UNIVERSITY OF CALIFORNIA
Santa Barbara

Search for supersymmetry in proton-proton collisions at 7 TeV in events with a single lepton, b-tagged jets and missing transverse momentum at the LHC using data collected from the CMS detector.

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Physics

by

Wing H. To

Committee in Charge:

Professor David Stuart, Chair

Professor Claudio Campagnari

Professor David Berenstein

March 2013

The Dissertation of
Wing H. To is approved:

Professor Claudio Campagnari

Professor David Berenstein

Professor David Stuart, Committee Chairperson

December 2012

Search for supersymmetry in proton-proton collisions at 7 TeV in events with a single lepton, b-tagged jets and missing transverse momentum at the LHC using data collected from the CMS detector.

by

Wing H. To

Curriculum Vitæ

Wing H. To

Education

2013	Doctor of Philosophy in Physics, University of California, Santa Barbara.
2006	Bachelor of Science in Physics, University of California, Santa Barbara
2003	Associate of Science, Citus College.

Experience

2007 – 2012	Graduate Research Assistant, High Energy Experiment, University of California, Santa Barbara.
2006 – 2006	Research Assistant, Seismology Earth Science Dept, University of California, Santa Barbara.
2005	Summer Undergraduate Research Fellow, Quantum Optics at Jet Propulsion Lab, California Institute of Technology.

Selected Publications

CMS Collaboration: "Search for supersymmetry in pp collisions at 7 TeV in events with a single lepton, jets, and missing transverse momentum," In *JHEP 08 (2011) 156,974, arXiv:1107.1870.*

CMS Collaboration: "Search for supersymmetry in pp collisions at 7 TeV in events with a single lepton, jets, and missing transverse momentum with 4.98 fb^{-1} of integrated luminosity," In *preprint: CERN-PH-EP-2012-348*

Abstract

Search for supersymmetry in proton-proton collisions at 7 TeV in events with a single lepton, b-tagged jets and missing transverse momentum at the LHC using data collected from the CMS detector.

Wing H. To

This thesis describes a search for supersymmetry (SUSY) at the Large Hadron Collider using data collected in Run 2011 at center of mass energy of 7 TeV by the Compact Muon Solenoid detector with an integrated luminosity of 4.98 fb^{-1}. First, the Large Hadron Collider is briefly described along with the Compact Muon Solenoid detector. Second, the Standard Model of particle physics is described which leads to the motivations for supersymmetry and its description. Finally, a search for supersymmetry in events with a single lepton, 2 b-tagged jets and missing transverse energy is documented in detail. The results are used to predict the yields from the standard model in bins missing transverse energy ($\not\!\!E_T$) and total transverse hadronic energy (HT). Limits are set using these results using a minimal supergravity SUSY model and simplified model spectra SUSY model.

Professor David Stuart

Dissertation Committee Chair

Contents

List of Figures **x**

List of Tables **xii**

1 Introduction **1**

2 Large Hadron Collider **5**

3 Compact Muon Solenoid (CMS) **9**

4 Theory **23**
4.1 The Standard Model 23
4.1.1 Hierarchy 25
4.1.2 Coupling constants 26
4.1.3 Dark Matter 28
4.2 Supersymmetry 29

5 Search for Supersymmetry **36**
5.1 Monte Carlo Samples 39
5.2 Datasets and Triggers 39
5.3 Physics Object Reconstruction 46
5.3.1 Electrons 46
5.3.2 Muon 50
5.3.3 Jets 52
5.3.4 $\not{E}_T$, Missing Transverse Energy and HT: 55
5.3.5 b-tagged jets: 55
5.4 Event Selections 56
5.5 Single lepton prediction 59

5.5.1 Single lepton systematics 67
5.6 Dilepton and Tau Prediction 70
5.6.1 Dilepton and tau systematics 85
5.7 QCD Prediction 87
5.7.1 Muon MET QCD prediction 90
5.7.2 Electron MET QCD prediction 96
5.7.3 QCD Muon p_T prediction 99
5.7.4 Electron pT QCD prediction 101
5.7.5 Combined QCD prediction 103
5.7.6 QCD systematics 105
5.8 $t\bar{t}$ and $W + jets$ cross section uncertainties 109
5.9 W Polarization in $t\bar{t}$ and $W + jets$ 114
5.9.1 W-boson polarization in $t\bar{t}$ decays 115
5.9.2 W-boson polarization in $W + jets$ events 116
5.10 Complete Prediction Results 116
5.11 mSUGRA and SMS Limits 117

6 Conclusion **127**

Bibliography **129**

List of Figures

2.1 LHC Diagram 8

3.1 Three dimensional diagram of CMS 11
3.2 Three dimensional diagram of Pixel 12
3.3 RZ view of silicon strip tracker 13
3.4 Silicon Strip Tracker Module Sizes 14
3.5 ECAL Lead Tungstate Crystal 14
3.6 ECAL 3D Diagram of Barrel, Endcap and Preshower 16
3.7 HCal Diagram for the HB, HE and HO. 18
3.8 Drift Tube Geometry in the XY plane. 20
3.9 Cathode Strip Chamber Geometry in RZ plane. 21
3.10 Resistive Plate Chamber geometry in XY plane. 22

4.1 Higgs propagator loop correction. 27
4.2 Running of the coupling constants. 32
4.3 Rotation curve of NGC 6503 33
4.4 Particles of the SM and SUSY 34
4.5 Feymann diagram for T1tttt. 35

5.1 Feymann diagram of a semileptonic TTBar decay. 38
5.2 The number of primary vertices before RW. 41
5.3 The number of primary vertices before RW. 41
5.4 The number of primary vertices after RW. 41
5.5 The number of primary vertices after RW. 41
5.6 Slice of the CMS Detector in XY Plane 47
5.7 Comparison of Neutrino PT and MET 60
5.8 Samples of MET Templates 61
5.9 Neutrino PT versus Lepton PT in Single Lepton MC 62

5.10 Kappa SL. 63
5.11 MC Closure in Single Lepton Prediction 64
5.12 Comparison of lepton PT with and without smearing 65
5.13 Single lepton only prediction in data. 66
5.14 Extrapolation of lepton pT 67
5.15 MET prediction for ignored lepton 75
5.16 MET prediction for lost lepton 76
5.17 Leptonic and Hadronic Tau Response Funtcions 78
5.18 MET prediction for lepton plus leptonic tau 79
5.19 MET prediction for lepton plus hadronic tau 81
5.20 MET prediction for single tau 83
5.21 MET prediction for all five dilepton and tau 84
5.22 Combined Rel Iso Preselection 90
5.23 Fraction of b-tagged events in QCD 91
5.24 Muon MET and d0BS Preselection 92
5.25 Muon Relative Isolation MET 100 to 150 94
5.26 Muon Relative Isolation MET 150 to 250 95
5.27 Muon Relative Isolation MET 250 to Infinity 95
5.28 Electron Relative Isolation MET 100 to 150 97
5.29 Electron Relative Isolation MET 150 to 250 98
5.30 Electron Relative Isolation MET 250 to infinity 98
5.31 CRIso pT Dependency 101
5.32 Muon Relative Isolation PT 100 to 150 103
5.33 Muon Relative Isolation PT 150 to 250 105
5.34 Muon Relative Isolation PT 250 to infinity 106
5.35 Electron Relative Isolation PT 100 to 150 107
5.36 Electron Relative Isolation PT 150 to 250 108
5.37 Electron Relative Isolation PT 250 to 350 109
5.38 Electron Relative Isolation PT 350 to 450 110
5.39 Electron Relative Isolation PT 450 to infinity 111
5.40 Dimuon Mass 112
5.41 Dielec Mass 112
5.42 TopChi2 Muon 114
5.43 TopChi2 Elec 114
5.44 Combined MET Prediction 120
5.45 mSUGRA Eff 123
5.46 mSUGRA limits 124
5.47 SMS-T1tttt Eff 125
5.48 SMS-T1tttt limits 126

List of Tables

5.1 Monte Carlo Samples 40
5.2 Data samples analyzed in the note. 43
5.3 Single Lepton Prediction $500 \leq HT < 750$ 68
5.4 Single Lepton Prediction $HT \geq 750$ 68
5.5 Systematic uncertainties for Single lepton $H_T \in [500, 750)$ 71
5.6 Systematic uncertainties for Single lepton $H_T \geq 750$ 71
5.7 Dilepton/tau prediction for $500 \leq HT < 750$ bin. 85
5.8 Dilepton/tau prediction for $HT \geq 750$ bin. 85
5.9 Systematic uncertainties for dilepton and tau $H_T \in [500, 750)$ GeV 87
5.10 Systematic uncertainties for dilepton and tau $H_T \geq 750$ GeV . . . 87
5.11 Selections for muon and electron $\not\!\!E_T$ QCD prediction. 92
5.12 Muon MET prediction for QCD 94
5.13 Elec MET prediction for QCD 97
5.14 Shape and signal selections lepton $p_T$prediction. 99
5.15 Muon PT prediction for QCD. 100
5.16 Electron PT prediction for QCD. 102
5.17 Combined QCD prediction. 104
5.18 Systematic uncertainties for QCD $H_T \in [500, 750)$ GeV 108
5.19 Systematic uncertainties for QCD $H_T \geq 750$ GeV 109
5.20 Data and MC comparison with different Z mass range. 113
5.21 Data and MC comparison at different χ^2 cut. 114
5.22 Combined MET Prediction $H_T \in [500, 750)$ GeV 118
5.23 Combined MET Prediction $H_T \in [500, 750)$ GeV 119

Chapter 1

Introduction

Modern particle physics pushes mankind's understanding of the fundamental constituents and interactions of our universe by probing matter at the smallest possible length scales. From John Dalton's "atomic theory," through Rutherford's discovery of the nucleus, up to the construction of the Standard Model, scientists have probed the constituents of matter from the atomic down to sub nucleon level, which is the length scale where quarks and gluons exist within protons and neutrons.

The Large Hadron Collider is the latest tool with which to measure and discover new physics at the smallest length scale. Four different detectors were built around the accelerator's 27 km circumference. Alice (A Large Ion Collider Experiment) was designed for heavy ion collisions to study quark-gluon plasma. LHCb (LHC beauty) specializes in bottom quark physics and the measurements of CP violating processes. There is Atlas (A Toroidal Lhc ApparatuS). The detector

used in this analysis is CMS (Compact Muon Solenoid) constructed 100 meters underground outside the small town of Cessy in the Jura foothills in France. This thesis will briefly describe the LHC (Chapter 2) and the CMS detector (Chapter 3) since multiple reference materials are available on these subjects[1] [4]. The theoretical motivation for supersymmetry at the LHC is discussed in chapter 4.

The main subject of this thesis will be on a search for supersymmetry (SUSY) in the single lepton, with b-tagged jets and missing transverse energy ($\not\!\!E_T$) using 4.98 fb^{-1} of integrated luminosity at the CMS detector (Chapter 5). This analysis is based upon a previously published analysis without the b-tagging requirement[12]. The single lepton, with b-tagged jets and $\not\!\!E_T$ analysis has sensitivity to many SUSY models where bottom (b) quarks participate in the decay cascade. The requirement of a single lepton [1] rejects many light flavor hadronic events (QCD) while isolating two predictable backgrounds, $t\bar{t}$ and $W + jets$ events. The analysis requires the event to have at least two b-tagged jets above 40 GeV and four or more total number of jets which are expected from the cascade decays of strongly produced massive SUSY particles such as squarks and gluinos. High $\not\!\!E_T$ is expected from SUSY signal models due to production of Lightest Supersymmetric Partners (LSP) in R-parity conserving SUSY.

[1] Throughout this thesis, "lepton" will refer to only electrons and muons unless otherwise specified.

The main background in this analysis is $t\bar{t}$ events that has two separate components, which are predicted separately. The two components are the single lepton events from semileptonic $t\bar{t}$ decays and dilepton-tau events from fully leptonic events or $t\bar{t}$ events involving taus. The dilepton and tau are lumped together since both channels produce additional undetected neutrinos which are predicted by a similar method. The single lepton component is the dominant background in most signal bins, which is predicted by the lepton spectrum method[11]. Although the analysis vetoes events with a second identified lepton, two lepton events (dilepton) contribute to our signal bins due to lepton identification inefficiency and detector coverage. There are also contributions from single tau events and lepton plus tau events. The dilepton and tau contributions are predicted by a separate method described in chapter 5. Light flavor hadronic event's (QCD) contribution is expected to be very small, but due to its contribution to our control sample a separate prediction is also made for the QCD contribution. The last and the smallest contribution is from $Z + jets$ events where one lepton is lost or ignored. This portion of the background is taken from Monte Carlo and found to be negligible. The final prediction is made in bins of $\not\!\!E_T$ and HT with all the separate components combined. Where HT is defined as the scalar sum of all transverse energy in the jets identified in each event. Since no significant

excess was found in the observed data, limits were made in terms of a minimal SUperGRAvity (mSUGRA) SUSY models and Simplified Model Spectra (SMS).

Chapter 2

Large Hadron Collider

The Large Hadron Collider (LHC) is currently the world's largest particle accelerator. It is located at the European Organization of Nuclear Research (CERN)[1]. The LHC tunnel has a circumference of 27 km and located approximately 100 meters underground, straddling the french-swiss border in Europe. It is a synchrotron accelerator capable of accelerating protons up to 7 TeV or lead nuclei up to 574 TeV per beam. The acceleration is done in steps shown schematically in figure 2.1. Protons are produced in a duoplasmatron source[2] and injected into the LINear ACcelerator (LINAC 2) where they are accelerated to 50 MeV. The LINAC feeds into the Proton Synchrotron Booster (PSB) which accelerates the protons to 1.4 GeV. The Proton Synchrotron (PS) takes the protons up to 26 GeV before injecting them into the Super Proton Synchrotron (SPS), which increases the protons' energy to 450 GeV. The SPS beam is finally steered into the main LHC ring where the protons can be accelerated up to 7 TeV of energy

where the two beams can be collided together at a center of mass energy up to 14 TeV. The proton beams are trapped in bunches by 8 Radiofrequency Cavities (RFC) and steered by 1232 dipole magnets and 392 quadrupole magnets operating at a field strength of up to 8.3 Tesla. The bunches are formed and accelerated using 8 Radiofrequency Cavities (RFC) delivering electric potential difference of 2 MV per cell at a frequency of 400 MHz. The RFC accelerates the proton up to an energy of 7 TeV, or a mere 3 meters per second slower than the speed of light. The dipole magnets are designed to keep the protons in a circular orbit while the quadrupole magnets are used to steer and focus the beams. Twenty four inner triplet magnets also squeeze the beams just before the beams collide at the detectors. The LHC is also designed to have high instantaneous luminosity of $1.8 \times 10^{34}\ cm^{-2}s^{-1}$. Each proton bunch contains up to 1.15×10^{11} protons and each beam contains up to 1404 bunches colliding at 40 MHz. The instantaneous luminosity in a symmetric collider is,

$$L = N^2/(t \times \sigma^2) \times B \tag{2.1}$$

where N is the number of particle per bunch, t is the time between crossing, σ is the effective transverse cross section of the beams and B is the number of bunches (2808). Due to their connection to experimental hardware parameters, luminosity is defined as,

$$L_{bb} = f \times N^2/(4\epsilon\beta^*) \times B \tag{2.2}$$

where f is the frequency of collision (40 MHz), ϵ is the emittance, which is the minimal width of the beam (3.75 μm), and β^* is the amplitude function, which is approximately the distance that the beam width doubles as it moves away from the interaction point (0.55 m).

$$\begin{aligned} L = &(40 \times 10^6 * 2808 \times (1.15 \times 10^{11})^2 / (4 \times 3.75 \times 10^{-4} \times 55)\ cm^{-2}s{-}1 \\ &= 1.8 \times 10^{34}\ cm^{-2}s^{-1}) \end{aligned} \tag{2.3}$$

This is equivalent to 10 $nb^{-1}s^{-1}$ leading to approximately 600 million proton-proton inelastic scattering per second. Multiple proton-proton collisions in a single crossing lead to "pile-up" effects which must be modeled in the Monte Carlo at this luminosity as discussed in the analysis part of this thesis. The analysis is based on the LHC 2011 Run data at a center of mass energy of 7 TeV, which ran at a peak luminosity of $3.5 \times 10^{33}\ cm^{-2}s^{-1}$ and comprises a total integrated luminosity of 4.98 fb^{-1}.

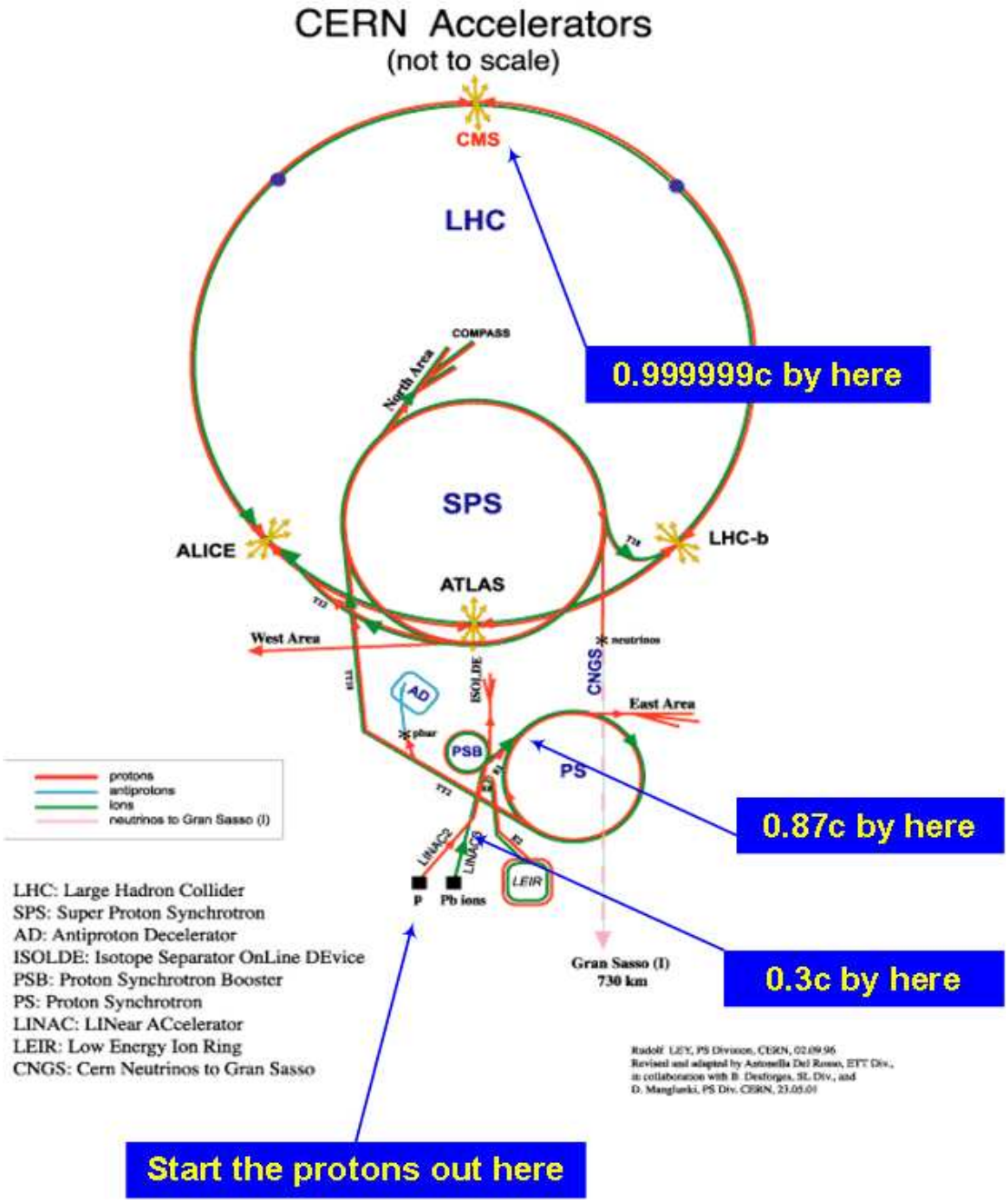

Figure 2.1: Diagram of the LHC acceleration sequence. Figure taken from [3]

Chapter 3

Compact Muon Solenoid (CMS)

The Compact Muon Solenoid (CMS) is one of two general purpose physics detectors at the LHC. CMS is designed for high precision momentum measurement of muons hence the name of the detector[4]. The solenoid comes in due to the requirement of large bending power of the superconducting solenoid magnet required to measure the momentum of very energetic charged particles such as muons. With an overall length of 21.6 m, and a diameter of 14.6 m and weighing in at 14500 tons shown in figure 3.1, CMS is compact only relative to the other experiment at the LHC.

In order to precisely measure the momentum of muons up to 1 TeV, the magnet must have a large transverse field to the muon's momentum. CMS's superconducting magnet has a field strength of 4 Tesla along the beam direction with an inner bore size of 5.9 m and length of 12.9 m. The coil is made of pure aluminum cooled to 4.2 K with liquid helium. The magnetic field is returned back in the

-Z direction through the magnetic yoke. The magnetic yoke is compose of both a cylindrical and endcap disks in layers with muon chambers dispersed between them. The magnetic flux density varies due to the structure of the magnetic yoke between 0.6 to 2.1 Tesla.

Starting from the interaction point, the first detector element in CMS is the silicon pixel tracker. The pixel detector is made of 150 x 100 μm^2 cells that detects charged particles as they pass through them. The pixel detector is composed of a barrel and endcap region as shown in figure 3.2. The barrel consists of three layers located at radii of 4.4, 7.3 and 10.2 cm with a length of 53 cm. The endcaps are located along on each side of the interaction point at $z = \pm 34.5\ cm$ and $z = \pm 46.5\ cm$. The modules on the endcap are arranged in a turbine-like geometry in order to create charge sharing between the pixel cells to improve positional resolution.

The second detector element in CMS is the silicon strip tracker. Although the strip and pixel combine to make the entire tracker, due to differences in their geometry, readout electronics and physics requirements they are treated as separate subdetectors. The silicon strip tracker is often just referred to within CMS as the Tracker. The strip tracker is made up of four different regions called partitions of the strip tracker. The tracker inner barrel/disks (TIB/TID), tracker outer barrel (TOB), tracker endcap plus (TEP, or TEC+ depending if you're in DAQ or

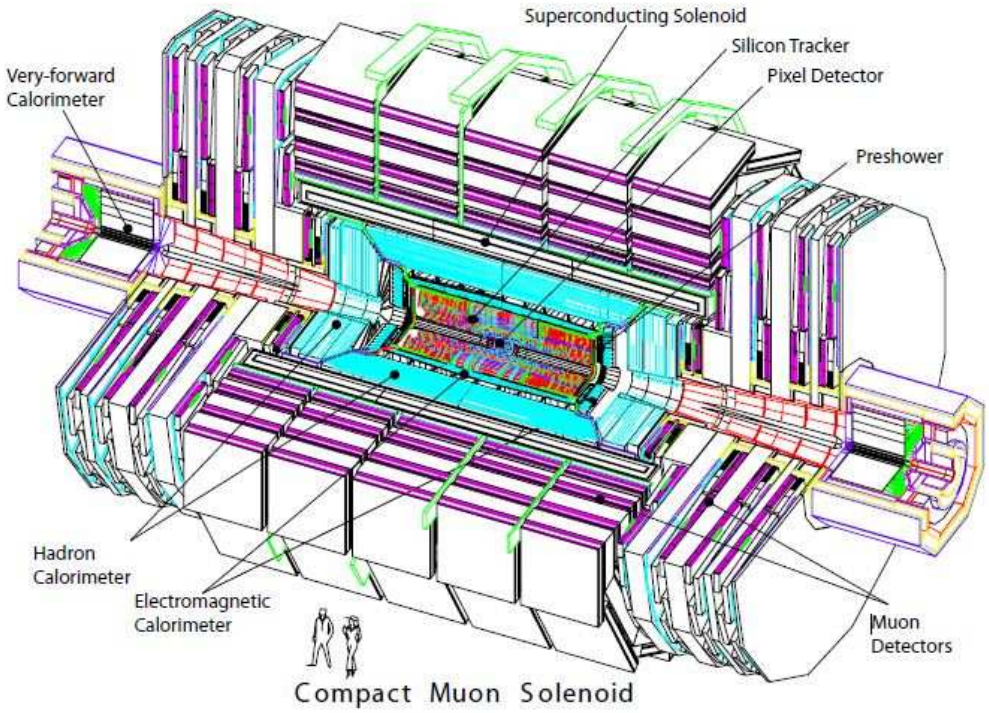

Figure 3.1: Three dimensional diagram of CMS with each sub detector element shown. Figure taken from [4].

DQM), and tracker endcap minus (TEM or TEC-). The modules in the the TIB are made up of silicon strip wafers of 12 cm in length, 6 cm wide and 320 μm thick. 256 strips of p+ implantation is done on the n- bulk with a w/p value of 0.25 as shown in [13]. Thus, in the TIB the width of the strips are 60 μm. Each module is assembled onto rods and placed in a cylindrical superstructure of the barrel at radii 25.5 cm, 33.9 cm, 41.85 cm and 49.8 cm extending from z = -70.0 cm to z = +70 cm. The inner two layers are double sided modules with a strip pitch of

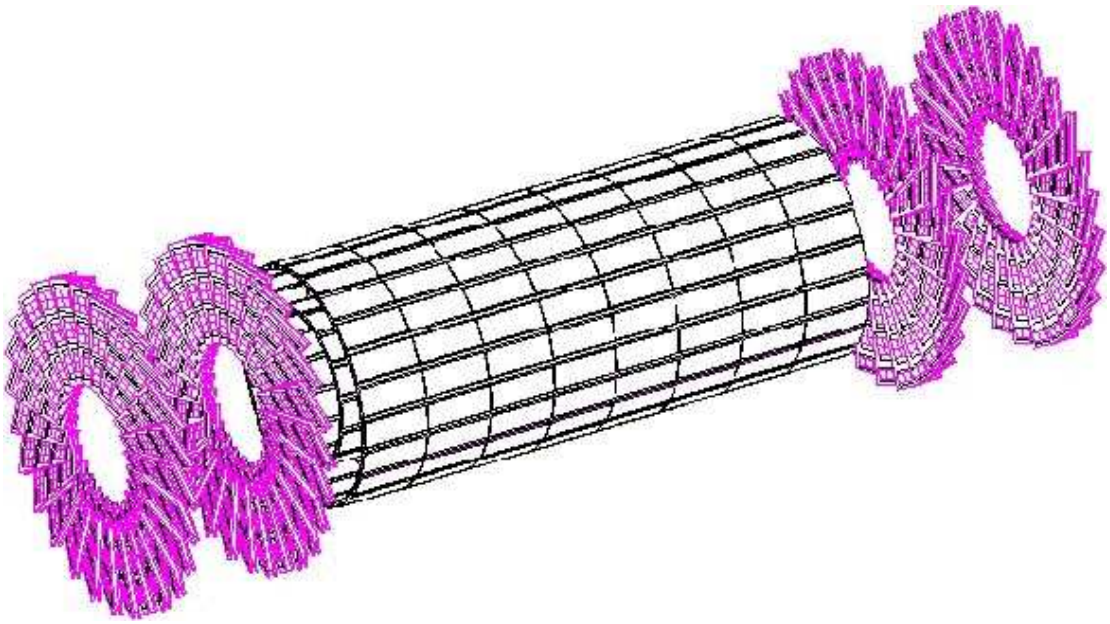

Figure 3.2: Three dimensional diagram of Silicon Pixel Detector (Pixel). The barrel region is shown in white, while the endcaps are shown in pink. Figure taken from [6].

80 μm for improved z resolution. The TIB also include six disks Tracker Inner Disks (TID). These six disks are located between $\pm$80.0 cm to $\pm$100 and the strip radii extend from a radius 20.0 cm to 50.0 cm as seen on figure 3.3. The sizes of disk modules varies between each ring's geometry as shown figure 3.4. The inner disk modules are also 320 μm thick with 256 strips implanted onto each module with the same w/p specification. Moving out from the TIB/TID, the TOB has 6 layers located at average radii of 60.8, 69.2, 78.0, 86.8, 96.5, and 108.0 cm from the interaction point. The TOB uses the thicker sensors of 500 μm thickness. The TEP and TEM are design symmetrically with coverage extending radially from

22 cm to 113.5 cm and located between ±124 cm and ±280 cm along the z-axis. The TEP and TEM use the same 320μm wafer as the TIB/TID made into shapes to fit in the endcap rings as shown in figure 3.4.

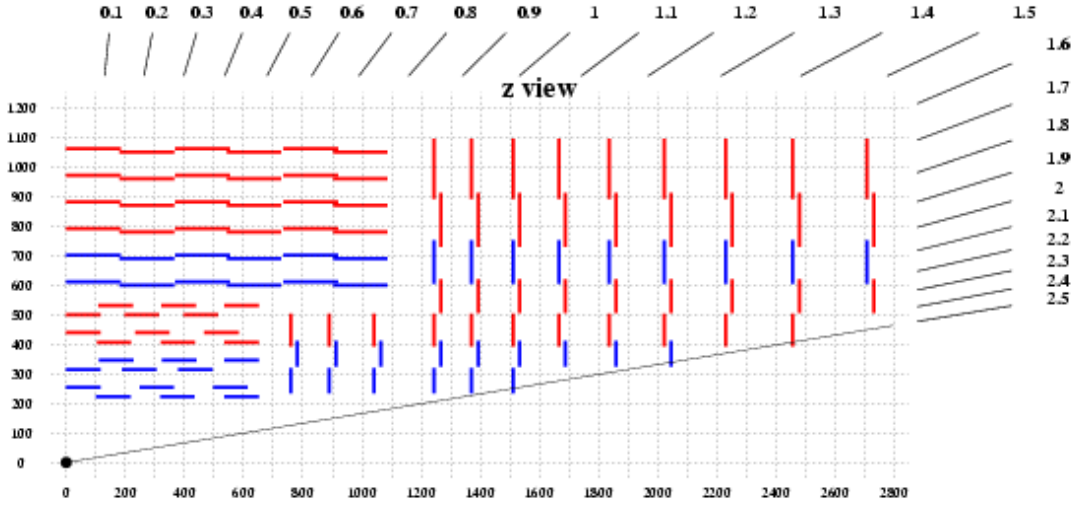

Figure 3.3: RZ view of silicon strip tracker. The modules in red double side modules while those in blue are single sided. Figure taken from [4].

CMS's Electromagnetic Calorimeter (ECAL) is made up of Lead Tungstate (PbWO4) crystals each with Avalanche Photo Didodes (APD) or Vacuum Phototriodes (VPT) attached to the ends of the crystals facing away from the interaction point as shown in figure 3.5. The Lead Tungstate has a high density (8.28 g/cm^3), and a very short radiation length of 0.89 cm and moliere radius of 2.2 cm. Each crystal is 22 cm in length which corresponds to 25.8 radiation lengths. The small moliere radius gives the crystal high positional resolution.

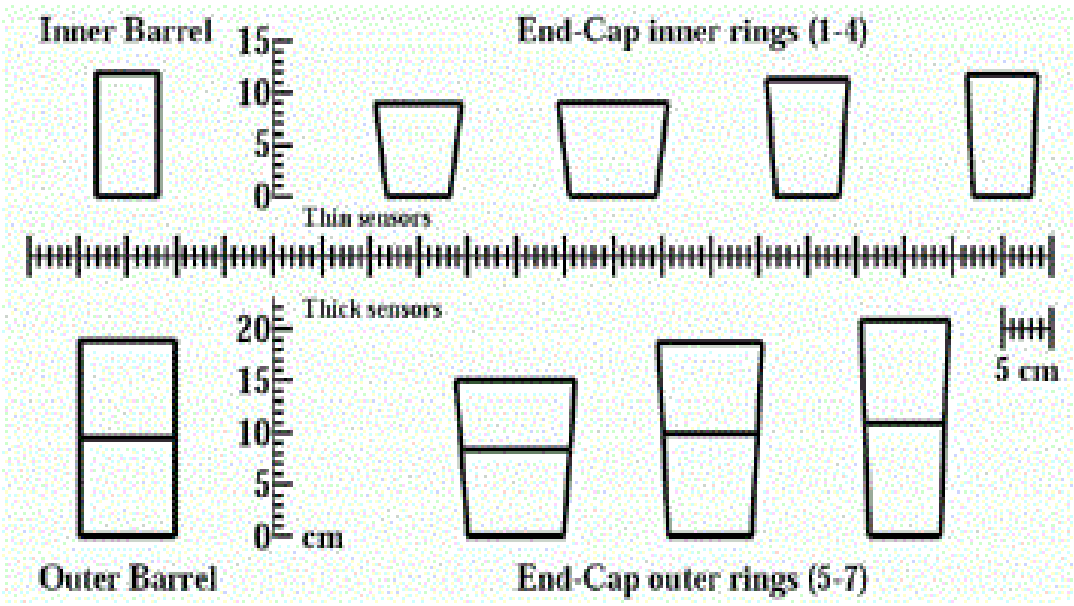

Figure 3.4: Silicon Strip Tracker Module Sizes. The modules are cut into shapes which fit the geometry of the modules as required to maximize detector coverage. Figure taken from [4].

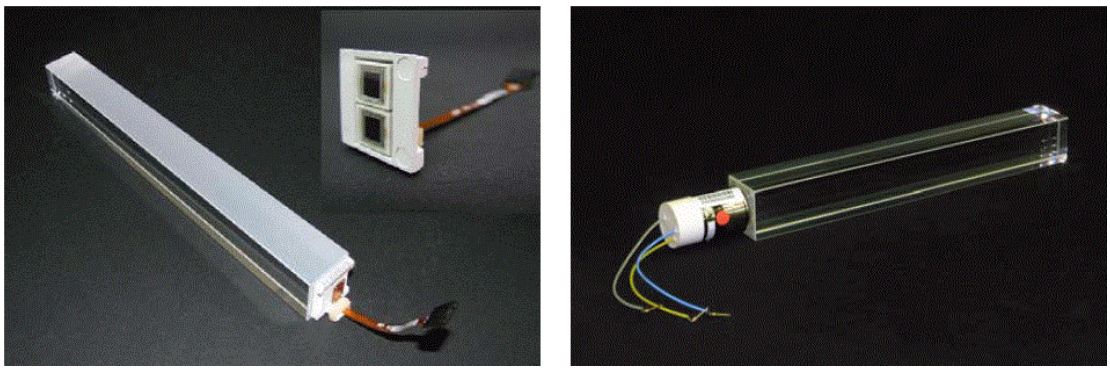

Figure 3.5: ECAL Lead Tungstate Crystal. Left: Barrel crystals with two APD attached to the end. Right: Endcap crystals with one VPT attached. Figure taken from [4].

