

UNIVERSITY OF CALIFORNIA, SAN DIEGO

**A Search for New Physics in Final States With Two Opposite-Sign  
Same-Flavor Leptons, Jets, and Missing Transverse Energy in  
Proton-Proton Collisions with a Center of Mass Energy of 13 TeV**

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

in

Physics

by

Charles Vincent Welke

Committee in charge:

Professor Avraham Yagil, Chair  
Professor Claudio Campagnari  
Professor John McGreevy  
Professor Yitzhak Tor  
Professor Frank Würthwein

2016

Copyright  
Charles Vincent Welke, 2016  
All rights reserved.

The Dissertation of Charles Vincent Welke is approved,  
and it is acceptable in quality and form for publication  
on microfilm and electronically:

---

---

---

---

---

Chair

University of California, San Diego

2016

## DEDICATION

This thesis is dedicated to my family, friends and the entire surf and  
turf high energy physics group.

## EPIGRAPH

*Everything in life is just for a while.*

—Philip K. Dick, A Scanner Darkly

## TABLE OF CONTENTS

Signature Page . . . . .	iii
Dedication . . . . .	iv
Epigraph . . . . .	v
Table of Contents . . . . .	vi
List of Figures . . . . .	ix
List of Tables . . . . .	xiv
Acknowledgements . . . . .	xvii
Vita . . . . .	xviii
Abstract of the Dissertation . . . . .	xix
Chapter 1      Introduction . . . . .	1
1.1     Particle Physics . . . . .	1
1.2     The Standard Model . . . . .	1
1.2.1     Problems with The Standard Model . . . . .	4
1.3     Supersymmetry (SUSY) . . . . .	4
1.4     Searches for SUSY in final states with a Z boson . . . . .	5
Chapter 2      The LHC and The CMS Experiment . . . . .	7
2.1     The Large Hadron Collider . . . . .	7
2.2     Physics Processes and Cross Sections . . . . .	9
2.3     Particle Interactions With Matter . . . . .	11
2.4     The Compact Muon Solenoid . . . . .	15
2.4.1     Silicon Tracker . . . . .	17
2.4.2     Electromagnetic Calorimeter . . . . .	18
2.4.3     Hadronic Calorimeter . . . . .	22
2.4.4     Muon System . . . . .	24
2.5     Physics objects . . . . .	27
2.5.1     Particle Flow . . . . .	27
2.5.2     Vertex Determination and Pileup . . . . .	29
2.5.3     Isolation . . . . .	30
2.5.4     leptons and photons . . . . .	31
2.5.5     jets . . . . .	32

Chapter 3	Missing Transverse Momentum . . . . .	35
3.1	raw $E_T^{\text{miss}}$ . . . . .	35
3.1.1	Sources of $E_T^{\text{miss}}$ . . . . .	36
3.2	Type-1 $E_T^{\text{miss}}$ . . . . .	36
3.2.1	Data vs. MC Comparison . . . . .	38
3.2.2	Data-driven $E_T^{\text{miss}}$ Predictions . . . . .	43
Chapter 4	Analysis Selections . . . . .	45
4.1	Triggers and Datasets . . . . .	45
4.2	Simulations from Monte-Carlo . . . . .	48
4.2.1	Data vs. MC Comparison . . . . .	51
4.3	Event Selection . . . . .	54
4.3.1	Vertex Selection . . . . .	54
4.3.2	MET filters . . . . .	54
4.3.3	Lepton Selection . . . . .	55
4.3.4	Lepton Isolation . . . . .	57
4.3.5	Lepton Scale Factors . . . . .	61
4.3.6	Electron Selection . . . . .	64
4.3.7	Muon Selection . . . . .	67
4.3.8	Photons . . . . .	69
4.3.9	MET . . . . .	71
4.3.10	Jets . . . . .	71
4.3.11	$H_T$ . . . . .	73
4.3.12	Signal Region Definitions . . . . .	74
4.3.13	Selection Summary . . . . .	76
Chapter 5	Background Estimation Methods . . . . .	79
5.1	Estimating the Z + jets Background with $E_T^{\text{miss}}$ Template . . . . .	79
5.1.1	Systematic Uncertainties in the $E_T^{\text{miss}}$ Template Prediction . . . . .	83
5.2	Estimating the Flavor-Symmetric Background . . . . .	86
5.2.1	$R_{SF/OF}$ . . . . .	87
5.2.2	Direct measurement of $R_{SF/OF}$ . . . . .	90
5.2.3	Measureing $R_{SF/OF}$ using the Factorization Method . . . . .	91
5.2.4	Results of factorizaton method . . . . .	97
5.2.5	Combination of the Methods . . . . .	98
5.2.6	MC closure test of the FS background prediction . . . . .	100
5.3	Estimating WZ, ZZ and other rare SM backgrounds with Simulation . . . . .	103
Chapter 6	Results . . . . .	107
6.1	Results in Signal Region A . . . . .	107
6.2	Results in Signal Region B . . . . .	109
6.3	ATLAS-like signal regions . . . . .	111

6.4	Systematic Uncertainties on Signal Model . . . . .	113
6.5	Interpretation of the Results . . . . .	116
Chapter 7	Summary and Conclusions . . . . .	118
Bibliography	. . . . .	120

## LIST OF FIGURES

Figure 1.1:	A diagram is shown with the particles making up the Standard Model along with the attributes of each particle.	3
Figure 1.2:	A diagram is shown with the particles making up the Standard Model on the left, and their SUSY counterparts on the right.	5
Figure 1.3:	A diagram showing a SUSY process where two protons collide and the result is pair-production of two gluinos, where each decays to a pair of quarks and a neutralino which subsequently decays to a Z boson and gravitino.	6
Figure 2.1:	The LHC can be seen along with the 4 main experiments at different points along the beamline.	8
Figure 2.2:	The cross section values for various standard model processes is shown in this figure as a function of the center of mass energy of two protons colliding.	10
Figure 2.3:	The peak instantaneous luminosity per day during the 2015 data taking period is shown on the left and the total integrated luminosity is shown on the right in this figure.	11
Figure 2.4:	Stopping power of copper as a function of muon momentum. Typical momenta of muons in this analysis are between 20-200 GeV.	14
Figure 2.5:	A diagram showing the pixel component of the tracker subdetector.	17
Figure 2.6:	A cross-sectional in the x-y plane of the CMS tracking system is shown here. The tracker provides excellent coverage for reconstruction of tracks from charged particles for $ \eta  < 2.5$ .	19
Figure 2.7:	A diagram showing material budget of the tracker as a function of $\eta$ on the x-axis, and radiation lengths (left) and interaction lengths (right).	20
Figure 2.8:	A cross-sectional view of the CMS ECAL is shown in this figure, with values of $\eta$ shown that determine the coverage of each subdetector.	20
Figure 2.9:	A feynman diagram is shown depicting electron-positron pair production from a photon initial state.	21
Figure 2.10:	A feynman diagram is shown of a process where an electron scatters off of an incoming photon, and then radiates a photon through bremsstrahlung.	22
Figure 2.11:	A cartoon depiction of a set of drift tubes is shown in this figure where a charged particle passing through the system ionizes the gas, and the ions are collected on the wires to detect a signal.	25
Figure 2.12:	A cartoon depiction of a cathode strip chamber is shown in this figure.	26

Figure 2.13: A cartoon depiction of a resistive plate chamber is shown in this figure. . . . .	27
Figure 3.1: $E_T^{\text{miss}}$ distributions for events passing the dijet selection without 2012 cleaning algorithms applied (open markers), with 2012 cleaning algorithms applied including the one based on jet identification requirements (filled markers), and events from MC (filled histograms). . . . .	37
Figure 3.2: The $E_T^{\text{miss}}$ distribution is shown for charged PF candidates only. The top row shows the barrel region on the left and transition region between the barrel and endcap on the right, the second row shows the endcap region including the tracker on the left and endcap region excluding the tracker on the right and the bottom row shows the HF region only. . . . .	40
Figure 3.3: The $E_T^{\text{miss}}$ distribution is shown for neutral electromagnetic PF candidates only. The top row shows the barrel region on the left and transition region between the barrel and endcap on the right, the second row shows the endcap region including the tracker on the left and endcap region excluding the tracker on the right and the bottom row shows the HF region only. . . . .	41
Figure 3.4: The $E_T^{\text{miss}}$ distribution is shown for neutral hadronic PF candidates only. The top row shows the barrel region on the left and transition region between the barrel and endcap on the right, the second row shows the endcap region including the tracker on the left and endcap region excluding the tracker on the right and the bottom row shows the HF region only. . . . .	42
Figure 3.5: The $E_T^{\text{miss}}$ distribution is shown with data vs. MC for events while only requirement two leptons with $ M_{\ell\ell} - M_Z  < 10\text{GeV}$ . The top row shows ee events on the left and $\mu\mu$ events on the left, and the bottom row shows ee+ $\mu\mu$ events together. The MC is normalized to data. . . . .	44
Figure 4.1: Data vs. MC comparison showing $m_{\ell\ell}$ with ee events on the left and $\mu\mu$ events on the right. For this analysis, a cut is required that $ M_{\ell\ell} - M_Z  < 10\text{GeV}$ . . . . .	51
Figure 4.2: data vs. MC comparison showing $m_{\ell\ell}$ with ee+ $\mu\mu$ events on the left and $e\mu$ events on the right. . . . .	51
Figure 4.3: data vs. MC comparison showing $N_{\text{jets}}$ with ee events on the left and $\mu\mu$ events on the right. . . . .	52
Figure 4.4: data vs. MC comparison showing $N_{\text{b-jets}}$ with ee events on the left and $\mu\mu$ events on the right. . . . .	52
Figure 4.5: data vs. MC comparison showing $H_T$ with ee events on the left and $\mu\mu$ events on the right. . . . .	53

Figure 4.6: data vs. MC comparison showing Z $p_T$ with ee events on the left and $\mu\mu$ events on the right. . . . .	53
Figure 4.7: The isolation distribution is shown for electrons on the left in simulated Z + jets MC events. On the right, the efficiency is shown when integrating all bins to the left of the value on the x-axis. . . . .	59
Figure 4.8: The isolation distribution is shown for muons on the left in simulated Z + jets MC events. On the right, the efficiency is shown when integrating all bins to the left of the value on the x-axis. . . . .	59
Figure 4.9: Electron scale factors as a function of $p_T$ and $\eta$ . The scale factor is applied as an event weight once per electron in the event. . . . .	62
Figure 4.10: Muon scale factors as a function of $p_T$ and $\eta$ . The left plot shows the scale factors associated with the ID selection criteria, and the right plot shows the scale factors associated with the isolation selection criteria. Each scale factor is applied as an event weight once per muon in the event. . . . .	62
Figure 4.11: Fastsim scale factors as a function of $p_T$ and $\eta$ . The left plot shows the scale factors applied per electron, and the right plot shows the scale factors applied per muon. Each scale factor is applied as an event weight once per lepton in the event. . . . .	63
Figure 4.12: Kinematic variables used to distinguish real electrons from fake electrons are shown with electrons in the barrel on the left and electrons in the endcap on the right. . . . .	65
Figure 4.13: The output of the electron MVA is shown comparing the distributions of electrons matched with generator level electrons (blue) vs. jets that fake electrons. The output in the barrel is shown in the left and endcap on the right. . . . .	66
Figure 4.14: Output of the electron MVA is shown comparing electrons in data vs. MC with electrons in the barrel on the left and the endcap on the right. The electrons shown are the probe leg when using a “Tag and Probe” technique. . . . .	67
Figure 4.15: Electron identification efficiency in data (top) and data to MC efficiency ratios (bottom) measured for the identification criteria based on a MVA discriminant with an average efficiency of 80%. The efficiency is measured with the tag and probe method and shown in three pseudorapidity ranges as a function of the electron transverse momentum. . . . .	68
Figure 4.16: A plot of the impact parameter for the secondary vertex with respect to the primary vertex which is used in order to differentiate jets from heavy and light flavor decay. . . . .	73
Figure 4.17: A plot of the 3-D flight distance of the secondary vertex with respect to the primary vertex which is used in order to differentiate jets from heavy and light flavor decay. . . . .	74

Figure 4.18: A plot of the mass of the jet originating at the secondary vertex which is used in order to differentiate jets from heavy and light flavor decay. . . . .	75
Figure 4.19: A plot showing a jet forming away from the primary vertex of a heavy flavor parton, for example a b or c quark. . . . .	76
Figure 4.20: The distribution showing the CSVv2 Discriminator is shown comparing data vs. MC. The MC is split by jet parton flavor. . . . .	77
 Figure 5.1: The $p_T$ distributions are shown for the Z+jets (black) and $\gamma$ +jets (red) events in signal region B where at least 1 jet is required to be b-tagged. The distributions are normalized to unit area. . . . .	82
Figure 5.2: The $E_T^{\text{miss}}$ distributions are shown before reweighting the $\gamma$ +jets sample in $p_T$ (red) and after reweighting (black) for $\gamma$ +jets events in signal region B where at least 1 jet is required to be b-tagged. The distributions are normalized to unit area. . . . .	83
Figure 5.3: The closure test for Signal region B where at least 1 jet is required to be b-tagged is shown where simulated $\gamma$ +jets events are used to predict the $E_T^{\text{miss}}$ for simulated Z+jets events. The distributions are normalized to $1 \text{ fb}^{-1}$ of data. . . . .	84
Figure 5.4: Direct measurement of $R_{SF/OF}$ in the control region for data (black) and $t\bar{t}$ MC (red). Values in the signal region are shown in blue for $t\bar{t}$ MC. Central rapidity is shown on the left and forward rapidity on the right. The green band represents the overall uncertainty on the measured value of $R_{SF/OF}$ . . . . .	91
Figure 5.5: $0.5(r_{\mu e} + 1/r_{\mu e})$ dependency for central leptons shown as a function of $N_{\text{jets}}$ , $N_{\text{vtx}}$ , lepton $p_T$ , $m_{\ell\ell}$ , $E_T^{\text{miss}}$ , and $N_{b-\text{tags}}$ . The uncertainty due to the assigned 10% systematic uncertainty on $r_{\mu/e}$ is indicated by the orange band. . . . .	94
Figure 5.6: $0.5(r_{\mu e} + 1/r_{\mu e})$ dependency for forward leptons shown as a function of $N_{\text{jets}}$ , $N_{\text{vtx}}$ , lepton $p_T$ , $m_{\ell\ell}$ , $E_T^{\text{miss}}$ , and $N_{b-\text{tags}}$ . The uncertainty due to the assigned 20% systematic uncertainty on $r_{\mu/e}$ is indicated by the orange band. . . . .	94
Figure 5.7: Ratio of measured to true trigger efficiency for combined SF and OF triggers as a function of $H_T$ measured in $t\bar{t}$ simulation. A value of 1 indicates no bias due to the choice of the supporting triggers.	95
Figure 5.8: $R_T$ dependency for central leptons shown as a function of $N_{\text{jets}}$ , $N_{\text{vtx}}$ , lepton $p_T$ , $m_{\ell\ell}$ , $E_T^{\text{miss}}$ , and $N_{b-\text{tags}}$ . The assigned 6.4% systematic uncertainty on $R_T$ is indicated by the orange band. . . . .	97
Figure 5.9: $R_T$ dependency for forward leptons shown as a function of $N_{\text{jets}}$ , $N_{\text{vtx}}$ , lepton $p_T$ , $m_{\ell\ell}$ , $E_T^{\text{miss}}$ , and $N_{b-\text{tags}}$ . The assigned 6.4% systematic uncertainty on $R_T$ is indicated by the orange band. . . . .	97

Figure 5.10: MC closure test for central (left) and forward (right) in the $m_{\ell\ell}$ variable. The blue histogram shows the MC prediction from the OF sample by multiplying with $R_{SF/OF}$ whereas the black markers correspond to the observation in the SF sample. This exercise is done on $t\bar{t}$ MC only, and is shown for all number of b-tagged jet bins. The top row shows the closure test for inclusive, the middle row 0 b-tagged jets and the bottom row $\geq 1$ jets. . . . .	102
Figure 5.11: The $E_T^{\text{miss}}$ and $N_{\text{jets}}$ distributions are shown for data vs. MC in the 3-lepton control region. We require $E_T^{\text{miss}} > 50$ GeV for the events shown in the $N_{\text{jets}}$ distribution. See Tables 5.11 and 5.12 for yields. . . . .	104
Figure 5.12: The $E_T^{\text{miss}}$ and $N_{\text{jets}}$ distributions are shown for data vs. MC in the 4-lepton control region. See Tables 5.13 and 5.14 for yields. . . . .	105
Figure 6.1: The $E_T^{\text{miss}}$ distribution is shown for data vs. the data-driven predictions in signal region B. The left plot shows the prediction when requiring $N_{\text{b-jets}} = 0$ , and the right plot shows the prediction when requiring $N_{\text{b-jets}} \geq 1$ . The dashed line in the plot represents the full uncertainty. See tables 6.1 and 6.2 for yields. . . . .	108
Figure 6.2: The $E_T^{\text{miss}}$ distribution is shown for data vs. the data-driven predictions in signal region B. The left plot shows the prediction when requiring $N_{\text{b-jets}} = 0$ , and the right plot shows the prediction when requiring $N_{\text{b-jets}} \geq 1$ . The dashed line in the plot represents the full uncertainty. See tables 6.3 and 6.4 for yields. . . . .	110
Figure 6.3: The $E_T^{\text{miss}}$ distribution is shown for data vs. the data-driven predictions in the ATLAS signal region. See Tables 6.1 and 6.5 for yields. . . . .	112
Figure 6.4: The ISR $p_T$ in data and MC is shown for $Z + \text{jets}$ (left) and $t\bar{t}$ (right). . . . .	115
Figure 6.5: Exclusion contours are shown when we interpret the results of this analysis in the SUSY model described in section 1.4. Everything to the left of the red (black) dotted line shows the masses which are excluded by the expected (observed) limit. . . . .	117

## LIST OF TABLES

Table 2.1:	Radiation ( $X_0$ ) and interaction lengths ( $\lambda_{int}$ ) for various materials that make up different sub-detectors in CMS. . . . .	13
Table 2.2:	Material budget of tracker and calorimeters in CMS at $ \eta  = 0$ as a function of radiation ( $X_0$ ) and interaction ( $\lambda_{int}$ ) lengths. . . . .	13
Table 3.1:	Definitions for the $\eta$ region chosen to define each subdetector region are listed in this table. . . . .	39
Table 4.1:	A list of all triggers used in this analysis is shown in the table below. The signal triggers all require at least two same flavor leptons, The control triggers require either at least one photon, or at least two different flavor leptons. . . . .	47
Table 4.2:	A list of MC samples used in this analysis are shown in this table. The MC samples used for background studies are all made with different generators, depending on which physics process that is being modeled. The signal MC is produced using the FastSim package.	50
Table 4.3:	Effective area values for muons derived in separate regions of $\eta$ are shown in this table. . . . .	58
Table 4.4:	Effective area values for electrons derived in separate regions of $\eta$ are shown in this table. . . . .	58
Table 4.5:	All the requirements for electrons with $p_T > 10$ GeV are shown in the table below. The electrons in the signal region are required to have $p_T > 20$ GeV. . . . .	67
Table 4.6:	A summary of the muons selection requirements is shown in this table for all muons used in this analysis. . . . .	69
Table 4.7:	A summary of the photon isolation requirements is shown in this table for all photons used in the template method prediction. . .	70
Table 4.8:	A summary of the photon selection requirements is shown in this table for all photons used in the template method prediction. . .	71
Table 4.9:	All the criteria for a jet to pass the loose ID requirement is shown in this table. . . . .	72
Table 4.10:	A summary of the jet selection requirements is shown in this table for all jets used in this analysis. . . . .	72
Table 4.11:	A list of all selections used to define the preselection and signal regions is shown in the following table. . . . .	78
Table 5.1:	A list of single- $\gamma$ triggers used in this analysis and their prescales is shown in this table. . . . .	81

Table 5.2: Results of the MC closure test shown for signal region B where at least 1 jet is required to be b-tagged. The systematic uncertainty for each region is chosen to cover the largest descrepancy between the $\gamma + \text{jets}$ prediction of the $Z + \text{jets}$ MC for each $E_T^{\text{miss}}$ region, or the the statistical uncertainty whichever is larger. All uncertainties shown are statistical only. . . . .	85
Table 5.3: Systematic uncertainties derived from the MC closure test shown for all the on-Z signal regions. The uncertainties for each region are chosen to cover the largest descrepancy between the $\gamma + \text{jets}$ prediction of the $Z + \text{jets}$ MC for each $E_T^{\text{miss}}$ region, or the statistical uncertainty whichever is larger. . . . .	85
Table 5.4: Systematic uncertainties on the normalization procedure are listed in the following table. . . . .	86
Table 5.5: Observed event yields in the control region and the resulting values of $R_{SF/OF}$ , $R_{ee/OF}$ , and $R_{\mu\mu/OF}$ . The results are shown separately for the central and forward lepton selection and the same quantities derived on simulation are shown for comaprison. . . . .	92
Table 5.6: Result of the calculation of $r_{\mu/e}$ . The observed event yields are shown in the Drell–Yan control region for the central and forward lepton selection in the ee and $\mu\mu$ channels and the resulting values of $r_{\mu/e}$ . The same quantaties derived from simulation are shown for comparison. . . . .	93
Table 5.7: Trigger efficiencies for data and MC with OS, $p_T > 20(20)$ GeV and $H_T > 400$ GeV for central and forward region seperated. . . . .	96
Table 5.8: All the calculated values for $r_{\mu/e}$ , $R_T$ , and $R_{SF/OF}$ are shown in this table. . . . .	98
Table 5.9: The results of the factorization method, and from direct measurement of $R_{SF/OF}$ are shown. The values from these two methods are combined using a weighted average to get the value for $R_{SF/OF}$ in the central and forward regions separately. These values are then combined to get the final value of $R_{SF/OF} = 1.05 \pm 0.04$ . . . . .	99
Table 5.10: Event yields in the signal region in simulation for both SF and OF lepton pairs. The OF yield is multiplied with $R_{SF/OF}$ . The quoted uncertainties are those of the MC counts in the signal region only. . . . .	101
Table 5.11: Yields in the 3-lepton control region binned in $E_T^{\text{miss}}$ . Uncertainties for each region are statistical only. . . . .	104
Table 5.12: Yields in the 3-lepton control region split by $N_{\text{jets}}$ . Uncertainties for each region are statistical only. . . . .	105
Table 5.13: Yields in the 4-lepton control region binned in $E_T^{\text{miss}}$ . Uncertainties for each region are statistical only. . . . .	106
Table 5.14: Yields in the 4-lepton control region split by $N_{\text{jets}}$ . Uncertainties for each region are statistical only. . . . .	106

Table 6.1: Results are shown for signal region A when requiring a b-veto. Systematic uncertainties for each region are included in the total uncertainty. . . . .	108
Table 6.2: Results are shown for signal region A where at least 1 b-tagged jet is required. Systematic uncertainties for each region are included in the total uncertainty. . . . .	108
Table 6.3: Results are shown for signal region B when requiring a b-veto. Systematic uncertainties for each region are included in the total uncertainty. . . . .	109
Table 6.4: Results are shown for signal region B when requiring at least 1 b-tagged jet. Systematic uncertainties for each region are included in the total uncertainty. . . . .	109
Table 6.5: Results are shown for the ATLAS signal region. Systematic uncertainties for each region are included in the total uncertainty. . . . .	111
Table 6.6: All systematic uncertainties of the expected signal yield are shown in this table. . . . .	113

## ACKNOWLEDGEMENTS

## VITA

2009	B. S. in Applied Mathematics, University of New Mexico
2011	M. S. in Physics, University of California, San Diego
2016	Ph. D. in Physics, University of California, San Diego

## PUBLICATIONS

Search for new Physics in Final States with two Opposite-Sign Same-Flavor leptons, jets, and MET in pp collisions at  $\sqrt{s} = 13$  TeV, *CMS Collaboration*, Physics Analysis Summary (2015), CMS-SUS-15-011

Search for physics beyond the standard model in events with two leptons, jets, and missing transverse momentum in pp collisions at  $\sqrt{s} = 8$  TeV, *CMS Collaboration*, JHEP 04 (2015) 124 [arXiv:1502.06031 [hep-ex]]

Searches for electroweak production of charginos, neutralinos, and sleptons decaying to leptons and W, Z, and Higgs bosons in pp collisions at 8 TeV, *CMS Collaboration*, Eur. Phys. J. C 74 (2014) 3036 [arXiv:1405.7570 [hep-ex]]

Search for electroweak chargino and neutralino production in channels with Higgs, Z, and W bosons in pp collisions at 8 TeV, *CMS Collaboration*, Phys. Rev. D 90 (2014) 092007 [arXiv:1409.3168 [hep-ex]]

## ABSTRACT OF THE DISSERTATION

### **A Search for New Physics in Final States With Two Opposite-Sign Same-Flavor Leptons, Jets, and Missing Transverse Energy in Proton-Proton Collisions with a Center of Mass Energy of 13 TeV**

by

Charles Vincent Welke

Doctor of Philosophy in Physics

University of California, San Diego, 2016

Professor Avraham Yagil, Chair

A search is performed in final states with  $Z \rightarrow \ell\ell$ , jets, and missing transverse momentum using data collected with the CMS detector at the CERN LHC in 2015 corresponding to an integrated luminosity of  $2.3 \text{ fb}^{-1}$  of pp collisions at  $\sqrt{s} = 13 \text{ TeV}$ . Signal regions are defined to maximize sensitivity to new physics including one to search where ATLAS reported a  $3.0 \sigma$  excess at 8 TeV. The observations in all regions are consistent with the predicted standard model backgrounds, and the results are interpreted in the context of simplified models of supersymmetry.

# **Chapter 1**

## **Introduction**

### **1.1 Particle Physics**

The goal of particle physics is to answer the question, what is matter composed of and how do different types of matter interact? Using the theory of the Standard Model, physicists have been very successful in answering this question for a variety of interactions. The standard model isn't able to explain everything however, for example the observation of dark matter and dark energy in the universe are not explained by the standard model. There are many theories that have been proposed which can explain these phenomena. Supersymmetry is an extension of the standard model which proposes an explanation for dark matter.

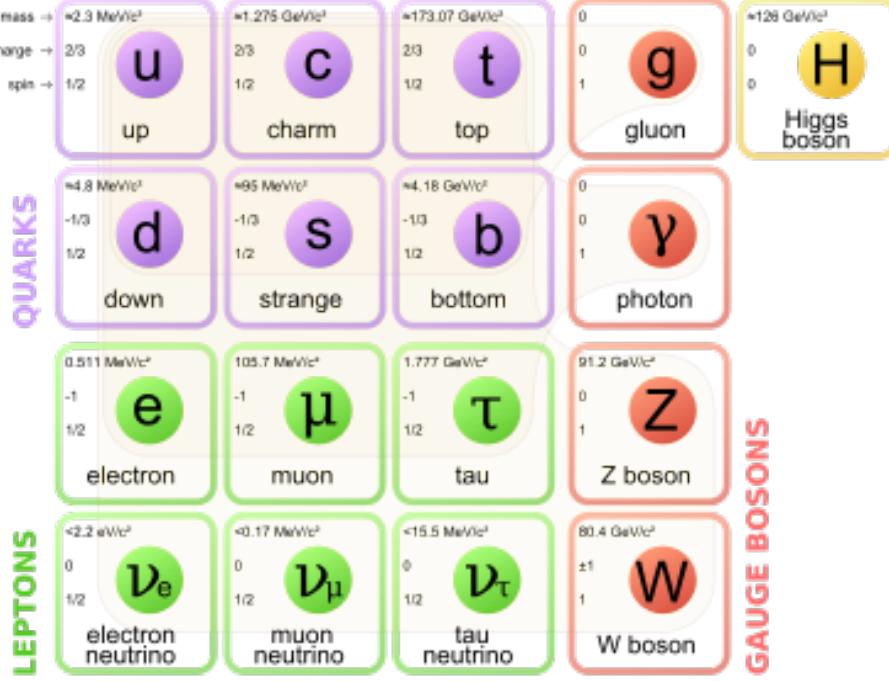
### **1.2 The Standard Model**

The standard model of particle physics is one of the most successful theories of all time. The theory is used to describe the composition of fundamental constituents

of matter and how they interact. According to the standard model, matter is made up of two types of elementary particles, fermions and bosons. Fermions are particles that have half-integer spin, and bosons have integer spin. Matter is made up of fermions, and matter interacts when fermions exchange bosons. There are two types of fermions, quarks and leptons, and each of these consist of 6 flavors. Quarks are divided into two categories, up-like (charge =  $2/3e$ ) and down-like (charge =  $-1/3e$ ). The up-like quarks are named up, charm and top and the down-like quarks are named down, strange, and bottom. Leptons are also divided into two categories, charged (charge =  $1e$ ) and neutral. Each charged lepton has neutral partner called a neutrino, and the lepton flavors are named electron, muon, tau. In addition, every particle in the standard model has an anti-particle partner. This anti-particle is exactly the same except the sign of all charges (electric and color) are the opposite value of the original charge.

There are 4 known forces in the universe, the gravitational force, the weak force, the electromagnetic force, and the strong force. With the exception of gravity, these forces are explained by the standard model as the exchange of bosons between fermions. According to the standard model, the weak force is governed by the  $Z^0$  and  $W^+$  and  $W^-$  bosons, the electromagnetic force is governed by the photon, and the strong force is governed by the gluon. In 2012, it was announced that a new particle was discovered at the LHC by the CMS [1] and ATLAS [2] experiments. This particle has properties consistent with that of a standard model Higgs boson. The Higgs boson is responsible for the mass of the massive standard model particles. A diagram of all the particles making up the standard model is shown in figure 1.1.

R-parity is a concept that is defined in equation 1.1 where  $B$  = baryon number,



**Figure 1.1:** A diagram is shown with the particles making up the Standard Model along with the attributes of each particle.

$L$  = lepton number, and  $s$  = spin. Baryon number is defined in equation 1.2, where  $N_q$  and  $N_{\bar{q}}$  are the number of quarks and anti-quarks respectively. Lepton number is defined in equation 1.3, where  $N_L$  and  $N_{\bar{L}}$  are the number of leptons and anti-leptons respectively. All particles in the standard model have  $R_P = 1$ .

$$P_R = (-1)^{3(B-L)+2s} \quad (1.1)$$

$$B = \frac{1}{3}(N_q - N_{\bar{q}}) \quad (1.2)$$

$$L = N_L - N_{\bar{L}} \quad (1.3)$$

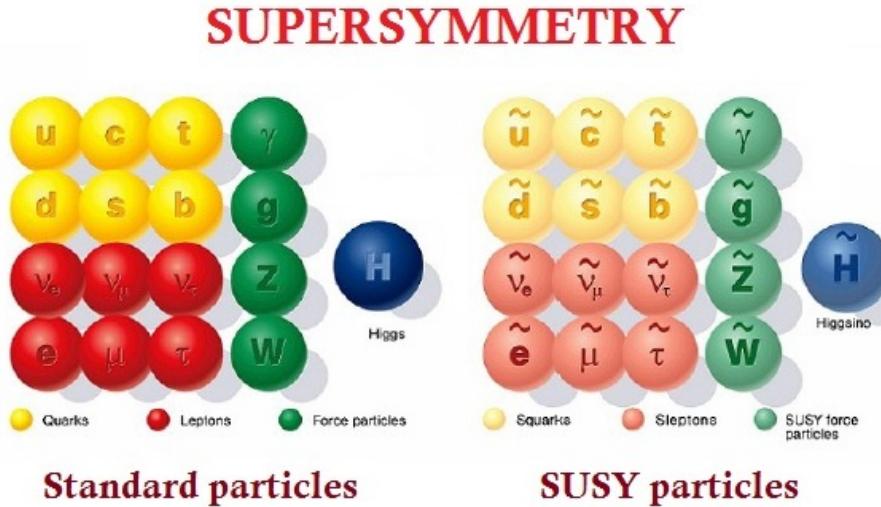
### 1.2.1 Problems with The Standard Model

The standard model is able to explain many things, but it is not complete. One example of where it fails is in explaining the existence of “dark matter”. Dark matter is a form of matter that does not interact electromagnetically and has been observed by astrophysicists [3]. Dark matter has been measured to be  $\sim 5$  times more abundant than visible matter in the universe. The hierarchy problem is another question that is not explained by the standard model, which essentially asks why is the weak force more than 1000 times stronger than gravity? One of the main goals of particle physics is to understand and provide explanations for these problems, and one way this can be done is within the context of supersymmetry (SUSY).

## 1.3 Supersymmetry (SUSY)

SUSY [4] is an idea which postulates that for every particle in the standard model, there exists a supersymmetric partner which has the property that the partners to bosons are fermionic, and the partners to fermions are bosonic. These particles are named sparticles, and for fermions, an s, short for scalar, is added to the beginning of the standard model particle name to represent the SUSY particle. A diagram of this is shown in figure 1.2. For example, a supersymmetric electron is known as a selectron, and a supersymmetric quark is known as a squark. For gauge bosons, the end of the name is changed to contain “ino”. For example, the supersymmetric partner to the W boson is known as the Wino and a supersymmetric gluon is known as a gluino. R-parity is an important concept within SUSY, and all SUSY particles have  $R_P = -1$ . This means that if R-parity is conserved, the lightest supersymmetric

particle (LSP) in a decay is stable. The LSP then makes a good candidate for dark matter. Using SUSY as a framework, one can create many models that can answer both the hierarchy problem as well as explain the existence of dark matter. This thesis focuses more on a search for new physics within SUSY that can be used as an explanation for dark matter.

