

UNIVERSITY OF CALIFORNIA, SAN DIEGO

**Searches for New Physics in Final States With Two Opposite-Sign Same-Flavor
Leptons, Jets, and Missing Transverse Energy in Proton-Proton Collisions at Center
of Mass Energies of 8 and 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 micro-
film and electronically:

Chair

University of California, San Diego

2016

DEDICATION

Add dedication

EPIGRAPH

something

translation

—source

TABLE OF CONTENTS

Signature Page	iii
Dedication	iv
Epigraph	v
Table of Contents	vi
List of Figures	ix
List of Tables	xiii
Acknowledgements	xv
Vita	xvi
Abstract of the Dissertation	xvii
Chapter 1 Introduction	1
1.1 Particle Physics	1
1.2 The Standard Model	2
1.2.1 Problems with The Standard Model	3
1.3 Supersymmetry (SUSY)	3
1.3.1 Gauge-Mediated Supersymmetry Breaking (GMSB)	4
1.4 Searches for SUSY in final states with Z bosons	4
Chapter 2 The LHC Accelerator and CMS Experiment	6
2.1 The Large Hadron Collider	6
2.2 The Compact Muon Solenoid	7
2.2.1 Silicon Tracker	9
2.2.2 Electromagnetic Calorimeter	10
2.2.3 Hadronic Calorimeter	13
2.2.4 Muon System	14
2.3 Physics objects	17
2.3.1 Particle Flow	17
2.3.2 Vertex Determination and Pileup	18
2.3.3 Isolation	19
2.3.4 leptons and photons	19
2.3.5 jets	20

Chapter 3	Missing Transverse Momentum	22
	3.1 raw MET	22
	3.2 Corrections to MET	23
	3.3 Type 1 MET	23
	3.3.1 Data vs. MC Comparison	24
Chapter 4	Analysis Selections	29
	4.1 Triggers and Datasets	29
	4.2 Simulations from Monte-Carlo	32
	4.3 Event Selection	34
	4.3.1 Vertex Selection	34
	4.3.2 MET filters	34
	4.3.3 Lepton Selection	35
	4.3.4 Lepton Isolation	36
	4.3.5 Lepton Scale Factors	40
	4.3.6 Electron Selection	40
	4.3.7 Muon Selection	41
	4.3.8 Photons	43
	4.3.9 MET	45
	4.3.10 H_T	45
	4.3.11 Jets	45
	4.3.12 Signal Region Definitions	47
	4.3.13 Selection Summary	48
Chapter 5	Background Estimation Methods	50
	5.1 Estimating the Z + jets Background with E_T^{miss} Template	50
	5.1.1 Systematic Uncertainties in the E_T^{miss} Template Prediction	54
	5.2 Estimating the Flavor-Symmetric Background	57
	5.2.1 $R_{SF/OF}$	58
	5.2.2 Direct measurement of $R_{SF/OF}$	61
	5.2.3 Measuring $R_{SF/OF}$ using the Factorization Method	62
	5.2.4 Results of factorization method	68
	5.2.5 Combination of the Methods	69
	5.2.6 MC closure test of the FS background prediction	71
	5.3 Estimating WZ, ZZ and other rare SM backgrounds with Simulation	74
Chapter 6	Results	78
	6.1 Results in Signal Region A	78
	6.2 Results in Signal Region B	80
	6.3 ATLAS-like signal regions	82
	6.4 Systematic Uncertainties on Signal Model	84
	6.5 Interpretation of the Results	87

Chapter 7	Summary and Conclusions	89
Bibliography	90

LIST OF FIGURES

Figure 1.1:	A feynman diagram for the T5ZZgmsb modeled is shown here. In this process, 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.	5
Figure 2.1:	The LHC can be seen along with the 4 main experiments at different points along the beamline.	7
Figure 2.2:	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$	10
Figure 2.3:	A cross-sectional view of the CMS ECAL is shown in this figure, with values of η shown that determine the coverage of each subdetector. . .	11
Figure 2.4:	A feynman diagram is shown depicting electron-positron pair production from a photon initial state.	12
Figure 2.5:	A feynman diagram is shown where an electron scatters off of an incoming photon, and then radiates a photon through bremsstrahlung. . .	12
Figure 2.6:	A cartoon depiction of a drift tube is shown in this figure.	15
Figure 2.7:	A cartoon depiction of a cathode strip chamber is shown in this figure.	16
Figure 2.8:	A cartoon depiction of a resistive plate chamber is shown in this figure.	16
Figure 2.9:	Stopping power of copper as a function of muon momentum. Typical momenta of muons in this analysis are between 20-200 GeV.	20
Figure 3.1:	The E_T^{miss} distribution is shown with data vs. MC for events with $Z \rightarrow \ell\ell$. 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.	25
Figure 3.2:	The E_T^{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.	26
Figure 3.3:	The E_T^{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.	27

Figure 3.4:	The E_T^{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.	28
Figure 4.1:	The isolation distribution is shown for electrons on the left in simulated $Z + \text{jets}$ MC events. On the right, the efficiency is shown when integrating all bins to the left of the value on the x-axis.	39
Figure 4.2:	The isolation distribution is shown for muons on the left in simulated $Z + \text{jets}$ MC events. On the right, the efficiency is shown when integrating all bins to the left of the value on the x-axis.	39
Figure 4.3:	Electron scale factors as a function of p_T and η . The scale factor is applied as an event weight once per electron in the event.	41
Figure 4.4:	Muon scale factors as a function of p_T and η . 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.	41
Figure 4.5:	Fastsim scale factors as a function of p_T and η . 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.	42
Figure 4.6:	The distribution showing the CSVv2 Discriminator is shown comparing data vs. MC. The MC is split by jet parton flavor.	47
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.	52
Figure 5.2:	The E_T^{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.	53
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^{miss} for simulated $Z + \text{jets}$ events. The distributions are normalized to 1 fb^{-1} of data.	55
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 $R_{SF/OF}$	62

Figure 5.5:	$0.5(r_{\mu e} + 1/r_{\mu e})$ dependency for central leptons shown as a function of N_{jets} , N_{vtx} , lepton p_T , $m_{\ell\ell}$, E_T^{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.	64
Figure 5.6:	$0.5(r_{\mu e} + 1/r_{\mu e})$ dependency for forward leptons shown as a function of N_{jets} , N_{vtx} , lepton p_T , $m_{\ell\ell}$, E_T^{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.	65
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.	66
Figure 5.8:	R_T dependency for central leptons shown as a function of N_{jets} , N_{vtx} , lepton p_T , $m_{\ell\ell}$, E_T^{miss} , and $N_{b\text{-tags}}$. The assigned 6.4% systematic uncertainty on R_T is indicated by the orange band.	68
Figure 5.9:	R_T dependency for forward leptons shown as a function of N_{jets} , N_{vtx} , lepton p_T , $m_{\ell\ell}$, E_T^{miss} , and $N_{b\text{-tags}}$. The assigned 6.4% systematic uncertainty on R_T is indicated by the orange band.	68
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 $R_{\text{SF/OF}}$ 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 ≥ 1 jets.	73
Figure 5.11:	The E_T^{miss} and N_{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_{jets} distribution. See Tables 5.11 and 5.12 for yields.	75
Figure 5.12:	The E_T^{miss} and N_{jets} distributions are shown for data vs. MC in the 4-lepton control region. See Tables 5.13 and 5.14 for yields.	76
Figure 6.1:	The E_T^{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 6.1 and 6.2 for yields.	79
Figure 6.2:	The E_T^{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.	81
Figure 6.3:	The E_T^{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.	83
Figure 6.4:	The ISR p_T in data and MC is shown for Z + jets (left) and $t\bar{t}$ (right).	86

Figure 6.5: Exclusion contours are shown when we interpret the results of this analysis in the T5ZZ model. Everything to the left of the red (black) dotted line shows the masses which are excluded by the expected (observed) limit. 88

LIST OF TABLES

Table 3.1:	Definitions for the η region chosen for each subdetector.	24
Table 4.1:	List of all triggers used in this analysis.	31
Table 4.2:	List of MC samples.	33
Table 4.3:	Effective area values for muons derived in separate regions of η	37
Table 4.4:	Effective area values for electrons derived in separate regions of η	38
Table 4.5:	Requirements for electrons with $p_T > 10$ GeV.	42
Table 4.6:	Summary of the muons selection requirements.	43
Table 4.7:	Summary of the photon isolation requirements.	44
Table 4.8:	Summary of the photon selection requirements.	45
Table 4.9:	Criteria for a jet to pass the loose ID requirement.	46
Table 4.10:	Summary of the jet selection requirements.	46
Table 4.11:	List of all selections used to define the preselection and signal regions.	49
Table 5.1:	List of single photon triggers used in this analysis and their prescales.	51
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 discrepancy between the γ +jets prediction of the Z+jets MC for each E_T^{miss} region, or the the statistical uncertainty whichever is larger. All uncertainties shown are statistical only.	56
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 discrepancy between the γ +jets prediction of the Z+jets MC for each E_T^{miss} region, or the statistical uncertainty whichever is larger.	56
Table 5.4:	Systematic uncertainties on the normalization procedure are listed in the following table.	57
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.	63
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.	64
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.	67
Table 5.8:	The values for $r_{\mu/e}$, R_T , and $R_{SF/OF}$ are shown.	69

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 $R_{SF/OF} = 1.05 \pm 0.04$	70
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.	72
Table 5.11:	Yields in the 3-lepton control region binned in E_T^{miss} . Uncertainties for each region are statistical only.	75
Table 5.12:	Yields in the 3-lepton control region split by N_{jets} . Uncertainties for each region are statistical only.	76
Table 5.13:	Yields in the 4-lepton control region binned in E_T^{miss} . Uncertainties for each region are statistical only.	77
Table 5.14:	Yields in the 4-lepton control region split by N_{jets} . Uncertainties for each region are statistical only.	77
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.	79
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.	79
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.	80
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.	80
Table 6.5:	Results are shown for the ATLAS signal region. Systematic uncertainties for each region are included in the total uncertainty.	82
Table 6.6:	Systematic uncertainties of the expected signal yield.	84

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

**Searches for New Physics in Final States With Two Opposite-Sign Same-Flavor
Leptons, Jets, and Missing Transverse Energy in Proton-Proton Collisions at Center
of Mass Energies of 8 and 13 TeV**

by

Charles Vincent Welke

Doctor of Philosophy in Physics

University of California, San Diego, 2016

Professor Avraham Yagil, Chair

A search is presented for physics beyond the standard model in final states with two opposite-sign same-flavor leptons, jets, and missing transverse momentum. The data sample corresponds to an integrated luminosity of 2.2fb^{-1} of proton-proton collisions at $\sqrt{s} = 13\text{TeV}$ collected with the CMS detector at the CERN LHC in 2015. The analysis focuses on the invariant mass distribution of the lepton pair, searching for a kinematic edge or a resonant-like excess compatible with the Z boson mass. The kinematic edge search includes phase-space regions matching the previous 8 TeV analysis where CMS reported

a 2.6σ excess. On the resonant Z boson peak, a region has been included to match one where ATLAS reported a 3.0σ excess at 8 TeV. Additional regions have been included in both searches beyond those in the 8 TeV analysis to increase sensitivity to new physics. The observations in all signal regions are consistent with the expectations from the standard model, 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 doesn't describe everything however, for example gravity cannot be explained by the standard model. The discovery of dark matter and dark energy in the universe is another example of something that is not explained by the standard model. There are many theories that have been proposed which can explain both of these phenomena. Supersymmetry is an extension of the standard model which can be used to fill in the gaps and explain both gravity and dark energy/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 uncharged. 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 for the electric charge is the negative 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 and ATLAS experiments. This particle has properties that make it look like a Higgs boson. The Higgs boson is responsible for particles to have mass.

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. The universe is composed of 5% visible matter, explained by the standard model, and the rest of the matter in the universe is in the form of dark matter and dark energy. A second example of where the standard model falls short is the hierarchy problem.

(FROM WIKIPEDIA) In particle physics, the most important hierarchy problem is the question that asks why the weak force is 10^{32} times stronger than gravity. Both of these forces involve constants of nature, Fermi's constant for the weak force and Newton's constant for gravity. Furthermore if the Standard Model is used to calculate the quantum corrections to Fermi's constant, it appears that Fermi's constant is surprisingly large and is expected to be closer to Newton's constant, unless there is a delicate cancellation between the bare value of Fermi's constant and the quantum corrections to it. (END FROM WIKIPEDIA)

One of the main goals of particle physics is to understand these shortcomings of the standard model. Theorists attempt to do this in many ways and one of these ways is using supersymmetry.

1.3 Supersymmetry (SUSY)

SUSY 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 is fermionic, and the partners to fermions are bosonic. These particles are named sparticles, and for fermions, an s is added to the beginning of the standard model particle name to rep-

represent the SUSY particle. 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. Using this idea, one can create a model that can answer both the hierarchy problem as well as explain the existence of dark matter.

1.3.1 Gauge-Mediated Supersymmetry Breaking (GMSB)

Gauge-Mediated Supersymmetry Breaking (GMSB) is a mechanism that allows supersymmetry to be broken. This mechanism postulates the existence of a particle that can be used to explain gravity, the graviton (\tilde{G}). Within SUSY, the hierarchy problem can be solved with one loop corrections coming from heavy stop quarks, and another consequence of GMSB is the proposal that stop quarks must have a mass greater than 2 TeV if the Higgs mass is 125 GeV. The current upper limits for excluding the observation of stop quarks set by CMS are around 650 GeV. (ADD STOP LIMIT PLOT)

1.4 Searches for SUSY in final states with Z bosons

Many models can be produced within the context of SUSY leading to a nearly endless variety of final states. There are many variables 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. One of these models nicknamed T5ZZgmsb, 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 as seen in figure 1.1. 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.

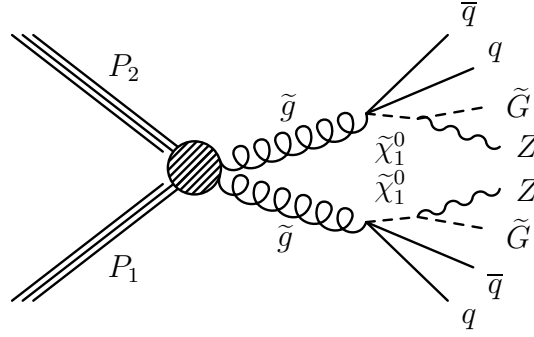


Figure 1.1: A feynman diagram for the T5ZZgmsb modeled is shown here. In this process, two protons collide and the result is pair-production of two gluinos, where each decays to a pair of quarks and a neutralino which subsequently decays to a Z boson and gravitino.