The ECAL barrel (EB) is made up of 61200 crystals, which cover the entire detector in ϕ, and extend from -268 cm to +268 cm along the z-axis ($|\eta|$ coverage of 1.479) shown in figure 3.6. Each crystal has a front face dimension of 2.2×2.2 cm^2 and 2.6×2.6 cm^2 at the rear face. The front face is located at a radius of 129 cm. Each crystal has two APDs attached each with an active area of 0.5×0.5 cm^2 and read out in parallel for each event. The ECAL endcaps (EE) are made of 14648 crystals located ± 315.4 cm from the interaction point along z-axis and offers coverage in the $|\eta|$ range between 1.479 to 3.0. The EE crystal has a front face dimension of 2.862×2.862 cm^2 and a rear face dimension of 3.0×3.0 cm^2. Each EE crystal has one VPT attach to the rear face, VPTs are photon multipliers with a single gain stage and has a circular active area of 2.80 cm^2. In order to improve positional resolution of photons, a pair of Preshower (PS) is installed between the tracker and the Ecal Endcaps. The PS offers a fiducial coverage of $1.653 < |\eta| < 2.6$. The radiator is made of layers of lead with silicon strip sensors placed between the layers. The total thickness of the PS is 20 cm.

The Hadronic Calorimeter (HCAL) is located around the ECAL. The barrel region of the HCal offers a fiducial coverage of $|\eta| < 1.3$. The HCAL Barrel (HB) is made up of alternating layers of brass, steel and plastic scintillitors and optical fibers to carry the light to read out towers. The steel and brass layers are stacked in the following order: a 4.0 cm steel front plate, 8 layers of 5.05 cm thick brass,

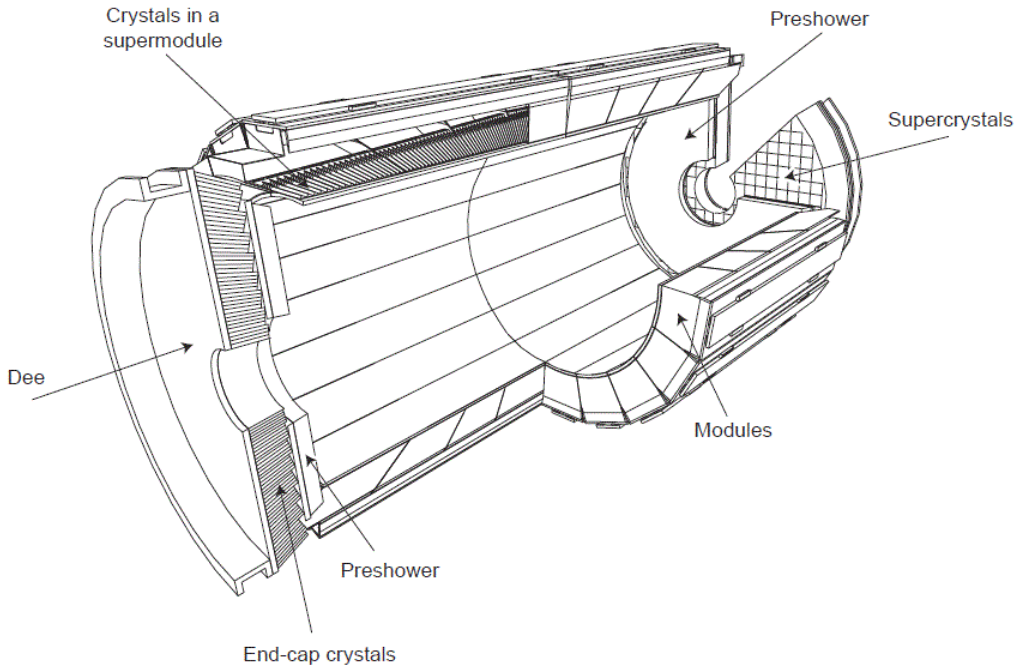

Figure 3.6: ECAL 3D Diagram of Barrel, Endcap and Preshower. The barrel region covers the center portion of the detector. Then the Preshower is placed in front of the endcap to give better positional resolution in the endcap. Finally each side is covered by the Endcap. Figure taken from [4].

6 layers of 5.65 cm thick brass and a 7.5 cm thick steel back plate. The HB has a total of 5.82 interaction length λ_i at $|\eta| = 0$ and 10.6 λ_i at $|\eta| = 1.3$. The plastic scintillitors are Kuraray SCSN81 which are placed between these layers with a thickness of 0.37 cm per layer. The plastic scintillators are divided into 32 η sector and 36 azimuthal wedges leading to a segmentation of 0.087×0.087 in $\eta \times \phi$ space, which corresponds to 5×5 crystals of the ECAL barrel. The plastic scintillators contains wavelength shifting fibers which guides the light to

an optical cable where it is read out by a hybrid photodiode (HPD). The HPD are photocathode biased at -8kV that accelerates photo-electrons onto a pixelated silicon photodiode. Due to the lack of interactions lengths at low $|\eta|$ an additional layer of the HCAL barrel is located outside the magnetic coil. This increases the total number of interaction lengths of the EB, HB, and HO together to 11.6 λ_i at $|\eta| = 0$. The endcap region of the HCAL (HE) covers the region of $1.3 < |\eta| < 3$. The HE is inserted into the ends of the CMS magnet, which requires its material to be non-magnetic. The HE is made up of brass plates and plastic scintillators. The scintillation lights are carried using wavelength shifting fibres to HPD read out modules. The total interaction length of the PS, EE and HE sum up to 10 λ_i. The HE has a segmentation of 0.087×0.087 for $|\eta| < 1.6$ and 0.17×0.17 for $|\eta| > 1.6$. The forward region of CMS is covered by the HCAL forward (HF) detector. Approximately 80% of the energy is deposited in the HF where $|\eta|$ goes from 3.0 to 5.0. High radiation tolerate materials are required at to build the HF. Steel plates are used as the absorber and quartz fibres are used to create Cherenkov light which are read out by photomultiplier tubes (PMT). The HF has an outer radius of 130.0 cm and located 1120 cm along the z-axis from the interaction point. Due to the large fraction of energy deposited in the HF, it is used as one of the luminosity monitors at CMS.

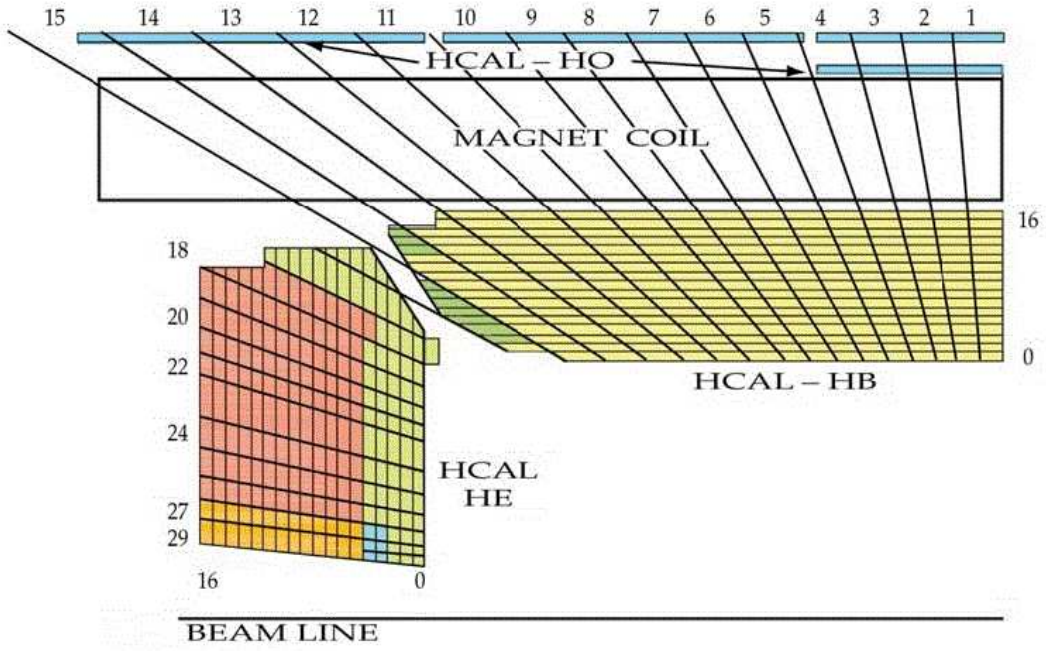

Figure 3.7: HCal Diagram for the HB, HE and HO. The HB and HO covers the barrel area up to $|\eta| = 1.3$, while the HE covers the endcap region in $1.3 < |\eta| < 3$. Figure taken from [4].

The outer most detector elements of CMS are the muon chambers. There are three different types of muon chambers in CMS, the Drift Tubes (DT), Cathode Strip Chambers (CSC) and Resistive Plate Chambers (RPC). The Drift Tubes (DT) occupy the central region of the detector with $|\eta| < 1.2$. The DTs are arranged in 4 layers in a cylindrical fashion, with 60 chambers in the inner 3 layers and 70 chambers in outer most layer as shown in figure 3.8. The active element of the DT are the 172000 sensitive wires each approximately 240 cm in length. The maximum drift time of the chamber filled with 85% Ar and 15%

CO_2 is 380 ns. The drift time is designed to limit the number of active channels while keeping occupancy low. In the endcap region, the Cathode Strip Chambers (CSC) covers the $0.9 < |\eta| < 2.4$, which has some overlap with the DTs as shown in figure 3.9. CMS contains 468 CSCs in the endcap range from $z = \pm 570\ cm$ to $z = \pm 1000\ cm$. The CSCs are multiwire proportional chambers where the wires run along the azimuthal direction allowing measurements of muons in the radial position. The ϕ measurement is made by interpolating the induced charge pulse read out on the wires. The third element of the muon system is the Resistive Plate Chambers (RPC). The RPCs are placed in both the barrel and the endcap region of CMS. The RPCs are parallel plate detectors with very good time resolution (< 25 ns) and only fair spatial resolution (track resolution of 150 μm). The main purpose of the RPCs are for triggering events which contains muons. The RPC made for CMS are basic double-gap modules with silicon strips which are read out. In the barrel region the RPCs are placed either in front of or on both sides of the DTs as shown in figure 3.10. In the endcap region the RPCs are mounted on both sides of all CSCs.

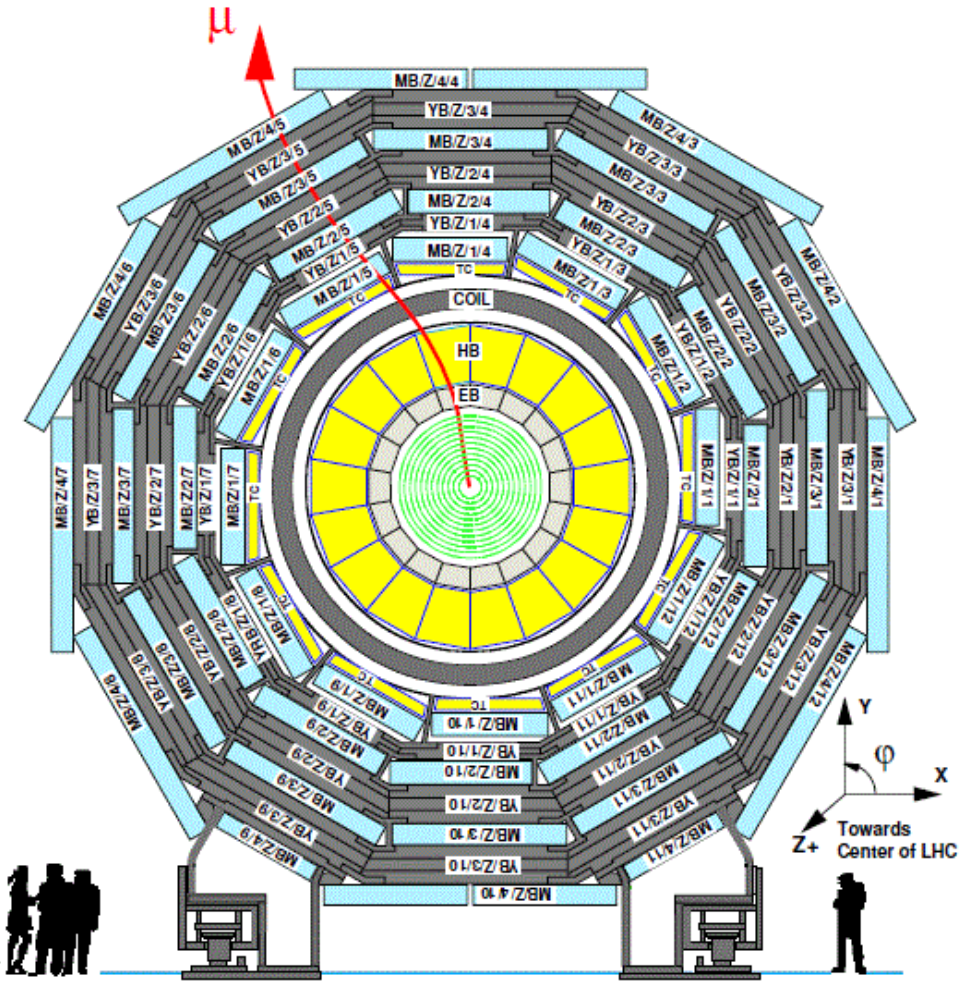

Figure 3.8: Drift Tube Geometry in the XY Plane. The DTs (blue) are stacked between the layers in the magnetic return yoke (Grey). Figure taken from [4].

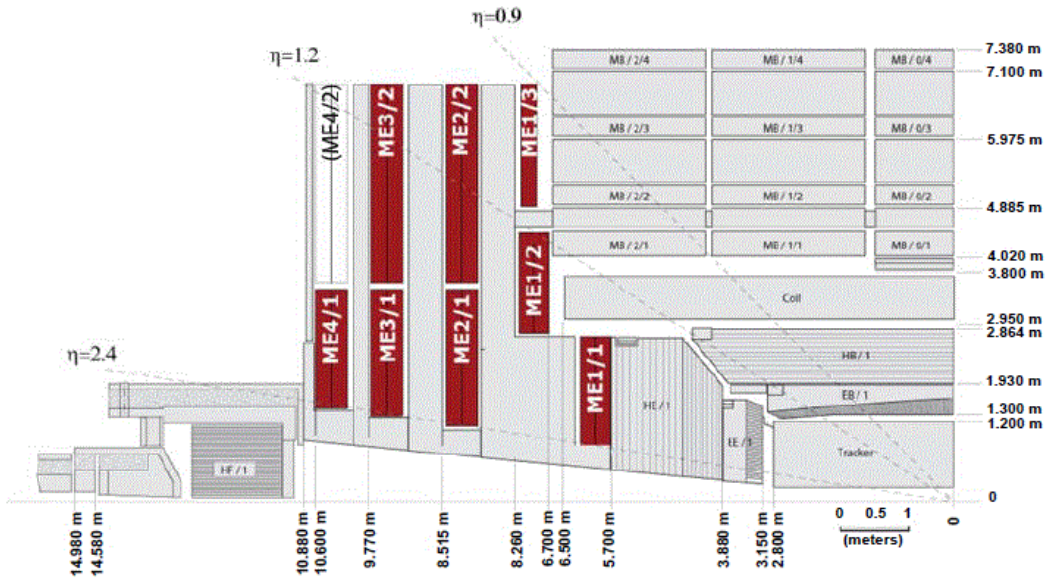

Figure 3.9: Cathode Strip Chamber Geometry in RZ plane. The CSCs are in red, while the DT and magnetic return yoke are in grey. Figure taken from [4].

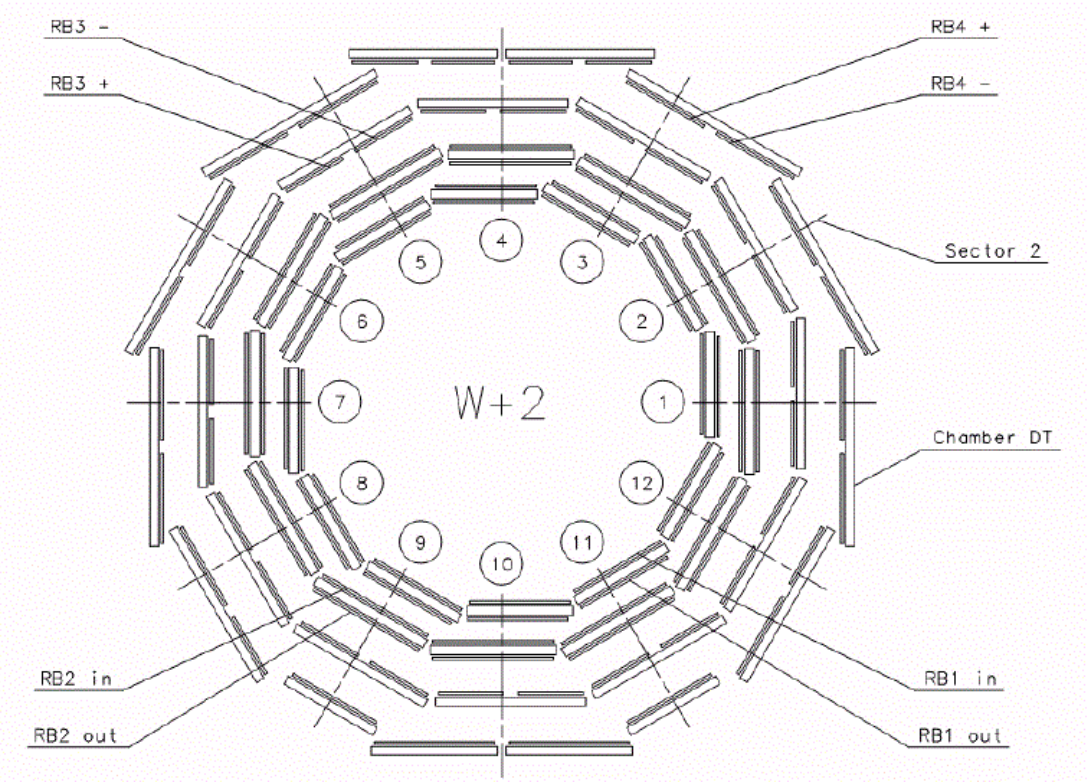

Figure 3.10: Resistive Plate Chamber geometry in XY plane. In the barrel region the RPCs are placed either in front of, or on both sides of, the DTs. While the endcap region the RPCs are mounted on both sides of all CSCs. Figure taken from [4].

Chapter 4

Theory

4.1 The Standard Model

The standard model (SM) of particle physics is the most well tested theory which describes the fundamental constituents of our universe to date. The standard model is described by the gauge theory of $SU(3)_C \times SU(2)_L \times U(1)_Y$. The $SU(3)_C$ portion describes the strong interactions of quarks and gluons which binds protons, neutrons and nuclei together. The C subscript denotes the color charges of the strong interaction. The $SU(2)_L$ group is part of the weak force which allows for weak decays such as those for neutrons and muons. The L subscript denotes that the weak force interacts only to left chiral particles. The $U(1)_Y$ group is the gauge theory which forms the electric force governing interaction of all electrically charged matter and photons. The Y subscript denotes the hypercharge, Y = 2(Q - T_3), where Q is the electric charge and T_3 is the z-component of weak isospin. The

electric and the weak force are unified into a single electroweak force by Salam, Glashow and Weinberg. This unification allows for spontaneous symmetry breaking which naturally generates masses through the Higgs boson for all particles except for photons and gluons using the Higgs mechanism[7].

Figure 4.4 left shows the contents of all elementary particles in the standard model. First to note are the spin-1 force carriers shown in green. γ is the photon which only carries the electric force; g is the gluon which carries the strong force; Z is the Z-boson which only carries the weak force; W is the W-boson which carries both the electric and weak forces. The spin-1/2 quarks which interact both electroweakly and strongly are listed in yellow. There are two archetypes of quarks, the up type and the down type. All up type quarks (up, charm and top) have an electric charge of +2/3, and all down type quarks (down, strange, bottom) have an electric charge -1/3. There are also U(1) charge conjugates of each quark which flips the electric charge. The spin-1/2 leptons are shown in red. Leptons interacts only electrically and weakly. There are also two archetypes of leptons, charged (electrons, muons, taus) with an electric charge of 1 and neutral (electron neutrinos, mu neutrinos, tau neutrinos) with an electric charge of 0. Note that there three generations of quarks and leptons of each archetype. This is the minimal number of generation required for the strong interaction to work.

Although, the standard model does not preclude a fourth generation, at this point a fourth generation quark or lepton has not be found.

The last but not least is H, which stands for the spin-0 Higgs particle. The addition of the Higgs field allows for spontaneous symmetry breaking[8]. Spontaneous symmetry breaking allows the Z and W-boson to be massive while keeping the photon massless [7] which are the properties we observed experimentally. As a spin-0 scalar with mass, the Higgs particle interacts with all other particle in loop correction of their respective propagators. This is how the Higgs generates masses for all other particles in the standard model. With the masses and coupling constants of the particles in the standard model one can predict the interactions of all fundamental particles using Feymann rules and diagrams.

As of this writing, an Higgs like particle has been observed at the LHC with a mass of 126 GeV which solidifies the predictive power of the SM. However, there are evidences that the SM is incomplete. Some of the problems with the SM are the hierarchy problem, unification of coupling constants, and existence of dark matter.

4.1.1 Hierarchy

In the standard model, the Higgs potential,

$$V = m_H^2|H|^2 + \lambda|H|^4 \tag{4.1}$$

must have a non zero minimum to create the Higgs' vacuum expectation value (VEV).

$$< H >= \sqrt{-m_H^2|H|^2/2\lambda} \approx 246GeV \tag{4.2}$$

$< H >$ places the Higgs mass at the electroweak scale of the order 100 GeV, but the Higgs couples to everything, which creates virtual loop correction to the Higgs mass, as shown in figure 4.1, from all fermions (top) and scalars (bottom) in the standard model where t can be any fermion where $\tilde{t}$ represents any scalar particle. The scalar loops increase the Higgs mass while the fermionic loops decrease the Higgs mass. Figure 4.1 only shows the leading order correction, all higher order contributions also modify the Higgs mass. Although fine tuning the loop correction factors in these two diagrams allows the SM to have an Higgs with mass near 100 GeV. This must be done to all higher order corrections which creates the hierarchy problem. The hierarchy problem does not break the SM but makes the the SM "disturbingly sensitive to the Higgs potential"[19].

4.1.2 Coupling constants

The SM is described by the gauge theory of $SU(3)_C \times SU(2)_L \times U(1)_Y$. The coupling constants in each gauge group changes as functions of the energy scale due to renormalization group (RG) cutoff scale, which is known as the running of the coupling constants. Figure 4.2 [20] top shows the coupling constants in the

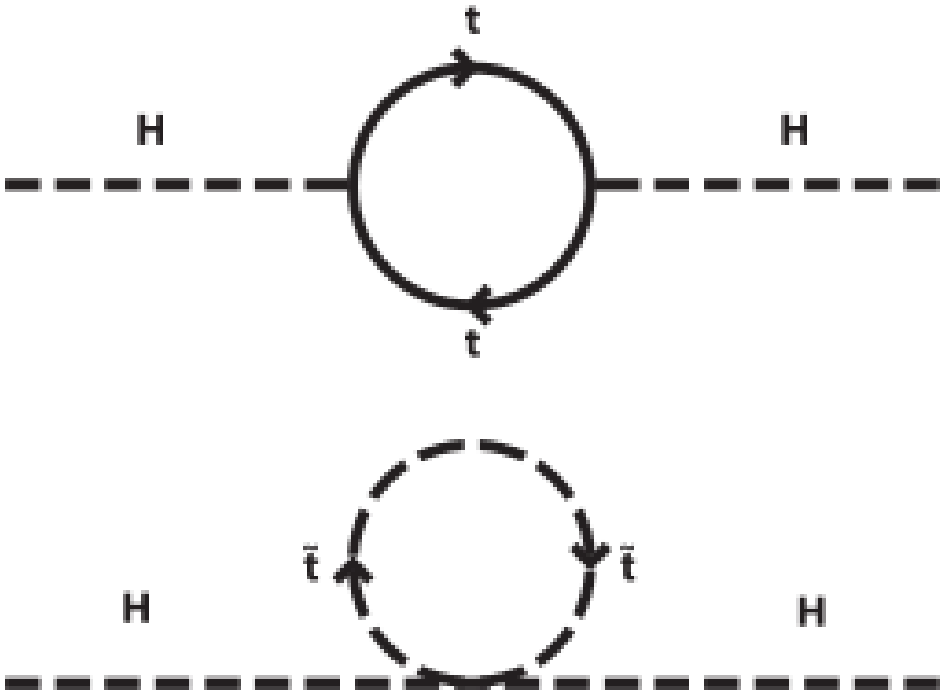

Figure 4.1: Higgs propagator loop correction. Top is the fermionic (t) loop correction to the Higgs mass. Bottom is the scalar ($\tilde{t}$) loop correction to the Higgs mass. Figure taken from [19].

SM as a function of the cut off scale. These three coupling constants does not unify at a high cut off scale, which will prevent the theory from satisfying any Grand Unification Theory (GUT). GUT is partly motivated by how electrons and protons charges cancel each other out exactly at a macroscopic level[9]. This is due to the fact that quarks have exactly 1/3 or 2/3 the charge of an electron which the SM does not explain. GUT unifies all three gauge theory under one theory

that makes $1/3 \times$ electron charge the elementary charge of all interactions, which requires the coupling constants between the gauge groups to unify at some energy level. This is not true in the standard model. The three coupling constants does not converge to a single point at any cut off scale as shown in top part of figure 4.2.

4.1.3 Dark Matter

In the SM, there is currently no prediction for the existence of dark matter in our universe. There are very strong evidences for dark matter's existence in our universe from astronomical observations. Direct detection has thus far eluded us at the time of this writing. The two strongest pieces of evidence can be found in the velocity of the galactic halo and estimation of the visible mass of nearby galaxies[21]. Figure 4.3 shows the rotational velocity of a nearby galaxy, NGC650. Fitting the actual observed rotational velocity to those modeled by gas, disk and dark matter halo composition shows that indeed an halo of unobserved mass exists. X-ray measurements also allow us to relate the amount of matter in another galaxy assuming hydrostatic equilibrium in spherical symmetry via the following equation from [21],

$$kT \approx (1.3 - 1.8)keV(\frac{M_r}{10^{14}M_{sun}})(\frac{1Mpc}{r}) \tag{4.3}$$

Where k is boltzmann's constant, T is the temperature, M_r is the mass of the galaxy as the function of radius r. Measurements of this function and gravitation lensing effect confirm that the visible mass of other nearby galaxies only makes up for about 10% of their total mass.

4.2 Supersymmetry

Motivated by the short comings of the standard model as described in the previous section, the SUperSYmmetric (SUSY) extension of the SM allows us to explain all three phenomena. The premise of SUSY is an added symmetry between fermions and bosons of our universe at a fundamental level. For each boson there exists a fermionic partner (Superpartner) allowing for the same number of degrees of freedom. For example, the top quark is a spin 1/2 fermion with two degrees of freedom. SUSY predicts there must exists two supersymmetric partners to the top quark (stop quark) each with spin 0. The particle zoo of SM and SUSY particles are shown in the diagram in figure 4.4. Note that even though only one stop quark is shown on the right, in the Minimal SuperSymmetric Models (MSSM) there will be two stop quark chiral states due to the number of degrees of freedom.

SUSY solves the hierarchy problem since for each fermion there exists a superpartner that will cancel out the loop correction from the Higgs mass ex-

actly. By adding SUSY particles to the RG calculations the coupling constants of $SU(3)_C \times SU(2)_L \times U(1)_Y$ in figure 4.2 bottom also become unified into one single force at the plank scale of 10^{16} GeV. The Lightest Supersymmetric Partner (LSP) from SUSY is a natural candidate for dark matter.

SUSY has a rich phenomenology due to its huge number of parameters. In the Minimal SuperSymmetric Models (MSSM), there are 120 new parameters with each generating a huge number of different signatures. In the following analysis, the results will be cast in two different SUSY frame work. The Minimal SUperGRAvity framework (mSUGRA) and Simplified Model Spectra (SMS)[22].

The details of mSUGRA is also described in Martin's SUSY Primer [19] and it has became the standard candle which SUSY analyses are measured against. mSUGRA is the first attempt to reduce the number of parameters in SUSY to 5 listed below [19]:

- m_0: Scalar SUSY particle parameter. This basically sets the starting point for mass of all SUSY scalar particles.
- $m_{1/2}$: Fermionic SUSY particle mass parameter. Specifically the Wino, Zino and photino the superpartners of W, Z and photon.
- $tan\beta$: Ratio between the coupling of the Higgs upper and lower doublet.
- A_0: Yukawa coupling constant to SUSY particles.

- $sign(\mu)$: Sign of Higgs mass parameter.

In CMS's SUSY searches for mSUGRA is done in the plane of m_0 versus $m_{1/2}$ with all other parameters set to a constant such as $tan\beta = 10$, $sign(\mu) > 0$, and $A_0 = 0$.

The other SUSY framework is the SMS [22] which casts SUSY searches in terms of phenomenological signature processes instead of theoretical parameters. SMS creates a set of SUSY particles that decays in a specific way in order to limit the vast number of signatures which SUSY can appear. The following analysis will be shown with limits for SMS-T1tttt model due to its b-rich phenomenology. T1tttt contains pair production of gluinos which decays into a pair of top quark and a LSP each as shown in figure 4.5. T1tttt is a natural SUSY model since this signature is mainly produced via an intermediary stop quark (the top's superpartner). Due to the top quark's large mass, the left and right chiral superpartners of the top quark, namely $\tilde{t_L}$ and $\tilde{t_R}$ will naturally have a large mass splitting. This will push one of the chiral state to a low mass that is reachable by the LHC.

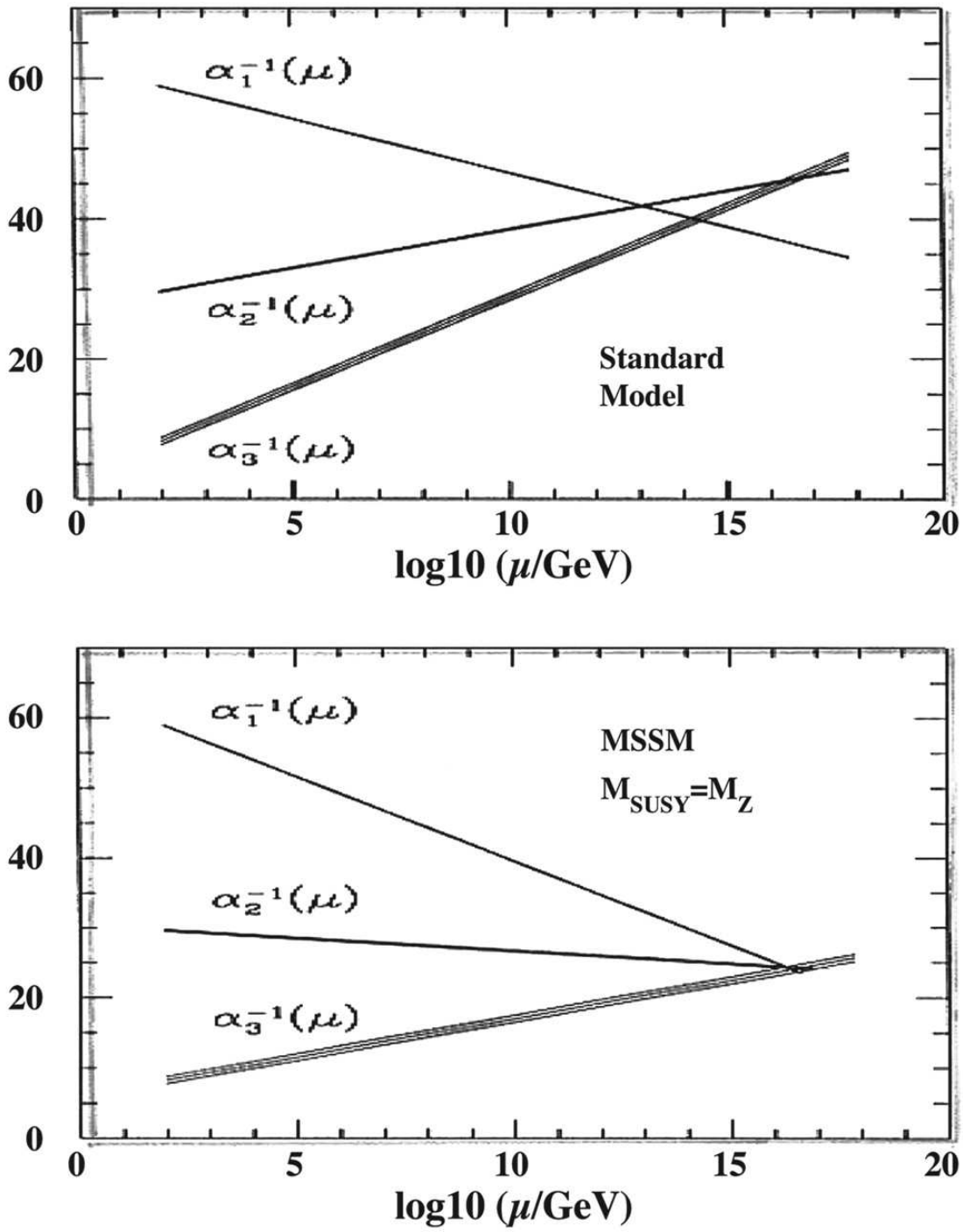

Figure 4.2: Running of the coupling constants. Top: α_1 corresponds to the U(1), α_2 corresponds to the SU(2), and α_3 corresponds to the SU(3). Where the x axis shows SM calculation for the cut off scale which the renormalization group is done. Bottom: Same as the top except adding one set SUSY particles (MSSM). Figure taken from [20].

