 0.12	< 0.2
$\lvert\eta\rvert$	< 2.1	< 2.5
ID	Global Muon	Global Muon or Tracker Muon
ID	PFMuon	PFmuon
N_{layers}(tracker)	> 5	
X^2 of track fit	< 10	
N_{layers}(pixel)	> 0	
$N_{segments}(\mu)$	> 1	
N_μ Hits	> 0	
$\lvert d0(\mathrm{PV})\rvert$	< 0.2 cm	
$\lvert dZ(\mathrm{PV})\rvert$	< 0.5 cm	

Table 6.7: Tight and loose muon definitions

are described in detail in section 6.4.1. Jets must pass the cuts described in Table 6.8. and defined here:

- p_T - component of the momentum transverse to the beam-line
- η - the pseudorapidity of the reconstructed jet
- CEF - Charged Electromagnetic Fraction: the ratio of energy measured from charged particles in the jet to the total number of particles in the jet
- NHF - Neutral Hadron Fraction: the ratio of energy measured from neutral particles to the total number of particles in the jet
- NEF - Neutral Electromagnetic Fraction: the ratio of energy measured from neutral particles in the ECAL (photons) to the total number of particles in the jet
- CHF - Charged Hadron Fraction: the ratio of energy from charged hadrons to the total number of particles in the jet
- NCH - Number of Charged Hadrons: raw charged hadron multiplicity
- N$_{constituents}$ - Number of constituents, which can be charged and neutral hadrons, as well as non-prompt photons and leptons.

All energy fraction cuts are performed on the raw jets (before energy corrections are applied). Jets must have $p_{\mathrm{T}} > 25$ GeV/c and $|\eta| < 2.4$ in addition to passing the energy fraction requirements in Table 6.8. Jets are 'cleaned' by rejecting jets that fall within a cone size of $\Delta R < 0.3$ from a lepton.

6.3.6 Analysis Cuts

In addition to the physics object ID cuts that were described in section 6.3, some additional cuts were implemented in this analysis in order to optimize the event selection for a Higgs particle

Cuts	Jet
Jet p_T	> 25 GeV/c
$\|\eta\|$	< 2.4
CEF, NHF, NEF	< 0.99
CHF, NCH	> 0
$N_{constituents}$	> 1

Table 6.8: Jet definition

with a mass of 125 GeV/c^2. These cuts were kept to a minimum so as to maximize the number of events in signal that made our final selection. Maximizing the amount of signal while still cutting out some of the impact of backgrounds on our selection was critical for our plan to use a multivariate analysis technique. The additional cuts are described below, the largest impact coming from a b-tag veto cut described in section 6.3.6.

Met Cut

A cut on MET requiring $\not{E}_T > 25$ was imposed in order to cut down on the impact from QCD processes in our final selection while preserving as much signal as possible.

Lepton Cuts

The electron p_{T} cut was raised to 30 GeV to provide separation from the trigger threshold. As described previously, the trigger used is HLT_Ele27_WP80_v* which has a min p_{T} requirement of 27 GeV. Slightly raising the required electron energy helps avoid events right on on trigger threshold, while only losing ~5% expected signal as seen in Figure 6.3.

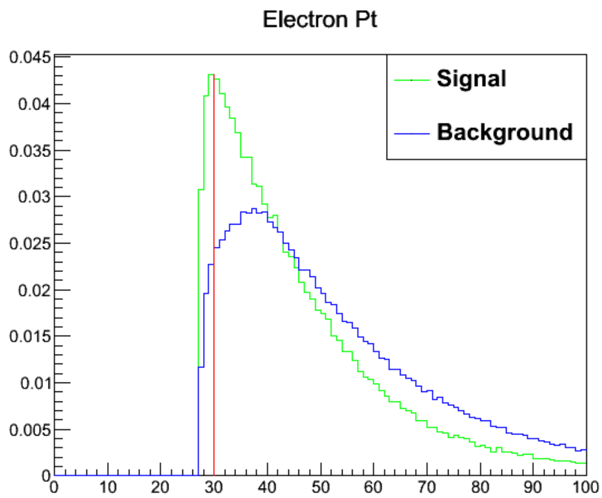


Figure 6.3: Normalized histograms of gluon-gluon fusion signal(green) W+jets background(blue). Red line shows cut level where 5% of signal is lost

When looking at muon selection the trigger used is HLT_IsoMu24_eta2p1_v*, which means it has a minimum energy requirement of 24 GeV for muons. As the selection and identification of muons in CMS is very good, we only impose a slight increase to this minimum in our analysis by requiring muons to have $p_{\mathrm{T}} > 25$ GeV. In addition, we require the electrons (muons) have an $|\eta| < 2.5$ (< 2.1).

Jet cuts

As described in section 6.3.5, we require all jets to have a minimum p_{T} of 25 GeV. In addition to this, we require that the leading jet in every event have a $p_{\mathrm{T}} > 30$ GeV. This additional selection helps to remove some of the multi-jet background while only minimally impacting signal acceptance. For events with more than 2 jets, we require that each jet beyond the first pass all of the jet criteria as outlined above. In our analysis we then split our sample up into 3 jets multiplicity bins: events with exactly 2 jets, exactly 3 jets, and 4 or more jets.

B-tag Veto Implementation

In addition to cuts on energy of the jet, it is useful if we can determine what kind of quark a jet came from. In order to do this, we employ an algorithm designed to tag jets as being from b quarks, specifically using the Combined Secondary Vertex (CSV) algorithm [82]. This algorithm relies on the ability of the tracking system in CMS to reconstruct secondary vertices. A jet deriving from a secondary vertex is a signature of b quark jets as the b quark lifetime is $\sim 1.5ps$, which corresponds to a flight distance in our detector of $\sim 450\mu m$, a distance measurable by the high granularity tracker that CMS employs.

The CSV algorithm uses a multi-variate approach which combines many input variables about the jet in order to generate a single discriminant that can be cut on in order to tag a being from a b quark. For this analysis we do not expect any b jets in our events, so we use this information to veto events with tagged b jets. Also, it's important to note that another analysis in CMS is searching for a Higgs via a Higgs produced in association with a vector boson where $H \to bb$ employs the same final state that we are looking for, but requires 2 b-tagged jets [91]. To ensure orthogonality to this analysis, only events with 1 or less b-tagged jet were considered.

Once we cut on the presence of more than 1 b-tagged jet, we separated events into two categories: events with 1 b-tagged jet, and events with zero b-tagged jets. We found that in looking at the events with 1 b-tag, there was a much larger impact of the $t\bar{t}$ background over events with no b-tags. Additionally, there was a significant expected yield in events where $H \to bb$. Together, these issues led us to use on the zero b-tagged events in our signal extraction and only use the events with 1 b-tag as cross checks. Tables showing the impact on the event yields in each of the categories are shown in section 6.5.

6.4 MC Corrections

While the modeling of MC is truly impressive in its breadth of calculation, ultimately it is limited by the models that are fed into it and the computing time required to calculate many orders of corrections to these perturbative theories. In addition, we are also limited by the instruments that measure our data, and while it is amazingly accurate in many respects, there are still limitations that need to be accounted for.

In our analysis jets and $\not\!\!E_T$ play a major role, and while every effort is made to measure or model the jets and $\not\!\!E_T$ as accurately as possible, this is not perfectly successful. Thus, several corrections must be made to these physics objects in order for data and MC to not only match each other, but to also describe what is going on within the detector. This section will discuss several sets of corrections and weights which were added to the samples after the data-taking and simulation steps.

The corrections we employ can be separated into two categories: corrections common to many CMS analyses, and corrections specifically designed for this analysis. The first category includes corrections to jet energy, jet resolution, $\not\!\!E_T$, b-tagging CSV discriminant weights, and top p_{T} weights. The second category includes corrections to the $\not\!\!E_\phi$ distribution, and corrective weights for our selected QCD sample. All corrections are applied either before or during the signal selection, while the weights are applied to the samples after selection (as they do not change any kinematic values).

6.4.1 Jet Energy Corrections

As we described earlier, the physics objects that we call jets are formed from the hadronization of quarks formed during a collision. There are many particles that make up this final jet object, and as such jets are not perfectly measured by the detector, nor perfectly reconstructed during processing. The response value of the jet,

$$R = \frac{p_{\mathrm{T}}^{RECO}}{p_{\mathrm{T}}^{ACTUAL}} \tag{6.1}$$

, where p_{T}^{RECO} is the reconstructed value of the jets momentum, and p_{T}^{ACTUAL} is the MC truth value of the jet momentum. This ratio is a measure of how well the detector measures the actual energy of the jet, and is very rarely 1. Thus, every analysis within CMS that uses jets must make use of the jet energy corrections (JEC) provided by the Jet Energy Resolution and Corrections (JERC) subgroup. These corrections seek to correct the response of the jets back to 1, on average.

As already stated, this analysis uses jets reconstructed using the anti-k$_T$ clustering algorithm with a cone size of 0.5 as part of particle flow jet reconstruction. The jet energy corrections we employ were designed for this type of jet. These corrections come from the global tags GR_R_53_V10 and START53_V7A for data and MC, respectively,which designate which data runs and which MC generation regimes the jets are coming from [92]. For MC we use the required L1FastJet, L2Residual and L3 Absolute corrections. For data, we use the equivalent levels as in MC plus the L2L3Residual corrections. These corrections should correct the jet responses back to 1 and make the responses for data and MC match.

The L1FastJet correction is a Charged Hadron Subtraction (CHS) correction which is implemented in the particle-flow algorithm, and involves subtracting the energy contributions from charged hadrons that are not associated with the jet from the energy cluster. The next stage, L2Residual correction, is a relative correction to make the measured jet response flat in η. The third stage, L3 Absolute, is a correction to the measured p_T of a jet in order to match the simulated jet p_T created using generator-level input and a similar jet-clustering algorithm. The L2 and L3 corrections are calculated using Monte Carlo, and thus when applying corrections to data a fourth correction factor is needed to fix the discrepancies between MC and data. This is called the L2L3 residual correction. These correction factors are described in reference [93], and are derived from 2011 7 TeV data, with a selection of dijet events near the Z-boson mass peak. A "tag-and-probe" procedure is applied to jets to determine the kinematic dependence (p_T and η) of the detector in both simulations and data. Additionally, a scale factor is needed to adjust for the difference in jet energy resolutionm, which will be described in the following section.

6.4.2 Jet Energy Resolution

One of the features of most MC samples that does not accurately represent what goes on in the detector is the jet resolution. Compared to the resolution of the real detector, the resolution in MC generated samples tends to be more sharply peaked with a smaller distribution of energies. This, in essence, means that the MC samples are simulating a better measure of the jet energies than we can actually measure with our detector. To correct this, the jet energies must be"smeared" such that the resolution in MC matches the resolution in data. There are multiple ways in which this 'jet smearing' can be employed; in this analysis we use a deterministic approach recommended by the JERC subgroup in which the reconstructed jet p_{T} is scaled based on the difference between matched, reconstructed, and generated jets [90]. The corrections are based on the jet η and can be found in table 6.9.

A multiplicative correction factor is calculated using this value of C_η as the η-based JER correction factor, seen in equation 6.2. The corrected jet then follows equation 6.3, where $\mathbf{X}_{jet}$ is the 4-vector of the jet. This corrected 4-vector contains the values used for the rest of the

Data/MC Correction Factors

$\lvert\eta\rvert$	Correction Factor C_η
	(factor +-stat. +syst.- syst.)
< 0.5	1.052±0.012+0.062-0.061
≥ 0.5 & < 1.1	1.057±0.012+0.056-0.055
≥ 1.1 & < 1.7	1.096±0.017+0.063-0.062
≥ 1.7 & < 2.3	1.134±0.035+0.087-0.085
≥ 2.3 & < 5.0	1.288±0.127+0.155-0.153

Table 6.9: Jet Energy Resolution (JER) correction scale factors by η

selection process.

$$C_{JER} = max\left(0.0, \frac{p_{\mathrm{T}}^{GEN}}{p_{\mathrm{T}}^{RECO}} + C_\eta \cdot \left(1 - \frac{p_{\mathrm{T}}^{GEN}}{p_{\mathrm{T}}^{RECO}}\right)\right) \tag{6.2}$$

$$\mathbf{X}_{Jet}^{corrected} = C_{JER} \cdot \mathbf{X}_{Jet}^{RECO} \tag{6.3}$$

Once the jet energy is corrected, it is important to remember that the measurement of the $\not\!\!E_T$ is intrinsically tied to the measurement of the jet energies. Scaling the jet energy changes the distribution of energy in the event, including the missing energy, meaning that $\not\!\!E_T$ must also be scaled appropriately. The two components of the $\not\!\!E_T$, $x_{\not E}$ and $y_{\not E}$, are corrected using equations 6.4 and 6.5.

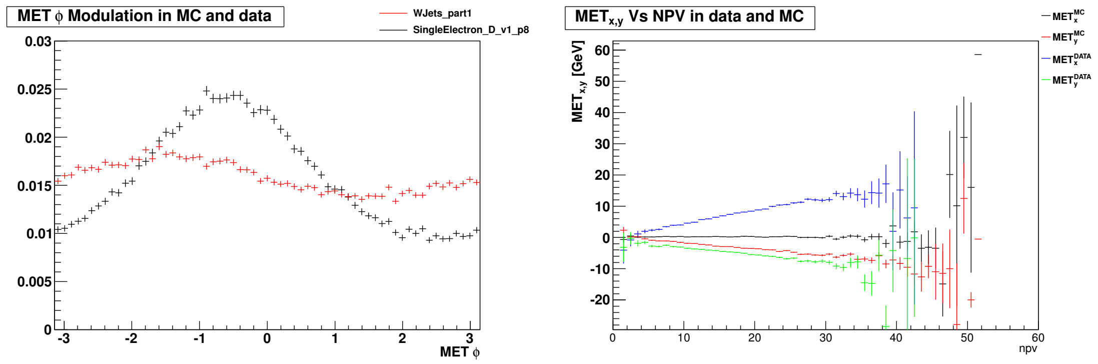
$$x_{\not E}^{corrected} = (1 - C_{JER})\, x_{Jet}^{RECO} + x_{\not E}^{RECO} \tag{6.4}$$

$$y_{\not E}^{corrected} = (1 - C_{JER})\, y_{Jet}^{RECO} + y_{\not E}^{RECO} \tag{6.5}$$

6.4.3 $\not\!\!E_{T\phi}$ Corrections

As described earlier, our analysis has an intrinsic contribution from $\not\!\!E_T$ as we expect a neutrino to be created in our signal decay that will not be measured by CMS. Since this is the case, we look carefully at the kinematics of this $\not\!\!E_T$ distribution to make sure that we are modeling and measuring it correctly. As the $\not\!\!E_T$ in our sample is attributed to a particle escaping detection, there should be no preferred direction (in the ϕ plane) for the decay to take place. Thus, any modulation seen in the distribution of ϕ of our measured $\not\!\!E_T$ must be an error in simulation or reconstruction and should be corrected.

As shown in figure 6.4a, there is a clear modulation in the ϕ distribution of $\not\!\!E_T$. The cause of this modulation is not known, though this effect could be seen if the collision of the proton beams was not head on. Any angle in the collision would produce a preferential scattering direction (which we do not want). This could also occur if there was an offset in the center of the proton bunches during collision. Though we do not know for sure what is causing this, we have established that this modulation is dependent upon the number of primary vertices in an event. This nPV dependence can be seen in figure 6.4b, where the x and y components of the $\not\!\!E_T$ scale with the nPV. Additionally, any cut on the p_T of the $\not\!\!E_T$ before the modulation is corrected will only exacerbate the problem, as the cut would preferentially select the events with $\not\!\!E_T$ on a specific side of the detector.



(a) (b)

Figure 6.4: (a) Distribution of $\not\!\!E_{T\phi}$ for both data and MC. Shown are W+jets MC in red and Single Electron data in black (b) The $\not\!\!E_{x,y}$ distributions as a function of nPV. Black and Red distributions show the MC $\not\!\!E_T$ X and Y distributions respectively, while blue and green show the data $\not\!\!E_T$ X and Y.

Though the modulations in data and MC are different, both need to be corrected to restore the expected 'flat' distribution in $\not\!\!E_\phi$. To correct for the modulation, each distribution of $\not\!\!E_{x,y}$ was fit with a first order polynomial. The parameters of this fit can be seen in table 6.10. These are then used to correct the $\not\!\!E_T$, for data and MC separately, using equations 6.6 and 6.7.

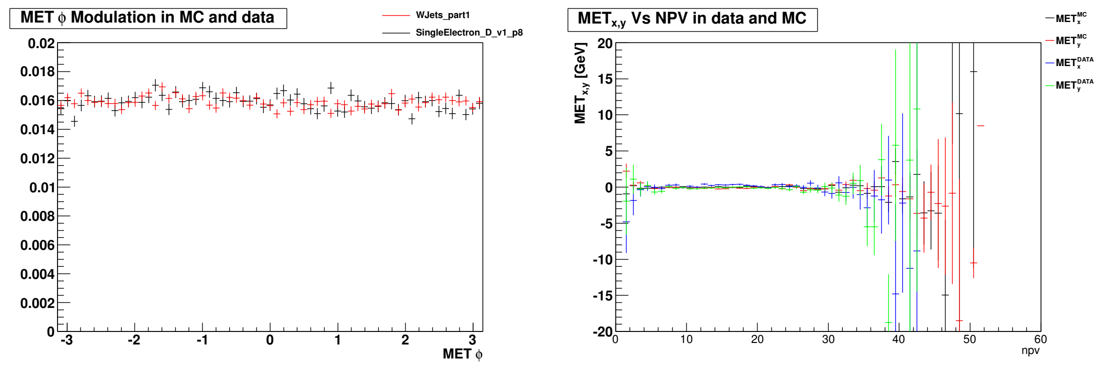
Sample	Parameter 0	Parameter 1
Data		
x	$2.0105E-01$	$4.2663E-01$
y	$-9.1350E-01$	$-2.3120E-01$
MC		
x	$2.9059E-01$	$-3.5293E-03$
y	$3.0183E-01$	$-1.9974E-01$

Table 6.10: List of parameter values for the $\not\!\!E_{T\phi}$ corrections.

$$\not\!\!E_x^{corrected} = \not\!\!E_x^{RECO} - ([0] + [1]{\cdot}nPV) \tag{6.6}$$

$$\not\!\!E_y^{corrected} = \not\!\!E_y^{RECO} - ([0] + [1]{\cdot}nPV) \tag{6.7}$$

Figure 6.5a shows the corrected $\not\!\!E_{T\phi}$ distributions while figure 6.5b shows the x and y distributions as a function of nPV. The modulations becomes negligible in both sets of distributions. At this point, it is safe to perform a p_T cut on the $\not\!\!E_T$ variable without a bias.



(a) (b)

Figure 6.5: $\not{E}_{T\phi}$ distributions with corrections applied. (a) Shown are W+jets MC in red and Single Electron data in black (b) The $\not{E}_{x,y}$ distributions as a function of nPV. Black and Red distributions show the MC $\not{E}_T$ X and Y distributions respectively, while blue and green show the data $\not{E}_T$ X and Y.

6.4.4 CSV Reweighting

In section 6.3.6, I introduced the identification criteria we use to tag a jet as coming from a b quark, the Combined Secondary Vertex (CSV) discriminant. The method of calculating these discriminants is described in full detail in AN-2006/014 [82] and a paper [94]. For this analysis we rely on rejecting events with b-tagged jets in our event selection, so any corrections to that discriminant that are needed should be applied. In CMS it has been noted [95] that a calibration is necessary to correct this CSV discriminant in order to make the data and MC distributions match.

This method corrects the per-jet CSV for both heavy and light flavor jets by calculating an event weight scale factor in exclusive bins of jet CSV output, jet p_{T}, and (in the case of light flavor jets) jet η. These weights are derived in [95] by comparing data and MC distributions, leading to the scale factors (SFs) that are binned by p_{T} and η. Using the three b-tag efficiency measurements described by the BTag Physics Object Group [96], there are three pairs of values to compare (CSV_{orig}, CSV_{equiv}). The reshaping function must satisfy $f(CSV_{\mathrm{equiv}}) = CSV_{\mathrm{orig}}$ for each of the operating points and for the upper and lower values of the CSV discriminant to make sure those values do not change (e.g., $CSV = 0.0$ and $CSV = 1.0$). The whole range of CSV discriminant values is found by linearly interpolating between these five points (the three working points, and upper and lower limit of the discriminate range).

The prescription that we used categorizes the jets into three flavors by checking the MC truth: heavy flavor (b jets), charm jets, and light flavor (anything else). The heavy flavor SFs are separated into 5 p_{T} bins with lower bounds at 25, 40, 60, 100, and 160 GeV with the lower bound being inclusive. The charm jets are given a flat scale factor of 1, and are described fully in a CMS Analysis Note (AN) [95]. For light flavor jets, a slightly different approach was taken.

There are only 3 p_{T} bins used; 25 to 40, 40 to 60, and > 60 GeV. Each of these p_{T} bins is then split into 3 bins by abs(η): < 0.8, ≥ 0.8 and < 1.6, ≥ 1.6 and < 2.41.

For each event, all of the jets in the event (that have passed our preselection criteria) are looped over and a weight value is calculated for that jet based on the flavor, p_{T}, and η of the jet. The individual jet weight is then combined multiplicatively with the weights of every jet in the event, and the resultant product is the CSV weight that is assigned to that event.

6.4.5 TTbar Reweighting

This section describes the procedure for calculating weights to correct the p_{T} spectrum of the TTbar MC sample we are using. In the normalized differential top-quark-pair cross section analysis, the shape of the p_{T} spectrum of the individual top quarks in data was found to be softer than predicted by the various simulations, resulting in an overestimation of the p_{T} of events with a top quark. This was described by the TOP-PAG [97], though they note that NNLO predictions [98] provide a reasonable description.

In this analysis, we use the results from the TOP-PAG referenced above for 8TeV single lepton events to generate a corrective weight based on the p_{T} of the top quarks in our TTbar MC sample. Using equations 6.8 and 6.9

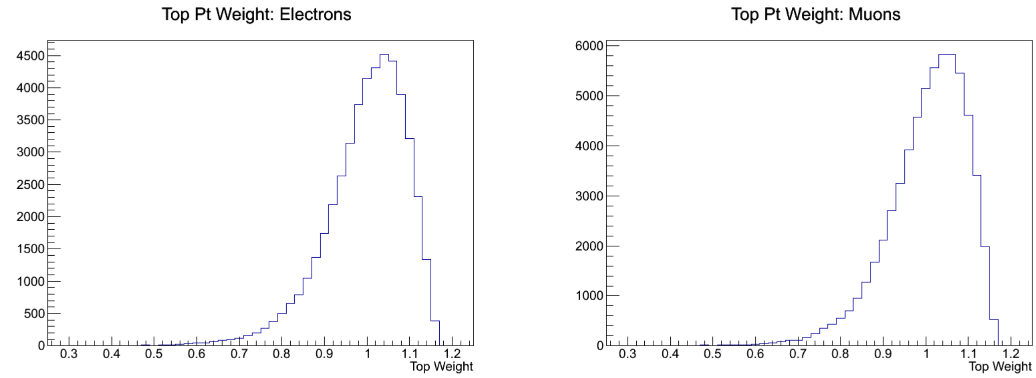
$$Weight = \sqrt{SF(t)SF(\bar{t})} \tag{6.8}$$

$$SF(x) = e^{a+b*x} \tag{6.9}$$

where A = 0.159 , B = -0.00141, and x = p_{T} of the top or anti-top quark in the event, a SF is calculated for each top quark in the event and combined to generate an event weight. Distributions of this weight are shown in figure 6.6 for both electron and muon samples. You can see from these plots that while the main peak is centered around 1, there is a longer tail on the side below 1 which results in scaling down of events with too high of a top quark p_{T}.