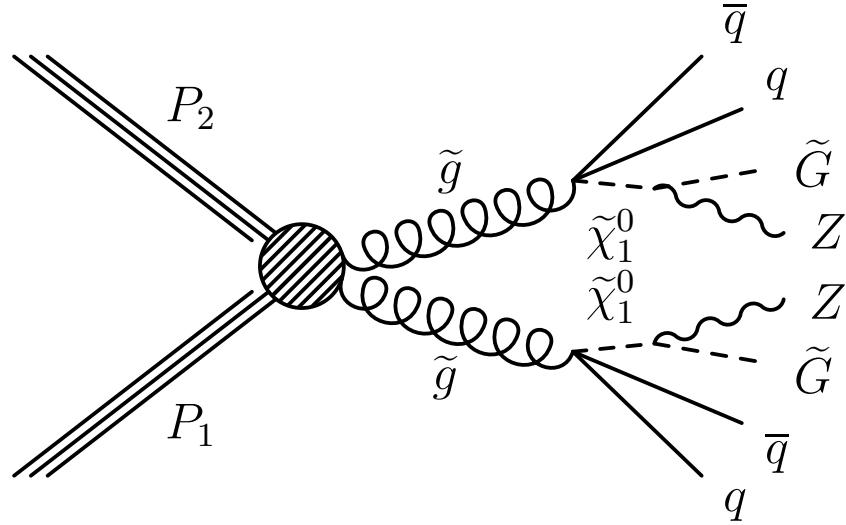


**Figure 1.2:** A diagram is shown with the particles making up the Standard Model on the left, and their SUSY counterparts on the right.

## 1.4 Searches for SUSY in final states with a Z boson

Many models can exist within the context of SUSY, leading to a nearly endless variety of final states. There are many parameters within SUSY that can be tuned to allow for the existence of different final states, for example the mass of the SUSY particles. A set of simplified models can be described to get a sense of what possible final states can be probed using the CMS detector at the LHC. These models are

named SMS models, where SMS stands for Simplified Model Space [5]. One of these models, shown in figure 1.3, is a model with a final state containing two Z bosons, four quarks and two invisible SUSY particles where the invisible particles in this model are the gravitino. This thesis focuses on a search done using data taken by the CMS detector at the LHC colliding protons with a center of mass energy of  $\sqrt{s} = 13\text{TeV}$ . The next chapter will explain in detail how particles are measured using the CMS detector.



**Figure 1.3:** A diagram showing a SUSY process where two protons collide and the result is pair-production of two gluinos, where each decays to a pair of quarks and a neutralino which subsequently decays to a Z boson and gravitino.

# Chapter 2

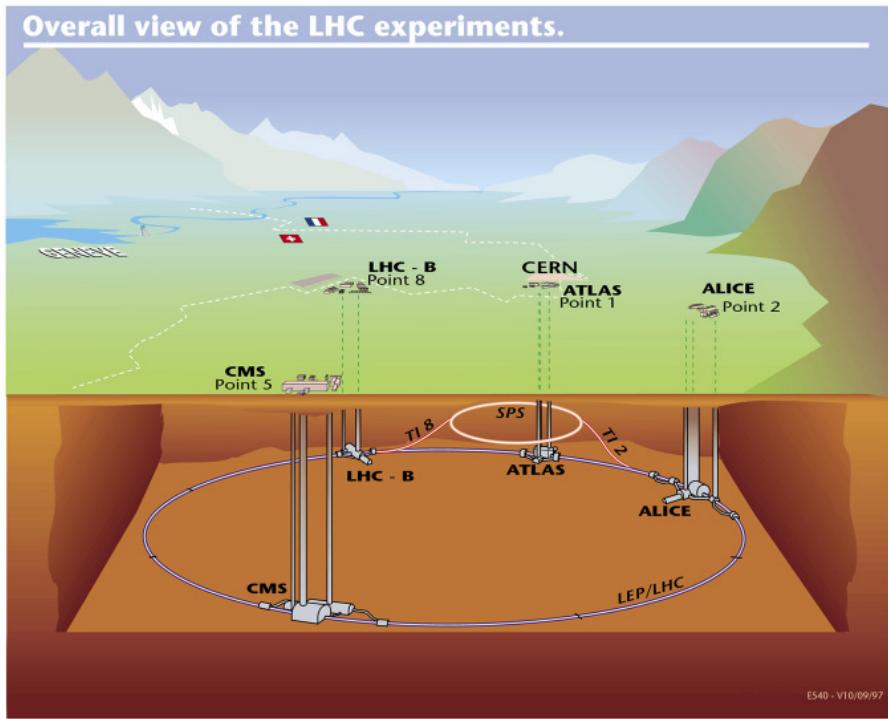
## The LHC and The CMS Experiment

### 2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the largest particle accelerator in the world [6]. It is designed to collide protons against other protons with a center of mass energy up to 14 TeV. The first collisions used for physics happened in March 2010, when the LHC was running at a center of mass energy of 7 TeV. The LHC was restarted the following year and operated at a higher energy of 8 TeV eventually delivering about four times the amount of data that was taken at 7 TeV. After the successful first run ended, the machine was shut down in order to upgrade many components. The next time the LHC was run was in May 2015, where protons were successfully collided at a center of mass energy of 13 TeV. The analysis shown in this thesis is based on the data taken during the 2015 running period at 13 TeV.

In order to achieve such large center of mass energies, the protons are accel-

erated in two separate circular beam pipes which are 27 km long and 100 m underground. These beam pipes are kept as close as possible to vacuum in order to minimize the interaction of protons with other particles. Since the pipe is circular, it's possible to accelerate two beams of protons in opposite directions and repeatedly collide them at various points around the ring. There are 4 points along the ring where protons are made to collide and the resulting collisions are recorded by various experiments as seen in figure 2.1. The Compact Muon Solenoid (CMS) experiment [7] is where the data was taken used for the results shown in this thesis.



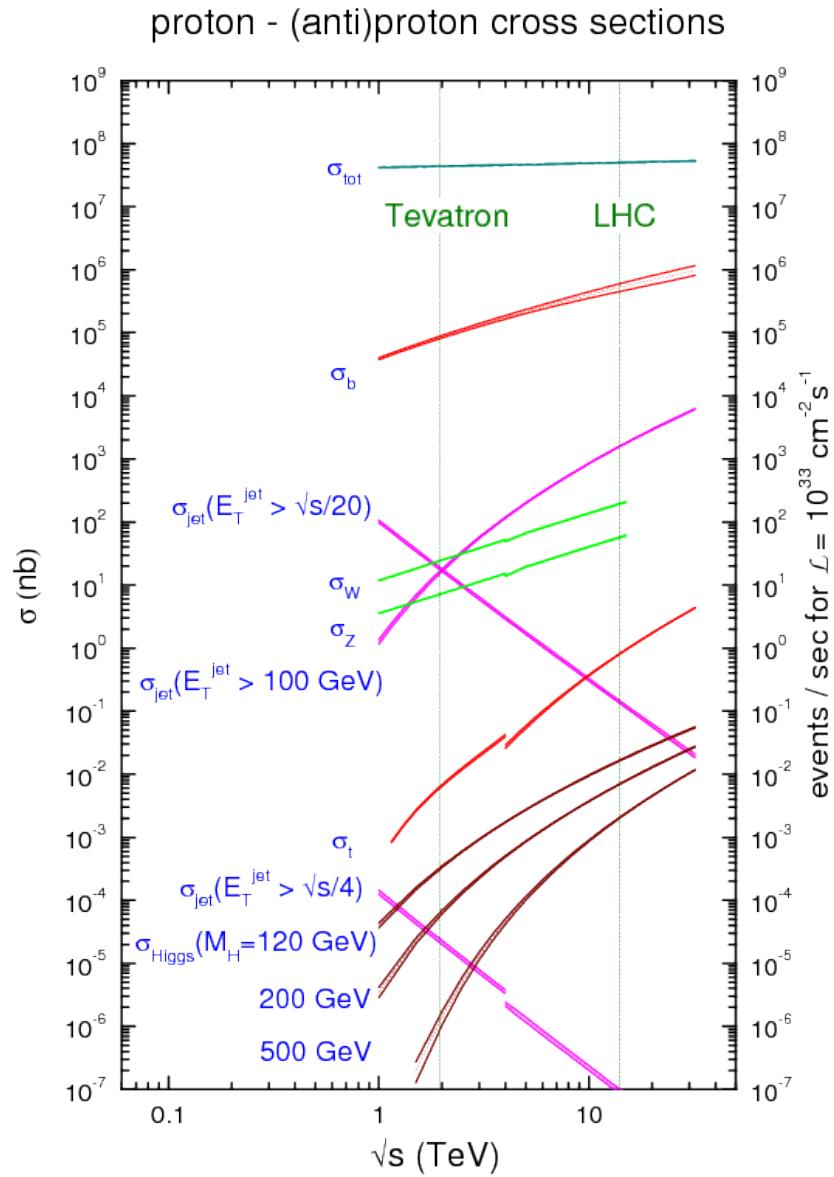
**Figure 2.1:** The LHC can be seen along with the 4 main experiments at different points along the beamline.

## 2.2 Physics Processes and Cross Sections

Particle physics processes are studied by measuring various quantities such as particle masses, lifetimes, and production rates both inclusively and differentially with respect to certain kinematic variables. When protons collide, it is not the protons themselves that interact but rather the constituents, or partons, that make up the proton. The formula describing this is given in equation 2.1. The number of events expected for a specific physics process is given by three quantities. The integrated luminosity ( $\mathcal{L}$ ) is dependent on the brightness of the proton beam. The cross section for the event to happen ( $\sigma$ ), which includes the branching fraction for a particle to decay to a specific final state, is a fundamental physics quantity. The units for cross section is “barns”, and a barn is defined as  $10^{-24} \text{ cm}^2$ . Lastly, the efficiency for the event to be reconstructed ( $\epsilon$ ) is completely determined by the detector.

$$N_{\text{events}} = \mathcal{L} * \sigma * \epsilon \quad (2.1)$$

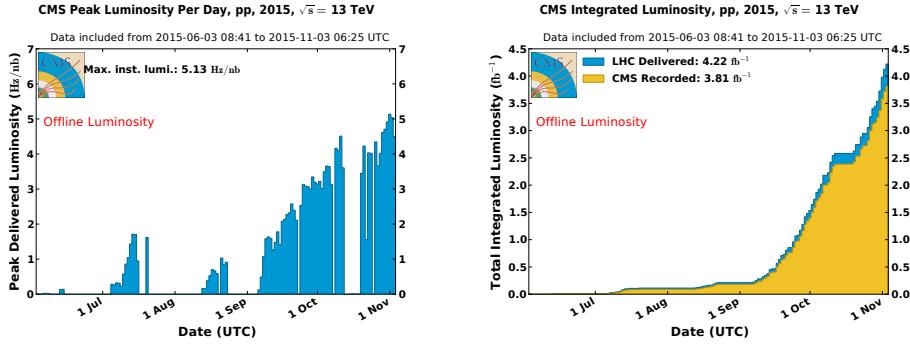
Protons can interact elastically (scattering) or inelastically, meaning they are destroyed in the process. The inelastic cross section for two protons to collide at a center of mass energy of 13 TeV is about 100 mb. This can be seen in figure 2.2 along with the cross section values for other interesting standard model processes. Other processes that are main backgrounds in the signal region of this analysis are Drell-Yan and  $t\bar{t}$  production which have a cross section at 13 TeV of roughly 4 nb, and 0.8 nb respectively. Taking into account all the branching fractions and detector efficiencies, these processes are reconstructed roughly once in a billion (trillion) times per inelastic proton collision for Z ( $t\bar{t}$ ) production.



**Figure 2.2:** The cross section values for various standard model processes is shown in this figure as a function of the center of mass energy of two protons colliding.

The peak instantaneous luminosity measured in a given day during the 2015 data taking period is shown on the left in figure 2.3, and the highest value is 5.13

$\frac{\text{Hz}}{\text{nb}}$ . This means that approximately  $10^8$  protons collide per second. The rate of collisions is much larger than the capacity for the detector to write out data, and as discussed previously, most of these events are just soft inelastic collisions of two protons. In order to maximize the efficiency of collecting data that is useful to any physics analysis, a triggering system is put in place which only writes out the event information if certain conditions are met. The maximum rate of all the concurrent triggers in CMS is about 1 kHz. The total luminosity collected in 2015 was  $3.81 \text{ fb}^{-1}$  which can be seen on the right in figure 2.3, but due to data quality issues the amount that was useable for physics analysis was  $2.26 \text{ fb}^{-1}$ .



**Figure 2.3:** The peak instantaneous luminosity per day during the 2015 data taking period is shown on the left and the total integrated luminosity is shown on the right in this figure.

## 2.3 Particle Interactions With Matter

The total amount of energy lost by a particle when interacting with matter is mainly due to two effects, interactions that lead to the particle decaying or “showering”, and energy loss by ionization. Two equations are used to describe energy loss due to showering, equations 2.2 and 2.3. Equation 2.2 represents the average energy

left over as a particle travels a distance  $x$  through a material with radiation length ( $X_0$ ). Equation 2.3 represents the average number of particles left when  $N_0$  particles traverse a distance  $x$  through a material with interaction length ( $\lambda_{int}$ ). The largest source of uncertainty when measuring a particles energy using the ECAL or HCAL is called the “stochastic term” and is mostly due to intrinsic statistical shower fluctuations and sampling fluctuations. The energy of a parton is distributed evenly among all of its decay products on average, so the initial particles energy is proportional to the number of decay products  $E_0 \propto N_0$ . The overall uncertainty due to the stochastic term becomes proportional to  $\sqrt{N_0}$  in the limit where  $N$  is large, therefore these two equations can be used to calculate the average energy lost due to electromagnetic interactions (radiation length) and strong interactions (interaction length).

$$E(x) = E_0 e^{\frac{-x}{X_0}} \quad (2.2)$$

$$N(x) = N_0 e^{\frac{-x}{\lambda_{int}}} \quad (2.3)$$

Typical values for radiation and interaction lengths are shown in table 2.1 for the different materials making up the tracker (silicon), ECAL (PbWO4) and HCAL (Copper). In table 2.2, the depth of each of these materials corresponding to  $|\eta| = 0$  is shown as a function of  $X_0$  and  $\lambda_{int}$ . The values in tables 2.2 and 2.1 are taken from a talk given at CERN in May 2016 [8]. Based on these values, an electron or photon loses at least  $\sim 40\%$  of its energy on average when traveling through the tracker before it is incident on the ECAL. This means either has a relatively high chance of interacting, where a photon might pair-produce and an electron might brem a photon.

These effects are taken into consideration when reconstructing photons and electrons and better-described in chapter 4.

Another effect that can be seen is in the HCAL, there are 16 layers of absorber material. This means that each layer has more than 4.5 radiation lengths. When a hadronic shower evolves, approximately 30% of the shower is in the form of EM energy.  $4.5X_0$  is essentially deep enough to completely contain the EM component of the hadronic shower before it can be measured. This leads to an increased uncertainty of the initial parton's energy.

**Table 2.1:** Radiation ( $X_0$ ) and interaction lengths ( $\lambda_{int}$ ) for various materials that make up different sub-detectors in CMS.

Material	$X_0$ [cm]	$\lambda_{int}$ [cm]
Silicon	9.36	45.5
Lead Tungstate ( $PbWO_4$ )	0.88	19.5
Copper	1.43	15.1

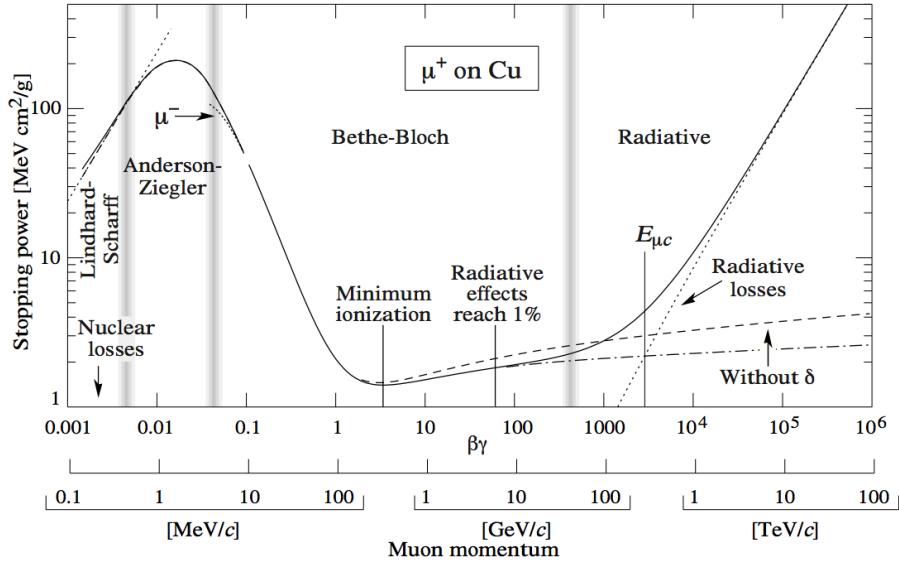
**Table 2.2:** Material budget of tracker and calorimeters in CMS at  $|\eta| = 0$  as a function of radiation ( $X_0$ ) and interaction ( $\lambda_{int}$ ) lengths.

Material	Radiation lengths [ $X_0$ ]	Interaction lengths [ $\lambda_{int}$ ]
Silicon	0.5	0.1
Lead Tungstate ( $PbWO_4$ )	25	1.1
Copper	73.9	7

Energy loss by ionization is mostly due to a particle having elastic collisions with electrons and can be characterized by the Bethe-Bloch formula shown in equation 2.4. In this equation,  $K$  is a constant equalling  $0.307 \text{ MeV g}^{-1} cm^2$ ,  $T_{\max}$  is the max energy transfer in a single collision,  $z$  is the charge of the incident particle,  $Z$  is the charge number of the medium,  $A$  is the atomic mass of the medium,  $m_e$  is the mass of an electron, and  $I$  is the mean excitation energy of the medium. Muons have

a low interaction cross section with the materials used to build CMS in the energy range that they are produced when coming from Z bosons produced at the LHC. They are referred to as minimum ionizing particles, and the stopping power of copper on muons can be seen in figure 2.4. Therefore, muons do not deposit very much energy in any of the calorimeters. Typical momenta of muons in this analysis are between 20-200 GeV. When muons have momenta greater than 200 GeV, radiative losses start to become the primary reason for energy loss.

$$-\frac{dE}{dx} = K z^2 \frac{Z}{A} \frac{1}{\beta^2} \left[ \frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 T_{max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right] \quad (2.4)$$



**Figure 2.4:** Stopping power of copper as a function of muon momentum. Typical momenta of muons in this analysis are between 20-200 GeV.

## 2.4 The Compact Muon Solenoid

The CMS detector consists of a triggering system and many sub detectors, where measurements taken using these sub detectors are used as input to the triggering system. When protons are collided at the interaction point within CMS, approximately N collisions happen per second. The rate of collisions is so high that it is impossible to record all of the data all of the time. The way this is controlled is by the use of triggers, the level 1 (L1) trigger and the high level triggers (HLT). The triggers are designed such that the full event information is stored when all the requirements of a trigger are met. For some physics processes, trigger rates are much higher than can be afforded by the allowed budget of the CMS experiment. These processes can still be studied by prescaling the triggers meaning the full information is saved once per N events. As the prescale N is increased, the rate of the trigger is reduced.

The CMS sub detector components are layered in such a way that one can take advantage of different particle interactions with different materials. The innermost layer is the tracker (described in section 2.4.1) which is used to measure the charge and momentum of charged particles. When a charged particle passes through the silicon tracker's layers, a small amount of energy is deposited. The deposit is large enough that a hit will be registered by the electronics, but small enough that the trajectory of the particle will be minimally affected.

The next layer is the Electromagnetic Calorimeter (described in section 2.4.2) which is used to contain and measure the energy of electrons and photons. The Electromagnetic Calorimeter (ECAL) is made of lead tungstate ( $\text{PbWO}_4$ ) scintilating crystals. Electrons and photons that enter the ECAL lose energy due to showering,

and the depth of the ECAL is sufficient enough to fully contain these particles.

Outside of the ECAL is the Hadronic Calorimeter (described in section 2.4.3) which is used to contain and measure the energy of hadronic particles. The Hadronic Calorimeter (HCAL) is designed in such a way that hadronic particles incident on the HCAL lose energy to showering. The goal is to fully contain these particles, although sometimes particles have sufficient energy that they are able to escape before losing sufficient energy to showering. This phenomenon is known as punch-through.

Beyond the HCAL is the solenoid which produces a magnetic field of 3.8 T at the center of the detector. This magnetic field is used to bend the tracks of charged particles in order to measure the momentum following the equation 2.5, where  $q$  is the particle charge,  $r$  is the radius of curvature, and  $B$  is the magnetic field magnitude.

$$p = qrB \quad (2.5)$$

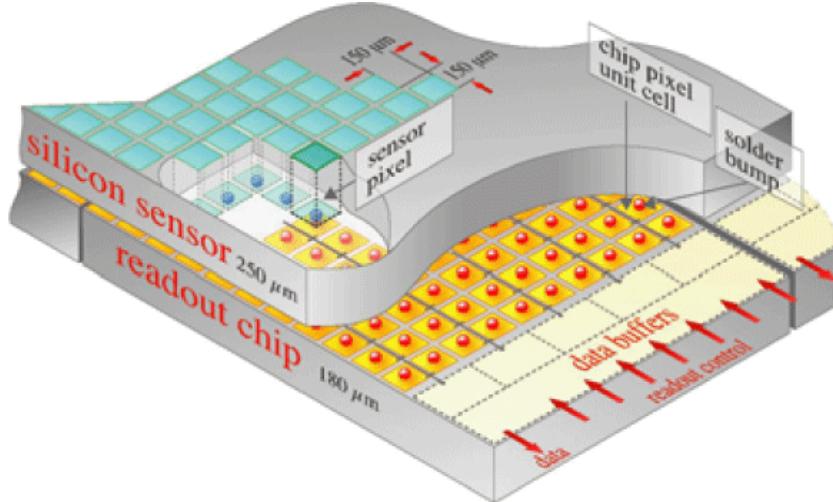
The only particles that should pass beyond the HCAL are muons and neutrinos. Neutrinos are measured indirectly by assuming that the momentum in the direction transverse to the beam is zero at the beginning of each collision, and then calculating the sum of the total transverse momentum of all particles directly measured in the detector. Since Neutrinos pass through the detector with no interactions, they contribute to missing transverse momentum, which will be referred to as ( $E_T^{\text{miss}}$ ) or MET.

Lastly, the muon system (described in section 2.4.4) is used to measure the charge and momentum of muons. At the energies that muons are most commonly produced in collisions at the LHC, muons are minimum ionizing particles (MIPs) meaning they do not lose very much energy as they pass through the detector com-

ponents.

### 2.4.1 Silicon Tracker

The tracker, shown in figure 2.6, has two main components, the pixels and the strips. The pixel component has three cylindrical layers very close to the beam pipe, which are positioned at 4, 7 and 11 cm from the center. The pixel detector contains over 65 million sensors (pixels) measuring 150 by 150  $\mu\text{m}$  which are connected directly to a chip used to read out the data. A diagram showing the pixels can be seen in figure 2.5. When a charged particle travels through the three pixel layers, the hits are recorded then used together with the measurement of the particle's trajectory through the strip component of the tracker.



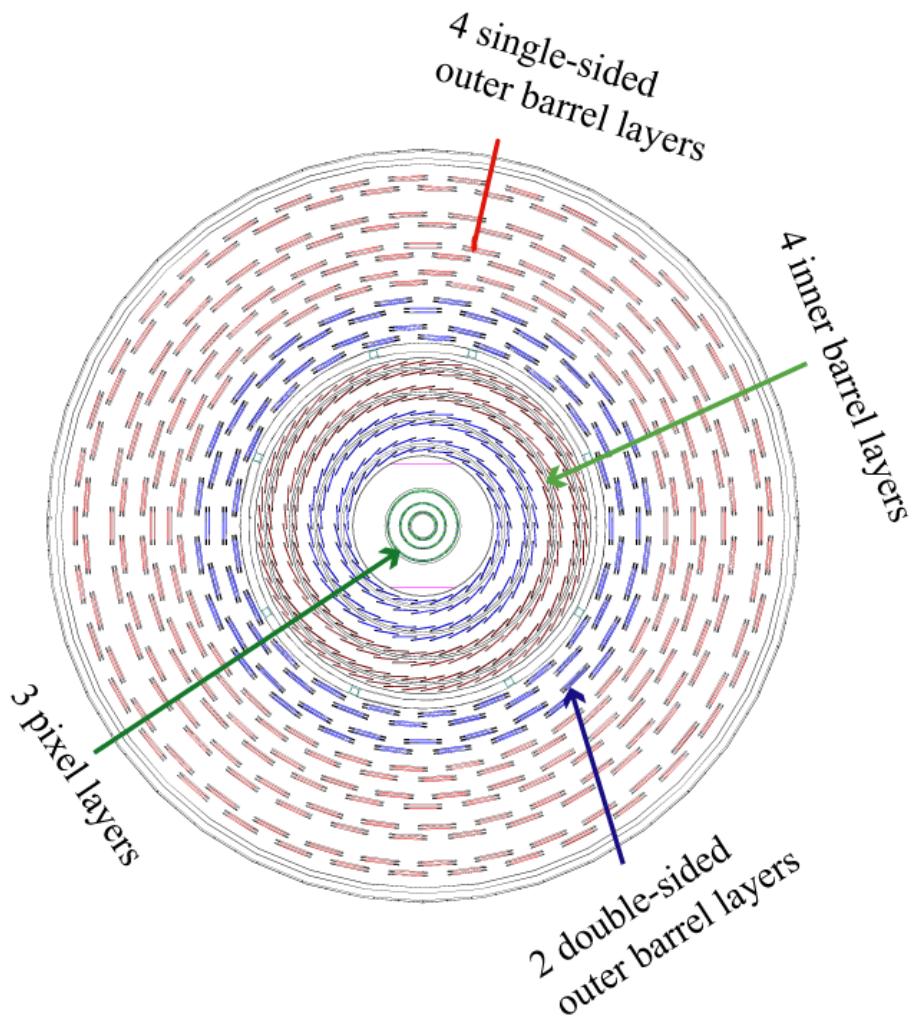
**Figure 2.5:** A diagram showing the pixel component of the tracker subdetector.

Immediately outside of the pixel layer is the strip layer. There are 10 layers of strips through which a charged particle passing through will leave deposits of energy that are read out. The very large magnetic field in this region from the solenoid is

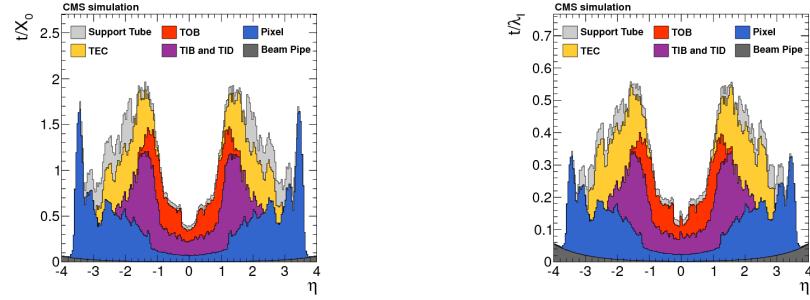
in the same direction as the beam, and this causes particle trajectories to curve in a direction tangent to the beam direction. The radius of this curvature is proportional to the particle's momentum and inversely proportional to the particle's charge according to equation 2.5. Using the hit information from the pixels and strips together, the momentum of all the charged particles that pass through the tracker is reconstructed and used as input to the particle flow algorithm 2.5.1 in order to identify particles produced in the collision. The total material budget of the tracker as a function of radiation and interaction length is shown in figure 2.7. The total depth is seen to be increasing with  $|\eta|$  up to where the tracker barrel meets the tracker endcap then decreasing after this point.

### 2.4.2 Electromagnetic Calorimeter

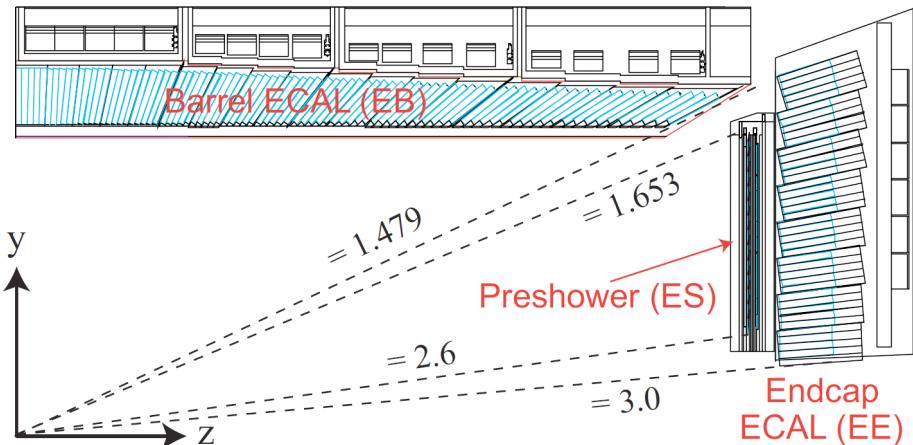
The ECAL is split into two regions, the ECAL barrel (EB), and ECAL endcap (EE). The EB occupies a region in the detector within  $|\eta| < 1.479$ , and the EE occupies the region with  $1.479 < |\eta| < 3.0$ . The geometry can be seen in figure 2.8.



**Figure 2.6:** A cross-sectional in the x-y plane of the CMS tracking system is shown here. The tracker provides excellent coverage for reconstruction of tracks from charged particles for  $|\eta| < 2.5$ .

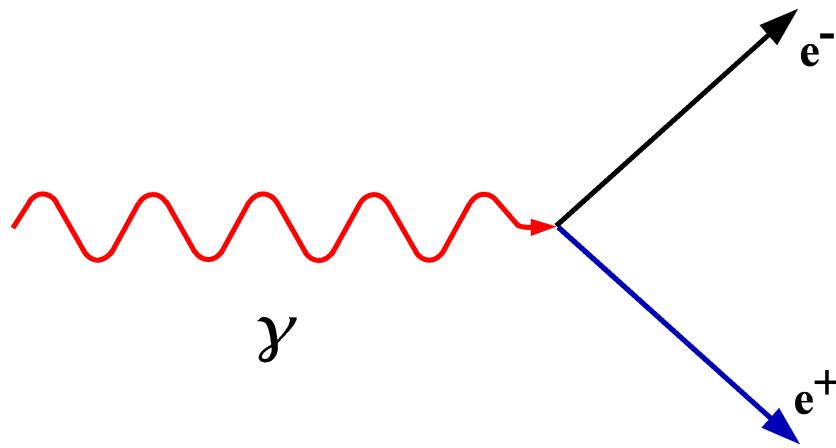


**Figure 2.7:** A diagram showing material budget of the tracker as a function of  $\eta$  on the x-axis, and radiation lengths (left) and interaction lengths (right).



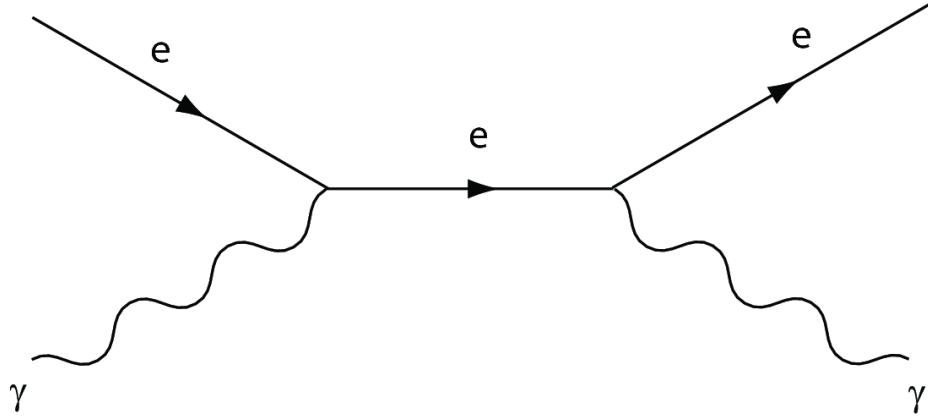
**Figure 2.8:** A cross-sectional view of the CMS ECAL is shown in this figure, with values of  $\eta$  shown that determine the coverage of each subdetector.

As written previously, the ECAL is made of lead tungstate crystals. These crystals act as absorbers as well as scintillators. What this means is that when electrons or photons are incident on the crystals, they lose energy due to showering. The showering process can be simply described as a combination of electron-positron pair production, and photon emission. Leading order diagrams of these processes are shown in figures 2.9 and 2.10.