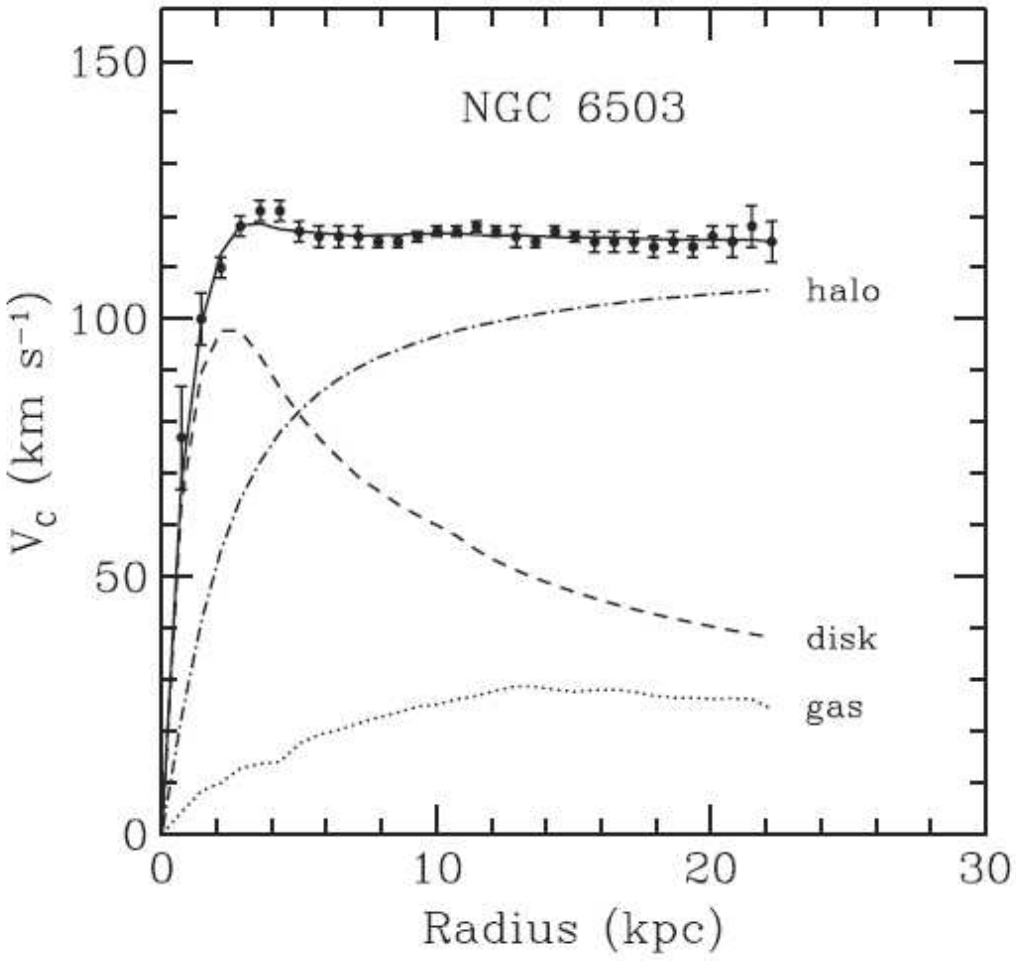

Figure 4.3: Rotation curve of a nearby galaxy (NGC 6503). Dot and dash lines are rotational velocities of gas, disk, and the dark matter halo fitted to the actual observed rotational velocity of the galaxy in solid points. Figure taken from [21].

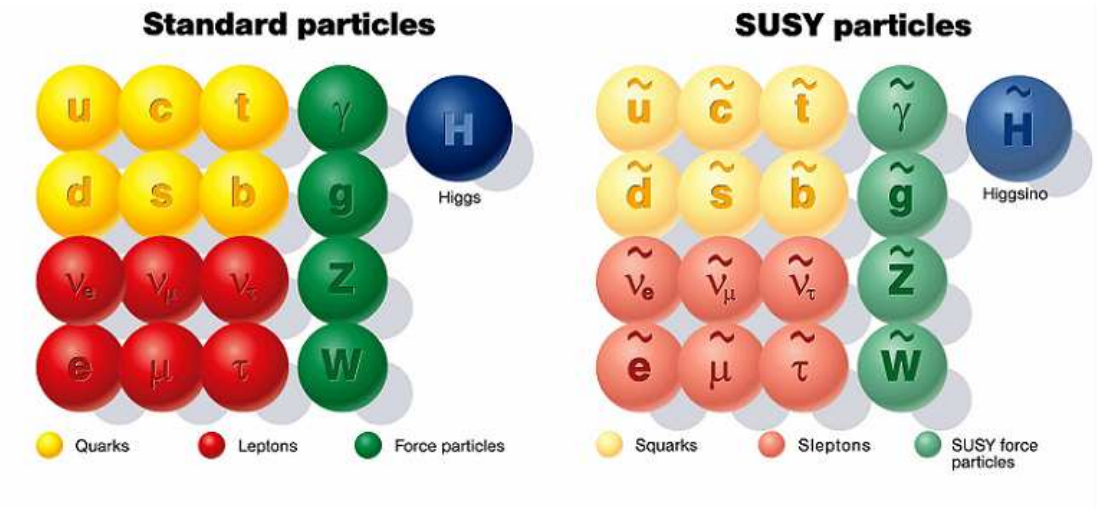

Figure 4.4: Particles of the SM and SUSY. Standard Model particles are shown on the left and SUSY particles are shown on the right with the tilda on top of them. Note that these partnerships are not 1 to 1 since the number of degrees of freedom must remain constant. Figure taken from [10].

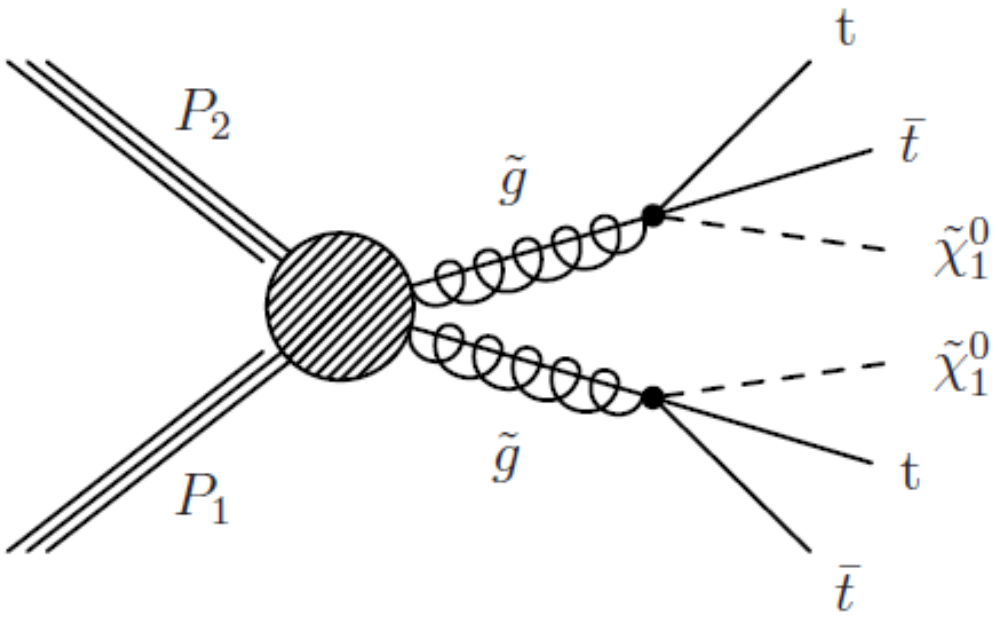

Figure 4.5: Feymann diagram for T1tttt. This particular signature produces a pair of gluinos which decays into two top quarks and an LSP each. Figure taken from [22]

Chapter 5

Search for Supersymmetry

Due to the vast theoretical parameter space of SUSY, the search for SUSY at CMS is broken up into signature reference analyses. This analysis is based on a signature of a single prompt lepton[1], two b-tagged jets with four or more total number of jets and $\not\!\!E_T$. The background composition of this search channel is primarily $t\bar{t}$ events. The single lepton requirement suppresses QCD events, while the requirement of two b-tagged jets suppresses the $W + jets$ and QCD backgrounds. The search is done in bins of $\not\!\!E_T$ since many SUSY models produce additional LSPs which are undetected by CMS, thus producing additional $\not\!\!E_T$ in an event.

The analysis presented in this thesis is based upon the lepton spectrum method [12][11], which predicts the $\not\!\!E_T$ tail of SM processes using the lepton p_T distribution in bulk. In CMS, prompt leptons are those decaying from W-boson and

[1]Note that in this analysis charged leptons are referred to as lepton and only include muons and electrons. Neutral leptons are referred to as neutrinos including all three generations.

Z-boson whose the mean lifetime is less than 10^{-24} second. In the single lepton case, only W boson decay will contribute to our signal and control sample since the analysis vetoes on events with a 2nd lepton. $t\bar{t}$ events decaying through a W-boson has an approximate branching fraction of 20% to decay leptonically on each side of the decay chain as shown in figure 5.1. Due to the V-A behavior of the weak force, the neutrinos will tend to have higher momentum than the leptons as shown in figure 5.9. This is a well understood and modeled effect in CMS Monte Carlo (MC) which is corrected for in the analysis using MC κ_{SL}-factors. The second component of the lepton spectrum method is the dilepton and tau prediction. The dilepton/tau prediction is made by emulating the $\not\!\!E_T$ distribution of these events using fully reconstructed dilepton events in data. The third component is the QCD contribution. QCD has a very small contribution even at the lowest $\not\!\!E_T$ bin, however QCD events contaminate our control sample and must be subtracted from the background prediction. The QCD contribution is estimated by extrapolating the combined relative isolation distribution. The last component is the Drell-Yan which is done at a MC level and found to be negligible.

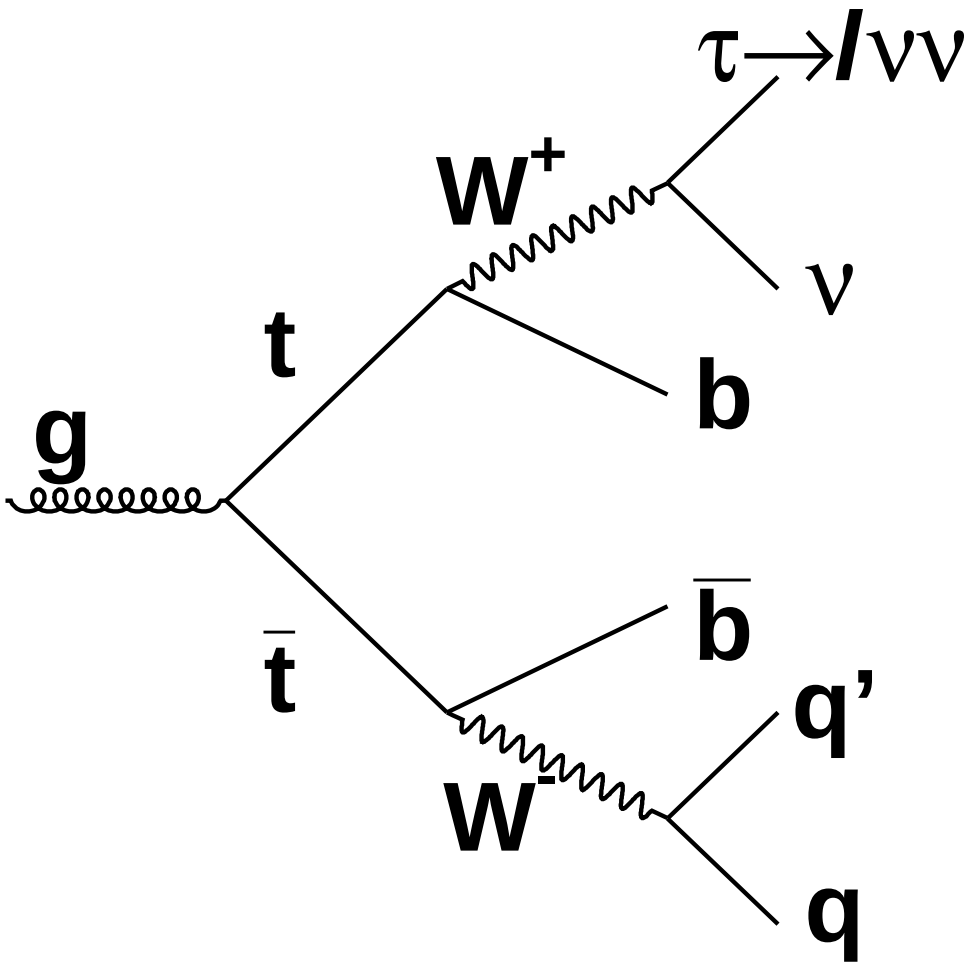

Figure 5.1: Feymann diagram of a semileptonic $t\bar{t}$ decay. The top half of the decay chain decayed leptonically producing both the lepton p_T spectrum and $\not\!\!E_T$. The bottom half of the decay chain decayed hadronically, which ideally has no contribution to $\not\!\!E_T$.

5.1 Monte Carlo Samples

CMS Monte Carlo is generated with various generators such as pythia, powheg, and madgraph[27][26]. The generators are integrated with CMSSW (CMS Soft-Ware) where detector simulation and reconstruction are done. MC samples are generated at a center of mass energy of 7TeV and processed using CMS SoftWare (CMSSW42X). These samples are listed in table 5.1. Due to the large number of pile-up events at the LHC, the number of pile-up events generated in the MC may not necessary agree to those found in the data. The pile-up effects are dealt with by reweighing the MC to have the same number of primary vertices as the data in 2011. Figure 5.2 and 5.3 show the number of primary vertices (NPV) before reweighing while figure 5.4 and 5.5 show NPV distribution after reweighing. The effects are quite small on the actual analysis since pile-up event generally does not affect high $\not\!\!E_T$ and lepton p_T events.

5.2 Datasets and Triggers

Dataset and triggering are combined to one section, since the data is parsed at the trigger level and distributed to various datasets for analysis purposes. This allows CMS to better coordinate the triggering requirements and get a smaller subsample of the data that interests the individual analysis groups. Table 5.2

Table 5.1: MC samples used in this analysis.

Object or process	Data set path
MC Samples:	
$t\bar{t}$	/TTJets_TuneZ2_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v2/AODSIM
W+jets	/WJetsToLNu_300_HT_inf_TuneZ2_7TeV-madgraph-tauola/Summer11-PU_S4_START42_V11-v1/AODSIM
Z+jets	/DYJetsToLL_TuneZ2_M-50_7TeV-madgraph-tauola/Summer11-PU_S4_START42_V11-v1/AODSIM
tW	/T_TuneZ2_tW-channel-DR_7TeV-powheg-tauola/Summer11-PU_S4_START42_V11-v1/AODSIM
	/Tbar_TuneZ2_tW-channel-DR_7TeV-powheg-tauola/Summer11-PU_S4_START42_V11-v1/AODSIM
t-channel	/T_TuneZ2_t-channel_7TeV-powheg-tauola/Summer11-PU_S4_START42_V11-v1/AODSIM
	/Tbar_TuneZ2_t-channel_7TeV-powheg-tauola/Summer11-PU_S4_START42_V11-v1/AODSIM
s-channel	/T_TuneZ2_s-channel_7TeV-powheg-tauola/Summer11-PU_S4_START42_V11-v1/AODSIM
	/Tbar_TuneZ2_s-channel_7TeV-powheg-tauola/Summer11-PU_S4_START42_V11-v1/AODSIM
QCD	/QCD_HT-250to500_TuneZ2_7TeV_pythia6/Summer11-PU_S4_START42_V11-v1/AODSIM
QCD	/QCD_HT-500to1000_TuneZ2_7TeV_pythia6/Summer11-PU_S4_START42_V11-v1/AODSIM
QCD	/QCD_HT-1000toInf_TuneZ2_7TeV_pythia6/Summer11-PU_S4_START42_V11-v1/AODSIM

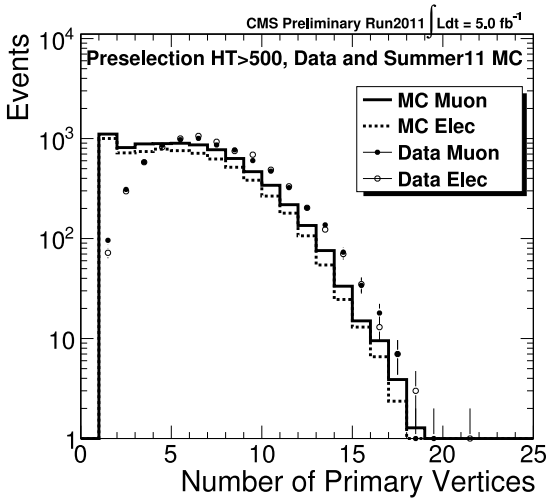

Figure 5.2: The number of primary vertices in Summer11 MC before pile-up reweighting.

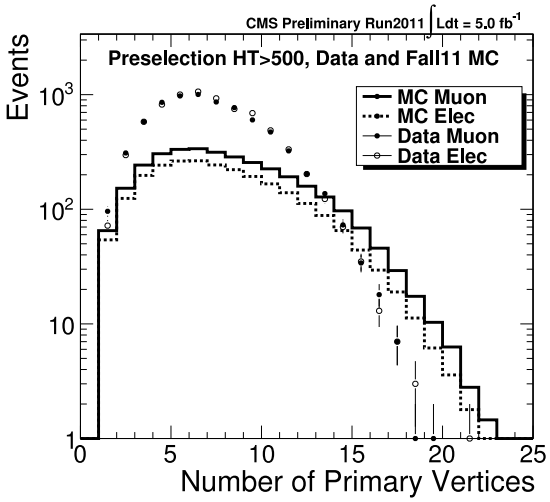

Figure 5.3: The number of primary vertices in Fall11 MC before pile-up reweighting.

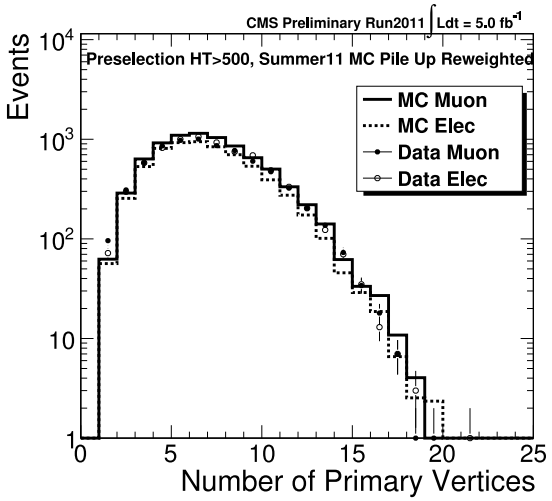

Figure 5.4: The number of primary vertices in Summer11 MC after pile-up reweighting.

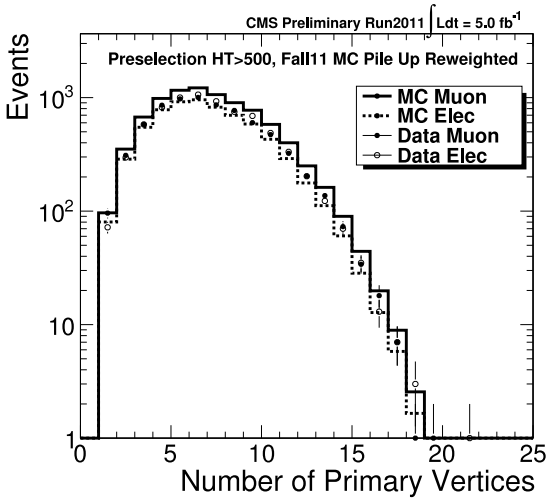

Figure 5.5: The number of primary vertices in Fall11 MC after pile-up reweighting.

shows the datasets used for this analysis. The MuHad and ElectronHad datasets are used to make the overall prediction, while the HT datasets are used for the $\not{E}_T$ smearing method in the single lepton prediction. Each event in the dataset is required to pass a trigger listed below within their respective dataset. The trigger requirement allows us to measure the trigger and reconstruction efficiency of our analysis cut. Every event in the MuHad and ElectronHad datasets is required to be unprescaled. Prescaled triggers are only used in the HT dataset since low HT triggers becomes prescaled quickly as the instantaneous luminosity of the LHC increased rapidly. In addition to the trigger requirement, a good luminosity section filter (JSON) is applied to every luminosity section of the datasets. The JSON filter is based on detector conditions during data taking at CMS. The data quality monitor (DQM) flags whether or not each subdetector is performing properly. Every subdetector must be flagged as good in order for a luminosity section to pass the JSON filter. The Golden JSON is used for all PromptReco events while the Aug05 and May10 ReReco correspond to the dataset labeled 05Aug2011 and May10ReReco.

- Muon triggers:
  - `HLT_Mu5_HT200`
  - `HLT_Mu8_HT200`

Table 5.2: Data samples analyzed in the note.

Object or process	Data set path
Datasets:	
Muons	/MuHad/Run2011A-May10ReReco-v1/AOD
Muons	/MuHad/Run2011A-PromptReco-v4/AOD
Muons	/MuHad/Run2011A-05Aug2011-v1/AOD
Muons	/MuHad/Run2011A-PromptReco-v6/AOD
Muons	/MuHad/Run2011B-PromptReco-v1/AOD
Electrons	/ElectronHad/Run2011A-May10ReReco-v1/AOD
Electrons	/ElectronHad/Run2011A-PromptReco-v4/AOD
Electrons	/ElectronHad/Run2011A-05Aug2011-v1/AOD
Electrons	/ElectronHad/Run2011A-PromptReco-v6/AOD
Electrons	/ElectronHad/Run2011B-PromtReco-v1/AOD
HT	/HT/Run2011A-May10ReReco-v1/AOD
HT	/HT/Run2011A-PromptReco-v4/AOD
HT	/HT/Run2011A-05Aug2011-v1/AOD
HT	/HT/Run2011A-PromptReco-v6/AOD
HT	/HT/Run2011B-PromptReco-v1/AOD

  - HLT_Mu15_HT200
  - HLT_HT250_Mu15_PFMHT20
  - HLT_HT250_Mu15_PFMHT40
  - HLT_HT300_Mu15_PFMHT40
  - HLT_HT300_Mu15_PFMHT50
  - HLT_Mu20_HT200 ($p_T(\mu)$ background control trigger)
  - HLT_Mu30_HT200 ($p_T(\mu)$ background control trigger)
  - HLT_Mu40_HT200 ($p_T(\mu)$ background control trigger)
  - HLT_Mu40_HT300 ($p_T(\mu)$ background control trigger)
  - HLT_Mu60_HT200 ($p_T(\mu)$ background control trigger)
  - HLT_Mu60_HT300 ($p_T(\mu)$ background control trigger)

- Electron triggers:
  - HLT_Ele10_CaloIdL_CaloIsoVL_TrkIdVL_TrkIsoVL_HT200
  - HLT_Ele10_CaloIdL_CaloIsoVL_TrkIdT_TrkIsoVL_HT200
  - HLT_Ele10_CaloIdT_CaloIsoVL_TrkIdT_TrkIsoVL_HT200
  - HLT_Ele15_CaloIdT_CaloIsoVL_TrkIdT_TrkIsoVL_HT200
  - HLT_Ele15_CaloIdT_CaloIsoVL_TrkIdT_TrkIsoVL_HT250

  - `HLT_Ele15_CaloIdT_CaloIsoVL_TrkIdT_TrkIsoVL_HT250_PFMHT25`
  - `HLT_Ele15_CaloIdT_CaloIsoVL_TrkIdT_TrkIsoVL_HT250_PFMHT40`
  - `HLT_Ele15_CaloIdT_CaloIsoVL_TrkIdT_TrkIsoVL_HT250_PFMHT50`
  - `HLT_HT350_Ele30_CaloIdT_TrkIdT` ($p_T(e)$ background control trigger)
  - `HLT_HT400_Ele60_CaloIdT_TrkIdT` ($p_T(e)$ background control trigger)
  - `HLT_HT450_Ele60_CaloIdT_TrkIdT` ($p_T(e)$ background control trigger)

- Jets triggers (artificial MET templates):
  - `HLT_HT300`
  - `HLT_HT350`
  - `HLT_HT400`
  - `HLT_HT450`
  - `HLT_HT500`
  - `HLT_HT550`
  - `HLT_HT600`
  - `HLT_HT650`
  - `HLT_HT700`
  - `HLT_HT750`

- JSON Filters:
  - Cert_160404-180252_7TeV_PromptReco_Collisions11_JSON.txt
  - Cert_160404-163869_7TeV_May10ReReco_Collisions11_JSON_v3.txt
  - Cert_160404-163869_7TeV_May10ReReco_Collisions11_JSON_v3.txt

5.3 Physics Object Reconstruction

Physics object reconstruction (RECO) in CMS is mainly done using the particle flow (PF) construct[25]. In this analysis all physics objects except muons and electrons utilize the particle flow reconstruction. There is also a cross check for particle flow version of the objects we use to ensure consistency in the overall event reconstruction.

5.3.1 Electrons

In CMS, electrons will create both tracks and deposit most of their energy into the Electromagnetic CALorimeter (ECAL) as shown in red on figure 5.6. Reconstruction is done by finding a seed which could be from an isolated ECAL energy deposit (ECAL driven) or tracks found in the pixel tracker (Tracker driven).[14] Electron tracks utilizes the Gaussian Sum Filter pattern recognition for track reconstruction which accounts for energy lost from bremsstrahlung[15]. Electrons

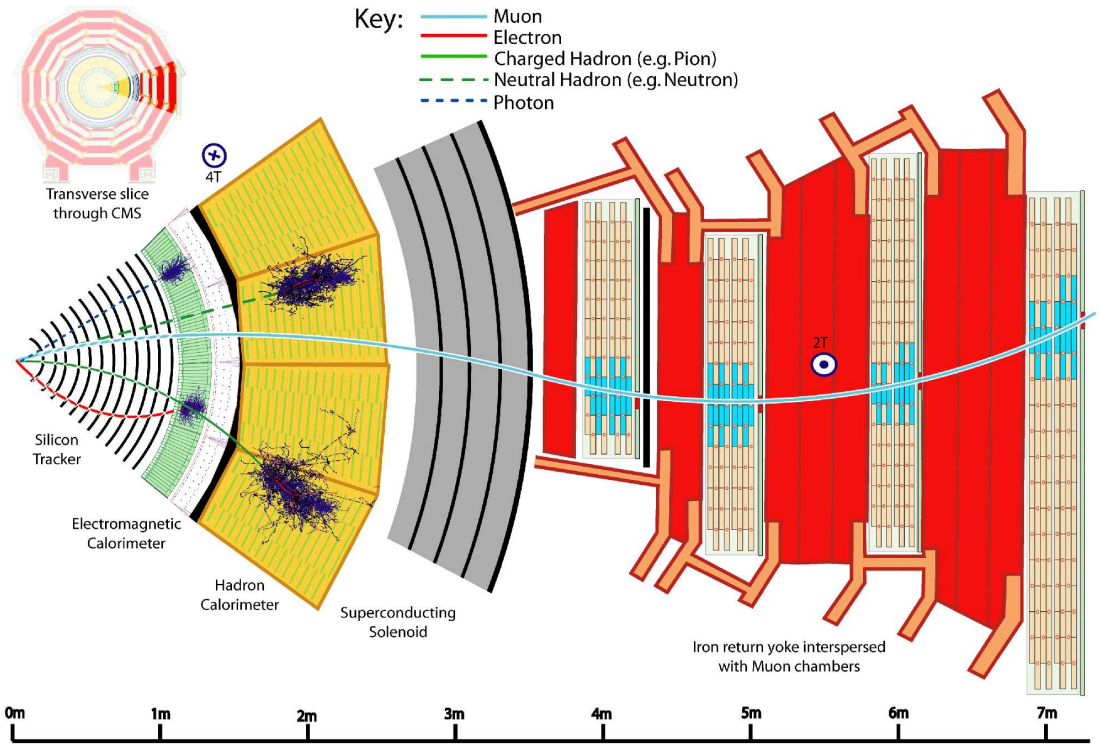

Figure 5.6: Slice of the CMS Detector in XY Plane. Where physics objects are shown in different colors. Muons are shown in blue, electrons are in red, charged hadrons ($\pi^{\pm}$ are in green, neutral hadrons (k_{long}) in dash green, and photons in dash blue. Figure taken from [5].

are identified at the physics object level using the EGamma Physics Object Group (POG) recommendation with the simple cut base algorithm[14]. The simple cut base algorithm uses a series of one dimensional cuts to reject fake electrons due to detector effects and decays in flight. There are two types of electrons in the analysis, Tight (for selection purposes, also known as WP70) and Veto electrons

(in order to veto on any additional leptons in the event, also known as WP90). These cuts are listed here:

- Electron p_T Transverse momentum of the electron. Low momentum electrons are more likely to be created by decay in flight and electron fake rate also increases at low momentum. Tight $\geq$ 20 GeV. Veto $\geq$ 15 GeV.

- Electron η Psuedorapidity of the electron. The fiducial coverage of the ECAL is limited to $|\eta < 2.5$. The region between the barrel and endcap often leads to mismeasurements and fakes. Tight and Veto: $|\eta| < 1.4442$ or $1.566 < |\eta| < 2.5$.

- Electron $|d0_{BS}|$: Impact parameter of the electron's track with respect to the beamspot. Prompt lepton decay should have small impact parameter where large impact parameter electrons are associated to those of decay in flight and conversions. Tight: $|d0_{BS}| < 0.02$ cm. Veto: $|d0_{BS}| < 0.08$ cm.

- Electron dZ(PV): The electron's track z position at the point of nearest approach with respect to the primary vertex. Same reason as above but the primary vertex is used because beamspot resolution along Z is horrible in hadron colliders. Tight: $|dZ(PV)| < 1.0$ cm. Veto: $|dZ(PV)| < 1.0$ cm.

- Conversion Rejection: missing hits: Any hits missing from the pixel tracker. Electrons from the interaction points should not have any missing hits as

it traverse through the pixel tracker. Missing hits are signs of a photon conversion in the pixel tracker. Tight = 0. Veto $\leq$ 1.

- Conversion Rejection: Dist and Δ Cot(θ): The distance of closest approach and difference in Cot(θ) with respect to all oppositely charged "good" tracks in the event where θ is the polar angle of the tracks. Since conversion results in a pair of electrons two tracks which originate from the same point in space. Tight: Dist $>$ 0.02 cm and Δ Cot(θ) $>$ 0.02. Veto: Dist $>$ 0.02 cm and Δ Cot(θ) $>$ 0.02. Conversion rejection requires both of these cuts to fail for the same electron.

- Combined Isolation: The sum of transverse energy in the tracker, ECAL, and HCal within a cone of $\Delta R < 0.3$ around the electron divided by the transverse momentum of the electron. Tight $<$ 0.07 (0.04) and Veto $<$ 0.10 (0.07). Where the values are for the barrel (endcap).

- Electron ID: $\sigma_{i\eta i\eta}$: This variable measures the width of the shower shape of the electron since a prompt electron should deposit most of energy in a signal ECAL crystal. Tight $<$ 0.01 (0.03) and Veto $<$ 0.01 (0.03). The difference between the barrel and endcap is driven by their respective resolution in η.

- Electron ID: $\Delta\phi$: The angular difference between the track propagated out to the ECAL and the center of the ECAL supercluster. Tight < 0.03 (0.02) and Veto < 0.80 (0.70).

- Electron ID: $\Delta\eta$: The psuedorapidity difference between the track propagated out to the ECAL and the center of the ECAL supercluster. Tight $<$ 0.004 (0.005) and Veto < 0.007 (0.009).

- Electron ID: HoE: The hadronic energy over the the electromagnetic energy deposited within the HCal tower where the electron's supercluster is located. Tight < 0.025 (0.025) and Veto < 0.12 (0.05).

5.3.2 Muon

Muons from W decay creates an isolated track as it traverses the tracker and a second track in the muon system after passing through both the ECAL and HCal shown in figure 5.6 in blue. Thus, a good fit between the two tracks both in momentum and direction allows CMS to identify muons coming from the interaction point. This is called the global fit for the global muon. Other muon reconstruction methods are also employed such as tracker muons where the object is seeded by the tracker and seeks out compatible muon segments in the muon system. In the

same way as electrons, there are two quality of muons used for the analysis Tight and Veto and their criteria are listed below:

- Muon p_T: Transverse momentum of the muon. Low momentum muons are more likely to be created by decay in flight. Tight $\geq$ 20 GeV. Veto $\geq$ 15 GeV.

- Muon η: Psuedorapidity of the muon. The fiducial coverage for muon is limited by η coverage of our tracker which only goes up to 2.4. Tight $\geq$ 20 GeV. Veto $\geq$ 15 GeV.

- Muon $|d0_{BS}|$: Impact parameter of the muon's track with respect to the beamspot. Prompt lepton decay should have a small impact parameter where large impact parameter electrons are associated to those of decay in flight. Tight: $<$ 0.02 cm. Veto: $|d0_{BS}| < 0.10$ cm.

- Muon dZ(PV): The muon's track z position at the point of nearest approach with respect to the primary vertex. Same reason as above but the primary vertex is used because beamspot resolution along Z is horrible in hadron colliders. Tight: $|dZ(PV)| < 1.0$ cm. Veto: $|dZ(PV)| < 1.0$ cm.

- Muon IDs: These are identification tag created by the muon POG. isGlobalMuon flags an muon created by the global fit algorithm. isTrackerMuon

flags a muon that is created by the tracker muon algorithm. Tight requires both to be true, veto only requires isGlobalMuon to be true.

- Muon Combined Relative Isolation, Iso_{Rel}^{Comb}: The sum of transverse energy in the tracker, ECAL, and HCal within a cone of $\Delta R < 0.3$ around the muon divided by the transverse momentum of the muon. Tight < 0.1 and Veto < 0.15.

- Muon Global χ^2 per degrees of freedom: χ^2 of the global fit between the inner tracks and the muon tracks of the muon trajectory. Tight < 10 and Veto no requirement.

- Muon Number of Valid Muon Hits and Stations: Number of valid hits in the muon track and the number of muon stations the muon track traverse. Tight: hits ≥ 1 and stations ≥ 2.

- Muon Number of Pixel Hits and Tracker layers: Number of hits in pixel detector and number of layers the tracker the muon track traverses. Tight: pixel hits ≥ 1 and layers ≥ 6.

5.3.3 Jets

Jets are created by the hadronization of highly energetic quarks or gluons. These partons are bounded by color confinement of the strong force which disallow

them to exist in free form. In high energy collisions, partons are fragmented into hadrons as the bounding energy between color anti-color pairs increases. These hadrons are the actual physical objects detected by the CMS detector. Due to the highly energetic nature of these quarks and gluons, they are collimated into a narrow cone of hadrons called jets.

The jet reconstruction algorithm used in this analysis is the Partical Flow (PF) algorithm. PF utilizes both tracking and calorimetric measurements to determine both the positions and energies of jets. Each individual jet is composed of different number of hadrons. Some of the hadrons are charged while others are neutral. Charged and neutral particles passing through the CMS calorimeter deposite different amount of energy in the ECAL and HCAL. PF associates tracks to the charged hadrons and neutral hadrons are those that have no track association. Separation of charged and neutral hadrons allows the reconstruction algorithm to apply different energy response in the ECAL and HCAL. This allows particle flow jets to have a much better resolution compared to traditional jets measured with only calorimetric detectors. The analysis uses the JETMET POG's recommendation of Loose ID [28] and jet energy correction [29] as itemized below:

- Jets p_T: Low energy jets often result from calorimeter noise, the threshold of 40 GeV is used also due to triggering paths which uses the same threshold. $p_T \geq 40$ GeV.