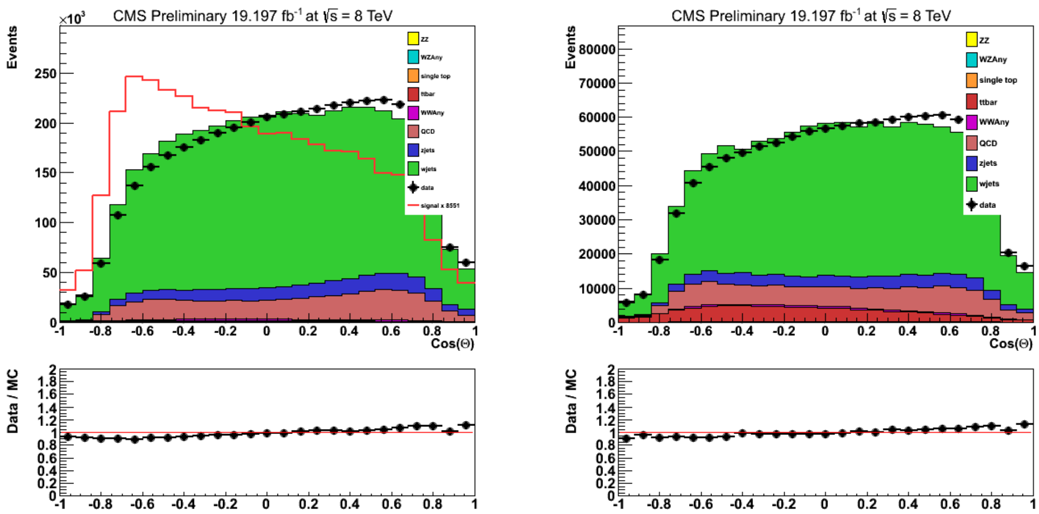
6.4.6 $cos(\theta_l)$ Weights

When looking at the comparison between data and MC we noticed a linear trend in the residual plot in the angular variable $cos(\theta_l)$. $cos(\theta_l)$ is one of the angular variables that describe the decay of the WW system. Specifically, $cos(\theta_l)$ is the cosine of the angle between the lepton in the decay and the WW decay plane. A diagram of this decay can be seen later in figure 7.4. Trends like this represent an error in the simulation, and as it is a linear trend we can quantify and correct for it. Figure 6.7a shows this variable in our 2-jet region, though the trend exists in



(a) (b)

Figure 6.6: Distribution of Top p_{T} weight for electron events (a) and muon events (b)



(a) (b)

Figure 6.7: Comparison of the Data to MC agreement in the 2-jet bin for $cos(\theta_l)$ in our signal region (a). Note the overestimation of data compared to MC in the low values, and the ovestimation in the high values, leading to a linear trend in the residual plot. (b) Comparison of $cos(\theta_l)$ for events with 1 b-tagged jet.

all of our jet bins.

This linear trend can be easily corrected for by generating a weight by comparing data and MC. We used only the W+jets MC sample, as it makes up the majority of the background and correcting this one background should improve overall agreement. In order to generate a correction without biasing our backgrounds by directly fitting to our signal region, we must find a control region that we can generate our new weights from. In this case, we used events that pass all of our signal selection except for the zero b-tag requirment. Instead, we use events that have exactly one b-tagged jet. Figure 6.7b shows the comparison of events in this sample, and it is clear that the same linear trend exists there.

In order to accurately generate corrective weights we must take into account the differences between our signal region and the control region (events with 1 b-tagged jet). Comparing the

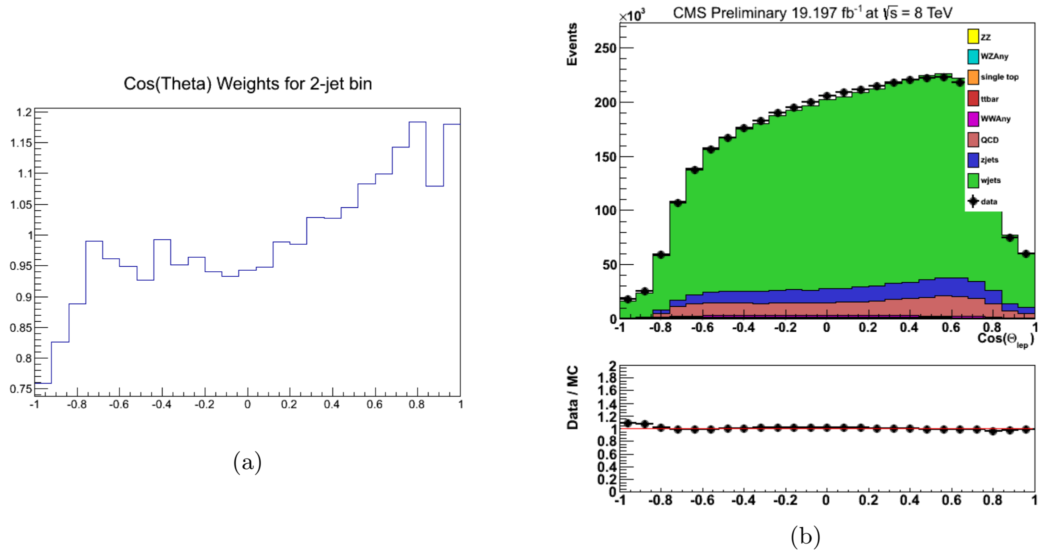


Figure 6.8: (a) Corrective weights generated by comparing data and W+jets MC in a 1 b-tag control region. (b) Data to MC agreement for $cos(\theta_l)$ in the signal region after weights have been applied.

impact of the MC backgrounds, it is clear that the $t\bar{t}$ sample has a much larger impact in the control region. This is expected, as $t\bar{t}$ events have at least 2 real b jets in them. In order to correct for this difference in background fraction, we scaled our $t\bar{t}$ MC sample to it's expected yield and subtracted it from the data sample. From there, we could directly compare the data and W+jets MC to generate the weights.

Using the weights shown in figure 6.8a, new event weights were calculated for our W+jets MC sample, which were applied multiplicatively with the pileup and CSV weights. The resulant disribution is shown in figure 6.8b. From that plot you can see that the linear disagreement has been flattened out by these weights. The same process was followed for the 3 and ≥ 4 jets bins.

6.4.7 QCD Reweighting

QCD η Weights

Our QCD sample is obtained from data selecting on anti-isolated leptons as described in 6.2. Selecting on this isolation gives us a sample of events that models QCD events well, but by design these events have different selection criteria that events in our signal region. We found that the ratio of the number of events in the signal region to those in the antiIso region varies dramatically over η and therefore the yields in our anti-isolated QCD region need to be transformed to represent the yields in the signal region. This led to the generation of weights based on the η of the QCD events that are applied after selection to correct the distribution to that seen in the signal region.

This η dependence of the QCD sample is clearly seen in QCD monte carlo events. We use

MC for this example because it allows us to compare QCD events that are in our signal region of isolation directly. While the QCD MC is lacking in statistics in some of the low p_{T} samples, the eta dependence is clearly seen in the higher pt samples. By combining six MC generated QCD samples according to their cross sections, we were able to obtain a single QCD sample to compare to. Table 6.11 delineates the ranges of $\hat{p_{\mathrm{T}}}$ that each sample covers as well as the cross sections used for combination.

$\hat{p_{\mathrm{T}}}$ Range (GeV)	Cross Section σ (pb)
20 to 30	2.866e+08
30 to 80	7.433e+07
80 to 170	1.191e+06
170 to 250	30990
250 to 350	4250
>350	810

Table 6.11: Jet Energy Resolution (JER) correction scale factors by η

Figure 6.9 shows the number of events in the signal region over the number of events in the anti-isolated region for each of these samples. The ratios are particularly high in $|\eta|$ regions that correspond to the endcaps of our detector ($|\eta| >\sim 1.5$) with values several times larger than those found in the central region.

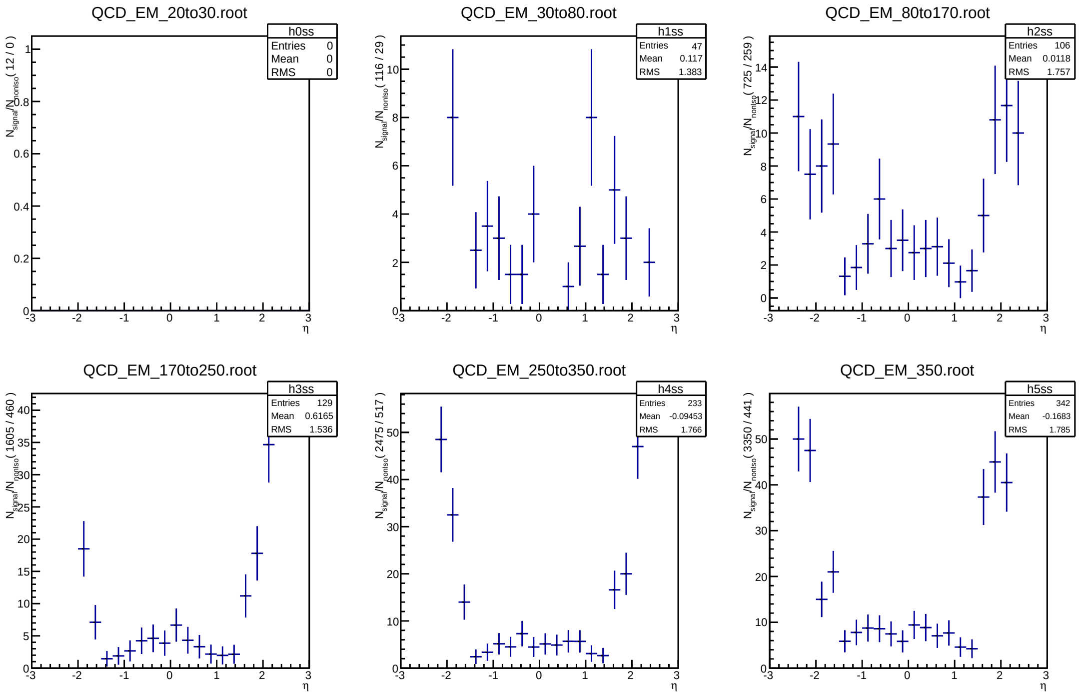


Figure 6.9: Ratio of number of events in the signal region to the events in the antiIso region for each of the six samples. The numbers in parentheses on the Y-axis reflect the number of entries in both regions. The first plot is empty due to a lack of MC statistics that passed our selection for low values of p_{T}.

For application to our analysis, we use a data-driven technique based on the zero intrinsic $\not\!\!E_T$ characteristic of the QCD sample. We separate our sample into 13 bins of $|\eta|$, where the

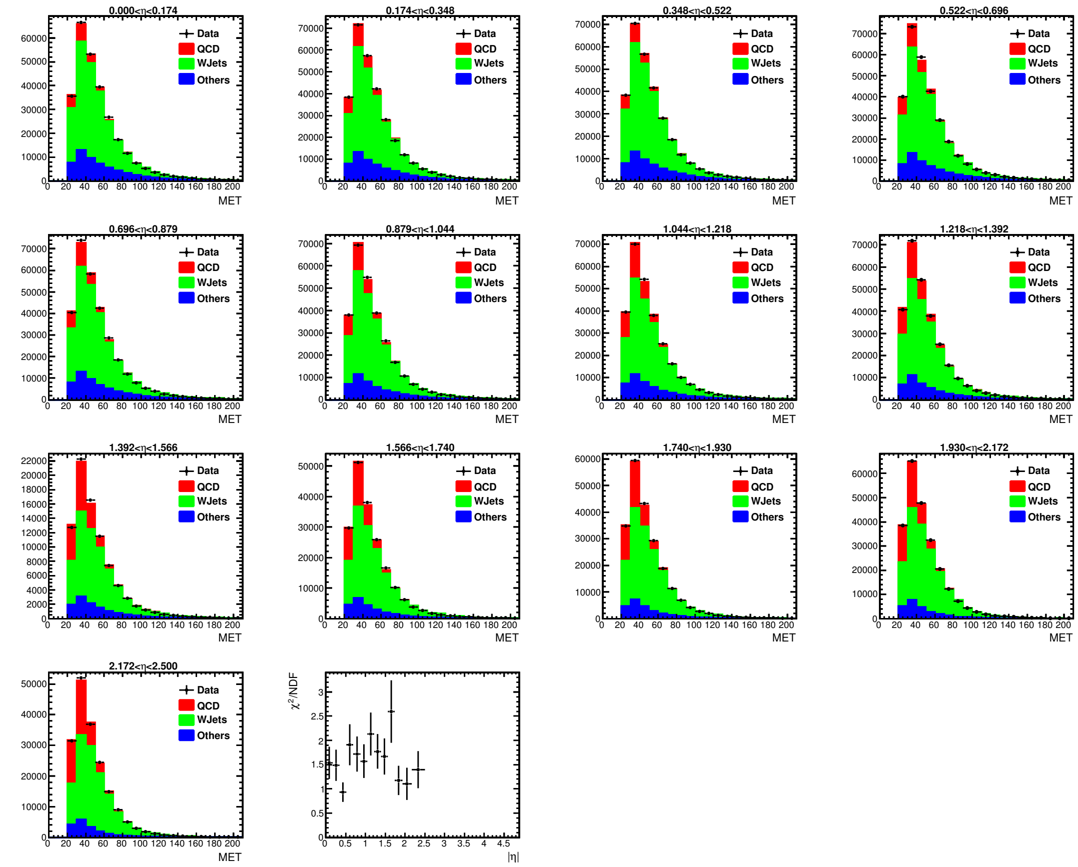


Figure 6.10: $\not\!\!E_T$ distribution for all 13 η bins. The distribution of QCD (solid red) and W+jets (solid green) has been fit to the data (dots) while keeping all other processes (blue) fixed to their SM expectation. The last plot shows the χ^2/NDF of the fits.

separations between bins are located at $|\eta| = \{0, 0.174, 0.348, 0.522, 0.696\ 0.879, 1.044, 1.218, 1.392, 1.566, 1.740, 1.930, 2.172, 2.5\}$, and the lower boundary of each bin is inclusive. In order to derive our correction factors we are interested in finding a function $s_{QCD}(\eta)$ such that:

$$N^{QCD}_{antiIso}(\eta) s_{QCD}(\eta) = N^{QCD}_{signal\ region}(\eta) \tag{6.10}$$

, where $N^{QCD}_{antiIso}$ and $N^{QCD}_{signal\ region}$ represent the number of QCD events in the anti-isolated and signal regions respectively, both for the same given luminosity. In each $|\eta|$ bin we measure the total QCD and W+jets yields by fitting their $\not\!\!E_T$ distribution to the data distribution in the signal region. The fit allows for the free variation of the QCD and W+jets normalization while keeping all other backgrounds fixed to the their SM expected normalization. In each η bin the fit returns the amount of data due to QCD ($N^{QCD}_{signal\ region}$) and the amount due to W+jets ($N^{W+jets}_{signal\ region}$). Figure 6.10 shows the fits in all of the η bins as well as the χ^2/NDF of all fits.

As defined in equation 6.10, we compute s_{QCD} for each η by dividing the measured $N^{QCD}_{signal\ region}$ by the number of data events in the anti-isolated region ($N^{QCD}_{antiIso}$). Figure 6.11 shows the resulting s_{QCD} as a function of absolute η derived using the full dataset signal and anti-isolated

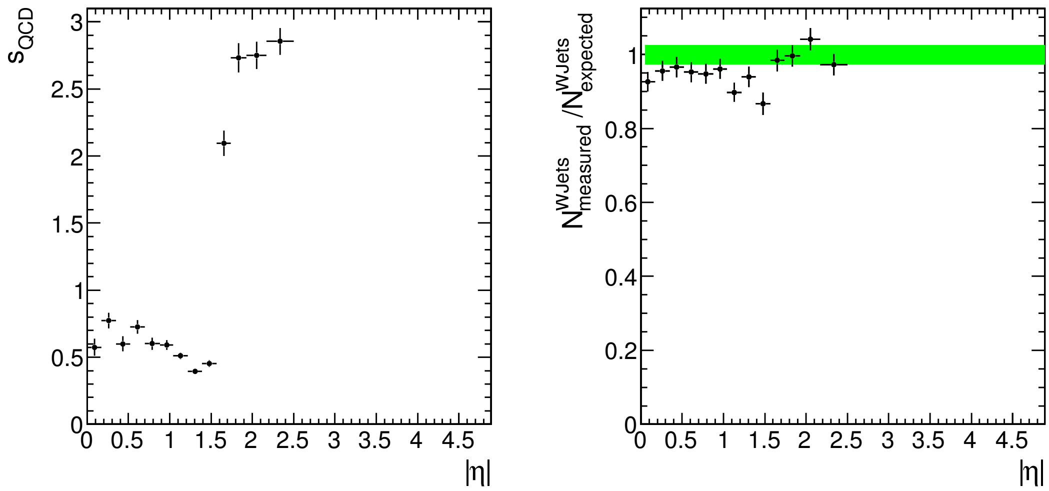


Figure 6.11: Left: Scale factors for QCD as a function of absolute lepton η. Right: Ratio of measured yield of W+jets events found from fitting to data to the SM expected yield of W+jets. The green band indicates the error on the expected SM W+jets cross section.

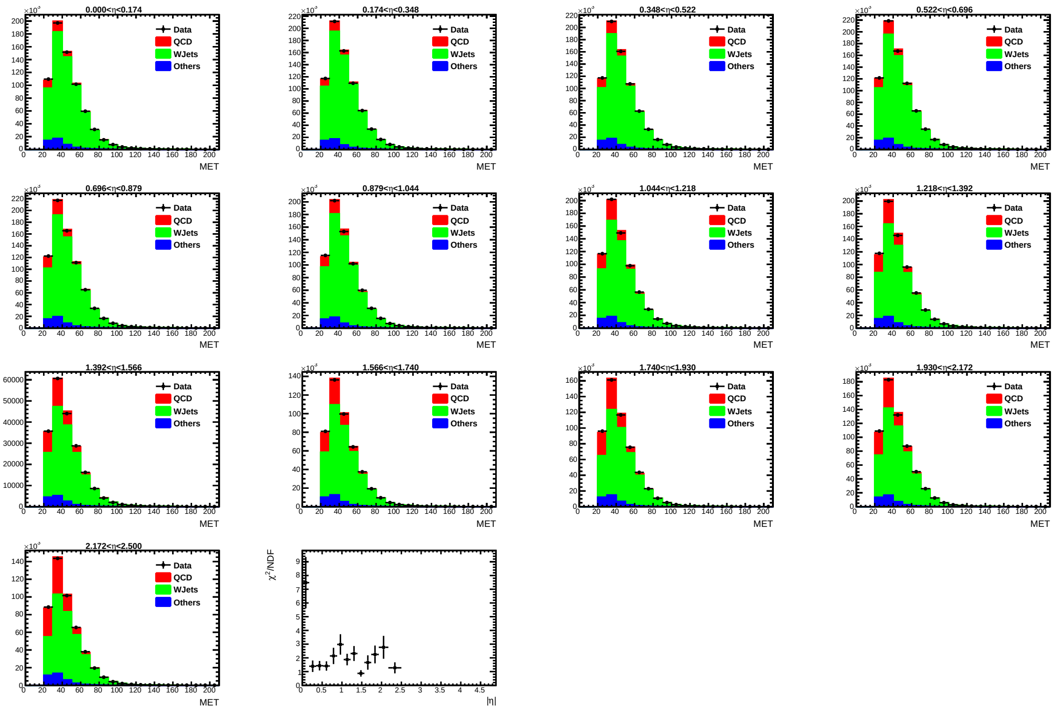
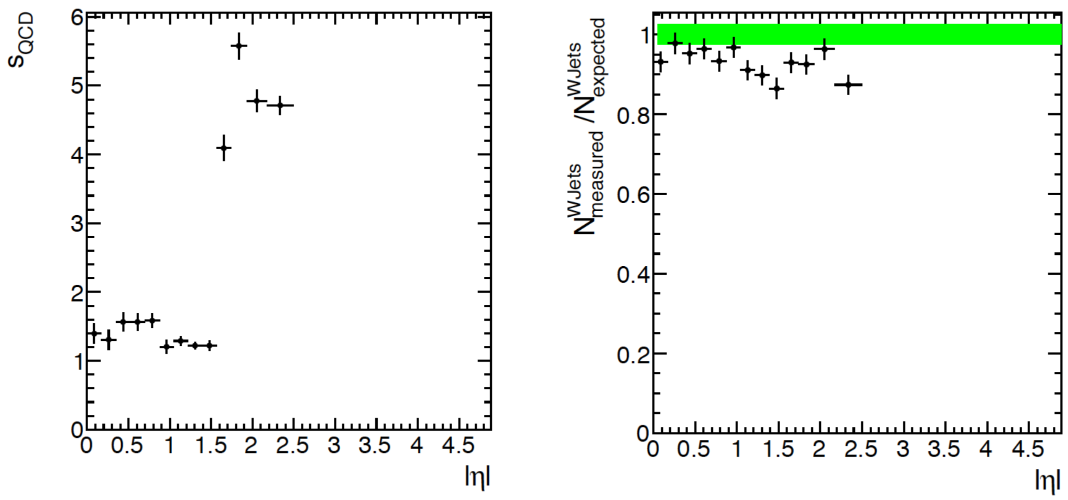


Figure 6.12: $\not\!\!E_T$ distributions for all η bins in the 1 jet control region electron QCD fits.

sample of $L_{antiIso} = 19148\ pb^{-1}$. These fits are shown using a data selection of 2+ jets in order to model the events in our signal region. Note that the shape of this distribution follows very closely to what we saw in the QCD monte carlo sample 6.9.



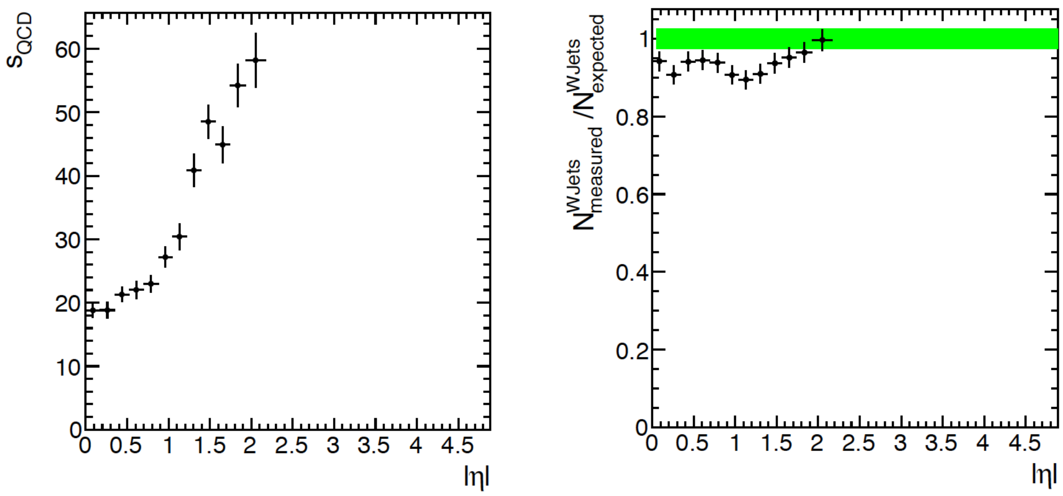
(a) (b)

Figure 6.13: Electron QCD scale factors (a) as a function of absolute lepton η in the 1 jet control region. (b) Ratio of the measured number of W+jets events to the SM expected number in the 1 jet control region.