Chapter 2

The LHC Accelerator and CMS

Experiment

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the largest particle accelerator in the world. 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, where 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.

In order to achieve such large center of mass energies, the protons are accelerated in a circular beam pipe which is 27 km long and 100 m underground. 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 is where the data was taken used for the results shown in this thesis.

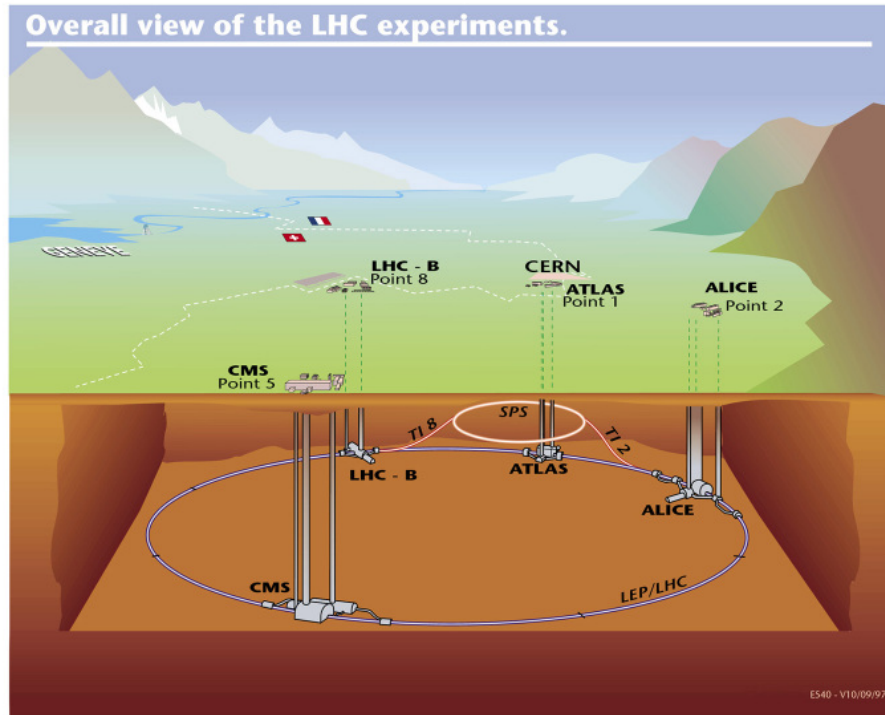


Figure 2.1: The LHC can be seen along with the 4 main experiments at different points along the beamline.

2.2 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 volume 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 for every set number of events. As the prescale 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 2.2.1) which is used to measure the charge and momentum of charged particles. When a charged particle passes through the silicon layers in the pixel detector, 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. Neutral particles do not interact with the tracker.

The next layer is the Electromagnetic Calorimeter (described in section 2.2.2) which is used to contain and measure the energy of electrons and photons. The Electromagnetic Calorimeter (ECAL) is made of lead tungstate (PbWO_4) scintillating crystals. The material has a radiation length of 0.89 cm and Molière radius ($R_M \simeq 2.19$ cm). 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 2.2.3) 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 2.1, where q is the particle charge, r is the radius of curvature, and B is the magnetic field magnitude.

$$p = qrB \quad (2.1)$$

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^{miss}) or MET.

Lastly, the muon system (described in section 2.2.4) 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 components.

2.2.1 Silicon Tracker

figure 2.2

Something about pixels and other stuff.

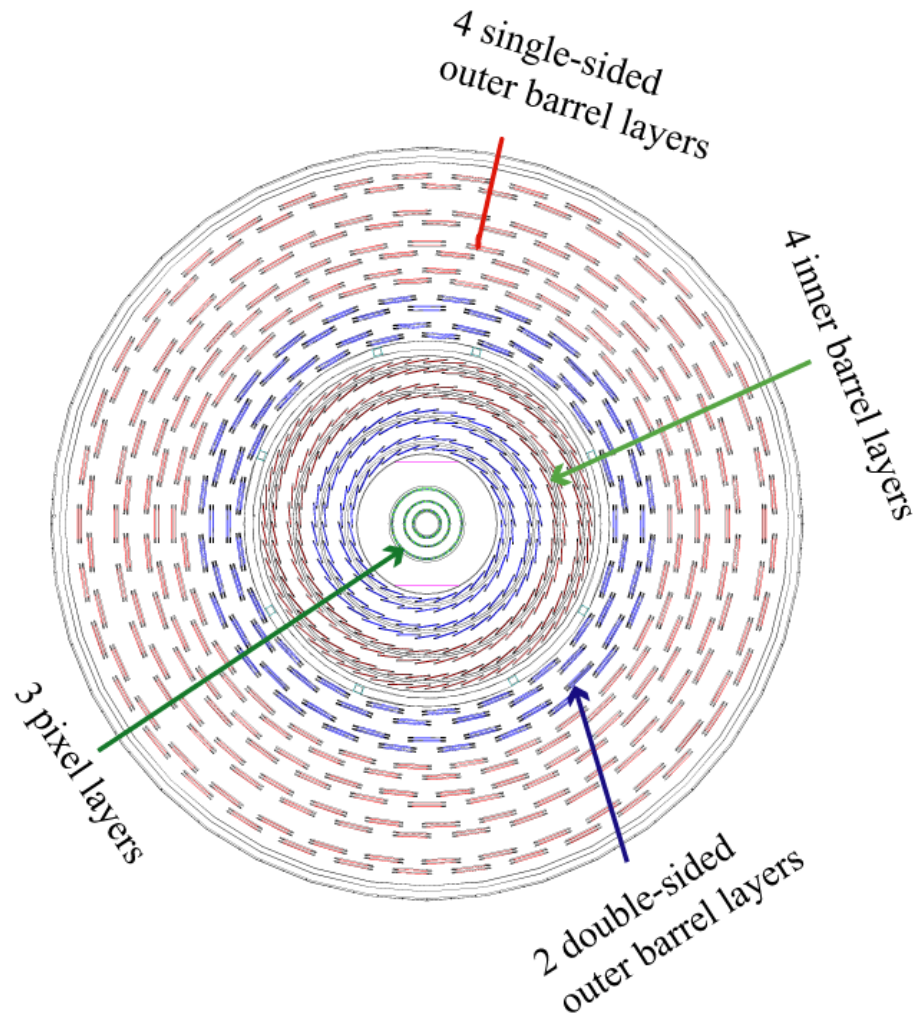


Figure 2.2: A cross-sectional in the x-y plane of the CMS tracking system is shown here. The tracker provides excellent coverage for reconstruction of tracks from charged particles for $|\eta| < 2.5$.

2.2.2 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 2.3.

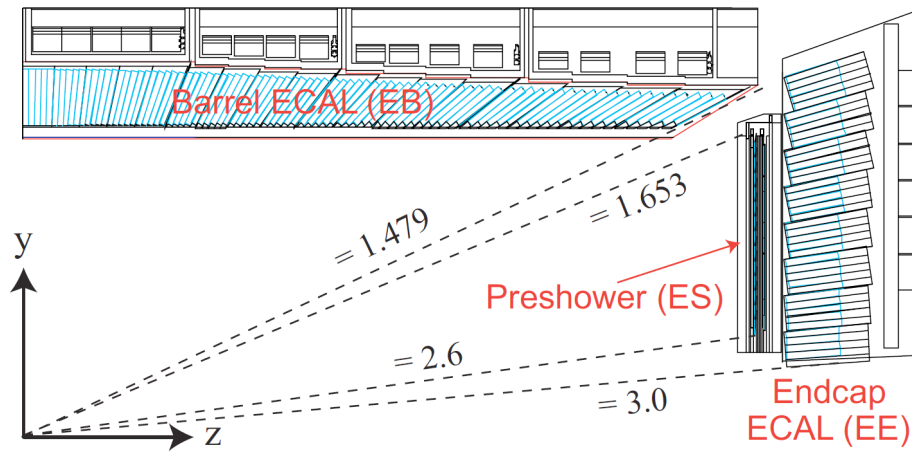


Figure 2.3: A cross-sectional view of the CMS ECAL is shown in this figure, with values of η shown that determine the coverage of each subdetector.

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 electromagnetic scattering, and electron-positron pair production. Leading order diagrams of these processes are shown in figures 2.4 and 2.5.

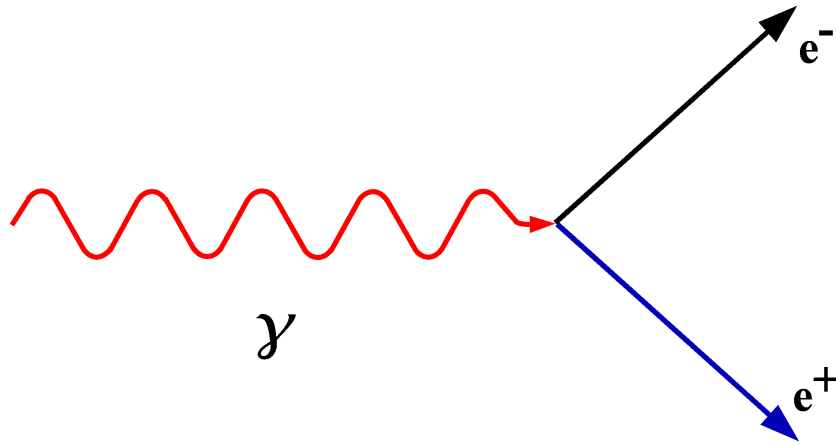


Figure 2.4: A feynman diagram is shown depicting electron-positron pair production from a photon initial state.

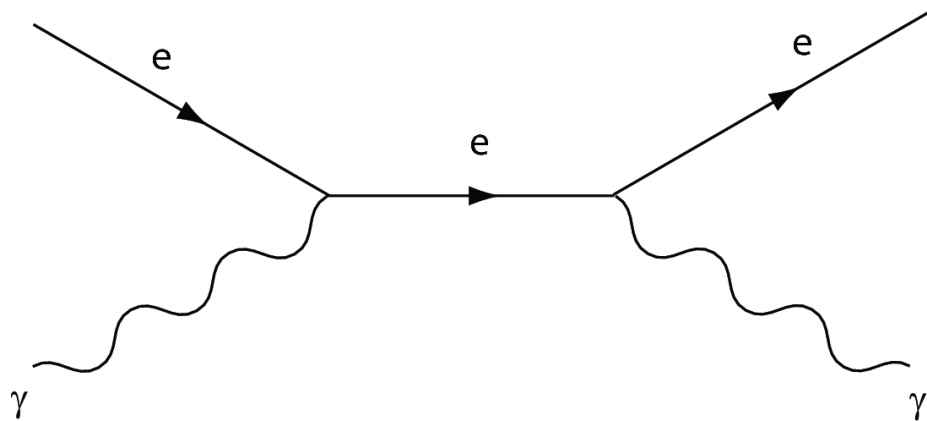


Figure 2.5: A feynman diagram is shown where an electron scatters off of an incoming photon, and then radiates a photon through bremsstrahlung.

As the initial particle passes through the crystal, it loses energy due to these processes, and this energy is converted to light by the scintillating 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.

2.2.3 Hadronic Calorimeter

The Hadronic Calorimeter (HCAL) is used to measure energies of hadronic particles, such as pions, gluons, protons, kaons and neutrons. 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 scintillating material and then measured by Hybrid-Photo Diodes (HPDs).

Approximately a third of the particles produced in a hadronic shower are π_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 particle's 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 (λ) in the HCAL starts at about $\lambda = 7$ at $\eta = 0$, and increases with eta to the end of the barrel up to $\lambda = 11$. There is then a slight decrease in the endcap of $\lambda = 10$.

2.2.4 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 completely outside of the solenoid. The information from these systems is used sometimes for triggering on events as well as help to identify muons and measure the muon 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 2.6. The DTs are individually filled with a mixture of Ar and CO₂ 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, 2.7.

There are 610 resistive plate chambers (RPCs), an example of which is shown in

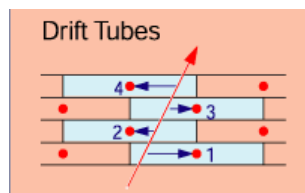


Figure 2.6: A cartoon depiction of a drift tube is shown in this figure.

2.8.

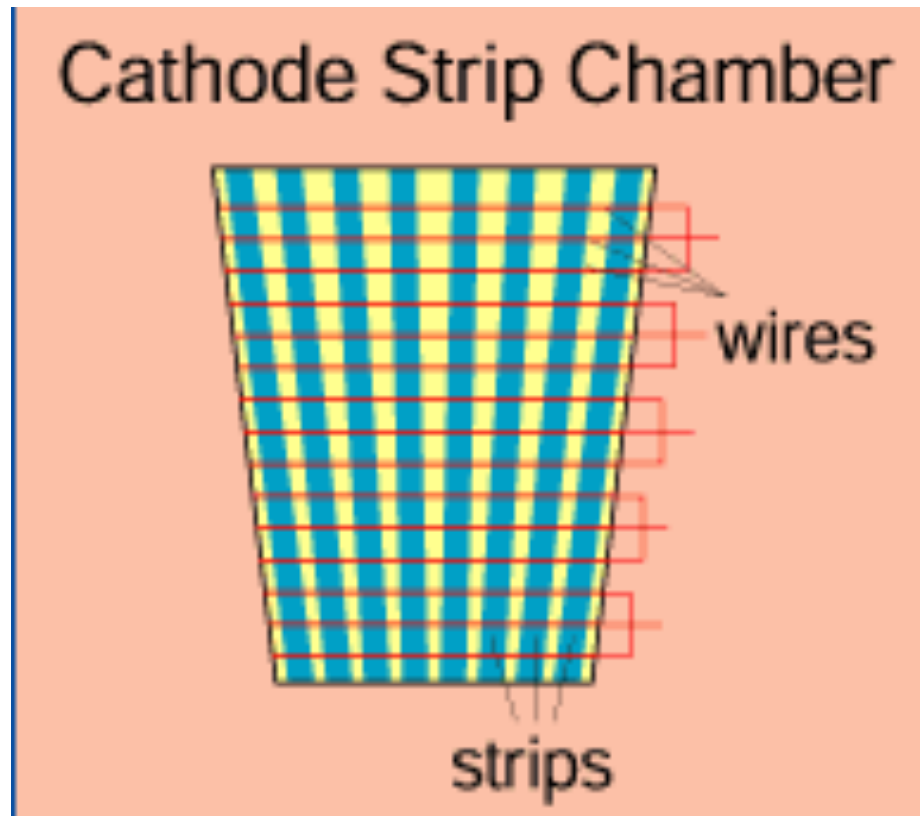


Figure 2.7: A cartoon depiction of a cathode strip chamber is shown in this figure.

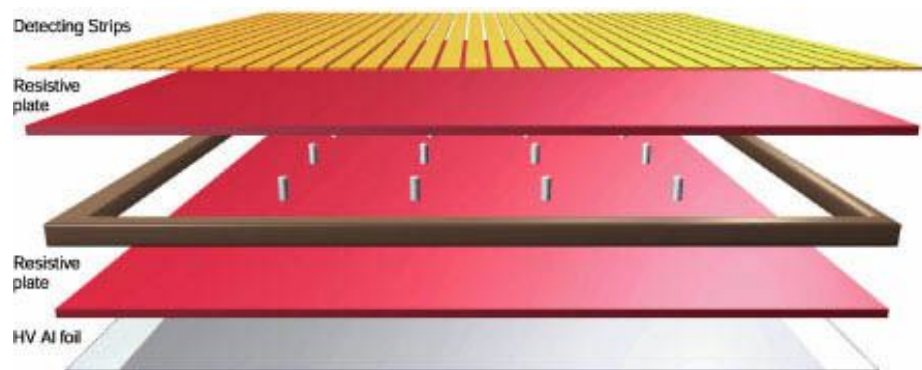


Figure 2.8: A cartoon depiction of a resistive plate chamber is shown in this figure.

2.3 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.

2.3.1 Particle Flow

In order to classify objects, an algorithm named particle flow (PF) is used [1]. The main goal when using the pf algorithm is to account for all energies measured in the detector. The way this is done is by clustering energies 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 energies are accounted for. 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

2.3.2 Vertex Determination and Pileup

The LHC collides protons in large bunches to maximize the probability of a hard collision. Multiple pairs of protons can interact at each bunch crossing, and the number of interactions depends on the beam density. (CITE PP COLLISION X-SECTION) It is important to identify which proton-pair interaction is responsible triggering the event being studied. In this analysis, the primary vertex is chosen using track information. Every 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 redefined to always be from the primary vertex. The energy of objects affected by this reassociation is recalibrated, and this calibration is described in section 2.3.5.

Pileup is defined as energy in the detector which does not come from the interaction being studied. There are two primary sources of pileup, in-time pileup and out-of-time pileup. In-time pileup comes from two main sources, multiple proton collisions in the event and “underlying event”. Pileup from the underlying event is the leftover energy that comes from what’s left of the two protons in the primary interaction. This energy is deposited in the detector blah blah, SHOULD THIS EVEN BE HERE? Pileup from collisions that are not from the primary vertex usually produce soft di-jet events. - REWORD Out-of-time pileup occurs when a collision happens during a different beam crossing than the event which is being recorded. The beam collision frequency is 25 ns, and information from some components of the detector is not read out faster than this. HOW IS THIS DEALT

WITH? SOMETHING ABOUT HCAL TIMING AND FIT

2.3.3 Isolation

Isolation is a concept that is very important when identifying leptons, and photons. It can be simply described as the total energy in the detector near an object of interest. The distance between two objects (ΔR) is defined by equation 2.2, where ϕ is the angle measured in plane transverse to the beam direction, and η is the pseudorapidity defined by equation 2.3. In the equation for η , the variable θ is defined as the angle measured from the beam axis to the axis of the transverse plane to the beam axis. In particle physics, η is preferred over θ when describing a particle's momentum along the beam axis because changes in η are lorentz invariant, whereas the same is not true for changes in θ .

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \quad (2.2)$$

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right) \quad (2.3)$$

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

2.3.4 leptons and photons

Z bosons can decay leptonically to $e\bar{e}$, $\mu\bar{\mu}$, and $\tau\bar{\tau}$ pairs, but identifying τ s is very complicated in CMS, and their resolution is not as good as electrons and muons. 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 ECAL is designed to contain Electrons and photons. briefly describe electrons have tracks, photons dont. then talk about muons, and then close with one thing in common with all three objects is isolation. reference later chapter with detailed selection.

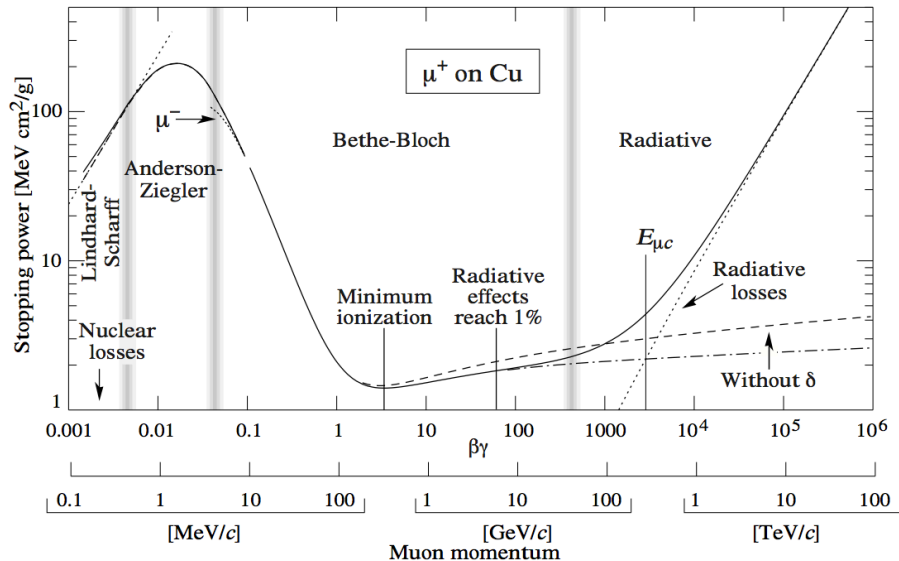


Figure 2.9: Stopping power of copper as a function of muon momentum. Typical momenta of muons in this analysis are between 20-200 GeV.

2.3.5 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 [2] 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 L1L2L3. Jets are

corrected such that the calibrated energy is as close to the parton that the jet came from as possible. 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 η , and L3 corrects the overall energy scale. The L1 correction is done by assuming a flat overall energy density 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 and the second jet is non-central η . Scale factors are then derived to correct the energy measurement for the non-central jet in separate regions of η . 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 L1L2L3 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.

add plot of closure with full JECs

The full description of cuts used to define jets for the analysis is defined in section

Chapter 3

Missing Transverse Momentum

In the final state targeted by this analysis, large values of E_T^{miss} can arise from the LSP escaping detection. It is important that E_T^{miss} from standard model processes is predicted accurately. E_T^{miss} can be calculated using the full collection of PF candidates described in section 2.3.1. In the following sections, E_T^{miss} calculations and corrections are described in detail.

3.1 raw MET

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

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

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

3.2 Corrections to MET

E_T^{miss} comes from three main sources, resolution effects due to mismeasurement of physics objects (referred to as fake E_T^{miss}), particles escaping detection in the detector (for example by being too far forward in η to be within the detector geometry), and real physics process where particles do not interact with the detector, thereby escaping detection. In the third case, the particles that are not detected 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^{miss} , it is very important to be able to identify and measure backgrounds with fake E_T^{miss} and real E_T^{miss} .

When calculating E_T^{miss} , the total contribution from pileup should be balanced since it is essentially random noise. Pileup energy is removed from each jet's cone when the L1 correction is applied to the jet which creates an imbalance of E_T^{miss} . The imbalance is corrected by adding this energy back in, and this is called the Type 1 correction.

add plot with metx mety etc

3.3 Type 1 MET

Type 1 E_T^{miss} refers to a version of E_T^{miss} where the corrections that are applied to jets are propagated to the E_T^{miss} calculation as described in section 2.3.5. 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^{miss} is not as straightforward as just adding the corrections vectorially with the E_T^{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 The jet corrections are not meant to be applied to lepton and photon objects. When propagating the jet corrections to the E_T^{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.

3.3.1 Data vs. MC Comparison

In this section we compare data to MC when constructing E_T^{miss} in events with $Z \rightarrow \ell\ell$. First we show the E_T^{miss} distribution after applying the Type-1 corrections which can be seen in figure 3.1. Then we show the E_T^{miss} components split into several categories of pf candidate type and η region which can be seen in figures 3.2, 3.3, and 3.4. This is done by taking the negative magnitude of the vector-sum of only the charged, neutral hadronic, or neutral electromagnetic PF candidates after separating in different regions of η defined in table 3.1.

Table 3.1: Definitions for the η region chosen for each subdetector.

Subdetector Region	η 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$

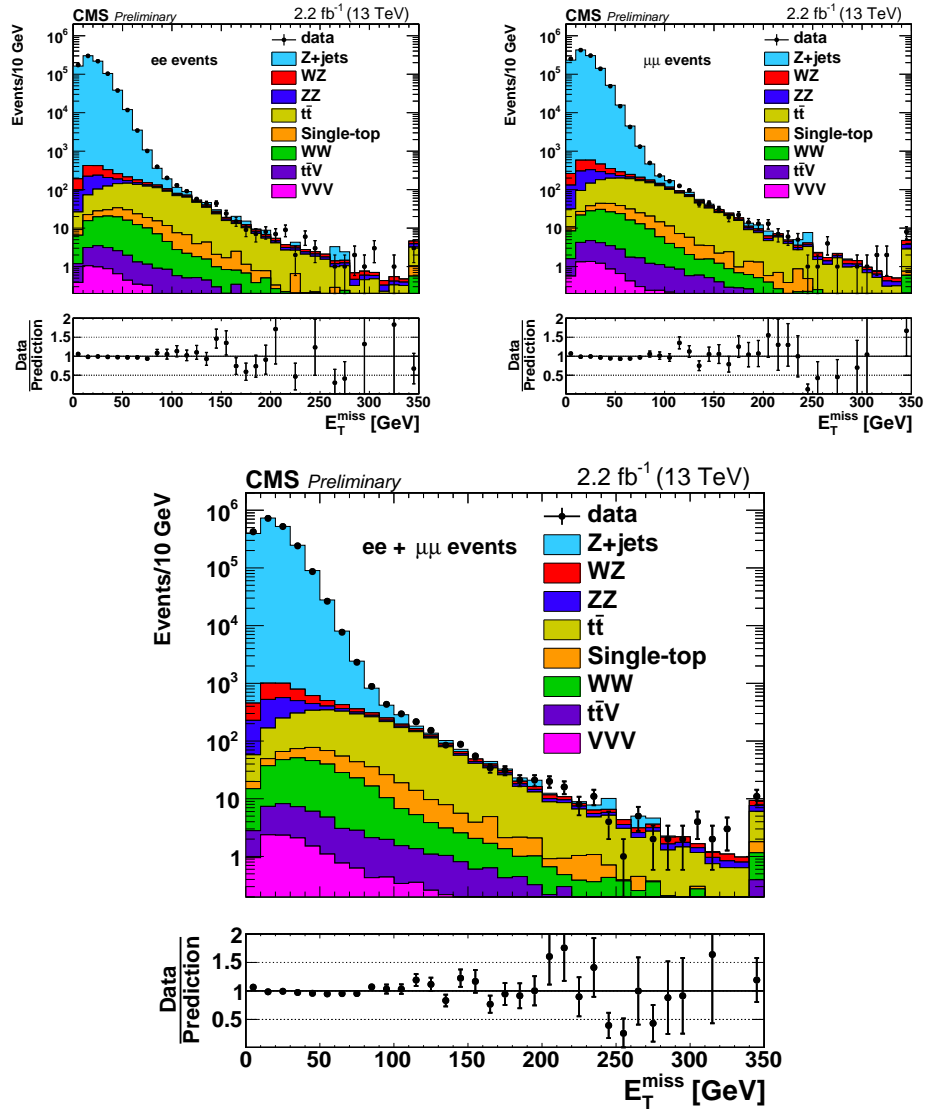


Figure 3.1: The E_T^{miss} distribution is shown with data vs. MC for events with $Z \rightarrow \ell\ell$. The top row shows ee events on the left and $\mu\mu$ events on the right, and the bottom row shows ee+ $\mu\mu$ events together.

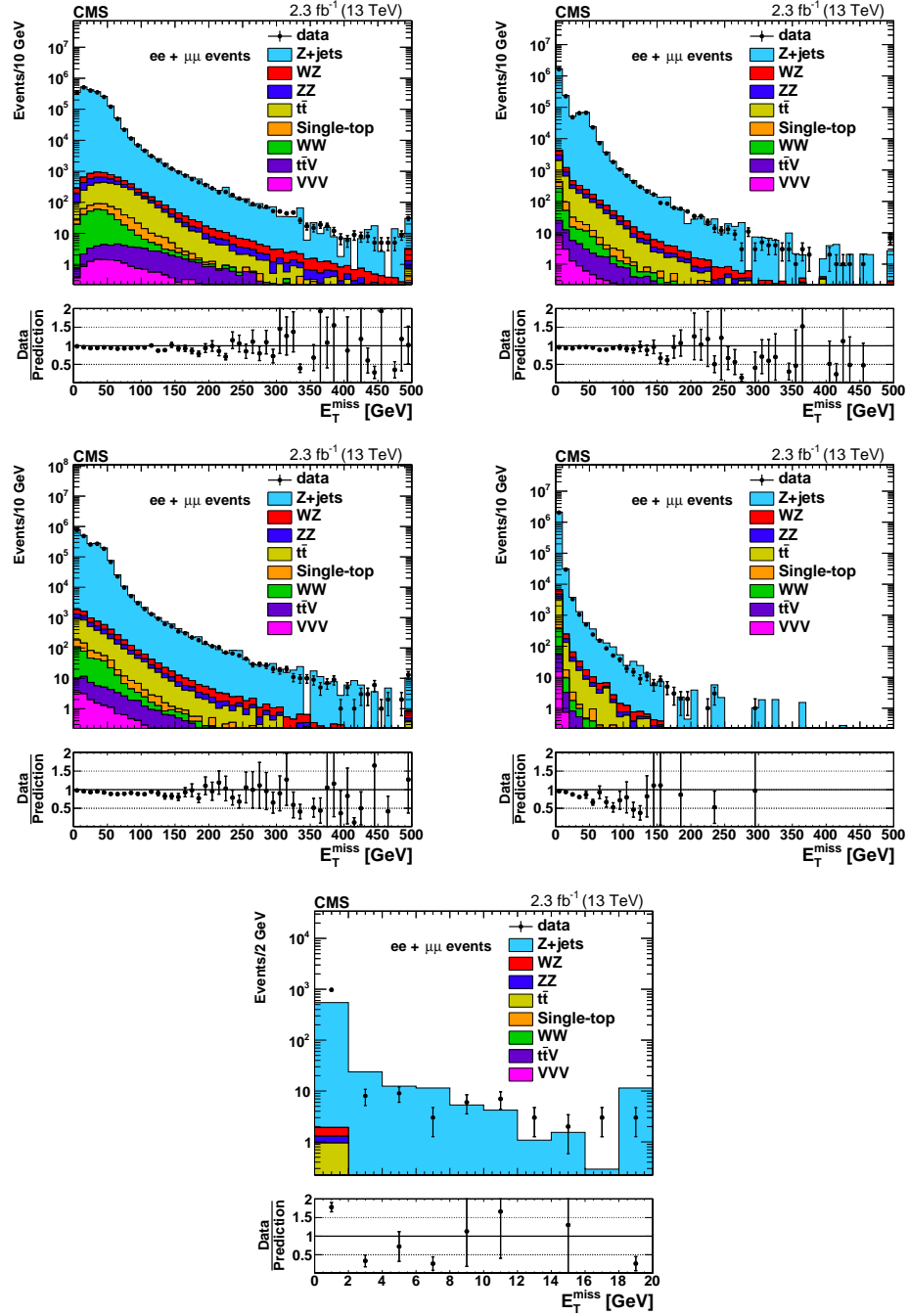


Figure 3.2: The E_T^{miss} distribution is shown for charged PF candidates only. The top row shows the barrel region on the left and transition region between the barrel and endcap on the right, the second row shows the endcap region including the tracker on the left and endcap region excluding the tracker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

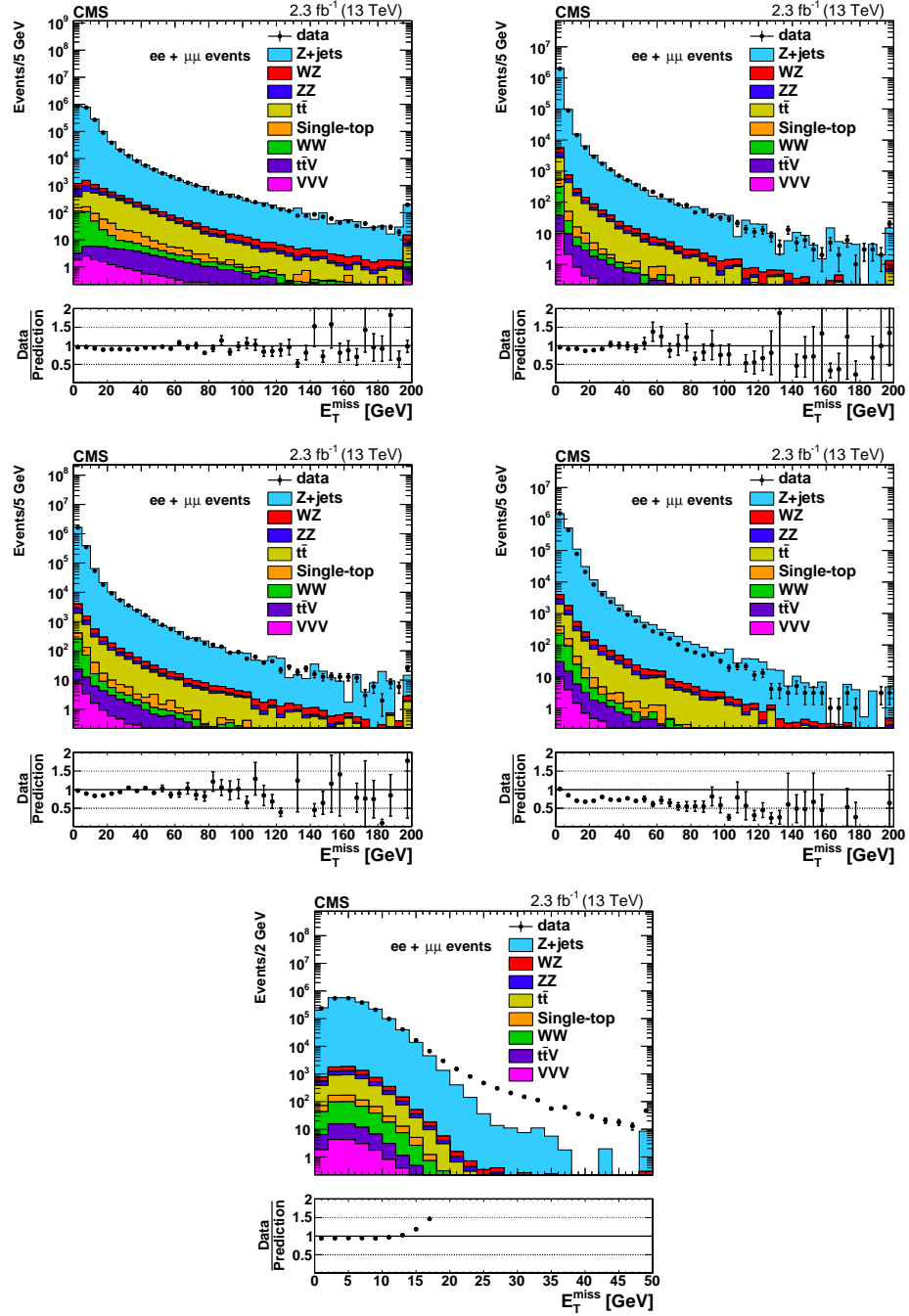


Figure 3.3: The E_T^{miss} distribution is shown for neutral electromagnetic PF candidates only. The top row shows the barrel region on the left and transition region between the barrel and endcap on the right, the second row shows the endcap region including the tracker on the left and endcap region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

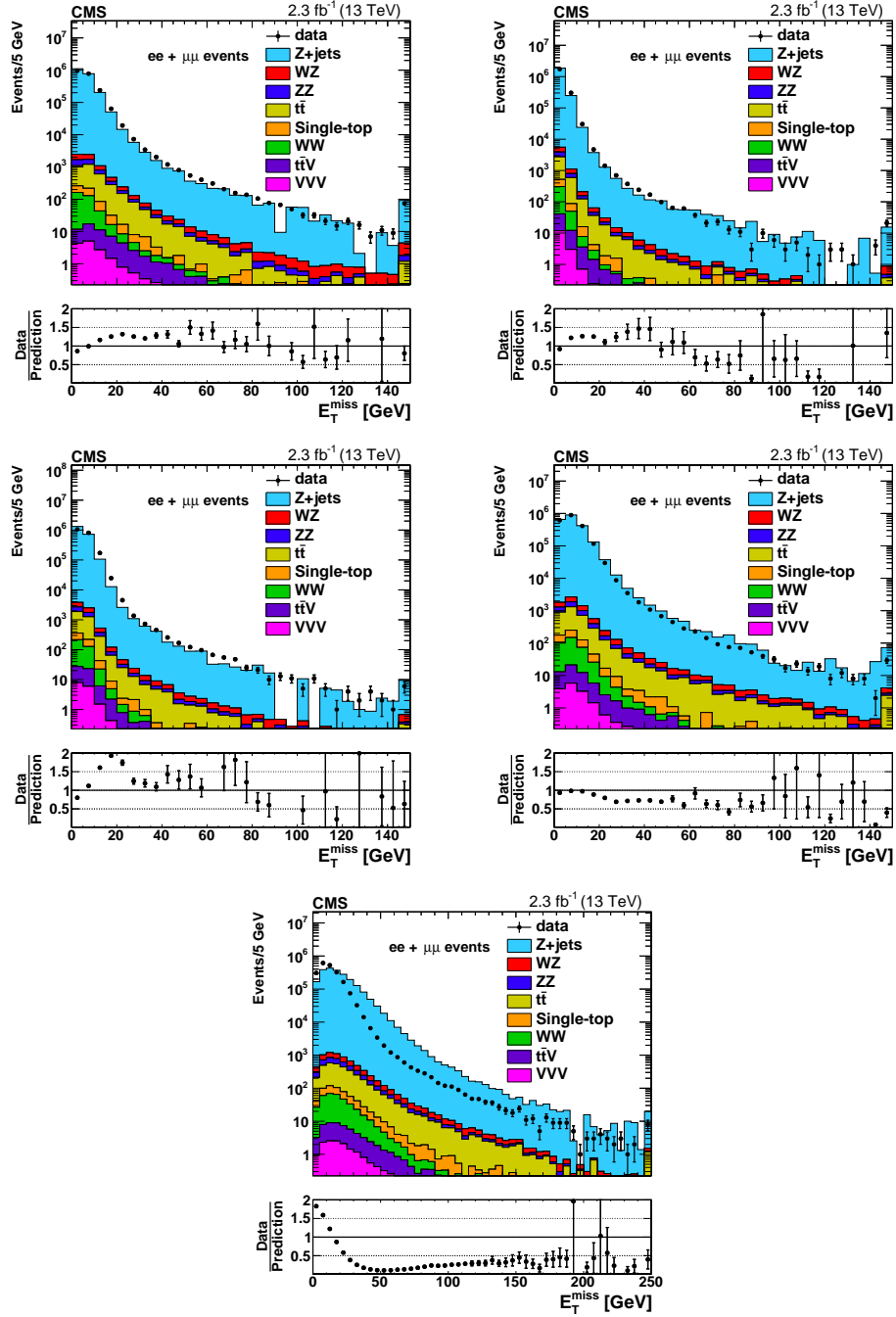


Figure 3.4: The E_T^{miss} distribution is shown for neutral hadronic PF candidates only. The top row shows the barrel region on the left and transition region between the barrel and endcap on the right, the second row shows the endcap region including the tracker on the left and endcap region excluding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

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: List of all triggers used in this analysis.

Signal triggers
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*
$e\mu$ triggers
HLT_Mu17_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v*
HLT_Mu8_TrkIsoVVL_Ele17_CaloIdL_TrackIdL_IsoVL_v*
HLT_Mu30_Ele30_CaloIdL_GsfTrkIdVL_v*
single-γ triggers
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*
H_T triggers
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

For the backgrounds predicted with Monte-Carlo (MC), the events are required to pass the same triggers as in data using the simulated triggers in the MC. This is seen in the inclusive data vs. MC plots in section **ADD SECTION WITH data vs MC** to show good agreement after requiring 2 leptons in the event. Cross sections and k-factors are applied to MC backgrounds to normalize them to the calculated luminosity of the signal triggers, which is 2.3 fb^{-1} . Cross section values are calculated using MadGraph software (CITATION), and k-factors are derived in different control regions in order to correct for beyond leading-order effects. In addition, the MC is reweighted to accurately represent the number of pileup events in data.

add a paragraph or two about generators, and what was used to make these samples with citations

Table 4.2: List of MC samples.

Process	Dataset Name	Cross Section [pb]
Z + jets	/DYJetsToLL_M-10to50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/*	18610
	/DYJetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/*	6104
	/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
	/ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8/*	38.09
single top	/ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8/*	38.09
	/TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX_madspin-pythia8/*	0.2043
	/TTWJetsToQQ_TuneCUETP8M1_13TeV-amcatnloFXFX_madspin-pythia8/*	0.40620
$t\bar{t}V$	/TTGJets_TuneCUETP8M1_13TeV-amcatnloFXFX_madspin-pythia8/*	3.697
	/TTZToQQ_TuneCUETP8M1_13TeV-amcatnlo-pythia8/*	0.5297
	/TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8/*	0.2529
	/ttHTtoNonbb_M125_13TeV-powheg-pythia8/*	0.5085
	/ZZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8/*	0.01398
VVV	/WWZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8/*	0.1651
	/WZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8/*	0.05565
	RunIISpring15MiniAODv2-74X_mcRun2_asymptotic_v2-v/MINIAODSIM	
Signal	SMS-T5ZZ_mGluino-*_mLSP-*_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	See [3]

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 2.3.2. For every event, we require the presence of at least one primary vertex satisfying the following criteria:

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

4.3.2 MET filters

A set of filters designated “ E_T^{miss} filters” are used in this analysis, where these filters are designed to remove events in data where large values of E_T^{miss} not due to physics are observed. Large values of E_T^{miss} in events that are due to physics effects are due to detector 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^{miss} vector aligns with an ECAL object in ϕ , 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^{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 gives a large p_T value, which then translates to large E_T^{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 μ . τ s are not considered for reasons described in 2.3.4. 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 an 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
- ΔR 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 ρ corrections scheme. In this scheme, ρ 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_{lepton} and ρ 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 η 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 η .

η 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

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

Table 4.4: Effective area values for electrons derived in separate regions of η .

η 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

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

The plots in figures 4.1 and 4.2 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. This distinction is made by requiring a generator level lepton to be within a cone of $\Delta R < 0.4$ for prompt leptons. Any lepton that passes the ID requirements that does not pass this matching requirement is designated as a non-prompt lepton.

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.

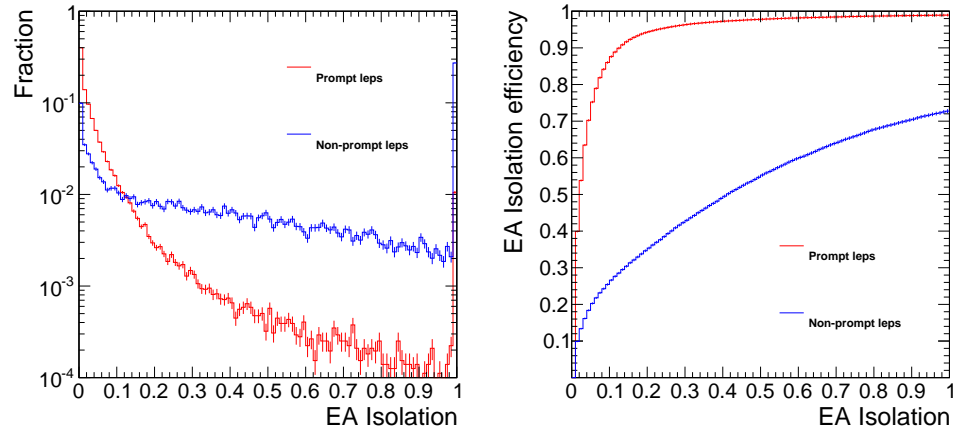


Figure 4.1: The isolation distribution is shown for electrons on the left in simulated $Z + \text{jets}$ MC events. On the right, the efficiency is shown when integrating all bins to the left of the value on the x-axis.

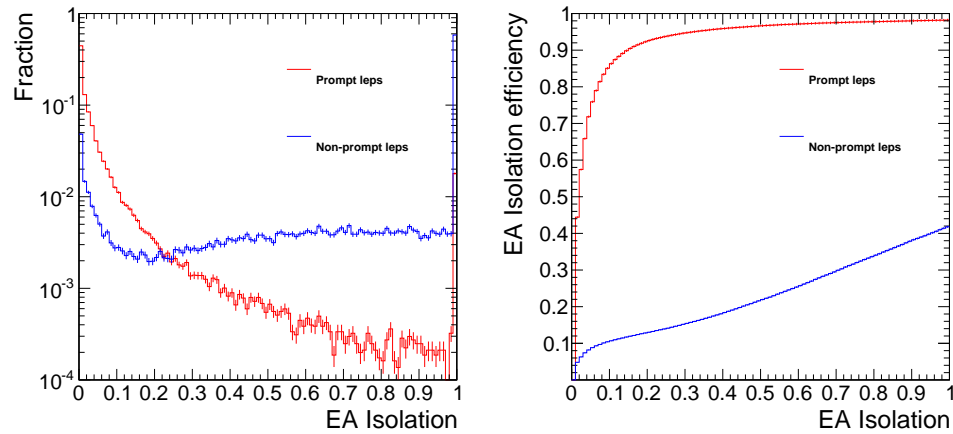


Figure 4.2: The isolation distribution is shown for muons on the left in simulated $Z + \text{jets}$ MC events. On the right, the efficiency is shown when integrating all bins to the left of the value on the x-axis.

4.3.5 Lepton Scale Factors

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. Additional scale factors are derived to correct for differences in simulation of leptons when using FastSim [4] 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 kinematic cut requirements can then be assessed on 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 4.3, scale factors for muon ID and isolation criteria are shown in 4.4, and FastSim scale factors are shown in 4.5.

4.3.6 Electron Selection

Output from an MVA is used to identify the electrons used in this analysis. This method is based on a previously approved method [5], and updated to be used to identify electrons in 13 TeV data. Electrons with $p_T > 15$ GeV and $|\eta| < 2.4$ are considered. The electron MVA requirements are listed in Table 4.5 for completeness.

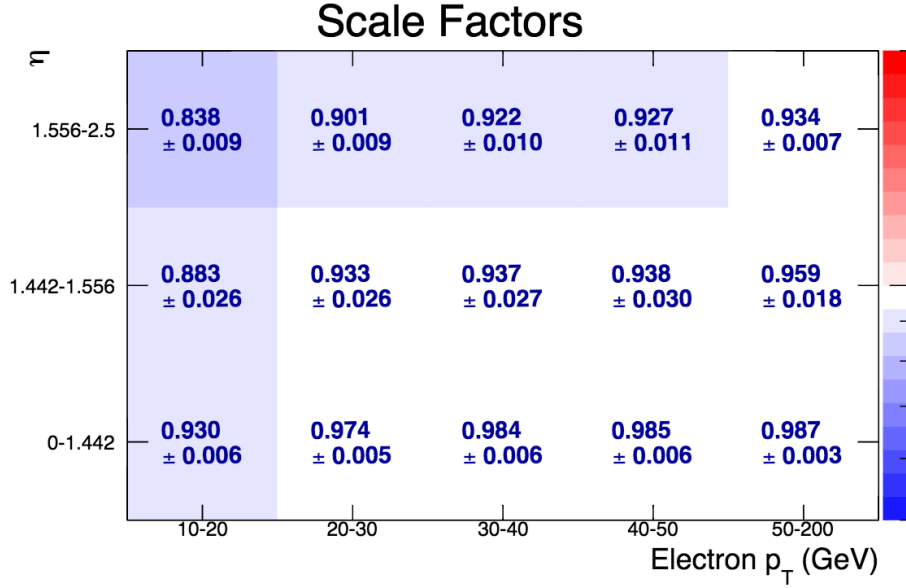


Figure 4.3: Electron scale factors as a function of p_T and η . The scale factor is applied as an event weight once per electron in the event.

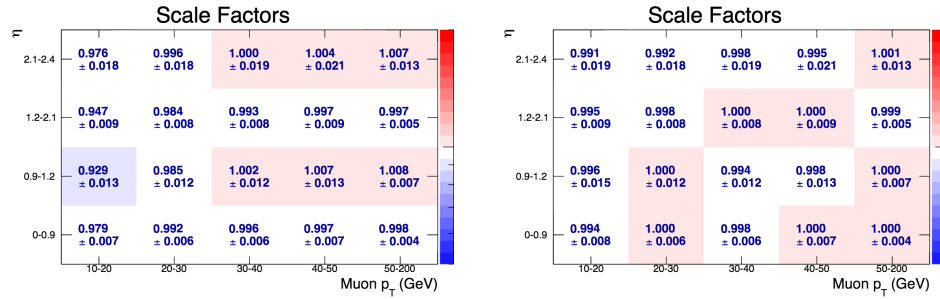


Figure 4.4: Muon scale factors as a function of p_T and η . The left plot shows the scale factors associated with the ID selection criteria, and the right plot shows the scale factors associated with the isolation selection criteria. Each scale factor is applied as an event weight once per muon in the event.

4.3.7 Muon Selection

The muon selection criteria is based on studies performed by the muon POG at CMS [6][7]. 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

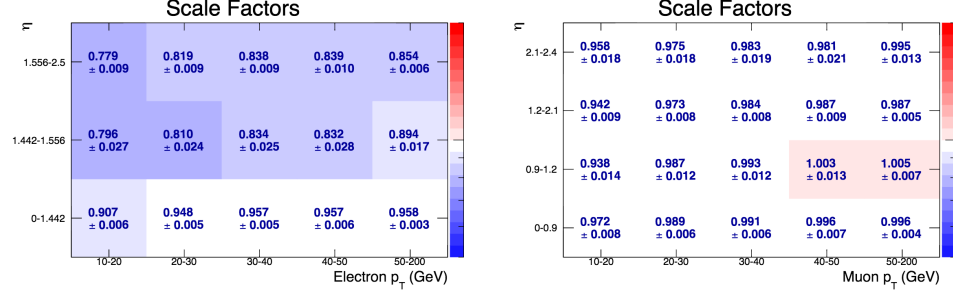


Figure 4.5: Fastsim scale factors as a function of p_T and η . The left plot shows the scale factors applied per electron, and the right plot shows the scale factors applied per muon. Each scale factor is applied as an event weight once per lepton in the event.

Table 4.5: Requirements for electrons with $p_T > 10$ 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

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. 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: Summary of the muons selection requirements.

good Global Muon Requirements	
Quantity	Requirement
Fraction of valid tracker hits	> 0.8
Normalized global-track χ^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(FIND SOME REFERENCE) 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 ρ 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: Summary of the photon isolation requirements.

PF objects	Requirement in barrel	Requirement in endcap
charged hadron	< 3.32	< 1.97
ρ corrected neutral hadron	$< 1.92 + 0.014 * p_T(\gamma)$ $+ 0.000019 * p_T^2(\gamma)$	$< 11.86 + 0.0139 * p_T(\gamma)$ $+ 0.000025 * p_T^2(\gamma)$
ρ 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^{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 π^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^{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, the electron would likely 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^{miss} having $\Delta\phi < 0.14$ are removed. The full photon requirement is listed in 4.8.

Table 4.8: Summary of the photon selection requirements.

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 γ within $dR < 0.2$ of an electron	
Matched to a pfjet with $\text{emfrac} > 0.7$	

4.3.9 MET

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

4.3.10 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 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.

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

4.3.11 Jets

The jets used in this analysis are described in section 2.3.5. Jets with $p_T > 35 \text{ GeV}$ and $|\eta| < 2.4$ are used with the L1L2L3 corrections applied to MC and the additional

L1L2L3ResidualL2L3 corrections applied to data. 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: Criteria for a jet to pass the loose ID requirement.

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: Summary of the jet selection requirements.

Quantity	Requirement
p_T	$> 35 \text{ GeV}$
$ \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 [8] is used to tag b-jets. This algorithm makes use of the fact that b-quarks have a larger lifetime than lighter partons due to having larger mass 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. **need to add a bit more to say what goes into the MVA** 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 4.6, and the cut value corresponding to the working point in this analysis is $CSV > 0.890$.

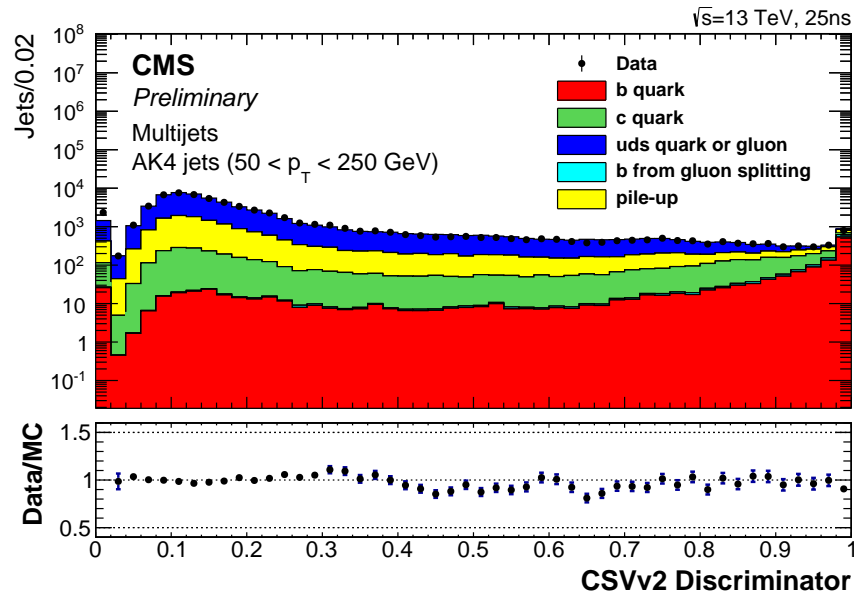


Figure 4.6: The distribution showing the CSVv2 Discriminator is shown comparing data vs. MC. The MC is split by jet parton flavor.

4.3.12 Signal Region Definitions

Multiple, orthogonal signal regions are based on cuts on the number of jets, H_T , and E_T^{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 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 3.0σ excess was seen in run-I by ATLAS [9]. The requirements of this signal region in this analysis are as follows:

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

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.

Table 4.11: List of all selections used to define the preselection and signal regions.

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

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 a data-driven method 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^{miss} Template

The Z + jets background is estimated using a method named the E_T^{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 system of jets, there should be no real E_T^{miss} . However, the resolution of the jet system is poor leading to events with E_T^{miss} due to these effects. The resolution of photons is similar to that of leptons, so

the assumption is made that E_T^{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^{miss} in the event, for example when an event containing a b-jet that decays semi-leptonically, since the E_T^{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, such that these events evenly sample the conditions over the full 2015 year period of data taking. 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.

Table 5.1: List of single photon triggers used in this analysis and their prescales.

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

In order to account for the fact that the Z-boson is massive whereas the photon is massless, the γ +jets sample is reweighted such that the boson p_T matches that of the Z+jets sample. A separate reweighting scheme is derived for each signal region, where the same cuts on N_{jets} and H_T are applied to the Z+jets and γ +jets data samples. The result of this reweighting can be seen in figures 5.1 and 5.2.

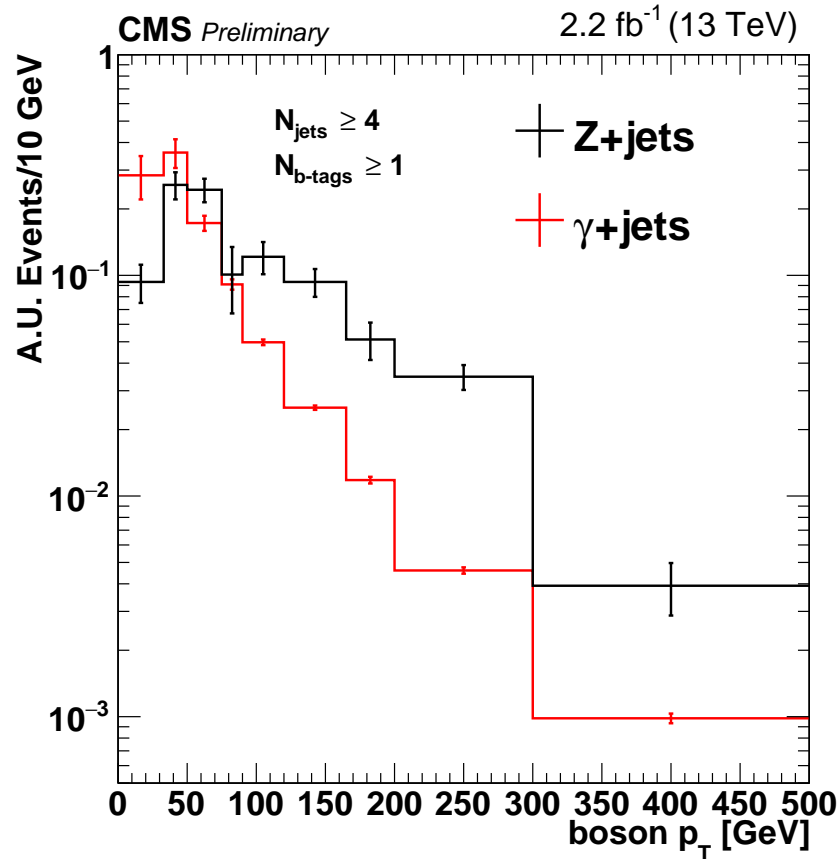


Figure 5.1: The p_T distributions are shown for the Z+jets (black) and γ +jets (red) events in signal region B where at least 1 jet is required to be b-tagged. The distributions are normalized to unit area.

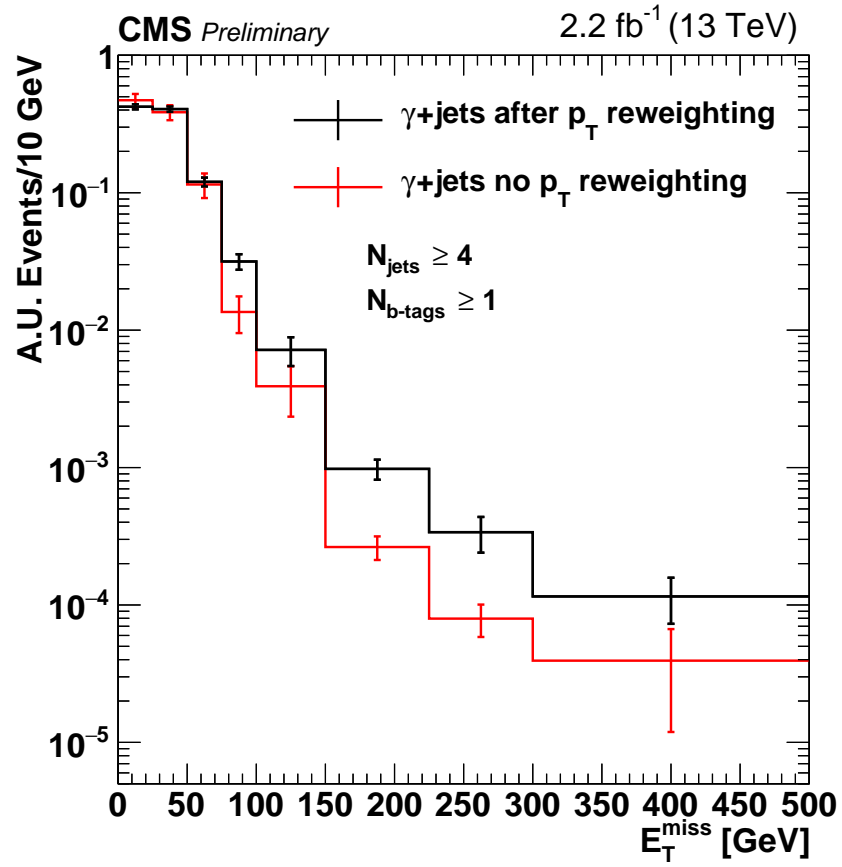


Figure 5.2: The E_T^{miss} distributions are shown before reweighting the γ + jets sample in p_T (red) and after reweighting (black) for γ + jets events in signal region B where at least 1 jet is required to be b-tagged. The distributions are normalized to unit area.

After reweighting in p_T , the resulting E_T^{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^{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.

5.1.1 Systematic Uncertainties in the E_T^{miss} Template Prediction

The systematic uncertainty associated with the E_T^{miss} template prediction comes from two sources: a MC closure study, and the uncertainty associated with normalizing in low E_T^{miss} . The largest uncertainty comes from the results of the closure study. This study is done by generating a E_T^{miss} template using simulated $\gamma + \text{jets}$ events and using this template to predict the E_T^{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^{miss} region, or the statistical uncertainty whichever is larger. The systematic uncertainty chosen for all regions is listed in table 5.3.

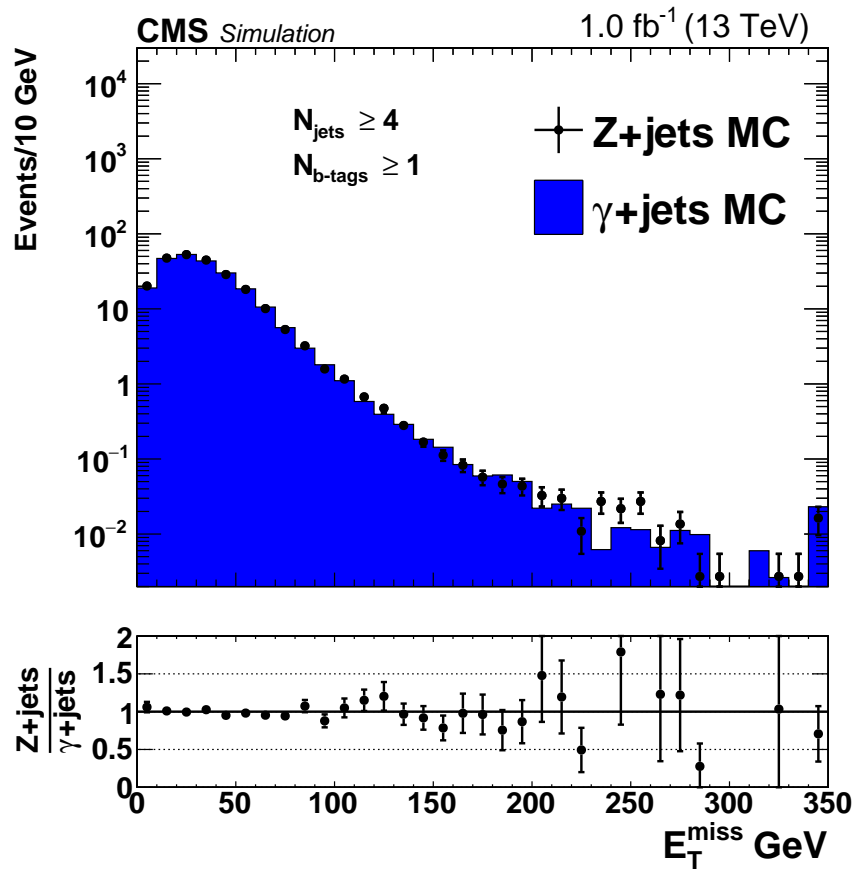


Figure 5.3: The closure test for Signal region B where at least 1 jet is required to be b-tagged is shown where simulated γ +jets events are used to predict the E_T^{miss} for simulated Z+jets events. The distributions are normalized to 1 fb^{-1} of data.

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 discrepancy between the $\gamma + \text{jets}$ prediction of the Z + jets MC for each E_T^{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 ± 2.9	40.5 ± 0.9	2.9 ± 0.2
$\gamma + \text{jets}$	190.2 ± 1.1	41.8 ± 0.5	2.7 ± 0.1
ratio	1.01 ± 0.02	0.97 ± 0.02	1.07 ± 0.07
Uncertainty	2 %	3 %	10 %
$E_T^{\text{miss}} [\text{GeV}]$	150 - 225	225 - 300	≥ 300
Z+jets	0.43 ± 0.03	0.11 ± 0.02	0.02 ± 0.01
$\gamma + \text{jets}$	0.47 ± 0.03	0.07 ± 0.01	0.03 ± 0.01
ratio	0.92 ± 0.10	1.50 ± 0.34	0.66 ± 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 discrepancy between the $\gamma + \text{jets}$ prediction of the Z + jets MC for each E_T^{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	≥ 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^{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%
≥ 1 b-tag	10%
SRB	
b-veto	3%
≥ 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^{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 the number of OF events comes from the combinatorics that arise from not distinguishing $e\mu$ from μe , 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.

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

These ratios are be 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 μ 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 $R_{SF/OF}$.

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

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

$R_{SF/OF}$ 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 control and signal region on

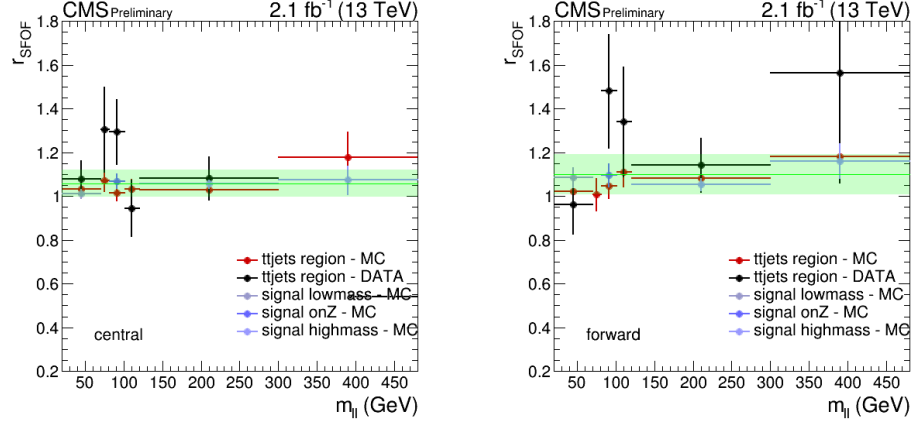


Figure 5.4: Direct measurement of $R_{SF/OF}$ in the control region for data (black) and $t\bar{t}$ MC (red). Values in the signal region are shown in blue for $t\bar{t}$ MC. Central rapidity is shown on the left and forward rapidity on the right. The green band represents the overall uncertainty on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

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 Measureing $R_{SF/OF}$ 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 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.

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

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_{jets} , N_{vtx} , lepton p_T , $m_{\ell\ell}$, E_T^{miss} , and N_{b-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 significant dependency on lepton p_T is observed, and the value of $r_{\mu/e}$ is especially stable with respect to N_{jets} and E_T^{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

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	$1.26 \pm 0.01 \pm 0.25$

based on these results.

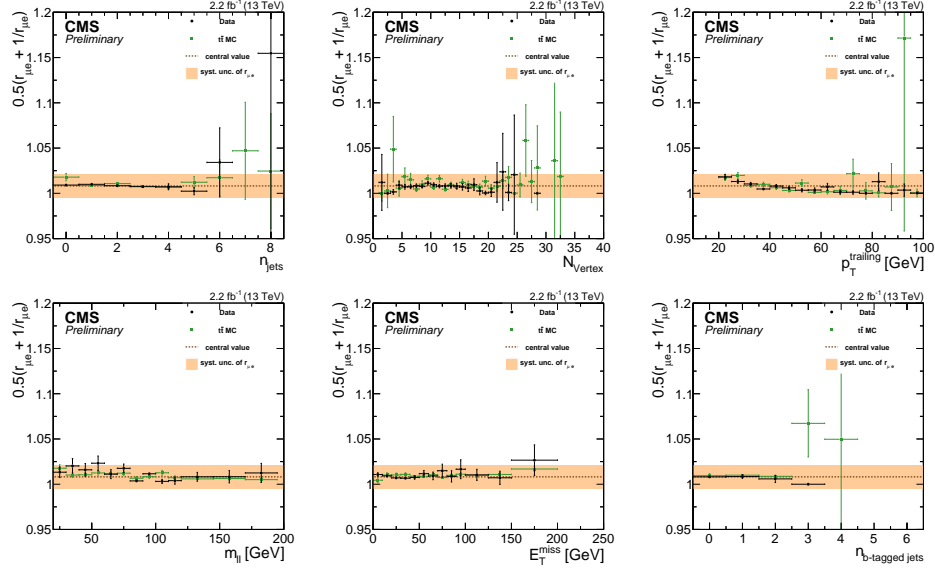


Figure 5.5: $0.5(r_{\mu e} + 1/r_{\mu e})$ dependency for central leptons shown as a function of N_{jets} , N_{vtx} , lepton p_T , $m_{\ell\ell}$, E_T^{miss} , and $N_{b\text{-tags}}$. The uncertainty due to the assigned 10% systematic uncertainty on $r_{\mu/e}$ is indicated by the orange band.

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

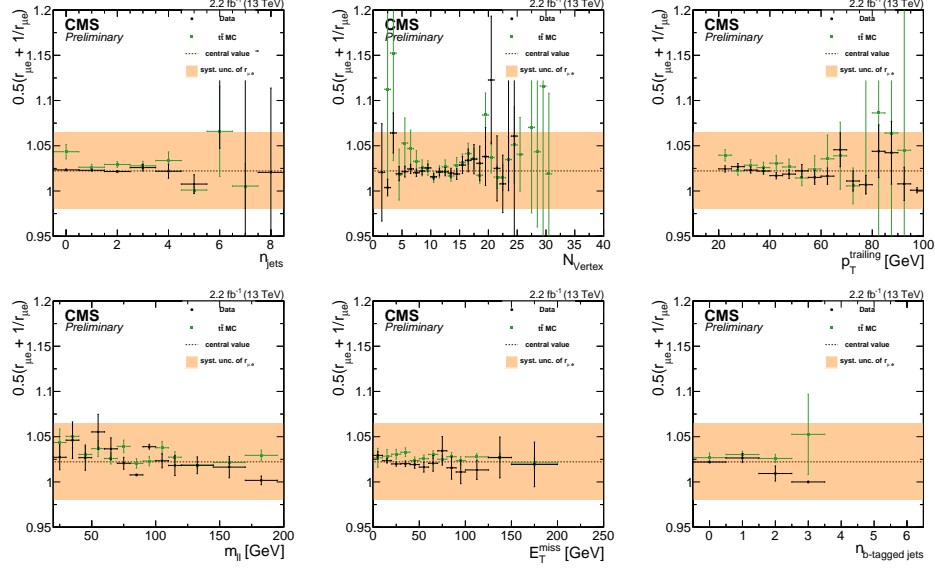


Figure 5.6: $0.5(r_{\mu e} + 1/r_{\mu e})$ dependency for forward leptons shown as a function of N_{jets} , N_{vtx} , lepton p_T , $m_{\ell\ell}$, E_T^{miss} , and $N_{b\text{-tags}}$. The uncertainty due to the assigned 20% systematic uncertainty on $r_{\mu/e}$ is indicated by the orange band.

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 $E_T^{\text{miss}} > 100\text{GeV}$ 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 400GeV is required to keep the PFHT triggers efficient. This is motivated in Fig. 5.7, 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 $\approx 400\text{GeV}$, 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 350GeV . Therefore

a higher offline H_T threshold is needed in the case of MC to match this trigger requirement.

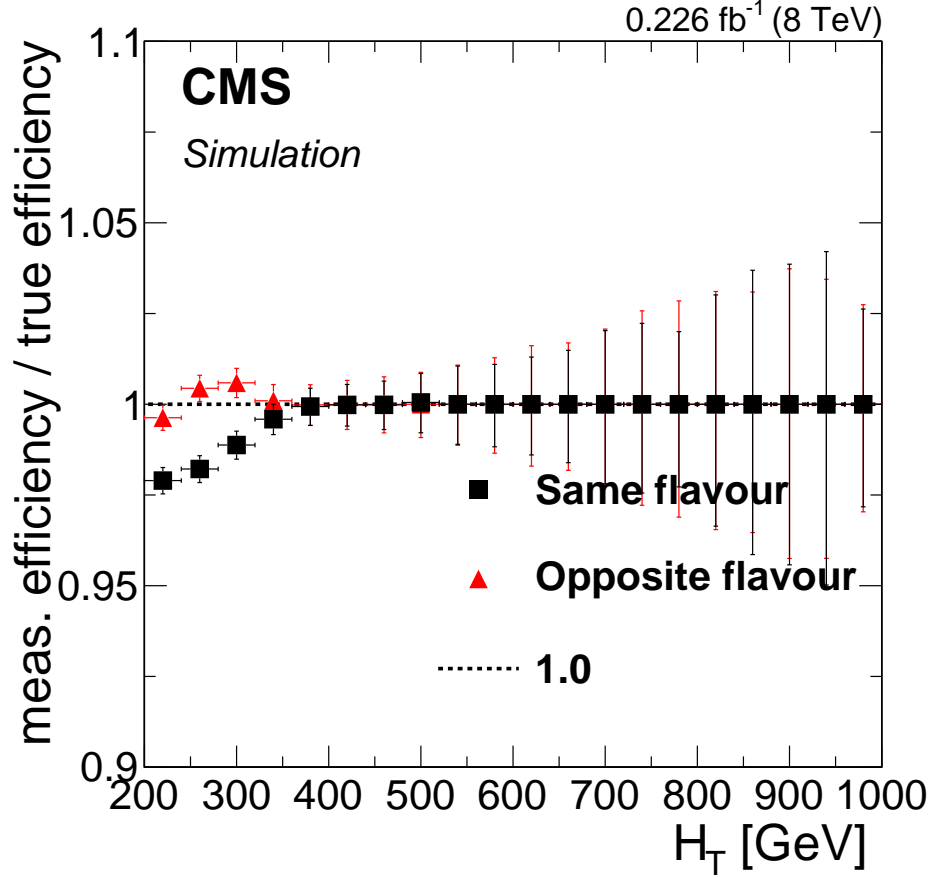


Figure 5.7: Ratio of measured to true trigger efficiency for combined SF and OF triggers as a function of H_T measured in $t\bar{t}$ simulation. A value of 1 indicates no bias due to the choice of the supporting triggers.

The resulting trigger efficiencies measured in data and MC are shown in table 5.7. All the trigger efficiencies are consistent with each other 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.

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

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 ± 0.01
$\mu\mu$	430	462	0.93 ± 0.01
$e\mu$	160	171	0.94 ± 0.03
	Forward		
ee	285	297	0.96 ± 0.02
$\mu\mu$	226	244	0.93 ± 0.02
$e\mu$	64	72	0.89 ± 0.05
	MC		
	Central		
ee	982.3	1041.5	0.943 ± 0.001
$\mu\mu$	623.0	664.1	0.938 ± 0.002
$e\mu$	1763.7	1926.1	0.916 ± 0.001
	Forward		
ee	363.3	390.8	0.930 ± 0.003
$\mu\mu$	284.7	307.2	0.927 ± 0.003
$e\mu$	720.4	798.9	0.902 ± 0.002

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 .

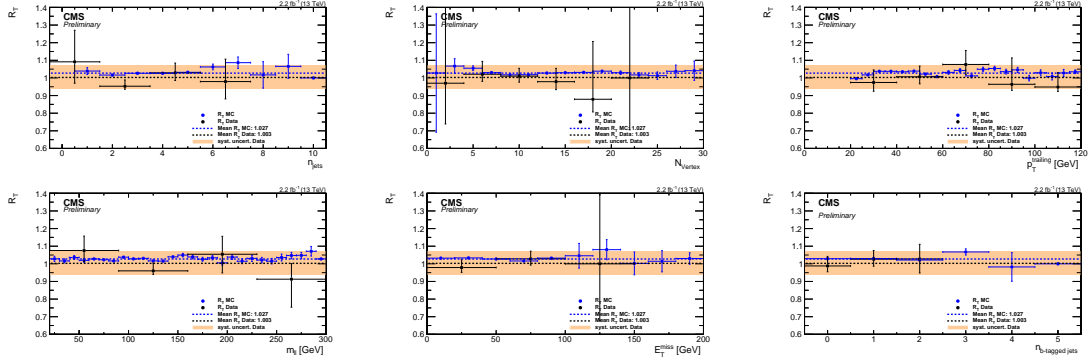


Figure 5.8: R_T dependency for central leptons shown as a function of N_{jets} , N_{vtx} , lepton p_T , $m_{\ell\ell}$, E_T^{miss} , and $N_{b\text{-tags}}$. The assigned 6.4% systematic uncertainty on R_T is indicated by the orange band.

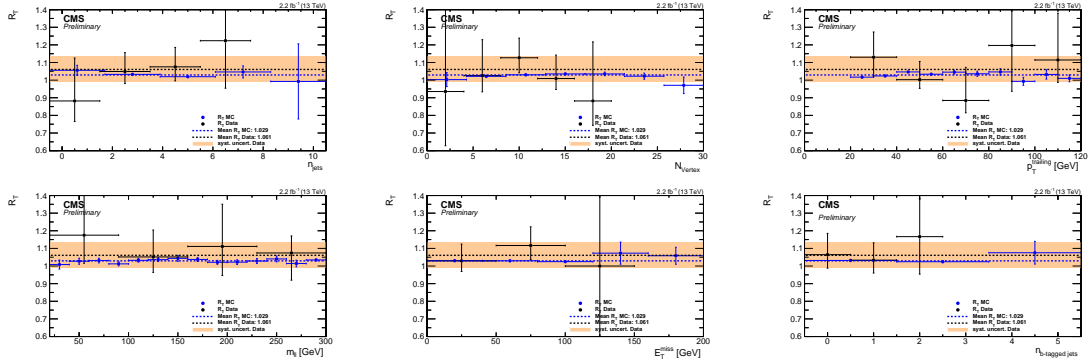


Figure 5.9: R_T dependency for forward leptons shown as a function of N_{jets} , N_{vtx} , lepton p_T , $m_{\ell\ell}$, E_T^{miss} , and $N_{b\text{-tags}}$. The assigned 6.4% systematic uncertainty on R_T is indicated by the orange band.

5.2.4 Results of factorization 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: The values for $r_{\mu/e}$, R_T , and $R_{SF/OF}$ are shown.

	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 a 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 ± 0.07	1.04 ± 0.07	1.08 ± 0.10	1.06 ± 0.09
Direct measurement	1.06 ± 0.06	1.05 ± 0.01	1.11 ± 0.14	1.08 ± 0.02
weighted average	1.04 ± 0.05	1.05 ± 0.01	1.10 ± 0.07	1.08 ± 0.02
	$R_{ee/OF}$			
	Central		Forward	
	Data	MC	Data	MC
Factorization method	0.44 ± 0.12	0.45 ± 0.12	0.43 ± 0.26	0.41 ± 0.26
Direct measurement	0.43 ± 0.04	0.45 ± 0.01	0.46 ± 0.11	0.44 ± 0.01
weighted average	0.43 ± 0.03	0.45 ± 0.01	0.46 ± 0.05	0.44 ± 0.01
	$R_{\mu\mu/OF}$			
	Central		Forward	
	Data	MC	Data	MC
Factorization method	0.57 ± 0.12	0.58 ± 0.12	0.65 ± 0.27	0.65 ± 0.26
Direct measurement	0.63 ± 0.05	0.60 ± 0.01	0.65 ± 0.12	0.64 ± 0.02
weighted average	0.63 ± 0.04	0.60 ± 0.01	0.65 ± 0.06	0.64 ± 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 ± 5	860 ± 5	2 ± 7	362 ± 3	356 ± 3	5 ± 4
DY + jets ($e\bar{e}, \mu\mu$)	19 ± 6	0 ± 0	18 ± 6	14 ± 4	0 ± 0	14 ± 4
DY + jets ($\tau\tau$)	11 ± 6	20 ± 5	-8 ± 8	6 ± 2	4 ± 3	2 ± 4
Single t	53 ± 1	52 ± 1	0 ± 2	21 ± 1	21 ± 1	0 ± 1
WW, ZZ, WZ	22 ± 0	15 ± 0	7 ± 0	10 ± 0	8 ± 0	1 ± 0
Other SM	9 ± 0	5 ± 0	3 ± 0	4 ± 0	3 ± 0	0 ± 0
Total MC simulation	980 ± 10	955 ± 7	24 ± 13	419 ± 6	394 ± 4	24 ± 8

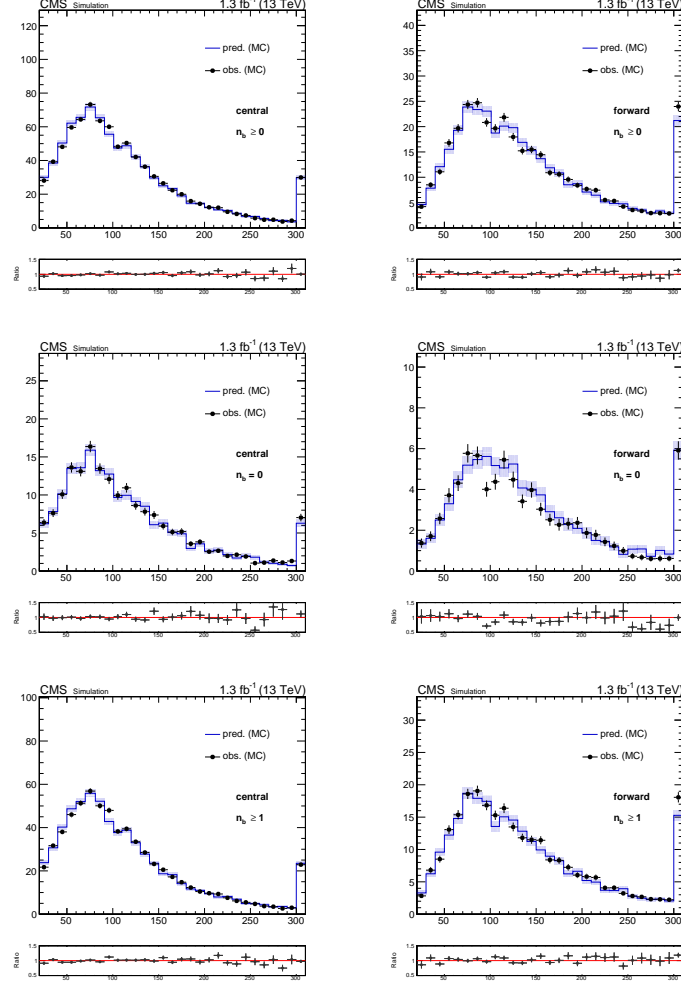


Figure 5.10: MC closure test for central (left) and forward (right) in the $m_{\ell\ell}$ variable. The blue histogram shows the MC prediction from the OF sample by multiplying with $R_{SF/OF}$ whereas the black markers correspond to the observation in the SF sample. This exercise is done on $t\bar{t}$ 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 ≥ 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^{miss} and N_{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 $m_{\ell\ell}$:81-101 GeV
- veto events with ≥ 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.

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

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

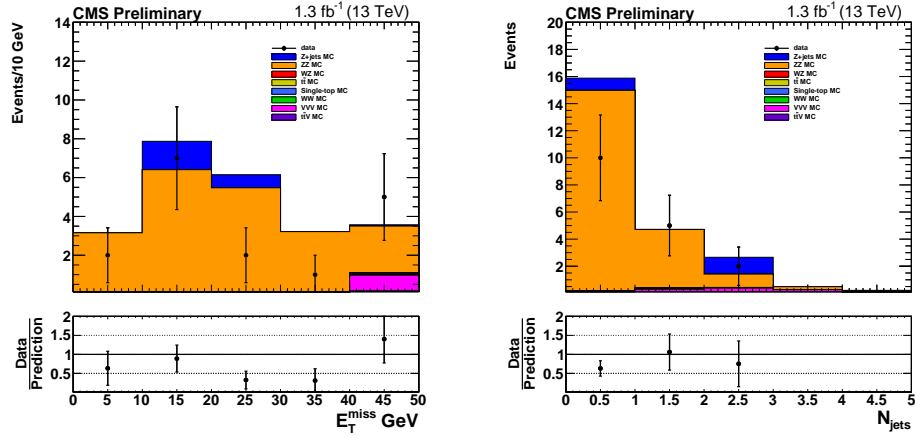


Figure 5.12: The E_T^{miss} and N_{jets} distributions are shown for data vs. MC in the 4-lepton control region. See Tables 5.13 and 5.14 for yields.

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

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

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

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

Chapter 6

Results

The results of this search are shown in the next section. The E_T^{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^{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^{miss} [GeV]	0 - 50	50 - 100	100 - 150	150 - 225	225 - 300	≥ 300
Z+jets	1333.1 ± 62.0	314.4 ± 24.3	24.3 ± 4.3	4.6 ± 0.6	1.5 ± 0.4	1.1 ± 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 ± 0.8	2.2 ± 1.1	1.6 ± 0.8	1.3 ± 0.6	0.8 ± 0.4	1.0 ± 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^{miss} [GeV]	0 - 50	50 - 100	100 - 150	150 - 225	225 - 300	≥ 300
Z+jets	182.6 ± 24.1	53.0 ± 11.4	4.5 ± 0.9	1.4 ± 0.4	0.7 ± 0.3	0.2 ± 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 ± 0.1	0.3 ± 0.1	0.3 ± 0.1	0.2 ± 0.1	0.1 ± 0.1	0.2 ± 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

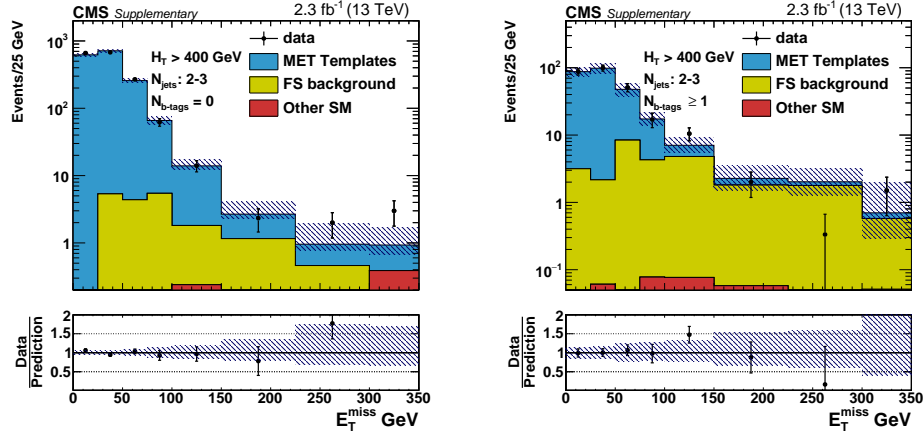


Figure 6.1: The E_T^{miss} distribution is shown for data vs. the data-driven predictions in signal region B. The left plot shows the prediction when requiring $N_{b\text{-jets}} = 0$, and the right plot shows the prediction when requiring $N_{b\text{-jets}} \geq 1$. The dashed line in the plot represents the full uncertainty. See tables 6.1 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^{miss} [GeV]	0 - 50	50 - 100	100 - 150	150 - 225	225 - 300	≥ 300
Z+jets	1907.2 ± 68.8	282.5 ± 16.4	10.0 ± 0.9	3.2 ± 0.6	0.3 ± 0.1	0.1 ± 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 ± 0.6	2.0 ± 0.8	1.0 ± 0.4	0.8 ± 0.3	0.5 ± 0.2	0.3 ± 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^{miss} [GeV]	0 - 50	50 - 100	100 - 150	150 - 225	225 - 300	≥ 300
Z+jets	415.4 ± 34.9	73.3 ± 8.0	5.0 ± 0.9	1.6 ± 0.3	0.4 ± 0.3	0.3 ± 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 ± 0.5	1.5 ± 0.7	0.8 ± 0.4	0.5 ± 0.2	0.2 ± 0.1	0.1 ± 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

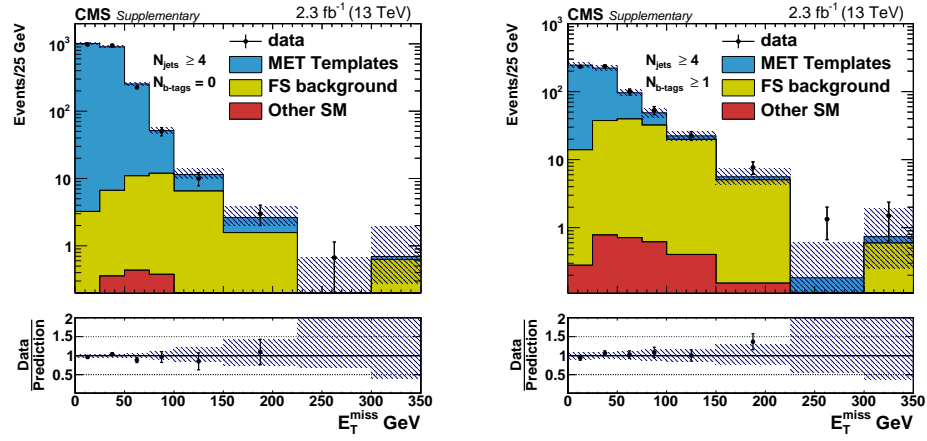


Figure 6.2: The E_T^{miss} distribution is shown for data vs. the data-driven predictions in signal region B. The left plot shows the prediction when requiring $N_{b\text{-jets}} = 0$, and the right plot shows the prediction when requiring $N_{b\text{-jets}} \geq 1$. The dashed line in the plot represents 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^{miss} [GeV]	0 - 50	50 - 100	100 - 150	150 - 225	≥ 225
Z+jets	1557.3 ± 56.6	386.2 ± 15.8	34.4 ± 5.0	8.1 ± 1.1	3.9 ± 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 ± 1.2	3.7 ± 1.6	2.3 ± 0.9	1.9 ± 0.8	2.1 ± 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

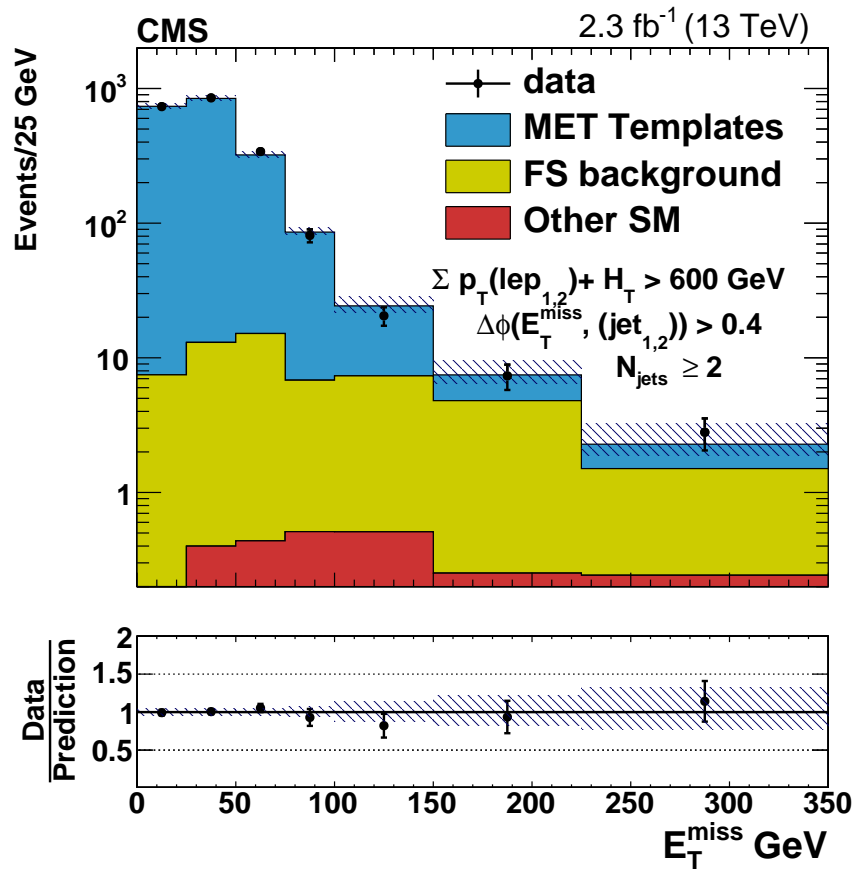


Figure 6.3: The E_T^{miss} distribution is shown for data vs. the data-driven predictions in the ATLAS signal region. See Tables 6.1 and 6.5 for yields.

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: Systematic uncertainties of the expected signal yield.

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	$\pm 2-5\%$
ISR Modeling	$\pm 1\%$
MC Statistics	$\pm 5-20\%$
Total uncertainty on signal	$\pm 13-24\%$

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 [10]. The analysis of the Van Der Meer scans gives an uncertainty value of 2.7% [11].

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 [12]. 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 Madgraph signal MC is generated using Fastsim [4], 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 η 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 [13][14], 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^{miss} , and more events at high E_T^{miss} . It varies from 5-20% depending on the signal region and the point in the scan.

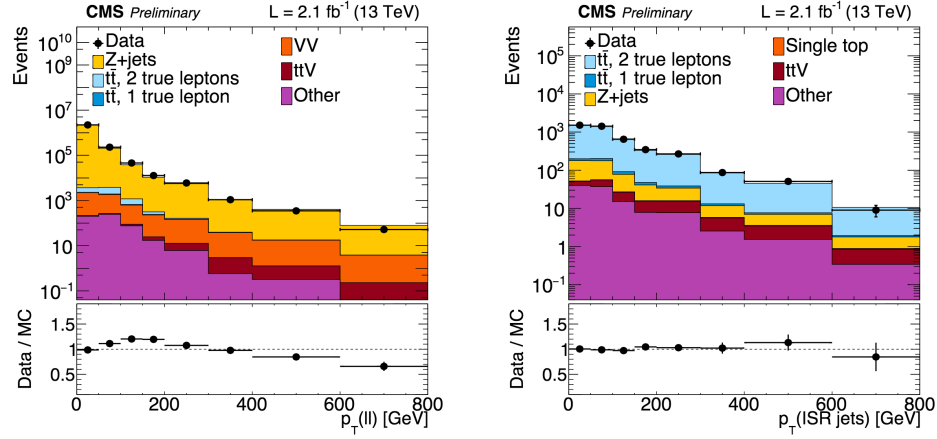


Figure 6.4: The ISR p_T in data and MC is shown for Z + jets (left) and $t\bar{t}$ (right).

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 the T5ZZgmsb model described in section 1.4. This particular SUSY model is expected to have many jets, so most of the sensitivity comes from SRB.

SOMETHING ABOUT 95% CL UPPER LIMITS

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.

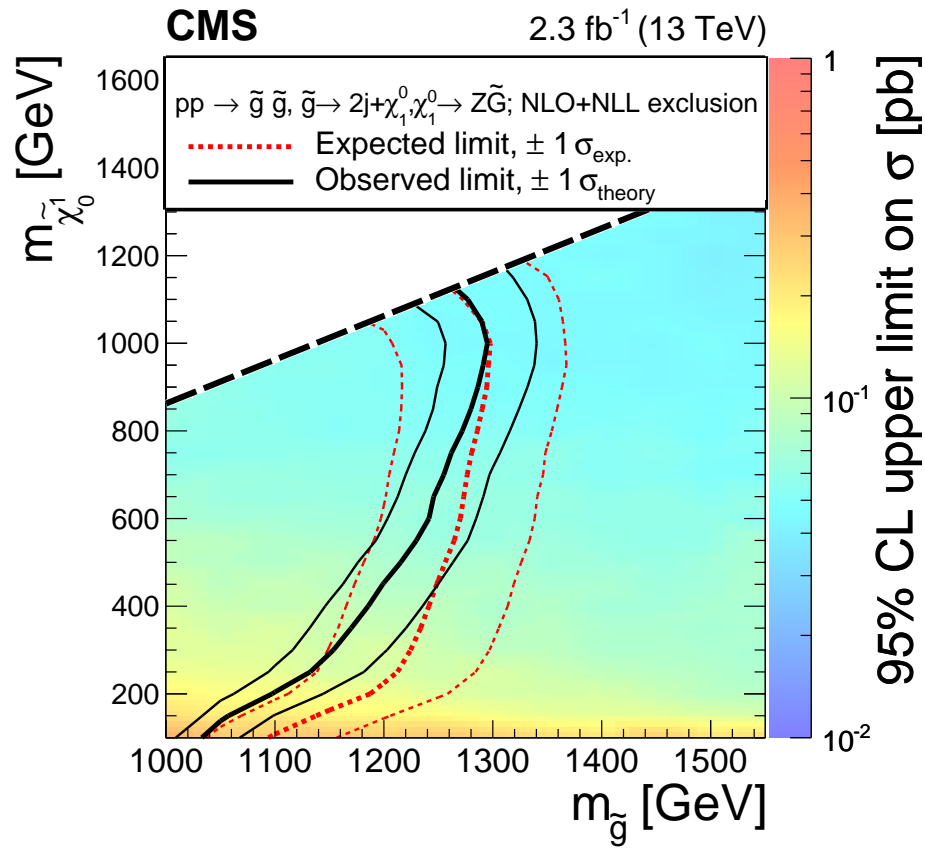


Figure 6.5: Exclusion contours are shown when we interpret the results of this analysis in the T5ZZ model. Everything to the left of the red (black) dotted line shows the masses which are excluded by the expected (observed) limit.

Chapter 7

Summary and Conclusions

Bibliography

- [1] 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).
- [2] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- k_t jet clustering algorithm”, *Journal of High Energy Physics* **4** (April, 2008) 63–+, arXiv:0802.1189. doi:10.1088/1126-6708/2008/04/063.
- [3] CMS Collaboration, “Gluino-gluino Cross Sections at 13 TeV”, (2013). twiki.cern.ch/twiki/bin/view/LHCPhysics/SUSYCrossSections13TeVgluglu.
- [4] S. Abdullin, P. Azzi, F. Beaudette et al., “The fast simulation of the CMS detector at LHC”, *J. Phys. Conf. Ser.* **331** (2011) 032049. doi:10.1088/1742-6596/331/3/032049.
- [5] 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.
- [6] G. Bayatian et al., “CMS Physics Technical Design Report Volume I: Detector Performance and Software”. Technical Design Report CMS. CERN, Geneva, 2006.
- [7] CMS Collaboration, “Performance of CMS muon reconstruction in cosmic-ray events”, *Journal of Instrumentation* **5** (March, 2010) 3022–+, arXiv:0911.4994. doi:10.1088/1748-0221/5/03/T03022.
- [8] CMS Collaboration, “Identification of b-quark jets with the CMS Experiment”, *Journal of Instrumentation* **8** (2013), no. 13, P04013.
- [9] ATLAS Collaboration, “Search for supersymmetry in events containing a same-flavour opposite-sign dilepton pair, jets, and large missing transverse momentum in $\sqrt{s} = 8$ TeV pp collisions with the ATLAS detector”, *The European Physical Journal C* **75** (2015), no. 7,. doi:10.1140/epjc/s10052-015-3518-2.

- [10] 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).
- [11] CMS Collaboration, “CMS Luminosity Measurement for the 2015 Data Taking Period”, *CMS Physics Analysis Summary* (2015).
- [12] CMS Collaboration, “Identification of b-quark Jets at the CMS Experiment in the LHC Run2 Startup”, *CMS Physics Analysis Summary* (2015).
- [13] CMS Collaboration, “Determination of jet energy calibration and transverse momentum resolution in CMS”, *Journal of Instrumentation* **6** (2011), no. 11, P11002.
- [14] CMS Collaboration, “Jet Energy Corrections: Official Software Tools for applying JEC Corrections and Uncertainties”, (2015).
<https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookJetEnergyCorrections>.