- Jet $|\eta|$: Limited by fiducial coverage of the tracker and HCal. $|\eta| < 2.4$.

- Jet Cone Size: The jet cone size used is 0.5 in $\eta - \phi$ space (ΔR).

- Jet ID: The JETMET POG loose ID allows us to reject fake jets from calorimetric noise while keeping high efficiency for real jets. The cuts are all applied to every jets. Neutral/Charged EM Fraction, Neutral Hadron Fraction < 0.99. Charged Hadron Fraction > 0. Charged Multiplicity > 0. (Charged Multiplicity + Neutral Multiplicity + Muon Multiplicity) > 1.

- Jet Energy Corrections (JEC): The MC and Data has three different JEC applied. L1Fastjet which corrects for pile up effects. L2Relative which is a scale factor applied to each jet. L3Absolute which is a constant amount of energy added to each jet as a function of p_T and η of the jet. The data also has Residual correction applied which corrects for the difference found between the data and MC after data taking.

5.3.4 $\not\!\!E_T$, Missing Transverse Energy and HT:

Real $\not\!\!E_T$ is created when weakly interacting particle passing through every layer of CMS detector go undetected. The neutrino is the prime example of the source of $\not\!\!E_T$, LSP created by SUSY models is another example. Fake $\not\!\!E_T$ is due to mismeasurement of the detector. This could be due to resolution of the calorimeters or due to gaps in the detector where particles escape detection. Just as particle flow improves the resolution of jets, particle flow is used to improve the resolution of $\not\!\!E_T$. $\not\!\!E_T$ vector is constructed out of the negative vector sum of all particle flow objects including muons and electrons with $p_T > 10$ GeV. HT is calculated as the scalar sum of all jets with $p_T > 40$ GeV with the identification shown in the jet subsection.

5.3.5 b-tagged jets:

This analysis utilizes the Combined Secondary Vertex (CSV) b-tag algorithm documented in the btagging and Vertexing paper[17]. CSV algorithm combines the information of reconstructed secondary vertices along with track-base lifetime information. The information used to construct the likelihood are vertex distance from primary vertex, the vertex's mass, the number of tracks, the fraction of energy in tracks relative to the jet, the $\Delta\eta$ of tracks relative to the jet, the 2D and 3D impact parameter of the tracks, and the number of the tracks. The likelihood

output the value of the CSV discriminator. This analysis uses a medium working point of 0.679, which has a light quark and gluon jets fake rate of approximately 1%. The charm quark fake rate is much higher at about 20% for this work point. The overall of b-tagging efficiency for b-quark jets is about 75% average of all jet p_T range.

5.4 Event Selections

The baseline selection is done to target high signal efficiency while isolating the analysis' main background contributor, $t\bar{t}$. Each event is filtered using recommendation by the JETMET POG[18] for the Run 2011 data. Each filter is applied successively and they are itemized below.

- Scraping Veto: Scraping events are those which the beam scraps the detector elements in CMS. This creates large $\not\!\!E_T$ but very few high quality tracks in the event. The veto ejects events by requiring the fraction of high quality tracks in each event to be greater than 25% if there are at least 10 track in the event.

- Primary Vertex (PV): The Primary Vertex is where the primary hard scattering occurs. The quality of the primary vertex is also importance since many other cut variables are based relative to the PV. The PV is required

to have at least 4 degrees of freedom, and maximum of (2, 24) cm in (dxy, dz) the transverse and longitudinal positions.

- HBHE Noise: Noise from the HCAL sometimes creates fake or greatly mismeasured events. This filter is based upon an isolation where isolated HCAL hits are more likely to be from noise rather than real hadronic showers. At the analysis level, this filter simply returns true or false, which is used to filter out HBHE noise events.

- Tracking Failure: Tracking failures occur when a large number of clusters are created in the tracker, which the track fitter is unable to reconstruct the individual tracks properly. We require the event to have at least 10% of its track to be flagged as high quality in order to remove the tracking failure events.

- CSC Beam Halo: Beam Halo comes from the proton's collision with either residual air molecules in the beam pipe or the tail of the beam profile colliding with the material in the accelerator. This will produce tracks in the CSC detector nearly along the beam direction. The filter rejects beam halo events by looking for these tracks in the CSC detector.

- Dead Ecal Cell: Some of the ECAL cells have broken data link which are considered dead. These cells are masked for reconstruction purposes. Some

of these cells have still trigger primitive links, which can be read out, while others do not. For the cells with trigger primitive we can measure the amount of energy loss within these dead cells. For the cell without trigger primitive a boundary energy filter is done where the amount of energy near the dead cell is used to approximate the energy loss in the dead cell. Events which have more than 10 GeV of energy lost in the dead cells are rejected.

- Lepton Muon or Electron: Each event is required to have 1 and only 1 "Tight" muon or "Tight" electron. Events with additional "Veto" leptons are rejected.

- jets and b-tagged jets: Each event has at least 4 or more jets, two or more must be b-tagged using the CSVM tagger.

- HT bins: HT is the scalar sum of the transverse energies in all the jets with p_T above 40 GeV in each event. The two bins are $500 \leq HT < 750$ and $HT \geq 750$ GeV.

- $\not\!\!E_T$ bins: There are four $\not\!\!E_T$ bins for each HT bin. This creates a total of 8 signal bins for this analysis. These bins are $\not\!\!E_T \in$ [150,250), [250,350), [350,450), [450,∞).

5.5 Single lepton prediction

The single lepton $\not\!\!E_T$ prediction is made using the muon and electron p_T distributions known as the lepton spectrum method[12]. $\not\!\!E_T$ and lepton p_T distribution differs due to three different effects. The difference between $\not\!\!E_T$ and lepton p_T resolution. Polarized W-bosons impart more momentum to the neutrino ($\not\!\!E_T$) than to the charged leptons. Leptons must also pass a p_T quality cut of 20 GeV where $\not\!\!E_T$ does not have this requirement.

The $\not\!\!E_T$ distribution is smeared out differently compared to the lepton p_T distribution. Figure 5.7 compares the difference between the neutrino's p_T to the actual $\not\!\!E_T$ measured in single lepton event. Under ideal conditions they should be identical. The two differ due to resolution effects, mismeasurements and other neutrinos produced by heavy flavor decays. These effects are corrected by using $\not\!\!E_T$ templates derived in situ from HT datasets. $\not\!\!E_T$ templates are taken from HT datasets in the HT trigger path from HLT_HT300 to HLT_HT750 in 50 GeV increments. L1 and HLT prescales are accounted for by weighing the prescaled events properly. Each event going into the template is required to have 0 leptons and binned into 50 GeV bins between HT of 500 to 750, then 750 to 850, 850 to 1000 and 1000 to infinity. A coarser binned version of the $\not\!\!E_T$ template is shown in figure 5.8, which shows the $\not\!\!E_T$ distribution's tail increases as a function of

HT. For each event (lepton) in our background control sample, the lepton's p_T is smeared out using the template's weight in each bin in $\not\!\!E_T$. The smeared lepton p_T is added into the total lepton distribution with weights and uncertainties that attribute to one event in a gaussian fashion.

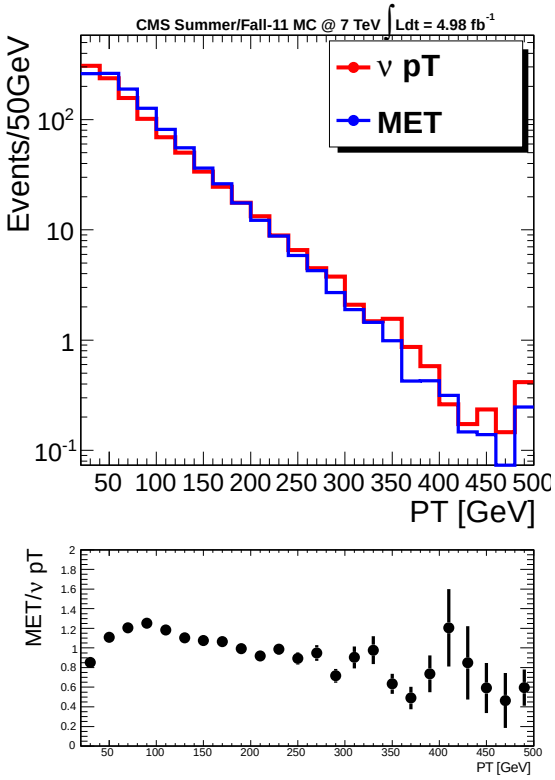

Figure 5.7: Comparison of neutrino p_T and $\not\!\!E_T$. Due to detector resolution effects, $\not\!\!E_T$ tends to have a larger tail than real neutrino p_T in single lepton events.

Due to polarization effects the W-boson's decay, the neutrino and lepton p_T distribution differs in case of polarized W decays. This effect is shown in

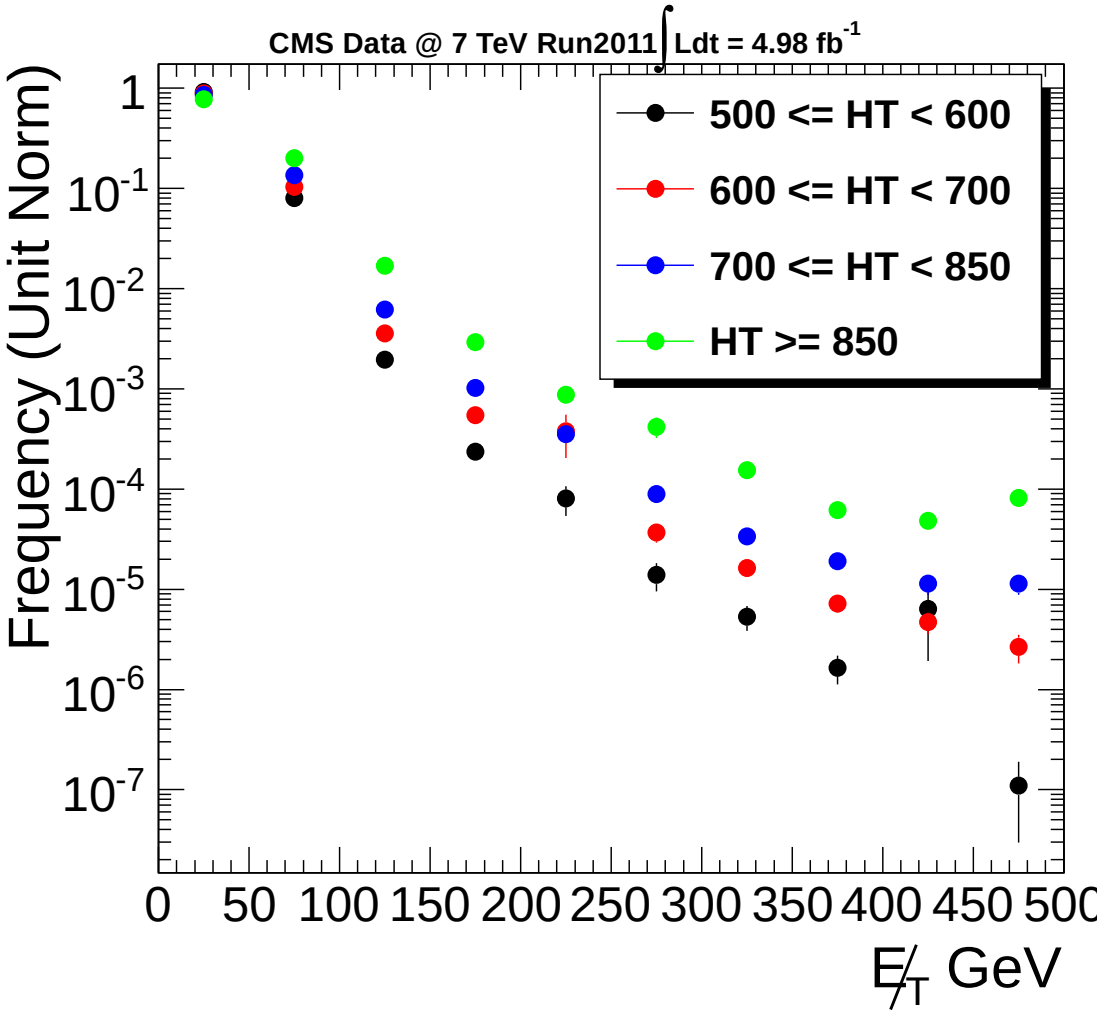

Figure 5.8: Samples of $\not\!\!E_T$ templates taken from data. Note that $\not\!\!E_T$ smearing is done in 50 GeV bins in HT. As the HT of the event increase so does the $\not\!\!E_T$ tail which is due to resolution effects.

5.9 where the neutrino will tend to have higher p_T than the charged lepton's p_T in both $t\bar{t}$ and $W + jets$ events. The lepton p_T threhold of 20 GeV also causes over prediction due to loss of low p_T lepton which corresponds to high

neutrino p_T events. These two effects are well understood based upon physics and reconstruction of events. They are corrected using MC single lepton κ factors. The MC single lepton κ factor is defined as,

$$\kappa_{SL} = \frac{\text{N TRUE SL in } \not{E}_T}{\text{N RECO SL in } p_T} \tag{5.1}$$

The numerator denotes the number of true single lepton events from MC infor-

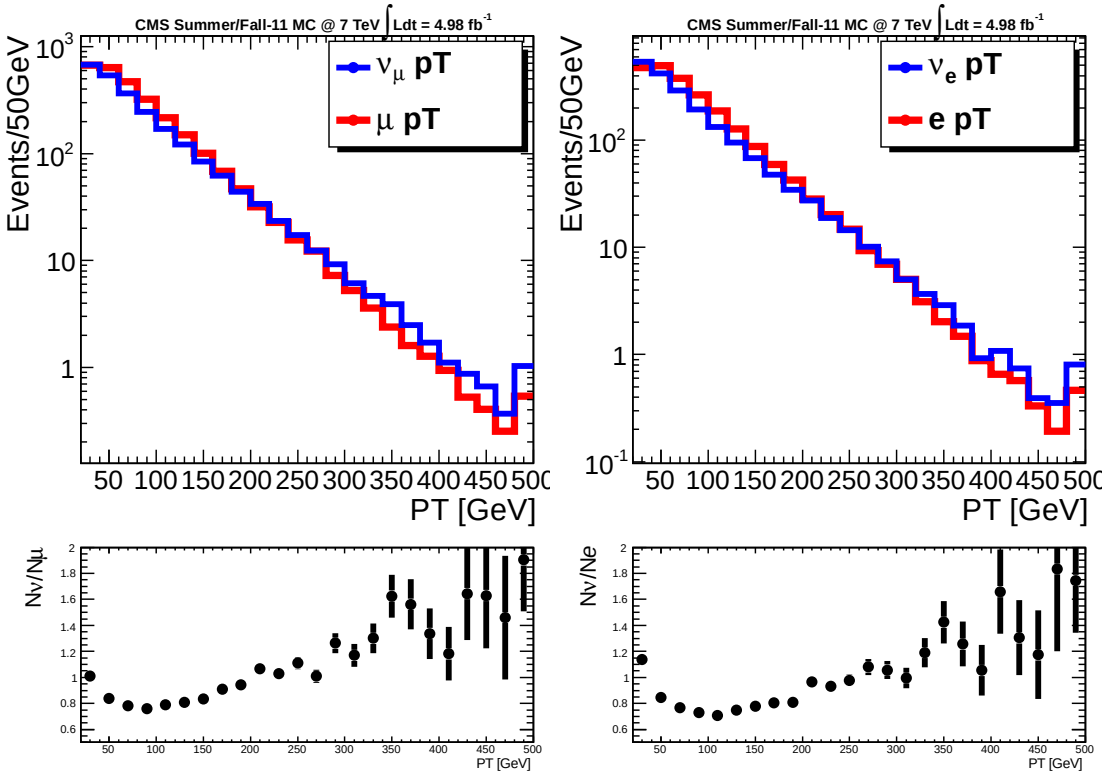

Figure 5.9: Neutrino p_T (blue) versus Lepton p_T (red) in Single Lepton MC. The ratio between the two distribution has the characteristic swoosh in both electrons and muons. This is due to lower neutrino p_T events migrating upward due to polarization effect.

mation in a particular bin of $\not{E}_T$. The denominator denotes the number of single

lepton events from using only reconstruction level information. This definition allows the analysis to decouple the single lepton and dilepton prediction in $\not\!\!E_T$, while keeping the dilepton in the lepton p_T bins which is small and not predicted separately. The κ_{SL} factors are shown in bins of HT and $\not\!\!E_T$ on figure 5.10. Combination of the $\not\!\!E_T$ smearing and κ_{SL} factor allows us to make the $\not\!\!E_T$ prediction using the lepton p_T spectrum in $t\bar{t}$ and $W + jets$ events. Figure 5.11 shows the smearing and κ_{SL} applied to MC in red. Since this is a MC correction applied to a MC the prediction the results come out exactly the same.

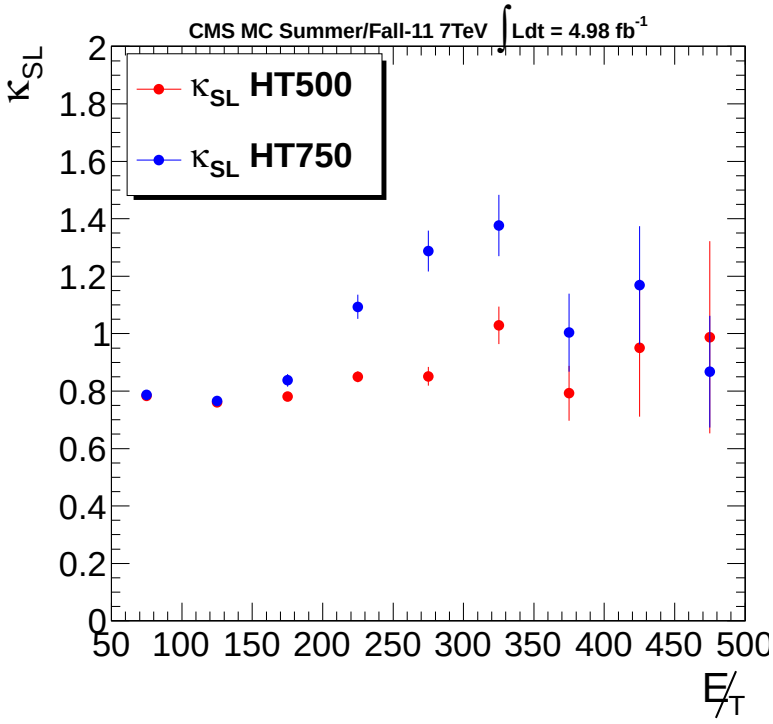

Figure 5.10: κ_{SL} as defined by Equation 5.1. The HT $\in$ [500,750) GeV bin is in red, while HT $\in$ [750,∞) is in blue.

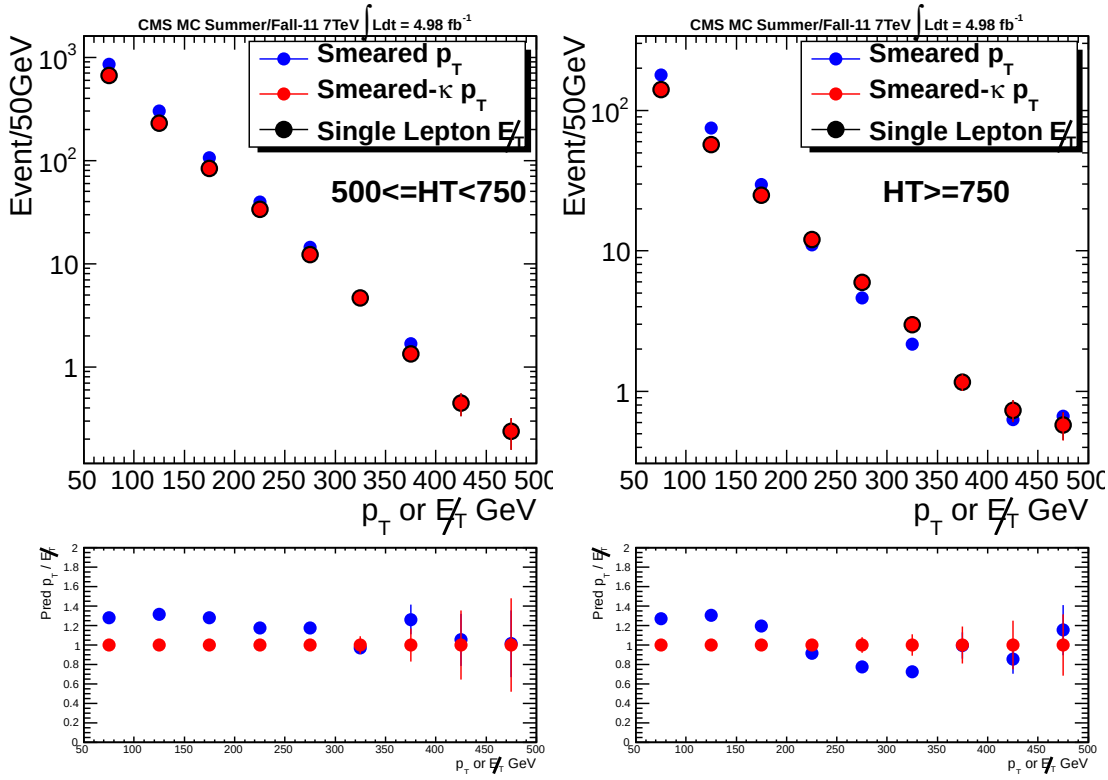

Figure 5.11: MC Closure in Single Lepton Prediction. For HT $\in$ [500,750) GeV (left) and HT $\in$ [750,∞) (right) the smeared lepton p_T spectrum is corrected using κ_{SL}, which creates the identical distribution since it is a MC to MC correction base on the same events.

Using the $\not\!\!E_T$ templates derived in data and κ_{SL} factors derived in MC, the prediction for the single lepton contribution to the $\not\!\!E_T$ distribution is made in data. First, the smearing is applied to the lepton p_T spectrum as shown in figure 5.12 with both muons and electrons combined for $500 \leq HT < 750$ on the left and $HT \geq 750$ on the right. The κ_{SL} factors are applied to the smeared lepton p_T distribution in 50 GeV bins on the plots and 100 GeV bins on the table.

Figure 5.13 shows the single lepton predictions in data for $500 \leq HT < 750$ and $HT > 750$ GeV bins. The results are also tabulated on table 5.3 and 5.4. The systematics are included in these tables which are evaulated in the next subsection.

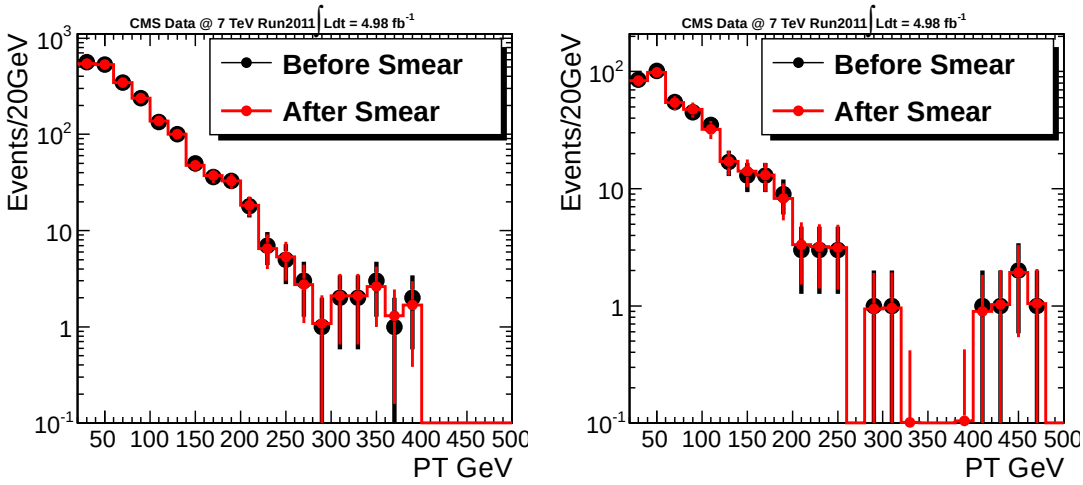

Figure 5.12: Comparison of lepton PT with and without smearing.

In case of the HT $\in$ [500,750) and $\not\!\!E_T > 450$ bin, the control sample has zero events as seen in figure 5.12. Having zero events in the control sample creates an ambiguous statistical uncertainty for the single lepton prediction. The statistical uncertainty in this bin is handled by extrapolating lepton $p_T$spectrum from 250 to 400 GeV using an exponential function

$$F = C \times e^{\alpha} \tag{5.2}$$

The fit parameters are determined by using a χ^2 goodness of fit method. The fit results in $C = 52.6 \pm 6.6$ and $\alpha = -0.07 \pm 0.06$. By integrating this function from

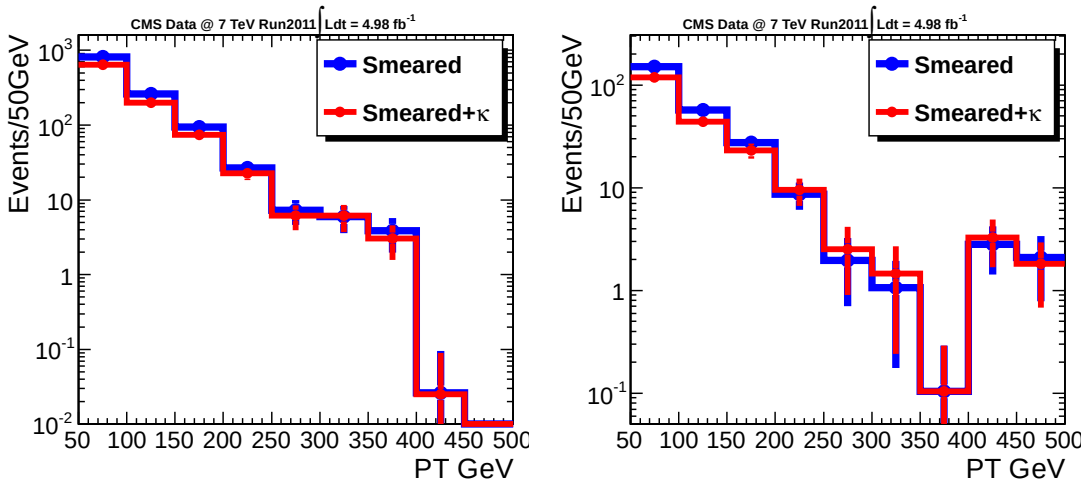

Figure 5.13: Single lepton only prediction in data. This is only the single lepton portion of the $\not\!\!E_T$ distribution so no comparison is made between the prediction and $\not\!\!E_T$ at this point.

450 to ∞, the extrapolation shows there could be 2.4 ± 1.5 raw events for lepton $p_T >$ 450 GeV. Since the lepton p_T smearing is done using binned histograms only, fractional events are added into bins between 450 to 900 GeV in 50 GeV bins where the fractions are weighted by the exponential fit function to give a total event weight of 2.4 events. The lepton p_T smearing and κ_{SL} is applied to these extrapolated events in 50 GeV bins. The total number of events after extrapolation, smearing and κ_{SL} factors is equal to 1.9 ± 1.1 events for HT $\in$ [500,750) and $\not\!\!E_T$ $>$ 450 bin. The central value of 1.9 is used as an upper limit on the statistical uncertainty of our prediction for this bin while the uncertainty due to the fit is added in quadrature into the systematic uncertainty.

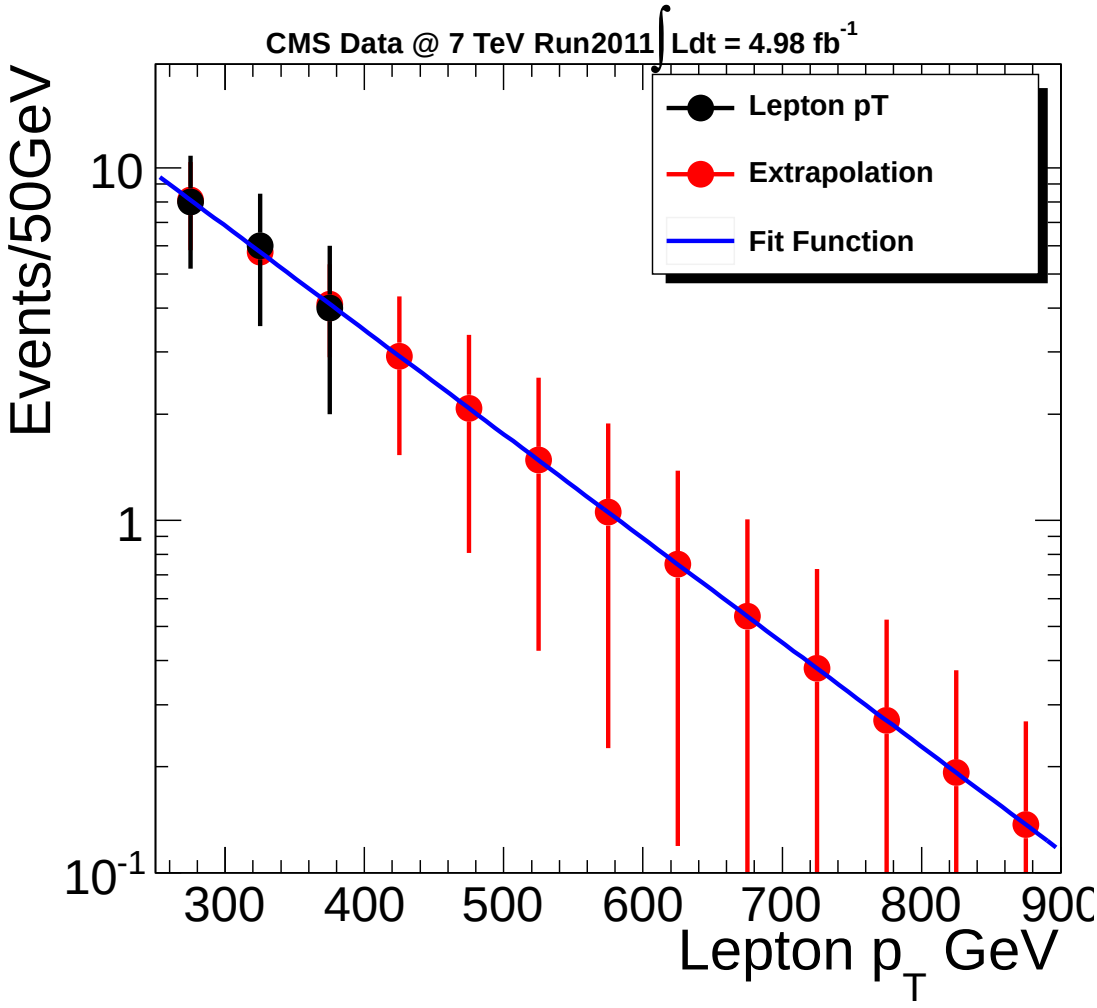

Figure 5.14: Extrapolation of lepton $p_T$for determining the uncertainties in the HT $\in$ [500,750) and $\not\!\!E_T$ $>$ 450 bin. Black points are lepton p_T spectrum from data. Red points are extrapolated points using an exponential function.

5.5.1 Single lepton systematics

The systematic uncertainties in the single lepton part of the prediction are itemized below:

- **Jet and $\not\!\!E_T$ energy scale** This is one of the dominant systematics in the single lepton prediction where the energies of jets hence $\not\!\!E_T$ is mismeasured by a certain scale. The uncertainties are modeled by varying the uncertainties in the energy of every jet in an event and recalculate the $\not\!\!E_T$ due to this change. This is the standard recipe given by the JETMET group in CMS.

- **MC κ_{SL} factor statistics** The other dominant uncertainty is due to the statistical uncertainties of the κ_{SL} factor. Due to limited amount of MC events available, the MC statistics in higher $\not\!\!E_T$ bins depreciate quickly. This uncertainty is taken as a systematic since it will not change with additional data but additional MC can reduce this uncertainty.

Table 5.3: Single Lepton Prediction $500 \leq HT < 750$

$\not\!\!E_T$ [GeV]	MC Expectation	κ_{SL}	Data Smeared	Data Prediction
150-250	116.8 ± 1.4	0.80 ± 0.01	121.6 ± 10.8	$96.8 \pm 8.6 \pm 13.8$
250-350	16.9 ± 0.5	0.93 ± 0.04	13.2 ± 3.5	$12.3 \pm 3.3 \pm 1.5$
350-450	1.78 ± 0.20	0.79 ± 0.10	3.89 ± 1.89	$3.08 \pm 1.49 \pm 0.74$
450+	0.24 ± 0.08	0.99 ± 0.33	0.00 ± 1.94	$0.00 \pm 1.94 \pm 0.92$

Table 5.4: Single Lepton Prediction $HT \geq 750$

$\not\!\!E_T$ [GeV]	MC Expectation	κ_{SL}	Data Smeared	Data Prediction
150-250	37.0 ± 0.8	0.90 ± 0.03	36.2 ± 5.3	$32.5 \pm 4.76 \pm 2.1$
250-350	8.95 ± 0.40	1.32 ± 0.06	3.02 ± 1.54	$3.99 \pm 2.03 \pm 1.77$
350-450	1.90 ± 0.20	1.16 ± 0.12	2.92 ± 1.39	$3.39 \pm 1.62 \pm 0.99$
450+	0.58 ± 0.13	0.87 ± 0.19	2.08 ± 1.30	$1.81 \pm 1.13 \pm 0.71$

- $\sigma(t\bar{t})$ **and** $\sigma(W)$**:** Uncertainty due to $W + jets$ cross section is measured using Drell-Yan events where two leptons are required to reconstruct a Z mass within 20 GeV. The difference between the data and MC is taken as the uncertainty in the $W + jets$ cross section. The $t\bar{t}$ cross section is measured by selecting single lepton events and reconstructing as $t\bar{t}$ χ^2 where χ^2 is defined in Equation 5.11 The difference between data and MC with various χ^2 cuts is taken as the uncertainty in the $t\bar{t}$ cross section. The details are found in Section 5.8.