As a cross check, the measured value of the W+jets yields ($N^{W+jets}_{signal\ region}(\eta)$) divided by the SM expectation are shown in figure 6.11 (right) as a function of absolute η. A flat linear fit over all η points in this distribution results in a value of 0.953 ± 0.008, which is not consistent with a value of one when considering the 2.56% error reported in the expected W+jets cross section. This shows that an additional, absolute scale factor is required to modify the yields for both the QCD and W+jets samples. So far only the shape of the QCD sample has been corrected by the weights seen in Figure 6.11 (left), but not the overall yields. The overall normalization correction is defined in the 'QCD and W+jets Yields' section below.

While these QCD scale factors would almost certainly correct the isolated/anti-isolated ratio, we would in effect be using the same signal events to both create the weights and to do a signal extraction. To avoid this, a control region was chosen that returned similar weights to those found in the signal region, but which contained a completely orthogonal set of events. The control region's selection was the same as the signal region except that it contained exactly 1 jet, as opposed to the signal region's 2+ jets selection. The $\not{E}_T$ fits for this control region can be seen in figure 6.12, and the scale factors can be found in figure 6.13. From these results it is clear that the the two regions return similar scale factors, a result that shows the control region scale factors will correct our signal region accurately while having no deleterious effects involved with fitting our signal region.

An identical setup and procedure to that described above was performed on the QCD muon sample as well. This included a separate set of fit and a separate set of event weights that were generated specifically for our muon QCD sample. The resultant weights for QCD and W+jets scaling factors are shown in figure 6.14. Note that while the weight values for muon QCD events are relatively large, this has little impact as it is the value of the weights relative to themselves that matter. This is because the QCD normalization will still be applied after the η weights are applied, which will scale the entire sample appropriately and correct any arbitrary inflation caused by the scaling.



(a) (b)

Figure 6.14: Muon QCD scale factors (a) as a function of absolute lepton η in the 1 jet control region. (b) Ratio of the measured number of W+jets events to the SM expected number in the 1 jet control region.

QCD Pile Up Weights

In addition to the (now fixed) η dependence, we noted that the distribution of primary vertices did not match that seen in data. For the MC generated samples we use the standard correction technique of generating weights based on the number of interactions in the event. This is possible for MC events as the true number of interactions generated can easily be ascertained, but the number of interactions in data is generated by a CMS macro [88] that uses information about the data run to calculate the expected number of interactions. As our QCD selection comes directly from data and the pile up tool does not account for selection bias, we instead use the number of primary vertices in the event to weight our sample.

Figure 6.15 shows the number of primary vertex distributions in our QCD and data selection as well as the weights generated for the QCD to correct for our selection bias. These weights are then applied in the same manner as the standard pileup weights are applied to our other

MC samples, which in the case of QCD means in combination multiplicatively with the above calculated η weights.

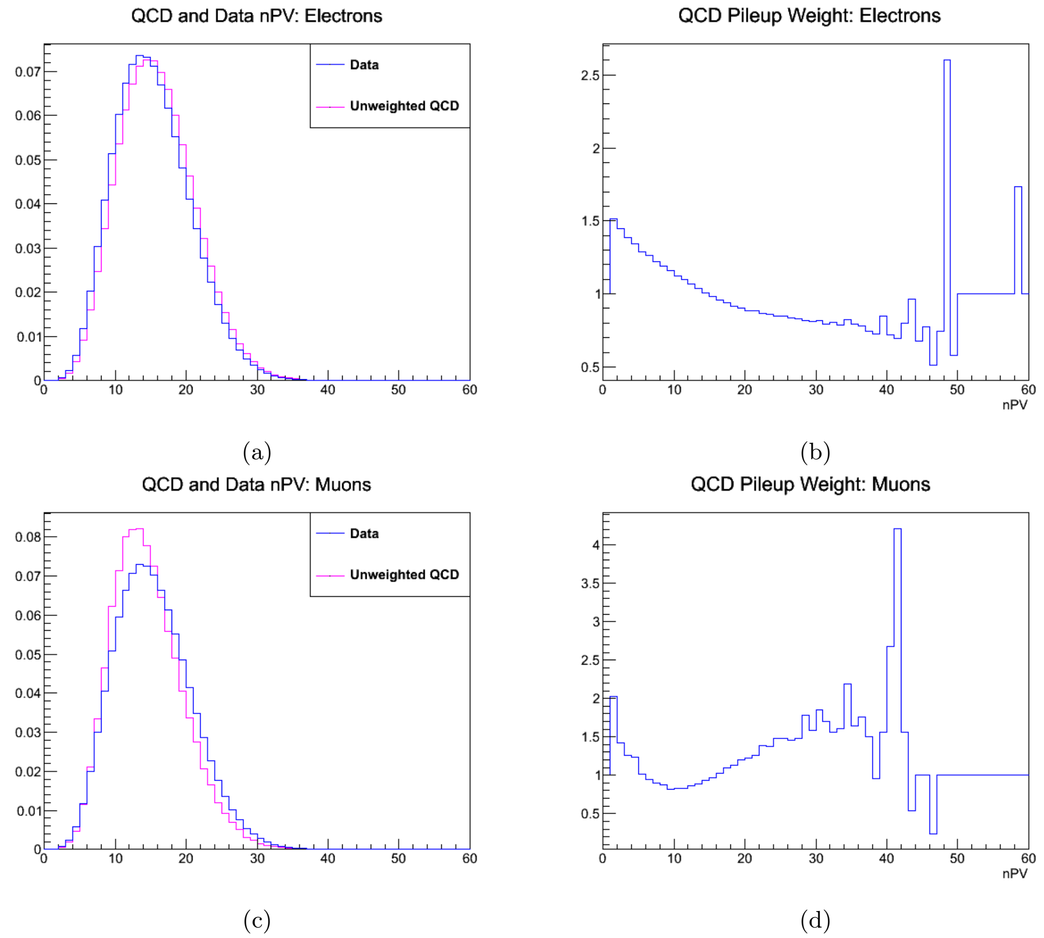


Figure 6.15: Number of Primary Vertex distribution comparisons (a,c) and associated weights (b,c) for data-driven QCD sample for electrons(a,b) and muons (c,d).

QCD and W+jets Yields

Using the corrections based on QCD η described above, we have fixed the shape of the QCD distribution. Also as described above, we have shown that we require a scale factor to be applied to the W+jets same so that it will more correctly match what we see in data. To do this we perform a two component fit to data without binning in η and allowing for the free variation of the W+jets and QCD. We take the cross sections of the other backgrounds as constants and scale their MC appropriately to the correct expected yields. Then we subtract these samples from the data leaving a distribution that should only contain contributions from W+jets and QCD events. By fitting the $\not{E}_T$ distribution using the two templates we can extract the fraction of that data that is predicted to come from each process. Comparing this yield to the MC predicted W+jets yield we get a scale factor to apply to the MC. Similarly, by taking the QCD yield as the correct expected yield, we are able to calculate what the cross section would have to be to yield the correct number of events. This fit is performed in each of the jet bins separately, with results shown in table 6.12.

Jet Bin	W+jets SF Ele	W+jets SF Mu	QCD xSec Ele	QCD xSec Mu
2 Jets	1.027 ± 0.004	0.99 ± 0.003	76.6 ± 3.1 pb	38.3 ± 1.5 pb
3 Jets	1.063 ± 0.010	0.995 ± 0.009	76.30 ± 6.7 pb	41.3 ± 2.7 pb
≥ 4 **Jets**	1.12 ± 0.021	1.013 ± 0.017	60.7 ± 13.5 pb	38.8 ± 4.9 pb

Table 6.12: List of W+jets Scale Factor values and QCD xSec values from $\not\!\!E_T$ fit.

6.5 MC Yields

Now that we have all of the MC corrections in place, we can look at the expected yields for each of our simulated signal and backgrounds. In this section I show the yields for both the zero b-tagged events and the 1 b-tag events, though we only use the zero b-tagged events in signal extraction. The 1 b-tag events are a useful cross check, and showing the expected yields helps illustrate why we chose not to use those events in this analysis.

Table 6.13 shows the yields for events with one b-tag, while table 6.14 shows yields for events with zero b-tagged jets. Additionally, the impact of each signal or background sample becomes readily apparent when they are viewed as percent yield tables instead of raw yields. Tables 6.15 and 6.16 show the percentage yields for events with 1 and 0 b-tags respectively, where the yields are normalized to the sum of events in their section (background, signal).

In order to better understand these tables, there are a few notations that need to be explained. Events from Higgs MC samples that are not $H \to WW$ events are referred to as 'volunteer signal', whereas events from all $H \to WW$ MC samples are shown as 'true signal'. Both types of signal events are normalized to the sum of $H \to WW$ events. In this way, we can compare how many events we would expect from these 'volunteer signal' events in respect to our true signal, and attempt to minimize the impact of this in our analysis.

From these tables we see that for the zero b-tag events (table 6.16), the dominant background for all jet bins is the W+jets sample. Also, the sum of the 'volunteer signal' events is at most 7% of the expected $H \to WW$ signal, showing that if we cut events with any b-tags, we can remove most of the contamination from these extraneous samples.

For the events with 1 b-tag, the story is different. Looking at table 6.15, we can see while W+jets is the dominant background for the 2 and 3 jet bin, when allowing 4 or more jets the TTbar background becomes dominant. Moreover, the 'volunteer signal' is as much as 87%, making it harder to distinguish the signal events we are looking for from the background. These reasons directly highlight why we chose not to use events with 1 b-tag, restricting our signal region to only events with zero b-tags.

Event Yield for 1 b-tag H→WW→lνjj 19.1 fb^{-1} Ele & Mu Sample

Process	**== 2**	**== 3**	**≥4**
Diboson	12028.09	5369.18	1967.63
W+jets	773253.48	272857.9	103508.87
Z+jets	64497.39	24237.81	9835.04
$t\bar{t}$	49612.48	86120.65	122073.6
Single t	40209.27	21303.23	10768.92
Multi-Jet	123928.96	43101.4	16061.17
Tot Bkg	**1063529.67**	**452990.17**	**264215.23**
ggH, H→WW $M_H = 125$	118.08	67.63	35.12
qqH, H→WW $M_H = 125$	22.46	16.92	8.19
WH_ZH_TTH, H→WW M_H125	35.76	34.35	49.09
Total H→WW	**176.3**	**118.9**	**92.4**
WH_ZH_TTH, H→ZZ M_H125	3.34	2.55	3.61
WH, H→ $b\bar{b}$ M_H125	148.12	53.31	15.35
TTH, H→ $b\bar{b}$ $M_H = 125$	2.1	5.94	22.7
Total 'Volunteer' Sig	**153.56**	**61.8**	**41.66**
Signal$_{H\to WW}$ / Bkg	0.000166	0.000262	0.000349
Signal$_{H\to WW}$ /$\sqrt{Bkg}$	0.171	0.177	0.179

Table 6.13: Expected event yield normalized to cross sections and luminosity. Top section shows background processes with all diboson processes combined as well as all single top processes combined. The middle section shows contributions from all $H \to WW$ processes that are considered as signal. Bottom section shows other Higgs processes that are not part of our signal that could contaminate our final state ('Volunteer Signal').

Event Yield for 0 b-tag H→WW→lνjj 19.1 fb^{-1} Ele & Mu Sample

Process	**== 2**	**== 3**	**≥4**
Diboson	39026.22	12612.58	3485.46
W+jets	3271138.31	726384.44	187723.52
Z+jets	272583.99	69588.32	19937.11
$t\bar{t}$	20005.51	24748.61	27686.99
Single t	16318.38	7096.2	3036.83
Multi-Jet	450503.85	119248.8	33681.6
Tot Bkg	**4069576.26**	**959678.95**	**275551.51**
ggH, H→WW $M_H = 125$	473.7	182.2	68.98
qqH, H→WW $M_H = 125$	92.06	45.17	16.51
WH_ZH_TTH, H→WW M_H125	124.51	77.97	42.95
Total H→WW	**739.27**	**323.51**	**137.22**
WH_ZH_TTH, H→ZZ M_H125	8.27	4.4	2.25
WH, H→ $b\bar{b}$ M_H125	40.2	12.63	3.39
TTH, H→ $b\bar{b}$ $M_H = 125$	0.53	1.14	3.14
Total 'Volunteer' Sig	**49.00**	**18.17**	**8.78**
Signal$_{H\to WW}$ / Bkg	0.000169	0.000318	0.000466
Signal$_{H\to WW}$/$\sqrt{Bkg}$	0.342	0.312	0.245

Table 6.14: Expected event yield normalized to cross sections and luminosity. Top section shows background processes with all diboson processes combined as well as all single top processes combined. The middle section shows contributions from all $H \to WW$ processes that are considered as signal. Bottom section shows other Higgs processes that are not part of our signal that could contaminate our final state ('Volunteer Signal').

Fractional Yield for 1 b-tag H→WW→lνjj 19.1 fb^{-1} Ele & Mu Sample

Process	**== 2**	**== 3**	**≥4**
Diboson	0.011	0.012	0.007
W+jets	0.727	0.602	0.392
Z+jets	0.061	0.054	0.037
$t\bar{t}$	0.047	0.190	0.462
Single t	0.038	0.047	0.041
Multi-Jet	0.117	0.095	0.061
Tot Bkg	**1.000**	**1.000**	**1.000**
ggH, H→WW $M_H = 125$	0.670	0.569	0.380
qqH, H→WW $M_H = 125$	0.127	0.142	0.089
WH_ZH_TTH, H→WW M_H125	0.203	0.289	0.531
Tot H→WW	**1.000**	**1.000**	**1.000**
WH_ZH_TTH, H→ZZ M_H125	0.019	0.021	0.039
WH, H→ $b\bar{b}$ M_H125	0.840	0.448	0.166
TTH, H→ $b\bar{b}$ $M_H = 125$	0.012	0.050	0.246
Tot 'Volunteer' / Tot H→WW	**0.871**	**0.520**	**0.451**

Table 6.15: Expected event yield normalized to total yield. Background samples are normalized to total background, while Higgs samples are normalized to total $H \to WW$ contribution. Dominant background highlighted for each jet bin, here W+jets is dominant for the 2 and 3 jet bin but $t\bar{t}$ is dominant for ≥ 4 jets.

Fractional Yield for 0 b-tag H→WW→lνjj 19.1 fb^{-1} Ele & Mu Sample

Process	**== 2**	**== 3**	**≥4**
Diboson	0.010	0.013	0.013
W+jets	0.804	0.757	0.681
Z+jets	0.067	0.073	0.072
W+jets	0.804	0.757	0.681
Z+jets	0.067	0.073	0.072
$t\bar{t}$	0.005	0.026	0.100
Single t	0.004	0.007	0.011
multi-Jet	0.111	0.124	0.122
Tot Bkg	**1.000**	**1.000**	**1.000**
ggH, H→WW $M_H = 125$	0.686	0.597	0.537
qqH, H→WW $M_H = 125$	0.133	0.148	0.129
WH_ZH_TTH, H→WW M_H125	0.180	0.255	0.334
Tot H→WW	**1.000**	**1.000**	**1.000**
WH_ZH_TTH, H→ZZ M_H125	0.012	0.014	0.018
WH, H→ $b\bar{b}$ M_H125	0.001	0.004	0.024
TTH, H→ $b\bar{b}$ $M_H = 125$	0.058	0.041	0.026
Tot 'Volunteer' / Tot H→WW	**0.071**	**0.060**	**0.068**

Table 6.16: Expected event yield normalized to total yield. Background samples are normalized to total background, while Higgs samples are normalized to total $H \to WW$ contribution. Dominant background highlighted for each jet bin, here W+jets is dominant for all jet bins.

Chapter 7

$H \to WW \to l\nu qq$ Analysis Part 2: MVA

In order to separate our signal sample ($H \to WW$) from our background samples, we utilize information contained in many different variables, as no single variable provides enough discriminating power on its own. By combining the information of several input variables in a multivariate analysis (MVA), a more powerful discrimination can be achieved. For this analysis, the MVA algorithm chosen was that of a Boosted Decision Tree (BDT). It has been implemented in the ROOT TMVA framework, available in all CMSSW releases. A BDT is trained for each jet category; each optimized separately for which input variables are used, the number of variables, and the BDT training parameters.

In this section I will describe the method for generating and optimizing our BDT using kinematic variables as the inputs. In section 7.2 I describe the selection of kinematic variables, in section 7.3 I describe the individual optimization of the the BDTs, and in 7.3.1 I describe the final optimization of BDT parameters.

7.1 Multivariate Analysis: Boosted Decision Tree

A decision tree is a binary tree structured classifier similar to the one shown in figure 7.1. Repeated left/right (yes/no) decisions are taken on one single variable at a time until a stop criterion is fulfilled. The phase space is split this way into many regions that are eventually classified as signal or background, depending on the majority of training events that end up in the final leaf node. The concept of 'boosting' a decision tree extends this concept from one single decision tree to many trees which form what is knows as a 'forest' of decision trees. Each tree is derived from the same set of training events, but allows for weighting so that each tree can learn from the previous one. The act of boosting helps to stabilize the response of the decision tree with respect to fluctuations in the training sample, and is able to considerably enhance the performance of the discriminant over that of a single tree [99].

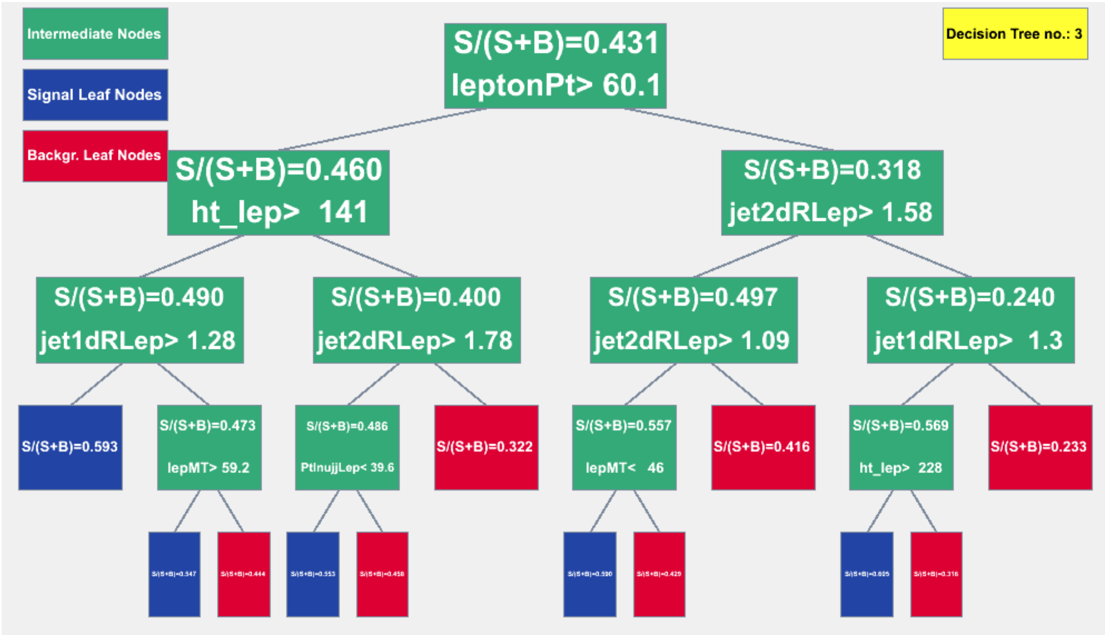


Figure 7.1: Example of a decision tree found in the kinematic BDT analysis. Starting with the root node (very top green box), a sequence of binary splits using the discriminating variables provided as input is applied to the data. Each split uses the variable that at this node gives the best separation between signal and background when being cut on. The same variable may thus be used at several nodes (as seen here with the variable jet2dRLep), while others might not be used at all. The leaf nodes at the bottom end of the tree are shown in blue for signal and red for background, depending on the majority of events that end up in the respective nodes.

Decision trees allow a straightforward interpretation as they can be visualized by a simple two-dimensional tree structure. In this respect, BDTs are similar to rectangular cuts; however, whereas cut-based analysis is able to select only one hypercube as a region of phase space, the decision tree is able to split the phase space into a large number of hypercubes, each of which is identified as either ‘signal-like’ or ‘background-like’. For decision trees, the path down the tree to each leaf node represents an individual cut sequence that selects signal or background depending on the type of the leaf node.

A shortcoming of decision trees is the instability of their output with respect to statistical fluctuations in the training sample from which the tree structure is derived. An example of this is if you had two input variables with very similar separation power. In this case, a fluctuation in the training sample can cause the tree to decide to split on a particular variable, while the other variable could have remained unaffected by this fluctuation. In such an example, the whole tree tree structure below the node in question is altered, possibly leading to a very different classifier response in the tree. To avoid this issue we need to construct a way in which small fluctuations in the training sample will not have a large effect on the resultant response. To overcome this problem we construct a forest of decision trees in which we classify an event using a majority vote of the classifications done by each tree in the forest.

In addition to creating a forest, each event is subjected to a boosting procedure while training. The boosting algorithm we employ is Adaptive Boost (AdaBoost) [100]. AdaBoost works by giving events that were misclassified during the training of a decision tree a higher event weight in the subsequent training tree. Starting with the original event weights (in our case 1) for the first decision tree, each tree is trained using an event sample with modified weights by multiplying the previous event weight by a common boost weight α. The boost weight is derived from the mis-classification rate (err) of the previous tree

$$\alpha = \frac{1 - err}{err}. \tag{7.1}$$

The weights of the entire event sample are then renormalized such that the sum of weights remains constant.

Using this boost we can assign a boost event classification, $y_{Boost}(\mathbf{x})$, where $(\mathbf{x})$ represents the group of input variables. Additionally, we define a single event classifier as h($\mathbf{x}$), with h($\mathbf{x}$) = +1 as signal and h($\mathbf{x}$) = -1 as background. Combining this we get

$$y_{Boost}(\mathbf{x}) = \frac{1}{N_{collection}} \cdot \sum_{i}^{N_{collection}} ln(\alpha_i) \cdot h_i(\mathbf{x}), \tag{7.2}$$

where $N_{collection}$ is the number of trees in the forest. This results in a classifier in which small (large) values of $y_{Boost}(\mathbf{x})$ indicate events that are more background (signal) like.

AdaBoost works well on trees with weak classifiers, specifically small individual trees with depths as short as 2 or 3 levels. Trees such as this have little discrimination power on their own but are much less likely to be overtrained, and as a group their performance is enhanced. Another way to enhance performance is to force the learning rate of the trees to be slow. This allows for a larger number of boost steps, and is accomplished by using a boost weight exponential parameter. This is achieved by letting $\alpha \to \alpha^{\beta}$, where β is boost weight exponent.

The training of a decision tree is the process that defines the splitting criteria for each node. Each time, training begins at the root node that contains the entire training sample, and an initial splitting criterion for that sample is determined. This split results in two subsets of training events that each undergo the same algorithm to determine the next splitting iteration. This process is repeated until the entire tree is built. At each node, the split value is determined by finding the best separation between signal and background that can be gained with a single cut on a single variable. Each level of nodes, beyond the root node, adds a layer of depth to the tree. The splitting of nodes then continues until the maximum depth allowed that the user has specified, or until the node does not contain enough events left to split again. This minimum number of events is also specified by the user.