**Figure 2.9:** A feynman diagram is shown depicting electron-positron pair production from a photon initial state.

As the initial particle passes through the crystal, it loses energy due to these processes, and this energy is converted to light by the scintilating properties of the crystal. This light is read out using a avalanche photo-diode (APDs) for each crystal in the EB, and a vacuum phototriode for each crystal in the EE. The amount of energy absorbed by the crystals can be parameterized by a quantity known as the radiation length ( $X_0$ ). The expected amount of energy left ( $E(x)$ ) after a particle with initial energy ( $E_0$ ) travels a distance  $x$  through an absorbing material is shown in equation 2.2. The CMS ECAL has a depth of  $\sim 25X_0$  at small ( $0-1.3$ ) and large ( $1.6-3.0$ ) values of  $\eta$ . In between these regions, the depth is larger due to the barrel-



**Figure 2.10:** A feynman diagram is shown of a process where an electron scatters off of an incoming photon, and then radiates a photon through bremsstrahlung.

like geometry of the ECAL. Additionally, the ECAL material has a Molière radius ( $R_M \simeq 2.19$  cm) where  $R_M$  is defined as the radius of a cylinder containing on average 90% of the shower's energy deposition. Because of the depth and intrinsic resolution, electron and photon energies are measured very well in the ECAL. There is one exception where the depth of the ECAL is small, and this is where the barrel and endcap meet. This leads to poor reconstruction of electrons and photons in this region.

### 2.4.3 Hadronic Calorimeter

The Hadronic Calorimeter (HCAL) is used to measure energies of hadronic particles, such as pions, gluons, protons, kaons and neutrons [9]. It is a sampling calorimeter meaning it has many layers of absorber material (brass) and scintillator material (plastic). As hadronic particles pass through the HCAL, they interact with the absorber layers by showering, meaning they decay to multiple lower energy particles. As these particles shower and the constituents pass through the scintillator,

photons are emitted by the scintilating material and then measured by Hybrid-Photo Diodes (HPDs).

Approximately a third of the particles produced in a hadronic shower are  $\pi_0$ s, which immediately decay to two photons most of the time. These particles are mostly stopped in the absorbing layer making it hard to accurately determine the energy lost due to this effect. Some detectors however are built to be able to accurately measure the energy lost to electromagnetic effects in the shower, and these are known as “compensating” calorimeters. When a particle is stopped by the HCAL, the energy deposited in all the layers is integrated over a full segment of the HCAL, and used to determine an incident particles energy. The HCAL at CMS is not able to compensate for these effects, so it must be calibrated offline using particles with a well-known initial energy.

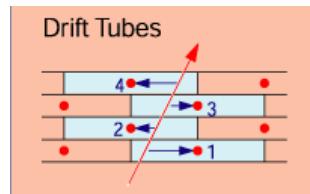
The HCAL geometry is such that it is fully contained within the solenoid. It is divided into towers, each of which lies behind an integer number of ECAL crystals. This helps in calibrating the energies of incident particles during reconstruction. The number of interaction lengths ( $\lambda_{int}$ ) in the HCAL starts at about  $\lambda_{int} = 7$  at  $\eta = 0$ , and increases with eta to the end of the barrel up to  $\lambda_{int} = 11$ . There is then a slight decrease in the endcap of  $\lambda_{int} = 10$ . The interaction length is used to characterize the longitudinal and transverse profile of hadronic showers. The expected number of secondary particles produced ( $N(x)$ ) by a number of particles traveling through the HCAL ( $N_0$ ) is written in equation 2.3 as a function of  $\lambda_{int}$  and the distance traveled ( $x$ ).

## 2.4.4 Muon System

The muon system is made up of 1400 chambers divided into 3 categories: 610 resistive plate chambers (RPCs), 250 drift tubes (DTs), and 540 cathode strip chambers (CSCs). It is the only detector subsystem which lies outside of the solenoid, but it is embedded within the return yoke of the solenoid. Therefore, the magnetic field in the muon system is non-uniform leading to a characteristic s-curve flight path in the R- $\phi$  plane for muons originating from the center of the detector. The B-field decreases in the return yoke as the distance from the solenoid increases radially outward leading to the bending radius of the muon decreasing as well. The information from these systems is used for triggering on events with muons as well as help to identify muons and measure their properties. Each system is located in different regions in the detector, with some systems overlapping others. The DTs cover a region of  $|\eta| < 1.2$ , the CSCs cover a region of  $0.9 < |\eta| < 2.4$ , and the RPCs cover a region of  $|\eta| < 2.1$ .

Each DT is 4 cm wide with a length of wire running all the way down the middle. A diagram of a DT is shown in figure 2.11. The DTs are individually filled with a mixture of Ar and CO<sub>2</sub> gas, and when muons pass through the DTs, this gas is ionized. A voltage difference is maintained between the wire and the outside of the DT such that when the gas is ionized, the ionized particles drift towards the charged components causing a voltage spike. The DTs are then arranged in such a way that a muon passing through the region where these are located will leave a hit in multiple DTs. The muon's trajectory can then be reconstructed using this information.

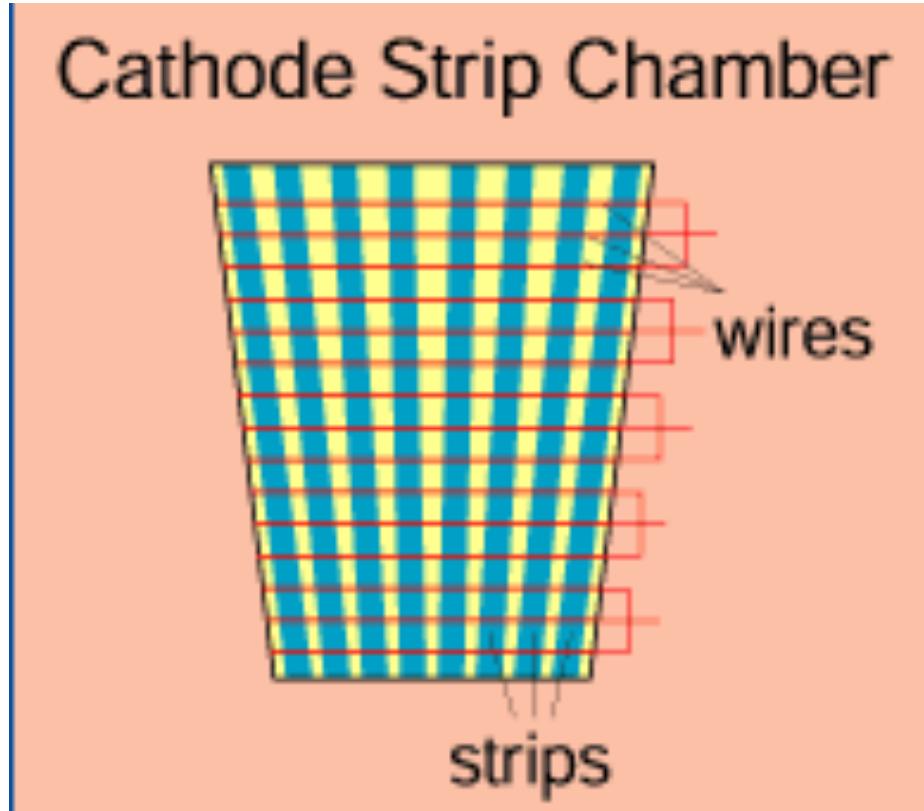
There are 540 CSCs shown in figure 2.12. The CSCs are made up of negatively charged wires (anodes) and positively charged strips (cathodes) within a gas volume.



**Figure 2.11:** A cartoon depiction of a set of drift tubes is shown in this figure where a charged particle passing through the system ionizes the gas, and the ions are collected on the wires to detect a signal.

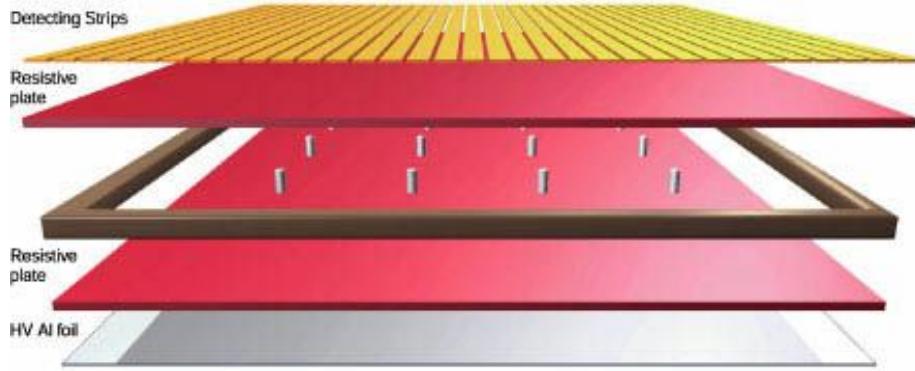
The gas molecules within this volume are ionized when a muon passes through the volume. This leads to a measurable hit on both the anodes and cathodes, and they are perpendicular to each other so that the muon's trajectory can be reconstructed accurately. There are six layers in each CSC module, which helps to accurately

identify muons.



**Figure 2.12:** A cartoon depiction of a cathode strip chamber is shown in this figure.

There are 610 resistive plate chambers (RPCs), an example of which is shown in figure 2.13. These are designed to have overlapping trigger coverage with the other components of the muon system for  $\eta < 2.1$ . Each RPC is made of two very high resistance plates, separated by a volume of gas. They are charged kept at a high voltage and when a muon passes through the gas, electrons are knocked off of the molecules. Metal strips are used to read out the position of the hit, and the muons momentum is measured with this information. The readout for the RPCs is very fast (1 ns) which makes it ideal to use for triggering events with good muons.



**Figure 2.13:** A cartoon depiction of a resistive plate chamber is shown in this figure.

## 2.5 Physics objects

The data taken by the CMS detector has to be processed in such a way that all the physics processes that happen in a single collision can be reconstructed. Each reconstructed collision is called an event. In each event, there are multiple physics objects that are reconstructed. The main objects that are pertinent to this analysis are electrons, muons, jets, and photons. Each of these objects has a unique signature in the detector, but it is still possible for this signature to be faked. The reconstruction of these objects is discussed in detail in the following sections.

### 2.5.1 Particle Flow

In order to classify objects, an algorithm named particle flow (PF) is used [10]. The main goal when using the PF algorithm is to account for all energy deposits measured in the detector in a consistent fashion. The way this is done is by clustering energy deposits measured across all subdetectors into separate PF candidates, and classifying these candidates. Once an object is classified, the clustering is done again with all the energy associated with the classified object removed. This process is

repeated until all measured energy deposits are accounted for. For example, a charged hadron would be reconstructed by first identifying a track in the tracker. Then a check is done to see if this track is consistent with an identified muon in the muon chambers. If not, all the energy that is associated with this track in the ECAL and HCAL is removed from the event, and the track measurement is used to describe the particle. This is done because the precision in the tracker is much better than in the calorimeters. This is done iteratively for all particles in the event until there are no tracks left, then all that remain are the neutral energy deposits. The use of particle flow leads to greatly improved measurements of jets since more of the jet energy is measured using the tracker. This also leads to a higher precision measurement of missing transverse energy for the same reasons.

The objects that are identified by this algorithm are listed below. In order to further reduce fakes from contaminating our signal region, we make additional cuts to help classify objects. These cuts are described in detail in the following section for each object we are interested in.

- muon
- photon
- charged hadron
- neutral hadron
- electron

### 2.5.2 Vertex Determination and Pileup

The LHC collides protons in large bunches to maximize the probability of a hard collision. Multiple pairs of protons interact at each bunch crossing, and the number of interactions is stochastic but depends on the beam density. Each proton-proton collision can produce particles that get reconstructed, and when these particles are reconstructed they are associated to a specific proton-proton collision using track information. It is important to identify which proton-pair interaction is responsible for triggering the event being studied. Every particle track is associated with a unique vertex, and the primary vertex is chosen by finding the vertex with the largest sum of  $(p_T)^2$  of tracks associated with that vertex. Neutral candidates do not have associated tracks, so they are defined to always be from the primary vertex. The energy of objects affected by this reassociation is recalibrated, and this calibration is described in section 2.5.5.

Pileup is defined as energy in the detector which does not come from the primary collision or primary interaction. There are two primary sources of pileup, in-time pileup (ITPU), and out-of-time pileup (OOTPU). ITPU occurs when protons other than those in the primary interaction interact in the same bunch crossing, and the particles produced are reconstructed. OOTPU occurs mainly due to the fact that protons are collided with a frequency of 25ns, and the timing resolution of some detector components is not fast enough to resolve which bunch crossing certain reconstructed particles came from. The number of expected pileup events can be determined using equation 2.6 where  $\mathcal{L}_{inst.}$  is the instantaneous luminosity,  $\sigma(p-p)_{inelastic}$  is the total inelastic scattering cross section for protons, and  $f_{collision}$  is the collision frequency. For 2015, this value is 12.75.

$$N_{\text{pileup}} = \mathcal{L}_{\text{inst.}} * \sigma(p - p)_{\text{inelastic}} * f_{\text{collision}} \quad (2.6)$$

The majority of energy from pileup comes from proton-proton interactions other than the primary interaction. These “soft” interactions produce predominantly hadronic final states, and the majority of these interactions are due to gluons scattering off of quarks or other gluons. These final states inherently do not have  $E_T^{\text{miss}}$  due to particles escaping detection due to them being hadronic, and this leads to a balance in the overall energy deposited in the detector. There are still contributions to  $E_T^{\text{miss}}$  from pileup due to the poor resolution of jets in CMS. A calibration is applied to account for pileup according to the method described in section 3.2.

### 2.5.3 Isolation

Isolation is a concept that is very important when identifying leptons, and photons. Isolation can be simply described as the total energy in the detector near an object of interest. Leptons and photons from the primary interaction are not produced together with colored objects, and are deemed prompt. Non-prompt leptons, such as those from a b-quark decaying semi-leptonically, and non-prompt photons such as those from  $\pi^0$  decay, are always produced in conjunction with a hadronic shower. These non-prompt objects are then very likely to be non-isolated due to the hadronic energy nearby. The distance between two objects ( $\Delta R$ ) is defined by equation 2.7, where  $\phi$  is the angle measured in plane transverse to the beam direction, and  $\eta$  is the psuedorapidity defined by equation 2.8. In the equation for  $\eta$ , the variable  $\theta$  is defined as the angle measured from the beam axis to the axis of the transverse plane to the beam axis. In particle physics,  $\eta$  is preferred over  $\theta$  when describing a particle’s

momentum along the beam axis because changes in  $\eta$  are lorentz invariant, whereas the same is not true for changes in  $\theta$ .

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \quad (2.7)$$

$$\eta = -\ln(\tan(\frac{\theta}{2})) \quad (2.8)$$

The choice of what to include when calculating isolation is defined differently for each object and will be described in the following sections.

#### 2.5.4 leptons and photons

Z bosons decay leptonically to  $e\bar{e}$ ,  $\mu\bar{\mu}$ , and  $\tau\bar{\tau}$  pairs  $3.363 \pm 0.004\%$ ,  $3.366 \pm 0.007\%$  and  $3.370 \pm 0.008\%$  of the time respectively. The rest of the time they decay to neutrinos, or to quark-antiquark pairs. This analysis focuses on decays to  $e\bar{e}$  and  $\mu\bar{\mu}$ . This is due to the fact that  $\tau$ s have a large mass and short lifetime which leads to them decaying. This is further complicated by the fact that  $\tau$ s decay to an electron ( $\mu$ ) and two neutrinos  $17.85 \pm 0.05\%$  ( $17.36 \pm 0.05\%$ ) of the time, and the rest of the time the  $\tau$  decays to hadronic final state particles. This leads to worse measurements in leptonic final states due to the missing energy from not detecting the neutrinos, and makes it hard to distinguish  $\tau$ s from jets in hadronic channels. In order to keep the analysis simple, this search is done in final states with  $Z \rightarrow ee$  and  $Z \rightarrow \mu\bar{\mu}$  only, and when referencing leptons, only electrons and muons are considered. It is also important to be able to identify photons in order to measure one of the main backgrounds.

The depth of the ECAL is  $\sim 25$  radiation lengths which means electrons and photons lose almost all of their energy in the ECAL. Another thing that helps to distinguish electrons from photons is that electrons will deposit energy in the silicon tracker whereas photons will not. Therefore, a track can be matched to energy deposits in the ECAL to help identify electrons. Similarly, a track found to correspond to a hit in the ECAL can be used to help determine that the energy deposited in the ECAL was not from a photon.

A different approach is taken to identify and measure muons in the CMS detector. Since muons are charged, they leave hits in the silicon detector which can be reconstructed as a track. The ECAL and HCAL are designed to contain all measurable energy, but due to the minimum ionizing nature of muons, muons are able to pass through these subdetectors without depositing much energy. The muon subdetector is the outermost layer of the CMS detector, and it is designed to identify muons. Muon tracks can be measured independently in the muon subdetector, and the measurement is matched to a set of hits in the tracker to determine the final muon energy.

The main thing all of these objects have in common is they are produced in interactions that do not have a hadronic component. Quantifying this energy is done by calculating the particles isolation, as described in section 2.5.3, and is used to discriminate against unwanted physics processes.

### 2.5.5 jets

A jet is an object used to represent a single parton, and it can be simply described as a substantial amount of hadronic activity concentrated in a single region

of the detector. The region is defined using the “anti- $k_T$ ” (ak) algorithm [11] with a jet radius of 0.4. For jets in this analysis, the charged candidates not associated with the primary vertex are removed before clustering. This removes some unwanted energy from pileup, and the jets are then calibrated in order to more accurately reflect the energy of the particles they represent.

The jet energy corrections are done in multiple levels, labeled as L1,L2 and L3. A detailed procedure of how the corrections are applied can be found here [12]. Jets are corrected such that the calibrated energy is as close as possible to the parton that produced the jet. The corrections are derived using MC, and each level corrects for different effects in the following way: L1 corrects for pileup, L2 corrects for different responses in the detector as a function of  $\eta$ , and L3 corrects the overall energy scale. The L1 correction is done by assuming a flat overall energy density from pileup in the detector which is calculated per event, and then subtracting this energy from the jet using the jet area to determine the overall energy contributed from pileup. The L2 correction is done by selecting di-jet events where one of the jets is central ( $\eta < 0.8$ ) and the second jet is non-central ( $\eta > 0.8$ ). Scale factors are then derived to correct the energy measurement for the non-central jet in separate regions of  $\eta$ . The L3 correction is done by deriving scale factors as a function of jet  $p_T$  which correct the jet energy to match the true energy of the parton that the jet came from.

In addition to the L1, L2 and L3 corrections, residual corrections are derived in data control regions to correct for differences in the detector response to jets between data and MC. This is done in a region with at least one photon or a Z boson which decays to two leptons, where the boson recoils off of a jet. The photon and leptons are measured with better resolution than the jets, so scale factors are derived as a

function of  $p_T$  to correct the detector response of jets. The full description of cuts used to define jets for the analysis is defined in sub-section 4.3.10.

# Chapter 3

## Missing Transverse Momentum

In the final state targeted by this analysis, a gravitino is proposed as the LSP. This particle is weakly interacting and makes a good dark matter candidate which lead to signal events having large values of  $E_T^{\text{miss}}$  due to this particle escaping detection. It is important that  $E_T^{\text{miss}}$  from standard model processes is predicted accurately.  $E_T^{\text{miss}}$  can be calculated using the full collection of PF candidates described in section 2.5.1. In the following sections,  $E_T^{\text{miss}}$  calculations and corrections are described in detail.

### 3.1 raw $E_T^{\text{miss}}$

Calculating  $E_T^{\text{miss}}$  using PF candidates is done by summing the transverse momentum vectors of all the PF candidates, described in equation 3.1.

$$\overrightarrow{E_T^{\text{miss}}} = - \sum_i \overrightarrow{p}_T^i \quad (3.1)$$

When  $E_T^{\text{miss}}$  is calculated without any calibration applied to the pf candidates, it is known as raw MET.

### 3.1.1 Sources of $E_T^{\text{miss}}$

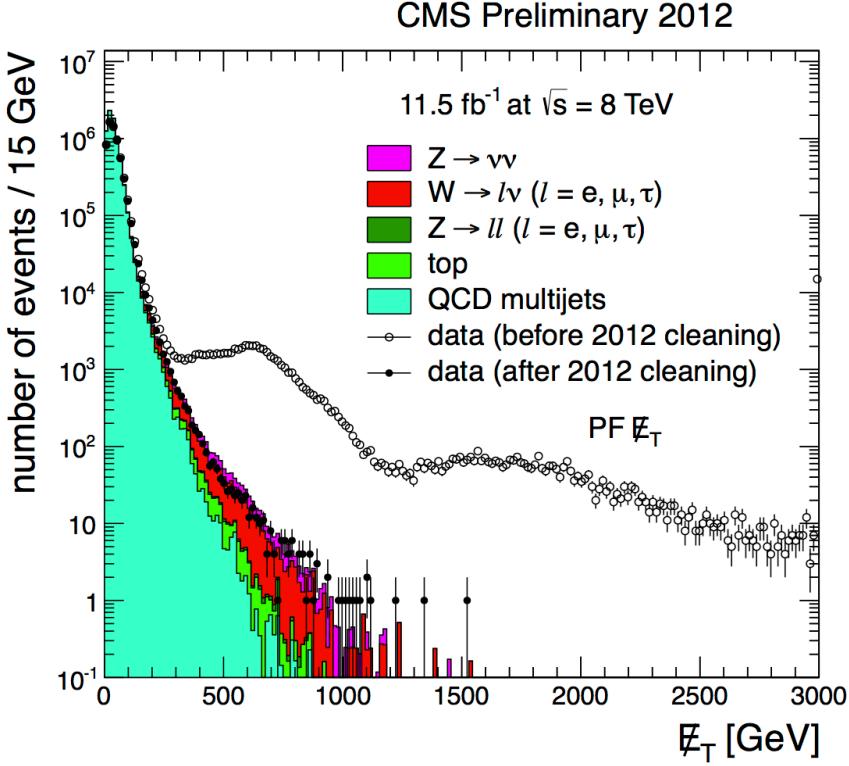
$E_T^{\text{miss}}$  comes from three main sources, noise in the detector from various sources, resolution effects due to mismeasurement of physics objects and visible particles from the hard interaction being too far forward in  $\eta$  to be within the detector geometry (referred to as fake  $E_T^{\text{miss}}$ ) and particles from real physics processes not interacting with the detector thereby escaping detection.

$E_T^{\text{miss}}$  due to noise in the detector can lead to huge tails in the  $E_T^{\text{miss}}$  distribution. The size of this effect is shown in figure 3.1 for the data taken in 2012. A set of filters is applied to mitigate these effects, and the details of these filters is described in section 4.3.2.

Particles that do not interact with the detector that lead to real  $E_T^{\text{miss}}$  can be standard model neutrinos, or particles from BSM theories such as the  $\tilde{G}$ . Since this analysis is searching in regions with large  $E_T^{\text{miss}}$ , it is very important to be able to identify and measure backgrounds with fake  $E_T^{\text{miss}}$  and real  $E_T^{\text{miss}}$ .

## 3.2 Type-1 $E_T^{\text{miss}}$

In order to reconstruct  $E_T^{\text{miss}}$  with high precision, corrections that are made to physics objects need to be propagated to the  $E_T^{\text{miss}}$  variable. The corrections with the largest impact on the  $E_T^{\text{miss}}$  resolution are the jet corrections, since jets have the worst resolution. Energy from pileup in each event is deposited randomly in the detector, and is expected to be balanced. It can happen that some pileup energy deposited overlaps with energy from a jet coming from the hard interaction. A correction is applied to the jet to account for this effect, however if the pileup energy removed from



**Figure 3.1:**  $E_T^{\text{miss}}$  distributions for events passing the dijet selection without 2012 cleaning algorithms applied (open markers), with 2012 cleaning algorithms applied including the one based on jet identification requirements (filled markers), and events from MC (filled histograms).

each jet's cone were to be propagated to the  $E_T^{\text{miss}}$ , it would create an imbalance. This is taken care of by adding the pileup energy back in, and the overall correction to  $E_T^{\text{miss}}$  is called the Type 1 correction.

Type 1  $E_T^{\text{miss}}$  refers to a version of  $E_T^{\text{miss}}$  where the corrections that are applied to jets are propagated to the  $E_T^{\text{miss}}$  calculation as described in section 2.5.5. The correction factors applied to the jets are derived for jets down to 10 GeV. The way these corrections are propagated to the  $E_T^{\text{miss}}$  is not as straightforward as just adding the corrections vectorially with the  $E_T^{\text{miss}}$ .

The jets we use in this analysis are made using all the pf candidates, except the charged pf candidates that are not associated with the primary vertex. Leptons and photons are included in the collection of pf candidates used when clustering jets. The jet corrections are meant to correct jets that come from hadronic objects, such as quarks and gluons. If the jet corrections are applied to lepton and photon objects, the jet corrections are not meant to be applied to lepton and photon objects. When propagating the jet corrections to the  $E_T^{\text{miss}}$ , we exclude corrections on jets with an electromagnetic fraction larger than 90%, and remove the pf muon candidates from the jet when deriving the corrections. This effectively removes photons and leptons from the objects that get corrected.

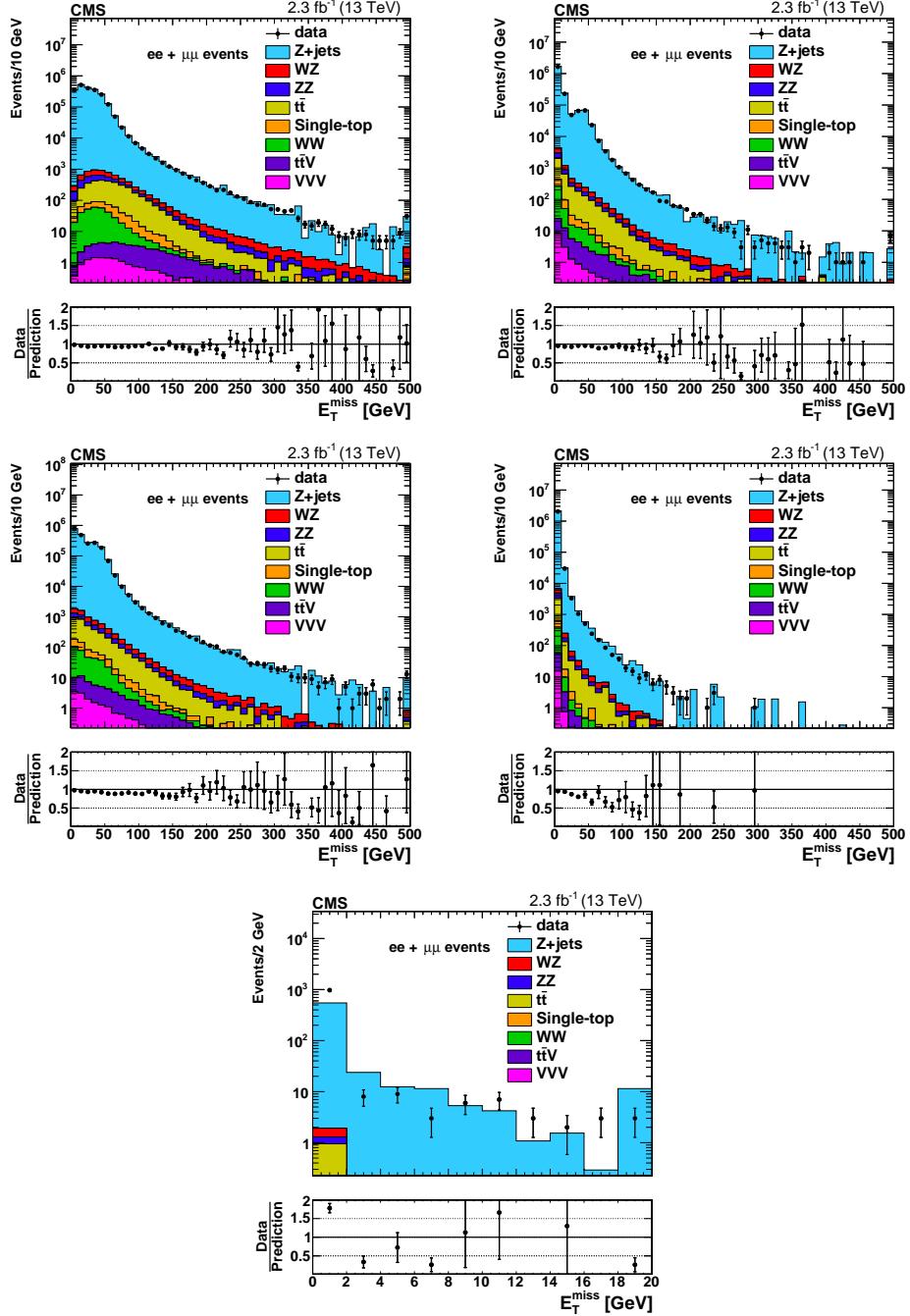
### 3.2.1 Data vs. MC Comparison

It is important to understand how well  $E_T^{\text{miss}}$  is simulated in MC in order to validate the MC samples used to simulate signal and certain background predictions. Studies were done to check the agreement in data vs. MC in different detector subsystems. The studied region is defined simply by having at least two leptons coming from a Z boson, so no real  $E_T^{\text{miss}}$  is expected in these events. The  $E_T^{\text{miss}}$  components are studied separately in different detector subsystems by choosing several categories of pf candidate type and  $\eta$  region defined in table 3.1. Charged candidates in the region with  $\eta < 2.4$ , can be used to study the tracker, neutral hadronic candidates can be used to study the HCAL, and neutral electromagnetic candidates can be used to study the ECAL. The  $E_T^{\text{miss}}$  in each region is defined by calculating the negative magnitude of the vector-sum of only the charged, neutral hadronic, or neutral electromagnetic PF candidates. The plots showing the agreement for each  $E_T^{\text{miss}}$  com-

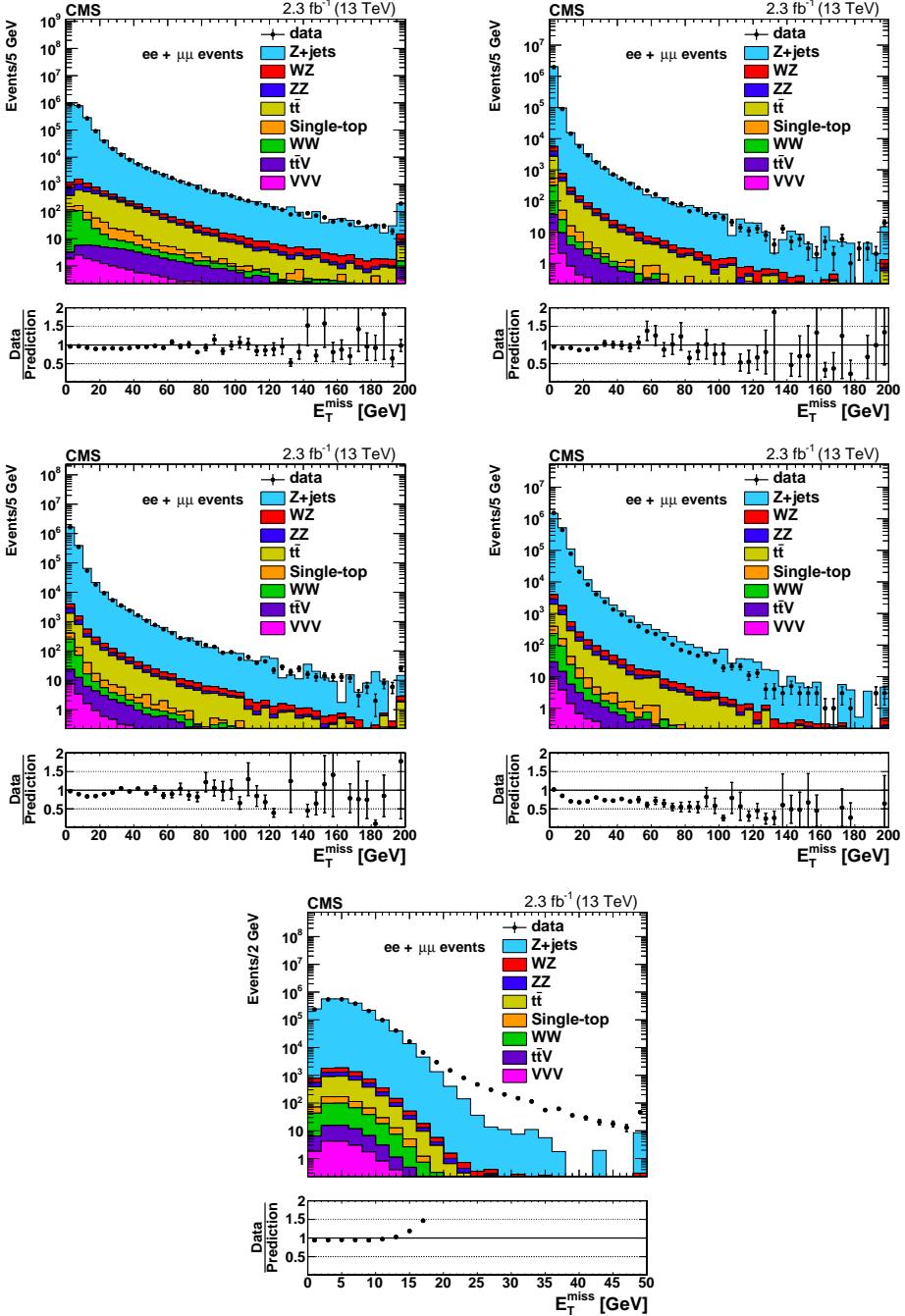
ponent are shown in figures 3.2, 3.3, and 3.4 respectively. The charged component of  $E_T^{\text{miss}}$  shows reasonable agreement over the entire CMS detector. The neutral electromagnetic component of  $E_T^{\text{miss}}$  shows good agreement in the tracker volume of CMS, and the agreement is bad for candidates outside of the tracker. The neutral hadronic component of  $E_T^{\text{miss}}$  shows poor agreement in the tracker volume of CMS, and the agreement gets much worse for candidates in the very forward region.

**Table 3.1:** Definitions for the  $\eta$  region chosen to define each subdetector region are listed in this table.

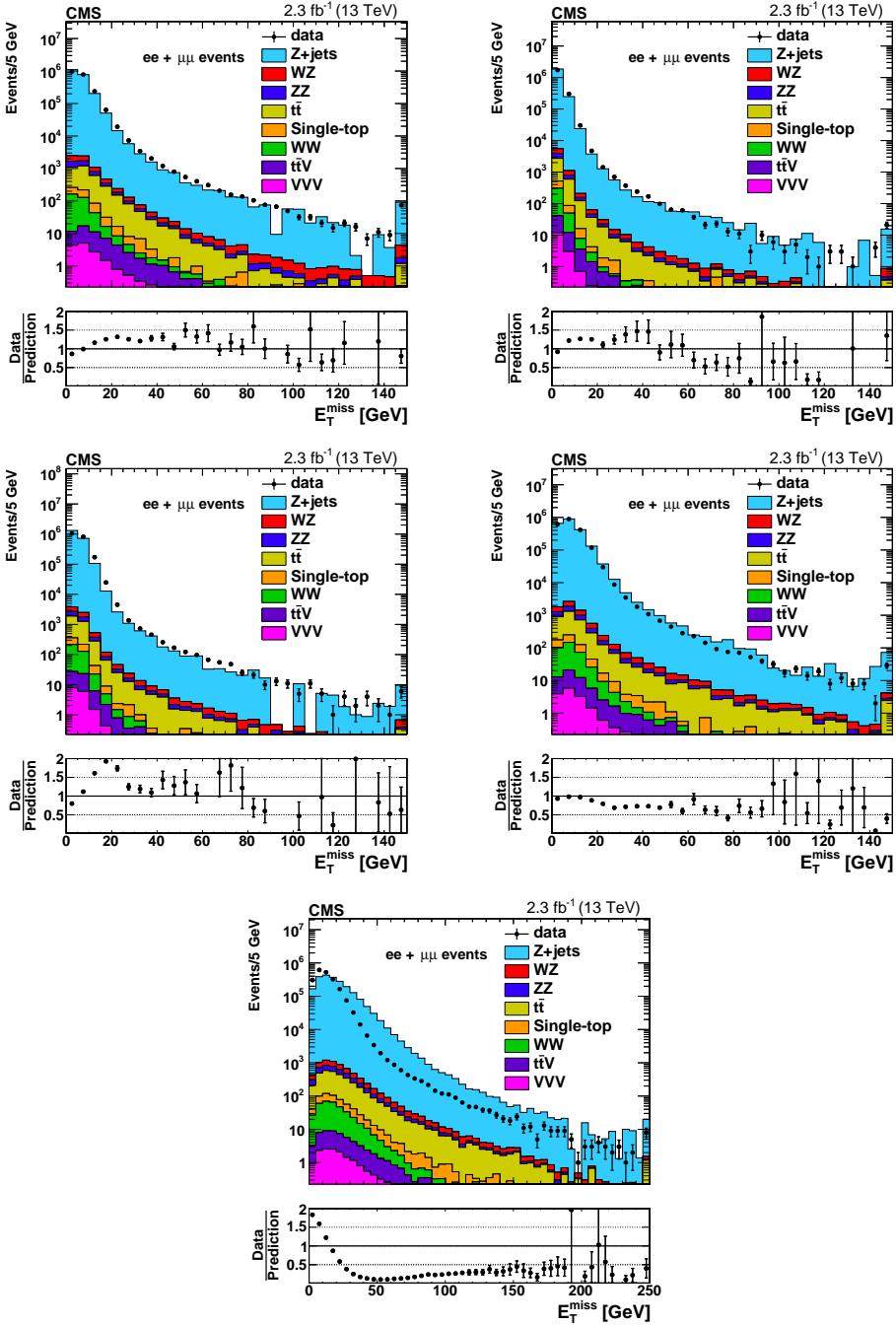
Subdetector Region	$\eta$ cut
Barrel	$ \eta  < 1.3$
Transition region	$1.3 <  \eta  < 1.6$
Endcap with tracker	$1.6 <  \eta  < 2.4$
Endcap without tracker	$2.4 <  \eta  < 3.0$
HF region	$ \eta  > 3.0$



**Figure 3.2:** The  $E_T^{\text{miss}}$  distribution is shown for charged PF candidates only. The top row shows the barrel region on the left and transition region between the barrel and endcap on the right, the second row shows the endcap region including the tracker on the left and endcap region excluding the tracker on the right and the bottom row shows the HF region only.



**Figure 3.3:** The  $E_T^{\text{miss}}$  distribution is shown for neutral electromagnetic PF candidates only. The top row shows the barrel region on the left and transition region between the barrel and endcap on the right, the second row shows the endcap region including the tracker on the left and endcap region excluding the tracker on the right and the bottom row shows the HF region only.

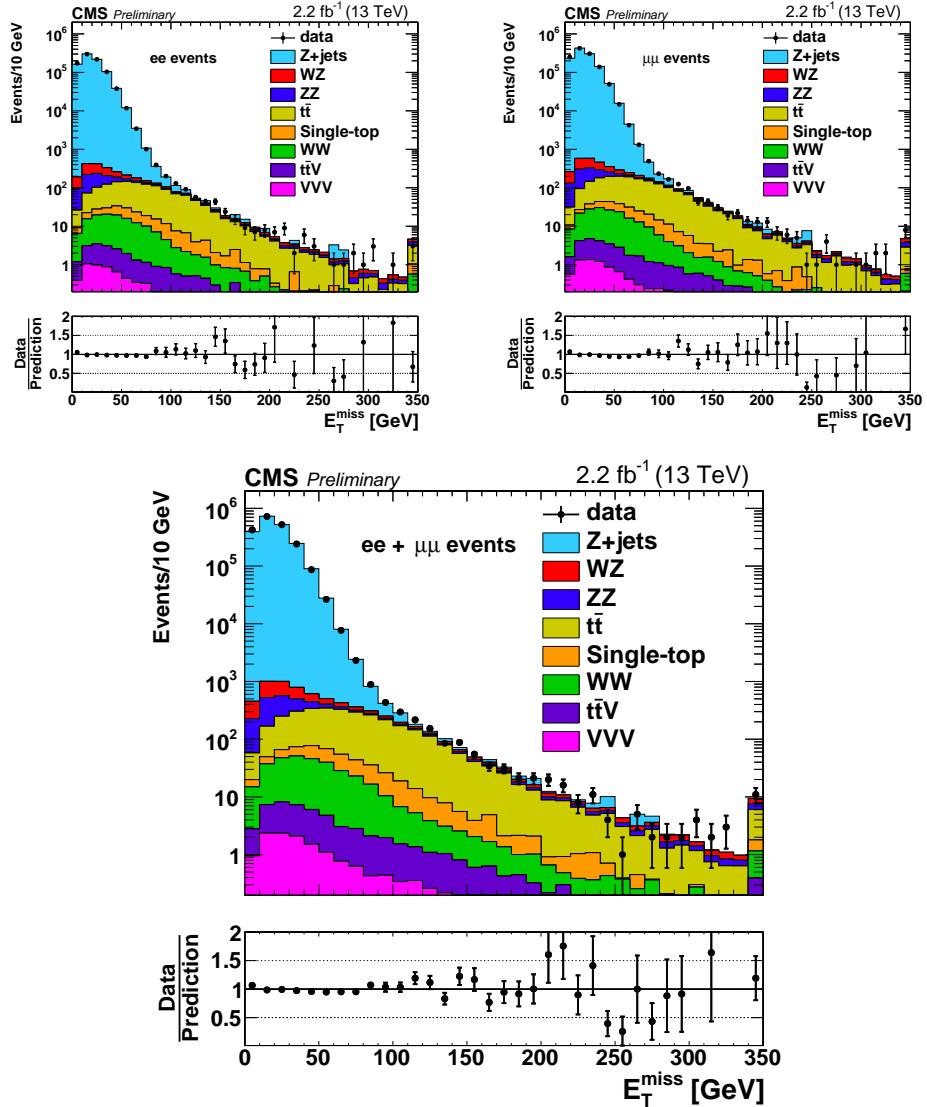


**Figure 3.4:** The  $E_T^{\text{miss}}$  distribution is shown for neutral hadronic PF candidates only. The top row shows the barrel region on the left and transition region between the barrel and endcap on the right, the second row shows the endcap region including the tracker on the left and endcap region excluding the tracker on the right and the bottom row shows the HF region only.

Finally, the  $E_T^{\text{miss}}$  distribution is shown after applying the Type-1 corrections as seen in figure 3.5. Two regions are considered in order to study events with real  $E_T^{\text{miss}}$  ( $E_T^{\text{miss}} > 100 \text{ GeV}$ ), and events with fake  $E_T^{\text{miss}}$  ( $E_T^{\text{miss}} < 100 \text{ GeV}$ ). Good agreement is seen in the real  $E_T^{\text{miss}}$  region, and reasonable agreement is seen in the fake  $E_T^{\text{miss}}$  region.

### 3.2.2 Data-driven $E_T^{\text{miss}}$ Predictions

Although good agreement is seen in data and MC after applying Type-1 corrections to  $E_T^{\text{miss}}$ , it's clear that there are still problems with the simulation based on the studies in section 3.2.1. Additionally, a new discovery in physics is important enough that it is imperative that the background predictions can be trusted. Therefore, it's important to make sure there is a reliable prediction for backgrounds coming from standard model processes in regions with large values of  $E_T^{\text{miss}}$ . Techniques were developed to predict the largest backgrounds using control regions in data, and these techniques are described in detail in chapter 5.



**Figure 3.5:** The  $E_T^{\text{miss}}$  distribution is shown with data vs. MC for events while only requirement two leptons with  $|M_{\ell\ell} - M_Z| < 10\text{GeV}$ . The top row shows ee events on the left and  $\mu\mu$  events on the right, and the bottom row shows ee+ $\mu\mu$  events together. The MC is normalized to data.

# Chapter 4

## Analysis Selections

### 4.1 Triggers and Datasets

The signal triggers used in this analysis were designed to be maximally efficient for events where a Z boson decays to two leptons. In order to pass the trigger, at least two leptons with a loose ID and isolation requirement are needed. Then lepton  $p_T$  thresholds of the triggers are low enough such that the triggers are fully efficient after making an offline  $p_T$  cut at 20 GeV for both the leading and sub-leading lepton in the event. In addition to the signal triggers, two other sets of triggers are used in order to collect data in control regions used for data-driven background predictions in the signal region, and a last set of triggers is used to collect data used to measure the efficiency of the dilepton triggers. The efficiencies of the ee, e $\mu$  and  $\mu\mu$  triggers with respect to the offline selection with two leptons with  $p_T > 20$  GeV have been measured as 0.95, 0.94 and 0.93 respectively. The way this measurement was made is described in section 5.2. All triggers used to collect data for this analysis are listed in table 4.1.

- Datasets
  - DoubleEG
  - DoubleMuon
  - MuonEG
  - SinglePhoton
  
- Datasets
  - Run2015C-05Oct2015-v1
  - Run2015D-05Oct2015-v1
  - Run2015D-PromptReco-v4

**Table 4.1:** A list of all triggers used in this analysis is shown in the table below. The signal triggers all require at least two same flavor leptons, The control triggers require either at least one photon, or at least two different flavor leptons.

<b>Signal triggers</b>
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v*
HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v*
HLT_Mu27_TkMu8_v*
HLT_Ele17_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*
HLT_DoubleEle33_CaloIdL_GsfTrkIdVL_MW_v*
<b>e<math>\mu</math> triggers</b>
HLT_Mu17_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v*
HLT_Mu8_TrkIsoVVL_Ele17_CaloIdL_TrackIdL_IsoVL_v*
HLT_Mu30_Ele30_CaloIdL_GsfTrkIdVL_v*
<b>single-<math>\gamma</math> triggers</b>
HLT_Photon22_R9Id90_HE10_IsoM_v*
HLT_Photon30_R9Id90_HE10_IsoM_v*
HLT_Photon36_R9Id90_HE10_IsoM_v*
HLT_Photon50_R9Id90_HE10_IsoM_v*
HLT_Photon75_R9Id90_HE10_IsoM_v*
HLT_Photon90_R9Id90_HE10_IsoM_v*
HLT_Photon120_R9Id90_HE10_IsoM_v*
HLT_Photon165_R9Id90_HE10_IsoM_v*
HLT_Photon165_HE10_v*
<b>H<sub>T</sub> triggers</b>
HLT_PFHT200_v*
HLT_PFHT250_v*
HLT_PFHT300_v*
HLT_PFHT350_v*
HLT_PFHT400_v*
HLT_PFHT475_v*
HLT_PFHT600_v*
HLT_PFHT650_v*
HLT_PFHT800_v*
HLT_PFHT900_v*

## 4.2 Simulations from Monte-Carlo

The CMS detector is simulated using the Geant4 toolbox [13], and various physics processes are simulated. This simulation keeps track of energies deposited by each particle, above a certain threshold, as the incident particle's shower progresses through detector materials. This process is very computationally intensive and time consuming, so an alternative method of simulation using parameterized shapes of particle showers, named FastSim, was developed. The simulation is composed of the following steps

- Modeling the Interaction Region
- Modeling the particles passage through the hierarchy of volumes that compose CMS detector and of the accompanying physics processes
- Modeling the effects of multiple interactions per beam crossing and/or the effect of events overlay (PileUp simulation)
- Modeling the detector's electronics response (digitization)

The actual simulation of the physics process is done using a variety of different software where the various packages used perform differently at simulating certain physics effects. In this analysis,  $t\bar{t}$  events are simulated using powheg [14] and Drell-Yan processes are simulated using MadGraph [15]. The matrix element calculations performed with these generators are then interfaced with pythia 8 [16] for the simulation of parton showering and hadronization.

After selecting events with two leptons, data is compared to MC and good agreement can be seen in the plots shown in section 4.2.1. Cross sections and k-factors

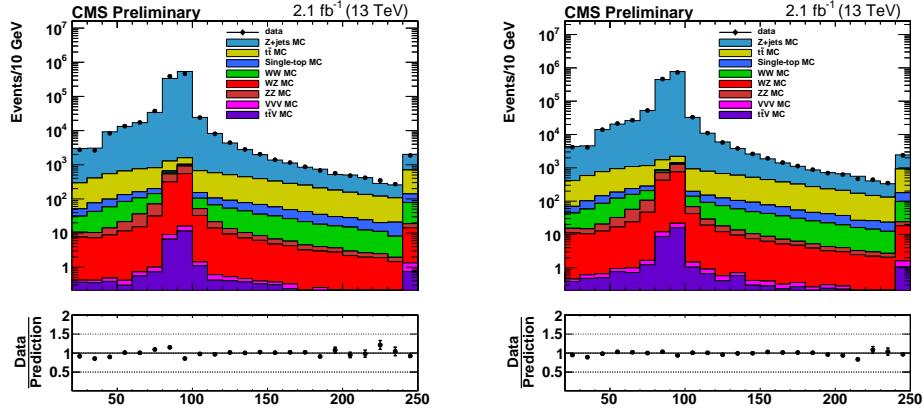
are applied to MC backgrounds to normalize them to the calculated luminosity of the signal triggers used to take data in 2015, which is  $2.3 \text{ fb}^{-1}$ . Cross section values are calculated using the FEWZ software package [17], and factors are derived by comparing data to MC in control regions in order to correct the MC for beyond leading-order effects. The events from MC are required to pass the same triggers as in data where the triggers are simulated in the MC, and the MC is then reweighted to accurately represent the number of pileup events in data.

**Table 4.2:** A list of MC samples used in this analysis are shown in this table. The MC samples used for background studies are all made with different generators, depending on which physics process that is being modeled. The signal MC is produced using the FastSim package.

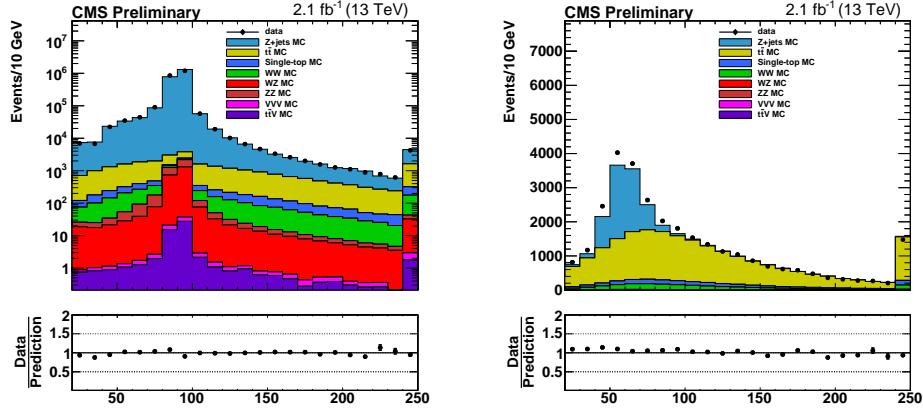
Process	Dataset Name	Cross Section [pb]
Z+jets	/DYJetsToLL_M-10to50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/*	18610
	/DYJetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/*	6104
$t\bar{t}$	/TTTo2L2Nu_13TeV-powheg/*	87.31483776
ZZ	/ZZTo4L_13TeV_powheg_pythia8/*	1.256
	/ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8/*	3.38
	/ZZTo2L2Nu_13TeV_powheg_pythia8/*	
WZ	/WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8/*	5.52
	/WZTo3LNu_TuneCUETP8M1_13TeV-powheg_pythia8/*	4.42965
WW	/WWTo2L2Nu_13TeV-powheg/*/	118.7
single top	/ST_tW_top_5f_inclusiveDecays_13TeV-powheg_pythia8/*	38.09
	/ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg_pythia8/*	38.09
$t\bar{t}V$	/TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8/*	0.2043
	/TTWJetsToQQ_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8/*	0.40620
	/TTGJets_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8/*	3.697
	/TTZToQQ_TuneCUETP8M1_13TeV-amcatnlo-pythia8/*	0.5297
	/TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8/*	0.2529
	/ttHToNonbb_M125_13TeV_powheg_pythia8/*	0.5085
VVV	/ZZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8/*	0.01398
	/WWZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8/*	0.1651
	/WZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8/*	0.05565
	*RunII Spring15 MiniAODv2-74X_mcRun2_asymptotic_v2-v*/MINIAODSIM	
Signal	SMS-T5ZZ_mGluino-*_mLSP-*_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	See [18]

### 4.2.1 Data vs. MC Comparison

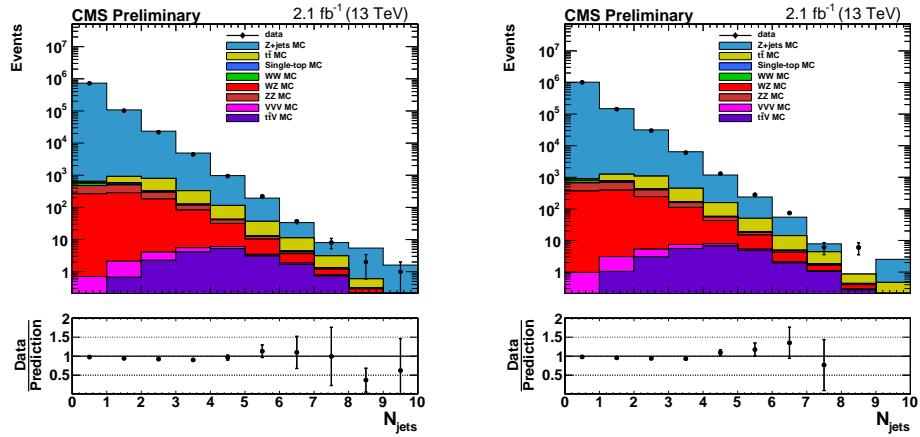
data is compared to MC simulation for many kinematic variables relevant to the search.



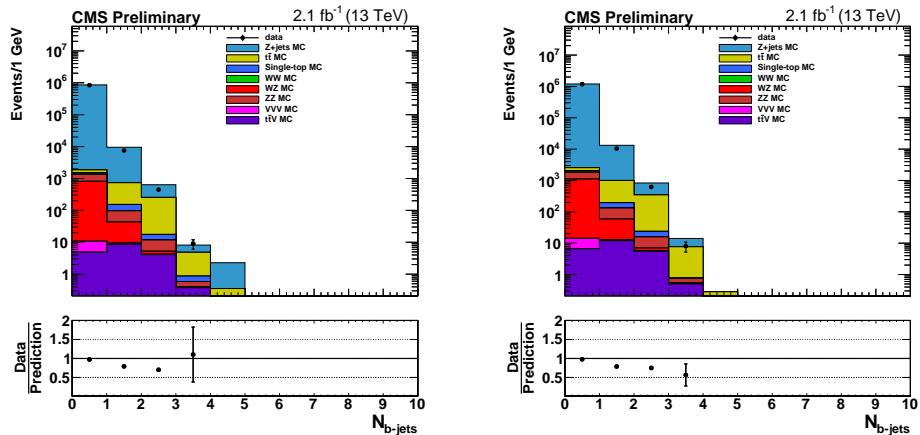
**Figure 4.1:** Data vs. MC comparison showing  $m_{\ell\ell}$  with ee events on the left and  $\mu\mu$  events on the right. For this analysis, a cut is required that  $|M_{\ell\ell} - M_Z| < 10\text{GeV}$ .



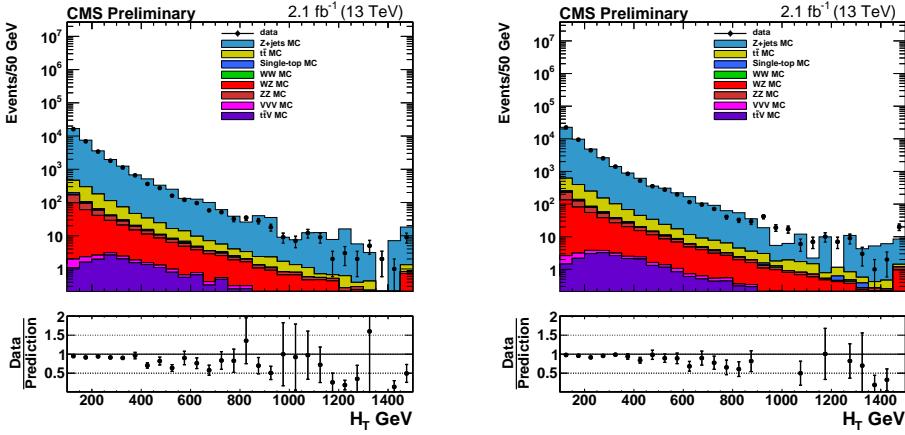
**Figure 4.2:** data vs. MC comparison showing  $m_{\ell\ell}$  with ee+ $\mu\mu$  events on the left and e $\mu$  events on the right.



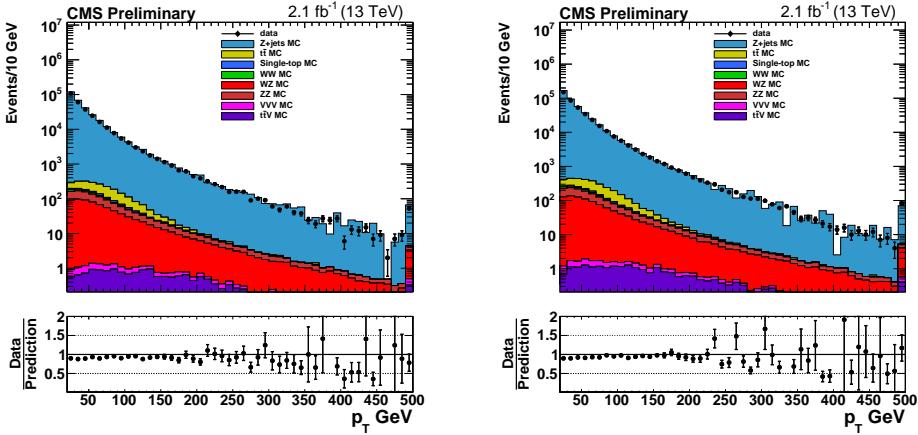
**Figure 4.3:** data vs. MC comparison showing  $N_{\text{jets}}$  with ee events on the left and  $\mu\mu$  events on the right.



**Figure 4.4:** data vs. MC comparison showing  $N_{\text{b-jets}}$  with ee events on the left and  $\mu\mu$  events on the right.



**Figure 4.5:** data vs. MC comparison showing  $H_T$  with ee events on the left and  $\mu\mu$  events on the right.



**Figure 4.6:** data vs. MC comparison showing  $Z p_T$  with ee events on the left and  $\mu\mu$  events on the right.

## 4.3 Event Selection

The object selections in this analysis are based on recommendations made by the relevant physics object groups (POGs) in CMS. These groups are responsible for providing analyzers using data from the CMS experiment with a guideline of how to select specific objects in the event. These object selections are presented to the experiment as a whole and approved before being commissioned. In the following sections, the specific object and event selections used in this analysis are described in detail.

### 4.3.1 Vertex Selection

The method to choose the primary vertices in the event is described in section 2.5.2. For every event, we require the presence of at least one primary vertex satisfying the following criteria:

- vertex is not fake; # of degrees of freedom(ndf)> 4; quality requirement
- $\rho < 2 \text{ cm}$  &  $|z| < 24 \text{ cm}$ ; require vertex is inside tracker volume

$\rho$  is the radial distance from the center of the beam spot, and  $|z|$  is the distance along the beam direction calculated from the center of the beam spot.

### 4.3.2 MET filters

A set of filters designated “ $E_T^{\text{miss}}$  filters” are used in this analysis, where these filters are designed to remove events in data where large values of  $E_T^{\text{miss}}$  occur that can come from effects such as noise in the detector or the reconstruction algorithm failing.

The first of these is simply a requirement that at least one good primary vertex is reconstructed in the event. The next filter is designed to remove events where hits are left in the CSC due to “beam halo” effects. Beam halo is caused by radiation coming from the interactions with the beam in the beam pipe before the beams are collided inside the detector volume. Additionally, a filter is applied to deal with energy spikes in the HCAL. These spikes come from beam halo effects where particles interact with the HCAL in a direction along the beamline leaving large energy signatures. A filter is then applied for situations where the ECAL readout is saturated. This happens in regions of the ECAL where there is no coverage due to bad crystals or lack of crystals. In these regions, the trigger readout is still available but saturates above a certain energy. When the trigger readout is saturated and the  $E_T^{\text{miss}}$  vector aligns with an ECAL object in  $\phi$ , these events are removed. The next filter applied has to do with known regions in the ECAL where the crystals were previously shown to give incorrect measurements in an inconsistent way. When events have large energies in any of these regions aligned with  $E_T^{\text{miss}}$ , the event is filtered. The last filter has to do with muon reconstruction. Muons are reconstructed using hits in the muon chamber and matching them with tracks in the tracker. When a low quality muon is reconstructed, there is a non-negligible chance for the muon to be matched incorrectly in such a way that it gives a large  $p_T$  value, which then translates to large  $E_T^{\text{miss}}$ .

### 4.3.3 Lepton Selection

The goal of this analysis is to study events with a Z boson that decays leptonically, and we consider only events with  $Z \rightarrow \ell\ell$ , where  $\ell = e$  or  $\mu$ .  $\tau$ s are not considered for reasons described in 2.5.4. In order to select events with this decay, it is required

there be at least two leptons passing the ID and isolation requirements described below, and that these leptons form a opposite-sign same-flavor (OSSF) pair with  $m_{\ell\ell}$  between 81-101 GeV. When more than one good OSSF di-lepton pair passes all the lepton requirements, the pair consisting of the two highest  $p_T$  leptons is chosen. The two highest  $p_T$  leptons in the event are required to have  $p_T > 20$  GeV. This requirement is chosen based on the expected di-lepton trigger  $p_T$  thresholds. In order to have high precision when identifying and reconstructing leptons, the leptons are required to be contained within the tracker volume. This is done by requiring both leptons to have  $|\eta| < 2.4$ . In order to avoid the transition region from the endcap to the barrel, where the muon and electron reconstruction efficiencies are very different, leptons having  $|\eta|$  in the range  $1.4 < |\eta| < 1.6$  are rejected. Events are vetoed if the two leading leptons are within a cone of  $\Delta R < 0.1$  from each other, in order to reduce the rate of events with fake leptons. These selections are summarized below.

- Require two leading leptons to be opposite-sign same-flavor (OSSF) ee and  $\mu\mu$  pairs
- The dilepton invariant mass is required to be consistent with the Z mass; namely  $81 < m_{\ell\ell} < 101$  GeV
- $p_T > 20$  GeV for the two leptons making up the di-lepton pair
- $|\eta| < 1.4$  or  $1.6 < |\eta| < 2.4$  for the two leptons making up the di-lepton pair
- $\Delta R$  between the two leading leptons must be greater than 0.1

#### 4.3.4 Lepton Isolation

The signal region targeted by this analysis is expected to have large  $H_T$  due to the many jets in the final state. This leads to an increased probability of a lepton overlapping with a jet in the event. When this happens, the lepton is no longer isolated and may fail a cut on the raw isolation value. The isolation variable is designed to reduce this inefficiency by using a variable cone size that depends on lepton  $p_T$ . For leptons with  $p_T < 50$  GeV, we use a cone with  $\Delta R = 0.2$ , for leptons with  $p_T$ : 50 - 200 GeV, we use a variable cone defined as  $\Delta R = \frac{10}{p_T(\text{lep})}$ , and for leptons with  $p_T > 200$  GeV, we use a cone with  $\Delta R = 0.05$ . Isolation is calculated according to equation 4.1, where  $p_T(i)$  represents all the pf candidates within the cone defined above. When cutting on isolation, the value is chosen relative to  $p_T(\text{lepton})$ . This quantity is designated mini-relative Isolation (miniRelIso).

$$I_{\text{lepton}} = \sum_i |p_T(i)| - |p_T(\text{lepton})| \quad (4.1)$$

Before cutting on isolation, corrections are made to account for pileup energy using the effective area  $\rho$  corrections scheme. In this scheme,  $\rho$  is the total energy density from pileup and is assumed to be uniform in the detector. The correction to isolation is done by calculating the energy from pileup in the isolation cone and subtracting it from the total energy in the cone.

This correction is not derived using the geometric area of the cone however due to the fact that the response of  $I_{\text{lepton}}$  and  $\rho$  are different with respect to the amount of pileup. Instead, the geometric area of each lepton's isolation cone is scaled in separate regions of  $\eta$  by a factor derived using the number of primary vertices in the event as a way to estimate the amount of pileup. These factors are derived

separately for electrons and muons and shown in tables 4.3 and 4.4.

**Table 4.3:** Effective area values for muons derived in separate regions of  $\eta$  are shown in this table.

$\eta$ region	Muon effective area
$ \eta  < 0.8$	0.0735
$0.8 <  \eta  < 1.3$	0.0619
$1.3 <  \eta  < 2.0$	0.0465
$2.0 <  \eta  < 2.2$	0.0433
$2.2 <  \eta  < 2.5$	0.0577

**Table 4.4:** Effective area values for electrons derived in separate regions of  $\eta$  are shown in this table.

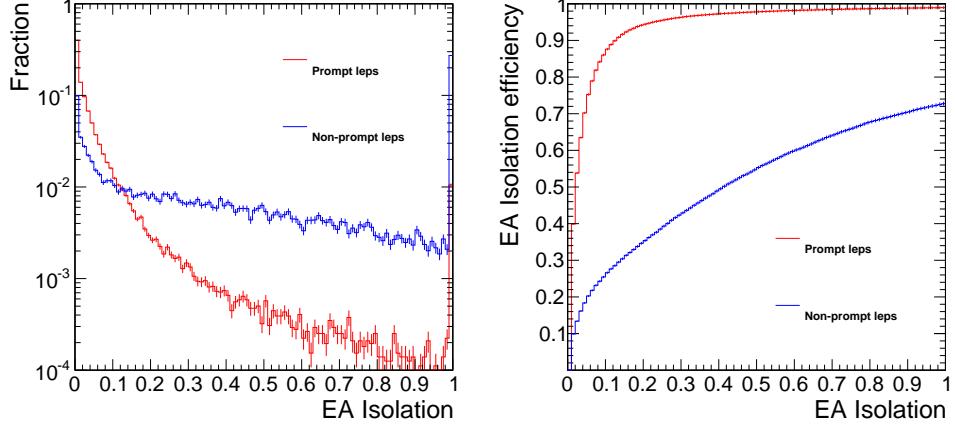
$\eta$ region	Electron effective area
$ \eta  < 1.0$	0.1752
$1.0 <  \eta  < 1.3$	0.1862
$1.479 <  \eta  < 2.0$	0.1411
$2.0 <  \eta  < 2.2$	0.1534
$2.2 <  \eta  < 2.3$	0.1903
$2.3 <  \eta  < 2.4$	0.2243
$2.4 <  \eta  < 2.5$	0.2687

The effective area is then defined by equation 4.2 where  $f$  is the scale factor, and  $R$  is the cone size defined above.

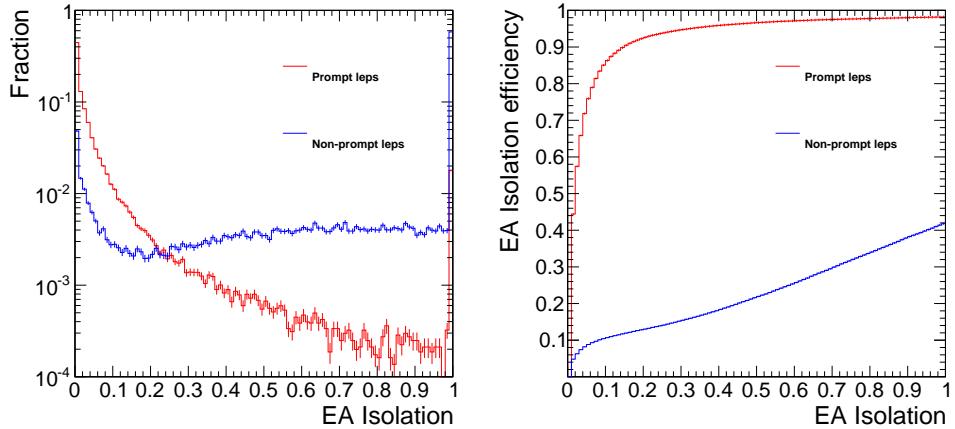
$$A_{\text{eff}} = f * \pi * R^2 \quad (4.2)$$

The plots in figures 4.7 and 4.8 show the isolation distribution for leptons that pass the ID selections made using simulated Z+jets MC. These leptons are separated into two categories, prompt and non-prompt leptons. Prompt leptons are defined to be leptons that come directly from the hard scatter process rather than a second order process, for example semi-leptonic b-decay. In order to study the difference in

expected behavior of prompt and non-prompt leptons, a distinction is made such that a prompt, reconstructed lepton is defined as being matched to a generator-level lepton within a cone of  $\Delta R < 0.4$ . Any lepton that passes the ID requirements that does not pass this matching requirement is designated as a non-prompt lepton.



**Figure 4.7:** The isolation distribution is shown for electrons on the left in simulated Z + jets MC events. On the right, the efficiency is shown when integrating all bins to the left of the value on the x-axis.

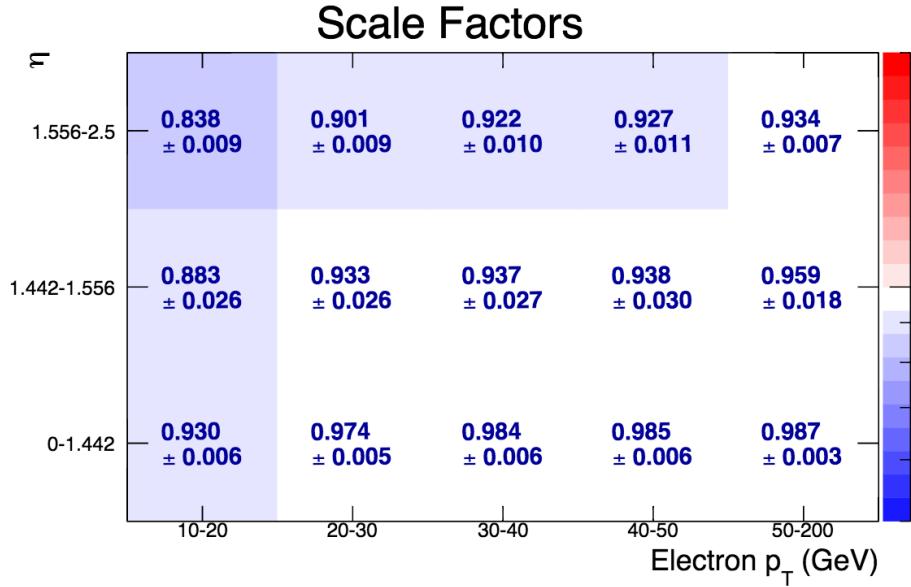


**Figure 4.8:** The isolation distribution is shown for muons on the left in simulated Z + jets MC events. On the right, the efficiency is shown when integrating all bins to the left of the value on the x-axis.

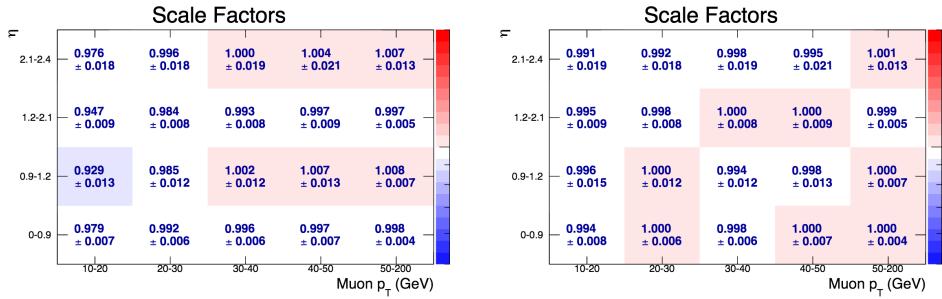
After applying the effective area corrections, the leptons in this analysis are required to have  $\text{miniRelIso} < 0.10$  ( $0.20$ ) for electrons (muons). This gives an efficiency for prompt leptons of  $> 90\%$  and rejection rate for non-prompt leptons of about 80% for electrons and 90% for muons.

### 4.3.5 Lepton Scale Factors

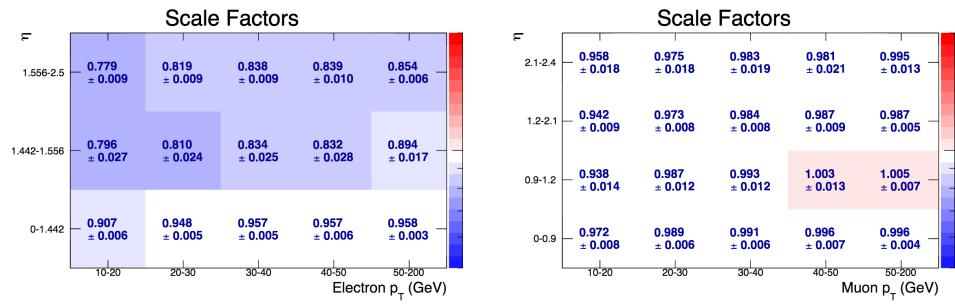
Scale factors are applied in order to correct for the differences in efficiencies in data and MC when applying lepton ID and isolation requirements. Additional scale factors are derived to correct for differences in simulation of leptons when using Fastsim [19] instead of fullsim. These scale factors are derived in data and MC using a “tag and probe” technique. The main premise of this technique is that leptons coming from a Z boson are produced in opposite-sign same-flavor pairs in very large quantities. In an event, a lepton with very tight ID and isolation requirements is chosen (the tag), and then a second lepton with very loose requirements (the probe) can be studied if it is found that the pair of leptons has a dilepton mass consistent with the Z mass. The efficiency of various cut requirements is then assessed for individual leptons, such as the efficiency of the ID and isolation requirements. This is done separately in data and MC, and then the MC is corrected based on these differences to match what is observed in data. Scale factors for electron ID and isolation criteria are shown in 4.9, scale factors for muon ID and isolation criteria are shown in 4.10, and Fastsim scale factors are shown in 4.11.



**Figure 4.9:** Electron scale factors as a function of  $p_T$  and  $\eta$ . The scale factor is applied as an event weight once per electron in the event.



**Figure 4.10:** Muon scale factors as a function of  $p_T$  and  $\eta$ . The left plot shows the scale factors associated with the ID selection criteria, and the right plot shows the scale factors associated with the isolation selection criteria. Each scale factor is applied as an event weight once per muon in the event.



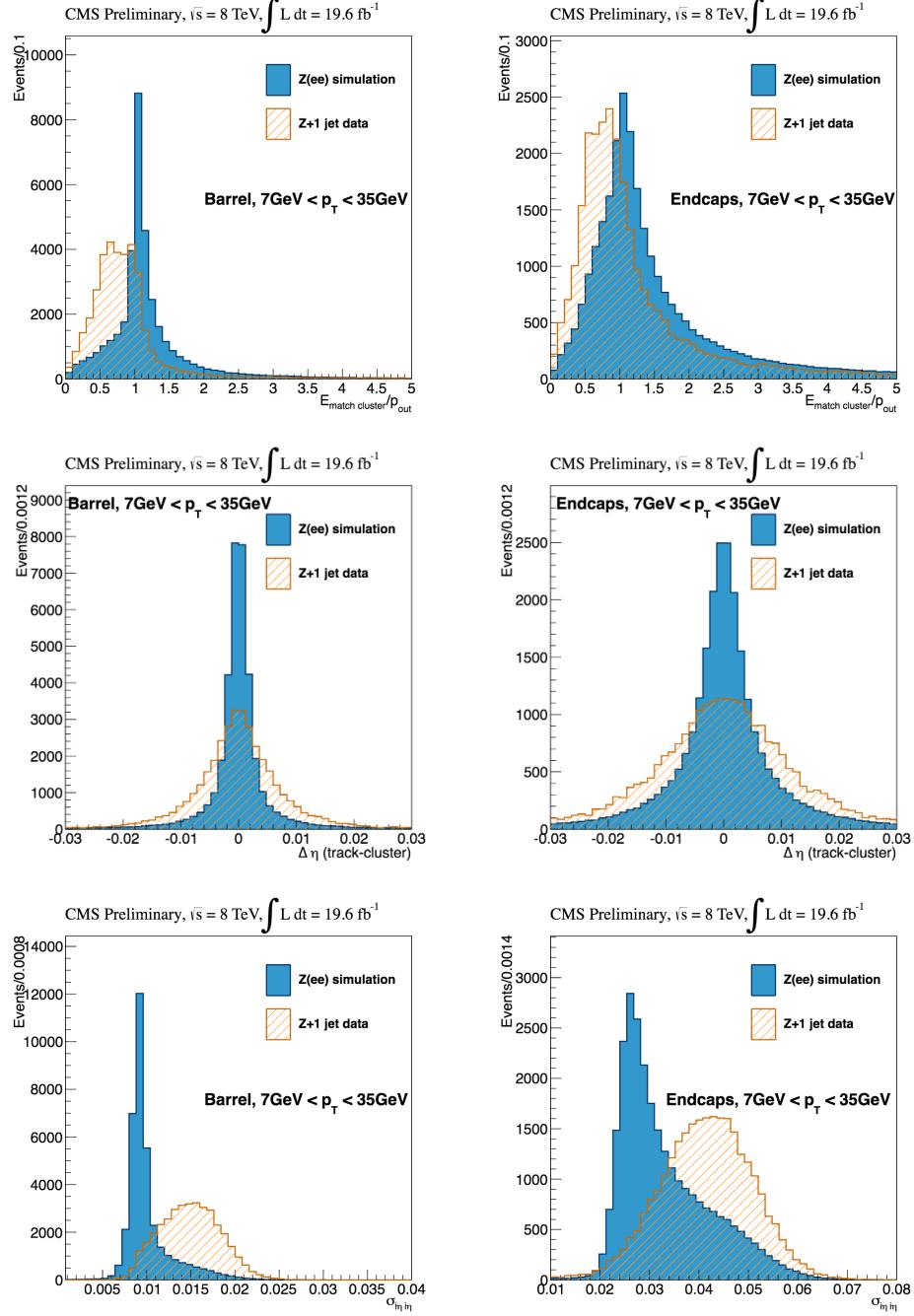
**Figure 4.11:** Fastsim scale factors as a function of  $p_T$  and  $\eta$ . The left plot shows the scale factors applied per electron, and the right plot shows the scale factors applied per muon. Each scale factor is applied as an event weight once per lepton in the event.

### 4.3.6 Electron Selection

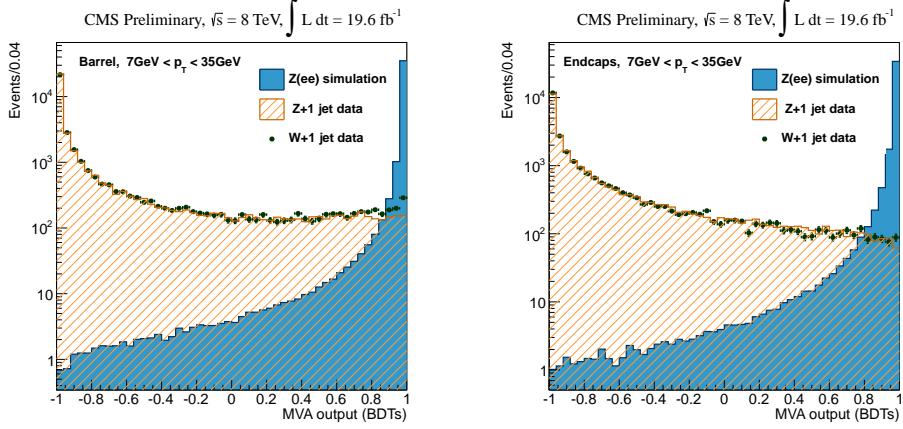
Output from a multivariate analysis (MVA) is used to identify the electrons used in this analysis [20] based on a method used previously [21]. This is done separately in the barrel and endcap using a variety of kinematic variables as input. Electrons with  $p_T > 15$  GeV and  $|\eta| < 2.4$  are considered. The training of the MVA is done using a Z+jets MC sample where the electrons used as signal input are required to be matched to generator level electrons, and electrons used as background are required to not be matched.

Variables used to differentiate electrons from fakes mainly use information from the tracker and ECAL. Measurements made by the ECAL and tracker are compared using the ratio of the electron cluster energy to the track momentum at the outermost track layer. This ratio is expected to be 1 for real electrons, but not necessarily for fakes.  $\sigma_{in\eta}$  is a purely calorimetric variable which uses the shower shape of the electron object to differentiate real and fake electrons. The main premise behind this variable is that electron energies are contained within 1-2 crystals in  $\eta$  where ECAL energy from charged hadrons is spread across many crystals leading to the shape for real electrons and fake electrons to be very different making it a good variable to differentiate the two. A way to differentiate real and fake electrons is to check the consistency of the track of the reconstructed electron object left in the tracker. This is done by measuring the difference in  $\eta$  of the track measured in the outer layer and the electron cluster in the ECAL ( $\Delta\eta(\text{track} - \text{cluster})$ ). An plot showing how each of these variables looks for real and fake electrons is shown in figure 4.12, and the output of the MVA variable in figure 4.13.

Electrons in data are compared to the electrons from MC using the probe leg



**Figure 4.12:** Kinematic variables used to distinguish real electrons from fake electrons are shown with electrons in the barrel on the left and electrons in the endcap on the right.



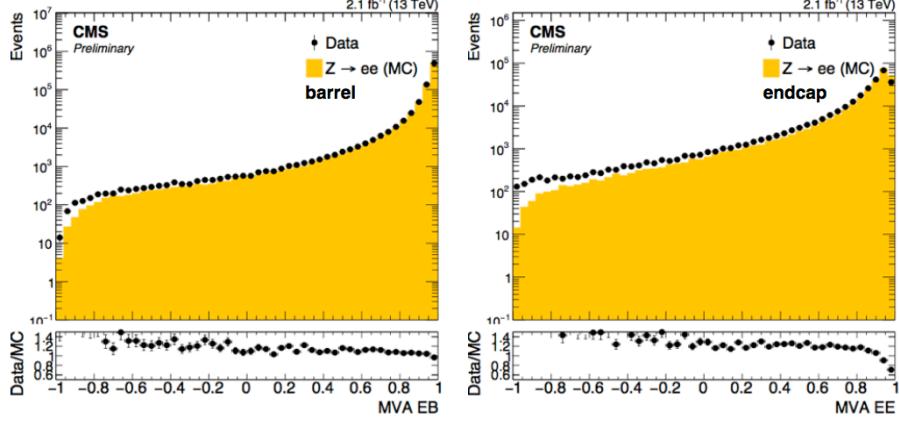
**Figure 4.13:** The output of the electron MVA is shown comparing the distributions of electrons matched with generator level electrons (blue) vs. jets that fake electrons. The output in the barrel is shown in the left and endcap on the right.

from a “tag and probe” technique similar to the one described in section 4.3.5. The requirements for the probe electron are as follows:

- $p_T(e) > 10 \text{ GeV}$
- $\eta(e) < 2.5$
- Loose identification criteria
- Relative isolation (pileup corrected)  $< 0.1$

The output of this MVA is shown in figure 4.14, and the MVA cut requirement for this analysis is shown in table 4.5.

The efficiency of the electron selection is measured to be on average 80% for electrons with  $p_T > 20 \text{ GeV}$  as seen in figure 4.15.



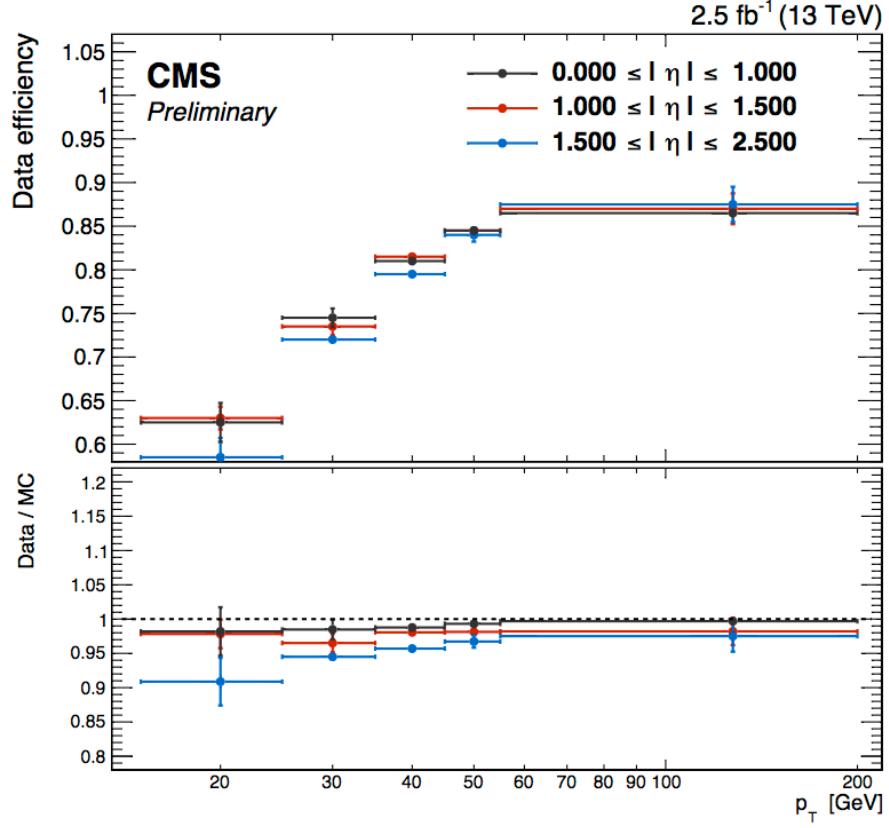
**Figure 4.14:** Output of the electron MVA is shown comparing electrons in data vs. MC with electrons in the barrel on the left and the endcap on the right. The electrons shown are the probe leg when using a “Tag and Probe” technique.

**Table 4.5:** All the requirements for electrons with  $p_T > 10$  GeV are shown in the table below. The electrons in the signal region are required to have  $p_T > 20$  GeV.

Region	MVA minimum cut value
Barrel $ \eta  < 0.8$	0.913286
Barrel $ \eta  > 0.8$	0.805013
Endcap	0.358969
Cut variable	Requirement
$d_0$ (w.r.t. 1st good PV)	< 0.05 cm
$d_z$ (w.r.t. 1st good PV)	< 0.1 cm
miniRelIso / $p_T$	< 0.10

### 4.3.7 Muon Selection

The muon selection criteria is based on studies performed by the muon POG at CMS [22][23]. Muons with  $p_T > 15$  GeV and  $|\eta| < 2.4$  are considered. The baseline muon selection requires that the muon be identified as a muon by the particle flow algorithm, as well as being identified as a tracker or global muon. To define a tracker muon, all tracks in the tracker with  $p_T > 0.5$  GeV and  $p > 2.5$  GeV are considered.



**Figure 4.15:** Electron identification efficiency in data (top) and data to MC efficiency ratios (bottom) measured for the identification criteria based on a MVA discriminant with an average efficiency of 80%. The efficiency is measured with the tag and probe method and shown in three pseudorapidity ranges as a function of the electron transverse momentum.

These tracks are extrapolated to the muon system and then required to have a matching hit in at least one muon segment. Global muons are defined by matching a tracker track with that of a track reconstructed independently by the muon system. After the two tracks are matched, a global muon track fit is done combining the hits from the two separate tracks. Muons in this analysis must either qualify as a global muon or pass a segment compatibility requirement. Segment compatibility is defined using a template based on simulation. In addition to these requirements, muons are required to be within 0.05 cm of the primary vertex in the x-y direction and within

0.1 cm in the z direction. The full muon selection requirements are listed in Table 4.6.

**Table 4.6:** A summary of the muons selection requirements is shown in this table for all muons used in this analysis.

good Global Muon Requirements	
Quantity	Requirement
Fraction of valid tracker hits	> 0.8
Normalized global-track $\chi^2$	< 3
Tracker-Standalone position match	< 12
Kick finder	< 20
Segment compatibility	> 0.303
tight segment compatibility	> 0.451 or passes above requirements
muon type	(global or tracker) and PF muon
$d_0$ (w.r.t. 1st good PV)	< 0.05 cm
$d_z$ (w.r.t. 1st good PV)	< 0.1 cm
miniRelIso / $p_T$	< 0.20

### 4.3.8 Photons

In this analysis it is not necessary that the photon sample is high in purity, it is only required that the photon-like object in each event is well-measured. The reason for this is discussed in more detail in section 5.1. Photons are required to pass a loose working point designed by the EGamma POG at CMS with some additional cuts. This working point was designed to be 90% efficient with a background rejection of greater than 80%, and photons with  $p_T > 22$  GeV and  $\eta < 2.4$  are considered. Photons that are in the transition region between the ECAL barrel and ECAL endcap are vetoed, specifically photons with  $1.4 < \eta < 1.6$ . Cuts made on different versions of the photon object's isolation calculated using independent sets of reconstructed particle flow objects. A cut is made on absolute pf charged hadron isolation with no corrections, and then cuts are made to the neutral hadron and photon isolations after

applying  $\rho$  corrections as described in section 4.3.4. Due to the different response of the detector in the barrel and endcap, the cuts are tuned separately for these regions. The cuts on neutral isolation are made as a function of the photon  $p_T$ , and all the isolation requirements are listed in table 4.7.

**Table 4.7:** A summary of the photon isolation requirements is shown in this table for all photons used in the template method prediction.

PF objects	Requirement in barrel	Requirement in endcap
charged hadron	$< 3.32$	$< 1.97$
$\rho$ corrected neutral hadron	$< 1.92 + 0.014*p_T(\gamma)$ $+0.000019*p_T^2(\gamma)$	$< 11.86 + 0.0139*p_T(\gamma)$ $+0.000025*p_T^2(\gamma)$
$\rho$ corrected photon	$< 0.81 + 0.0053*p_T(\gamma)$	$< 0.83 + 0.0034*p_T(\gamma)$

In addition to the isolation requirements, further cuts are made on photons to maximize the energy resolution of the photon and also remove events with real  $E_T^{\text{miss}}$ . The photon is required to be mostly electromagnetic, and a cut is made on the ratio of hadronic energy to electromagnetic energy ( $\frac{H}{E}$ ). In order to reject fake photons from  $\pi^0$  decays, a requirement is made of  $\sigma_{i\eta i\eta} < 0.0102$  (0.274) for photons in the barrel (endcap). It is possible to end up with real  $E_T^{\text{miss}}$  in photon events in the process of  $W \rightarrow e + \bar{\nu}_e$  where the electron immediately radiates most of its energy away in the form of a photon. Before radiating away its energy, it is possible for the electron to leave a track in the pixel layer, therefore in order to remove these events, it is required that there be no pixel track matched to the photon. These events are also removed by rejecting photons which have an electron of at least  $p_T > 10$  GeV matched within a cone of  $dR < 0.2$ . It is required that a pfjet with  $p_T > 10$  GeV be matched to the photon with a cone of  $\Delta R < 0.3$ . The matched jet is then required to have an overall electromagnetic energy fraction of at least 70%. Lastly, events with a photon that is aligned with  $E_T^{\text{miss}}$  having  $\Delta\phi < 0.14$  are removed. The full photon

requirement is listed in 4.8.

**Table 4.8:** A summary of the photon selection requirements is shown in this table for all photons used in the template method prediction.

Quantity	Requirement
$p_T$	$> 22 \text{ GeV}$
$ \eta $	$< 2.4$
veto photons in transition region	$ \eta :1.4 - 1.6$
$\frac{H}{E}$	$< 0.5$
$\sigma_{i\eta i\eta}$ (barrel)	$< 0.0102$
$\sigma_{i\eta i\eta}$ (endcap)	$< 0.274$
$\Delta\phi(\gamma, E_T^{\text{miss}})$	$> 0.14$
No matching pixel track (pixel veto)	
reject $\gamma$ within $dR < 0.2$ of an electron	
Matched to a pfjet with emfrac $> 0.7$	

### 4.3.9 MET

Type 1 corrected  $E_T^{\text{miss}}$  is used in this analysis, and a full description can be found in chapter 3.

### 4.3.10 Jets

The jets used in this analysis are described in section 2.5.5. Jets with  $p_T > 35 \text{ GeV}$  and  $|\eta| < 2.4$  are used with the L1, L2 and L3 corrections applied to MC and the additional ResidualL2L3 corrections applied to data. The residualL2L3 correction is used to correct the scale of the jets in MC to match what is measured in data as a function of the jet's  $p_T$ , and  $\eta$ . A loose selection criteria, listed in table 4.9, is applied to each jet to remove jets coming from noise in the detector. In addition to this, any jet that is found to be near a lepton within a cone of  $dR = 0.4$  is removed from the

event. This is done to avoid possibly double counting energy. The full jet selection is listed in table 4.10.

**Table 4.9:** All the criteria for a jet to pass the loose ID requirement is shown in this table.

Quantity	Requirement
Neutral hadron fraction	< 0.99
Neutral EM fraction	< 0.99
Charged hadron fraction	> 0
Charged multiplicity	> 0
Charged EM fraction	< .99
Number of constituents	> 1

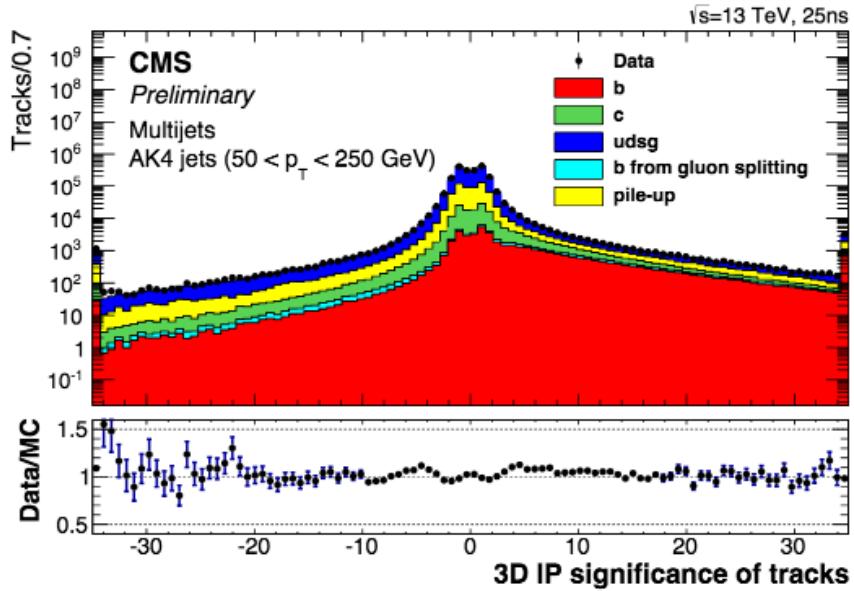
**Table 4.10:** A summary of the jet selection requirements is shown in this table for all jets used in this analysis.

Quantity	Requirement
$p_T$	> 35 GeV
$ \eta $	< 2.4
Pass loose jet ID	See table 4.9

In addition to using the above requirements to classify jets, it is also useful to separate jets into two categories, light-flavored and b-tagged. This distinction is useful since b-tagged jets tend to come from processes having top quarks in the decay chain, and these backgrounds can be reduced by removing events with b-tagged jets. In order for the search to remain inclusive, signal regions containing b-tagged jets are analyzed separately.

The Combined Secondary Vertex (CSV) algorithm [24][25] is used to tag b-jets. This algorithm makes use of the fact that b-quarks have a larger average lifetime than lighter partons leading to the possibility to identify a second vertex away from the primary vertex where the tracks in the b-jet are associated with this secondary

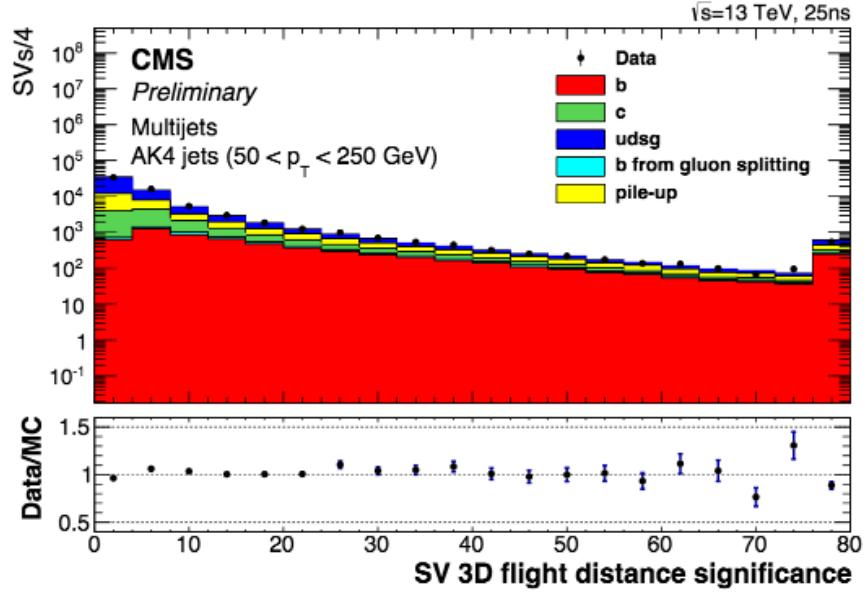
vertex. A diagram showing this process is shown in figure 4.19. Other variables that are used to discriminate heavy flavor jets from light are the track multiplicity, jet mass and jet energy. These variables can be seen in figures 4.16, 4.17, and 4.18. In this analysis, a working point was chosen which corresponds to the misidentification rate for b-jets to be 1%. For this working point, the b-tagging efficiency is measured to be 65% per b-tag for the working point used. The b-tag discriminator distribution can be seen in figure 4.20, and the cut value corresponding to the working point in this analysis is  $\text{CSV} > 0.890$ .



**Figure 4.16:** A plot of the impact parameter for the secondary vertex with respect to the primary vertex which is used in order to differentiate jets from heavy and light flavor decay.

### 4.3.11 $H_T$

The signal model targeted by this analysis is expected to have large amounts of hadronic activity. A simple way to quantify the amount of hadronic activity in an



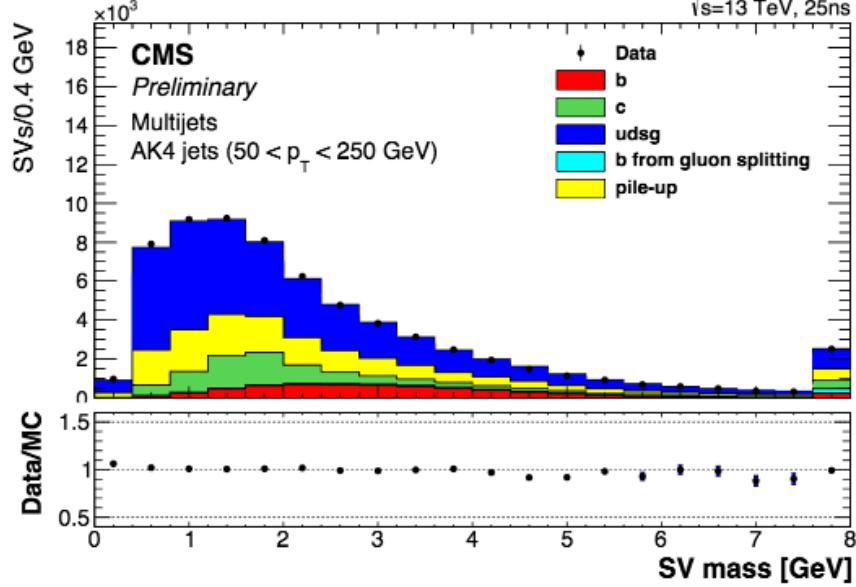
**Figure 4.17:** A plot of the 3-D flight distance of the secondary vertex with respect to the primary vertex which is used in order to differentiate jets from heavy and light flavor decay.

event is to use the variable “ $H_T$ ”.  $H_T$  can be simply described as the scalar sum of the  $p_T$  of the jets in the event and is defined by the equation 4.3.

$$H_T = \sum |\mathbf{p}_T(\text{jet})| \quad (4.3)$$

### 4.3.12 Signal Region Definitions

Multiple, orthogonal signal regions are defined based on cuts on the number of jets,  $H_T$ , and  $E_T^{\text{miss}}$ . These regions are then divided up into two categories, with and without b-tags. The signal regions are classified into two separate categories called the “A” signal regions and “B” signal regions. The “A” regions are defined as having  $H_T > 400$  GeV and 2-3 jets, whereas the “B” regions are defined as having at least 4 jets. This leads to 16 orthogonal signal regions. These regions are analyzed

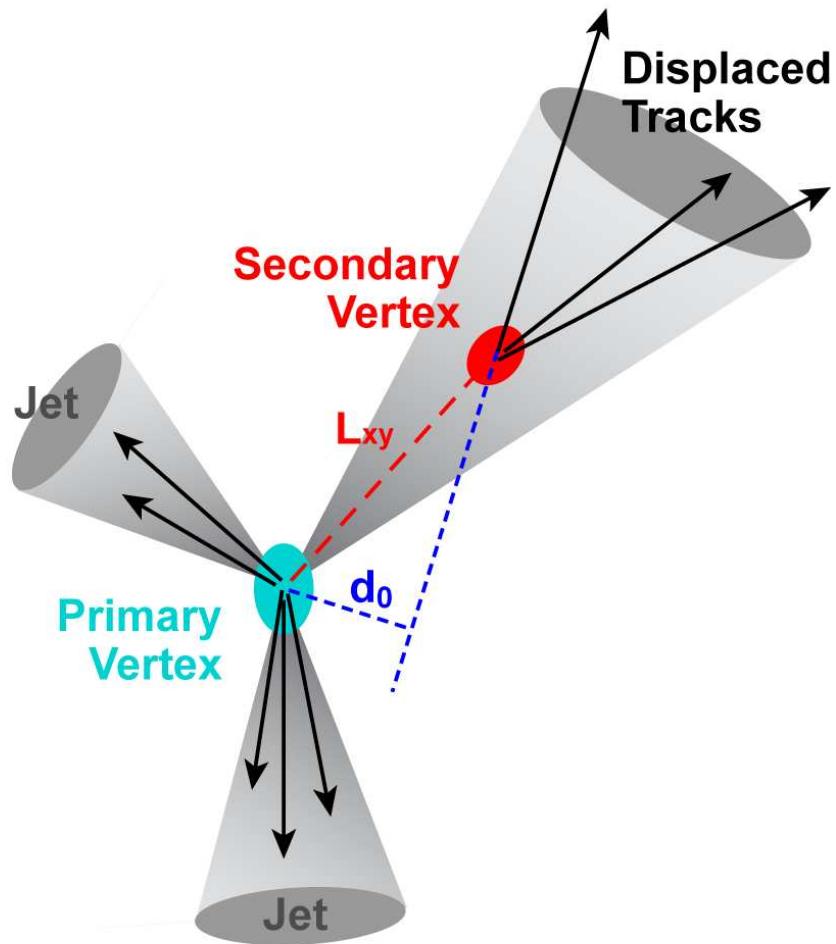


**Figure 4.18:** A plot of the mass of the jet originating at the secondary vertex which is used in order to differentiate jets from heavy and light flavor decay.

separately and then eventually combined to interpret results using the signal model described in 1.4.

In addition to the inclusive signal regions listed, a special signal region is defined which is designed to probe the region where a  $3.0\sigma$  excess was seen in run-I by ATLAS [26]. The requirements of this signal region in this analysis are as follows:

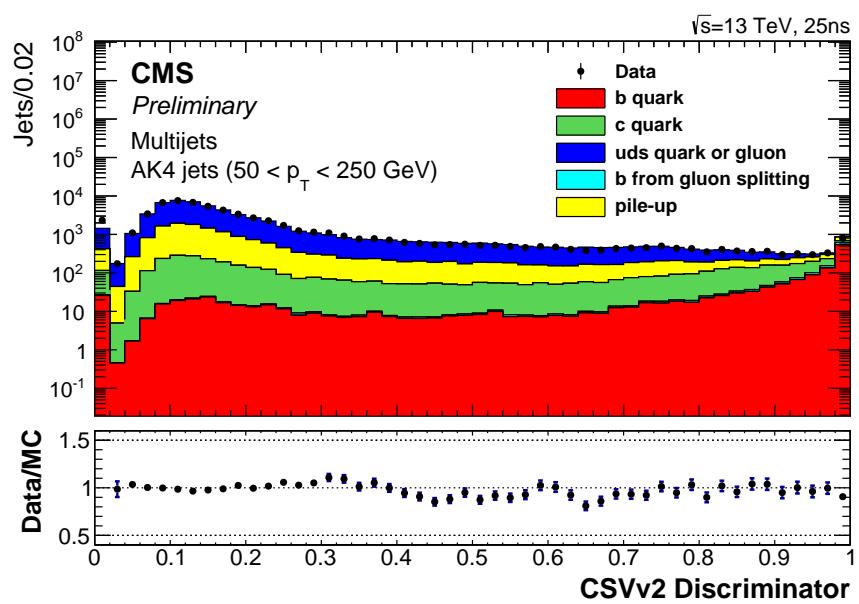
- at least 2 jets
- $H_T + p_T(\ell_1) + p_T(\ell_2) > 600$  GeV
- $E_T^{\text{miss}} > 225$  GeV
- $\Delta\phi(E_T^{\text{miss}}, \text{jet}_{1,2}) > 0.4$



**Figure 4.19:** A plot showing a jet forming away from the primary vertex of a heavy flavor parton, for example a b or c quark.

#### 4.3.13 Selection Summary

This section summarizes all of the object selections used in this analysis. The preselection and the signal region selections are summarized in table 4.11.



**Figure 4.20:** The distribution showing the CSVv2 Discriminator is shown comparing data vs. MC. The MC is split by jet parton flavor.

**Table 4.11:** A list of all selections used to define the preselection and signal regions is shown in the following table.

preselection	
2 OSSF leptons ( $ee$ , $\mu\mu$ )	$p_T > 20$ GeV, $ \eta  < 2.4$
Dilepton invariant mass	$81 < m_{\ell\ell} < 101$ GeV
Jets	$p_T > 35$ GeV, $ \eta  < 2.4$
Minimum number of jets	$\geq 2$
b-tag requirement (medium)	$CSVv2IVF > 0.890$
Signal selections	
b-tag requirements	
$N_{b\text{-tags}}$	$= 0$
$N_{b\text{-tags}}$	$\geq 1$
$E_T^{\text{miss}}$ binning	
$E_T^{\text{miss}}$	100 - 150 GeV
$E_T^{\text{miss}}$	150 - 225 GeV
$E_T^{\text{miss}}$	225 - 300 GeV
$E_T^{\text{miss}}$	$> 300$ GeV
A Signal regions	$N_{\text{jets}} = 2$ or $3$ and $H_T \geq 400$ GeV
B Signal regions	$N_{\text{jets}} \geq 4$
ATLAS Signal region	$N_{\text{jets}} \geq 2$ , $H_T + p_T(\ell_1) + p_T(\ell_2) > 600$ GeV, $\Delta\phi(E_T^{\text{miss}}, jet_{1,2}) > 0.4$ , $E_T^{\text{miss}} > 225$ GeV

# Chapter 5

## Background Estimation Methods

The techniques used to estimate the SM backgrounds in the signal regions are described in detail in this chapter. The SM backgrounds fall into three categories, Z+jets, Flavor symmetric, and other SM backgrounds. The first two backgrounds are estimated using data from a control region orthogonal to the signal region with techniques developed for this analysis, and the backgrounds falling into the third category are estimated using MC simulation. The MC samples used to estimate these backgrounds are studied in a control region which is orthogonal to the signal regions.

### 5.1 Estimating the Z+jets Background with $E_T^{\text{miss}}$ Template

The Z+jets background is estimated using a method named the  $E_T^{\text{miss}}$  template method. The estimate is done using a control sample of data events gathered using a suite of single photon triggers. In events with a Z boson recoiling off of a sys-

tem of jets, there should be no real  $E_T^{\text{miss}}$ . However, the resolution of the jet system is poor leading to events with  $E_T^{\text{miss}}$  due to these effects. The resolution of photons is similar to that of leptons, so the assumption is made that  $E_T^{\text{miss}}$  in  $\gamma + \text{jets}$  events is from the same sources as in  $Z + \text{jets}$  events. This method also works in cases where there is real  $E_T^{\text{miss}}$  in the event, for example in events containing a b-jet that decays semi-leptonically, since the  $E_T^{\text{miss}}$  source is not from the leptons or photon. The main difference between these samples is the fact that the Z-boson is massive whereas the photon is not. To account for this difference, the photon's  $p_T$  distribution is reweighted to match that of the dilepton system's  $p_T$ .

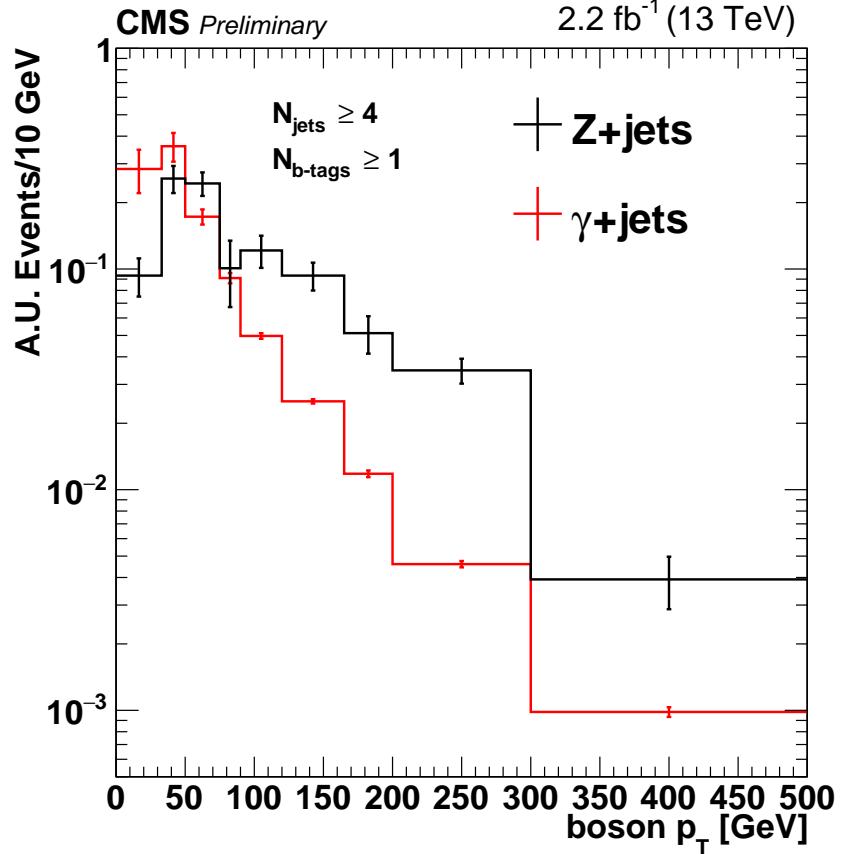
This method does not require that the photon-like object used in the  $\gamma + \text{jets}$  events be extremely pure, it only matters that the photon-like object is well-measured. The selection described in Sec. 4.3.8 is used to select the photon-like object. The  $\gamma + \text{jets}$  events are selected with a suite of single photon triggers with  $p_T$  thresholds varying from 22–165 GeV. These triggers as well as their average prescales are listed in table 5.1. Each  $\gamma + \text{jets}$  event is weighted by the trigger prescale, in order to remove the bias in photon  $p_T$  before the reweighting is done. When an event passes a trigger, the trigger information is stored for all triggers that pass. Due to the way the prescale software is written, the triggers with larger prescales must be prescaled by an integer multiple of a lower prescale value, otherwise the trigger information will not be saved.

In order to account for the fact that the Z-boson is massive whereas the photon is massless, the  $\gamma + \text{jets}$  sample is reweighted such that the boson  $p_T$  matches that of the  $Z + \text{jets}$  sample. A separate reweighting scheme is derived for each signal region, where the same cuts on  $N_{\text{jets}}$  and  $H_T$  are applied to the  $Z + \text{jets}$  and  $\gamma + \text{jets}$  data

**Table 5.1:** A list of single- $\gamma$  triggers used in this analysis and their prescales is shown in this table.

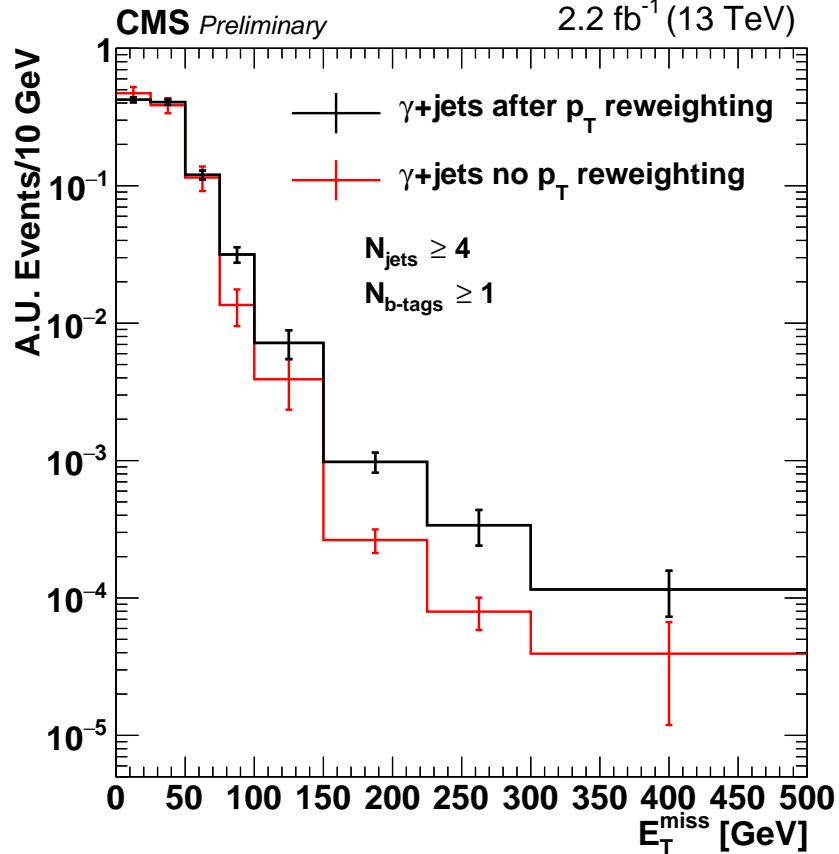
$p_T$ threshold [GeV]	trigger name	prescale
22	HLT_Photon22_R9Id90_HE10_IsoM_v*	1008
30	HLT_Photon30_R9Id90_HE10_IsoM_v*	504
36	HLT_Photon36_R9Id90_HE10_IsoM_v*	168
50	HLT_Photon50_R9Id90_HE10_IsoM_v*	24
75	HLT_Photon75_R9Id90_HE10_IsoM_v*	4
90	HLT_Photon90_R9Id90_HE10_IsoM_v*	2
120	HLT_Photon120_R9Id90_HE10_IsoM_v*	1
165	HLT_Photon165_R9Id90_HE10_IsoM_v*	1
165	HLT_Photon165_HE10_v*	1

samples. The result of this reweighting can be seen in figures 5.1 and 5.2.



**Figure 5.1:** The  $p_T$  distributions are shown for the  $Z + \text{jets}$  (black) and  $\gamma + \text{jets}$  (red) events in signal region B where at least 1 jet is required to be b-tagged. The distributions are normalized to unit area.

After reweighting in  $p_T$ , the resulting  $E_T^{\text{miss}}$  distribution is normalized in a region where  $Z + \text{jets}$  is the dominant background, namely,  $E_T^{\text{miss}} < 50$  GeV. In order to assess the systematic uncertainties associated with the template method, a closure test is done using MC simulation where simulated  $\gamma + \text{jets}$  events are used to predict the  $E_T^{\text{miss}}$  in simulated  $Z + \text{jets}$  events. The closure is assessed separately in each signal region, and uncertainties are assigned based on the results of this closure test. The uncertainties of this method are described in detail in subsection 5.1.1.

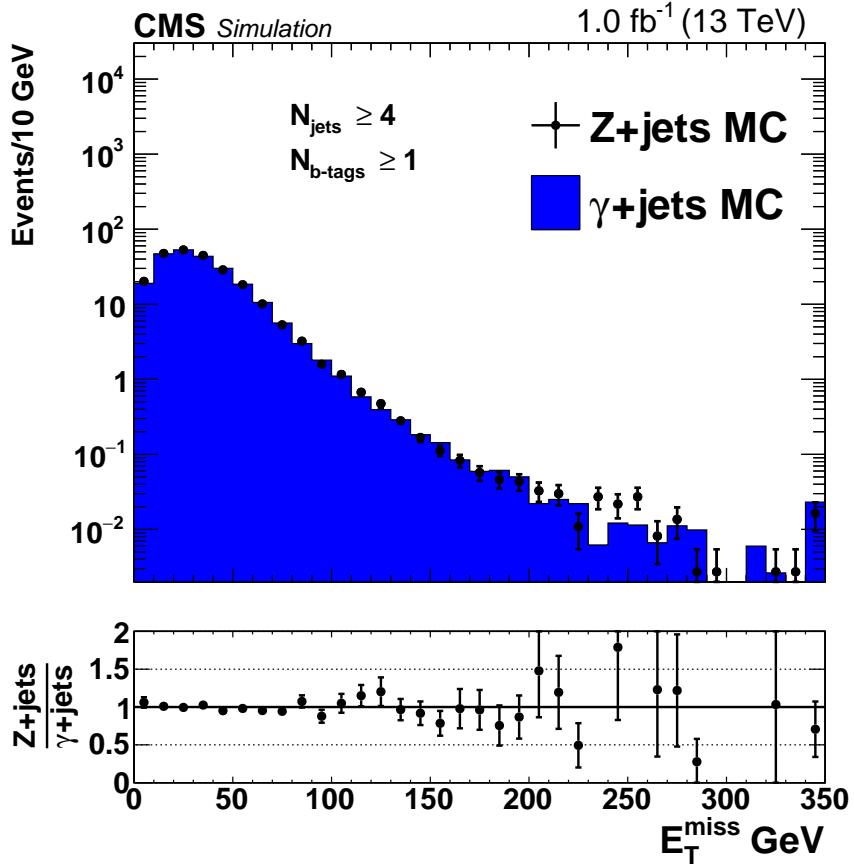


**Figure 5.2:** The  $E_T^{\text{miss}}$  distributions are shown before reweighting the  $\gamma + \text{jets}$  sample in  $p_T$  (red) and after reweighting (black) for  $\gamma + \text{jets}$  events in signal region B where at least 1 jet is required to be b-tagged. The distributions are normalized to unit area.

### 5.1.1 Systematic Uncertainties in the $E_T^{\text{miss}}$ Template Prediction

The systematic uncertainty associated with the  $E_T^{\text{miss}}$  template prediction comes from two sources: a MC closure study, and the uncertainty associated with normalizing in low  $E_T^{\text{miss}}$ . The largest uncertainty comes from the results of the closure study. This study is done by generating a  $E_T^{\text{miss}}$  template using simulated  $\gamma + \text{jets}$  events and using this template to predict the  $E_T^{\text{miss}}$  for events in simulated  $Z + \text{jets}$

events. This is done using the exact same procedure described in section 5.1. For the MC closure test, we evaluate the systematic uncertainty separately in various regions including our signal regions. The results of the closure test are shown for signal region B where at least 1 jet is required to be b-tagged is shown in figure 5.3 and table 5.2. The uncertainty for each region is chosen to cover the largest discrepancy between the  $\gamma + \text{jets}$  prediction of the  $Z + \text{jets}$  MC for each  $E_T^{\text{miss}}$  region, or the statistical uncertainty whichever is larger. The systematic uncertainty chosen for all regions is listed in table 5.3.



**Figure 5.3:** The closure test for Signal region B where at least 1 jet is required to be b-tagged is shown where simulated  $\gamma + \text{jets}$  events are used to predict the  $E_T^{\text{miss}}$  for simulated  $Z + \text{jets}$  events. The distributions are normalized to  $1 \text{ fb}^{-1}$  of data.

**Table 5.2:** Results of the MC closure test shown for signal region B where at least 1 jet is required to be b-tagged. The systematic uncertainty for each region is chosen to cover the largest descrepancy between the  $\gamma + \text{jets}$  prediction of the Z+jets MC for each  $E_T^{\text{miss}}$  region, or the the statistical uncertainty whichever is larger. All uncertainties shown are statistical only.

$E_T^{\text{miss}}[\text{GeV}]$	0 - 50	50 - 100	100 - 150
Z+jets	$191.4 \pm 2.9$	$40.5 \pm 0.9$	$2.9 \pm 0.2$
$\gamma + \text{jets}$	$190.2 \pm 1.1$	$41.8 \pm 0.5$	$2.7 \pm 0.1$
ratio	$1.01 \pm 0.02$	$0.97 \pm 0.02$	$1.07 \pm 0.07$
Uncertainty	2 %	3 %	10 %
$E_T^{\text{miss}}[\text{GeV}]$	150 - 225	225 - 300	$\geq 300$
Z+jets	$0.43 \pm 0.03$	$0.11 \pm 0.02$	$0.02 \pm 0.01$
$\gamma + \text{jets}$	$0.47 \pm 0.03$	$0.07 \pm 0.01$	$0.03 \pm 0.01$
ratio	$0.92 \pm 0.10$	$1.50 \pm 0.34$	$0.66 \pm 0.30$
Uncertainty	10 %	50 %	50 %

**Table 5.3:** Systematic uncertainties derived from the MC closure test shown for all the on-Z signal regions. The uncertainties for each region are chosen to cover the largest descrepancy between the  $\gamma + \text{jets}$  prediction of the Z+jets MC for each  $E_T^{\text{miss}}$  region, or the statistical uncertainty whichever is larger.

$E_T^{\text{miss}}[\text{GeV}]$	0 - 50	50 - 100	100 - 150	150 - 225	225 - 300	$\geq 300$
SRA, b-veto	1 %	4 %	4 %	5 %	15 %	35 %
SRA, with b-tags	1 %	3 %	5 %	10 %	30 %	40 %
SRB, b-veto	1 %	2 %	4 %	10 %	20 %	25 %
SRB, with b-tags	2 %	3 %	10 %	10 %	50 %	50 %
ATLAS Region	2 %	2 %	10 %	10 %	10 %	-

The systematic uncertainty associated with the normalization of the template prediction in data in the region with  $E_T^{\text{miss}}$  from 0-50 GeV is assessed in the following way. The normalization factor is derived by requiring the total background prediction in this region to add up to the yield observed in data. The uncertainty associated with the ratio of data/prediction is taken as a systematic uncertainty over the entire template prediction. Table 5.4 shows the uncertainties for each signal region.

**Table 5.4:** Systematic uncertainties on the normalization procedure are listed in the following table.

Signal region	Uncertainty
SRA	
b-veto	4%
$\geq 1$ b-tag	10%
SRB	
b-veto	3%
$\geq 1$ b-tag	6%
ATLAS Region	3%

## 5.2 Estimating the Flavor-Symmetric Background

The flavor symmetric background method is used to predict standard model backgrounds where ee and  $\mu\mu$  events are equally likely to be produced as  $e\mu$  events. This happens for events where the two leptons each come from a separate W boson. The largest background that falls into this category is  $t\bar{t}$ , but other processes that contribute are single top, WW pair production, and events with  $Z \rightarrow \tau\tau$  where the taus both decay leptonically. One important feature of events in this category is that these events all have real  $E_T^{\text{miss}}$  coming from the associated neutrino production when the W decays leptonically.

This prediction is done using a data control sample of  $e\mu$  events. These events are collected using the  $e\mu$  triggers listed in table 4.1, and a correction is made to this sample to account for the difference in efficiencies of reconstructing and triggering on electrons and muons. The prediction for the number of same-flavor (SF) events observed ( $N_{SF}^{obs.}$ ) can be obtained by using the ratio of SF to opposite-flavor (OF) events ( $R_{SF/OF}$ ) multiplied by the number of OF events observed in data( $N_{OF}^{obs.}$ ). This is done in the following way.

### 5.2.1 $R_{SF/OF}$

The number of observed events in data for each dilepton region is the number of events produced multiplied by the trigger and reconstruction efficiencies, Therefore the true number of events produced can be obtained from this relation as seen in equation 5.2.

$$N_{ee,\mu\mu,e\mu}^{observed} = N_{ee,\mu\mu,e\mu}^{true} \times (\epsilon_{ee,\mu\mu,e\mu}^{trig.} * \epsilon_{ee,\mu\mu,e\mu}^{reco.}) \quad (5.1)$$

$$N_{ee,\mu\mu,e\mu}^{true} = \frac{N_{ee,\mu\mu,e\mu}^{observed}}{\epsilon_{ee,\mu\mu,e\mu}^{trig.} * \epsilon_{ee,\mu\mu,e\mu}^{reco.}} \quad (5.2)$$

The number of SF events is the sum of ee and  $\mu\mu$  events, and the number of OF events is  $2*$  the number of  $e\mu$  events. The factor 2 in in the number of OF events comes from the combinatorics that arise from not distinguishing  $e\mu$  from  $\mu e$ , and this also leads to the total number of OF events to be equal to the total number of SF events.

$$N_{OF}^{true} = N_{SF}^{true} \quad (5.3)$$

$$N_{SF} \equiv N_{ee} + N_{\mu\mu} \quad (5.4)$$

$$N_{OF} \equiv N_{e\mu} + N_{\mu e} = 2 * N_{e\mu} \quad (5.5)$$

Combining this information,  $R_{SF/OF}$  can be written in terms of purely the trigger and reconstruction efficiencies as seen in equation 5.6.

$$R_{SF/OF} = \frac{N_{SF}^{obs.}}{N_{OF}^{obs.}} = \frac{\epsilon_{ee}^{\text{reco}} \cdot \epsilon_{ee}^{\text{trig.}} + \epsilon_{\mu\mu}^{\text{reco}} \cdot \epsilon_{\mu\mu}^{\text{trig.}}}{2\epsilon_{e\mu}^{\text{reco}} \cdot \epsilon_{e\mu}^{\text{trig.}}} \quad (5.6)$$

These ratios are measured in appropriate SF and OF control regions.

The value for  $R_{SF/OF}$  can also be calculated in a separate way which underlines the advantage of using the combined SF sample compared to the separate ee and  $\mu\mu$  samples and is done by measuring two quantities,  $r_{\mu/e}$  and  $R_T$ .  $r_{\mu/e}$  is the ratio of the number of  $\mu s$  reconstructed to the number of reconstructed electrons, shown in equation 5.7.  $R_T$  is the square root of the ratio of the number SF events to OF events, shown in equation 5.8. While  $R_{\ell\ell/OF}$  is directly affected by the differences in reconstruction and trigger efficiencies by the factors  $r_{\mu/e}$  or  $r_{\mu/e}^{-1}$ , these differences partially cancel out in  $R_{SF/OF}$ .

$$r_{\mu/e} \equiv \sqrt{\frac{N_{\mu\mu}}{N_{ee}}} = \sqrt{\frac{\epsilon_{\mu\mu}^{\text{reco.}} \cdot \epsilon_{\mu\mu}^{\text{trig.}}}{\epsilon_{ee}^{\text{reco.}} \cdot \epsilon_{ee}^{\text{trig.}}}} \quad (5.7)$$

$$R_T \equiv 2 \sqrt{\frac{N_{ee} N_{\mu\mu}}{N_{OF}}} = \frac{\sqrt{\epsilon_{ee}^{\text{reco.}} \epsilon_{ee}^{\text{trig.}} \cdot \epsilon_{\mu\mu}^{\text{reco.}} \epsilon_{\mu\mu}^{\text{trig.}}}}{\epsilon_{OF}^{\text{reco.}} \cdot \epsilon_{OF}^{\text{trig.}}}, \quad (5.8)$$

$R_{SF/OF}$  is then derived using these equations according to equation 5.9.

$$\begin{aligned} R_{ee/OF} &= \frac{N_{ee}}{N_{OF}} = \frac{1}{2} \sqrt{\frac{N_{ee}}{N_{\mu\mu}}} \times 2 \frac{\sqrt{N_{ee} N_{\mu\mu}}}{N_{OF}} = \frac{1}{2} r_{\mu/e}^{-1} \times R_T \\ R_{\mu\mu/OF} &= \frac{N_{\mu\mu}}{N_{OF}} = \frac{1}{2} \sqrt{\frac{N_{\mu\mu}}{N_{ee}}} \times 2 \frac{\sqrt{N_{ee} N_{\mu\mu}}}{N_{OF}} = \frac{1}{2} r_{\mu/e} \times R_T \\ R_{SF/OF} &= R_{ee/OF} + R_{\mu\mu/OF} = \frac{1}{2} (r_{\mu/e} + r_{\mu/e}^{-1}) \times R_T. \end{aligned} \quad (5.9)$$

The final value of  $R_{SF/OF}$  is calculated in two ways. The first is a direct measurement of the ratio in data in a control region enriched in  $t\bar{t}$  described in subsection 5.2.2, while the second consists of the separated estimation of the  $r_{\mu/e}$  and  $R_T$  factors described in subsection 5.2.3.

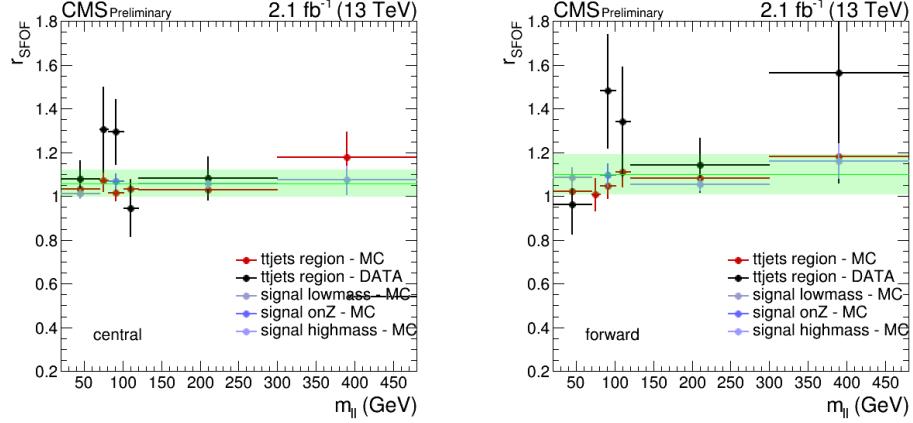
### 5.2.2 Direct measurement of $R_{SF/OF}$

$R_{SF/OF}$  is measured directly in a control region enriched in  $t\bar{t}$ . This region is orthogonal to the signal regions and defined below.

- the same lepton selection
- exactly two jets
- $\cancel{E}_T$  between 100 and 150 GeV

The results of this direct measurement of  $R_{SF/OF}$  are displayed in figure 5.4 and table 5.5 where values of MC simulation and data are compared for the aforementioned control region. The values from  $t\bar{t}$  MC are shown for comparison in the signal region and in low-mass, and high-mass regions. Additionally, the leptons are classified into two regions, “central” and “forward”. For events to be considered central, both leptons must have  $|\eta| < 1.4$ , and for events to be considered forward, at least one lepton must have  $|\eta| > 1.6$ . This is done since the leptons measured in the barrel region have better resolution, so it can be checked that the resolution does not degrade when leptons are measured in the forward region. The two values are combined using a weighted average for the final result. The values measured in data for  $R_{SF/OF}$  shown in figure 5.4 agree with the measurement made in MC within the uncertainties, and both measured values are close to 1.

The numerical values of the correction factors are obtained in the low-mass and high-mass regions combined, excluding the on-Z region because of the contamination with Drell–Yan backgrounds. They are shown in Tab. 5.5. The results from data and MC agree well within their uncertainties and are close to unity. To study the extrapolation of the measured value into the signal region, the ratio of  $R_{SF/OF}$  in the



**Figure 5.4:** Direct measurement of  $R_{SF/OF}$  in the control region for data (black) and  $t\bar{t}$  MC (red). Values in the signal region are shown in blue for  $t\bar{t}$  MC. Central rapidity is shown on the left and forward rapidity on the right. The green band represents the overall uncertainty on the measured value of  $R_{SF/OF}$ .

control and signal region on simulation is studied. It is found to be compatible with unity within the statistical uncertainty of the simulation. This statistical uncertainty is therefore assigned as the systematic uncertainty of this method.

### 5.2.3 Measuring $R_{SF/OF}$ using the Factorization Method

$r_{\mu/e}$  and  $R_T$  are measured in data in order to calculate  $R_{SF/OF}$  according to Eq. 5.9.

The measurement of  $r_{\mu/e}$  is performed in a Drell–Yan enriched control region which takes advantage the large number of ee and  $\mu\mu$  pairs from  $Z$  boson decays. This region is defined as having  $E_T^{\text{miss}} < 50$  GeV, and at least two jets. The invariant dilepton mass is then required to be near the  $Z$  boson mass,  $60 \text{ GeV} < m_{\ell\ell} < 120 \text{ GeV}$ .

The observed event yields in the two channels are shown in table 5.6, together with the resulting value of  $r_{\mu/e}$ . The results on MC are shown for comparison. It can

**Table 5.5:** Observed event yields in the control region and the resulting values of  $R_{SF/OF}$ ,  $R_{ee/OF}$ , and  $R_{\mu\mu/OF}$ . The results are shown separately for the central and forward lepton selection and the same quantities derived on simulation are shown for comparison.

	$N_{SF}$	$N_{OF}$	$R_{SF/OF} \pm \sigma_{stat}$	Transfer factor $\pm \sigma_{stat}$
Central				
Data	668	631	$1.059 \pm 0.059$	—
MC	790.9	753.2	$1.050 \pm 0.013$	$0.980 \pm 0.016$
Forward				
Data	339	306	$1.108 \pm 0.087$	—
MC	389.3	360.9	$1.079 \pm 0.021$	$1.018 \pm 0.026$
	$N_{ee}$	$N_{OF}$	$R_{ee/OF} \pm \sigma_{stat}$	Transfer factor $\pm \sigma_{stat}$
Central				
Data	269	631	$0.426 \pm 0.031$	—
MC	339.0	753.2	$0.450 \pm 0.007$	$0.988 \pm 0.020$
Forward				
Data	141	306	$0.461 \pm 0.047$	—
MC	157.4	360.9	$0.436 \pm 0.011$	$1.023 \pm 0.036$
	$N_{\mu\mu}$	$N_{OF}$	$R_{\mu\mu/OF} \pm \sigma_{stat}$	Transfer factor $\pm \sigma_{stat}$
Central				
Data	399	631	$0.632 \pm 0.040$	—
MC	451.9	753.2	$0.600 \pm 0.009$	$0.974 \pm 0.019$
Forward				
Data	198	306	$0.647 \pm 0.059$	—
MC	232.0	360.9	$0.643 \pm 0.015$	$1.013 \pm 0.030$

be seen that the efficiency for muons is higher than for electrons by 10% for central leptons and by 20% in the forward region. Consistent results are observed in the simulation.

The function in equation 5.9 where  $r_{\mu/e}$  is used is studied as a function of kinematic variables that are relevant to the search. Specifically, the function  $0.5(r_{\mu e} + 1/r_{\mu e})$  is plotted as a function of  $N_{\text{jets}}$ ,  $N_{\text{vtx}}$ , lepton  $p_T$ ,  $m_{\ell\ell}$ ,  $E_T^{\text{miss}}$ , and  $N_{b\text{-tags}}$ . The result of these studies is shown for central and forward leptons in figures 5.5 and 5.6. The dashed line illustrates the central value observed in data. No

**Table 5.6:** Result of the calculation of  $r_{\mu/e}$ . The observed event yields are shown in the Drell–Yan control region for the central and forward lepton selection in the ee and  $\mu\mu$  channels and the resulting values of  $r_{\mu/e}$ . The same quantities derived from simulation are shown for comparison.

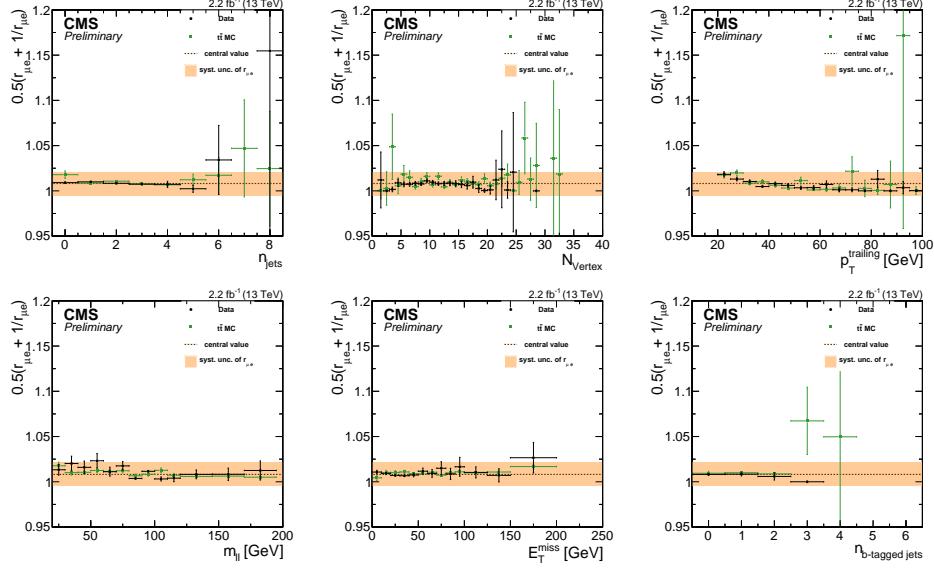
	$N_{\mu\mu}$	$N_{ee}$	$r_{\mu/e} \pm \sigma_{\text{stat.}} \pm \sigma_{\text{syst.}}$
Central			
Data	23533	18238	$1.14 \pm 0.01 \pm 0.11$
MC	30400	23711	$1.13 \pm 0.00 \pm 0.11$
Forward			
Data	14937	9807	$1.23 \pm 0.01 \pm 0.25$
MC	19449	12287	$1.26 \pm 0.01 \pm 0.25$

significant dependency on lepton  $p_T$  is observed, and the value of  $r_{\mu/e}$  is especially stable with respect to  $N_{\text{jets}}$  and  $E_T^{\text{miss}}$ . No dependency is observed on the number of reconstructed vertices showing that  $r_{\mu/e}$  is insensitive to pileup. A systematic uncertainty of 10%(20%) is assigned in the central (forward) lepton selection based on these results.

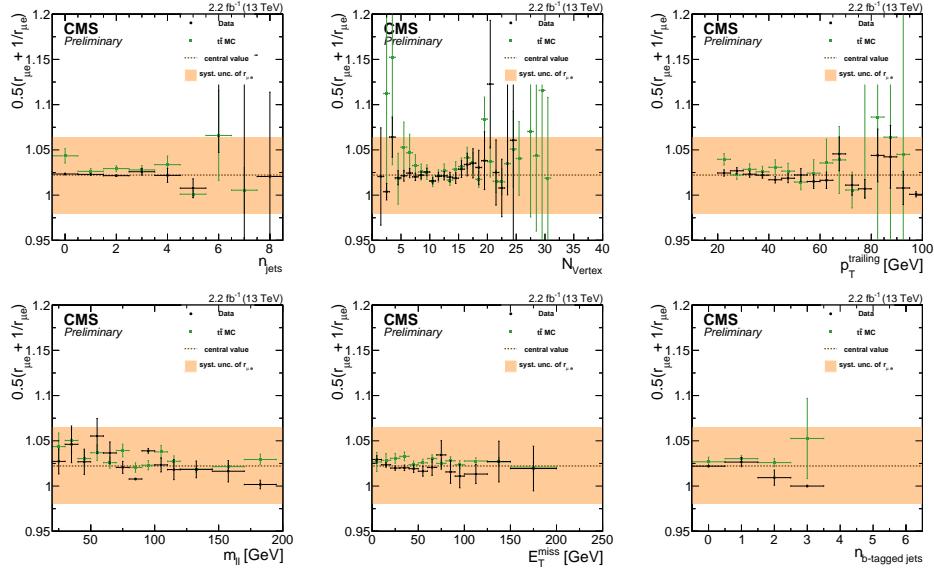
The measurement of  $R_T$  is done in the following way. First, trigger efficiencies are measured using a control sample of dilepton events, collected with the Particle Flow HT triggers with thresholds between 200 GeV and 900 GeV listed in table 4.1. The efficiency is calculated as the fraction of events in this sample that also passes the dilepton triggers for the given flavor combination as shown in equation 5.10.

$$\epsilon_{\text{trigger}} = \frac{\text{Lepton pair } \cap \text{PF}H_T \text{ trigger } \cap \text{Dilepton trigger}}{\text{Lepton pair } \cap \text{PF}H_T \text{ trigger}} \quad (5.10)$$

It is required that all events with  $N_{\text{jets}} \geq 2$  and  $E_T^{\text{miss}} > 100$  GeV are vetoed to exclude the signal region and ensure the orthogonality of the factorization method and the direct measurement of  $R_{SF/OF}$  in the control region. A minimum  $H_T$  value of 400 GeV is required to keep the PFHT triggers efficient. This is motivated in Fig. 5.7,

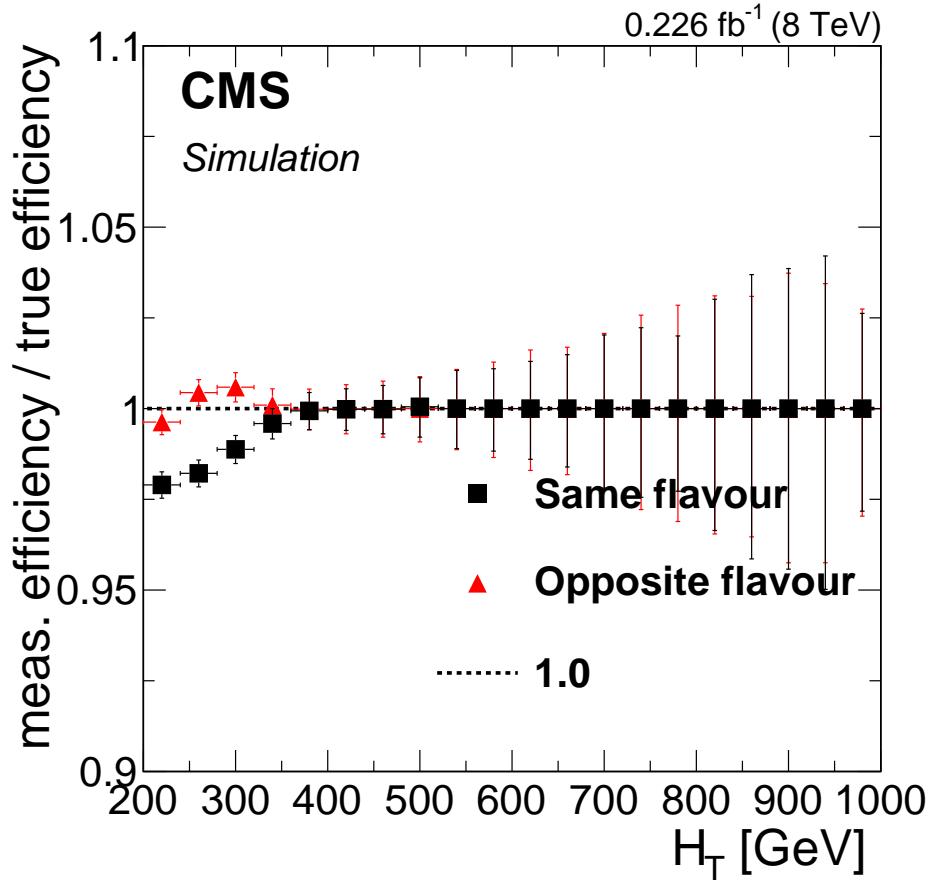


**Figure 5.5:**  $0.5(r_{\mu e} + 1/r_{\mu e})$  dependency for central leptons shown as a function of  $N_{\text{jets}}$ ,  $N_{\text{vtx}}$ , lepton  $p_{\text{T}}$ ,  $m_{\ell\ell}$ ,  $E_{\text{T}}^{\text{miss}}$ , and  $N_{\text{b-tags}}$ . The uncertainty due to the assigned 10% systematic uncertainty on  $r_{\mu/e}$  is indicated by the orange band.



**Figure 5.6:**  $0.5(r_{\mu e} + 1/r_{\mu e})$  dependency for forward leptons shown as a function of  $N_{\text{jets}}$ ,  $N_{\text{vtx}}$ , lepton  $p_{\text{T}}$ ,  $m_{\ell\ell}$ ,  $E_{\text{T}}^{\text{miss}}$ , and  $N_{\text{b-tags}}$ . The uncertainty due to the assigned 20% systematic uncertainty on  $r_{\mu/e}$  is indicated by the orange band.

where the ratio between the measured and the true efficiencies in  $t\bar{t}$  simulation are shown as a function of  $H_T$ . For low values of  $H_T$ , there is a significant deviation from unity, more evident in the case of OF triggers. The value is close to unity above  $\approx 400$  GeV, indicating the absence of bias in the trigger efficiency measurement due to the use of PF- $H_T$  triggers. It is important to note that the lowest  $H_T$  threshold included in the trigger simulation is 350 GeV. Therefore a higher offline  $H_T$  threshold is needed in the case of MC to match this trigger requirement.



**Figure 5.7:** Ratio of measured to true trigger efficiency for combined SF and OF triggers as a function of  $H_T$  measured in  $t\bar{t}$  simulation. A value of 1 indicates no bias due to the choice of the supporting triggers.

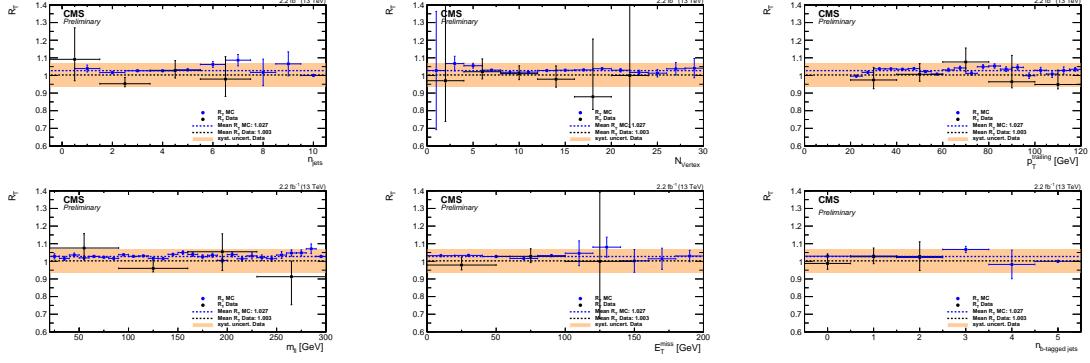
The resulting trigger efficiencies measured in data and MC are shown in

table 5.7. All the trigger efficiencies are consistent with eachother in the central region, and the  $e\mu$  trigger efficiencies are measured to be slightly lower in the forward regions. The values measured in data are limited by statistics, but the values measured are consistent with the values measured in MC.

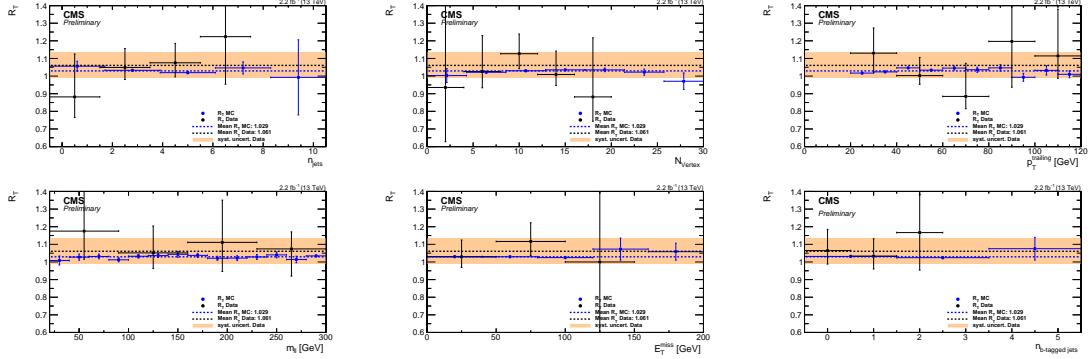
**Table 5.7:** Trigger efficiencies for data and MC with OS,  $p_T > 20(20)$  GeV and  $H_T > 400$  GeV for central and forward region seperated.

	numerator	denominator	$\epsilon_{trigger} \pm \sigma_{stat}$
Data			
Central			
ee	837	885	$0.95 \pm 0.01$
$\mu\mu$	430	462	$0.93 \pm 0.01$
$e\mu$	160	171	$0.94 \pm 0.03$
Forward			
ee	285	297	$0.96 \pm 0.02$
$\mu\mu$	226	244	$0.93 \pm 0.02$
$e\mu$	64	72	$0.89 \pm 0.05$
MC			
Central			
ee	982.3	1041.5	$0.943 \pm 0.001$
$\mu\mu$	623.0	664.1	$0.938 \pm 0.002$
$e\mu$	1763.7	1926.1	$0.916 \pm 0.001$
Forward			
ee	363.3	390.8	$0.930 \pm 0.003$
$\mu\mu$	284.7	307.2	$0.927 \pm 0.003$
$e\mu$	720.4	798.9	$0.902 \pm 0.002$

Similarly to  $r_{\mu/e}$ , the dependency of  $R_T$  on different observables is studied. The results are shown in figures 5.8 and 5.9. No significant dependency of  $R_T$  on any event property is observed, and a systematic uncertainty of 5% is assigned to each of the trigger efficiencies, which translates to an uncertainty of 6% on  $R_T$ .



**Figure 5.8:**  $R_T$  dependency for central leptons shown as a function of  $N_{\text{jets}}$ ,  $N_{\text{vtx}}$ , lepton  $p_T$ ,  $m_{ll}$ ,  $E_T^{\text{miss}}$ , and  $N_{b\text{-tags}}$ . The assigned 6.4% systematic uncertainty on  $R_T$  is indicated by the orange band.



**Figure 5.9:**  $R_T$  dependency for forward leptons shown as a function of  $N_{\text{jets}}$ ,  $N_{\text{vtx}}$ , lepton  $p_T$ ,  $m_{ll}$ ,  $E_T^{\text{miss}}$ , and  $N_{b\text{-tags}}$ . The assigned 6.4% systematic uncertainty on  $R_T$  is indicated by the orange band.

### 5.2.4 Results of factorizaton method

The result of the factorization method is summarized in table 5.8. The value and statistical uncertainty of  $R_{SF/OF}$  are both driven by  $R_T$ . Good agreement between data and MC is observed within the uncertainties.

**Table 5.8:** All the calculated values for  $r_{\mu/e}$ ,  $R_T$ , and  $R_{SF/OF}$  are shown in this table.

	Central		Forward	
	Data	MC	Data	MC
$r_{\mu/e}$	1.14±0.11	1.13±0.11	1.23±0.25	1.26±0.25
$R_T$	1.00±0.07	1.03±0.07	1.06±0.09	1.03±0.07
$R_{SF/OF}$	1.01±0.07	1.04±0.07	1.08±0.1	1.06±0.09
$R_{ee/OF}$	0.44±0.12	0.45±0.12	0.43±0.26	0.41±0.26
$R_{\mu\mu/OF}$	0.57±0.12	0.58±0.12	0.65±0.27	0.65±0.26

### 5.2.5 Combination of the Methods

The results of the two methods to determine  $R_{SF/OF}$  are shown in table 5.9.

The value for  $R_{SF/OF}$  in the central and forward regions is calculated using a weighted average of the two independent measurements. This assumes that the uncertainties are sufficiently gaussian, which is justified as they are mostly statistical in nature. The resulting values in data are  $R_{SF/OF} = 1.04 \pm 0.05$  for central and  $R_{SF/OF} = 1.10 \pm 0.07$  for forward lepton pairs. This results in an systematic uncertainty of 5% and 7% on the dominant background of the analysis in most bins. The final value uses the weighted average of  $R_{SF/OF}$  from the central and forward regions, and is calculated to be  $R_{SF/OF} = 1.05 \pm 0.04$ .

**Table 5.9:** The results of the factorization method, and from direct measurement of  $R_{SF/OF}$  are shown. The values from these two methods are combined using a weighted average to get the value for  $R_{SF/OF}$  in the central and forward regions separately. These values are then combined to get the final value of  $R_{SF/OF} = 1.05 \pm 0.04$ .

	$R_{SF/OF}$			
	Central		Forward	
	Data	MC	Data	MC
Factorization method	1.01 $\pm$ 0.07	1.04 $\pm$ 0.07	1.08 $\pm$ 0.10	1.06 $\pm$ 0.09
Direct measurement	1.06 $\pm$ 0.06	1.05 $\pm$ 0.01	1.11 $\pm$ 0.14	1.08 $\pm$ 0.02
weighted avarage	1.04 $\pm$ 0.05	1.05 $\pm$ 0.01	1.10 $\pm$ 0.07	1.08 $\pm$ 0.02
	$R_{ee/OF}$			
	Central		Forward	
	Data	MC	Data	MC
Factorization method	0.44 $\pm$ 0.12	0.45 $\pm$ 0.12	0.43 $\pm$ 0.26	0.41 $\pm$ 0.26
Direct measurement	0.43 $\pm$ 0.04	0.45 $\pm$ 0.01	0.46 $\pm$ 0.11	0.44 $\pm$ 0.01
weighted avarage	0.43 $\pm$ 0.03	0.45 $\pm$ 0.01	0.46 $\pm$ 0.05	0.44 $\pm$ 0.01
	$R_{\mu\mu/OF}$			
	Central		Forward	
	Data	MC	Data	MC
Factorization method	0.57 $\pm$ 0.12	0.58 $\pm$ 0.12	0.65 $\pm$ 0.27	0.65 $\pm$ 0.26
Direct measurement	0.63 $\pm$ 0.05	0.60 $\pm$ 0.01	0.65 $\pm$ 0.12	0.64 $\pm$ 0.02
weighted avarage	0.63 $\pm$ 0.04	0.60 $\pm$ 0.01	0.65 $\pm$ 0.06	0.64 $\pm$ 0.02

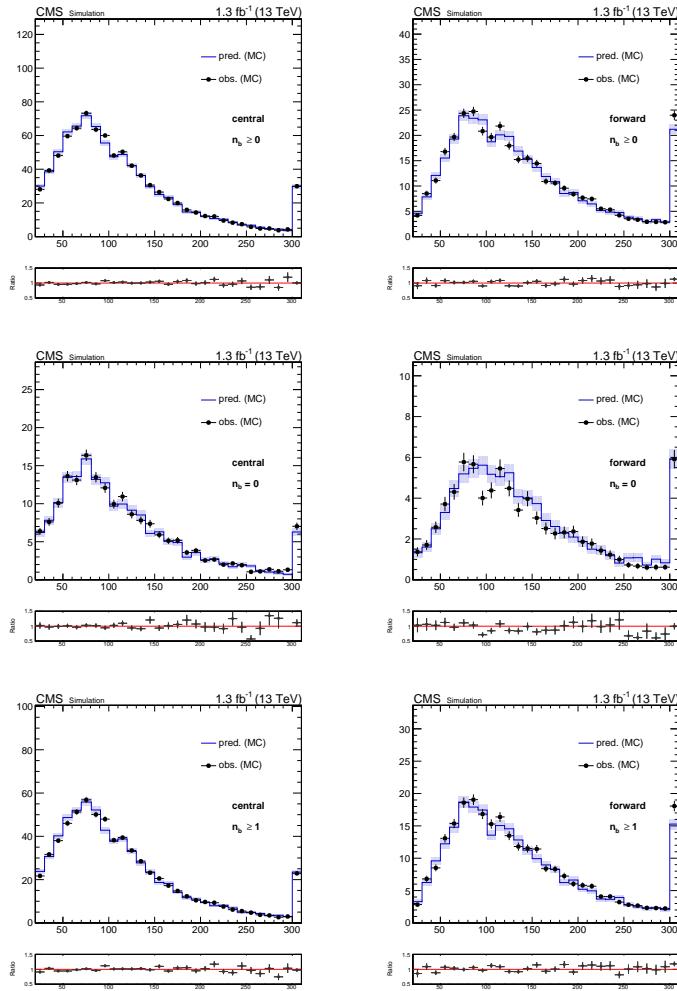
### 5.2.6 MC closure test of the FS background prediction

In order to validate the FS background prediction method, a closure test is performed in MC in the following way.  $R_{SF/OF}$ ,  $r_{\mu/e}$ , and  $R_T$  are calculated using MC simulation and subsequently used to predict the SF spectrum in  $m_{\ell\ell}$  from the OF control sample. The closure results are shown in figure 5.10 where the blue histogram shows the predicted SF spectrum obtained using the OF sample, and the black points show the same-flavor observation in MC. Good closure is seen in both central and forward regions, and the shape is seen to agree as well. It is also clear from the same figure that this method works regardless of the number of required b-tagged jets.

Results of the closure test are shown in table 5.10. The observed values of SF events are compared to the number of OF events scaled by the  $R_{SF/OF}$  value obtained using MC simulation for different processes. Only the statistical uncertainty on the event yield is shown. The dominant background is from  $t\bar{t}$  as expected and observed to be well-predicted using the FS method. Good closure is also seen in sub-dominant backgrounds such as single top quark production and  $DY + \text{jets} \rightarrow \tau\tau$ . Non FS backgrounds, such as  $DY + \text{jets} \rightarrow ee$ ,  $\mu\mu$  or rare processes involving  $Z$  boson production, are not predicted well, as expected.

**Table 5.10:** Event yields in the signal region in simulation for both SF and OF lepton pairs. The OF yield is multiplied with  $R_{SF/OF}$ . The quoted uncertainties are those of the MC counts in the signal region only.

	Central			Forward		
	SF	OF	SF-OF	SF	OF	SF-OF
$t\bar{t}$	$863 \pm 5$	$860 \pm 5$	$2 \pm 7$	$362 \pm 3$	$356 \pm 3$	$5 \pm 4$
DY + jets ( $ee, \mu\mu$ )	$19 \pm 6$	$0 \pm 0$	$18 \pm 6$	$14 \pm 4$	$0 \pm 0$	$14 \pm 4$
DY + jets ( $\tau\tau$ )	$11 \pm 6$	$20 \pm 5$	$-8 \pm 8$	$6 \pm 2$	$4 \pm 3$	$2 \pm 4$
Single t	$53 \pm 1$	$52 \pm 1$	$0 \pm 2$	$21 \pm 1$	$21 \pm 1$	$0 \pm 1$
WW, ZZ, WZ	$22 \pm 0$	$15 \pm 0$	$7 \pm 0$	$10 \pm 0$	$8 \pm 0$	$1 \pm 0$
Other SM	$9 \pm 0$	$5 \pm 0$	$3 \pm 0$	$4 \pm 0$	$3 \pm 0$	$0 \pm 0$
Total MC simulation	$980 \pm 10$	$955 \pm 7$	$24 \pm 13$	$419 \pm 6$	$394 \pm 4$	$24 \pm 8$



**Figure 5.10:** MC closure test for central (left) and forward (right) in the  $m_{\ell\ell}$  variable. The blue histogram shows the MC prediction from the OF sample by multiplying with  $R_{SF/OF}$  whereas the black markers correspond to the observation in the SF sample. This exercise is done on  $t\bar{t}$  MC only, and is shown for all number of b-tagged jet bins. The top row shows the closure test for inclusive, the middle row 0 b-tagged jets and the bottom row  $\geq 1$  jets.

## 5.3 Estimating WZ, ZZ and other rare SM backgrounds with Simulation

In this section, we consider the systematic uncertainty in the WZ and ZZ background predictions both of which are taken from MC. We do this by defining two control regions where we can study these backgrounds. To study the WZ sample, we require there be exactly 3 leptons in the event with  $p_T > 20$  GeV. We then require 2 of the 3 leptons makes an OSSF pair which has  $m_{\ell\ell}$  81 to 101 GeV. To study the ZZ sample, we require there be exactly 4 leptons in the event with  $p_T > 20$  GeV. We then require 2 of the 4 leptons makes an OSSF pair which has  $m_{\ell\ell}$  81 to 101 GeV.

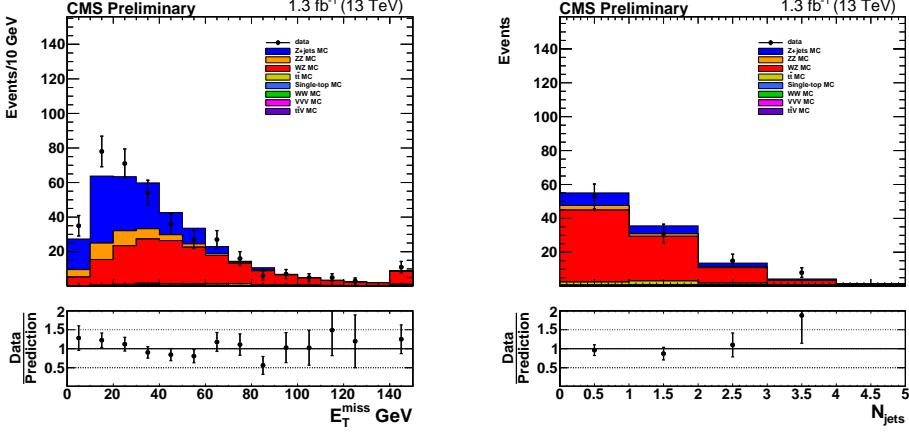
The results of this study is shown in sections 5.3 and 5.3. Because of the limited amount of data, we only show the  $E_T^{\text{miss}}$  and  $N_{\text{jets}}$  distributions.

### 3 lepton Control Region

We define a control region where we can compare the WZ MC to data by selecting events with exactly 3 leptons. The full requirements are listed below:

- exactly 3 leptons with  $p_T > 20$  GeV
- at least 2 leptons form an OSSF pair with  $m_{\ell\ell}:81\text{-}101$  GeV
- veto events with  $\geq 1$  b-tagged jet
- $E_T^{\text{miss}} > 50$  GeV

From the limited statistics we have in this region, we see good closure for the MC to predict data. We assign an systematic uncertainty of 50% to the background predictions from this MC in our signal region.



**Figure 5.11:** The  $E_T^{\text{miss}}$  and  $N_{\text{jets}}$  distributions are shown for data vs. MC in the 3-lepton control region. We require  $E_T^{\text{miss}} > 50$  GeV for the events shown in the  $N_{\text{jets}}$  distribution. See Tables 5.11 and 5.12 for yields.

#### 4 lepton Control Region

We define a control region where we can compare the ZZ MC to data by selecting events with exactly 4 leptons. The full requirements are listed below:

- exactly 4 leptons with  $p_T > 20$  GeV
- at least 2 leptons form an OSSF pair with  $m_{\ell\ell}: 81\text{-}101$  GeV

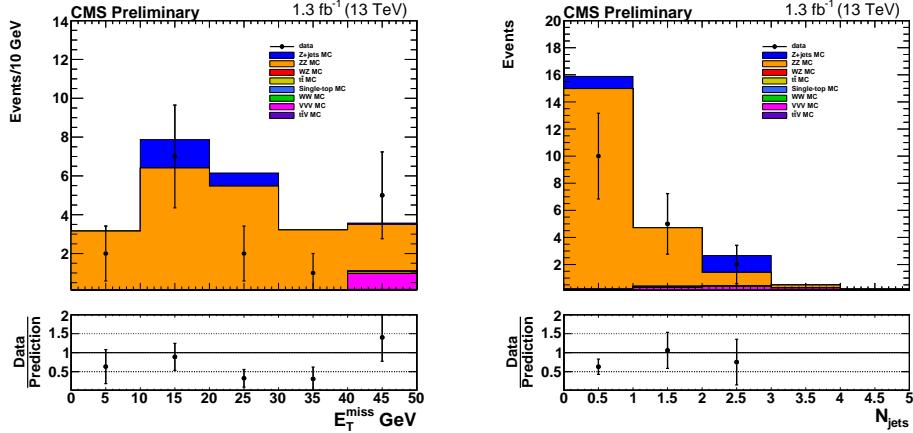
**Table 5.11:** Yields in the 3-lepton control region binned in  $E_T^{\text{miss}}$ . Uncertainties for each region are statistical only.

$E_T^{\text{miss}}[\text{GeV}]$	0 - 50	50 - 100	100 - 150	$\geq 150$
Z+jets	$127.2 \pm 13.1$	$15.0 \pm 4.0$	$0.1 \pm 0.1$	$< 0.1$
FS bkg	$2.6 \pm 0.3$	$3.8 \pm 0.4$	$1.2 \pm 0.2$	$0.4 \pm 0.1$
WZ + ZZ bkg	$125.0 \pm 0.7$	$67.5 \pm 0.5$	$11.9 \pm 0.2$	$6.5 \pm 0.2$
ttW SM BG	$0.8 \pm 0.0$	$0.7 \pm 0.0$	$0.3 \pm 0.0$	$0.3 \pm 0.0$
vvV SM BG	$0.8 \pm 0.1$	$1.1 \pm 0.1$	$0.6 \pm 0.1$	$0.3 \pm 0.0$
total BG	$256.4 \pm 13.1$	$88.0 \pm 4.1$	$14.0 \pm 0.3$	$7.4 \pm 0.2$
Data	274	83	15	9
Data/BG	$1.07 \pm 0.08$	$0.94 \pm 0.11$	$1.07 \pm 0.28$	$1.21 \pm 0.41$

**Table 5.12:** Yields in the 3-lepton control region split by  $N_{\text{jets}}$ . Uncertainties for each region are statistical only.

$N_{\text{jets}}$	0	1	2	3	$\geq 4$
Z+jets	$7.4 \pm 2.8$	$4.7 \pm 2.3$	$2.2 \pm 1.7$	$0.8 \pm 0.7$	$< 0.1$
FS bkg	$1.7 \pm 0.2$	$2.2 \pm 0.3$	$0.9 \pm 0.2$	$0.3 \pm 0.1$	$0.1 \pm 0.1$
WZ + ZZ bkg	$45.2 \pm 0.4$	$27.8 \pm 0.4$	$9.6 \pm 0.2$	$2.6 \pm 0.1$	$0.7 \pm 0.1$
ttv SM BG	$0.2 \pm 0.0$	$0.4 \pm 0.0$	$0.3 \pm 0.0$	$0.2 \pm 0.0$	$0.1 \pm 0.0$
vvv SM BG	$0.1 \pm 0.0$	$0.4 \pm 0.0$	$0.6 \pm 0.1$	$0.4 \pm 0.1$	$0.5 \pm 0.1$
total BG	$54.6 \pm 2.8$	$35.5 \pm 2.3$	$13.6 \pm 1.7$	$4.3 \pm 0.7$	$1.5 \pm 0.1$
Data	53	31	15	8	0
Data/BG	$0.97 \pm 0.14$	$0.87 \pm 0.17$	$1.10 \pm 0.31$	$1.88 \pm 0.73$	—

From the limited statistics we have in this region, we see reasonable closure for the MC to predict data. We assign an systematic uncertainty of 50% to the background predictions from this MC in our signal region.



**Figure 5.12:** The  $E_T^{\text{miss}}$  and  $N_{\text{jets}}$  distributions are shown for data vs. MC in the 4-lepton control region. See Tables 5.13 and 5.14 for yields.

**Table 5.13:** Yields in the 4-lepton control region binned in  $E_T^{\text{miss}}$ . Uncertainties for each region are statistical only.

$E_T^{\text{miss}}[\text{GeV}]$	0 - 10	10 - 20	20 - 30	30 - 40	$\geq 40$
Z+jets	$< 0.1$	$1.5 \pm 1.0$	$0.7 \pm 0.7$	$< 0.1$	$0.1 \pm 0.1$
FS bkg	$< 0.1$	$< 0.1$	$< 0.1$	$< 0.1$	$0.1 \pm 0.0$
WZ + ZZ bkg	$3.1 \pm 0.0$	$6.3 \pm 0.1$	$5.4 \pm 0.1$	$3.1 \pm 0.0$	$2.4 \pm 0.0$
ttv SM BG	$< 0.1$	$< 0.1$	$< 0.1$	$< 0.1$	$0.2 \pm 0.0$
vvv SM BG	$< 0.1$	$< 0.1$	$< 0.1$	$0.1 \pm 0.0$	$0.8 \pm 0.1$
total BG	$3.2 \pm 0.0$	$7.9 \pm 1.0$	$6.1 \pm 0.7$	$3.2 \pm 0.0$	$3.6 \pm 0.1$
Data	2	7	2	1	5
Data/BG	$0.63 \pm 0.45$	$0.89 \pm 0.36$	$0.33 \pm 0.23$	$0.31 \pm 0.31$	$1.40 \pm 0.63$

**Table 5.14:** Yields in the 4-lepton control region split by  $N_{\text{jets}}$ . Uncertainties for each region are statistical only.

$N_{\text{jets}}$	0	1	2	3	$\geq 4$
Z+jets	$0.9 \pm 0.7$	$< 0.1$	$1.2 \pm 1.0$	$< 0.1$	$0.1 \pm 0.1$
FS bkg	$< 0.1$	$0.1 \pm 0.0$	$< 0.1$	$< 0.1$	$< 0.1$
WZ + ZZ bkg	$14.8 \pm 0.1$	$4.3 \pm 0.0$	$1.0 \pm 0.0$	$0.2 \pm 0.0$	$< 0.1$
ttv SM BG	$0.1 \pm 0.0$	$0.1 \pm 0.0$	$< 0.1$	$< 0.1$	$< 0.1$
vvv SM BG	$< 0.1$	$0.2 \pm 0.0$	$0.3 \pm 0.0$	$0.3 \pm 0.0$	$0.1 \pm 0.0$
total BG	$15.9 \pm 0.7$	$4.7 \pm 0.1$	$2.7 \pm 1.0$	$0.5 \pm 0.0$	$0.2 \pm 0.1$
Data	10	5	2	0	0
Data/BG	$0.63 \pm 0.20$	$1.06 \pm 0.47$	$0.75 \pm 0.60$	–	–

# Chapter 6

## Results

The results of this search are shown in the next section. The  $E_T^{\text{miss}}$  distribution is shown for each of the five signal regions, along with a yield table that shows the predictions for each of the three major backgrounds binned in  $E_T^{\text{miss}}$  and the observed data yield. No significant excesses were seen with respect to the standard model prediction. In the final sections of this chapter, the results are interpreted in the context of the signal model described in section 1.4.

### 6.1 Results in Signal Region A

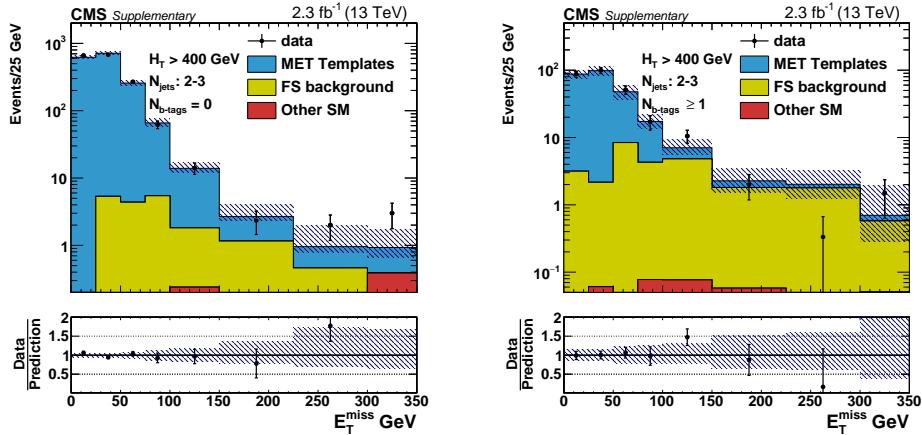
The results for signal region A are shown separately in the b-veto region, and in the signal region when requiring at least 1 b-tagged jet. The uncertainties in the tables and plots include the systematic uncertainties described in chapter 5.

**Table 6.1:** Results are shown for signal region A when requiring a b-veto. Systematic uncertainties for each region are included in the total uncertainty.

$E_T^{\text{miss}}[\text{GeV}]$	0 - 50	50 - 100	100 - 150	150 - 225	225 - 300	$\geq 300$
Z+jets	$1333.1 \pm 62.0$	$314.4 \pm 24.3$	$24.3 \pm 4.3$	$4.6 \pm 0.6$	$1.5 \pm 0.4$	$1.1 \pm 0.5$
FS bkg	$5.3^{+3.6}_{-2.3}$	$9.5^{+4.3}_{-3.1}$	$3.2^{+3.1}_{-1.7}$	$3.2^{+3.1}_{-1.7}$	$1.1^{+2.4}_{-0.9}$	$0.0^{+1.2}_{-0.0}$
Other SM	$1.7 \pm 0.8$	$2.2 \pm 1.1$	$1.6 \pm 0.8$	$1.3 \pm 0.6$	$0.8 \pm 0.4$	$1.0 \pm 0.5$
total BG	$1340.0^{+62.1}_{-62.1}$	$326.1^{+24.7}_{-24.6}$	$29.1^{+5.3}_{-4.7}$	$9.1^{+3.2}_{-1.9}$	$3.4^{+2.5}_{-1.0}$	$2.1^{+1.4}_{-0.7}$
Data	1340	332	28	7	6	6

**Table 6.2:** Results are shown for signal region A where at least 1 b-tagged jet is required. Systematic uncertainties for each region are included in the total uncertainty.

$E_T^{\text{miss}}[\text{GeV}]$	0 - 50	50 - 100	100 - 150	150 - 225	225 - 300	$\geq 300$
Z+jets	$182.6 \pm 24.1$	$53.0 \pm 11.4$	$4.5 \pm 0.9$	$1.4 \pm 0.4$	$0.7 \pm 0.3$	$0.2 \pm 0.2$
FS bkg	$5.3^{+3.6}_{-2.3}$	$12.6^{+4.8}_{-3.6}$	$9.5^{+4.3}_{-3.1}$	$5.3^{+3.6}_{-2.3}$	$5.3^{+3.6}_{-2.3}$	$1.1^{+2.4}_{-0.9}$
Other SM	$0.2 \pm 0.1$	$0.3 \pm 0.1$	$0.3 \pm 0.1$	$0.2 \pm 0.1$	$0.1 \pm 0.1$	$0.2 \pm 0.1$
total BG	$188.0^{+24.4}_{-24.2}$	$65.9^{+12.4}_{-12.0}$	$14.3^{+4.4}_{-3.2}$	$6.9^{+3.6}_{-2.3}$	$6.1^{+3.6}_{-2.3}$	$1.5^{+2.4}_{-0.9}$
Data	188	68	21	6	1	3



**Figure 6.1:** The  $E_T^{\text{miss}}$  distribution is shown for data vs. the data-driven predictions in signal region B. The left plot shows the prediction when requiring  $N_{\text{b-jets}} = 0$ , and the right plot shows the prediction when requiring  $N_{\text{b-jets}} \geq 1$ . The dashed line in the plot represents the full uncertainty. See tables 6.1 and 6.2 for yields.

## 6.2 Results in Signal Region B

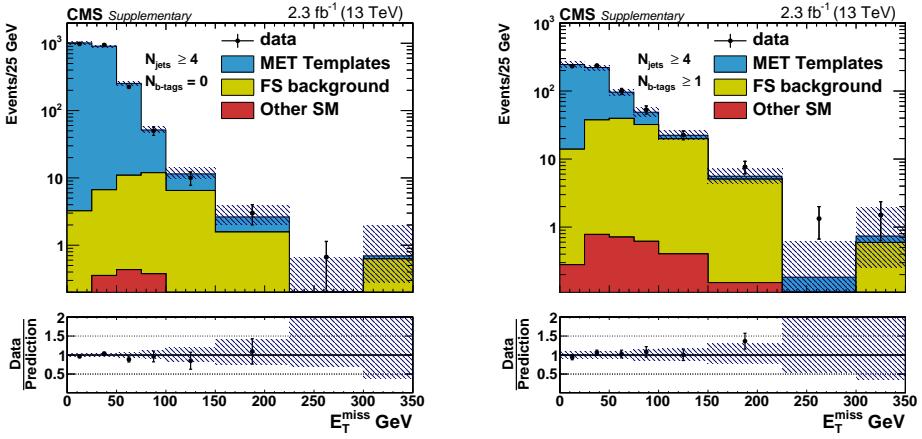
The results for signal region A are shown separately in the b-veto region, and in the signal region when requiring at least 1 b-tagged jet. The uncertainties in the tables and plots include the systematic uncertainties described in chapter 5.

**Table 6.3:** Results are shown for signal region B when requiring a b-veto. Systematic uncertainties for each region are included in the total uncertainty.

$E_T^{\text{miss}}[\text{GeV}]$	0 - 50	50 - 100	100 - 150	150 - 225	225 - 300	$\geq 300$
Z+jets	$1907.2 \pm 68.8$	$282.5 \pm 16.4$	$10.0 \pm 0.9$	$3.2 \pm 0.6$	$0.3 \pm 0.1$	$0.1 \pm 0.1$
FS bkg	$9.5^{+4.3}_{-3.1}$	$22.1^{+6.0}_{-4.9}$	$12.6^{+4.8}_{-3.6}$	$4.2^{+3.3}_{-2.0}$	$0.0^{+1.2}_{-0.0}$	$1.1^{+2.4}_{-0.9}$
Other SM	$1.4 \pm 0.6$	$2.0 \pm 0.8$	$1.0 \pm 0.4$	$0.8 \pm 0.3$	$0.5 \pm 0.2$	$0.3 \pm 0.1$
total BG	$1918.0^{+68.9}_{-68.9}$	$306.6^{+17.5}_{-17.1}$	$23.6^{+4.9}_{-3.7}$	$8.2^{+3.4}_{-2.1}$	$0.8^{+1.2}_{-0.2}$	$1.5^{+2.4}_{-0.9}$
Data	1918	275	20	10	2	0

**Table 6.4:** Results are shown for signal region B when requiring at least 1 b-tagged jet. Systematic uncertainties for each region are included in the total uncertainty.

$E_T^{\text{miss}}[\text{GeV}]$	0 - 50	50 - 100	100 - 150	150 - 225	225 - 300	$\geq 300$
Z+jets	$415.4 \pm 34.9$	$73.3 \pm 8.0$	$5.0 \pm 0.9$	$1.6 \pm 0.3$	$0.4 \pm 0.3$	$0.3 \pm 0.2$
FS bkg	$50.4^{+8.6}_{-7.5}$	$70.4^{+10.1}_{-9.0}$	$38.9^{+7.6}_{-6.5}$	$14.7^{+5.1}_{-3.9}$	$0.0^{+1.2}_{-0.0}$	$1.1^{+2.4}_{-0.9}$
Other SM	$1.2 \pm 0.5$	$1.5 \pm 0.7$	$0.8 \pm 0.4$	$0.5 \pm 0.2$	$0.2 \pm 0.1$	$0.1 \pm 0.1$
total BG	$467.0^{+35.9}_{-35.7}$	$145.2^{+12.9}_{-12.0}$	$44.7^{+7.7}_{-6.6}$	$16.8^{+5.1}_{-3.9}$	$0.6^{+1.2}_{-0.3}$	$1.5^{+2.4}_{-0.9}$
Data	467	152	45	23	4	3



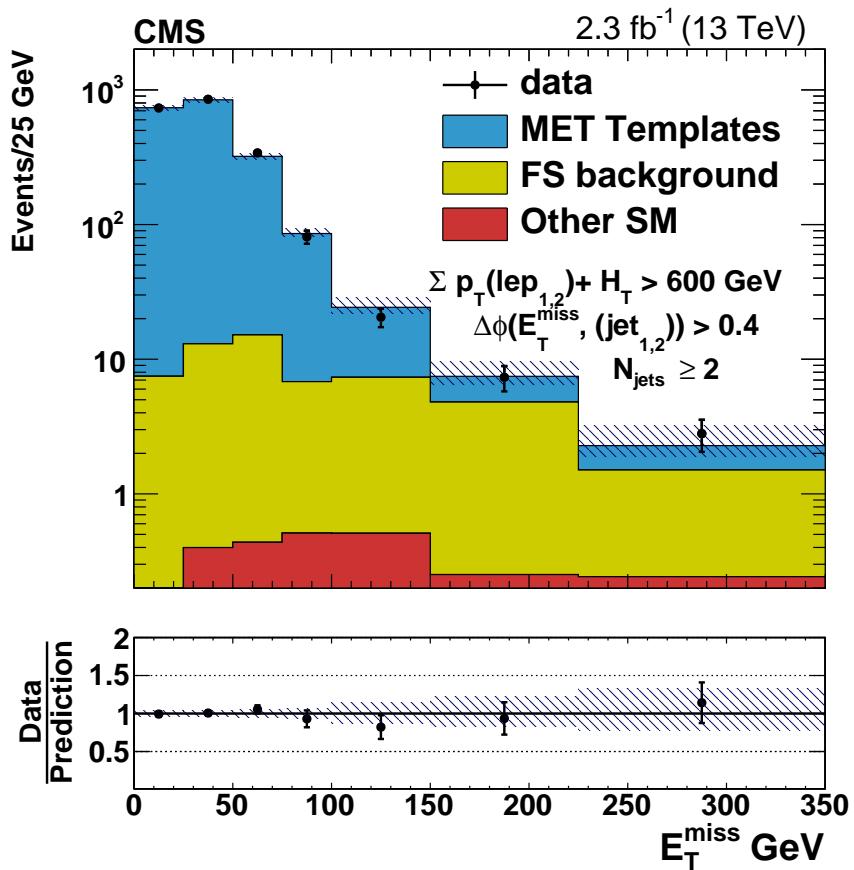
**Figure 6.2:** The  $E_T^{\text{miss}}$  distribution is shown for data vs. the data-driven predictions in signal region B. The left plot shows the prediction when requiring  $N_{b\text{-jets}} = 0$ , and the right plot shows the prediction when requiring  $N_{b\text{-jets}} \geq 1$ . The dashed line in the plot represents the full uncertainty. See tables 6.3 and 6.4 for yields.

### 6.3 ATLAS-like signal regions

The results for the ATLAS-like signal region in the on-Z search are shown in this section. The signal regions defined as  $E_T^{\text{miss}} \geq 225$  GeV. The uncertainties in the tables and plots include the systematic uncertainties derived from the MC closure test.

**Table 6.5:** Results are shown for the ATLAS signal region. Systematic uncertainties for each region are included in the total uncertainty.

$E_T^{\text{miss}}$ [GeV]	0 - 50	50 - 100	100 - 150	150 - 225	$\geq 225$
Z+jets	$1557.3 \pm 56.6$	$386.2 \pm 15.8$	$34.4 \pm 5.0$	$8.1 \pm 1.1$	$3.9 \pm 0.7$
FS bkg	$20.0^{+5.8}_{-4.6}$	$21.0^{+5.9}_{-4.7}$	$13.7^{+5.0}_{-3.8}$	$13.7^{+5.0}_{-3.8}$	$6.3^{+3.8}_{-2.5}$
Other SM	$2.8 \pm 1.2$	$3.7 \pm 1.6$	$2.3 \pm 0.9$	$1.9 \pm 0.8$	$2.1 \pm 1.0$
total BG	$1580.0^{+57.0}_{-56.8}$	$410.9^{+16.9}_{-16.5}$	$50.4^{+7.1}_{-6.3}$	$23.6^{+5.2}_{-4.0}$	$12.3^{+4.0}_{-2.8}$
Data	1580	420	41	22	14



**Figure 6.3:** The  $E_T^{\text{miss}}$  distribution is shown for data vs. the data-driven predictions in the ATLAS signal region. See Tables 6.1 and 6.5 for yields.

## 6.4 Systematic Uncertainties on Signal Model

This section summarizes the systematic uncertainties assessed for the expected signal yields. The sources of uncertainty are listed in table 6.6, and they are described in detail in the section below.

**Table 6.6:** All systematic uncertainties of the expected signal yield are shown in this table.

Source	Value (%)
Luminosity	$\pm 2.7\%$
PU reweighting	$\pm 5\%$
B-tag eff, heavy flavor	$\pm 5\%$
B-tag eff, light flavor	$\pm 2\%$
Lepton ID/Iso Efficiency	$\pm 3\%$
Fastsim Modeling	$\pm 6\%$
Lepton Trigger Efficiency	$\pm 5\%$
Jet Energy Scale	$\pm 2\text{-}5\%$
ISR Modeling	$\pm 1\%$
MC Statistics	$\pm 5\text{-}20\%$
Total uncertainty on signal	$\pm 13\text{-}24\%$

**Luminosity:** Special runs are performed during the running of the LHC in order to assess the accuracy of the measured luminosity value. These runs are named Van Der Meer scans, after the physicist who first described this method [27]. The analysis of the Van Der Meer scans gives an uncertainty value of 2.7% [28].

**PU reweighting:** The signal MC sample is reweighted such that the pileup distribution accurately reflects what is seen in data. When the nominal result is compared to what is seen when the pileup reweighting is not applied, differences of up to 5% are seen for some mass points in the plane, so an uncertainty of 5% is applied to account for these effects.

**B-tagging Efficiency:** Scale factors are applied in order to correct for the differences in efficiencies in data and MC when selecting b-tagged jets. The b-tagging scale factors are varied up and down using the uncertainties measured by the b-tagging POG at CMS [29]. The scale factors are varied according to light-flavored and heavy-flavored jets separately. The uncertainty for heavy(light)-flavored jets is 5(2)%.

**Lepton ID/isolation Efficiency:** Scale factors are applied in order to correct for the differences in efficiencies in data and MC when selecting applying lepton ID and isolation requirements. When the nominal result is compared to what is seen when the scale factors are not applied, differences of up to 3% are seen for some mass points in the plane, so an uncertainty of 3% is applied to account for these effects.

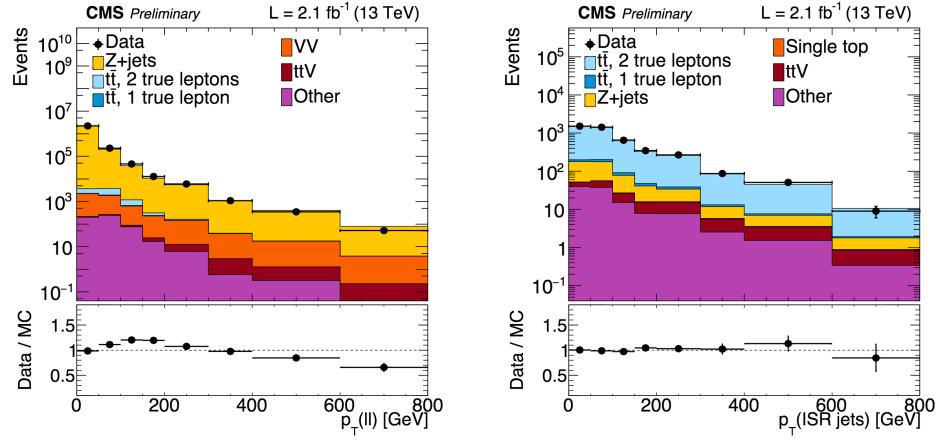
**Modeling using FastSim:** The signal MC is generated using Madgraph and Fastsim [19], and scale factors are derived to correct for the differences in efficiency when comparing to Fullsim when selecting leptons. The scale factors are measured as a function of  $p_T$  and  $\eta$  of the lepton. When the nominal result is compared to what is seen when the scale factors are not applied, differences of up to 6% are seen for some mass points in the plane, so an uncertainty of 6% is applied to account for these effects.

**Lepton Trigger Efficiency:** A flat scale factor is applied to the signal sample to account for the trigger efficiency using the efficiency measurements listed in table 5.7. An uncertainty of 5% is assessed, which is meant to cover the difference in efficiency in the turn-on curve of each trigger.

**Jet Energy Scale:** The jet energy scale is varied up and down within the uncertainties derived by the JETMET POG in CMS [30][31], and then these values are propagated through to all the objects used in the full analysis. This gives a variation of up to 2(5)% in expected signal yields in regions with  $E_T^{\text{miss}} < 300$  ( $> 300$ ) GeV.

**Modeling of Initial State Radiation:** The signals for this analysis tend to have small ISR boost, so studies were performed in inclusive Z+jets and  $t\bar{t}$  regions to test the modeling of the initial state radiation in MC. The results of this are shown in figure 6.4, and scale factors were derived as a function of ISR  $p_T$  and used to derive an uncertainty of 1% due to this effect.

**MC Statistics:** The MC statistical uncertainty is taken into account, and after applying the signal selection less signal events are seen for lower values of  $E_T^{\text{miss}}$ , and more events at high  $E_T^{\text{miss}}$ . It varies from 5-20% depending on the signal region and the point in the scan.

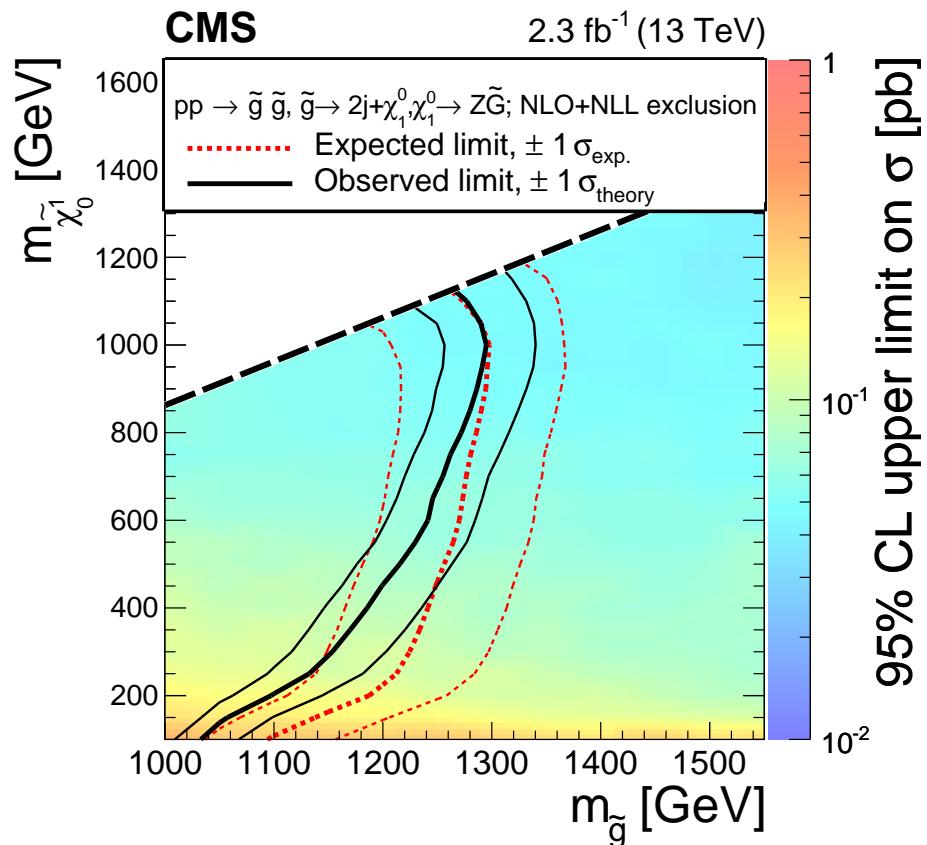


**Figure 6.4:** The ISR  $p_T$  in data and MC is shown for Z+jets (left) and  $t\bar{t}$  (right).

## 6.5 Interpretation of the Results

No significant excess is seen in any of the signal regions, so upper limits are set on the production cross section of a specific SUSY model. The model used to interpret the results of this analysis is described in section 1.4. This particular SUSY model is expected to have many jets, so most of the sensitivity comes from SRB.

The interpretation of these results is done by assigning 95% confidence level upper limits according to the CLs technique [32][33]. These limits are obtained using the higgsCombine tool [34]. The expected (observed) upper limit where this analysis is sensitive is for gluinos with mass up to 1.3 TeV when the neutralino mass is large, and when the neutralino mass is small the expected (observed) upper limit is around 1.2 (1.1) TeV. These results show a significant improvement over the 8 TeV result where we saw an observed and expected limit for gluino masses from 1 to 1.1 TeV.



**Figure 6.5:** Exclusion contours are shown when we interpret the results of this analysis in the SUSY model described in section 1.4. Everything to the left of the red (black) dotted line shows the masses which are excluded by the expected (observed) limit.

# Chapter 7

## Summary and Conclusions

In 2010, a search was performed with the first data taken at CMS [35].  $5 \text{ fb}^{-1}$  of data was taken with a center of mass energy of 7 TeV. No new physics was discovered in this dataset, and the search was performed again with the larger dataset taken in 2012 [36]. This dataset consisted of  $20 \text{ fb}^{-1}$  of data with a center of mass energy of 8 TeV. The search was also extended to be sensitive to electroweak SUSY [37, 38]. There was no new physics discovered in any of these channels. After 2012, the LHC was shut down for upgrades, and data taking resumed in 2015 with a higher center of mass energy of 13 TeV.

A search for new physics beyond the standard model was performed using data collected by the CMS detector in 2015. No significant excess beyond the standard model background predictions was observed, and upper limits were set on the cross section in the context of a SUSY model where two gluinos are pair produced, and subsequently decay to quarks and a neutralino. The neutralino then decays to a Z boson, and a gravitino in this model, where the gravitino is expected to be massless. Models where the gluino has a mass range of 1050 to 1300 GeV are excluded by these

results.

# Bibliography

- [1] CMS Collaboration, “Observation of a new boson with a mass near 125 GeV”, *CMS Physics Analysis Summary CMS-PAS-HIG-12-020* (2012).
- [2] T. A. Collaboration, “Observation of an Excess of Events in the Search for the Standard Model Higgs boson with the ATLAS detector at the LHC”, Technical Report ATLAS-CONF-2012-093, CERN, Geneva, (Jul, 2012).
- [3] J. S. G. Bertone, D. Hooper, “Particle Dark Matter: Evidence, Candidates, Constraints”, *Phys.Rept.* **405** (2004).
- [4] S. P. Martin, “A Supersymmetry primer”,. [Adv. Ser. Direct. High Energy Phys.18,1(1998)].
- [5] LHC New Physics Working Group Collaboration, “Simplified models for LHC new physics searches”, *J. Phys. G* **39** (2012) 105005. Model of Section IV.E with topology (B+B).
- [6] L. Evans and P. Bryant, “LHC Machine”, *Journal of Instrumentation* **3** (2008), no. 08, S08001.
- [7] CMS Collaboration, “The CMS experiment at the CERN LHC”, *JINST* **3** **S08004** (2008) 361.
- [8] CERN Collaboration, “CERN Academic Training Lectures 2016: Calorimetry by Examples”,. <https://indico.cern.ch/event/526766>.
- [9] CMS HCAL Collaboration, “Design, performance, and calibration of CMS hadron-barrel calorimeter wedges”, *Eur. Phys. J.* **C55** (2008) 159–171.
- [10] CMS Collaboration, “Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET”, *CMS Physics Analysis Summary CMS-PAS-PFT-09-001* (2009).
- [11] M. Cacciari, G. P. Salam, and G. Soyez, “The anti-k<sub>t</sub> jet clustering algorithm”, *Journal of High Energy Physics* **4** (April, 2008) 63–.

- [12] C. Collaboration, “CMS JEC Run I legacy performance plots”,.
- [13] GEANT4 Collaboration, “GEANT4—a simulation toolkit”, *Nucl. Instrum. Meth. A* **506** (2003) 250.
- [14] S. Frixione, P. Nason, and C. Oleari, “Matching NLO QCD computations with parton shower simulations: the POWHEG method”, *JHEP* **11** (2007) 070.
- [15] J. Alwall, R. Frederix, S. Frixione et al., “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”, *JHEP* **07** (2014) 079.
- [16] T. Sjöstrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 physics and manual”, *JHEP* **05** (2006) 026.
- [17] Y. Li and F. Petriello, “Combining QCD and electroweak corrections to dilepton production in FEWZ”, *Phys. Rev. D* **86** (2012) 094034.
- [18] CMS Collaboration, “Gluino-gluino Cross Sections at 13 TeV”, (2013). twiki.cern.ch/twiki/bin/view/LHCPhysics/SUSYCrossSections13TeVgluglu.
- [19] S. Abdullin, P. Azzi, F. Beaudette et al., “The fast simulation of the CMS detector at LHC”, *J. Phys. Conf. Ser.* **331** (2011) 032049.
- [20] CMS Collaboration, “Electron and Photon performance using data collected by CMS at  $s = \sqrt{13}$  TeV and 25ns”, *CMS-DP-2015-067* (Dec, 2015).
- [21] CMS Collaboration, “Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at  $\sqrt{s} = 8$  TeV”, *Journal of Instrumentation* **10** (2015), no. 6, P06005.
- [22] G. Bayatian et al., “CMS Physics Technical Design Report Volume I: Detector Performance and Software”. Technical Design Report CMS. CERN, Geneva, 2006.
- [23] CMS Collaboration, “Performance of CMS muon reconstruction in cosmic-ray events”, *Journal of Instrumentation* **5** (March, 2010) 3022–+.
- [24] CMS Collaboration, “Identification of b-quark jets with the CMS Experiment”, *Journal of Instrumentation* **8** (2013), no. 13, P04013.
- [25] CMS Collaboration, “Identification of b quark jets at the CMS Experiment in the LHC Run 2”, *CMS Physics Analysis Summary CMS-PAS-BTV-15-001* (2015).
- [26] ATLAS Collaboration, “Search for supersymmetry in events containing a same-flavour opposite-sign dilepton pair, jets, and large missing transverse momentum in  $\sqrt{s} = 8\text{TeV}$  pp collisions with the ATLAS detector”, *The European Physical Journal C* **75** (2015), no. 7,.

- [27] S. van der Meer, “Calibration of the effective beam height in the ISR. oai:cds.cern.ch:296752”, Technical Report CERN-ISR-PO-68-31. ISR-PO-68-31, CERN, Geneva, (1968).
- [28] CMS Collaboration, “CMS Luminosity Measurement for the 2015 Data Taking Period”, *CMS Physics Analysis Summary* (2015).
- [29] CMS Collaboration, “Identification of b-quark Jets at the CMS Experiment in the LHC Run2 Startup”, *CMS Physics Analysis Summary* (2015).
- [30] CMS Collaboration, “Determination of jet energy calibration and transverse momentum resolution in CMS”, *Journal of Instrumentation* **6** (2011), no. 11, P11002.
- [31] CMS Collaboration, “Jet Energy Corrections: Official Software Tools for applying JEC Corrections and Uncertainties”, (2015). twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookJetEnergyCorrections.
- [32] A. L. Read, “Presentation of search results: The  $CL_s$  technique”, *J. Phys. G* **28** (2002) 2693.
- [33] T. Junk, “Confidence level computation for combining searches with small statistics”, *Nucl. Instrum. Meth. A* **434** (1999) 435.
- [34] ATLAS and CMS Collaborations Collaboration, “Procedure for the LHC Higgs boson search combination in summer 2011”, ATL-PHYS-PUB-2011-011, CMS NOTE-2011/005, CERN, (2011).
- [35] CMS Collaboration, “Search for physics beyond the standard model in events with a Z boson, jets, and missing transverse energy in pp collisions at  $\sqrt{s} = 7$  TeV”, *Physics Letters B* **716** (2012), no. 2, 260 – 284.
- [36] CMS Collaboration, “Search for physics beyond the standard model in events with two leptons, jets, and missing transverse momentum in pp collisions at  $\sqrt{s} = 8$  TeV”, *Journal of High Energy Physics* **2015** (2015), no. 4, 1–41.
- [37] CMS Collaboration, “Searches for electroweak production of charginos, neutralinos, and sleptons decaying to leptons and W, Z, and Higgs bosons in pp collisions at 8 TeV”, *The European Physical Journal C* **74** (2014), no. 9, 1–42.
- [38] CMS Collaboration, “Searches for electroweak neutralino and chargino production in channels with Higgs, Z, and W bosons in pp collisions at 8 TeV”, *Phys. Rev. D* **90** (Nov, 2014) 092007.