- W **polarization in** $t\bar{t}$**:** The $t\bar{t}$ polarization uncertainty is handle by reweighing the events in MC to have either $\pm 5\%$ variation in the fraction of W helicity state in $\lambda(W^+) = -1$ and $\lambda(W^+) = 0$. This variation is 10 times larger than those found in theoretically fitted uncertainties. The details are described in section 5.9.

- W **polarization in** W**+jets:** The $W + jets$ polarization uncertainty is handled in a similar way as the $t\bar{t}$ polarization uncertainty. However, the variations are broken up in three different bins in rapidity $\in$ [0-1), [1,2], and [2,5). Then the variation is done in three different ways. 1) $\pm$ 10% change to $f_{-1} - f_{+1}$ in both W^+ and W^-. 2) $\pm$ 5% change to f_{-1}, f_{+1} for only W^+ or only W^-. 3) $\pm$ 10% change to the polarization fraction in longitudinal

direction, f_0, for W^+ and W^-. The maximal of these three variation is taken to reweigh the MC events to produce the systematic uncertainties. The details are described in section 5.9.

- **Lepton efficiency:** The lepton efficiency was measured using tag and probe methods in Z events. The lepton efficiency was found to be between 1% to 5%[30]. The analysis uses a flat ± 5% uncertainty to vary the prediction in MC to reproduce this uncertainty.

The results of all the single lepton prediction systematic uncertainties are listed in the table 5.5 and 5.6 for $500 \leq HT < 750$ and $HT \geq 750$ GeV in bins of $\not\!\!E_T$ respectively. These systematics are those listed in the final results in table 5.3 and 5.4 in an absolute fashion. These systematics are propagated to the full (single lepton , dilepton/tau, and QCD) prediction assuming they are uncorrelated.

5.6 Dilepton and Tau Prediction

Due to the significant contribution to the $\not\!\!E_T$ tail from $t\bar{t} \rightarrow b\bar{b}l\nu l`\nu`$ (dilepton) and $t\bar{t} \rightarrow b\bar{b}\tau\nu_\tau X$ (tau) events, a separate data driven method is designed to predict their contributions to our signal regions[12]. The method predicts each of the dilepton and tau contribution individually (5 total). They are combined at the end to obtain the final dilepton/tau prediction.

Table 5.5: Systematic uncertainties for Single lepton prediction for $H_T \in [500, 750)$ GeV in the search $\not{E}_T$ bins. Each uncertainty is expressed as the percentage (%) change in the ratio of predicted to the true number of events (evaluated in Monte Carlo) in the single lepton prediction.

E_T^{miss} :	$[150; 250)$	$[250; 350)$	$[350; 450)$	≥ 450 GeV
SL $\not{E}_T$ and jet energy scale	12.9	9.2	19.9	20.9
SL MC κ Stat	1.0	3.9	9.5	23.4
SL MC κ x-sec	0.7	0.8	4.1	5.4
SL Polarization $t\bar{t}$	3.4	5.6	6.3	7.9
SL Polarization $W + jets$	0.1	0.5	1.2	3.8
SL Lepton Efficiency	5	5	5	5
Single Lepton Total	14.3	12.5	23.9	33.4

Table 5.6: Systematic uncertainties for Single lepton prediction for $H_T \geq 750$ GeV in the search $\not{E}_T$ bins. Each uncertainty is expressed as the percentage (%) change in the ratio of predicted to the true number of events (evaluated in Monte Carlo).

E_T^{miss} :	$[150; 250)$	$[250; 350)$	$[350; 450)$	≥ 450 GeV
SL $\not{E}_T$ and jet energy scale	2.0	1.3	20.1	32.4
SL MC κ Stat	2.5	6.3	19.8	19.4
SL MC κ x-sec	1.3	0.2	2.7	7.6
SL Polarization $t\bar{t}$	2.6	4.0	5.0	5.2
SL Polarization $W + jets$	0.3	0.2	0.6	0.3
SL Lepton Efficiency	5	5	5	5
Single Lepton Total	6.6	9.1	29.2	39.2

The dilepton contribution comes from two categories, ignored leptons and lost leptons. Ignored leptons are those detected and reconstructed by the CMS detector, however failed our lepton quality cuts, thus ignored in our analysis. Lost leptons are those which completely escaped our detector due to fiducial coverage

or detector inefficiency. The tau contribution comes in three different components. Single tau events where the tau decays leptonically; lepton plus tau event where the tau decay hadronically; lepton plus tau event where the tau decay leptonically but fails our lepton quality cuts. All five contributions are measured using data distribution which are corrected by Monte Carlo κ_{DL} and α factors that are defined below. The ignored leptons $\not\!\!E_T$ distribution prediction is described in steps at first. The other four types of dilepton/tau events use a similar method. So only the differences in their methods will be described in detail.

First, the ignored leptons are dilepton events from $t\bar{t} \rightarrow b\bar{b}l\nu l`\nu`$ or $DY \rightarrow l\bar{l}$ events in which the second lepton is detected and reconstructed, but does not pass our analysis lepton veto requirement. In this case, the entire event is assumed to be reconstructed correctly and no further corrections to the event's $\not\!\!E_T$ is needed.

- Step 1: In data, the $\not\!\!E_T$ distribution for ignored lepton events are predicted by real dilepton events where the number of reconstructed good leptons is required to be exactly two instead of one. All three channels contributes to the predictor, $\mu\mu$, ee and μe.

- Step 2: In MC, the same predictor distribution is created with identical cuts.

- Step 3: In MC, the $\not\!\!E_T$ distribution for ignored lepton is made using MC truth information. The MC truth ignored lepton event criteria are as follow,
  - a) event passing the analysis cuts in MC with a single lepton.
  - b) a second reconstructed lepton with no requirement is matched to a MC truth lepton from a W decay within a ΔR of 0.1

  This distribution is shown in figure 5.15 top left ($500 \leq HT < 750$) and right ($HT \geq 750$) in red.

- Step 4: The distribution from step 2 is normalized to same number of events as the distribution from step 3 with $50 < \not\!\!E_T < 150$. This factor is call the α factor. The number of events found by having two leptons vastly differ from the actual ignore lepton rate. The α factor allows the MC to tell us the effects of branching ratio $\times$ acceptance of ignored lepton events, while the shape of ignored lepton $\not\!\!E_T$ distribution is still taken from data. This distribution is shown in figure 5.15 top-left in red.

- Step 5: The distribution from step 4 is further corrected bin by bin (50 GeV width) using the MC truth distribution. This factor is call the κ_{DL} factor even though the same symbol is used for tau prediction. The κ_{DL} corrects for any bias in the method within the information contained from the Monte Carlo. Some of these biases derive from selecting good lepton in

step 1 and step 2, while not doing so in step 3. The statistical uncertainty due to limited MC sample size is taken as a systematic uncertainty in the method. The κ_{DL} for ignored lepton is shown in 5.15 bottom-left in blue.

- Step 6: Both α and κ_{DL} factors are applied to the data dilepton $\not\!\!E_T$ distribution from step 1. The final data prediction is shown in figure 5.15 top-left in black. The data to MC prediction ratio is shown in figure 5.15 bottom-left in black.

Second, the prediction for lost lepton is very similar to ignored leptons. Lost leptons are dilepton events in which the second lepton escapes detection or is not reconstructed due to inefficiency. In lost lepton events the $\not\!\!E_T$ must be corrected to to account for the lost of lepton. This is modifies the distribution in step 3 of the ignored lepton with a $\not\!\!E_T$ distribution with a corrected $\not\!\!E_T$. The corrected $\not\!\!E_T$ distribution is found by taking the vector difference of the event's reconstructed $\vec{\not\!\!E}_T$ and lepton vector. The step 3b requirement is also inverted since the second lepton is never reconstructed. The distribution for lost leptons are shown the same schema in figure 5.16.

Third, the prediction for lepton plus a leptonic tau (lep tau lep) is described. In this decay channel, one of the two W boson decays into a tau which then decay leptonically show in the equation below.

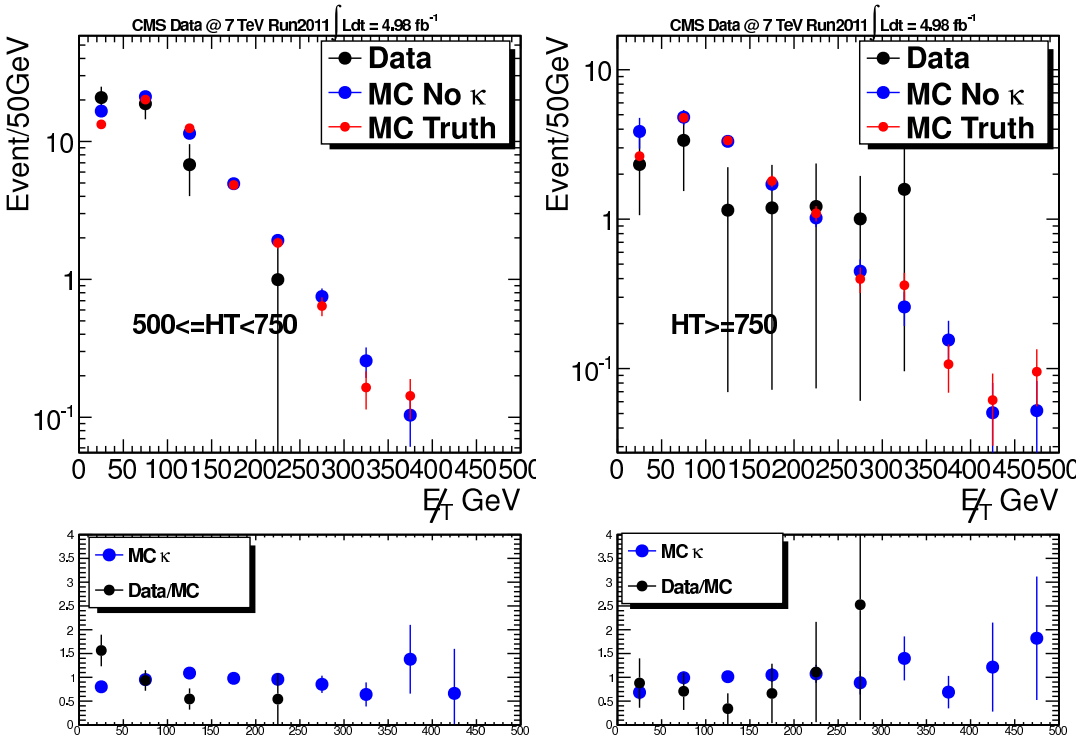

Figure 5.15: Top Left: $\not\!\!E_T$ prediction for ignored lepton for $500 \leq HT < 750$. Black points show the final prediction from data; blue points show the prediction in MC before κ factors are applied; red points show the MC truth $\not\!\!E_T$ distribution. Top right: $\not\!\!E_T$ prediction for ignored lepton for $HT \geq 750$. Bottom row: κ factors and data/MC ratios for $500 \leq HT < 750$ (left) and $HT \geq 750$ (right)

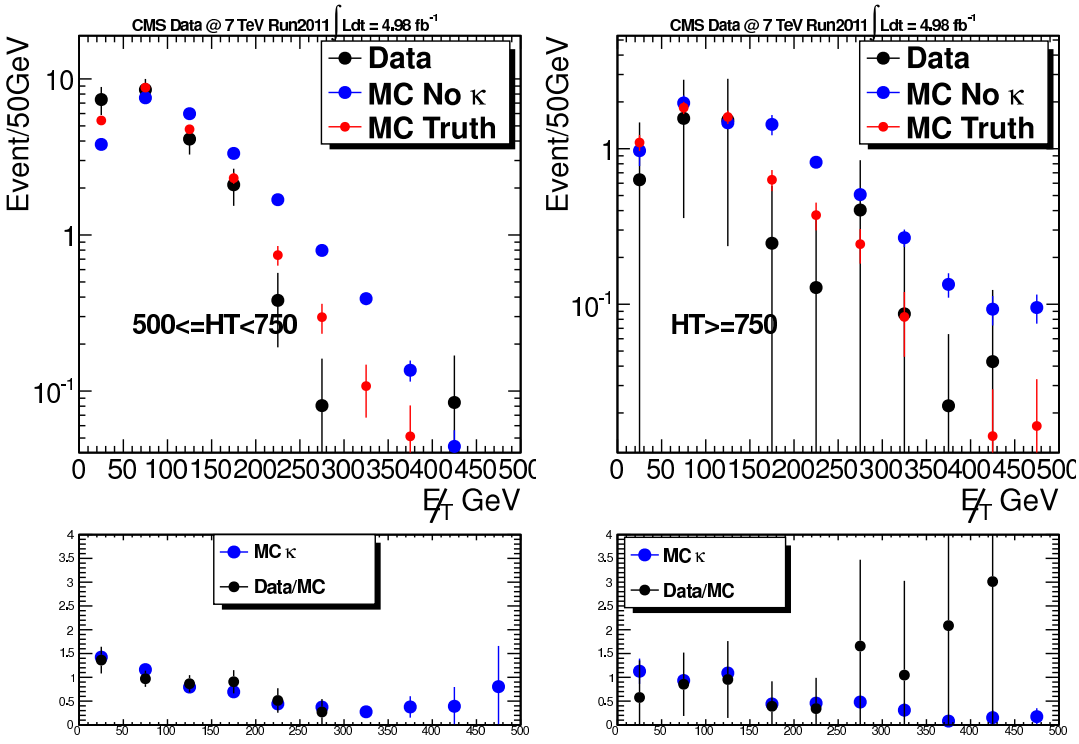

Figure 5.16: Top Left: $\not\!\!E_T$ prediction for lost lepton for $500 \leq HT < 750$. Black points show the final prediction from data; blue points show the prediction in MC before κ factors are applied; red points show the MC truth $\not\!\!E_T$ distribution. Top right: $\not\!\!E_T$ prediction for lost lepton for $HT \geq 750$. Bottom row: κ factors and data/MC ratios for $500 \leq HT < 750$ (left) and $HT \geq 750$ (right)

$$W^+ \rightarrow \bar{\tau}\nu_\tau \rightarrow \bar{\nu_\tau} \bar{l} \nu_l \nu_\tau (\text{or charge conjugates}) \tag{5.3}$$

This decay channel is model using dilepton (ll) events found in data by turning one of the two leptons into a tau decaying leptonically. The $\not\!\!E_T$ distribution in lep tau lep and ll differs mainly due to the two additional neutrinos presence in the lep tau lep decay. ll to lep tau lep transformation is done using a leptonic tau response (r_{lep}) function derived from MC. r_{lep} is defined as the following,

$$r_{lep} = \frac{p_T(visible\ lepton)}{p_T(original\ tau)} \tag{5.4}$$

Where p_T (original tau) is the p_T of a tau identified to decay leptonically using MC truth information and the visible lepton is the electron or muon decay product of the tau at MC truth level. This distribution is created in bins of p_T of the original tau shown on the left side of figure 5.17. For each ll event, one lepton at a time is turned into a leptonic tau by sampling the r_{lep} function randomly 1000 times. The $\not\!\!E_T$ vector of the lep tau lep event is then recomputed,

$$\vec{\not\!\!E}_{T_{leptaulep}} = (1 - r_{lep}) \times \vec{p}_{Tl} + \vec{\not\!\!E}_{Tll} \tag{5.5}$$

Statistical uncertainties of resampling the same event 1000 times is done by taking each of the 1000 samples as a whole to have a total event weight of 1. The

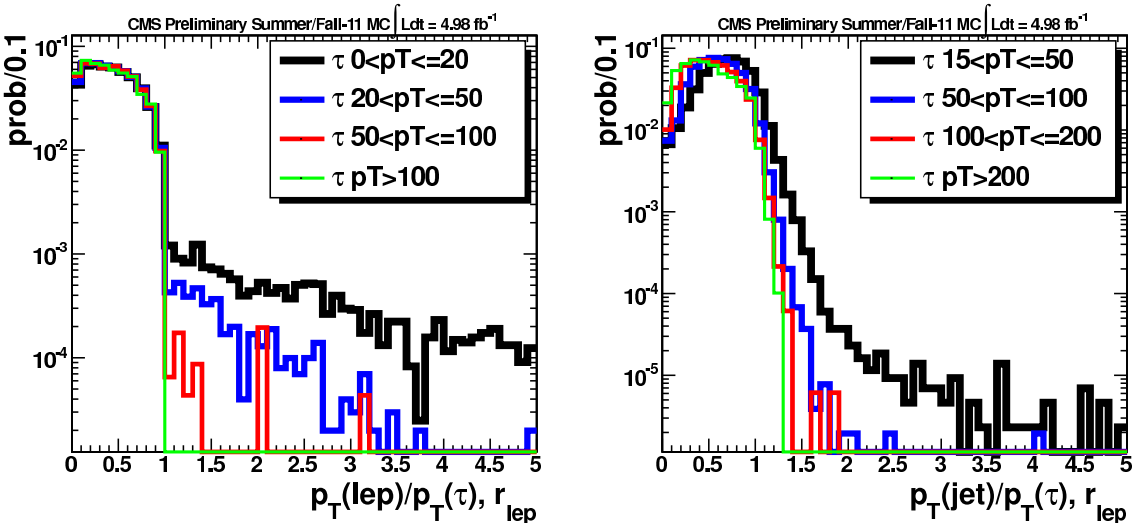

Figure 5.17: Left: Leptonic Tau Response. The histograms are for taus with $p_T \in$ [0,20)(black), [20,50)(blue), [50,100)(red) and 100+(green). Right: Hadronic Tau Response. The histograms are for taus with $p_T \in$ [15,50)(black), [50,100)(blue), [100,200)(red) and 200+(green).

1000 samples produces a new $\not\!\!E_T$ distribution with a spread similar to that of the response function. This procedure creates the $\not\!\!E_T$ distribution in step 1 and 2 just as the ignored lepton case. The $\not\!\!E_T$ distribution in step 3 is taken from MC truth information as it was done previously. The same steps are followed to create α and κ_{DL} to create the prediction in data as shown in figure 5.18.

The fourth contribution is the lepton plus an hadronic tau (lep tau had) channel as shown in the equation below.

$$W^+ \rightarrow \bar{\tau}\nu_\tau \rightarrow \bar{\nu_\tau}\nu_\tau \; jet(\text{or charge conjugates}) \tag{5.6}$$

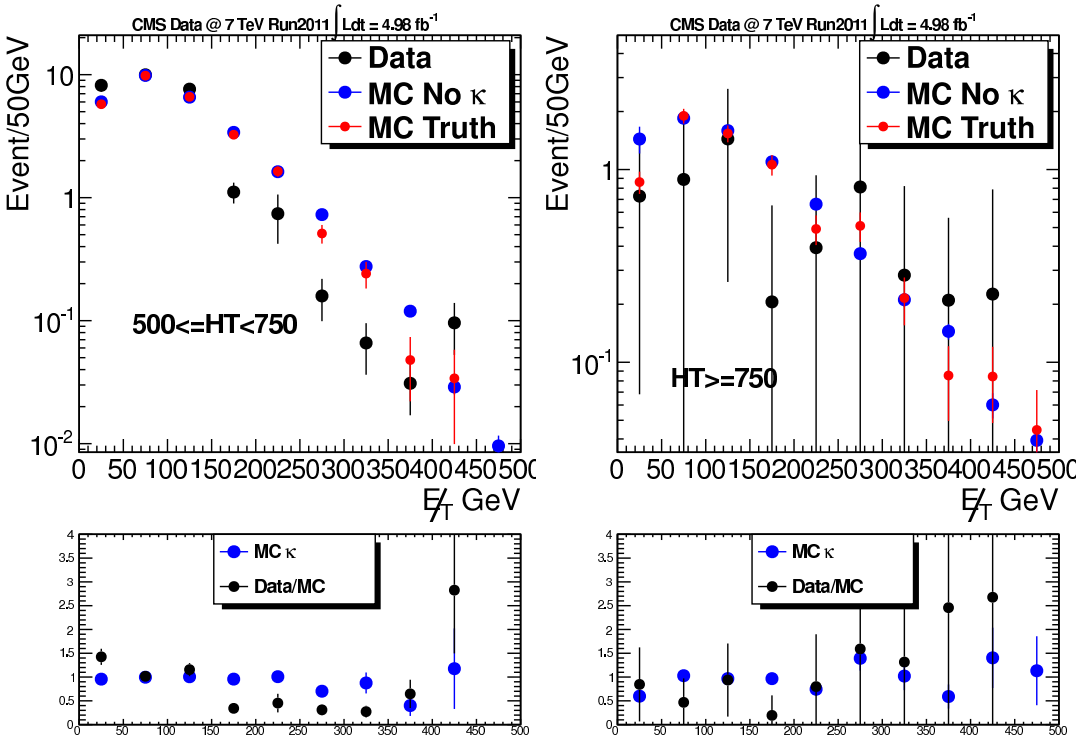

Figure 5.18: Top Left: $\not\!\!E_T$ prediction for lepton plus leptonic tau for $500 \leq HT < 750$. Black points show the final prediction from data; blue points show the prediction in MC before κ factors are applied; red points show the MC truth $\not\!\!E_T$ distribution. Top right: $\not\!\!E_T$ prediction for lepton plus leptonic tau for $HT \geq 750$. Bottom row: κ factors and data/MC ratios for $500 \leq HT < 750$ (left) and $HT \geq 750$ (right)

This channel is treated in a similar away as the lep tau lep channel except an hadronic response function is used. The response function (r_{had}) is created by selecting lep tau had events at the MC truth level. In this case, the resulting p_T of the jet is used instead of the lepton. This is done by taking a jet within a $deltaR$ of 0.3 of the hadronic tau as shown below,

$$r_{had} = \frac{p_T(jet_{\Delta R<0.3})}{p_T(original\ tau)} \tag{5.7}$$

The distribution is shown in figure 5.17 right, and each lepton in the event sample the r_{had} function randomly 1000 times. Once again, the $\not\!\!E_T$ of each event is modified using the r_{had} function in the following equation,

$$\vec{\not\!\!E}_{T_{leptauhad}} = (1 - r_{had}) \times \vec{p}_{Tl} + \vec{\not\!\!E}_{T_{ll}} \tag{5.8}$$

The statistics of the sampling is handle the same way as lepton plus a lepton tau case. Since the one of the two lepton is turned into a tau jet, the preselection in step 1 and 2 requires only 3 jets in the event. If the tau jet's p_T is over 40 GeV, the tau jet is added to the event's H_T and jet count, otherwise the event is not modified. The final 4 jets and H_T selection is applied after the lepton is turned into a tau jet. The procedure flows normally from this point and the resulting distributions are shown in 5.19.

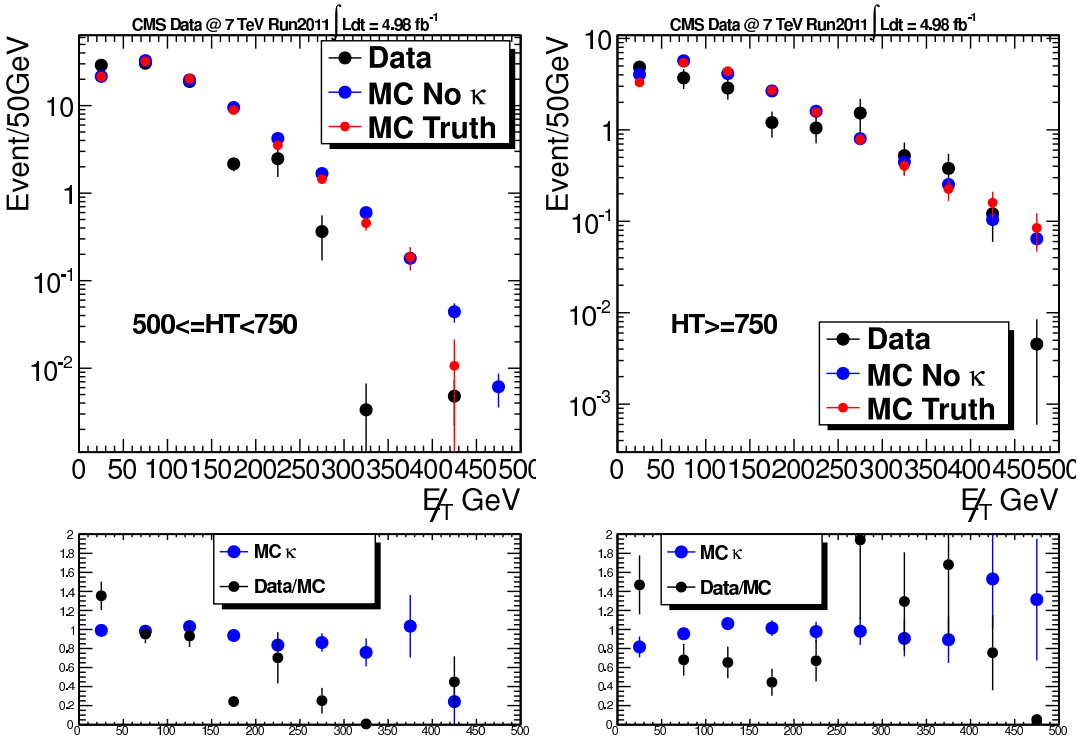

Figure 5.19: Top Left: $\not\!\!E_T$ prediction for lepton plus hadronic tau for $500 \leq HT < 750$. Black points show the final prediction from data; blue points show the prediction in MC before κ factors are applied; red points show the MC truth $\not\!\!E_T$ distribution. Top right: $\not\!\!E_T$ prediction for lepton plus hadronic tau for $HT \geq 750$. Bottom row: κ factors and data/MC ratios for $500 \leq HT < 750$ (left) and $HT \geq 750$ (right)

The fifth and last contribution is the single tau events shown in the equation below

$$t\bar{t} \rightarrow b\bar{b}W^+W^- \rightarrow b\bar{b}\ jet\ jet\ \bar{\nu_\tau}\bar{l}\nu_l\nu_\tau \tag{5.9}$$

These events are modelled using single lepton events with 4 jets and the appropriate H_T cut. The single lepton is turned in to a tau via the same r_{lep} response function as the lep tau lep prediction. The $\not\!\!E_T$ distribution in step 1 and 2 are modified using the r_{lep} response function in this equation,

$$\vec{\not\!\!E}_{T_{single\ tau}} = (1 - r_{lep}) \times \vec{p}_{Tl} + \vec{\not\!\!E}_{T_{single\ lepton}} \tag{5.10}$$

The statistics of the sampling is handled the same way as lepton plus a lepton tau case. Step 3 - 6 is the same as above with the results shown in figure 5.20.

All five of these contributions are combined together to make the final dilepton/tau prediction for the $\not\!\!E_T$ distribution. The combination is done by first applying only the α factors individually to get a combined distribution before κ_{DL} is applied. The final combined prediction is done with a combined κ_{DL} for each individual 50 GeV bins. The results are shown in figure 5.21 in 50 GeV bins with statistical uncertainties only, and are tabulated on table 5.7 and 5.8 in 100 GeV bins with systematic uncertainties included.

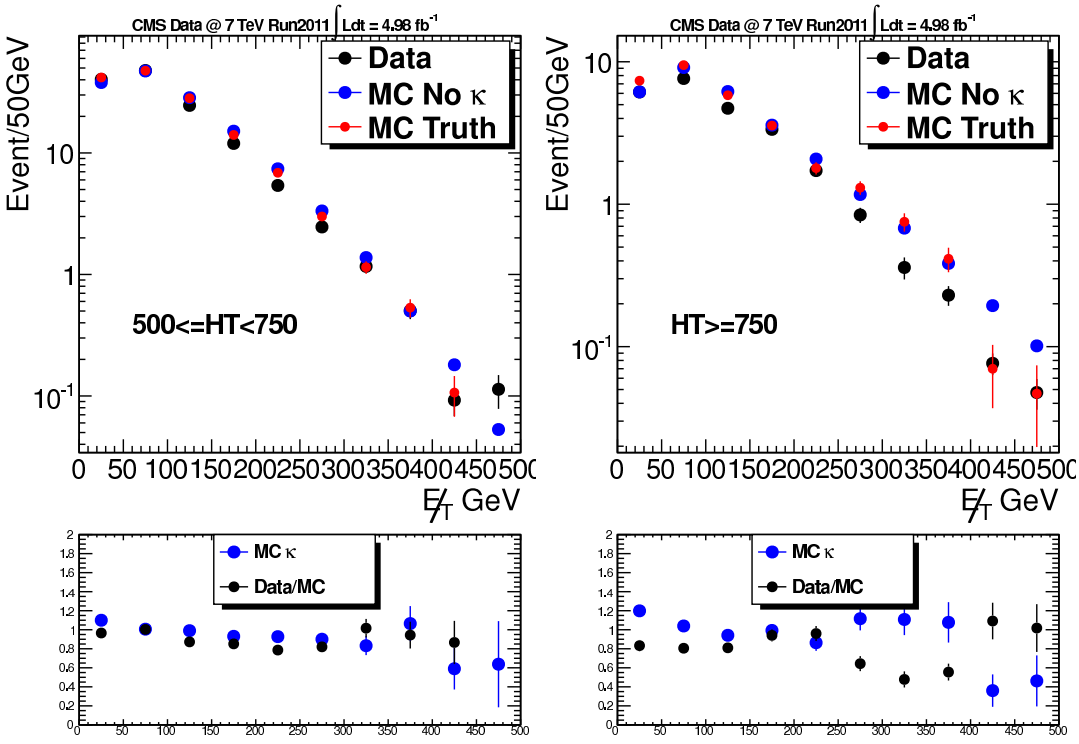

Figure 5.20: Top Left: $\not\!\!E_T$ prediction for single tau for $500 \leq HT < 750$. Black points show the final prediction from data; blue points show the prediction in MC before κ factors are applied; red points show the MC truth $\not\!\!E_T$ distribution. Top right: $\not\!\!E_T$ prediction for single tau for $HT \geq 750$. Bottom row: κ factors and data/MC ratios for $500 \leq HT < 750$ (left) and $HT \geq 750$ (right)

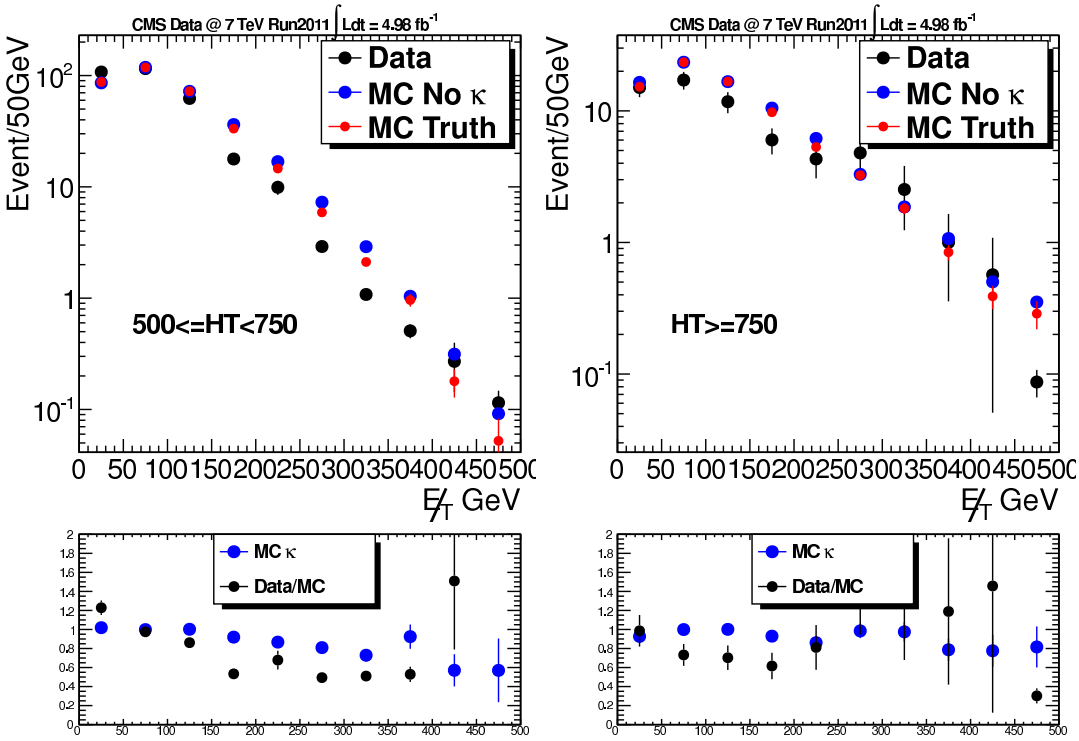

Figure 5.21: Top Left: $\not\!\!E_T$ prediction for combining all 5 dilepton and tau contributions for $500 \leq HT < 750$. Black points show the final prediction from data; blue points show the prediction in MC before κ factors are applied; red points show the MC truth $\not\!\!E_T$ distribution. Top right: $\not\!\!E_T$ prediction for combining all 5 dilepton and tau contributions for $HT \geq 750$. Bottom row: κ factors and data/MC ratios for $500 \leq HT < 750$ (left) and $HT \geq 750$ (right)

Table 5.7: Dilepton/tau prediction for $500 \leq HT < 750$ bin.

E_T^{miss} :	[150; 250)	[250; 350)	[350; 450)	≥ 450 GeV
Ignored	1.0 ± 1.0	< 0.01	< 0.01	< 0.01
Lost	2.5 ± 0.6	0.08 ± 0.08	0.08 ± 0.08	0.02 ± 0.03
$l + \tau_{lep}$	2.26 ± 0.02	0.16 ± 0.01	0.08 ± 0.01	0.05 ± 0.06
$l + \tau_{had}$	7.48 ± 0.06	0.88 ± 0.02	0.04 ± 0.01	0.03 ± 0.04
Single τ	17.7 ± 0.1	3.73 ± 0.02	0.56 ± 0.01	0.12 ± 0.07
MC Expect	33.1 ± 0.4	10.2 ± 0.2	1.35 ± 0.09	0.09 ± 0.09
Data Total	$31.1 \pm 1.2 \pm 2.1$	$4.85 \pm 0.13 \pm 0.28$	$0.76 \pm 0.08 \pm 0.12$	$0.22 \pm 0.12 \pm 0.09$

Table 5.8: Dilepton/tau prediction for $HT \geq 750$ bin.