Each leaf node (a final node, or node that is not subsequently split) is classified as a signal or background node depending on the purity value of that leaf. Purity is calculated as $\frac{S}{S+B}$, with values > 0.5 classified as signal nodes, and values < 0.5 classified as background nodes. The separation value used to assess the performance of a variable with a specific cut is known as the Gini Index, defined as $p \cdot (1-p)$, where p is the purity already defined. This has a maximum when the samples are fully mixed, and falls off to zero when the sample consists of only one class of event. This is important, as a cut that selects primarily for background events is just as important as one that selects for signal events.

Each split of a node is defined as a single cut on a single variable, where the training procedure selects the variable and cut value that optimizes the increase in separation index between the parent node and the sum of the indices of the two daughter nodes (weighted by their relative fraction of events). The cut values are chosen by scanning over the variable range with a user specified granularity. The granularity must be large enough to allow for many cut options, but not so large that the computing time taken to scan the region becomes unmanageable.

In principle, when creating a decision tree, the node splitting process could continue until each leaf contained only a single signal or background event. With boosted decision trees, as I mentioned above, we never approach this limit as a possibility due to limitations on the depth and the relatively large number of minimum events we require for a node to split. This is important to note as allowing nodes with too few events can result in overtraining. Overtraining is a bias in the BDT discriminant response by overcontstraining the sample. In this case a small fluctuation in the input variable distribution would lead to incorrect classification of events. An example of this is the theoretical limit I described above with only one event in each leaf. Such a tree would imply that there are choices that lead to perfect signal and background identification, but this is not the case.

To avoid overtraining we split our simulated signal and background events in half, using one half of the events for training and the other half to test the classification response of the BDT algorithm. The figure of merit we use to quantify overtraining is the Kolomogrov-Smirnoff test, which computes the probability that two distributions have been sampled from the same underlying probability distribution. The results of the training and testing for each of the jet categories are described in table 7.3 and shown in section 7.4.

BDTs allow you to specify many of the parameters that control the growth of the tree and the method of boosting. For each of our categories we optimized the BDT parameters, which are defined here:

- NTrees: the number of trees generated in training that are part of the forest.
- nEventsMin: the minimum number of events allowed in a leaf node allowed after splitting.
- MaxDepth: the maximum number of node levels allowed (not including the root node).
- BoostType: the method of boosting used. We use Adaptive Boost (AdaBoost).
- AdaBoostBeta: the exponent of the AdaBoost weight value used to control BDT learning. We use a value of $\beta = 0.5$.
- SeparationType: The algorithm used to measure separation of signal and background. We use the Gini Index.
- nCuts: the Number of grid points in variable range used in finding optimal cut in node splitting. We use value of nCuts = 20, as finer stepping values did not increase noticeably the performance of the BDTs.
- PruneMethod: no pruning is necessary for our trees as we are using a boosted procedure that already limits the depth of the trees.
- NodePurityLimit: nodes with purity > NodePurityLimit are signal. We use a purity limit of 0.5.

7.2 Kinematic Variable Selection and Definition

As I described in section 7.1, BDTs work by making individual cuts on variables that help to separate the known signal and background samples that are provided. In order to do this, many input variables are used to provide distributions of known signal(s) and background(s). In this analysis there is one dominant background and a group of contributing signal processes (as shown in table 6.16 of section 6.5). In order to properly train the BDT, we must then provide it with distributions representing the signal and background we expect in our data sample.

To provide useful information to the BDT, we want to give it input distributions in which the signal and background act differently. To find these, we began by comparing the normalized distributions of W+jets MC to that of a combined $H \to WW$ signal MC. Looking at the distributions by eye, we were able to quickly pick out a few of the variables whose kinematic shape differed between signal and background. In order to quantify this difference though, two different Figures of Merit (FOMs) were calculated for each distribution. These are described in equation 7.3 and 7.4

$$FOM1 = \sum_{i=1}^{nBins} (Signal - Background)^2 \tag{7.3}$$

$$FOM2 = \sum_{i=1}^{nBins} \frac{(Signal - Background)^2}{(Signal + Background)^2}, \tag{7.4}$$

where i denotes a single bin in the distribution. These values are quite small by virtue of using normalized histograms, so were multiplied by 10^5 for ease of reading in the plots. Figure 7.2 shows a few of these distributions along with their FOMs.

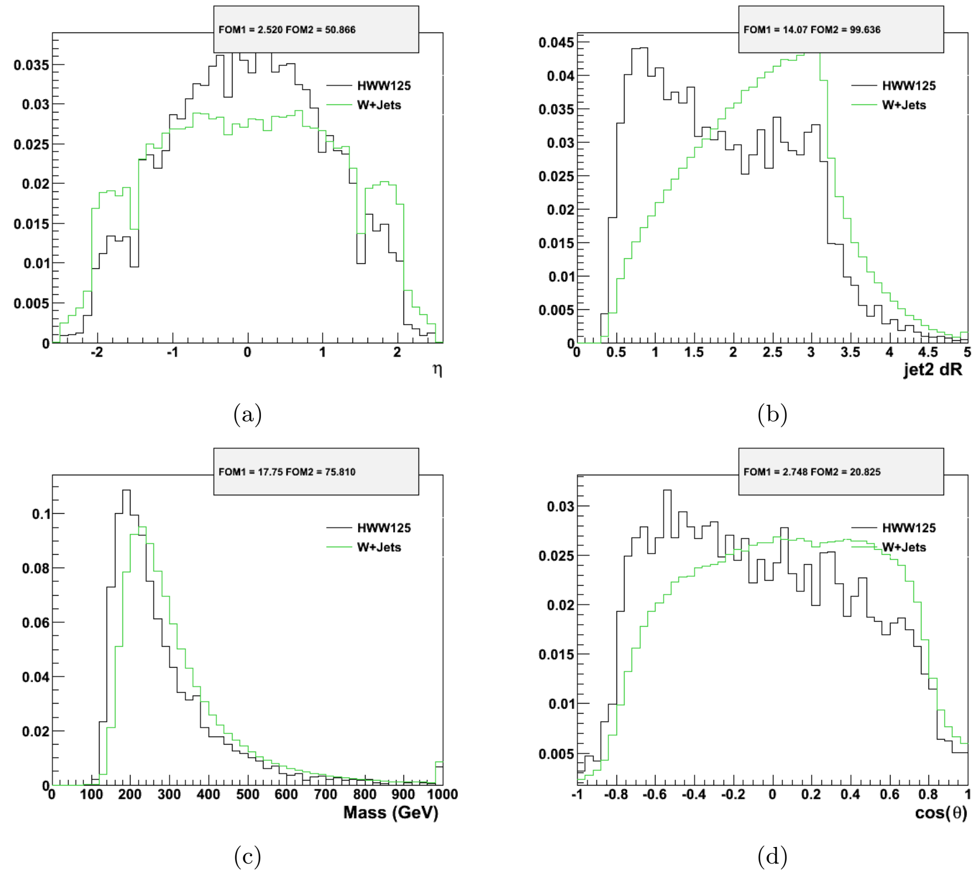


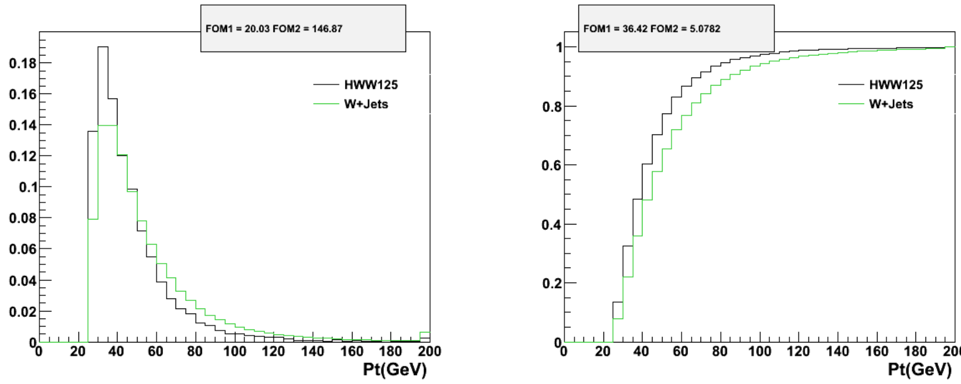
Figure 7.2: Normalized histograms of ggH M125 signal (black) and W+jets (green). Also shown are the two FOMs calculated from the distribution. The variables shown are lepton η (a), ΔR(lep,jet2) (b) , M_{lvjj} (c), and Cos(Θ_{lep}) (d).

Additionally, for every kinematic variable a Cumulative Distribution Function (CDF) was generated in order to give another way to discriminate between the signal and background distribution. CDFs were built from each of the variable distributions by filling a new histogram bin by bin, setting the bin contents equal to sum of all bins before it in the nominal distribution:

$$(CDF\ Bin)_i = \sum_{0}^{i} (Nominal\ Dist\ Bin)_i \tag{7.5}$$

To illustrate this, figure 7.3 shows the normal distribution for one of our variables (lepton p_{T})

and its corresponding PDF By looking for distributions that have the maximum area between the signal and background CDF curves we can identify the variables that have maximal differences. For this reason, we again calculate FOM1 and FOM2 for each CDF distribution. Using the results of these 4 calculations, all of the potential input variables were assigned a ranking based on the two FOMs, and the top 20 highest ranked variables were identified in each jet bin. To achieve a final ranking, an average of the ranking from each of the 4 ranking options was used. This helped to reduce any bias that one method had over another.



(a) (b)

Figure 7.3: Nominal input histogram (a) for lepton p_{T} showing signal (black) and background (green), and the corresponding CDF (b) for lepton p_{T}

.

Inputs chosen for the individual jet bin BDTs will be discussed in 7.3. A list of all of the variables considered are shown in table 7.1 below. In each variable name, 'lep' refers to either the electron or muon in the event, while any jet number refers to jet leading in p_{T}. Thus, the jet in each event with the highest p_{T} is always known as 'jet1', the jet with the second highest p_{T} is 'jet2', and so on. Additional definitions follow below:

- $P_{T_{lep}}$: the p_{T} of the single lepton in the event.
- m_T: the traverse mass of the leptonic W.
- ΔR(lep,jet1): the ΔR between the lepton and the leading p_{T} jet where $\Delta \mathrm{R} = \sqrt{\Delta\Phi^2 + \Delta\eta^2}$.
- Ht: the scalar sum of the lepton p_{T} and E_{T} of all jets in the event.
- M_{lvjj}: the 4-body mass derived by combining the 4-vectors of the lepton, met, and two leading jets in the event.
- $P_{T_{lnujj}}$: the p_{T} of the reconstructed 4 body system.
- ΔR(lep,jj):the ΔR as defined above between the lepton the the di-jet system (consisting of the 2 leading jets).

- $\Delta\phi$(met,jet): the $\Delta\phi$ between the $\not\!\!E_T$ and the leading jet.
- $\Delta\phi$(jet,jet): the $\Delta\phi$ between the two leading jets.
- min$\Delta\phi$(l,j): the smallest value of $\Delta\phi$ between the lepton and any single jet in the event.
- η: the η of the lepton, or any jet in the event.
- ϕ: the ϕ of the lepton, or any jet in the event.
- Charge: the charge of the lepton.
- $\Delta\eta$(jet,jet): the $\Delta\eta$ between the two leading jets.
- CSV$_{discr}$(jet1): the value of the b-tag CSV discriminant for the leading jet.
- Met: the $\not\!\!E_T$ of the event.

Table 7.1: List of Variables Considered for MVA

VarName	VarName
Cos($\Delta\Phi_{WH}$)	Cos($\Delta\Phi_{WW}$)
Cos(Θ_{jet})	Cos(Θ_{lep})
Cos(Θ_{WH})	$\Delta\eta$(jet,jet)
$\Delta\phi$(jet,jet)	$\Delta\phi$(met,jet)
$\Delta\phi$(met,lep)	ΔR(lep,jj)
η(jet,jet)	ht
CSV$_{discr}$(jet1)	CSV$_{discr}$(jet2)
ΔR(lep,jet1)	ΔR(lep,jet2)
η_{jet1}	η_{jet2}
ϕ_{jet1}	ϕ_{jet2}
jet1 P_T	jet2 P_T
ΔR(lep,jet3)	ΔR(lep,jet4)
Charge$_{lep}$	η_{lep}
Charge$\times\eta_{lep}$	$P_{T_{lep}}$
Met	ϕ_{met}
min$\Delta\phi$(lep,jet)	min$\Delta\phi$(met,jet)
$M_{(jet,jet)}$	$M_{l\nu jj}$
m_T	nBTags$_{CSVm}$
nJets	nLowJets
nPV	$P_{T_{lnujj}}$
ΣJet$_{E_T}$	$P_{T_{jet,jet}}$

Finally, there are angular variables that define the kinematics of the decay for our signal, $H \to WW \to l\nu jj$. These are shown in figure 7.4. To describe the Higgs decay we use information about the decaying daughter particles to reconstruct the event. The angular relation of these particles [101] gives us useful relations to help distinguish signal and background events. The invariant mass of the leptonic W, $m_{l\nu}$, is constrained in a kinematic fit to compute the longitudinal momentum of the neutrino. The angular variables are correlated and defined in 7.4. The angle θ^* is the polar angle between the parton collision axis z and the X decay axis z' as defined in the rest frame of particle X. The angle Φ_1 is the azimuthal angle between the zz' plane and the decay plane of hadronic W.

The angle Φ is the angle between the decay planes of the two W systems in the X rest frame. The angle θ_2 is the angle between the direction of the lepton from the leptonically decaying W and the axis denoting the direction normal to the WW system rest frame. The angle θ_1 is analogous to θ_2 except that it refers to the hadronic W, and it is ambiguous as to which jet is originating from the fermion anti-fermion. As a results the angle is defined from 0 to π for the leading p_{T} jet.

For ease of use in the analysis we have taken the cosine of these angles and named them such that they are easily identifiable. Thus, Φ corresponds to Cos(dΦ_{WW}), Φ_1 corresponds to Cos(dΦ_{WH}), θ_1 corresponds to Cos(θ_j), θ_2 corresponds to Cos(θ_l), and θ^* corresponds to Cos(θ_{WH}). For this analysis all angles are calculated for each event, but the use of an individual angle is based on performance in the BDT itself.

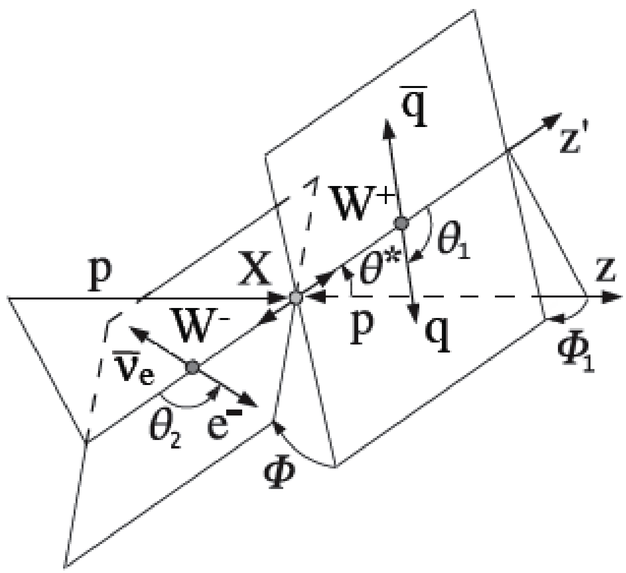


Figure 7.4: Defining the angular variables in a $H \to WW \to l\nu jj$ decay process. Reprinted from [23]

7.3 BDT Input Optimization

For each jet bin in our analysis (2,3,$\geq$ 4 jets) we identified the variables with the best discrimination power using the procedure described in section 7.2. The following procedure was followed to train and optimize the BDT in each jet bin, so I will describe the procedure using examples from just one BDT training, and then show the results for all.

As I have shown, W+jets is the clearly dominant background, so we are able to train the BDT using only W+jets MC as a background input. This simplifies the selection and training, and is a good representation of our expected background. To train our signal we used a combination of all of the $H \to WW$ samples normalized to their respective expected yields. This choice was made because each of the $H \to WW$ samples provides a noticeable impact in the expected yields, and also the addition of extra MC events benefits the BDT training. When training a BDT, absolute normalization of the samples does not matter, so all of the signals were normalized to the $ggH \to WW$ sample so they would be represented in their proper expected fractions in the training. These values used to scale the signal inputs are shown in table 7.2. As W+jets was the only MC sample used as background, a global scale factor of 1.0 was used for all jet bins.

Process	== 2	== 3	≥4
ggH, H→WW $M_H = 125$	1.0	1.0	1.0
qqH, H→WW $M_H = 125$	0.195	0.248	0.239
WH_ZH_TTH, H→WW M_H125	0.256	0.416	0.608

Table 7.2: Global scale factors for BDT signal inputs, by jet bin.

Using the 20 best kinematic inputs ranked above, we trained the BDT. After training, the BDT output is checked for evidence of overtraining and correlation amongst the input variables. Samples with very low Kolomogrov-Smirnoff test values were rejected as overtrained. Very low K-S values indicate that it is very unlikely that the training and test sample came from the same underlying distribution (which we know they do), and thus show that the trained BDT will not give consistent results. Once a BDT is trained, the input variables are ranked by order of importance by the BDT itself. This is done by showing which variables were used the most to distinguish signal from background. We then removed the two variables ranked the lowest, and retrained the BDT with a smaller set of input variables. Examples of BDT output and input variable correlations can be seen in figure 7.5. In this example there are 11 variables used, the minimal correlation showing there are no redundant variables.

Once a BDT is trained it is possible to generate an ROC curve using the discriminant output. ROC stands for Receiver Operating Characteristic, and it serves to illustrate the performance of a binary classifier system as you vary the threshold. In our case, we use it to show the background rejection versus the signal acceptance for each possible cut value on the BDT discriminant. The ROC curve is a useful tool for quantifying which BDT to use, as there are many different FOMs that could be calculated from an ROC curve. We chose the minimum distance from any point on the ROC curve to the 'perfect point' of (1,1) which denotes 100% signal acceptance and 100% background rejection as our FOM. A curve that minimizes this distance therefore shows the maximum descrimination between signal and background.

To optimize the BDT, we followed the procedure outlined above of training a BDT, then removing the two lowest ranking variables. This process was continued down to ∼3 variables.

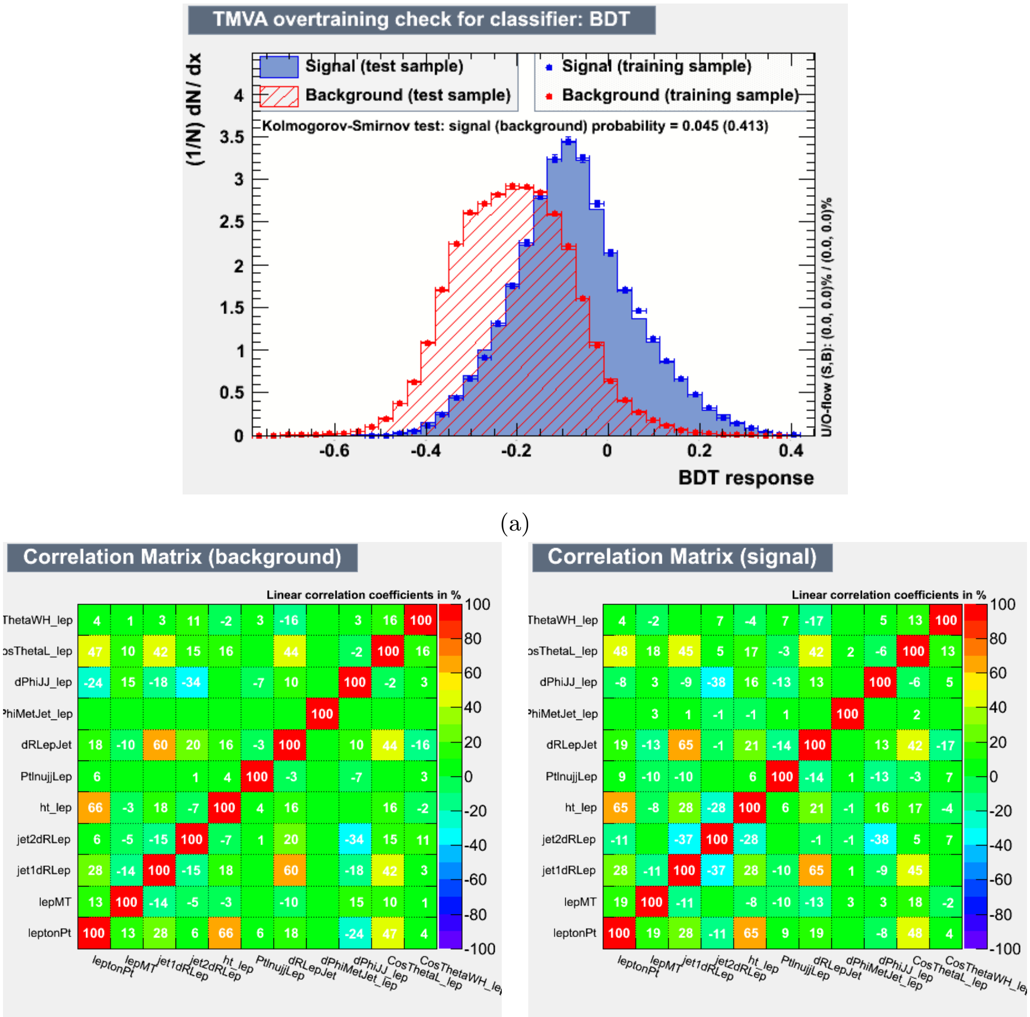
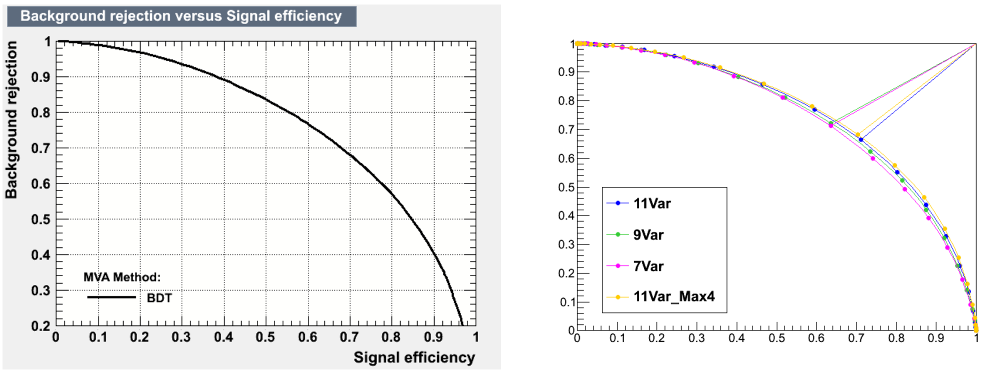


Figure 7.5: Example output plots from a BDT Training. A comparison of the training and test samples with K-S results (a), correlation matrix of background (b) and signal (c) input variables.