E_T^{miss} :	[150; 250)	[250; 350)	[350; 450)	≥ 450 GeV
Ignored	2.4 ± 1.6	2.6 ± 1.8	< 0.01	< 0.01
Lost	0.4 ± 0.4	0.5 ± 0.5	0.07 ± 0.09	< 0.01
$l + \tau_{lep}$	0.5 ± 0.6	1.4 ± 1.4	0.38 ± 0.64	0.01 ± 0.07
$l + \tau_{had}$	3.6 ± 1.9	2.2 ± 1.4	0.53 ± 0.70	< 0.02
Single τ	4.6 ± 0.1	1.5 ± 0.1	0.20 ± 0.01	0.05 ± 0.01
MC Expect	16.7 ± 0.3	5.2 ± 0.1	1.57 ± 0.08	0.35 ± 0.04
Total	$10.9 \pm 2.5 \pm 1.4$	$8.1 \pm 2.6 \pm 1.4$	$1.18 \pm 0.94 \pm 0.19$	$0.06 \pm 0.07 \pm 0.03$

5.6.1 Dilepton and tau systematics

Systematics uncertainties for the dilepton/tau prediction are handled separately from the single lepton prediction. Figures 5.15 through 5.21 are shown

with statistical uncertainties only. The systematic uncertainties are calculated separately in 100 GeV bins with the combined dilepton/tau prediction which are listed on table 5.7 and 5.8.

- **Jet and $\not\!\!E_T$ energy scale** Just as in the single lepton case, the jet and $\not\!\!E_T$ energy scale uncertainty from the MC contributes to an uncertainty in the dilepton and tau prediction. The uncertainties are modeled by varying the uncertainties in the energy of every jet in an event and recalculating the $\not\!\!E_T$ due to this change. This is the standard recipe given the JETMET group in CMS.

- **MC κ_{DL} factor statistics** The other dominant uncertainty is due to the statistical uncertainties of the κ_{DL} and α factor. Since the uncertainty in $\kappa_{DL} >> \alpha$ these two uncertainties are treated as one item. Due to limited amount of MC events available, the MC statistics in higher $\not\!\!E_T$ bins degrade quickly. This uncertainty is taken as a systematic since it will not change with additional data but additional MC can reduce this uncertainty.

The systematic uncertainties for the dilepton and tau prediction are combined and summarized on table 5.9 and 5.10. The absolute uncertainties are tabulated on table 5.7 and 5.8 in the final data prediction. These uncertainties are added

Table 5.9: Systematic uncertainties for dilepton and tau prediction for $H_T \in [500, 750)$ GeV in the search $\not{E}_T$ bins. Each uncertainty is expressed as change in the ratio of predicted to the true number of events (evaluated in Monte Carlo) in the single lepton prediction.

E_T^{miss} :	$[150; 250)(\%)$	$[250; 350)$ (%)	$[350; 450)$ (%)	≥ 450 GeV (%)
DL $\not{E}_T$ and JES	6.5	4.2	4.5	20.3
DL MC κ Stat	1.9	4.0	10.6	33.5
DL+τ Total	6.8	5.8	15.2	39.2

Table 5.10: Systematic uncertainties for dilepton and tau prediction for $H_T \geq 750$ GeV in the search $\not{E}_T$ bins. Each uncertainty is expressed as change in the ratio of predicted to the true number of events (evaluated in Monte Carlo).

E_T^{miss} :	$[150; 250)$ (%)	$[250; 350)$ (%)	$[350; 450)$ (%)	≥ 450 GeV (%)
DL $\not{E}_T$ and JES	13.0	16.1	8.2	21.1
DL MC κ Stat	1.9	6.0	10.6	23.6
DL+τ Total	13.1	17.2	13.4	31.7

into the full total (single lepton, dilepton/tau, and QCD) prediction assuming the uncertainties are uncorrelated between the predictions.

5.7 QCD Prediction

Light flavor hadronic events (QCD) do not produce any prompt leptons from on shell W-boson or Z-boson decays. QCD events contaminate our signal and control samples through fake leptons. Fake leptons could be from the noise in

the detector, such as a noisy muon chamber getting associated with an isolated track in the tracker with a small random probability. Fake leptons could also be from a real physics object that the detector misidentifies, such as a kaon punch through into the muon chambers. The last source of fake leptons are not fake leptons at all, these are leptons created through decays of hadrons, such as those from bottom meson decays. All of these fake leptons are predicted in the same way due to a common characteristic between them. The isolation distribution for fake leptons is much flatter than that of prompt leptons. Since these events do not produce any real $\not\!\!E_T$ or only a small amount in case of hadron decays, their contribution to the $\not\!\!E_T$ tail is negligible. The analysis simply sets an upper limit on the number of QCD events in most bin of $\not\!\!E_T$. The more important component is QCD contamination of the control sample in lepton p_T. Fake lepton events produce random leptons in the entire p_T range. Since the lepton spectrum method is based upon the lepton p_T distribution, these random events will bias our single lepton measurement. This analysis uses a shape and scale method to estimate the contribution of QCD into both $\not\!\!E_T$ and lepton p_T. These predictions are made separately for muons and electrons due to the difference in their isolation requirements and trigger biases.

The overall method utilizes the shape and normalization of combined relative isolation of the lepton to make our QCD prediction. Combined relative isolation

(Iso^{Rel}_{Comb}) is defined to be the sum of the transverse energy in the tracker, ECAL, and HCAL within a cone of $\Delta R < 0.3$ of the lepton in $\eta - \phi$ space divided by the lepton p_T. The Iso^{Rel}_{Comb} distribution for various MC and data are shown in figure 5.22. The binning in Iso^{Rel}_{Comb} distribution is made such that the first bin is the same cut applied in our single lepton analysis cut. In the MC, events with prompt leptons such as $t\bar{t}$ and $W + jets$ has a sharp peak near zero while QCD events have a much flatter distribution. The shapes between muons and electrons are vastly different which force us to make their prediction separately. Note that the disagreement between the data and MC on the electron side is due to the isolation requirement at the trigger level. This is a substantial difference that will be corrected for the electron side of this analysis. The QCD prediction is made by measuring the shape of Iso^{Rel}_{Comb} in QCD events using data and finding a normalization factor to scale this shape to predict the amount of QCD events in our signal or control sample. Each of the four prediction (muon $\not\!\!E_T$, muon p_T, electron $\not\!\!E_T$ and electron p_T) has slightly different shapes and cuts applied. So the details of each prediction will be described in their individual subsections. Finally, due to lack of QCD events in these regions the prediction is made without the b-tagging requirement and made for HT > 500 without binning applied. The final predictions are made by computing a scale factor between the inclusive sample to the exclusive b-tagged sample in data. Figure 5.23 shows the Iso^{Rel}_{Comb} distribution

of a signal sample with and without the 2-btagged jets and HT bin requirements. By taking the ratio of these distribution in the high Iso^{Rel}_{Comb} regions. The scale factor is computed for both muons and electrons combined, $SF_{500} = 0.062 \pm 0.003$ for $500 < HT < 750$ & 2-btagged jets, and $SF_{750} = 0.013 \pm 0.001$ for $HT > 750$ & 2-btagged jets. These values do not differ significantly between muons and electrons since they are only dependent on the hadronic part of the event.

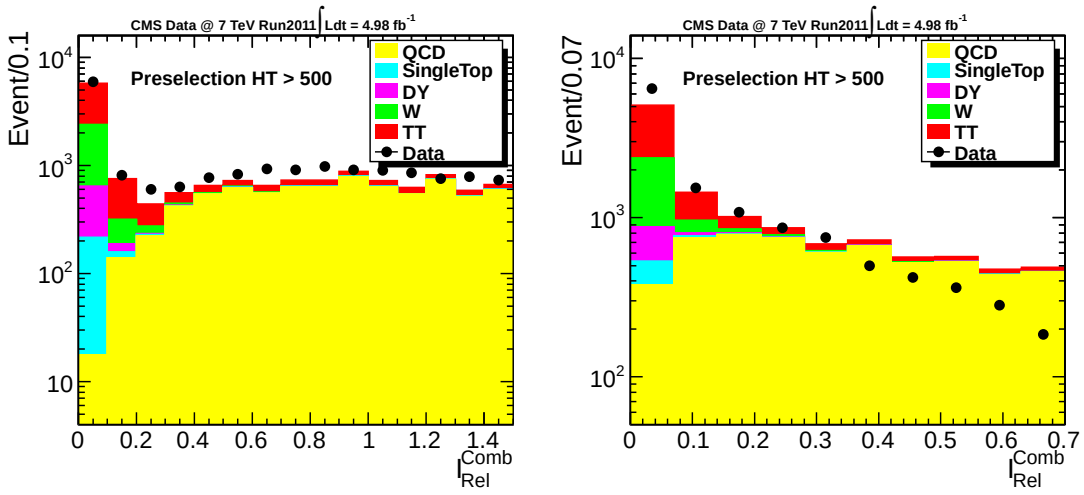

Figure 5.22: Left: Muon Iso^{Rel}_{Comb} distribution for various MC sample overlaid with data with only preselection applied. Right: Electron Iso^{Rel}_{Comb} distribution for MC and data with only preselection applied.

5.7.1 Muon MET QCD prediction

The QCD contribution to muon sample's $\not\!\!E_T$ distribution is expected to be tiny since the muon fake rate is much smaller than electrons and even a moderate

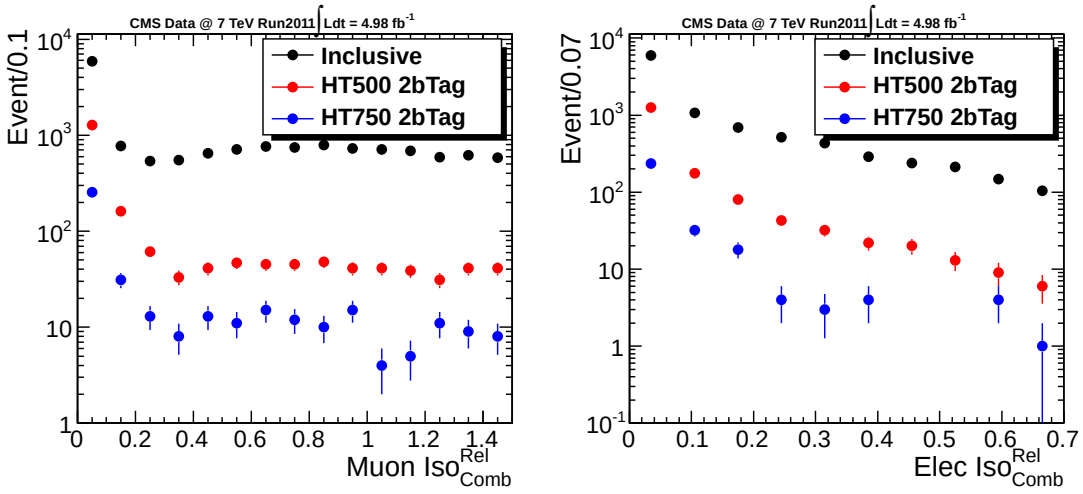

Figure 5.23: Relative isolation distributions for muon (left) and electron (right) events for signal samples with inclusive cuts ($HT \geq 500$, 0 or more b-tagged jets) in black. Low HT bin cut ($500 \leq HT < 750$, 2 or more b-tagged jets) in red. High HT bin cut ($HT \geq 750$, 2 or more b-tagged jets) in blue. The high relative isolation regions are used to determine the inclusive to HT500,2-btag and HT750,2-btag ratios for all muon QCD prediction.

$\not\!\!E_T$ cut removes most of the QCD events. For completeness, the prediction will be documented here thoroughly. First, the shape and signal region must be defined. Figure 5.24 left shows that the low $\not\!\!E_T$ region is dominated by QCD events once the Iso^{Rel}_{Comb} cut is removed. Since QCD events resides in the low $\not\!\!E_T$ region, a requirement of $\not\!\!E_T$ < 100 GeV is made for the shape sample. Figure 5.24 right shows that leptons with large impact parameters are also more likely to be from QCD. Table 5.11 shows the cuts for both shape and signal samples used to predict QCD contribution to $\not\!\!E_T$ in the muon channel. The signal cuts are made to emulate those used in the analysis' event selection and signal bins. A

bin of 100 < $\not\!\!E_T$< 150 GeV is made for validation purposes due to the larger QCD contribution in this bin. The normalization region for muons is chosen to be between 0.5 < Iso^{Rel}_{Comb} < 1.5. This region is dominated by QCD events in the background shape which minimizes signal contamination.

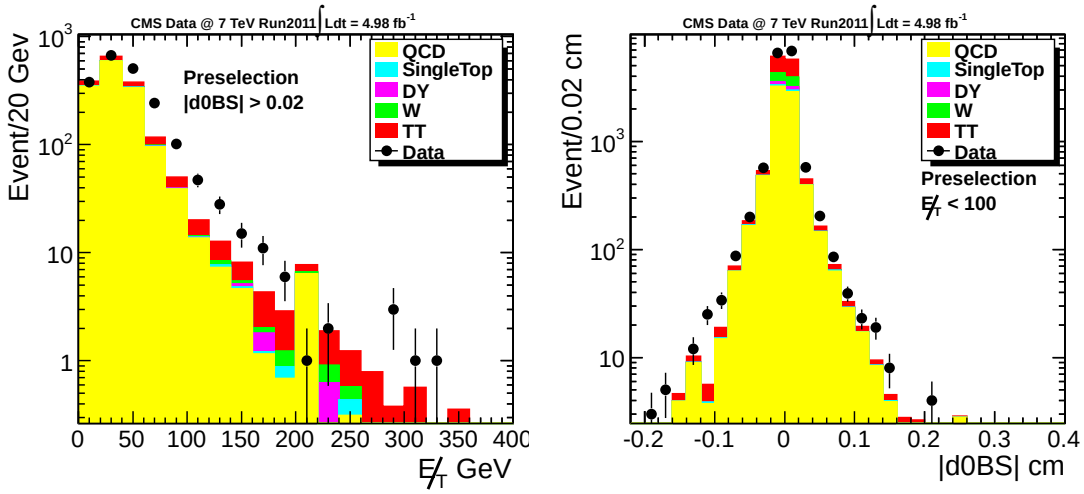

Figure 5.24: Muon channel: $\not\!\!E_T$ distribution in data and MC with d0(BS) and relative isolation removed from preselection cut with $H_T > 500$. Muon channel: d0 relative to beamspot distribution in data and MC with d0(BS) and relative isolation removed from preselection cut with $H_T > 500$.

Table 5.11: Selections for muon and electron $\not\!\!E_T$ QCD prediction.

Cuts	$\lvert d0_{BS}\rvert$ cm	$\not\!\!E_T$ GeV
Shape	> 0.02	< 100
Signal $\not\!\!E_T$ 100	< 0.02	100 to 150
Signal $\not\!\!E_T$ 150	< 0.02	150 to 250
Signal $\not\!\!E_T$ 250	< 0.02	250 to ∞

The QCD prediction method is illustrated in figure 5.25 for the MC on the left and in data on the right. In the MC case, the shape cuts are applied which produces the black points before normalization, while the signal cuts produces the stacked histograms. Using the MC, the method subtracts out the $t\bar{t}$ and $W + jets$ contributions to both the shape and signal Iso^{Rel}_{Comb} distribution. The shape distribution is normalized to have the equal number of events as the signal distribution in the high Iso^{Rel}_{Comb} region between 0.5 to 1.5. The prediction for QCD contribution to our single muon control sample in this $\not\!\!E_T$ bin the number of events with $Iso^{Rel}_{Comb} < 0.1$ in the normalized shape distribution. In the data (right), the signal shape are no longer stacked histogram and changed to red points for the raw distribution, and blue points for the MC subtracted distribution. The same normalization procedure is done between the blue and black distributions. The same normalized shape distribution (black point) predicts the number of QCD events in data with $Iso^{Rel}_{Comb} < 0.1$. The SF_{500} and SF_{750} factors are applied to divide the prediction into the HT $\in$ [500,750) and [750,∞) bins which are tabulated on table 5.12. The same plots are shown for $\not\!\!E_T \in$ [150,250) and [250,∞) in figure 5.26 and 5.27 respectively. The prediction stops at $\not\!\!E_T \in$ [250,∞) due to lack of data events in the high Iso^{Rel}_{Comb} region in the signal samples. In this case, upper limits on $\not\!\!E_T \in$ [350,450) and [450,∞) are derived from the [250,∞) bin.

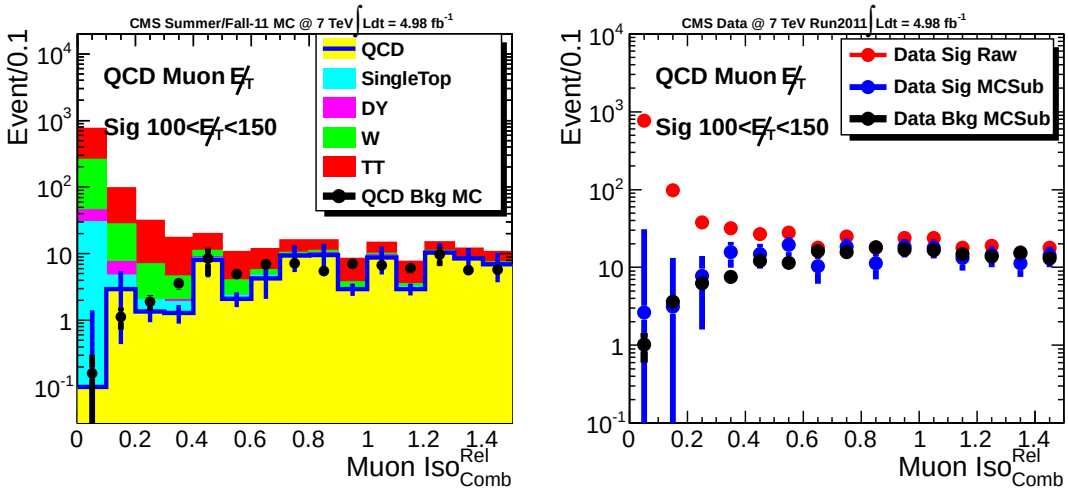

Figure 5.25: Muon relative isolation distributions for muon events with $100 < \not{E}_T < 150$ GeV. Left, MC signal shapes in stacked histograms and predicted background shape in solid black histogram which is normalized in the region of $0.5 < Iso^{Rel}_{Comb} < 1.5$. Right, data raw signal shape(red), MC subtracted data signal shape(blue), and predicted background shape(black) which is normalized by the same region of $0.5 < Iso^{Rel}_{Comb} < 1.5$.

Table 5.12: Muon MET prediction for QCD

$500 \leq HT < 750$			
$\not{E}_T$ Cut [GeV]	MC Truth	MC Prediction	Data Prediction
100-150	< 0.001	0.098 ± 0.088	0.062 ± 0.026
150-250	< 0.001	0.003 ± 0.003	0.008 ± 0.003
> 250	< 0.001	< 0.001	< 0.001
$HT \geq 750$			
$\not{E}_T$ Cut [GeV]	MC Truth	MC Prediction	Data Prediction
100-150	< 0.001	0.023 ± 0.021	0.015 ± 0.006
150-250	< 0.001	< 0.001	0.002 ± 0.001
> 250	< 0.001	< 0.001	< 0.001

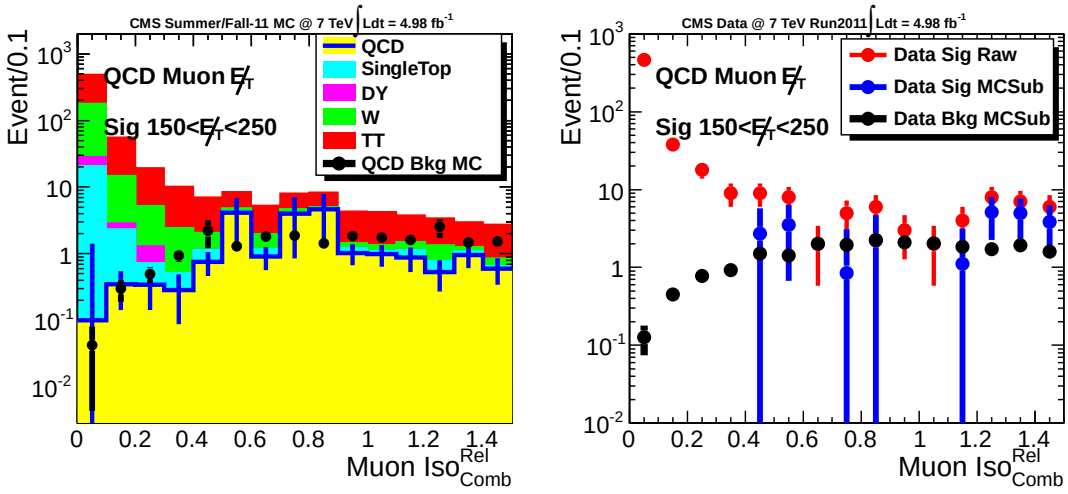

Figure 5.26: Muon relative isolation distributions for muon events with $150 < \not{E}_T < 250$ GeV. Left, MC signal shapes in stacked histograms and predicted background shape in solid black histogram. Right, data raw signal shape(red), MC subtracted data signal shape(blue), and predicted background shape(black).

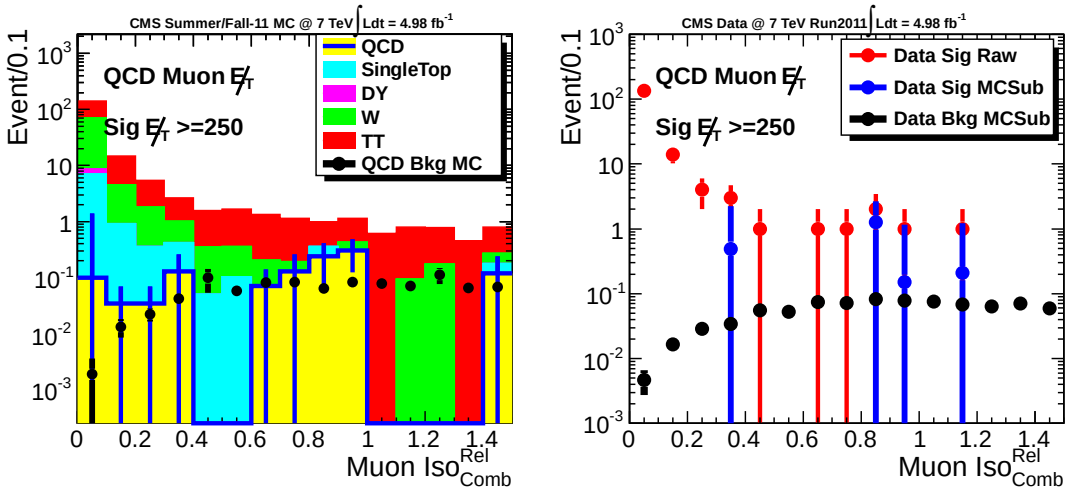

Figure 5.27: Muon relative isolation distributions for muon events with $\not{E}_T \geq 250$ GeV. Left, MC signal shapes in stacked histograms and predicted background shape in solid black histogram. Right, data raw signal shape(red), MC subtracted data signal shape(blue), and predicted background shape(black).

5.7.2 Electron MET QCD prediction

The QCD prediction in the electron $\not\!\!E_T$ channel follows exactly as those found in the muon $\not\!\!E_T$ channel. In fact, the same signal and shape regions are followed between the two channels. The only difference between the electron and muon channel is the binning of the Iso^{Rel}_{Comb} due to difference in the isolation cut in the single lepton analysis. The bin size is reduce from 0.1 to 0.07 in Iso^{Rel}_{Comb} in the electron case. The other modification to the electron $\not\!\!E_T$ channel is due to trigger inefficiency. At the trigger level, no requirements were made on the muons which can be see by the agreement between the data and MC in the high Iso^{Rel}_{Comb} region in figure 5.22 left. For the electron case, a trigger requirement of tracker isolation of 0.2 and calometric isolation of 0.2 is applied to the electron triggers. This biases the Iso^{Rel}_{Comb} distribution as shown in figure 5.22 right. Due to the MC subtraction portion of the QCD prediction a trigger efficiency scale factor must be applied. This is done by comparing the number of events in data over the MC in the region of interests which is between Iso^{Rel}_{Comb} of 0.35 to 0.7. The computed trigger scale factor is $SF_{trig} = 0.67 \pm 0.02$. This scale factor is only applied to the MC subtraction step of the method. The rest of the method in this channel follows the same as those prescribed in the muon $\not\!\!E_T$ prediction. The predictions are made using the data and MC histograms found in figure 5.28, 5.29, and 5.30. The final results in bins of HT are tabulated in table 5.13.

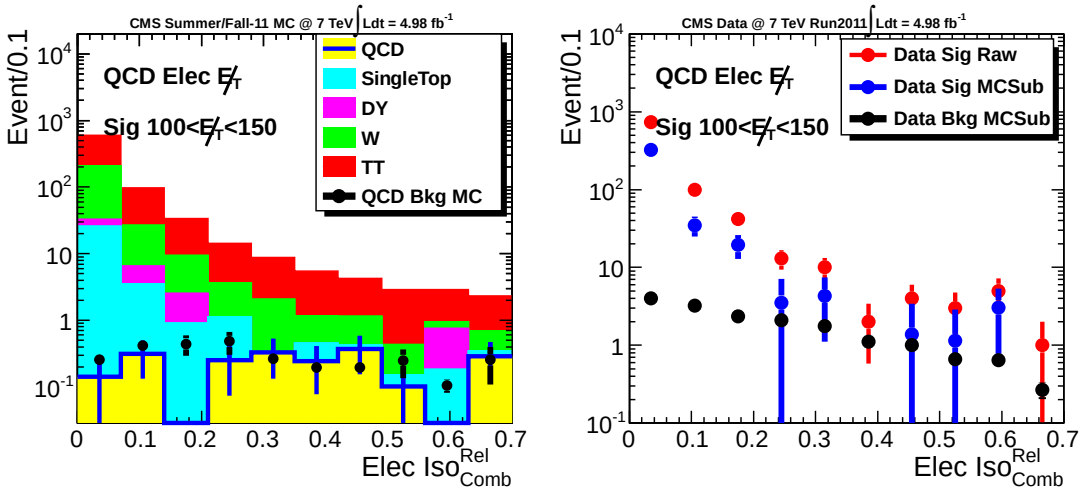

Figure 5.28: Electron relative isolation distributions for Electron events with $100 < \not{E}_T < 150$ GeV. Left, MC signal shapes in stacked histograms and predicted background shape in solid black histogram which is normalized in the region of $0.35 < Iso^{Rel}_{Comb} < 0.7$. Right, data raw signal shape(red), MC subtracted data signal shape(blue), and predicted background shape(black) which is normalized in the region of $0.35 < Iso^{Rel}_{Comb} < 0.7$.

Table 5.13: Elec MET prediction for QCD

$500 \leq HT < 750$			
$\not{E}_T$ Cut [GeV]	MC Truth	MC Prediction	Data Prediction
100-150	0.009 ± 0.008	0.016 ± 0.002	0.248 ± 0.015
150-250	0.009 ± 0.008	0.002 ± 0.001	0.160 ± 0.010
> 250	0.008 ± 0.008	0.002 ± 0.001	0.004 ± 0.001
$HT \geq 750$			
$\not{E}_T$ Cut [GeV]	MC Truth	MC Prediction	Data Prediction
100-150	0.002 ± 0.002	0.003 ± 0.001	0.052 ± 0.003
150-250	0.002 ± 0.002	0.001 ± 0.001	0.034 ± 0.002
> 250	0.002 ± 0.002	0.001 ± 0.001	0.001 ± 0.001

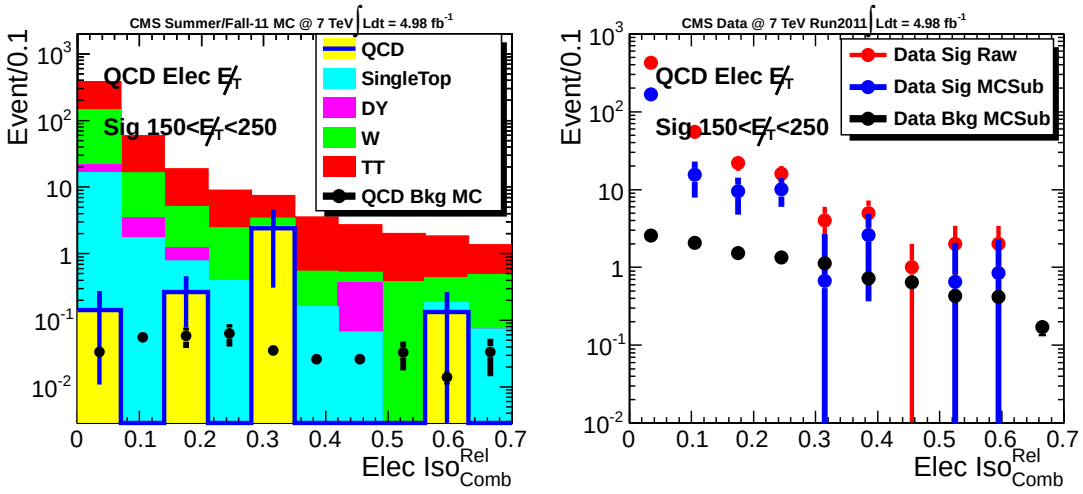

Figure 5.29: Electron relative isolation distributions for Electron events with 150 $< \not{E}_T <$ 250 GeV. Left, MC signal shapes in stacked histograms and predicted background shape in solid black histogram. Right, data raw signal shape(red), MC subtracted data signal shape(blue), and predicted background shape(black).

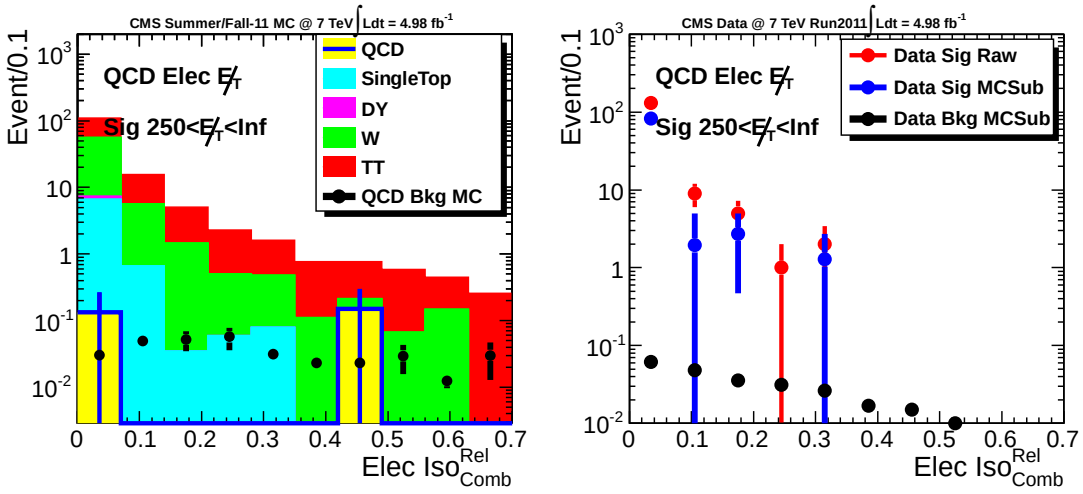

Figure 5.30: Electron relative isolation distributions for Electron events with $\not{E}_T \geq$ 250 GeV. Left, MC signal shapes in stacked histograms and predicted background shape in solid black histogram. Right, data raw signal shape(red), MC subtracted data signal shape(blue), and predicted background shape(black).

5.7.3 QCD Muon p_T prediction

The QCD prediction in muon p_T distribution is more difficult due to the dependence of the Iso^{Rel}_{Comb} distribution on the lepton's p_T. Figure 5.31 left shows the lepton p_T dependence of the Iso^{Rel}_{Comb} in QCD Monte Carlo. This is a natural dependency since the lepton p_T is explicitly in the denominator of Iso^{Rel}_{Comb}. In order to correct for this dependency MC correction factors, κ_{QCD}, are applied to the muon p_T prediction. κ_{QCD}s are defined as the ratio between the number of MC truth QCD events and number of MC predicted QCD events. Figure 5.32, 5.33 and 5.34 shows the muon p_T prediction in the same methodology as in the muon $\not\!\!E_T$ prediction. Table 5.15 shows the results of the QCD prediction in muon p_T distribution with the additional κ_{QCD} applied to the data prediction to correct for the p_T dependency in Iso^{Rel}_{Comb}.

Table 5.14: Shape and signal selections lepton $p_T$prediction.

Cuts	$\lvert d0_{BS}\rvert$ cm	$\not\!\!E_T$ GeV	p_T GeV
Shape	> 0.02	< 100	> 100
Signal p_T 100	< 0.02	none	100 to 150
Signal p_T 150	< 0.02	none	150 to 250
Signal p_T 250	< 0.02	none	> 250

Table 5.15: Muon $p_T$prediction for QCD. *Data predictions are shown with statistical uncertainty from data raw, while the systematic uncertainty is due to the limited statistics of the κ_{QCD} factor. See section 5.7.6 for full systematic uncertainty treatment.

$500 \leq HT < 750$					
$\not\!\!E_T$ Cut [GeV]	MC Truth	MC Prediction	κ_{QCD}	Data Raw	Data Pred*
100-150	0.007 ± 0.007	0.007 ± 0.007	0.97 ± 1.38	0.078 ± 0.087	$0.076 \pm 0.085 \pm 0.105$
150-250	0.021 ± 0.013	0.001 ± 0.001	15.8 ± 18.4	0.012 ± 0.014	$0.194 \pm 0.216 \pm 0.221$
> 250	0.002 ± 0.002	< 0.001	40.1 ± 62.9	< 0.001	$0.001 \pm 0.001 \pm 0.002$
$HT \geq 750$					
$\not\!\!E_T$ Cut [GeV]	MC Truth	MC Prediction	κ_{QCD}	Data Raw	Data Pred*
100-150	0.001 ± 0.001	0.002 ± 0.002	0.97 ± 1.38	0.016 ± 0.018	$0.016 \pm 0.018 \pm 0.022$
150-250	0.004 ± 0.003	0.0003 ± 0.0003	15.8 ± 18.4	0.003 ± 0.003	$0.041 \pm 0.045 \pm 0.055$
> 250	0.0003 ± 0.0004	< 0.001	40.1 ± 62.9	< 0.001	< 0.001

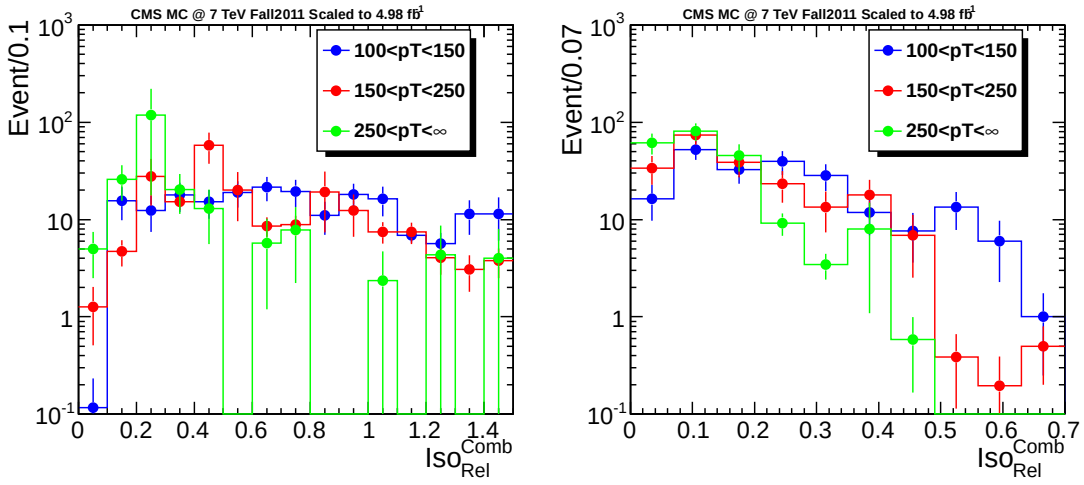

Figure 5.31: Left: Muon Iso^{Rel}_{Comb} distribution with various muon p_T bins. The Iso^{Rel}_{Comb} distribution pulls towards the lower values as the muon p_T increases. Right: Electron Iso^{Rel}_{Comb} distribution with various electron p_T bins. Samm notation as the muon.