From each BDT discriminant we generated an ROC curve, and an example of a single curve is shown in figure 7.6a. Additionally, figure 7.6b shows the ROC output from a number of the trained BDT options overlaid with one another. As shown in this example, there is a point where reducing the number of input variables begins to degrade the efficacy of the BDT. For each jet bin this whole procedure was followed resulting in an optimized BDT setup for each one. The variables chosen by this process for each jet bin are shown in table 7.3, and were previously defined in section 7.2.



(a) (b)

Figure 7.6: ROC output from example BDT training (a) and comparison of ROCs from multiple BDT trainings of the same jet bin (b). As shown in (b), the samples with 11 variables provided the greatest discrimination power.

Table 7.3: BDT Variable choices optimized by jet bin

Var. Name	**2-Jets**	**3-Jets**	≥ 4 **Jets**
$P_{T_{lep}}$	★	✓	
Charge$\times\eta_{lep}$		✓	✓
m_T	✓		
$P_{T_{lnujj}}$	✓		
M_{lvjj}		✓	✓
ht	✓	★	★
ΔR(lep,jet1)	✓		
ΔR(lep,jet2)	✓	✓	✓
ΔR(lep,jet3)		✓	✓
ΔR(lep,jj)	✓	✓	
min$\Delta\phi$(lep,jet)		✓	
$\Delta\eta$(jet,jet)		✓	
$\Delta\phi$(met,jet)	✓	✓	✓
$\Delta\phi$(met,lep)			✓
$\Delta\phi$(jet,jet)	✓		
Cos(Θ_{lep})	✓	✓	
Cos(Θ_{WH})	✓	✓	
Cos(Θ_{jet})		✓	

List of kinematic BDT input variables for the 2, 3, and ≥ 4 jet bins.

7.3.1 BDT Parameter Optimization

Once each jet bin had an optimized BDT, we investigated the BDT control parameters which were defined at the end of section 7.1 to find the optimal working parameters for the BDT. The options tested included the number of trees used in a BDT(nTrees), the β-factor used in boosting(adaBoostBeta), the maximum depth allowed to the trees (MaxDepth), the minimum number of events allowed per node (nEventsMin), and the relative fraction of signal to background events trained on.

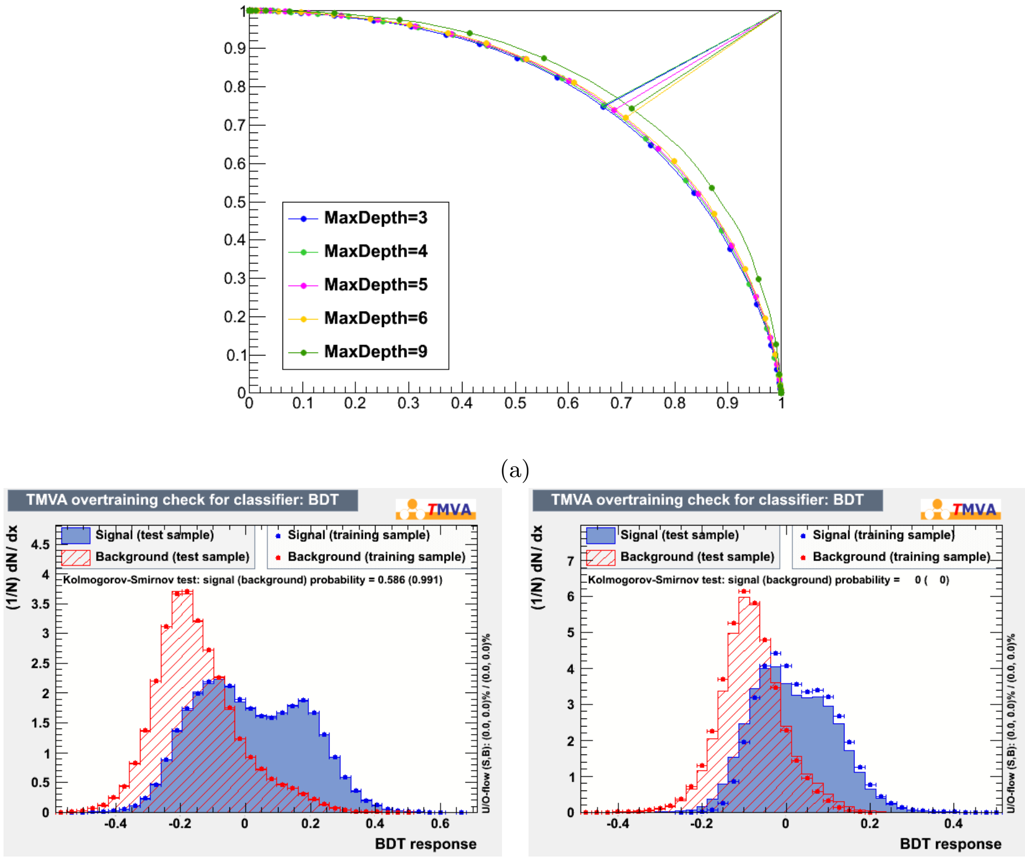
To get the best results we took our BDT that was already optimized for the number of input variables as described in table 7.3 and then proceeded to vary the control parameters one at at time to test their impact on the BDT output. The first parameter that we varied was the MaxDepth of the trees. Figure 7.7 shows the results from this. While it is clear from the ROC curve that increasing the MaxDepth improves the discrimination power, keeping track of the overtraining is a necessity as that also increases drastically. We found that the best solution was to have the largest value for MaxDepth that did not result in overtraining.

Analogous tests were performed for the boost factor (adaBoostBeta) and number of trees in the forest (nTrees) as shown in figure 7.8. It is clear in these test that the values labeled as 'default' produce the best results, with variations up or down on the initial value leading to decreased sensitivity in the BDT.

The actual values used our BDT training were nTrees = 850, adaBoostBeta = 0.5, and nEventsMin = 100. The optimal MaxDepth was BDT dependent, with a deeper tree providing more discrimination power but also much more likely to overtrain. Balancing these two issues led us to MaxDepths of 3-4. Additionally found that using the maximum number of input signal and background events that we had was best, with the samples split equally between test and training samples

7.4 BDT Input Variables: Data to Monte Carlo Comparisons

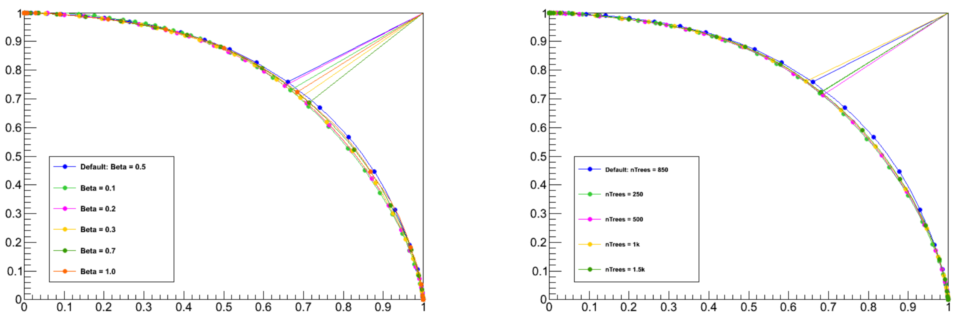
To assess the quality of the modeling provided by the MC simulation, and to ensure that the distributions we trained our BDT on accurately reflect what we see in the data, we make comparisons between the MC distributions and the data. For background, we consider all of the MC processes described in section 6.1.3: W+jets, Z+jets, $t\bar{t}$, QCD, Diboson (WW, WZ, and ZZ), and single top processes. All MC samples are scaled to the expected yield using their NLO σ, and have had all analysis cuts and MC corrections applied. Figures 7.9 and 7.10 show



(a)

(b) (c)

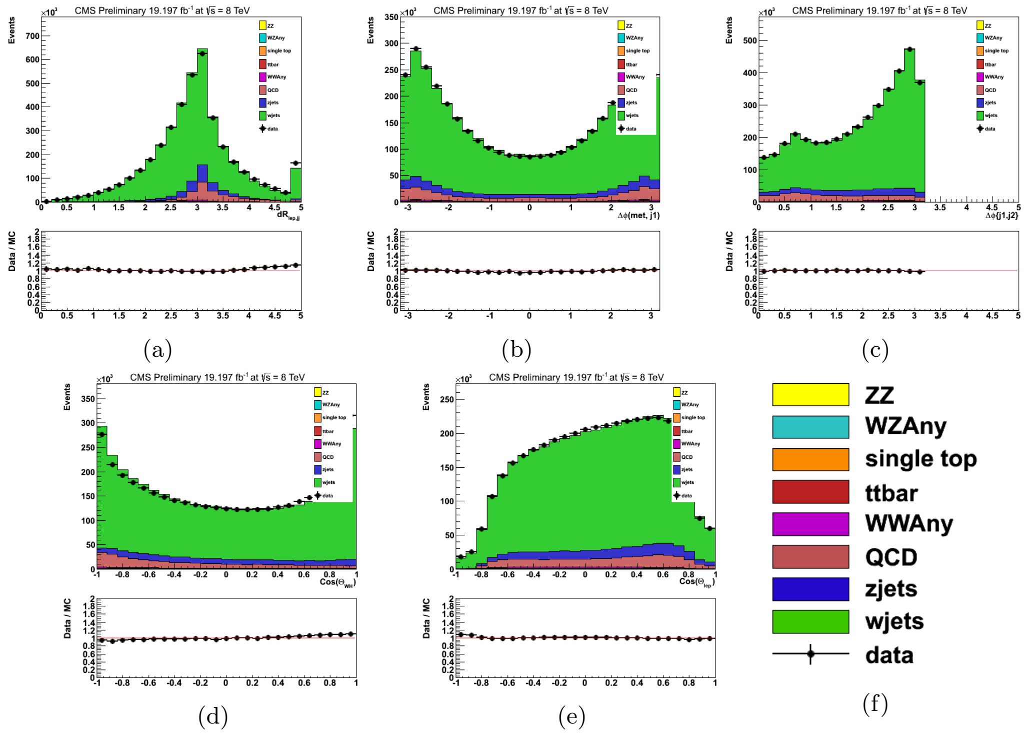
Figure 7.7: Overlay of ROC curves from a BDT trained with 5 different MaxDepth Values (a). BDTs trained with a MaxDepth of 3(b) and 9 (c). Note the K-S test results showing good agreement (values >> 1) for MaxDepth = 3 and severe overtraining (values ~ 0) for MaxDepth = 9.



(a) (b)

Figure 7.8: Overlay of ROC curves from a BDT trained with different boost factors β, and different numbers of trees. In both cases the 'nominal' value shown in dark blue performed best.

comparisons for the input variables used in the 2-jet BDT training. Figures 7.11, 7.12 and 7.13 show comparisons for the input variables used in the 3-jet BDT training. Figures 7.14 and 7.15 show comparisons for the input variables used in the $\geq$ 4-jet BDT training. We see reasonable agreement across all of the samples which gives us confidence in the qualitative aspects of the MC modeling.



(a) (b) (c) (d) (e) (f)

Figure 7.9: Data-MC comparison plots for the input variables used for the 2-jet BDT training. Shown here are ΔR(lep,jj) (a), $\Delta\phi$(met,jet) (b), $\Delta\phi$(jet,jet) (c), Cos(Θ_{WH}) (d), and Cos(Θ_{lep}) (e).

7.5 BDT Output: Data to Monte Carlo Comparisons

Using the inputs shown in section 7.4 for each jet bin, we can generate a BDT discriminant output for each one. Figures 7.16, 7.17, and 7.18 show the BDT training output and the data to MC comparison plots for the 2-jet, 3-jet, and $\geq$ 4-jet categories respectively. Also shown on these plots is the BDT signal shape (in red), which is scaled to roughly the size of the background so it can be easily seen.

7.6 Systematic Uncertainties

There are three types of systematic uncertainties leading to uncertainties considered in this analysis: uncertainties that affect the rate, shape, or rate and shape of signal or background processes. Rate uncertainties affect the number of expected events for a particular signal or

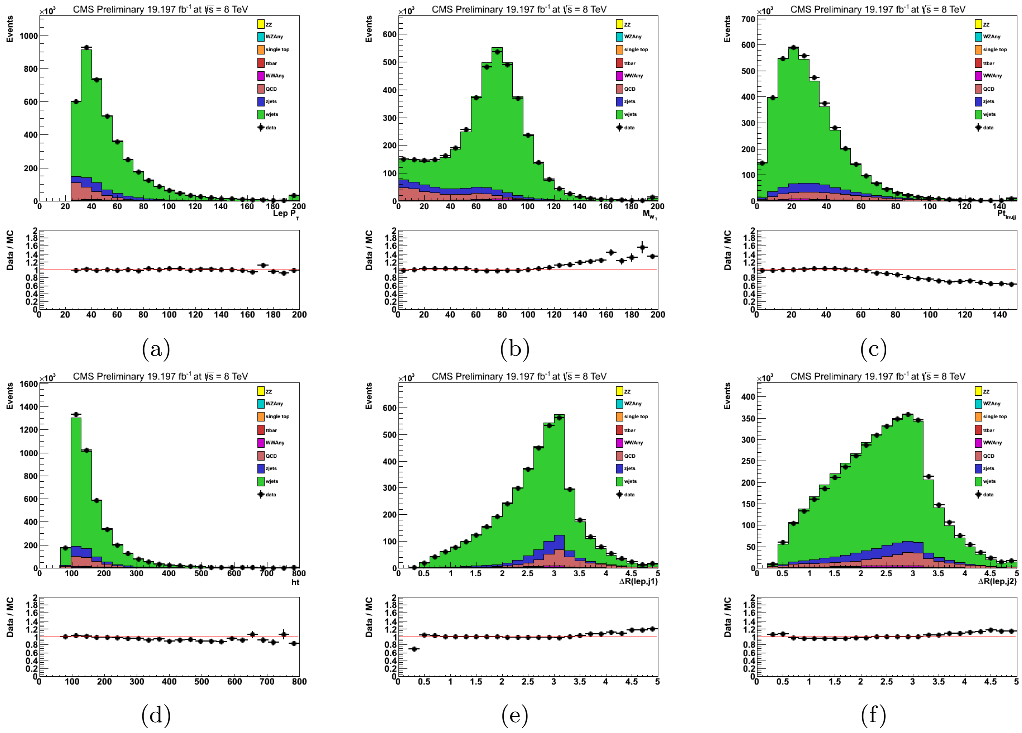


Figure 7.10: Data-MC comparison plots for the input variables used for the 2-jet BDT training. Shown here are (a), m_T (b), $P_{T_{lnujj}}$ (c), ht (d), ΔR(lep,jet1) (e), and ΔR(lep,jet2) (f).

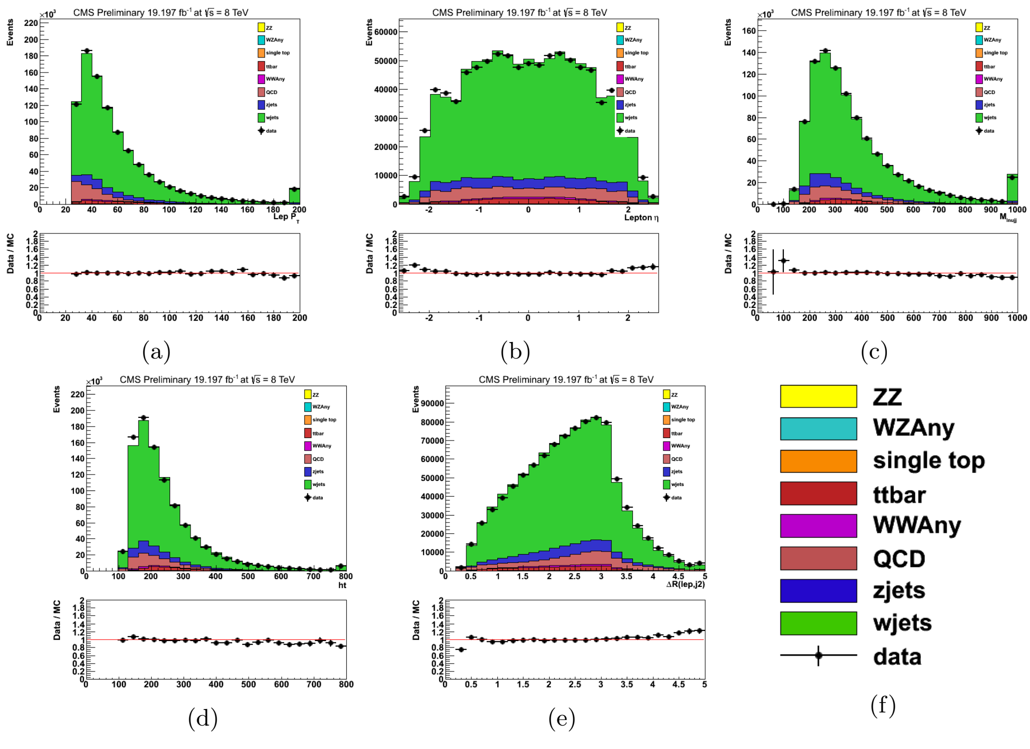


Figure 7.11: Data-MC comparison plots for the input variables used for the 3-jet BDT training. Shown here are $p_{\mathrm{T}\ lep}$ (a), Charge×η_{lep} (b), M_{lvjj} (c), ht (d), and ΔR(lep,jet2) (e).

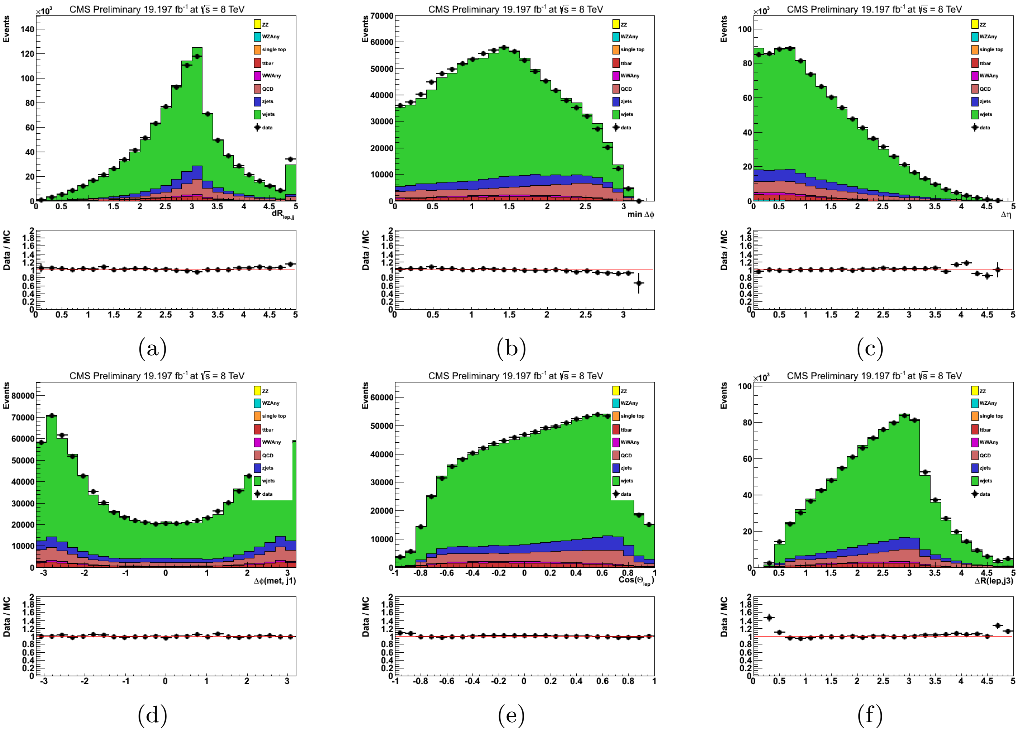


Figure 7.12: Data-MC comparison plots for the input variables used for the 3-jet BDT training. Shown here are ΔR(lep,jj) (a), minΔϕ(lep,jet) (b), Δη(jet,jet) (c), Δϕ(met,jet) (d), Cos(Θ_{lep}) (e), and ΔR(lep,jet3) (f).

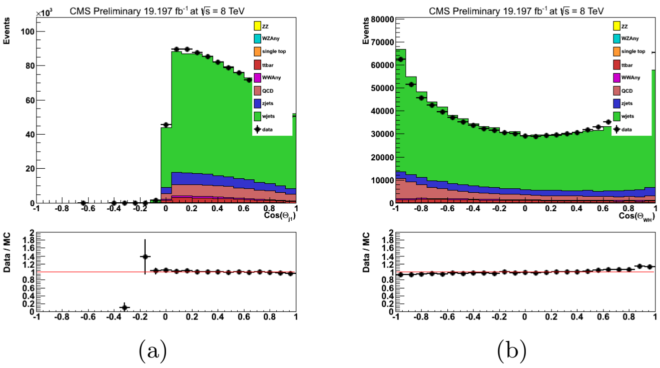


Figure 7.13: Data-MC comparison plots for the input variables used for the 3-jet BDT training. Shown here are Cos(Θ_{jet}) (a), and Cos(Θ_{WH}) (b).

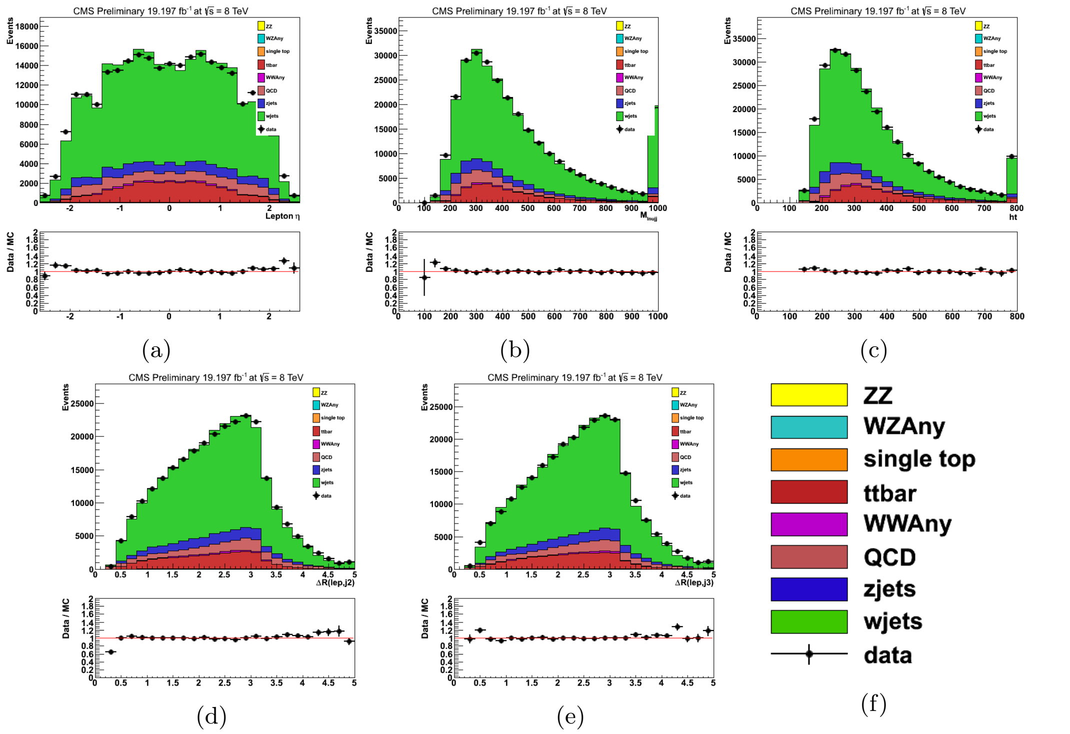


Figure 7.14: Data-MC comparison plots for the input variables used for the $\geq$ 4-jet BDT training. Shown here are Charge$\times\eta_{lep}$ (a), $M_{l\nu jj}$ (b), ht (c), ΔR(lep,jet2) (d), and ΔR(lep,jet3) (e).

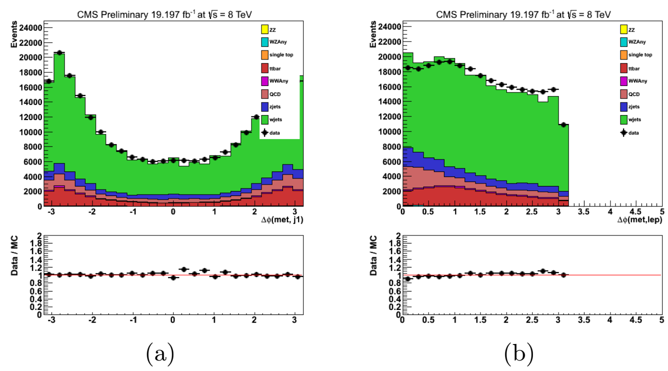
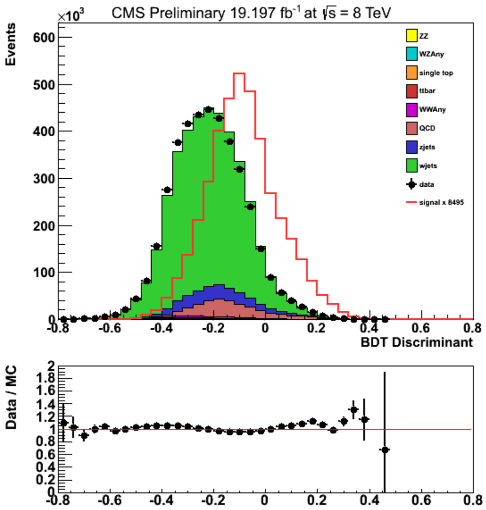
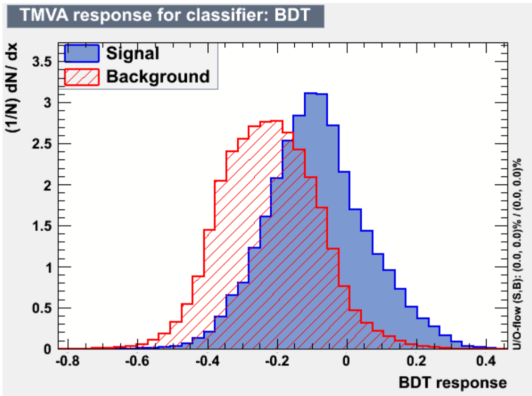


Figure 7.15: Data-MC comparison plots for the input variables used for the $\geq$ 4-jet BDT training. Shown here are $\Delta\phi$(met,jet) (a), and $\Delta\phi$(met,lep) (b).

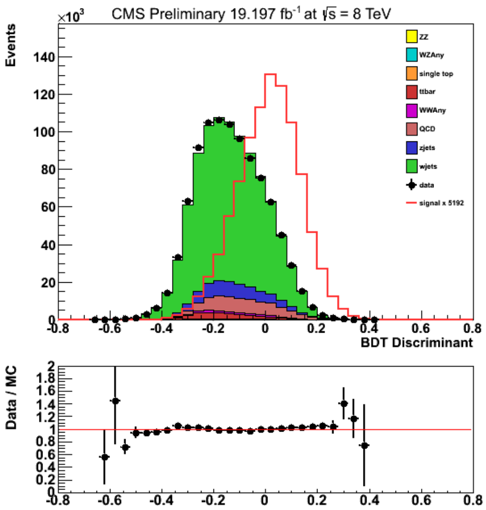


(a)

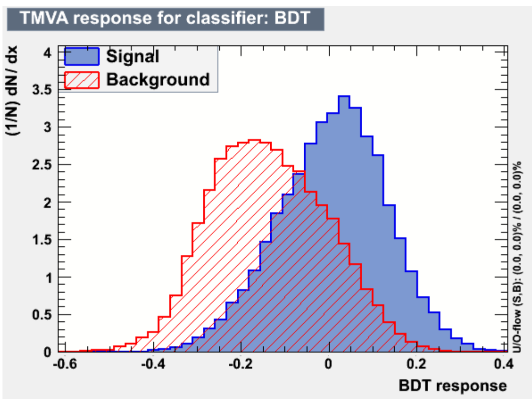


(b)

Figure 7.16: (a) Output discriminant from the training of the 2-jet BDT. (b) Data-MC comparison plot for 2-jet bin BDT discriminant.

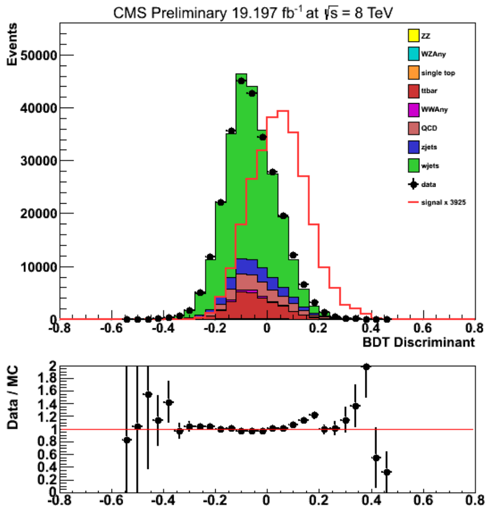


(a)

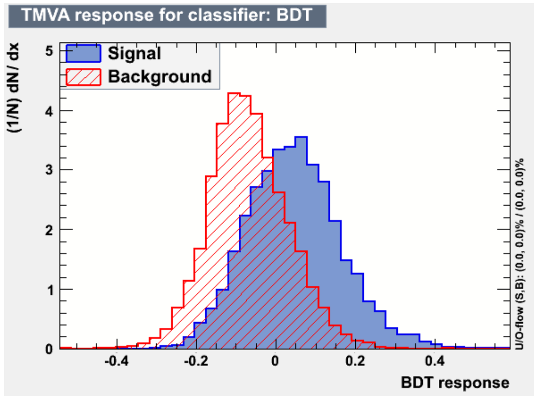


(b)

Figure 7.17: (a) Output discriminant from the training of the 3-jet BDT. (b) Data-MC comparison plot for 3-jet bin BDT discriminant.



(a)



(b)

Figure 7.18: (a) Output discriminant from the training of the $\geq$ 4-jet BDT. (b) Data-MC comparison plot for $\geq$ 4-jet bin BDT discriminant.

background process. Shape uncertainties do not affect the rate of the process, but rather the shape of the BDT output discriminant for that particular process. It is possible for an uncertainty to affect both rate and shape for a process, but for this analysis we have decoupled this effect so for uncertainties that affect both rate and shape we report a separate uncertainty for each part.

Table 7.4 shows a list of the systematic effects considered in this analysis. The first column shows the uncertainty name, with one row per line of uncertainty applied in the limit setting procedure (described in section 8.1. The next column is the rate uncertainty, followed by noting which uncertainties are shape rather than rate. As we decouple these processes, there are no lines in that table with both a rate and shape uncertainty. The column labeled 'Limit Impact' shows the effect on the limit calculation if that uncertainty is removed from the calculation. This shows the impact each uncertainty has on the analysis; uncertainties having low values here have little impact, and high values show which uncertainties our analysis is most sensitive to. The largest single source of uncertainty is from our QCD Multi-jet $|\eta|$ weights, both the rate and shape.

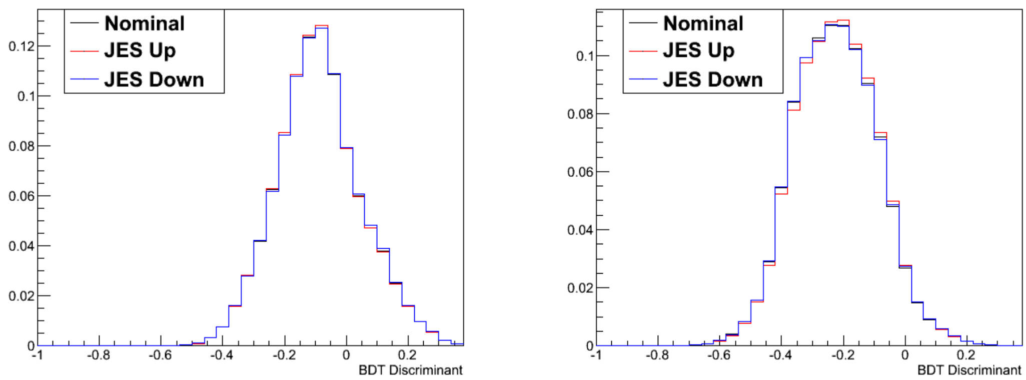
Jet Energy Scale (JES): The Jet Energy Scale systematic is based on the uncertainty on the L1, L2, L3, and L2L3 residual corrections to the reconstructed jet energy, as described in section 6.4.1. To evaluate the effect on the BDT discriminant output, the jet energy scale is shifted by one standard deviation up and down using the standard JetMET procedure [102], [103]. For each variation, the jet energies are recalculated, allowing for new jets to pass the selection where they once failed, or fail the selection where they once passed, resulting in a migration of events into and out of our selection, or across jet categories. Finally, the BDT response is recalculated, and the effect for signal and the W+jets

List of Systematic Errors

Source	Rate Uncertainty	Shape?	Limit Impact	Remarks
QCD Scale (ggH)	7-8%	No	$< 1\%$	ggH signal only
QCD Scale (qqH)	0.2%	No	$< 1\%$	qqH signal only
QCD Scale (ZH)	1%	No	$< 1\%$	ZH signal only
QCD Scale (WH)	3.1%	No	$< 1\%$	WH signal only
QCD Scale (ttH)	4-9%	No	$< 1\%$	ttH signal only
PDF (gg)	6-7%	No	$< 1\%$	ggH signal only
PDF ($q\bar{q}$)	2.6-2.8%	No	$< 1\%$	qqH signal only
QCD Scale ($t\bar{t}$)	4%	No	$< 1\%$	$t\bar{t}$ only
QCD Scale (Z+jets)	3.4%	No	$< 1\%$	Z+jets only
QCD Scale (Single t)	5%	No	$< 1\%$	All single t samples
QCD Scale (VV)	3%	No	$< 1\%$	WW, WZ, ZZ samples
Luminosity 8 TeV	2.6%	No	$< 1\%$	All samples
ME matching	-	Yes	$< 1\%$	W+jets only
Q^2scale	-	Yes	$< 1\%$	W+jets only
$\not{E}_T$	0.2%	No	$< 1\%$	All MC samples
Lepton Efficiency	2%	No	1-2%	All MC samples
puWeight	0-8%	No	$< 1\%$	All MC samples
CSV Weight	0-17%	No	$< 1\%$	All samples
Top p_{T} Weight	-	Yes	$< 1\%$	$t\bar{t}$ only
Top p_{T} Weight (Rate)	0.5-2%	No	$< 1\%$	$t\bar{t}$ only
Jet Energy Scale	-	Yes	$< 1\%$	All MC samples
Jet Energy Scale (Rate)	0-20%	No	2.3%	All MC samples
QCD Multi-jet η-Weight	-	Yes	13.6%	QCD and W+jets only
QCD Multi-jet η-Weight (Rate: QCD,W+jets)	6-30%, 0.5-1%	No	6.8%	QCD, W+jets only
CosThetaL Weight	-	Yes	2%	W+jets only

Table 7.4: Summary of the systematic uncertainties considered on the inputs to the limit calculation.

background is shown in figure 7.19, with uncertainties shown below in table 7.5.



(a) (b)

Figure 7.19: Output discriminant from the training of the 2-jet BDT and the output for samples with JES scaled up and down for ggH signal sample (a) and W+jets sample (b).

Table 7.5: Systematic Uncertainty Due to JES shift

Process	== 2	== 3	≥4
Diboson	1-2%	2%	2%
Z+jets	0-5.5%	$<1\%$	$<1\%$
$t\bar{t}$	8-19%	4-7%	2-4%
Single t	2-0%	$<1\%$	$<1\%$
ggH, H→WW $M_H = 125$	0-5%	0-2%	0-3%
qqH, H→WW $M_H = 125$	$<1\%$	4%	7%
WH_ZH_TTH, H→WW M_H125	2-3%	0-5%	5-8%
WH_ZH_TTH, H→ZZ M_H125	1.5%	0-6%	4-5%
WH, H→b$\bar{b}$ M_H125	8-9%	1-10%	2-13%
TTH, H→b$\bar{b}$ $M_H = 125$	4-17%	11-24%	18-21%

Lepton Selection and Trigger Efficiency: Systematic uncertainties on the trigger efficiencies are on the order of 1% [29]. Systematic uncertainties for lepton selection are on the order of 2%. Both of these systematic uncertainties are accounted for in our limits.

W+jets Shape Uncertainties: In order to best model the W+jets MC sample we need to account for uncertainties in Q^2 and matrix element parton matching. Samples are generated using the MADGRAPH generator which is a matrix element level generator and includes tree-level calculations for processes with multiple additional jets, matched to the PYTHIA parton shower to model additional soft and collinear radiation. Since the MADGRAPH + PYTHIA is tree-level, the choice of the renormalization and factorization scales in this calculation has a significant impact. To include the effects of this uncertainty, the factorization and renormalization scales are varied by a factor of two. As W+jets is our dominant background, we used new samples generated under conditions with Q^2 and matrix element parton matching scaled up and down from the nominal. These new samples are listed in

7.6. The new samples were subject to the same analysis cuts as all MC backgrounds and all weights were applied. For the shape uncertainty, inputs were normalized and provided to the combine tool (discussed in section 8.3. There is no rate uncertainty associated with this shift as the W+jets sample scaling is corrected from the QCD η weight fits and all rate uncertainties are taken into account there.

Sample	Dataset	Cross Sect.
ME Matching Up	/WJetsToLNu_matchingup_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	37509pb
ME Matching Down	/WJetsToLNu_matchingdown_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	37509pb
Q^2 Scale Up	/WJetsToLNu_scaleup_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v2/AODSIM	37509pb
Q^2 Scale Down	/WJetsToLNu_scaledown_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	37509pb

Table 7.6: List of samples used for W+jets systematic uncertainty shape.

Pileup Weights: The uncertainty due to event pileup are needed because MC events are simulated under an assumed pileup scenario that does not perfectly match what is seen in data. We apply weights to correct this as described in section 6.1.4. The uncertainty in the weights arises from the uncertainty on the number of pileup interactions in a particular bunch crossing:

$$N_i = \frac{\mathcal{L} \cdot \sigma_{min.bias}}{v_{orbit}} \tag{7.6}$$

where $\mathcal{L}$ is the instantaneous luminosity, $\sigma_{min.bias}$ is the cross-section of minimum bias interactions and v_{orbit} is the LHC orbit frequency (11246 Hz). Uncertainty on the pileup weight is calculated by assuming a $\pm 7\%$ shift on the $\sigma_{min.bias}$ of 69.4 mb. The shape changes produced by this are negligible and therefore are not considered. Rate uncertainties by sample are shown in table 7.7.

Table 7.7: Systematic Uncertainty Due to Pileup Weights

Process	== 2	== 3	≥4
Diboson	2-5%	3-6%	3.5-7%
W+jets	3%	4%	4%
Z+jets	7-8%	7-8%	7-8%
$t\bar{t}$	2%	2%	2%
Single t	1-3%	2-8%	2-9%
QCD Multi-Jet	0-2%	0-3%	0-4%
ggH, H→WW $M_H = 125$	2-3%	3%	3.5%
qqH, H→WW $M_H = 125$	0.5-3%	1-3.5%	2.5-4%
WH_ZH_TTH, H→WW M_H125	0-3%	1-3%	2-3.5%
WH_ZH_TTH, H→ZZ M_H125	0.5-3%	2-4%	2-4%
WH, H→b$\bar{b}$ M_H125	0.5-3%	2-4%	3.5-4.5%
TTH, H→b$\bar{b}$ $M_H = 125$	1.5-4.5%	0-2.5%	2-4%

CSV Weights: The note detailing the derivation of CSV weights describes a detailed lists of systematic uncertainties to be applied to the CSV weights [95]. For this analysis the CSV corrections are so minor that this is unnecessary. We instead overestimate the error by using the weight×weight for the σ_{up} and no weight for σ_{down}. This results in an uncertainty of 0 - 16% as shown in table 7.8.

Table 7.8: Systematic Uncertainty Due to CSV Weights

Process	== 2	== 3	≥4
Diboson	0.5-2%	1-3.5%	1-5%
W+jets	0-3%	0-5.5%	0-8.5%
Z+jets	2-5%	0-5.5%	2-5%
$t\bar{t}$	5-11%	6-14%	6-17%
Single t	4-9%	4-12%	5-16%
ggH, H→WW $M_H = 125$	1-3%	1-5%	1-7%
qqH, H→WW $M_H = 125$	0-2%	1.5-2.5%	2-4%
WH_ZH_TTH, H→WW M_H125	$< 1\%$	$< 1\%$	$< 1\%$
WH_ZH_TTH, H→ZZ M_H125	$< 1\%$	$< 1\%$	$< 1\%$
WH, H→b$\bar{b}$ M_H125	$< 1\%$	$< 1\%$	$< 1\%$
TTH, H→b$\bar{b}$ $M_H = 125$	$< 1\%$	$< 1\%$	$< 1\%$

Top p_{T}: Following the prescription of the TOP PAG [97] the uncertainty on top p_{T} reweighting is calculated by using 2× the weight for the σ_{up} and no weight for σ_{down}. This results in an uncertainty of 0.5 - 2.1% on the TTbar MC sample.

LHC Luminosity: The uncertainty on the luminosity 2.6% is applied to all MC samples [84].

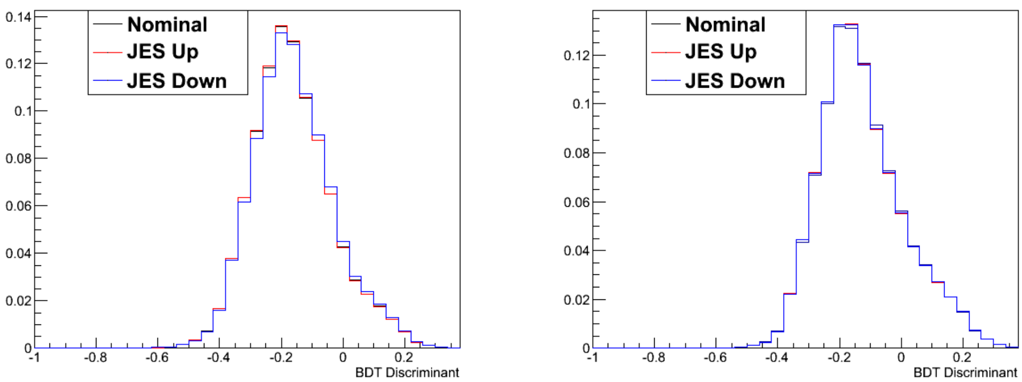
Sample Cross Sections: Cross section (σ) uncertainties for background samples were taken from CMS Standard Model calculations [104] and uncertainties on the signal samples are from the CERN yellow page Report 3 [105] with background cross sections ranging from 2-5% uncertainty and signal cross section 10-11% uncertainty. Details of the uncertainties on σ due to QCD scale and Parton Distribution Function (PDF) uncertainties are shown below in table 7.9.

Process	pdf		QCD Scale					QCD Scale			
	gg	$q\bar{q}$	ggH	qqH	WH	ZH	$t\bar{t}H$	$t\bar{t}$	V	VV	Single t
Single top											5%
Z+jets									3.4%		
Dibosons										3%	
$t\bar{t}$								4%			
ggH	7-7.5%		7-8%								
qqH		2.6-2.8%		0.2%							
WH_ZH_TTH					1%	3.1%	3.8-9%				

Table 7.9: Cross section uncertainties used for the limit settings

E_{T} Uncertainty: E_{T} directly affects our signal acceptance as we employ a hard cut of $E_{\mathrm{T}} > 25$ GeV. Using the result from the high mass $\ell^{\pm}\nu$jj group as a conservative estimate on this uncertainty we employ their 0.2% uncertainty.

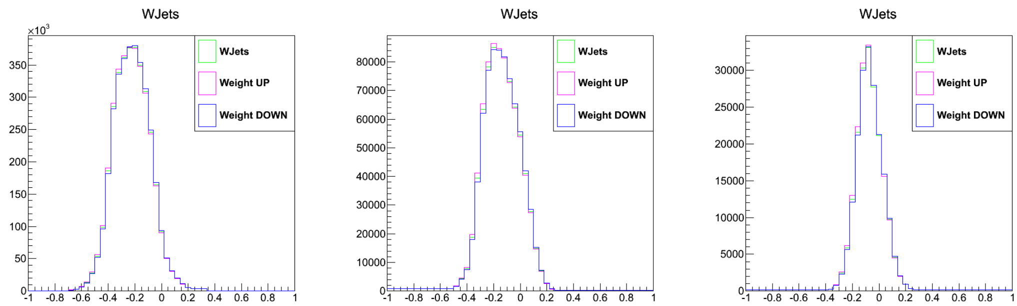
QCD η Weight Uncertainty: Uncertainty on the η weight applied to the QCD sample was generated by varying the selection criteria for our data-driven QCD sample. Our QCD sample selection was described in 6.2 as having a window cut on the pfIsolation. To generate the alternate QCD samples we relaxed one side of the window at a time and used the new selection of events as the sample with which to generate η weights from. We then followed the same procedure outlined in 6.4.7 to generate weights for a varied 'up' and 'down' QCD sample. Applying these new weights leads to a shape uncertainty on the QCD sample and a rate uncertainty for the QCD(6-30%) and W+jets (0.1-0.5%) samples. Shape variation due to QCD η weight shifts up and down in the 2-jet bin are shown in figure 7.20.



(a) (b)

Figure 7.20: Output discriminant from the training of the 2-jet BDT and the output for samples with QCD η weights scaled up and down for electron sample (a) and muon sample (b).

Cos(θ_{lep}) Weight Uncertainty: Uncertainty on the Cos(θ_{lep}) weight applied to the W+jets sample was generated by using the standard method of setting the weight uncertainty σ_{up} equal to weight*weight, and σ_{down} equal to no weight. These new weights were applied to the W+jets sample as before, and the resultant shapes used as uncertainties on the W+jets sample. As this does not change our selection, no rate uncertainty is applied here.



(a) (b) (c)

Figure 7.21: Shape uncertainty histograms for variations of the Cos(θ_{lep}) weights for (a) 2-jets, (b) 3-jets, and (c) $\geq$ 4-jets.

Chapter 8

Results

In general, the results of a search for a new physics process such as the Higgs can have two outcomes; the signal is directly observed, or, confident that the SM Higgs exists, we can set an upper limit on how much signal coul be accomodated by our data. In the case that no significant deviation from the SM predictions is seen, it is common practice to set upper limits on the Higgs production cross section in relation to its SM expectation $\sigma^{95\%}/\sigma^{SM}$. We look at two different approaches in this analysis: a counting experiment [106] using the BDT discriminant value as a variable to cut on, and a modified frequentist approach (also called CL_s) which takes advantage of the BDT shape using a binned discriminant.