5.7.4 Electron pT QCD prediction

Electron p_T QCD prediction is done using the same method as the muon p_T QCD prediction. They both suffer from the same lepton p_T dependency on the Iso^{Rel}_{Comb} distribution as shown in figure 5.31 right for electrons. Due to the larger fake rate of electrons relative to muons, the QCD contamination to the electron p_T distribution can be done for all signal bins. Figure 5.35 through 5.39 shows the QCD prediction for bins of electron $p_T \in$ [100,150), [150,250), [250,350), [350,450), and [450,∞). It is nice to note that the method does predict the QCD contamination when it has a substantial contribution. The results with all 5 bins in electron p_T are summarized in table 5.16.

Table 5.16: Electron $p_T$prediction for QCD. *Data predictions are shown with statistical uncertainty from data raw, while the systematic uncertainty is due to the limited statistics of the κ_{QCD} factor. See section 5.7.6 for full systematic uncertainty treatment.

$500 \leq HT < 750$					
$\not\!\!E_T$ Cut [GeV]	MC Truth	MC Prediction	κ_{QCD}	Data Raw	Data Pred*
100-150	1.0 ± 0.4	6.8 ± 1.4	0.15 ± 0.07	9.4 ± 1.1	$1.4 \pm 0.2 \pm 0.7$
150-250	1.4 ± 0.5	2.6 ± 0.5	0.53 ± 0.20	3.3 ± 0.4	$1.8 \pm 0.2 \pm 0.7$
250-350	0.89 ± 0.35	0.41 ± 0.08	2.17 ± 0.96	0.66 ± 0.08	$1.44 \pm 0.18 \pm 0.63$
350-450	0.33 ± 0.10	0.06 ± 0.01	5.85 ± 2.17	0.19 ± 0.02	$1.10 \pm 0.13 \pm 0.41$
450+	0.295 ± 0.048	0.019 ± 0.004	15.9 ± 4.2	0.009 ± 0.001	$0.147 \pm 0.018 \pm 0.038$
$HT \geq 750$					
$\not\!\!E_T$ Cut [GeV]	MC Truth	MC Prediction	κ_{QCD}	Data Raw	Data Pred*
100-150	0.21 ± 0.09	1.43 ± 0.30	0.15 ± 0.07	1.98 ± 0.24	$0.29 \pm 0.04 \pm 0.14$
150-250	0.30 ± 0.10	0.56 ± 0.11	0.53 ± 0.20	0.70 ± 0.08	$0.37 \pm 0.05 \pm 0.14$
250-350	0.19 ± 0.07	0.09 ± 0.02	2.17 ± 0.96	0.14 ± 0.02	$0.30 \pm 0.04 \pm 0.13$
350-450	0.068 ± 0.021	0.012 ± 0.002	5.85 ± 2.17	0.039 ± 0.005	$0.230 \pm 0.028 \pm 0.085$
450+	0.062 ± 0.010	0.004 ± 0.001	15.9 ± 4.2	0.002 ± 0.001	$0.031 \pm 0.004 \pm 0.008$

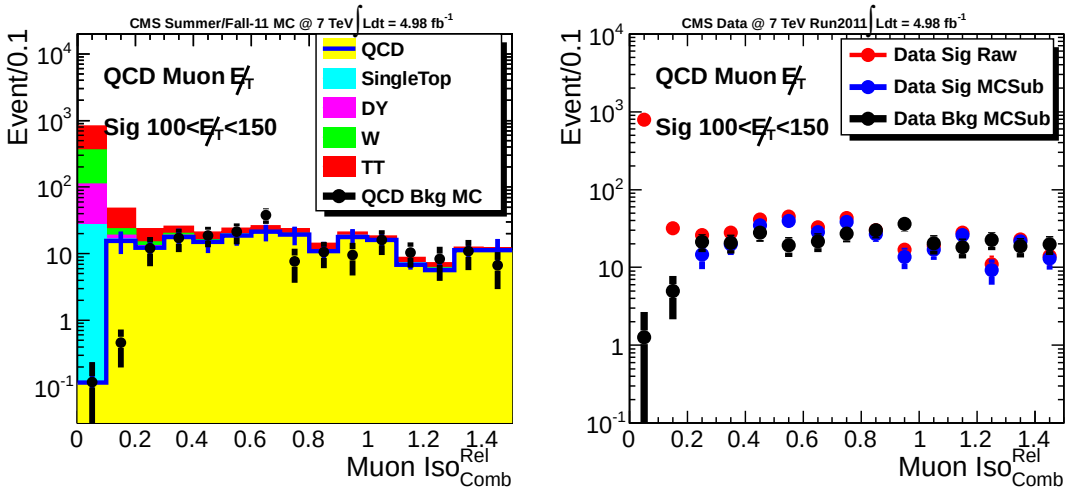

Figure 5.32: Muon relative isolation distributions for muon events with 100 $<p_T<$ 150 GeV. Left, MC signal shapes in stacked histograms and predicted background shape in solid black histogram. Right, data raw signal shape(red), MC subtracted data signal shape(blue), and predicted background shape(black) before κ_{QCD}.

5.7.5 Combined QCD prediction

The QCD prediction are combined between the muon and electron channels with $\not\!\!E_T$ prediction adding to the analysis' final $\not\!\!E_T$ prediction. The QCD p_T contribution actually creates an over prediction in the analysis' $\not\!\!E_T$ prediction. Instead of correcting the observed value in $\not\!\!E_T$, the QCD p_T prediction is subtracted from the final $\not\!\!E_T$ prediction. This is due to the fact that the lepton p_T tail from QCD does not contribute to the $\not\!\!E_T$ distribution observed in data. Table 5.17 shows the complete QCD results with systematics evaulated in the next subsection.

Table 5.17: Combined QCD prediction. Data predictions are shown with statistical uncertainties. Systematic uncertanties on all four QCD channels are treated together in the combined results. See section 5.7.6 for full systematic uncertainty treatment.

$500 \leq HT < 750$				
$\not\!\!E_T$ Cut [GeV]	Muon $\not\!\!E_T$	Electron $\not\!\!E_T$	Muon p_T	Electron p_T
150-250	0.008 ± 0.003	0.002 ± 0.001	0.194 ± 0.216	1.8 ± 0.2
250-350	< 0.001	0.002 ± 0.001	0.001 ± 0.001	1.44 ± 0.18
350-450	< 0.001	< 0.002	< 0.002	1.10 ± 0.13
450+	< 0.001	< 0.002	< 0.002	0.147 ± 0.018
$HT \geq 750$				
$\not\!\!E_T$ Cut [GeV]	Muon $\not\!\!E_T$	Electron $\not\!\!E_T$	Muon p_T	Electron p_T
150-250	0.002 ± 0.001	0.002 ± 0.002	0.041 ± 0.045	0.37 ± 0.05
250-350	< 0.001	0.002 ± 0.002	< 0.001	0.30 ± 0.04
350-450	< 0.001	< 0.002	< 0.001	0.230 ± 0.028
450+	< 0.001	< 0.002	< 0.001	0.031 ± 0.004

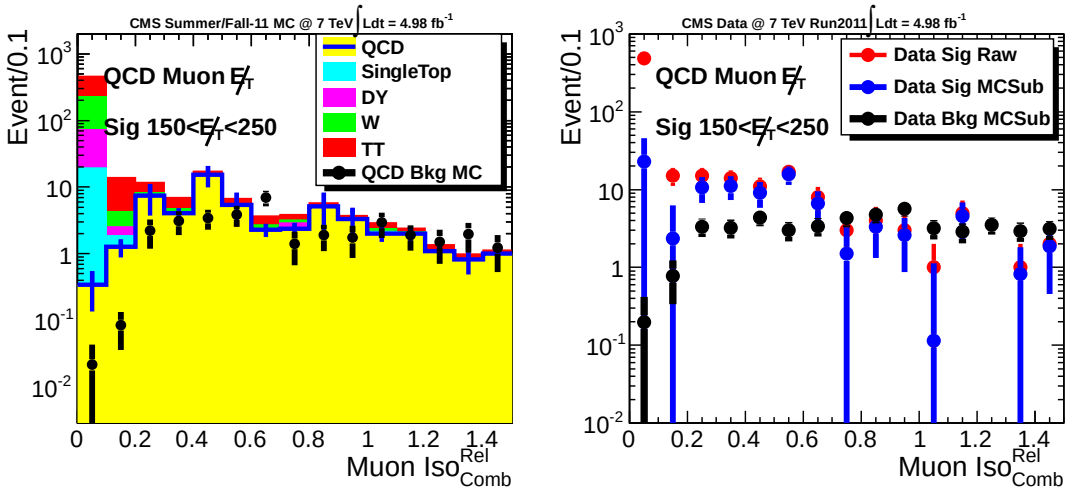

Figure 5.33: Muon relative isolation distributions for muon events with 150 $<p_T<$ 250 GeV. Left, MC signal shapes in stacked histograms and predicted background shape in solid black histogram. Right, data raw signal shape(red), MC subtracted data signal shape(blue), and predicted background shape(black) before κ_{QCD}.

5.7.6 QCD systematics

The list of systematics in the QCD prediction are listed as follows:

- $\sigma(t\bar{t})$ **and** $\sigma(W)$**:** The uncertainties in $t\bar{t}$ and W cross sections must be taken into account since these processes are subtracted out from data using MC. See section 5.8 for more details.

- **b-tagging Fake/Eff:** Due to limited data and MC statistics, QCD prediction is made using an inclusive (no b-tag requirement) sample then scaled using a QCD dominated region to find the fraction of b-tagged events in QCD. Dominated uncertainty is the charms faking bottoms in an event.

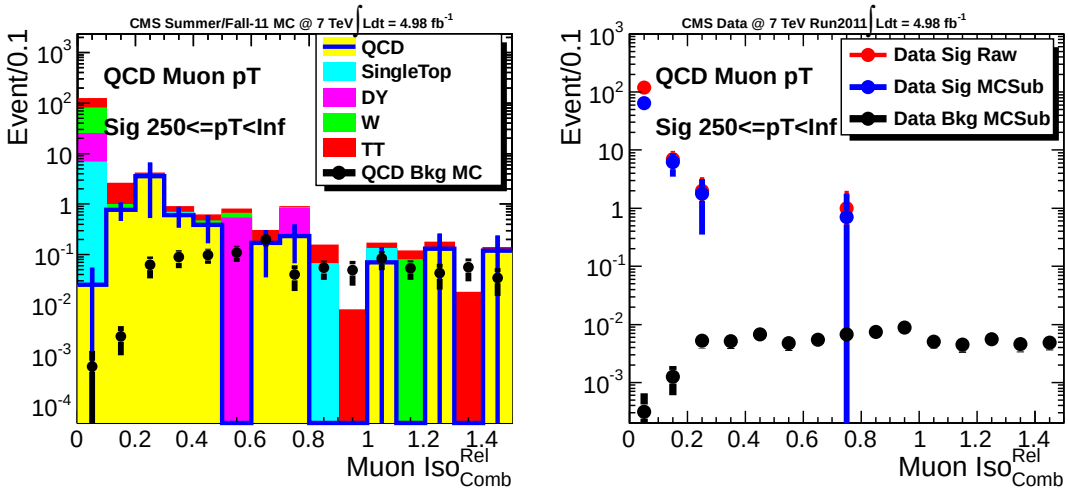

Figure 5.34: Muon relative isolation distributions for muon events with $p_T \geq 250$ GeV. Left, MC signal shapes in stacked histograms and predicted background shape in solid black histogram. Right, data raw signal shape(red), MC subtracted data signal shape(blue), and predicted background shape(black) before κ_{QCD}.

This uncertainty is taken in account by increasing the chance for a charm quark faking a bottom by 8%. The change in the final QCD prediction is taken as the uncertainty.

- **QCD p_T κ_0:** Due to the dependency of relative isolation on lepton's p_T, correction factors (κ) are used. This systematic is evaulated by varying the background sample's lepton p_T cut from the norminal value of 100 GeV up and down by 50 GeV in 10 GeV steps. The largest variation in the QCD p_T prediction is taken as the uncertainty for each bin. The uncertainty of these κ factors are added into only the p_T portion of the QCD prediction systematic uncertainties.

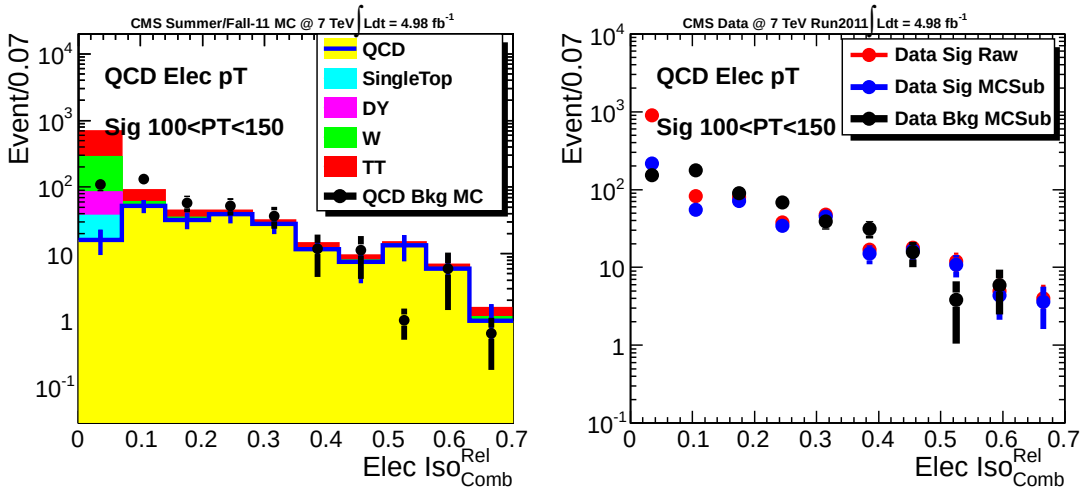

Figure 5.35: Electron relative isolation distributions for Electron events with 100 $<p_T<$ 150 GeV. Left, MC signal shapes in stacked histograms and predicted background shape in solid black histogram. Right, data raw signal shape(red), MC subtracted data signal shape(blue), and predicted background shape(black) before κ_{QCD}.

- **QCD** p_T $\sigma(\kappa)$**:** The κ_{QCD} depends on MC statistics from QCD, which is limited due to the analysis' lepton requirements. The statistical uncertainty from calculating κ_{QCD} from MC is propagated into the systematic uncertainties.

The QCD prediction systematics are tabulated as fraction of the prediction on table 5.18 and 5.19 for HT $\in$ [500,750) and [750,∞) respectively. The absolute value of these uncertainties are shown together with the total QCD prediction results on table 5.17.

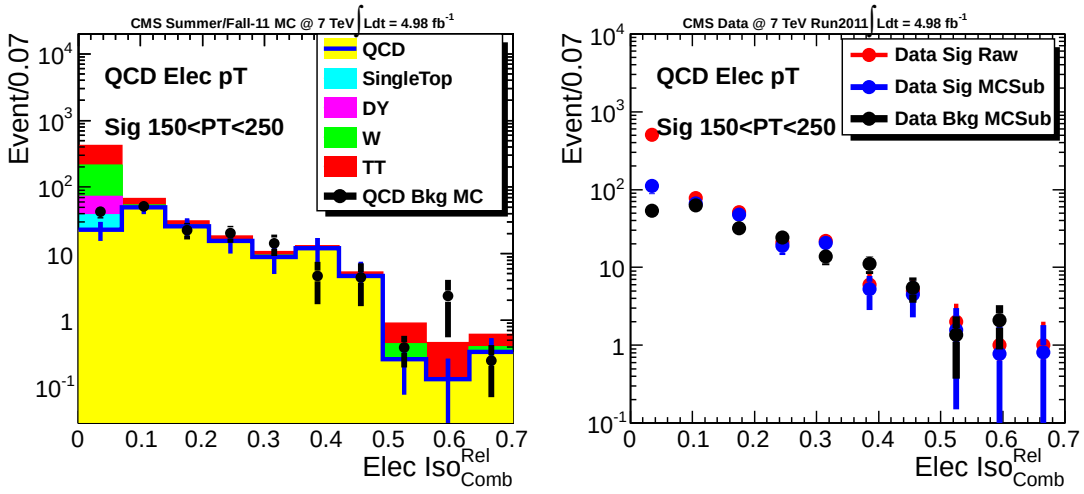

Figure 5.36: Electron relative isolation distributions for Electron events with 150 $<p_T<$ 250 GeV. Left, MC signal shapes in stacked histograms and predicted background shape in solid black histogram. Right, data raw signal shape(red), MC subtracted data signal shape(blue), and predicted background shape(black) before κ_{QCD}.

Table 5.18: Systematic uncertainties for QCD prediction for $H_T \in [500, 750)$ GeV in the search $\not\!\!E_T$ bins. Each uncertainty is expressed as change in the ratio of predicted to the true number of events (evaluated in Monte Carlo) in the QCD prediction.

E_T^{miss} :	$[150; 250)(\%)$	$[250; 350)$ (%)	$[350; 450)$ (%)	≥ 450 GeV (%)
QCD $\sigma(t\bar{t})$,$\sigma(W)$	5.0	6.7	7.6	30.5
QCD b-tagging	4.5	4.5	4.5	4.5
QCD p_T κ_0	1.9	3.8	15.9	40.2
QCD p_T $\sigma(\kappa)$	44.2	44.0	37.1	26.4
QCD Total	44.7	44.9	41.3	57.1

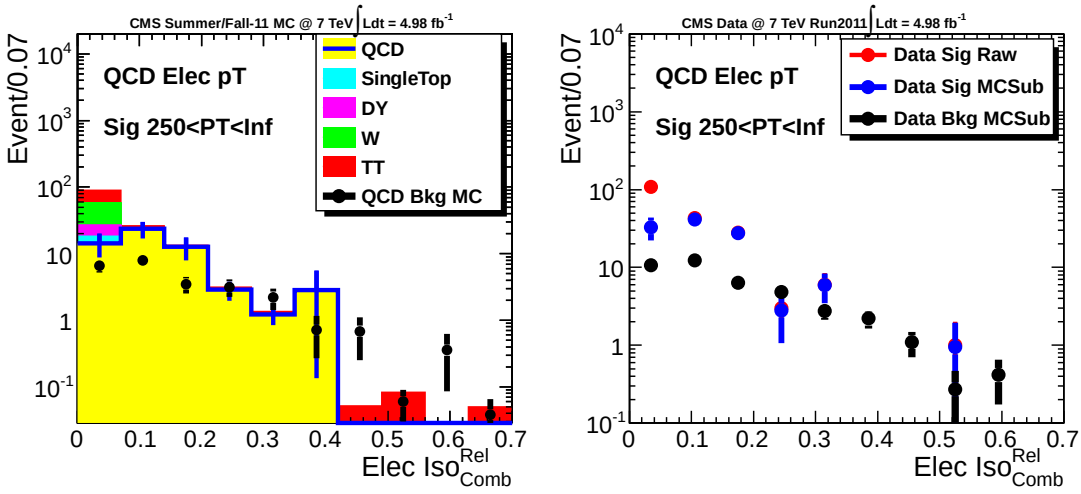

Figure 5.37: Electron relative isolation distributions for Electron events with 250 $<p_T<$ 350 GeV. Left, MC signal shapes in stacked histograms and predicted background shape in solid black histogram. Right, data raw signal shape(red), MC subtracted data signal shape(blue), and predicted background shape(black) before κ_{QCD}.

Table 5.19: Systematic uncertainties for QCD prediction for $H_T \geq 750$ GeV in the search $\not{E}_T$ bins. Each uncertainty is expressed as change in the ratio of predicted to the true number of events (evaluated in Monte Carlo).

E_T^{miss} :	[150; 250) (%)	[250; 350) (%)	[350; 450) (%)	$\geq$ 450 GeV (%)
QCD $\sigma(t\bar{t})$,$\sigma(W)$	5.0	6.7	7.6	30.5
QCD b-tagging	5.8	5.8	5.8	5.8
QCD p_T κ_0	1.9	3.8	15.9	40.2
QCD p_T $\sigma(\kappa)$	44.2	44.0	37.1	26.4
QCD Total	44.9	45.0	41.5	57.2

5.8 $t\bar{t}$ and $W + jets$ cross section uncertainties

In the calculation of the correction factors for the single lepton prediction and QCD prediction method, MC corrections are applied to the data in order to

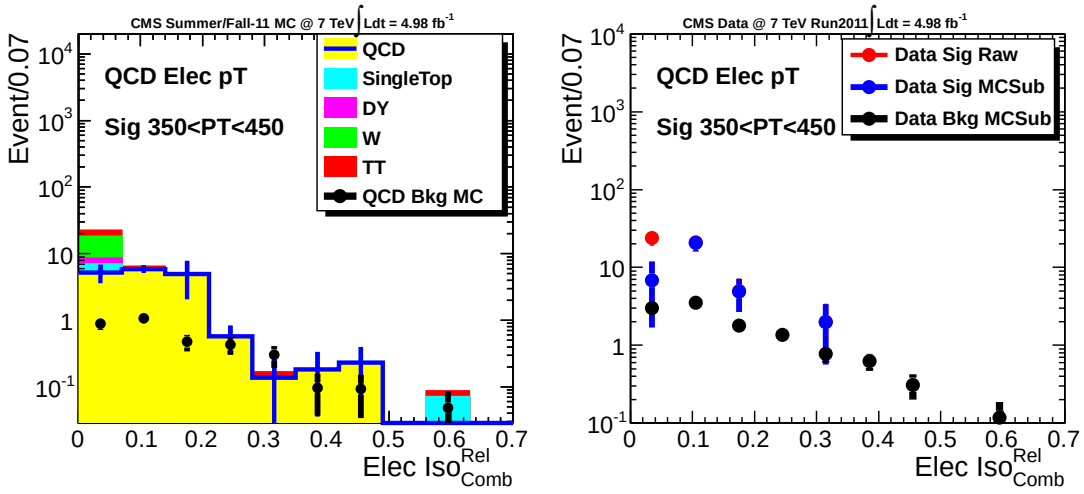

Figure 5.38: Electron relative isolation distributions for Electron events with 350 $<p_T<$ 450 GeV. Left, MC signal shapes in stacked histograms and predicted background shape in solid black histogram. Right, data raw signal shape(red), MC subtracted data signal shape(blue), and predicted background shape(black) before κ_{QCD}.

correct for well known effects of W polarization and prompt lepton contamination. These corrections have a 1st or 2nd order dependency on the cross section and the integrated luminosity in both scale and resolution. In this section of the analysis, both the scale bias and resolution are estimated for $W + jets$ and $t\bar{t}$. The scale bias is added in quadrature with the resolution uncertainty to make an estimate of the overall uncertainty of the cross section and integrated luminosity.

The $W + jets$ cross section uncertainty is measured using $Z + jets$ events by comparing data to MC with the analysis preselection cut of 4 jets and $H_T > 500$. Figure 5.40 and 5.41 show the dimuon and dielectron mass distribution in events with two opposite sign tight leptons, 4 jets and $H_T > 500$ GeV. The number of

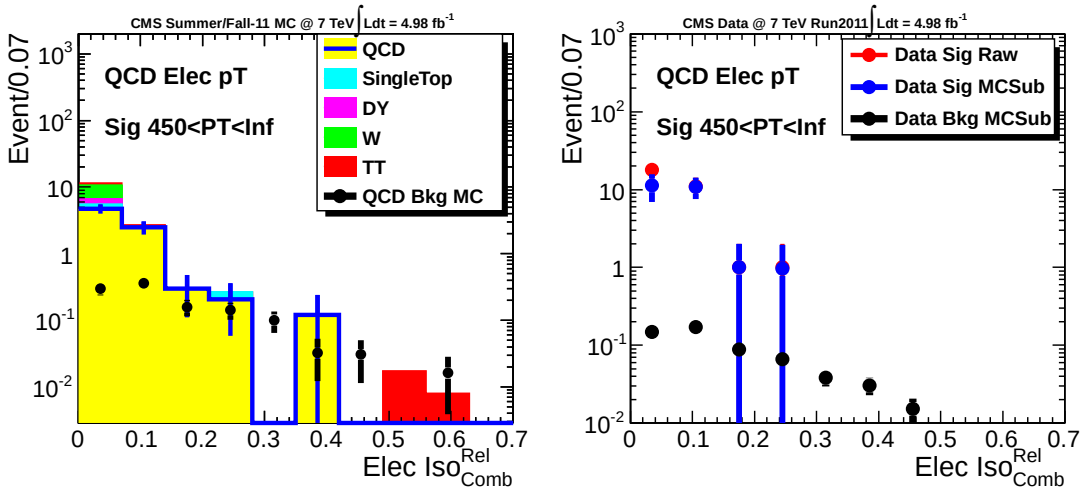

Figure 5.39: Electron relative isolation distributions for Electron events with $p_T \geq 450$ GeV. Left, MC signal shapes in stacked histograms and predicted background shape in solid black histogram. Right, data raw signal shape(red), MC subtracted data signal shape(blue), and predicted background shape(black) before κ_{QCD}.

events are counted using three different mass bins of 75 to 95, 80 to 100, and 85 to 105 GeV. The total uncertainty for each channel is calculated by summing in quadrature the statistical uncertainty, the difference from 1 for the ratio, and the maximal difference in the variation of the mass range. From table 5.20, this value is determined to be 11% for the muon channel and 13% for the electron channel.

The $t\bar{t}$ cross section uncertainty is determined in a similar way by using the topbox χ^2 variable defined by the hadronic M2, hadronic M3, and leptonic M3 differences with respect to their standard model measured values. Equation 5.11 shows the χ^2 variable as a function of $M3_{had}$, M_{top}, M2, M_W, $M3_{lep}$.

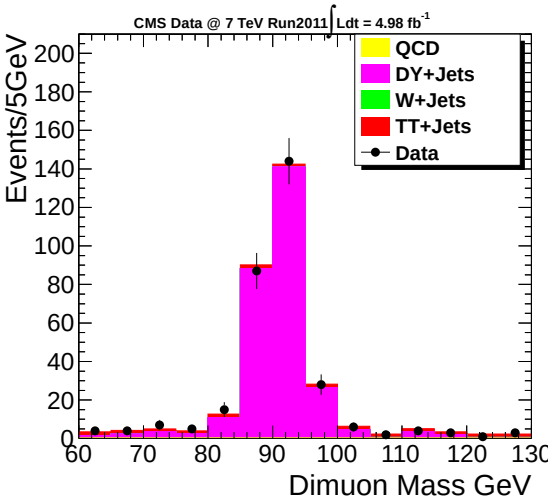

Figure 5.40: Dimuon mass distribution for events two oppositely signed muons with 4 jets and $HT > 500$.

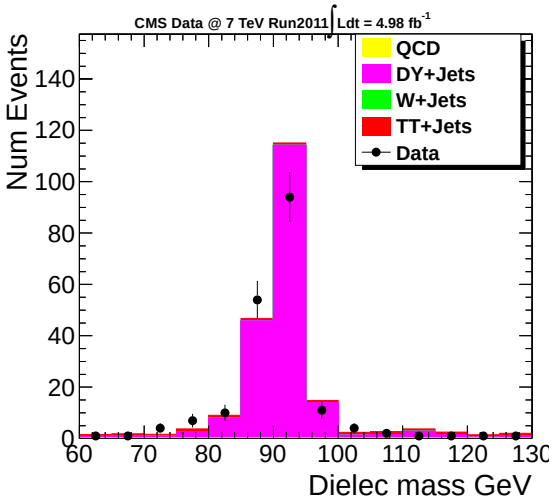

Figure 5.41: Dielectron mass distribution for events two oppositely signed electrons with 4 jets and $HT > 500$.

$$\chi^2 = \frac{(M3_{had} - M_{top})^2}{\sigma_{M3had}} + \frac{(M2 - M_W)^2}{\sigma_{M2}} + \frac{(M3_{Lep} - M_{top})^2}{\sigma_{M3lep}} \tag{5.11}$$

The uncertainties (σ) are taken from the width of these distributions in the MC, where truth matching has been applied in $t\bar{t}$. Figure 5.42 and 5.43 show the topbox χ^2 distributions in events with 1 muon or electron, 4 jets and HT > 500. The same strategy for evaluating the uncertainty is applied from the W + 4 jets. Three different cuts are applied to the χ^2 distributions $< 3, < 4$,and< 5 and the ratio between the data and MC are computed. The total uncertainty for each channel is calculated by summing in quadrature the statistical uncertainty, the difference from 1 for the ratio, and the maximal difference in the variation of the

mass range. From table 5.21, the total uncertainty is calculated to be 8% for the muon channel and 25% for the electron channel.

The uncertainties of both $t\bar{t}$ and $W+jets$ are combined in quandrature for an overall cross sectioinal uncertainty of 30%. In the κ factor calculation, the $t\bar{t}$ and $W+jets$ cross sections are varied by 30% and the combination with the maximal difference between the original and varied prediction is taken as the uncertainty in the prediction due to $t\bar{t}$ and $W+jets$ cross section uncertainties. In the QCD prediction, the MC subtraction includes the uncertainty of $t\bar{t}$, $Z+jets$, and $W+jets$ cross sections. The cross sections uncertainties are added into the systematic uncertainties of the QCD prediction by varying the MC subtraction cross section by 30% and the maximal difference in the prediction is taken as a systematic uncertainty in the QCD prediction.

Table 5.20: Data and MC comparison with different Z mass range.

Channel and Process	Mass Range	Data	MC	Ratio
Muon , Z+4 jets	75 to 95 GeV	251 ± 16	249 ± 11	1.01 ± 0.08
Muon , Z+4 jets	80 to 100 GeV	274 ± 17	274 ± 11	1.00 ± 0.07
Muon , Z+4 jets	85 to 105 GeV	265 ± 16	267 ± 11	0.99 ± 0.08
Electron , Z+4 jets	75 to 95 GeV	165 ± 13	168 ± 9	0.94 ± 0.09
Electron , Z+4 jets	80 to 100 GeV	169 ± 13	176 ± 9	0.91 ± 0.08
Electron , Z+4 jets	85 to 105 GeV	163 ± 13	163 ± 9	0.91 ± 0.09

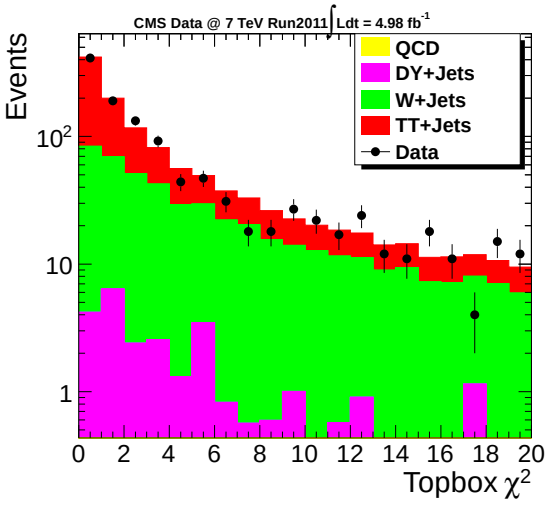

Figure 5.42: chi^2 distribution for single muon events with 4 jets and $HT > 500$.

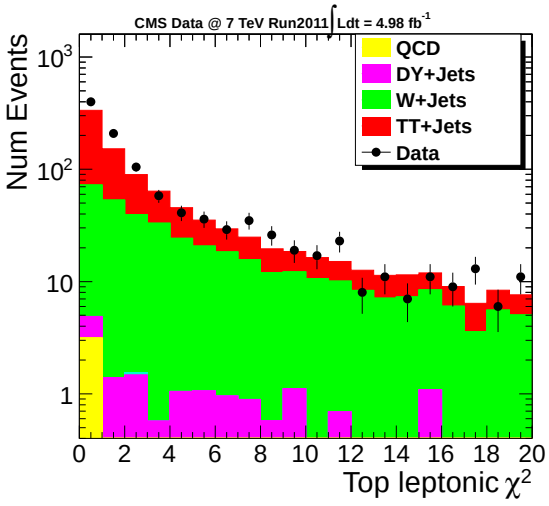

Figure 5.43: chi^2 distribution for single electron events with 4 jets and $HT > 500$.

5.9 W Polarization in $t\bar{t}$ and $W + jets$

Due to the vector-axial coupling of the electroweak interaction, polarized W-boson impart higher fraction of its momentum into the neutrino than the charged lepton. A fraction of W-boson produced directly in $W+jets$ events and $t\bar{t}$ events

Table 5.21: Data and MC comparison at different χ^2 cut.

Channel and Process	χ^2	Data	MC	Ratio
Muon, TTBar	< 3	736 ± 27	736 ± 5	1.00 ± 0.04
Muon, TTBar	< 4	828 ± 29	818 ± 6	1.01 ± 0.04
Muon, TTBar	< 5	872 ± 30	874 ± 6	1.00 ± 0.03
Electron, TTBar	< 3	714 ± 27	567 ± 6	1.24 ± 0.05
Electron, TTBar	< 4	772 ± 28	641 ± 6	1.21 ± 0.04
Electron, TTBar	< 5	813 ± 29	686 ± 6	1.18 ± 0.04

are polarized at the LHC. This section describes how the uncertainties due to W-polarization is measured in this analysis for the single lepton prediction.