8.1 Statistical Methods: Limit Setting

In high energy physics, and especially at the LHC, it is common to use the CL_s method to determine the limits on a particular production cross section. This method works by defining a likelihood function for a particular distribution, and evaluating that likelihood function for two hypotheses: signal + background, and background only. First, we define a likelihood function $\mathcal{L}(\text{data}|\mu,\theta)$, as

$$\begin{aligned}
\mathcal{L}(\text{data}|\mu,\theta) &= Poisson(\text{data}|\mu\cdot s(\theta)+b(\theta))\cdot p(\tilde{\theta}|\theta) && (8.1)\\
&= \prod_i \frac{(\mu s_i+b_i)^{n_i}}{n_i!}e^{-(\mu s_i+b_i)}\cdot p(\tilde{\theta}|\theta) && (8.2)
\end{aligned}$$

where μ is the signal strength modifier and θ represents the full suite of nuisance parameters [107]. Nuisance parameters are included in our calculation to represent the systematic uncertainties in our analysis, with one nuisance parameter per source of uncertainty. In our definition above ‘data’ represents either actual experimentally observed data or pseudo-data used to construct the sampling distributions. $s(\theta)$ and $b(\theta)$ represent the expected number of signal and background events respectively, or for a binned likelihood they are s_i and b_i.

The Probability Distribution Function (pdf) of a nuisance parameter $p\left(\tilde{\theta}|\theta\right)$, where $\tilde{\theta}$ is the default value, reflects the degree of confidence in what the true value of θ is. For rate uncertainties we use a log-normal distribution given by

$$\rho(\theta) = \frac{1}{\sqrt{2\pi}\ln\kappa}\exp\left(-\frac{(ln(\theta/\tilde{\theta}))^2}{2ln(\kappa)^2}\right)\frac{1}{\theta} \tag{8.3}$$

where κ is the parameter used to determine the width of the uncertainty, and $\tilde{\theta}$ is the nominal value of the distribution.

For shape uncertainties, a different method is needed. Uncertainties that change the shape are often due to a shift that affects selection, leading to a new set of efficiencies for the process in question. A good example of this is the jet energy scale, where shifting this scale up and down 1 σ leads to events migrating into and out of our selected sample (due to our jet cuts). We can apply this uncertainty and generate three sample distributions; nominal, and $\pm 1\sigma$. We are then faced with the problem of turning our new shapes into a continuous estimate of uncertainty in each bin. To do this, a process known as “vertical morphing” [108] is employed where the systematic is associated to a nuisance parameter taken from a unit Gaussian distribution, which is used to parametrize a quadratic interpolation for shifts below the 1σ value of a given bin, and linear interpolation for values beyond.

In order to compare the compatibility of our data with the signal + background and background only hypotheses, we construct a test statistic based on the profile likelihood ratio:

$$\tilde{q}_\mu = -2\ln\frac{\mathcal{L}(\text{data}|\mu,\hat{\theta}_\mu)}{\mathcal{L}(\text{data}|\hat{\mu},\hat{\theta})} \qquad ,0 \leq \hat{\mu} \leq \mu \tag{8.4}$$

where $\hat{\theta}_\mu$ refers to the conditional maximum likelihood estimators of θ, given the signal strength parameter μ and the data. In this calculation the signal is allowed to be scaled by μ, and ‘data’ refers to either actual experimental data or generated pseudo-data. The pair of parameter estimators $\hat{\mu}$ and $\hat{\theta}$ correspond to the global maximum of the likelihood.

The lower constraint ($0 \leq \hat{\mu}$) is dictated by physics (requiring that the signal rate be positive), while the upper constraint ($\hat{\mu} \leq \mu$) is imposed by hand in order to guarantee a one-sided confidence interval. This also results in the assumption that fluctuations in the data such that $\hat{\mu} > \mu$ are not considered evidence against the signal hypothesis; instead, such a case indicates a lack of sensitivity by the model to the signal in question.

To perform the full CL$_S$ technique, a number of calculations must be performed:

1. Calculate the observed value of the test statistic $\hat{\theta}_\mu^{obs}$ for the given signal strength modifier μ being tested.

2. Find values for the nuisance parameters $\hat{\theta}_0^{obs}$ and $\hat{\theta}_\mu^{obs}$ best describing the observed data. These are found by maximizing the values in equation 8.1 for the background-only and signal+background hypotheses respectively.

3. Generate toy MC pseudo-data to construct $pdf's$ of the $background-only$, $f(\tilde{q}_\mu|\mu,\hat{\theta}_\mu^{obs})$, and $signal{+}background$, $f(\tilde{q}_\mu|0,\hat{\theta}_\mu^{obs})$ hypotheses. For the purposes of generating a pseudo-dataset the values of $\hat{\theta}_0^{obs}$ and $\hat{\theta}_\mu^{obs}$ are fixed to the values obtained by fitting the observed data, but are allowed to float in fits needed to evaluate the test statistic.

4. Once we have constructed $f(\tilde{q}_\mu|\mu,\hat{\theta}_\mu^{obs})$ and $f(\tilde{q}_\mu|0,\hat{\theta}_\mu^{obs})$, we define two p-values that are associated with the actual observation for the signal+background and background-only hypotheses, p_μ and p_b:

$$p_\mu = P(\tilde{q}_\mu \geq \tilde{q}_\mu^{obs}|signal + background) = \int_{\tilde{q}_\mu^{obs}}^{\inf} f(\tilde{q}_\mu|\mu,\hat{\theta}_\mu^{obs})d\tilde{q}_\mu \tag{8.5}$$

$$1 - p_b = P(\tilde{q}_\mu \geq \tilde{q}_\mu^{obs}|background - only) = \int_{\tilde{q}_0^{obs}}^{\inf} f(\tilde{q}_\mu|0,\hat{\theta}_0^{obs})d\tilde{q}_\mu \tag{8.6}$$

5. CL$_s(\mu)$ is calculated as a ratio of these p-values:

$$CL_s(\mu) = \frac{p_\mu}{1 - p_b} \tag{8.7}$$

If, for $\mu = 1$, CL$_s \leq \alpha$, we would state that the signal is excluded with a (1-α)CL$_s$ confidence level. To quote the 95% upper limit on μ, $\mu^{95\%CL}$, the value of μ is adjusted until CL$_s$ = 0.05.

To calculate the expected limit using the frequentist CL$_s$ approach described above, the most straightforward approach would be to generate a large set of background-only pseudo-data and calculate CL$_s$ and $\mu^{95\%CL}$ for each of them as if they were real data. This would allow you to build a cumulative probability distribution of the results. In practice though, this is very computationally expensive so it is useful to find another method to approximate this. We use what is known as the 'asymptotic approach' with makes an analytic approximation of the full CL$_s$ technique to avoid generating so many pseudo-experiments [109]. In this approach, the $pdfs$, $f(\tilde{q}_\mu|\mu,\hat{\theta}_\mu^{obs})$, and $f(\tilde{q}_\mu|0,\hat{\theta}_\mu^{obs})$ are approximated as a falling exponential below $q_{\mu,A}$, and a Gaussian above, where $q_{\mu,A}$ is the test statistic of the Asimov dataset (background only hypothesis with nominal nuisance value parameters).

8.2 Counting Experiment Results

As shown earlier in section 7.5, the BDT algorithm provides a discriminant values between -1 and +1 for every event. Events that are more 'background-like' will have values closer to -1, and events that are more 'signal-like' will have values closer to +1. This output provides an ideal distribution in which to place a cut to separate signal from background. For a counting experiment, we look at a the yield in data for a specific cut on the events, and compare that to the expected number of signal and background events.

We want to maximize the number of signal / background events, so placing a cut near the 'signal-like' side of the BDT distribution will achieve this goal. In order to optimize the value at which to place our cut, a range of cut values from BDT discriminant = 0 to 1 were tested. The Figure Of Merit (FOM) used to choose the optimal cut was the a priori limit on the Asimov dataset (a priori limits do not depend on the observed dataset), which is more correlated to limit performance than simple calculations of signal / background. An optimal cut value of BDT discriminant > 0.24 was chosen, and yield results by jet bin are shown in table 8.1.

In the limit calculations for this analysis, the backgrounds are composed of the following categories: W+jets, Z+jets, $t\bar{t}$, diboson (WW, WZ and ZZ combined), single-top (s-channel, t-channel, and tW-channel combined), and QCD. The rates of these background processes, as well as the signal, are allowed to vary according to a set of nuisance parameters, and the values of these nuisance parameters are constrained according to the uncertainties summarized in table 7.4. Each row in that table represents a single nuisance parameter, which is assumed to be completely correlated across all categories and processes to which it applies.

For this analysis we used the information from all 3 jet bins, combining their results in one limit calculation. Table 8.2 shows the results of the Asymptotic CL$_s$ limits, with values corresponding to the 95% upper limit on μ, and the 1 and 2 σ bands shown for expected limits. As you can see from the table, while each individual jet bin does not have a large discrimination power, combining them increases the discrimination power. The results of this signal extraction technique do not reach standard model sensitivities, but the sensitivity it shows is an attestation of the power of our BDT. In a separate analysis test I performed the same counting experiment optization by using single kinematic input variables instead of the BDT discriminant output. I found that by using the BDT discriminant to choose a cut between signal and background we gain a factor of ~ 3 in sensitivity over that of a single kinematic variable (results not shown).

Event Yield (± stat. uncert.) for H→WW→lνjj after BDT > 0.24 cut

Process	== 2	== 3	≥4
Diboson	33.7 ± 5.81	7.47 ± 2.73	9.17 ± 3.00
W+jets	2291 ± 47.9	462 ± 21.5	279.6 ± 16.7
Z+jets	803 ± 28.3	79.9 ± 8.94	126.9 ± 11.3
$t\bar{t}$	28.7 ± 5.36	4.01 ± 2.00	75.9 ± 8.71
Single t	16.8 ± 4.10	4.48 ± 2.12	9.83 ± 3.14
Multi-Jet	1967 ± 44.4	372.2 ± 19.3	251.3 ± 15.9
Tot Bkg	5163 ± 71.9	944.2 ± 30.7	762.7 ± 27.6
ggH, H→WW $M_H = 125$	7.40 ± 2.72	4.55 ± 2.13	3.50 ± 1.87
qqH, H→WW $M_H = 125$	3.68 ± 1.92	2.23 ± 1.49	0.70 ± 0.84
WH_ZH_TTH, H→WW M_H125	0.35 ± 0.59	0.32 ± 0.57	0.69 ± 0.83
WH_ZH_TTH, H→ZZ M_H125	0.02 ± 0.14	0.01 ± 0.10	0.02 ± 0.14
WH, H→ $b\bar{b}$ M_H125	0.05 ± 0.22	0.01 ± 0.10	0.01 ± 0.10
TTH, H→ $b\bar{b}$ $M_H = 125$	0.00 ± 0.00	0.00 ± 0.00	0.02 ± 0.14
Total Signal	11.50 ± 3.39	7.12 ± 2.67	4.94 ± 2.22
Total Sig + Bkg	5174.8 ± 71.9	951.34 ± 30.8	767.6 ± 27.7
Total Data Events	5349	994	762
Total Expected MC / Data	0.967	0.957	1.007

Table 8.1: Shows expected event yield normalized to cross sections and luminosity for all signal and background processes after the optimized BDT discriminant cut of > 0.24. Uncertainties shown are statistical only. Data yields are also shown for direct comparison.

A Priori Limit Results with the Asymptotic Method

		Expected Limit		
Jet Bin	**Observed**	**Median**	**68% C.L. Range**	**95% C.L. Range**
2 jets	96.61	90.25	{67.66 , 122.98}	{52.70 , 161.44}
3 jets	34.66	31.37	{23.67 , 42.26}	{18.57 , 55.34}
≥ 4 jets	35.21	33.88	{24.26 , 48.05}	{18.19 , 65.87}
Combined Jet Bin	16.42	13.91	{9.89 , 19.84}	{7.36 , 27.29}

Table 8.2: Results of the expected and observed a priori limit using the Asymptotic CL_S method for a counting experiment using a cut of BDT > 0.24.

8.3 Shape-based Analysis

8.3.1 Yields and Limits Using Statistical Uncertainties

Fitting the simulated samples to the measured data will test for the presence of signal, but in the absence of measured signal we will set an upper limit on the Higgs cross section. This upper limit is reported at a confidence level of 95%, so the upper limit measured tells us the maximum amount of signal that we could see in this sample, given our selections, approach, and uncertainties. As a reminder, table 8.3 shows the results of our full selection expected yields in each signal and background category, as well as that measured in data. As shown above, a counting experiment does not have enough discrimination power to reach standard model level sensitivity. One method that can be used to improve on this is to utilize the entire shape of the BDT output distribution in order to separate signal and background. In this method, each bin in the shape histogram acts as its own counting experiment, and the results of all of those are combined in a final result.

Shape analyses benefit from more than just having additional bins in which to conduct counting experiments. As described in section 8.1, the uncertainties on each sample are taken as nuisance parameters, and the statistical method of combining each of the channels seeks the best fit for these parameters. This means that bins with very little signal can be used to constrain the backgrounds, so that in bins with higher signal content we have a much more constrained estimate for these backgrounds than previously. Thus, a shape analysis benefits from having more information about each sample, and a larger number of channels to attempt signal extraction from (3 jet channels times the number of bins in each histogram rather than just 3 jet channels).

Event Yield (± stat. uncert.) for H→WW→lνjj 19.1 fb^{-1} Ele & Mu Sample

Process	== 2	== 3	≥4
Diboson	39027 ± 198	126133 ± 112	3485 ± 59
W+jets	3417692 ± 1849	766107 ± 875	199983 ± 447
Z+jets	272587 ± 522	69589 ± 264	19937 ± 141
$t\bar{t}$	22467 ± 150	27793 ± 167	31092 ± 176
Single t	16318 ± 128	7097 ± 84.2	3037 ± 55.1
Multi-Jet	278270 ± 528	78131 ± 280	21009 ± 145
Tot Bkg	4046361 ± 2012	961327 ± 980.	278543 ± 528
ggH, H→WW $M_H = 125$	548 ± 23.4	211 ± 14.5	79.9 ± 8.9
qqH, H→WW $M_H = 125$	106 ± 10.3	52.5 ± 7.2	17.4 ± 4.2
WH_ZH_TTH, H→WW M_H125	124 ± 11.2	77.6 ± 8.81	42.9 ± 6.55
WH_ZH_TTH, H→ZZ M_H125	8.25 ± 2.87	4.38 ± 2.09	2.24 ± 1.50
WH, H→b$\bar{b}$ M_H125	40.1 ± 6.33	12.59 ± 3.55	3.38 ± 1.84
TTH, H→b$\bar{b}$ $M_H = 125$	0.53 ± 0.73	1.13 ± 1.06	3.12 ± 1.77
Total Signal	827.8 ± 28.8	359.1 ± 18.95	148.9 ± 12.20
Total Sig + Bkg	4047189 ± 2012	961686 ± 981	278692 ± 528
Total Data Events	4024809	946065	270664
Total Expected MC / Data	1.0056	1.0165	1.0297

Table 8.3: Shows expected event yield normalized to cross sections and luminosity for all signal and background processes. Uncertainties shown are statistical only. Data yields are also shown for direct comparison.

For the shape based analysis, the signals, backgrounds, and uncertainties are treated the same as for the counting experiment. The one exception is the addition of shape uncertainties described in table 7.4. Also, the binning of the BDT discriminant shape must be carefully chosen to minimize the impact of MC statistics. This is done by ensuring that each bin in the BDT discriminant shape has a statistical uncertainty for background of $\leq$ 10%. This ensures that there can be no bins with zero background events, a situation that could lead to spurious signal significance if even one event was found in that bin in the experimental dataset. To accomplish this we took the BDT discriminant output (a value restricted to be -1 < BDT discriminant < 1) and created a very finely binned histogram. Then, starting with the lowest bin, we calculated the statistical uncertainty on that bin. If the statistical uncertainty was $\leq$ 10% that bin was merged with the next bin in the distribution and the calculation was redone.

This method resulted in a relatively large number of bins in our distribution (which is desired, as more bins correlates with better discrimination power), while ensuring that we do not have any bins that will result in spurious signal significances. This method results in a variable bin widths, and different numbers of bins for each of our BDT trainings. The end result is a distribution with 27, 21, and 18 bins for the 2-jet 3-jet, and $\geq$ 4-jet categories respectively.

Using this setup for the input BDT discriminant shape histograms we calculate the a priori expected limits using the asymptotic CL_s method with the Higgs Combine Tool [110]. This is a software package that uses Roo-Stats[111] to compute the CL_s limits as described above. A statistics only uncertainty approach shows the theoretical limit of the sensitivity of the analysis, as addition of systematic uncertainties can only cause the limit to go up in value.

8.3.2 Limits Using Full Systematic Uncertainties

Using the same binning described in section 8.3 we can calculate the full asymptotic CL_S limits using the BDT discriminant as input shape, and accounting for all of our systematic uncertainties. To do this we again use the Higgs Combine Tool, providing as input three categories of information:

1. The expected number of events passing our selection criteria for each signal and background process (in each jet/lepton category).

2. For each systematic rate uncertainty, a nuisance parameter with the values described in table 7.4 is provided.

3. For each systematic shape uncertainty, two histograms defining the $\pm\sigma$ change on the BDT discriminant shape are provided.

The results of the limit calculations using the Asymptotic CL_S method for a Higgs Mass of 125 GeV/c with all of our systematic uncertainties is shown in table 8.4. All 'Expected Limit' results are reported from the Asimov dataset.

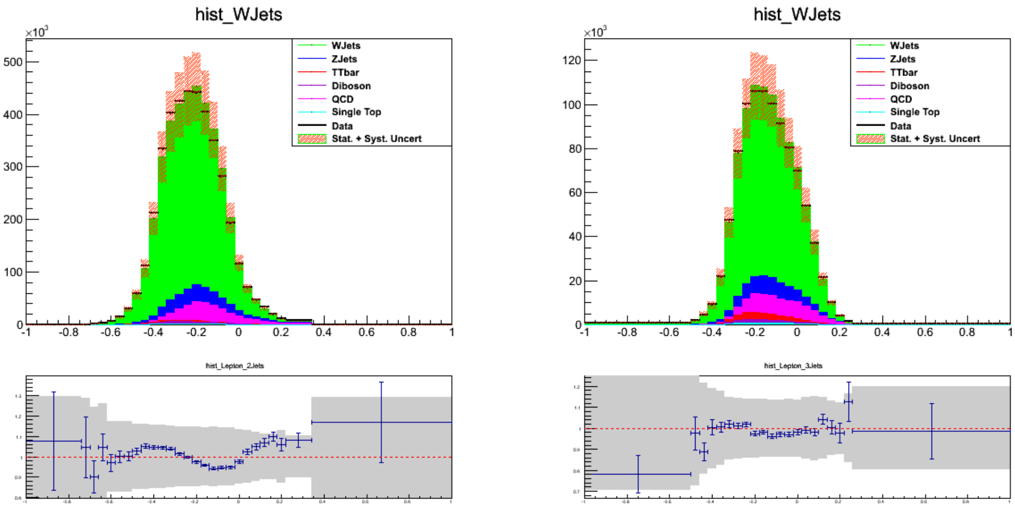
The results seen when looking at the data are quite surprising. As the results in data did not closely match (showing a deviation of $< 2\sigma$), the expected results in the 2 and 3 jet bin categories, this warranted further investigation. The first thing to look at was the distribution of the BDT discriminant, in order to see if there was any visual discrepancy. Figure 8.1a(8.1b) shows a plot of this BDT output distribution for the 2-jet (3-jet) category. Also shown in these plots is the total uncertainty (statistical + rate + shape) for each of the background combined, shown as red hatched lines. In order to see the impact on the edge bins a ratio plot has also

A Priori Limit Results with the Asymptotic Method

		Expected Limit		
Jet Bin	**Observed**	**Median**	**68% C.L. Range**	**95% C.L. Range**
2 jets	**1.45**	13.06	{9.53 , 18.32}	{7.19 , 24.76}
3 jets	**18.84**	10.03	{9.36 ,15.10 }	{7.49 , 19.10}
$\geq$ 4 jets	**19.00**	19.97	{14.36 , 27.85}	{10.88 , 37.22}
Combined Jet Bin	**8.86**	4.98	{3.58 , 7.03}	{2.69 , 9.54}

Table 8.4: Results of the expected a priori limit using the Asymptotic CL_S method for the kinematic BDT with all systematic uncertainties included. Expected results use the Asimov dataset hypothesis for calculation.

been included, which shows the values for data / simulation for each bin. The gray band shown here is the extent of the uncertainties from the backgrounds.



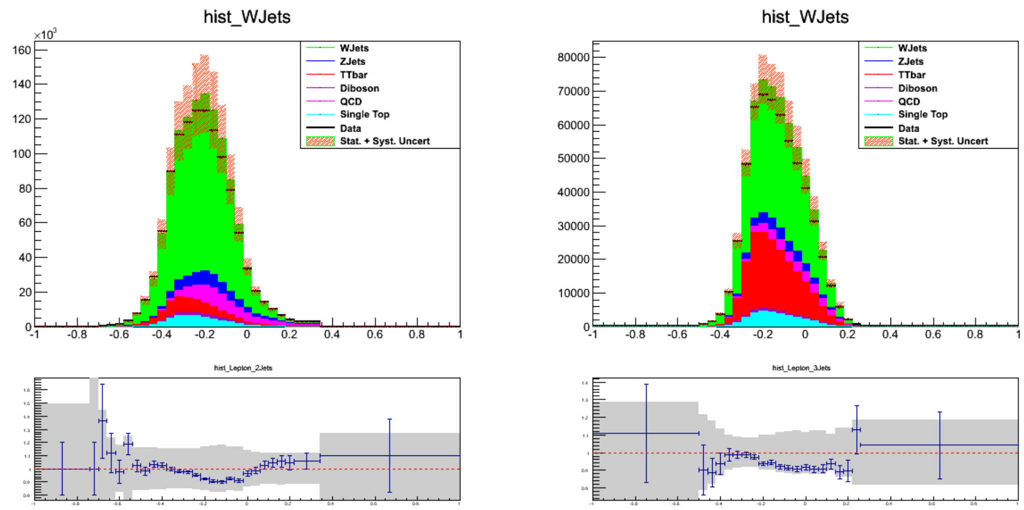
(a) 2-jet bin BDT Output

(b) 3-jet bin BDT Output

Figure 8.1: BDT Discriminant value distribution for the 2-jet (a) and 3-jet(b) categories. Each plot shows the full BDT distribution (top) with red hatched lines showing total uncertainty on backgrounds, and a ratio of data / simulation (bottom) with uncertainty shown as a gray band.

Figure 8.1 clearly shows that the data fall within our uncertainty bands, though the 2-jet bin plot shows a modulation that could be described by a shift in the BDT peak. The uncertainty in the shape of the MC backgrounds should cover this difference, as shown in the plots, but the results from the data indicate that while the raw uncertainty on each bin is sufficient, the shape difference is not accounted for.

In an effort to understand this many different methods were used. An example of one such method was to use a control region to see if this same shape difference was seen between data and simulation there. We used events with 1 b-tag in order to analyze this. Figure 8.2 shows the analogous plots in this control region to the plots shown in 8.1. The lower statistics in the control region resulted in insufficient statistics in the edge bins to generate a proper comparison of Data and simulation, so the ratios were set to 1 for these few bins, and the uncertainty was set at 100%. Figure 8.2 shows that the relationship between data and MC is seen in our control region is very similar to that seen in the signal region. This indicates that there might be an additional shape systematic uncertainty that needs to be taken into account between data and MC.



(a) 2-jet bin BDT Output

(b) 3-jet bin BDT Output

Figure 8.2: BDT Discriminant value distribution for the 2-jet (a) and 3-jet(b) categories in the b-tag control region. Each plot shows the full BDT distribution (top) with red hatched lines showing total uncertainty on backgrounds, and a ratio of data / simulation (bottom) with uncertainty shown as a gray band.

The control region was used to generate this shape uncertainty, which was then applied to all of the MC backgrounds. Although the shape appears similar, the addition of this uncertainty did not help to reconcile the effect seen in the fits. Many other control regions were also tested, but nothing that was tried served to explain the difference in shape between data and simulation.

Given unlimited time to address this matter I'm confident that the source of this uncertainty could be fully understood, but given the constraints of this analysis I must accept what I have measured. There is a real difference seen between the data and MC that is evident in our BDT output distributions. This difference was not seen in the comparisons of data and MC for individual kinematic variables, nor in two dimensional plots of the input variables and the BDT discriminant. As a result, I must conclude that some combination of the kinematic variables that are used in the BDT combine to create this phenomenon. While the individual jet bins do not have enough power individually to constrain this difference, by combining them together we are able to constrain each parameter further and see a result in data that agrees within 2σ to that expected from simulation.

8.4 Summary of Results

Using the entire 19.1 fb^{-1} of data collected at $8TeV$ no direct observation of the Higgs was seen in the $H \to WW \to l\nu jj$ decay channel. Due to the large amount of background, while Higgs events certainly exist in our data, we do not achieve the sensitivity needed to discriminate it from our backgrounds. Thus, in the absence of a significant excess of events in data indicative of our signal, we can set upper limits on the production rate of $H \to WW \to l\nu jj$. Two methods of setting limits were employed using the information from our trained and optimized Boosted Decision Tree (BDT)s. By placing a cut on the BDT discriminant output we were able to set an upper limit on the production cross section of 16.42, using the statistical methods described above. From simulations alone the expected factor was 13.91, a difference of less than 1-σ from the observed value.

Using the full BDT output shape we were also able to set limits on the production cross section. As noted above, a large uncertainty in the shape between data and simulation produced some curious results. By using the combined information in the 2, 3, and ≥ 4 jet bin shapes, this uncertainty was better constrained. Using this method an upper limit of 8.86 times the production cross section is measured, which falls within 2-$sigma$ of the expected value of 4.98 seen from simulations alone.

Chapter 9

Conclusion / Summary

The results for a search for the Higgs Boson in the H→WW→lνjj in pp collisions at $\sqrt{s} = 8$ TeV center of mass energy have been presented. This analysis begins with the production of protons in the LHC accelerator complex, traveling through many complex systems on their way to a collision at $\sqrt{s} = 8$ TeV at the center of the CMS detector. The superior tracking and reconstruction of particles in CMS led to over 19 fb^{-1} of data collected in 2012 that was used in this analysis.

Once collected, a search was performed for our signal in a final state that included one isolated lepton, one neutrino (indicated by $\not\!\!E_T$), and two jets. We further required that the jets not be b-tagged, restricting our sample to light flavor jets that are more common from a W decay. The search region was divided into categories based on the number of jets in the event, using categories of 2, 3, or $\geq$4 jets. For each category we trained a Boosted Decision Tree (BDT) by using kinematic variables as inputs, with each category optimized for maximum signal extraction potential.

We looked at two methods for signal extraction, a counting experiment that took advantage of the BDT by using it as a superior discrimination variable to cut away background, and as a shape analysis using the entire BDT output shape to separate signal from background. No significant excess was seen seen using either method, so an upper limit on the production cross section was placed. Using a counting experiment we set a limit of 16.4 times the standard model, and using shape based signal extraction we were able to lower this limit to 8.86.

Though this analysis did not have the sensitivity to observe the Higgs directly, I am optimistic that in the future the increase of data will make this possible. With Run II at the LHC just beginning, and an increase in the Higgs production cross section at $\sqrt{s} = 13$ TeV, there will definitely be more signal out there to find. The increase in luminosity and pileup will require new and unique ways to reduce the backgrounds seen in this channel, but though careful background modeling I believe it's possible.

Finally, combining this analysis with others looking for the same final state could increase the sensitivity. Use of Matrix Element values for particle production could serve as a good addition to the kinematic information of the event, producing a result more sensitive to probing the limits of the Standard Model.

Bibliography

[1] "Image, The Standard Model of Elementary Particles". URL http://www.quantumdiaries.org/wp-content/uploads/2014/03/2000px-Standard_Model_of_Elementary_Particles.svg_.jpg.

[2] CDF Collaboration, "Search for Standard Model Higgs Boson Production in Association with a W Boson Using a Matrix Element Technique at CDF in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV", *Phys.Rev.* **D85** (2012) 072001, doi:10.1103/PhysRevD.85.072001, arXiv:1112.4358.

[3] "LHC Underground Diagram of Experiements". URL http://www3.nd.edu/~cjessop/research/CMS/CMS.html.

[4] "LHC Aerial View". URL http://www.physics.ucr.edu/research/particle.html.

[5] "LHC, Delivered Luminosity". URL http://cms-service-lumi.web.cern.ch/cms-service-lumi/publicplots/int_lumi_cumulative_pp_2.png.

[6] L. Evans, "The Large Hadron Collider", *New Journal of Physics* **9** (2007) 335, doi:10.1088/1367-2630/9/9/335.

[7] J.-L. Caron, "The LHC injection complex.. L'ensemble d'injection du LHC.", (May, 1993). AC Collection. Legacy of AC. Pictures from 1992 to 2002.

[8] L. Evans and P. Bryant, "LHC Machine", *JINST* **3** (2008) S08001, doi:10.1088/1748-0221/3/08/S08001.

[9] M. Barnes, J. Borburgh, B. Goddard, and M. Hourican, "Injection and extraction magnets: septa", arXiv:1103.1062.

[10] J.-L. Caron, "Superconducting cavity with its cryostat. Cavit supraconductrice dans son cryostat.", (May, 1994). AC Collection. Legacy of AC. Pictures from 1992 to 2002.

[11] "Image, CMS Detector". URL https://cms-docdb.cern.ch/cgi-bin/PublicDocDB/RetrieveFile?docid=11514&version=1&filename=cms_120918_03.png.

[12] M. Schott and M. Dunford, "Review of single vector boson production in pp collisions at $\sqrt{s} = 7$ TeV", *Eur.Phys.J.* **C74** (2014) 2916, doi:10.1140/epjc/s10052-014-2916-1, arXiv:1405.1160.

[13] J. Nielsen, "Fundamentals of LHC Experiments", arXiv:1106.2516.

[14] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[15] CMS Collaboration, "CMS Tracker Detector Performance Results", cms twiki, CERN, 2014.

[16] "Image,ECAL response to laser monitoring light in CMS during 2011 and 2012". URL https://twiki.cern.ch/twiki/pub//CMSPublic/EcalDPGResultsCMSDP2012027/laser_history_2011-12.png.

[17] "Image,Russian navy shells re-used in the CMS Hadron Calorimeter". URL http://cms.web.cern.ch/sites/cms.web.cern.ch/files/styles/large/public/field/image/navy-shells-image_0095.jpg?itok=cnSm1r4Z.

[18] CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker", *JINST* **9** (2014), no. 10, P10009, doi:10.1088/1748-0221/9/10/P10009, arXiv:1405.6569.

[19] "Image,2011 data collected by early July, corresponding to an integrated luminosity of 1.1 fb-1, superposition of various dimuon trigger paths". URL https://twiki.cern.ch/twiki/pub/CMSPublic/PhysicsResultsMUO/dimuMass_2011Run_1fb_20July2011.pdf.

[20] CMS Collaboration, "Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector",.

[21] M. Cacciari, G. P. Salam, and G. Soyez, "The Anti-k(t) jet clustering algorithm", *JHEP* **0804** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[22] A. B. Galtieri, F. Margaroli, and I. Volobouev, "Precision measurements of the top quark mass from the Tevatron in the pre-LHC era", *Rept.Prog.Phys.* **75** (2012) 056201, doi:10.1088/0034-4885/75/5/056201, arXiv:1109.2163.

[23] Y. Gao et al., "Spin determination of single-produced resonances at hadron colliders", *Phys.Rev.* **D81** (2010) 075022, doi:10.1103/PhysRevD.81.075022, arXiv:1001.3396.

[24] "CERN Brochure, LHC the guide". URL http://cds.cern.ch/record/1165534/files/CERN-Brochure-2009-003-Eng.pdf.

[25] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys.Lett.* **B716** (2012) 30–61, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

[26] 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, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.

[27] CMS Collaboration, "Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV", *Eur.Phys.J.* **C75** (2015), no. 5, 212, doi:10.1140/epjc/s10052-015-3351-7, arXiv:1412.8662.

[28] "Image, CMS Higgs Physics Results,Higgs to Gamma Gamma". URL https://twiki.cern.ch/twiki/pub/CMSPublic/PhysicsResultsHIG/Fig_19.png.

[29] CMS Collaboration Collaboration, "Search for the Standard Model Higgs boson in the H to WW to lnujj decay channel in pp collisions at the LHC", Technical Report CMS-PAS-HIG-13-027, CERN, Geneva, 2012.

[30] H. Weyl, "Electron and Gravitation. 1. (In German)", *Z.Phys.* **56** (1929) 330–352, doi:10.1007/BF01339504.

[31] C. N. Yang and R. L. Mills, "Conservation of Isotopic Spin and Isotopic Gauge Invariance", *Phys. Rev.* **96** (Oct, 1954) 191–195, doi:10.1103/PhysRev.96.191.

[32] IceCube Collaboration, "Measurement of Atmospheric Neutrino Oscillations with IceCube", *Phys.Rev.Lett.* **111** (2013), no. 8, 081801, doi:10.1103/PhysRevLett.111.081801, arXiv:1305.3909.

[33] S. Dawson, "Introduction to electroweak symmetry breaking", arXiv:hep-ph/9901280.

[34] M. E. Peskin and D. V. Schroeder, "An Introduction to Quantum Field Theory". Westview Press, URL http://www.westviewpress.com, 1995.

[35] C. Quigg, "The Electroweak theory", arXiv:hep-ph/0204104.

[36] M. Gell-Mann, "A Schematic Model of Baryons and Mesons", *Phys.Lett.* **8** (1964) 214–215, doi:10.1016/S0031-9163(64)92001-3.

[37] G. Zweig, "An SU(3) model for strong interaction symmetry and its breaking. Version 1",.

[38] O. W. Greenberg, “Spin and Unitary-Spin Independence in a Paraquark Model of Baryons and Mesons”, *Phys. Rev. Lett.* **13** (Nov, 1964) 598–602, doi:10.1103/PhysRevLett.13.598.

[39] D. Griffiths, “Introduction to Elementary Particles”. John Wiley & Sons, New York, USA, 1987.

[40] UA1 Collaboration, “Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at s**(1/2) = 540-GeV”, *Phys.Lett.* **B122** (1983) 103–116, doi:10.1016/0370-2693(83)91177-2.

[41] UA1 Collaboration, “Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c**2 at the CERN SPS Collider”, *Phys.Lett.* **B126** (1983) 398–410, doi:10.1016/0370-2693(83)90188-0.

[42] P. W. Higgs, “Broken Symmetries and the Masses of Gauge Bosons”, *Phys. Rev. Lett.* **13** (Oct, 1964) 508–509, doi:10.1103/PhysRevLett.13.508.

[43] N. Cabibbo, “Unitary Symmetry and Leptonic Decays”, *Phys. Rev. Lett.* **10** (Jun, 1963) 531–533, doi:10.1103/PhysRevLett.10.531.

[44] M. Kobayashi and T. Maskawa, “CP Violation in the Renormalizable Theory of Weak Interaction”, *Prog.Theor.Phys.* **49** (1973) 652–657, doi:10.1143/PTP.49.652.

[45] J. Bjorken and S. Glashow, “Elementary Particles and SU(4)”, *Phys.Lett.* **11** (1964) 255–257, doi:10.1016/0031-9163(64)90433-0.

[46] S. L. Glashow, J. Iliopoulos, and L. Maiani, “Weak Interactions with Lepton-Hadron Symmetry”, *Phys. Rev. D* **2** (Oct, 1970) 1285–1292, doi:10.1103/PhysRevD.2.1285.

[47] J. E. Augustin et al., “Discovery of a Narrow Resonance in e^+e^- Annihilation”, *Phys. Rev. Lett.* **33** (Dec, 1974) 1406–1408, doi:10.1103/PhysRevLett.33.1406.

[48] J. J. Aubert et al., “Experimental Observation of a Heavy Particle J”, *Phys. Rev. Lett.* **33** (Dec, 1974) 1404–1406, doi:10.1103/PhysRevLett.33.1404.

[49] M. Kobayashi and T. Maskawa, “CP Violation in the Renormalizable Theory of Weak Interaction”, *Prog.Theor.Phys.* **49** (1973) 652–657, doi:10.1143/PTP.49.652.

[50] “FNAL Website, Discoveries at Fermilab - Discovery of the Bottom Quark”. URL http://www.fnal.gov/pub/inquiring/physics/discoveries/bottom_quark_pr.html.

[51] “CERN Website, Achievements with Antimatter”. URL http://cern-discoveries.web.cern.ch/cern-discoveries/courier/heavylight/heavylight.html.

[52] The ALEPH, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the SLD Electroweak and Heavy Flavour Groups, “Precision Electroweak Measurements on the Z Resonance”, *Phys. Rept.* **427** (2006) 257, arXiv:hep-ex/0509008.

[53] The ALEPH, DELPHI, L3, OPAL Collaborations, the LEP Electroweak Working Group, “Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP”, *Phys. Rept.* **532** (2013) 119, arXiv:1302.3415.

[54] F. Abe et al., “Observation of Top Quark Production in $\overline{p}p$ Collisions with the Collider Detector at Fermilab”, *Phys. Rev. Lett.* **74** (Apr, 1995) 2626–2631, doi:10.1103/PhysRevLett.74.2626.

[55] S. Abachi et al., “Search for High Mass Top Quark Production in $p\overline{p}$ Collisions at $\sqrt{s}$ = 1.8 TeV”, *Phys. Rev. Lett.* **74** (Mar, 1995) 2422–2426, doi:10.1103/PhysRevLett.74.2422.

[56] “CERN Courier, The W and Z at LEP”. URL http://cerncourier.com/cws/article/cern/29076.

[57] T. e. a. Aaltonen, “Higgs boson studies at the Tevatron”, *Phys. Rev. D* **88** (Sep, 2013) 052014, doi:10.1103/PhysRevD.88.052014.

[58] E. Boos and L. Dudko, “The Single Top Quark Physics”, *Int.J.Mod.Phys.* **A27** (2012) 1230026, doi:10.1142/S0217751X12300268, arXiv:1211.7146.

[59] “CERN Website, The accelerator complex”. URL http://home.web.cern.ch/about/accelerators.

[60] “CERN Website, Radiofrequency Cavities”. URL http://home.web.cern.ch/about/engineering/radiofrequency-cavities.

[61] “CERN Linac2 Duoplasmatron Schematic”. URL http://linac2.home.cern.ch/linac2/sources/source.htm.

[62] P. Service, “Linac 2. Linac 2”, (Feb, 1979).

[63] K. S. Klaus Hanke, and Michel Chanel.

[64] “Image, CERN PSbooster 2 batch filling scheme of the PS”. URL http://ps-div.web.cern.ch/ps-div/LHC-PS/schindlrf.gif.

[65] “Image, CERN PS Complex”. URL http://ab-div-abp-hei-ps.web.cern.ch/ab-div-abp-hei-ps/images/pscomplex[1].gif.

[66] "Image, CERN PS Dipoles". URL http://home.web.cern.ch/sites/home.web.cern.ch/files/styles/medium/public/image/accelerator/2013/01/accelerators-ps.jpg?itok=5E3Tp4Q0.

[67] P. TRACZYK, "SPS tunnel", (Nov, 2013).

[68] C. Simoes, "Large Hadron Collider (LHC)", *astronoo* (June, 2013).

[69] "Image, CERN Accelerator Complex". URL http://www.cernmg.free.bg/en/cern03.jpg.

[70] "CERN Website, Super Proton Synchrotron". URL http://home.web.cern.ch/about/accelerators/super-proton-synchrotron.

[71] "CERN Website, Cryogenics: Low temperatures, high performance". URL http://home.web.cern.ch/about/engineering/cryogenics-low-temperatures-high-performance.

[72] "CERN Website, A vacuum as empty as interplanetary space". URL http://home.web.cern.ch/about/engineering/vacuum-empty-interplanetary-space.

[73] "Image, LHC Dipole being lowered into cavern". URL http://lhc-machine-outreach.web.cern.ch/lhc-machine-outreach/images/cern-photos/CE0072M.jpg.

[74] "Image, LHC Dipole Magnetic Field". URL http://www.quantumdiaries.org/wp-content/uploads/2011/04/LHC-Dipole-Magnetic-Field.png.

[75] M. Barnes et al., "Injection and extraction magnets: Kicker magnets", arXiv:1103.1583.

[76] "CERN Website, CMS". URL http://home.web.cern.ch/about/experiments/cms.

[77] CMS Collaboration, "Precise Mapping of the Magnetic Field in the CMS Barrel Yoke using Cosmic Rays", *JINST* **5** (2010) T03021, doi:10.1088/1748-0221/5/03/T03021, arXiv:0910.5530.

[78] "CERN Website, Using Russian navy shells". URL http://cms.web.cern.ch/news/using-russian-navy-shells.

[79] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET",.

[80] W. Adam, R. Frhwirth, A. Strandlie, and T. Todor, "Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC",.

[81] H. Bethe and W. Heitler, "On the Stopping of fast particles and on the creation of positive electrons", *Proc.Roy.Soc.Lond.* **A146** (1934) 83–112, doi:10.1098/rspa.1934.0140.

[82] C. Weiser, "A Combined Secondary Vertex Based B-Tagging Algorithm in CMS", CMS Note 2006/014, 2006.

[83] A. Schalicke et al., "Geant4 electromagnetic physics for the LHC and other HEP applications", *J.Phys.Conf.Ser.* **331** (2011) 032029, doi:10.1088/1742-6596/331/3/032029.

[84] CMS Collaboration Collaboration, "CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update", Technical Report CMS-PAS-LUM-13-001, CERN, Geneva, 2013.

[85] T. Sjöstrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 physics and manual", *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.

[86] J. Alwall et al., "MADGRAPH 5: Going Beyond", *JHEP* **1106** (2011) 37, doi:10.1007/JHEP06(2011)128, arXiv:arXiv:1106.0522.

[87] P. Nason, "A New method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* **0411** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.

[88] CMS Collaboration, "Utilities for Accessing Pileup Information for Data", cms twiki, CERN, 2012.

[89] CMS Collaboration, "Multivariate Electron Identification", cms twiki, CERN, 2014.

[90] CMS Collaboration, "Jet Energy Resolution", cms twiki, CERN, 2015.

[91] CMS Collaboration, "Search for the standard model Higgs boson produced in association with a W or a Z boson and decaying to bottom quarks", *Phys.Rev.* **D89** (2014), no. 1, 012003, doi:10.1103/PhysRevD.89.012003, arXiv:1310.3687.

[92] CMS Collaboration, "Recommended Jet Energy Corrections and Uncertainties For Data and MC", cms twiki, CERN, 2015.

[93] CMS Collaboration, "Jet Energy Resolution in CMS at sqrt(s)=7 TeV",.

[94] CMS Collaboration, "Identification of b-quark jets with the CMS experiment", *JINST* **8** (2013) P04013, doi:10.1088/1748-0221/8/04/P04013, arXiv:1211.4462.

[95] D. M. Puigh, "Calibration of the Combined Secondary Vertex b-Tagging discriminant using dileptonic ttbar and Drell-Yan events", CMS Note 2013/130, 2013.

[96] CMS Collaboration, "Operating Points for b-tagging Algorithms", cms twiki, CERN, 2015.

[97] CMS Collaboration, "pt(top-quark) based reweighting of ttbar MC", cms twiki, CERN, 2014.

[98] N. Kidonakis, "NNLL threshold resummation for top-pair and single-top production", *Phys.Part.Nucl.* **45** (2014), no. 4, 714–722, `doi:10.1134/S1063779614040091`, `arXiv:1210.7813`.

[99] A. Hocker et al., "TMVA - Toolkit for Multivariate Data Analysis", *PoS* **ACAT** (2007) 040, `arXiv:physics/0703039`.

[100] Y. Freund and R. E. Schapire, "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting", *Journal of Computer and System Sciences* **55** (1997) 119 – 139, `doi:http://dx.doi.org/10.1006/jcss.1997.1504`.

[101] B. A. Dobrescu and J. D. Lykken, "Semileptonic decays of the standard Higgs boson", *JHEP* **1004** (2010) 083, `doi:10.1007/JHEP04(2010)083`, `arXiv:0912.3543`.

[102] CMS Collaboration, "Jet Energy Corrections: Official Software Tools for applying JEC Corrections and Uncertainties", cms twiki, CERN, 2015.

[103] CMS Collaboration, "Jet energy scale uncertainty sources", cms twiki, CERN, 2015.

[104] CMS Collaboration, "Standard Model Cross Sections for CMS at 8 TeV", cms twiki, CERN, 2015.

[105] CMS Collaboration, "SM Higgs production cross sections at s = 8 TeV (update in CERN Report3)", cms twiki, CERN, 2014.

[106] T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl.Instrum.Meth.* **A434** (1999) 435–443, `doi:10.1016/S0168-9002(99)00498-2`, `arXiv:hep-ex/9902006`.

[107] ATLAS Collaboration, "Procedure for the LHC Higgs boson search combination in summer 2011",.

[108] J. Conway, "Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra", `arXiv:1103.0354`.

[109] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur.Phys.J.* **C71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0,10.1140/epjc/s10052-013-2501-z, arXiv:1007.1727.

[110] CMS Collaboration, "Documentation of the RooStats-based statistics tools for Higgs PAG", cms twiki, CERN, 2015.

[111] RooStats Team Collaboration, "RooStats for Searches", arXiv:1203.1547.

List of Acronyms

$\not\!\!E_{\mathrm{T}}$ Missing Transverse Energy

AdaBoost Adaptive Boost

ALICE A Large Ion Collider Experiment

AN Analysis Note

ATLAS A Toroidal LHC Apparatus

BDT Boosted Decision Tree

BR Branching Ratio

CERN European Center for Nuclear Research

CMS Compact Muon Solenoid

DAQ Data Acquisition

ECAL Electromagnetic Calorimeter

EW Electro-Weak

FOM Figure Of Merit

gg-F Gluon-Gluon Fusion

HCAL Hadronic Calorimeter

HLT High Level Trigger

IP Interaction Point

LEP Large Electron-Positron Collider

LHC Large Hadron Collider

LHCb Large Hadron Collider beauty

LHCf Large Hadron Collider forward

MC Monte-Carlo

ME Matrix Element

MoEDAL Monopole and Exotics Detector At the LHC

MVA Multi-Variate Analysis

NLO Next to Leading Order

NNLO Next to Next to Leading Order

PDF Parton Distribution Function

QCD Quantum Chromodynamics

QED Quantum Electrodynamics

QFT Quantum Field Theory

RF Radio-Frequency

SM Standard Model

TOTEM Total Cross Section, Elastic Scattering and Diffraction Dissociation

VBF Vector Boson Fusion

VEV Vacuum Expectation Value