5.9.1 W-boson polarization in $t\bar{t}$ decays

For the $t\bar{t}$'s W-boson polarization uncertainty measurement, the analysis utilizes the results from measurements made at D0 and CDF at the Tevatron. [23] The polarization of W-bosons from $t\bar{t}$ decays are parameterized by,

$$\omega(cos(\theta^*)) \propto 2(1 - cos^2\theta^*)f_0 + (1 - cos\theta^*)^2 f_- + (1 + cos\theta^*)^2 f_+ \qquad (5.12)$$

where f_0 is the longitudinal and f_- , f_+ are the left and right handed components of the W-boson's polarization. $cos(\theta^*)$ is the angle charged lepton relative to the W-boson in the top's rest frame. f_0's constrained to 0.688 ± 0.004, $f_- = 0.310 \pm 0.004$ and $f_+ = 0.0017 \pm 0.0001$. The systematic treatment ignores f_+ since variation in the other polarization will also vary f_+. For f_0 and f_-, they are each varied by 5%. This variation produces a weight for each event, which we can reweight the MC with in order to induce a MC with a 5% $t\bar{t}$ W-boson polarization uncertainty.

5.9.2 W-boson polarization in $W + jets$ events

For W-bosons in $W + jets$ events the uncertainty is also measured by varying the fraction of longitudinal and transverse polarized W-boson in the MC. The W-boson's polarization is also already measured at the LHC by CMS. [24] The polarization of W-boson are parameterized by,

$$\frac{dN}{dcos\theta^*} \propto (1 + cos^2\theta^*) + \frac{1}{2} f_0(1 + 3cos^2\theta^*) \pm (f_- - f_+)cos\theta^* \tag{5.13}$$

Where cosθ^* is the W-polarization the angle between the charged lepton and W-boson in the W's rest frame. The values for $f_0 = 0.183 \pm 0.087 \pm 0.123$ and $(f_- - f_+) = 0.240 \pm 0.036 \pm 0.031$.The fraction contribution of f_0 and $(f_- - f_+)$ are varied as function of η and W p_T . The $|\eta|$ bins are (0,1] , (1,2], (2,5], and there p_T bin are (50,100], (100,300], (300,500], and 500+. A 5% and 10% variation in the W-polarization is done and the greatest variation in the single lepton prediction in the MC is kept.

5.10 Complete Prediction Results

The three separate background predictions are combined together and compared to the number of observed yields on table 5.22 for HT $\in$ [500,750) and table 5.23 for HT $>$ 750. The results are also shown in figure 5.44 with HT $\in$ [500,750) on the left and HT $>$ 750 on the right. The three backgrounds are the

single leptons using the lepton spectrum method with $\not\!\!E_T$ smearing, the dileptons/taus prediction, and the QCD prediction. Note that QCD p_T contribution creates an over prediction. This is due to the fact that the lepton p_T tail from QCD does not contribute to the $\not\!\!E_T$ distribution observed in data. The final prediction is calculated by summing up the single lepton, dilepton/tau, and QCD $\not\!\!E_T$ predictions, but QCD p_T contribution is subtracted from the final prediction. Both the statistical and systematic uncertainties added in quadrature between the different prediction assuming they are uncorrelated. Note that single lepton contribution dominates in low $\not\!\!E_T$ bins of 150-250 and 250-350. Dilepton/tau contributes mostly to the high $\not\!\!E_T$ bins of 350-450 and 450+. QCD contributes fairly uniformly to all lepton p_T bins which produces a negative contribution the prediction. This is most prominent in the higher $\not\!\!E_T$ bins since the other contributions become smaller. From the data predictions and the observed number of event yields, no significant excess in $\not\!\!E_T$ is observed. This allows us to create exclusion limits based upon these results as shown in the next section.

5.11 mSUGRA and SMS Limits

No significant excess beyond the standard model predictions were observed in this analysis, limits on SUSY models are made in mSUGRA [19] and SMS [22] with

Table 5.22: Complete Prediction Results for $500 \leq HT < 750$. The final prediction is calculated by summing up the single lepton, dilepton/tau, and QCD $\not\!\!E_T$ predictions, but the QCD p_T contribution is subtracted from the final prediction. DY+jets contributions are taken directly from MC but contributes zero events in all signal bins. QCD systematic uncertainties are evaluated with both $\not\!\!E_T$ and p_T prediction combined and shown only on the QCD p_T row.

$500 \leq HT < 750$				
$\not\!\!E_T$ Cut [GeV]	150-250	250-350	350-450	450+
MC Expectation	92.7 ± 1.2	14.3 ± 0.5	1.71 ± 0.18	0.11 ± 0.04
Single l	$96.8 \pm 8.6 \pm 13.8$	$12.3 \pm 3.3 \pm 1.5$	$3.08 \pm 1.49 \pm 0.74$	$0.00 \pm 1.94 \pm 0.92$
Dilepton+τ	$31.1 \pm 1.2 \pm 2.1$	$4.9 \pm 0.1 \pm 0.3$	$0.76 \pm 0.08 \pm 0.12$	$0.22 \pm 0.12 \pm 0.09$
QCD $\not\!\!E_T$	0.010 ± 0.003	0.002 ± 0.001	< 0.002	< 0.002
QCD p_T	$2.0 \pm 0.3 \pm 0.9$	$1.44 \pm 0.18 \pm 0.6$	$1.10 \pm 0.13 \pm 0.5$	$0.15 \pm 0.02 \pm 0.09$
DY+jets (MC)	0	0	0	0
Pred Data	$125.9 \pm 8.7 \pm 14.0$	$15.8 \pm 3.3 \pm 1.5$	$2.74 \pm 1.50 \pm 0.78$	$0.07 \pm 1.94 \pm 1.19$
Obs Data	140	26	4	0

Table 5.23: Complete Prediction Results for $HT > 750$. The final prediction is calculated by summing up the single lepton, dilepton/tau, and QCD $\not\!\!E_T$ predictions, but the QCD p_T contribution is subtracted from the final prediction. DY+jets contributions are taken directly from MC but contributes zero events in all signal bins. QCD systematic uncertainties are evaluated with both $\not\!\!E_T$ and p_T prediction combined and shown only on the QCD p_T row.

$HT \geq 750$				
$\not\!\!E_T$ Cut [GeV]	150-250	250-35	350-450	450+
MC Expectation	30.2 ± 0.7	8.3 ± 0.4	1.72 ± 0.18	0.63 ± 0.12
Single l	$32.5 \pm 4.8 \pm 2.1$	$4.0 \pm 2.0 \pm 1.8$	$3.39 \pm 1.62 \pm 0.99$	$1.81 \pm 1.13 \pm 0.71$
Dilepton+τ	$10.9 \pm 2.5 \pm 1.4$	$8.1 \pm 2.6 \pm 1.4$	$1.45 \pm 1.35 \pm 0.25$	$0.15 \pm 0.15 \pm 0.03$
QCD $\not\!\!E_T$	0.004 ± 0.002	0.002 ± 0.002	< 0.002	< 0.002
QCD p_T	$0.41 \pm 0.07 \pm 0.18$	$0.30 \pm 0.04 \pm 0.14$	$0.23 \pm 0.03 \pm 0.09$	$0.031 \pm 0.004 \pm 0.017$
DY+jets (MC)	0	0	0	0
Pred Data	$43.0 \pm 5.4 \pm 2.5$	$11.8 \pm 3.3 \pm 2.3$	$4.61 \pm 2.11 \pm 1.02$	$1.93 \pm 1.14 \pm 0.71$
Obs Data	42	7	0	1

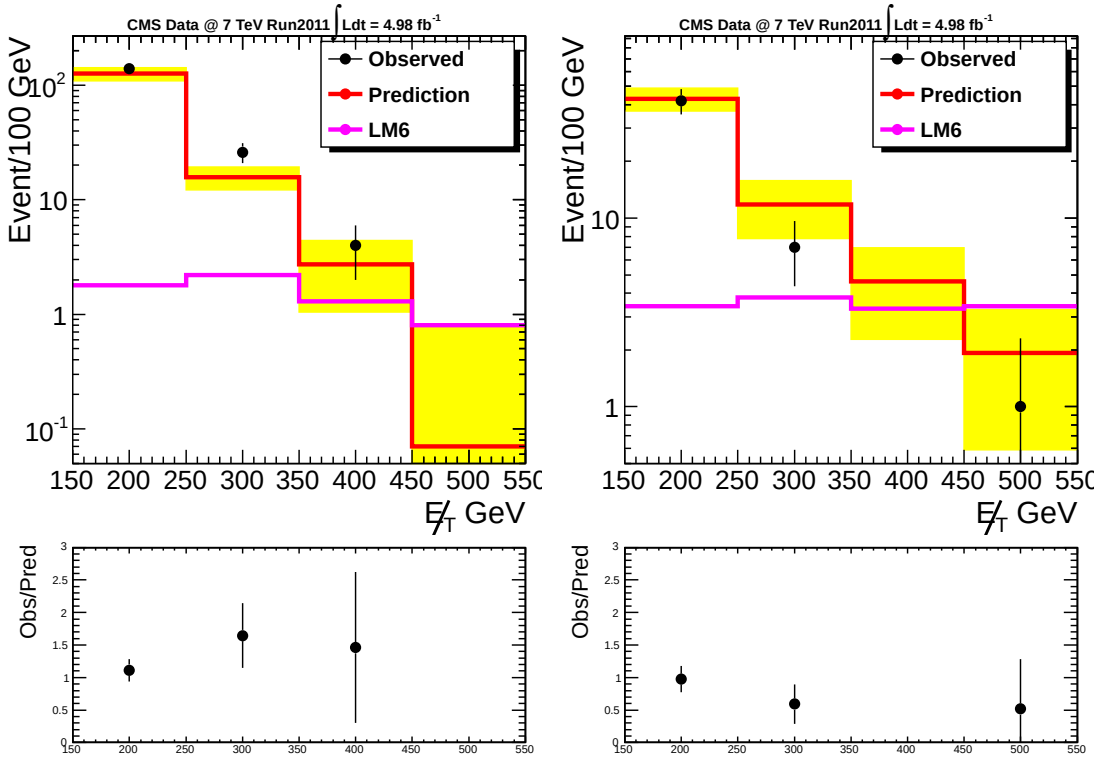

Figure 5.44: Top: Combined $\not\!\!E_T$ Prediction is shown in red histogram and yellow uncertainty bands for $500 \leq HT < 750$ (left) and $HT \geq 750$ (right). The observed number of events are shown in black points. The number of LM6 events expected from MC is in magenta. Bottom: Ratio between the observed number of events and the prediction.

the T1tttt model. The limit setting is done using CMS's Higgs combined limit tool [34] running in asymptotic mode. The limits are set using the asymptotic CL_S method. The asymptotic method approximates of the full CL_S calculation fairly well and requires much less computing time. The asymptotic CL_S method utilizes inputs from data cards which contains the SM background predictions

and their uncertainties along with the yield of the test hypothesis (the model point). The Higgs limit tool in asymptotic CL_S method mode also combines all 8 bins of the analysis result and create a single result in the form of the signal strength for that particular model. The background prediction for this analysis is separated into the three main backgrounds: single lepton, dilepton and taus, and QCD. Each background is inputted separately along its uncertainty. The uncertainties due to luminosity and trigger efficiency are taken to be 5% total for all signal points. The theory uncertainties on each model's production cross section are also included into the limit calculation which varies between 5% to 40% depending on the model point. The 95% confidence level (CL) upper limits on cross section are produced for each model point by multiplying the signal strength to the theory cross section. By comparing the 95% upper limit on cross sections compared to theory cross sections, limit curves are constructed for those using the expected and observed limits along with their respective $\pm 1\sigma$ uncertainty bands.

The mSUGRA limits in $m_{1/2}$ versus m_0 plane with $\tan\beta = 10$, $A_0 = 0$, and $sign(\mu) > 0$ are made in 20 GeV by 20 GeV bins. Figure 5.45 left shows the best acceptance $\times$ efficiency out of all 8 signal bins in the analysis in the $m_{1/2}$ versus m_0 plane. Since mSUGRA with $\tan\beta = 10$ is not a particularly bottom quark rich model, the signal efficiency of this analysis is fairly low, 2% to 0.2%. Figure 5.46 shows the 95% CL limit curves in the $m_{1/2}$ versus m_0 plane. The solid black line is

the observed limit along with its $\pm 1\sigma$ uncertainty band in dashed black. The solid red line is the expected limit along with its $\pm 1\sigma$ uncertainty band in dashed red. The best observed limit curve from our previous analysis (SUS12010) [12] without b-tagging is shown in solid blue. The observed limits from both analyses are very similar to each other. Although this analysis does not significantly increase the exclusion in mSUGRA, it is a great cross check with SUS12010 to show that the analysis performs similarly in the b-tagged case.

In the SMS model of T1tttt, each of the four tops in figure 4.5 decay into bottom quarks giving us a bottom rich environment. Figure 5.47 shows the acceptance $\times$ efficiency of this analysis in the plane of mass of the light supersymmetric partner (LSP) versus mass of the gluino. The analysis' efficiency is an order of magnitude higher than mSUGRA since each event produces four bottom quarks thus increasing the analysis' selection efficiency. Figure 5.48 shows the 95% CL upper expected limit on cross section in m_{lsp} versus $m_{\tilde{g}}$ plane. The expected limit is shown in red and the observed limit is shown in black along with their respective $\pm 1\sigma$ uncertainty bands. The best expected limit curve from a hadronic search using bottom quark jets and $\not\!\!E_T$ [31] is shown in solid blue. This analysis has a better exclusion limit in the SMS-T1tttt model due to the large branching fraction the signal events into single lepton channel.

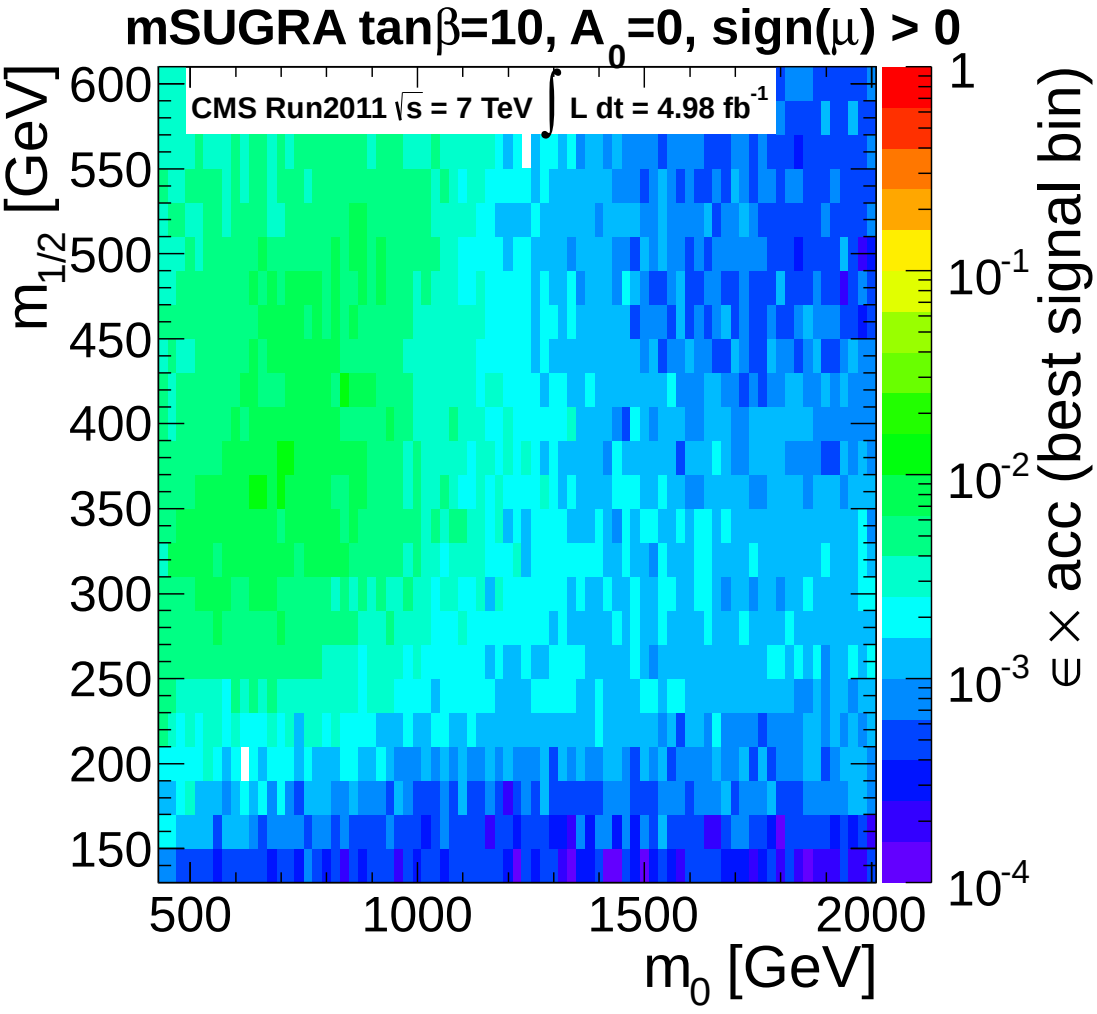

Figure 5.45: Efficiency × Acceptance for the most efficient $\not\!\!E_T$ and H_T signal bin for mSUGRA model points with $tan\beta = 10$, $A_0 = 0$, and sign(μ) > 0.

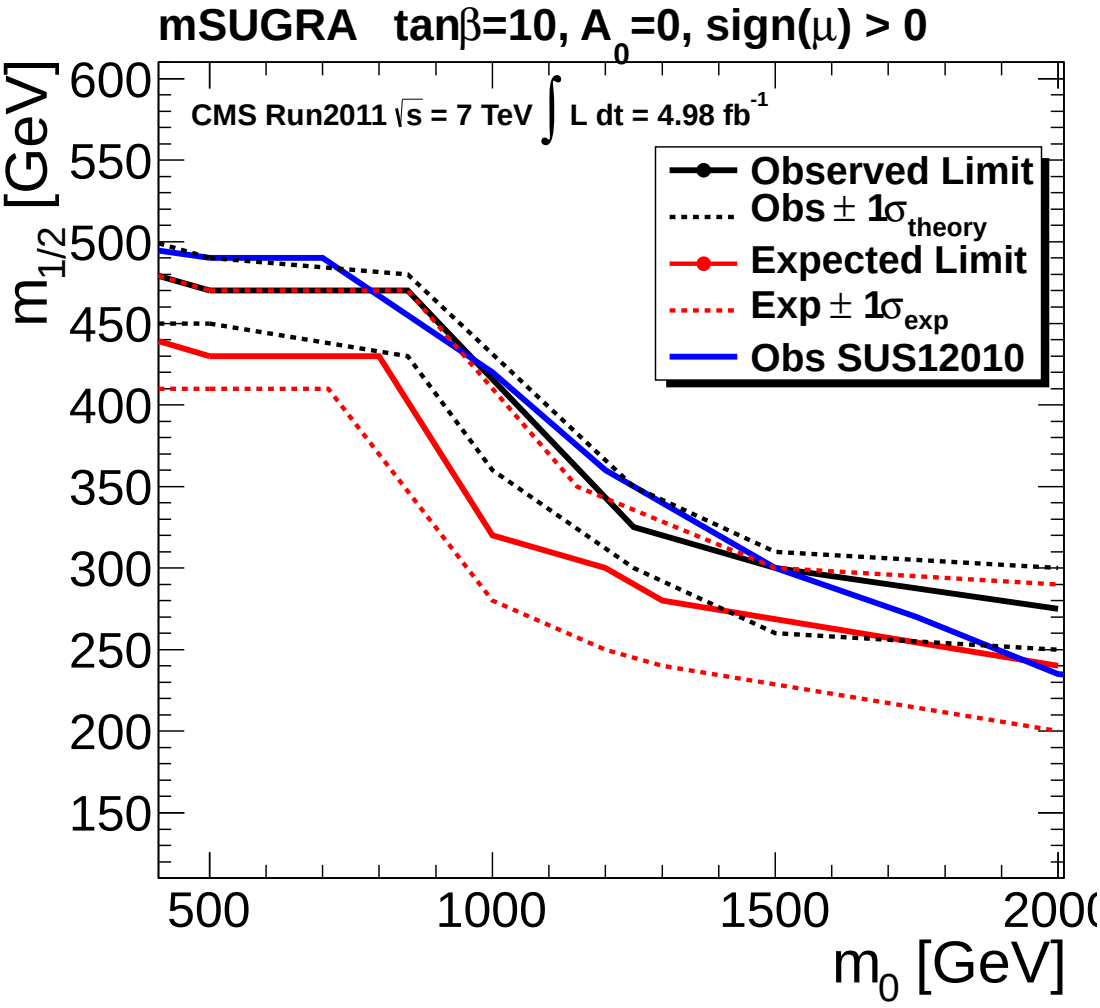

Figure 5.46: The 95% CL exclusion limit for mSUGRA models with $tan\beta = 10$, $A_0 = 0$, sign(μ) > 0. Expected limit is shown in red along with its $\pm 1\sigma$ bands in dashed red and the observed limit is shown in black along $\pm 1\sigma$ bands in dashed black.

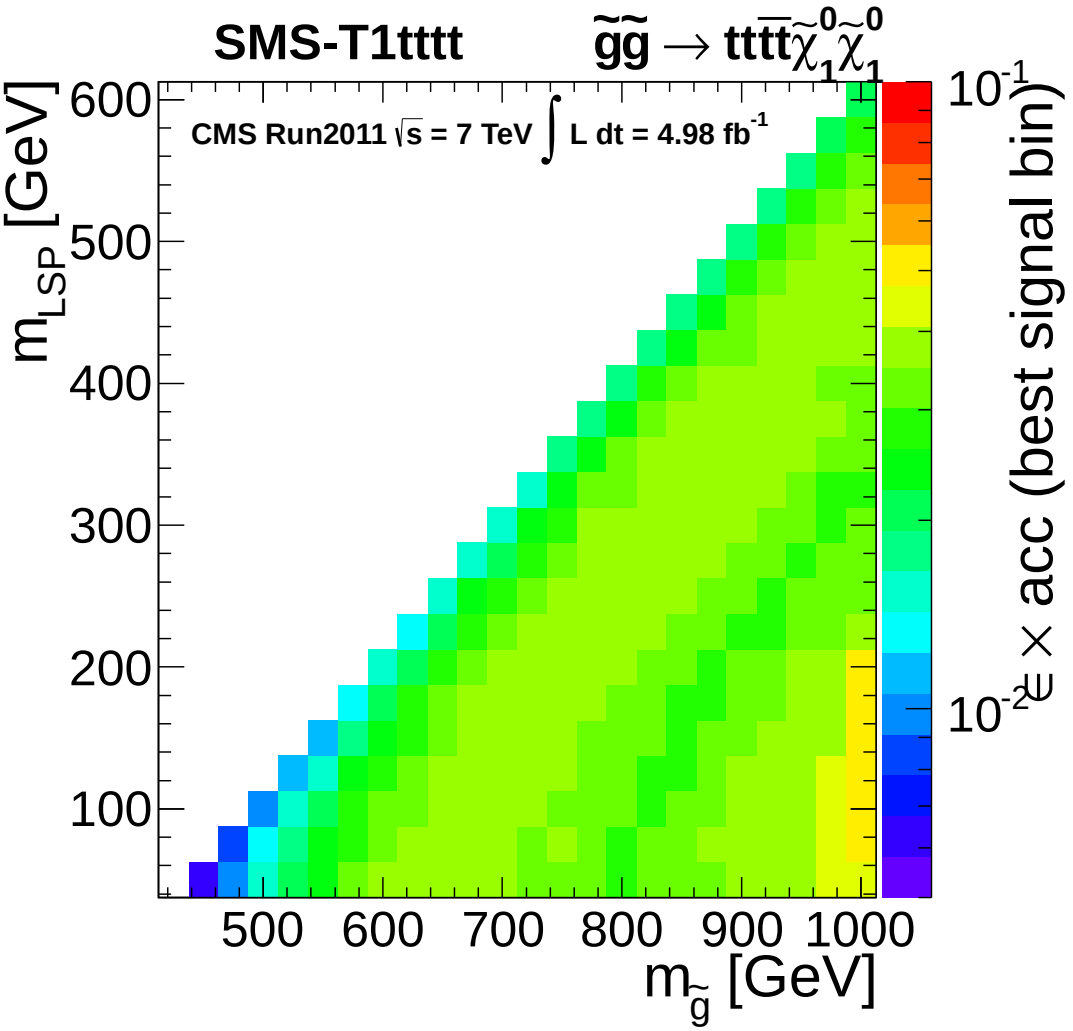

Figure 5.47: Efficiency × Acceptance for the most efficient $\not{E}_T$ and H_T signal bin for SMS-T1tttt model points.

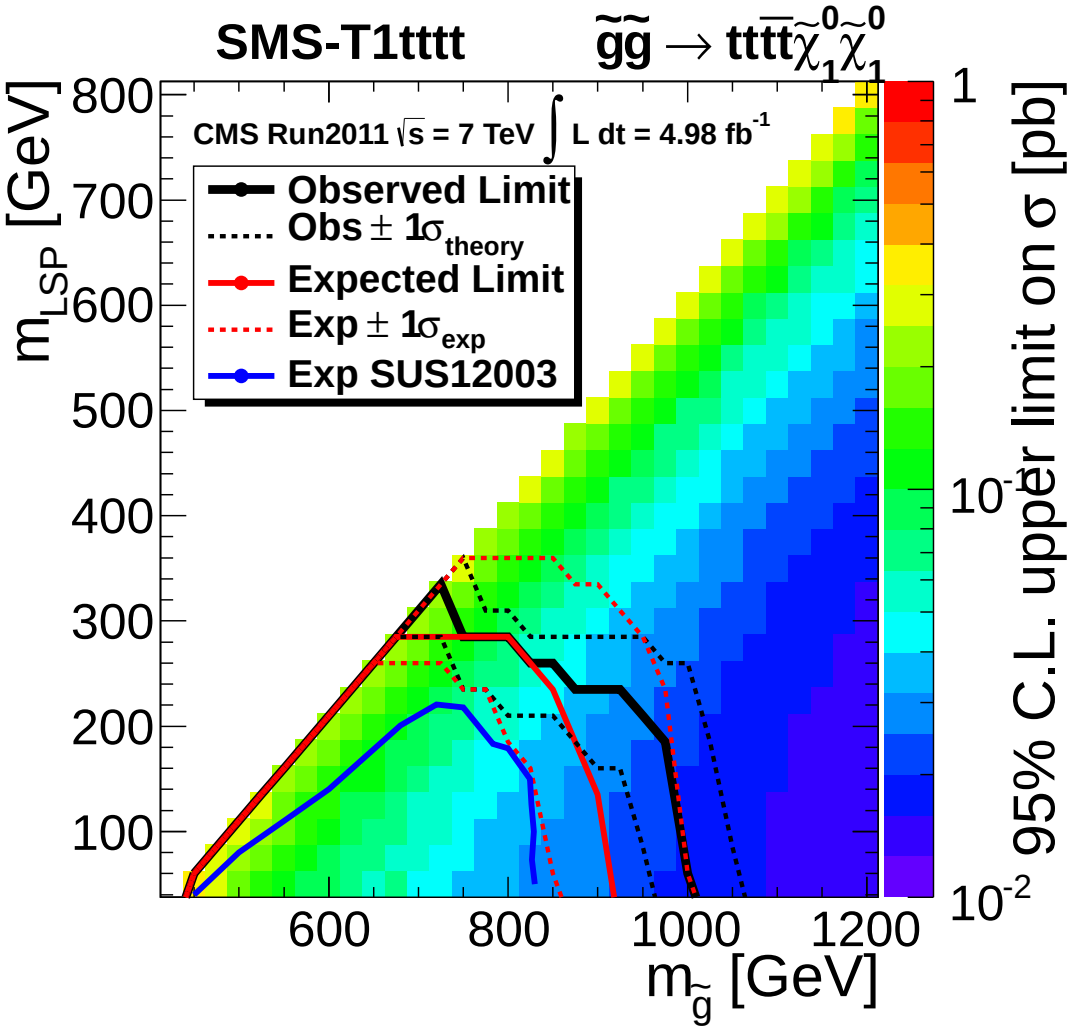

Figure 5.48: The 95% CL upper limit on cross section for each T1tttt models points. Expected limit is shown in red along with its $\pm 1\sigma$ bands in dashed red and the observed limit is shown in black along $\pm 1\sigma$ bands in dashed black.

Chapter 6

Conclusion

This thesis has presented an analysis for a search for supersymmetry in the single lepton, two b-tagged jets and $\not\!\!E_T$ at the LHC using data collected by the CMS detector. Each of the three separate background prediction methods target a specific subset of the Standard Model background. The single lepton prediction was done using the lepton spectrum method. The dilepton and tau prediction is done by emulating lost lepton, ignored lepton, and tau decay using response functions taken from Monte Carlo. The QCD prediction is made by extrapolating QCD contribution from high to low relative isolation. There were no significant excesses observed using these methods to predict the Standard Model contribution to the $\not\!\!E_T$ tail in bins of both HT and $\not\!\!E_T$. 95% confidence level limits on SUSY models in both mSUGRA and SMS-T1tttt were set using the observation and prediction from data. Due to b-tagging requirement of this analysis, this analysis is shown to be less sensitive to models in the mSUGRA plane since fewer bottom

quarks are in the decay chain of the most mSUGRA model. In a bottom rich environment such as T1tttt, this limit is driven by the lower $\not\!\!E_T$ bins in the analysis. Due to the nature of the T1tttt future analysis could benefit from lower $\not\!\!E_T$ requirements while increasing the number of b-tagged jets and HT in the event. As a generic single lepton search, the method of breaking up the single, dilepton, tau and QCD prediction gave us a clear and easy understanding of the backgrounds in a single lepton search. Each of the three method can be changed without affecting the other two allowing us to quickly reproduce future results using these methods. This analysis formed the foundation for a similar b-tagged search using the 8 TeV data collected during CMS Run2012 [32].

Bibliography

[1] LHC Project Report. http://lhc.web.cern.ch/lhc/lhc-designreport.html

[2] Duoplasmatron http://linac2.home.cern.ch/linac2/sources/source.htm

[3] LHC Machine Outreach http://lhc-machine-outreach.web.cern.ch/lhc-machine-outreach/images/complex/Cern-complex.gif

[4] CMS Collaboration. The CMS experiment at the CERN LHC, JINST 2008, 3/08/S08004

[5] CMS Colloboration. http://cdsweb.cern.ch/record/922757/files/lhcc-2006-001.pdf

[6] CMS Collaboration. http://www.phy.olemiss.edu/HEP/cms/pixel/index.html

[7] F. Halzen and A.D. Martin Quarks and Leptons: An Introductory Course in Modern Particle Physics, Wiley; 1st ed, 1984

[8] P. Higgs Broken Symmetries and the mass of gauge bosons, Physical Review Letters; vol 13, num 16, 1964

[9] W. de Boer Grand Unified Theories and Supersymmetry in Particle Physics and Cosmology; hep-ph/9402255, March 1994.

[10] DESY Collaboration. http://scienceblogs.com/startswithabang/2012/11/14/is-there-any-particle-physics-beyond-the-standard-model/.

[11] Victor Pavlunin. Modeling missing transverse energy in V+jets at CERN LHC, Phys. Rev. D 81, 035005 (2010)

[12] CMS Collaboration. Search for supersymmetry in pp collisions at 7 TeV in events with a single lepton, jets, and missing transverse momentum with 4.98 fb^{-1} of integrated luminosity, *preprint: CERN-PH-EP-2012-348*.

[13] S. Braibant et al. Investigation of design parameters for radiation hard silicon microstrip detectors Nucl. Instrum. Meth. A485 (2002) 343-361.

[14] CMS Collaboration. https://twiki.cern.ch/twiki/bin/view/ CMSPublic/SWGuideCutBasedElectronID.

[15] W. Adam, R. Frhwirth, *et al.*. Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC,Journal of Physics G: Nuclear and Particle Physics, **31** no. 9, N9, 2005.

[16] CMS Collaboration. https://twiki.cern.ch/twiki/bin/view/CMS/SusyRA4SingleLeptonDocument .

[17] CMS Collaboration. Identification of b-quark jets with the CMS experiment, arXiv:1211.4462 ; CMS-BTV-12-001 ; CERN-PH-EP-2012-262.

[18] CMS Collaboration. MET Optional Filters. https://twiki.cern.ch/twiki/bin/view/CMS/MissingETOptionalFilters.

[19] Steven Martin. A Supersymmetry Primer, updated Sept. 6, 2011. http://arxiv.org/pdf/hep-ph/9709356v6.pdf.

[20] Frank Wilczek. Asymptotic freedom: From Paradox to paradigm, PNAS, June 14, 2005 vol 102 no. 24.

[21] G. Bertone et al. Particle dark matter: evidence, candidates and constraints. Physics Reports 405 (2005) 279-390.

[22] CMS Collaboration. Interpretation of Searches for Supersymmetry with Simplified Models, http://cdsweb.cern.ch/record/1445580/files/SUS-11-016-pas.pdf.

[23] CDF and D0 Collaborations Combination of CDF and D0 measurements of the W boson helicity in top quark decays, Physical Review D85, 071106(R) 2012.

[24] CM Collaboration. Measurement of the Polarization of W Bosons with Large Transverse Momenta in W+Jets Events at the LHC, CERN-PH-EP/2011-043, 2011/04/20.

[25] Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and Missing E_T. *CMS Physics Analysis Summary*, PFT-09-001, 2009.

[26] J. Alwall et al. MadGraph/MadEvent v4: The New Web Generation. *JHEP*, 09:028, 2007.

[27] T. Sjöstrand, S. Mrenna, and P. Z. Skands. PYTHIA 6.4 Physics and Manual; v6.420, tune D6T. *JHEP*, 05:026, 2006.

[28] CMS Collaboration. https://twiki.cern.ch/twiki/bin/view/CMS/JetIDDocumentation.

[29] CMS Collaboration. https://twiki.cern.ch/twiki/bin/view/ CMSPublic/WorkBookJetEnergyCorrections.

[30] W. Adam et al.. Reconstruction, identification, and trigger efficiencies for SUSY searches in the single lepton channel in 2011, AN-2011/536, 2011.

[31] CMS Collaboration. https://twiki.cern.ch/twiki/bin/view/ CMSPublic/PhysicsResultsSUS12003.

[32] CMS Collaboration. http://cms-physics.web.cern.ch/cms-physics/public/SUS-13-007-pas.pdf.

[33] Greg Landsberg. Bayesian 95% CL limit and expected (average) 95% CL limit calculator for cross-section with Poisson statistics. http://ihp-lx.ethz.ch/Stamet/lectureNotes/cl95cms.c.

[34] CMS Collaboration. https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideHiggsAnalysisCombinedLimit.