Search for a Heavy Right-Handed W Boson and Heavy Neutrino of the Left-Right Symmetric Standard Model

A THESIS
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Nathaniel Joseph Pastika

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTORATE OF PHILOSOPHY

Jeremiah Mans

August, 2014

© Nathaniel Joseph Pastika 2014
ALL RIGHTS RESERVED

Acknowledgements

First and foremost I would like to acknowledge the support of my parents Diane and Peter Pastika without whom I never would have made it to this point on my life. Similarly, I thank all my friends and family for the support and encouragement they have shown me over the years.

Thanks to Bryan Dahmes, without whom I never would have been able to finish this thesis. Also many thanks to Phillip Dudero for laying the ground work for my thesis with his own. Thanks to Alex Gude and all the members of the Minnesota CMS group. You have all contributed greatly, from coding advice to the hours spent in conversation about nothing.

Never would I have stayed sane without my friends, Benjamin Borgeson, Tianran Chen, Kevin Christie, Alexey Finkel, Chad Geppert, Ryo Namba, Joffrey Peters, Adam Schreckenberger, and Pamela Vo Sherwood. You all left your mark on my life and I thank you all for it.

I would also like to thank all my collaborators in CMS and at the LHC who have made this research possible through their dedication. The excellence of the CMS and LHC physics programs speaks for itself, but none the less I am grateful to all those involved.

Last, but not least, I want to thank my adviser Jeremiah Mans for all his help and guidance. From the first time I met him as the professor of my modern experimental physics II class I knew that he was a man worth working for.

Abstract

The search for direct production of a heavy right-handed W_R boson, which decays to a right-handed neutrino $N_l(l = e, \mu)$, with a final state containing two same-flavor leptons and two jets, consistent with a left-right symmetric extension to the standard model, is presented. The search was conducted using the full $19.7\,\text{fb}^{-1}$ of proton-proton collision data collected at a center of mass energy of $\sqrt{s} = 8\,\text{TeV}$ by the Compact Muon Solenoid detector at the Large Hadron Collider. No significant deviation from the standard model is observed in either channel. 95% confidence level exclusion limits in the plane of W_R mass versus right-handed neutrino mass are set for the electron and muon channel that extends up to 3 TeV in W_R mass and exclude most neutrino masses below the mass of the W_R. The electron and muon channel data are also combined assuming degenerate right-handed neutrino masses. A brief discussion of the efforts to upgrade the readout electronics of the hadronic calorimeter of the Compact Muon Solenoid detector is also presented.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

One of today biggest questions in physics is the source of the matter/antimatter asymmetry in the universe. This thesis works to address a small piece of the puzzle. Cosmological and astrophysical measurements confirm that the universe at large is dominated by matter, but the known fundamental laws of physics have trouble accounting for this [1]. One of the key ingredients required to explain the matter/antimatter asymmetry is violation of charge-parity (CP) symmetry. The Weinberg-Glashow-Salam model [2, 3], also called the standard model of particle physics, provides CP violation through the weak nuclear force. The weak force breaks parity (P) symmetry maximally and CP symmetry as well. Unfortunately, this CP violation is insufficient to explain the matter dominated universe we live in. Furthermore, as an assumption in the construction of the standard model, a theoretical framework to understand the source of parity violation remains a mystery. These and other concerns hint to possible new physics beyond the standard model.

A class of models called left-right symmetric (LRS) extensions to the standard model [4] attempt to provide explanations to these questions. These models restore parity as a conserved symmetry in the standard model which is spontaneously broken at some higher energy scale. They also provide additional sources for CP violation sufficient to explain the matter/antimatter asymmetry [5]. Additionally, LRS models provide an explanation for the small but non-vanishing masses of the standard model neutrinos through the see-saw mechanism [6].

LRS models predict several new particles which may be discovered through direct

production at particle accelerators like the Large Hadron Collider (LHC). These include a right-handed partner to the W boson, called W_R, a massive neutral gauge boson, called Z', and massive unstable right-handed neutrinos, called N_l, which may be light enough to be discovered in the $\sqrt{s} = 8\,\text{TeV}$ center of mass energy collisions of the LHC. Discovery of any of these particles would be an exciting event, but the decay $W_R \to lN_l$ would be particularly valuable in learning about a possible LRS model. This search channel would not only provide a measurement of the mass of the W_R but also the mass of the heavy neutrino and would allow a measurement of properties of both particles.

This thesis details the search for evidence of an LRS model through the process $W_R \to lN_l$ with the Compact Muon Solenoid (CMS) detector at the LHC. Results are presented using the full dataset collected from $\sqrt{s} = 8\,\text{TeV}$ proton-proton collisions collected during LHC Run-I. The decays of the W_R to electron and muon flavor heavy neutrinos are considered. The heavy neutrinos are themselves unstable and decay to a charged lepton and two quarks, which appear as hadronic jets in the detector. This leads to the final states for each search channel of two same flavor charged leptons and two jets, $eejj$ and $\mu\mu jj$ for the electron and muon channels respectively. No statistically significant excess over the standard model expectations is observed so exclusion limits on possible masses of a W_R boson, N_e, and N_μ are presented. The results of this search are published in [7, 8].

Chapter 2 gives a brief overview of the standard model and an introduction to the salient features of LRS models. Chapter 3 details the CMS detector and the LHC. Chapter 4 serves as a short interlude briefly describing the efforts to upgrade the hadronic calorimeter in CMS. Chapter 5 describes the process of event selection and object reconstruction. Chapters 6 and 7 cover the background determinations in the signal region and the systematic uncertainties assigned thereon. Finally, the final exclusion limits are presented in chapter 8.

Chapter 2

Theory of Left-Right Symmetric Models

2.1 History of the Standard Model

Our understanding of modern particle physics is firmly based on the advancements made over the last century and a half. The most stunning of these achievements was the formulation of the Weinberg-Glashow-Salam Model, more commonly referred to as the standard model of particle physics or just the standard model (SM) [2, 3]. The SM combines the separate forces of nature starting with the theory of electromagnetism formulated by Maxwell in the 19th century [9].

At the dawn of the 20th century Lorentz and Einstein developed relativistic mechanics and showed that electromagnetism was inherently Lorentz invariant [10, 11]. The classical theory of electromagnetism was updated in the 1920s by Dirac to be both a relativistic and quantum mechanical theory [12]. This relativistic treatment predicted anti-particles and explained the source of the spin of fundamental particles. The existence of antimatter was confirmed in 1932 by Anderson when he observed positrons in decays of cosmic rays [13]. In the 1940's the full theory of quantum electrodynamics was completed by Feynman and many others as they formulated a complete and self-consistent theory of electromagnetic interactions [14].

The development of the theory of the weak force started with Pauli and his study of beta decay in the 1930s. He proposed that there existed a very light undetectable

particle, which he called the neutrino, to account for apparent energy non-conservation in beta decay. Fermi shortly thereafter proposed a simple model for beta decay where the neutron decayed through a four fermion contact interaction into a proton, electron, and a neutrino which successfully predicted the decay spectrum of beta decay [15].

In the 1950's a conflict of sorts with the weak force came from the so called $\tau - \theta$ puzzle. Two particles, the τ and the θ, were found which appeared to be identical in every way except for their parity [16]. In 1956, Lee and Yang proposed that the weak force was not parity symmetric and the τ and the θ were in fact a single particle with a parity violating decay [17]. This was confirmed the next year by Wu in cryogenic cobalt-60 decays and others studying meson decays to muons in a storage ring [18, 19]. This confirmation of parity violation lead Marshak and Sudarshan, as well as Feynman and Gell-Mann, to propose the $V-A$ model for the weak interaction which incorporated parity violation into the weak force [20, 21].

Meanwhile, in 1954, Yang and Mills developed a non-Abelian gauge theory to describe the weak force which replaced the contact interaction of Fermi with an interaction mediated by a charged, spin-1 boson [22]. At the time this theory was not of broad interest because it required a massless mediating particle which would have made the weak force a long range force. However, in 1960 Glashow refined the theory proposed by Yang and Mills, adding the $V-A$ model of Marshak and Sudarshan, to combine electromagnetism and the weak force into a single unified theory described by the gauge group of $\mathrm{SU}(2)_L \times \mathrm{U}(1)$ [23]. Here the subscript L refers to the purely left-handed coupling of the weak SU(2) group. In 1967, the next piece of the puzzle, added by Weinberg and Salam [2, 3], was the Brout, Englert, and Higgs (BEH) mechanism [24, 25]. The BEH mechanism gave mass to the vector bosons in Yang-Mills theory and thereby explained the short range nature of the weak force.

The final piece of the SM, developed in parallel, was the addition of the strong force. Since 1947 particles such as the kaon were discovered in cosmic ray events in cloud chambers [26]. Kaons and other particles like them were called "strange" particles because they seemed to posses a common quantum number which other particles did not. These "strange" particles were not thoroughly understood until the 1960s when Gell-Mann and Zweig proposed the quark model based on the $\mathrm{SU}(3)_C$ symmetry group [27, 28]. Their model explained that strongly interacting particles were actually made

up of smaller particles called quarks (at this time, up, down, charm, and strange) and predicted the existence of the yet to be discovered charm quark. The subscript C on the $\mathrm{SU}(3)_C$ group indicates that this group refers to the color charge. Upon experimental confirmation of their prediction of quarks, the SM was expanded to include the full $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)$ symmetry group.

Experimental verification of the SM has been tremendous. The first major test was the indirect observation of the neutral current in at CERN in 1973 [29]. Shortly thereafter the charm quark was discovered at BNL and SLAC and the bottom quark by the E288 Experiment at FNAL [30, 31, 32]. A decade later, in 1983, the UA1 and UA2 experiments operating on the SPS at CERN had the first direct detection of $W^{\pm}$ and Z bosons [33, 34, 35, 36]. The top quark, the final charged fermion of the SM, was discovered in 1995 by the CDF and D0 collaborations at FNAL [37, 38]. The final piece of the SM, the Higgs boson, was discovered by the CMS and ATLAS Collaborations at CERN in 2012 [39, 40]. The discovery of every particle predicted by the SM is not its only achievement as it proves to have unparalleled precision and accuracy in calculation of electroweak measurement results. For example, the muon anomalous magnetic moment can be calculated to predict the gyromagnetic ratio for the muon to an accuracy of one part in 10^{10} and agrees with the measured value to better than one part in 10^9 [41].

2.2 Structure of the Standard Model

The SM describes a self-consistent (if not complete) picture of properties and interactions of matter. The particles of the SM are shown schematically in figure 2.1 [42]. The matter portion of the SM contains spin-1/2 fermions which are split into two categories, quarks and leptons, which are each split into three generations of particles and three generations of their corresponding anti-particles. Quarks are particles which interact through all the SM forces and have non-integer charge ($-\frac{1}{3}e$ or $+\frac{2}{3}e$). Leptons interact only through the weak and electromagnetic forces and have integer charge (0 or $-1e$). Each generation has the same properties and quantum numbers as the previous with the exception of their mass. In each generation, the mass of the particles is greater than the corresponding particle in the previous generation. The quarks and leptons of

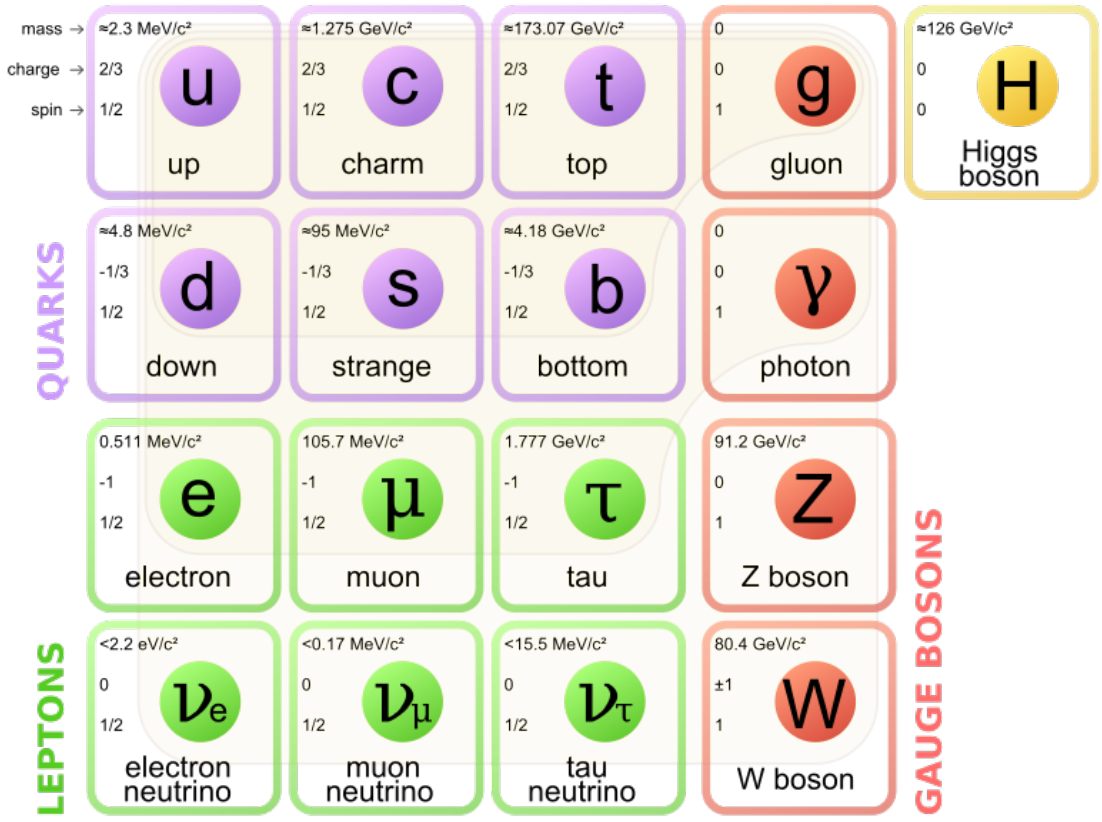


Figure 2.1: The particles of the standard model grouped by type and generation with basic particle information shown for each. The matter particles, quarks and leptons are show on the left, and the gauge bosons, responsible for interaction, and the Higgs boson, responsible for mass, are shown on the right.

each generation are each divided into doublets containing particles which differ in electric charge by $1e$. The exceptions to this rule are the right-handed leptons, for which there is no left-handed neutral component (neutrino) to form a doublet. This lack of a right-handed neutrino results from the built in parity violation of the SM. Each particle described by the standard model has an anti-particle with opposite quantum numbers which may be the same particle for some neutral particles such as the photon.

The matter particles interact with each other through the vector bosons. These are each described with the help of the gauge groups which define the standard model. The strong force is described by the $\mathrm{SU}(3)_C$ group (C for color). The strongly interacting

quarks all carry a quantum number called color which acts as the charge for the strong interaction and can take three charge and three anti-charge values. The three charge values are referred to by the primary colors red, green, and blue. A neutral (or colorless) object can be made by combining all three charge types (or anti-charge types) or by combining a color with its anti-color. The force between color charged particles is carried by the gluons, which are spin-1 vector bosons that each carry a color and an anti-color. This means that there are eight distinct gluons (representations of the eight generators of $\mathrm{SU}(3)_C$). An important feature of the strong force, called confinement, is that only color neutral combinations are allowed. If an object with color charge breaks apart it will create more colored particles to keep each fragment color neutral. This leads to the formation of hadronic "jets," a spray of strongly interacting particles which originates from the initial quark or gluon. Observed strongly bound particles made up of a quark and an anti-quark are called mesons, while those made of three quarks or anti-quarks are called baryons. Both categories are referred to as hadrons.

The weak force and electromagnetism are described by the $\mathrm{SU}(2)_L \times \mathrm{U}(1)$ groups. Both quarks and leptons interact through these forces which are mediated by four vector bosons: the photon and the Z boson, which are neutral, and the $W^{\pm}$ bosons which carry electric charge. These are not the particles described directly by the $\mathrm{SU}(2)_L$ or $\mathrm{U}(1)$ groups, but linear combinations created through the Higgs mechanism. The exact combinations will be discussed along with the BEH mechanism in the next section. An important feature of the SM is that the $\mathrm{SU}(2)_L$ group is strictly left-handed which is an explicit breach of parity. This is manifested as a purely left-handed interaction of the $W^{\pm}$ and the lack of right-handed neutrinos.

The weak force brings a further complication to the SM. The eigenstates of the weak interaction are different from the mass eigenstates for SM fermions. This leads to mixing between the generations of quarks and neutrinos. In the quark sector, this mixing is described by the Cabibbo-Kobayashi-Maskawa (CKM) matrix. The mixing between neutrinos is described by the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix. In addition to describing the generational mixing, the CKM and PMNS matrices each contain a single phase which describes the CP violation in the quark and lepton sectors. While the CKM matrix is largely diagonal, with small off diagonal terms to describe the inter-generational mixing, the PMNS matrix contains large off-diagonal terms with

near maximal mixing between some generations.

2.3 Mass and the BEH Mechanism

The final sector of the standard model is the Higgs sector, described by the Brout-Englert-Higgs (BEH) mechanism. This adds only a single new particle, the Higgs boson, but it has implications for all the other particles of the SM. The basic purpose of the Higgs boson is to generate the masses of the fundamental particles, including itself. Naively it is possible to write a mass term in a Lagrangian for a particle such as

$$\mathcal{L}_m = -\frac{1}{2}m_A^2 A^2 - m_f f^2 \tag{2.1}$$

for a spin-1 particle A_μ with mass m_A and a spin-1/2 particle f with mass m_f. This is perfectly acceptable for the spin-1/2 particle, but in the case of the gauge bosons this term expressly breaks Lorentz invariance. In order to preserve Lorentz invariance in the model, the mass of the particles must be introduced in a more circuitous way. This is accomplished by introducing a spin-0 field with a potential term which provides a continuous but degenerate series of minima and which interacts with the spin-1 particle through a local gauge coupling. The scalar field then develops a non-zero vacuum expectation value (VEV) due to a spontaneous transition to a minima of the potential and this, with the gauge coupling, causes the spin-1 particle to develop a mass. This particle, which is introduced to explain mass, must be a scalar to not break Lorentz invariance itself through its condensate. Through this mechanism the Lagrangian of the model contains only Lorentz invariant terms, while still providing for the mass of each particle. A more detailed explanation of this example can be found in [43] in the context of scalar QED. It is also possible to use the Higgs mechanism to create the mass of the fermion f as well, although it is not required to preserve Lorentz invariance.

In the case of the SM, the situation is a little more complicated because the Abelian U(1) group of QED is replaced with the non-Abelian $\mathrm{SU(2)}_L \times \mathrm{U(1)}$ group of the SM. In this case we have four vector bosons: A_μ^i where $i \in \{0, 1, 2\}$ from the $SU\left(2\right)_L$ group, and B_μ from the U(1) group. If there was no BEH Mechanism these would be the massless gauge bosons that mediate the weak and electromagnetic forces. In reality though, the physical states of the gauge bosons are linear combinations of A_μ^i and B_μ.

In the SM, the Higgs sector is represented by a complex doublet of scalar particles

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}. \tag{2.2}$$

The Lagrangian which describes the Higgs doublet is of the form

$$\mathcal{L}_H = (D_\mu \Phi)^\dagger D^\mu \Phi - V(\Phi) \tag{2.3}$$

where $V(\Phi) \equiv \frac{1}{2}\left(|\Phi|^2 - v^2/2\right)$ is the Higgs potential and $D_\mu \equiv \partial_\mu + i g_L \tau^i A^i_\mu + i\frac{g'}{2} Y B_\mu$. Here v is the VEV of the neutral component of the uncharged Higgs field, giving $\langle\Phi\rangle = \begin{pmatrix} 0 & v/\sqrt{2} \end{pmatrix}$, τ^i are the generators of SU(2), Y is the generator of U(1), and g_L and g' are the couplings for the vector fields A^i_μ and B_μ respectively. Ignoring the terms containing partial derivatives and the potential term in (2.3) the Lagrangian becomes

$$\mathcal{L}_{HK} = \left[\left(i g_L \tau^i A^i_\mu + i\frac{g'}{2} Y B_\mu\right)\Phi\right]^\dagger \left(i g_L \tau^i A^{i\mu} + i\frac{g'}{2} Y B^\mu\right)\Phi. \tag{2.4}$$

Plugging in the explicit forms of the generators τ^i and Y, and substituting $\langle\Phi\rangle$ in place of Φ, 2.4 becomes

$$\mathcal{L}_{HK} = \frac{v^2}{8}\left[g_L^2\left(A^1_\mu + iA^2_\mu\right)\left(A^{1\mu} - iA^{2\mu}\right) + \left(g' B_\mu - g_L A^3_\mu\right)^2\right]. \tag{2.5}$$

If we now make the field redefinitions

$$W^\pm_\mu \equiv \frac{1}{\sqrt{2}}\left(A^1_\mu \pm iA^2_\mu\right),$$

$$Z_\mu \equiv \frac{1}{\sqrt{g'^2 + g_L^2}}\left(g' B_\mu - g_L A^3_\mu\right),\ A_\mu \equiv \frac{1}{\sqrt{g'^2 + g_L^2}}\left(g_L B_\mu + g' A^3_\mu\right)$$

the Lagrangian becomes

$$\mathcal{L}_{HK} = \left(\frac{v g_L}{2}\right)^2 W^+_\mu W^{-\mu} + \frac{1}{2}\left(\frac{v\bar{g}}{2}\right)^2 Z_\mu Z^\mu \tag{2.6}$$

where $\bar{g} \equiv \sqrt{g'^2 + g_L^2}$. Here we can read off the masses of our redefined fields of the form $m^2 W^+_\mu W^{-\mu}$ for a complex field and $\frac{1}{2}m^2 Z_\mu Z^\mu$ for the real field. This gives us $m_W = v g_L/2$, $m_Z = v\bar{g}/2$, and $m_A = 0$. This choice of fields results in the generation of masses for the weak vector bosons, $W^\pm$ and Z, while leaving us with a massless photon, A. Generally, the Z and photon states are written in terms of the Weinberg angle

$$Z_\mu = \sin\theta_W B_\mu - \cos\theta_W A^3_\mu, \qquad A_\mu = \cos\theta_W B_\mu - \sin\theta_W A^3_\mu \tag{2.7}$$

where $\sin\theta_W = g'/\bar{g}$ and $\cos\theta_W = g_L/\bar{g}$.

If we turn our attention to the fermions of the SM we have the option of explaining their mass through the Higgs mechanism, or asserting their masses directly in the Lagrangian of the SM. A effective mass term like that shown in 2.1 results in either case. Taking a closer look at the fermion mass term, $-mf^2$, in the Weyl basis yields

$$\mathcal{L}_D = -m\bar{f}f = -m\chi_L^\dagger\chi_R - m\chi_R^\dagger\chi_L \tag{2.8}$$

where χ_L and χ_R are the two component Weyl spinors representing the left and right-handed components of f where $f = (\chi_L, \chi_R)$. This mass term, called a Dirac mass, contains the product of left and right-handed fields. This is not an issue for the charged fermions because they have both left and right-handed fields. However, by construction the standard model contains no right-handed neutrinos. If neutrinos were really massless this would not be an issue, but there is a large body of evidence, through neutrino oscillation, that demonstrates neutrinos must have a non-zero, if small, mass [44, 45]. A Dirac style mass term is therefore not possible without making some extension to the SM.

Another possibility to explain the mass of the neutrinos exists if they are Majorana particles, meaning that they are their own anti-particles [46]. This is a possibility that is not available to the charged fermions and allows for an effective mass term of the form

$$\mathcal{L}_M = -m_L\chi_L^\dagger\chi_L - m_R\chi_R^\dagger\chi_R. \tag{2.9}$$

A mass term of this form cannot be generated by the minimal Higgs doublet common in the SM, but if such a term is added, either directly to the Lagrangian or through another Higgs boson, there are a few theoretical benefits. These will be described in more detail in section 2.4.

The SM of particle interactions is a resounding success in describing all of the known fundamental interactions of nature, bar gravity. On the other hand, the SM has shortcomings as well. On the purely theoretical, side many theorists are dissatisfied with the way in which parity violation is introduced to the SM as an explicit assumption. From a more experimentally driven side, the SM is also challenged to explain the source and scale of neutrino mass and CP violation. These issues have been addressed in many

ways, both before and after the inception of the SM, but a particularly interesting set of models for LHC physics are the Left Right Symmetric (LRS) extensions to the SM.

2.4 Left-Right Symmetric Models

The dissatisfaction with the SM started before even the model was fully realized with the prediction and discovery of parity violation. No sooner had parity violation been confirmed in the weak decay of cobalt-60 by Wu, than theorists (including Lee and Yang themselves) began looking for ways to restore parity as a fundamental symmetry of the universe. In the 70s this question was addressed by a group of models which I will refer to as Left Right Symmetric extensions to the SM (or simply LRS models).

LRS extensions to the SM were first proposed by Pati and Salam in 1974 [4]. These models extend the SM to restore parity as an exact symmetry which is then broken at some intermediate mass scale. In addition, if neutrinos are Majorana, LRS models can also explain the smallness of the left-handed neutrino's masses through the see-saw mechanism. Similarly, LRS models can provide possible paths for lepton flavor and CP violation with either Majorana or Dirac neutrinos [47].

LRS extensions can come by embedding the SM in a large symmetry group (such as SU(5) or O(10) popular in grand unified theories). Although there are many advantages and disadvantages to embedding the SM in a larger symmetry group, we will limit the discussion here to the simple extension $G_{LR} = \mathrm{SU}(2)_R \times \mathrm{SU}(2)_L \times \mathrm{U}(1)$. This will allow us to explore all relevant features of the LRS models without the complexity of larger models.

The LRS model represented by G_{LR} extends the SM with an additional neutral and two charged vector bosons, Z' and $W_R^{\pm}$ respectively, as well as right handed neutrinos, N_l, to complete the lepton doublet structure. To expressly satisfy the restoration of parity we assume that the left and right coupling constants are equal ($g \equiv g_R = g_L$).

The multiplet structure is

$$Q_{iL,R} = \begin{pmatrix} u_i \\ d_i \end{pmatrix}_{L,R}, \qquad l_{iL} = \begin{pmatrix} e_i \\ \nu_i \end{pmatrix}_{L}, \qquad l_{iR} = \begin{pmatrix} e_i \\ N_i \end{pmatrix}_{R} \tag{2.10}$$

where here the index i stands for the different generations. The "representation content"

of these doublets is

$$Q_{iL} = \left(\frac{1}{2}, 0, \frac{1}{3}\right), \qquad Q_{iR} = \left(0, \frac{1}{2}, \frac{1}{3}\right), \qquad l_{iL} = \left(\frac{1}{2}, 0, -1\right), \qquad l_{iR} = \left(0, \frac{1}{2}, -1\right) \tag{2.11}$$

where the three numbers represent the quantum-numbers in the $\mathrm{SU}(2)_L$, $\mathrm{SU}(2)_R$, and U(1) symmetry groups respectively.

In addition to the new fermions and vector bosons, we must add several new particles to the Higgs sector. As a simple example let us take a toy model with only the following Higgs doublets

$$\chi_{L,R} = \begin{pmatrix} \chi^+_{L,R} \\ \chi^0_{L,R} \end{pmatrix} \tag{2.12}$$

which are responsible to couple the left and right handed sectors independently [48]. In this model we can write the most general Higgs potential consistent with gauge symmetry, parity symmetry, and renormalizability as

$$V = -\mu^2 \left(\chi_R^\dagger \chi_R + \chi_L^\dagger \chi_L\right) + \lambda_1 \left[\left(\chi_R^\dagger \chi_R\right)^2 + \left(\chi_L^\dagger \chi_L\right)^2\right] + \lambda_2 \chi_R^\dagger \chi_R \chi_L^\dagger \chi_L. \tag{2.13}$$

Although this is expressly symmetric under parity, we can find an asymmetric solution for the VEV of the fields such that

$$\left\langle \chi_L^0 \right\rangle = 0 \text{ and } \left\langle \chi_R^0 \right\rangle = U_R. \tag{2.14}$$

The minima in the potential are found as

$$\frac{\partial V}{\partial \chi_L^\dagger} = \left(-\mu^2 + 2\lambda_1 \chi_L^\dagger \chi_L + \lambda_2 \chi_R^\dagger \chi_R\right) \chi_L = 0 \tag{2.15}$$

$$\frac{\partial V}{\partial \chi_R^\dagger} = \left(-\mu^2 + 2\lambda_1 \chi_R^\dagger \chi_R + \lambda_2 \chi_L^\dagger \chi_L\right) \chi_R = 0 \tag{2.16}$$

Combining these with the desired VEV, (2.14), we get the constraint $\mu^2 = 2\lambda_1 U_R^2$. To confirm that this is indeed a minimum we can look at the second derivatives

$$\frac{\partial^2 V}{\partial \chi_L^\dagger \partial \chi_L} = -\mu^2 + 4\lambda_1 \chi_L^\dagger \chi_L + \lambda_2 \chi_R^\dagger \chi_R > 0 \tag{2.17}$$

$$\frac{\partial^2 V}{\partial \chi_R^\dagger \partial \chi_R} = -\mu^2 + 4\lambda_1 \chi_R^\dagger \chi_R + \lambda_2 \chi_L^\dagger \chi_L > 0 \tag{2.18}$$

This leads to the conditions $\lambda_2 > 2\lambda_1$. Not only is this a minimum but it is in fact the absolute minima as is shown in [49].

To find the masses under these assumptions we need to examine the Lagrangian

$$\mathcal{L}_H = \frac{1}{2}\left(D_\mu \chi_L\right)^\dagger D^\mu \chi_L + \frac{1}{2}\left(D_\mu \chi_R\right)^\dagger D^\mu \chi_R - V \tag{2.19}$$

where we have $D_\mu = \partial_\mu + ig_L \tau^i W^i_{L\mu} + ig_R \tau^i W^i_{R\mu} + i\frac{g'}{2} Y B_\mu$. Here g_L and g_R are the SU(2) gauge couplings (which are equal in this model and will be called $g \equiv g_R = g_L$) and g' is the $U(1)$ gauge coupling constant.

The relevant parts of the above Lagrangian are the terms containing W_R, B, and χ_R. Left-handed terms will disappear when acting on the right handed singlet χ_R and terms containing χ_L are uninteresting as it develops no VEV. These interesting terms give us the Yukawa term of the Lagrangian

$$\mathcal{L}_Y = \frac{1}{2}\left[\left(g\tau^i W^i_{L\mu} + g\tau^i W^i_{R\mu} + i\frac{g'}{2} Y B_\mu\right)\chi_R\right]^\dagger \left(g\tau^j W^{j\mu}_L + g\tau^j W^{j\mu}_R + i\frac{g'}{2} Y B_\mu\right)\chi_R \tag{2.20}$$

Plugging in (2.14) and diagonalizing the Lagrangian we find the Yukawa term of the Lagrangian after symmetry breaking to be

$$\mathcal{L}_Y = \left(\frac{1}{4}U_R^2 g^2\right) W_R^+ W_R^- + \frac{1}{2}\left(\frac{1}{4}U_R^2\left(g^2 + g'^2\right)\right) Z'_\mu Z'^\mu \tag{2.21}$$

where we have, similarly to the left-handed case, $W_R^\pm = \frac{1}{\sqrt{2}}\left(W_R^1 \mp iW_R^2\right)$ and $Z' = \left(g^2 + g'^2\right)^{-1/2}\left(gW^3_{R\mu} - g'B_\mu\right)$. From here we can see W_R develops a mass of $m_{W_R} = \frac{1}{2}gU_R$ and Z' develops a mass of $m_{Z'} = \frac{1}{2}U_R\sqrt{g^2 + g'^2}$ while all other vector bosons will remain massless.

The mass splitting between the massive right-handed and the massless left-handed vector bosons demonstrates the spontaneous breaking of left-right symmetry. Here we have demonstrated that parity can be spontaneously broken in a model with an expressly LRS Lagrangian. This is unfortunately an unrealistic model. The left-handed sector remains massless and the SM $\mathrm{SU(2)}_L \times \mathrm{U(1)}$ symmetry is unbroken.

It is possible to make this model more realistic by adding the analog to the SM Higgs doublet, a bi-doublet which interacts with both the left and right-handed gauge groups [48, 49]. This multiplet would take the form

$$\Phi = \begin{pmatrix} \phi_1^0 & \phi_2^+ \\ \phi_1^- & \phi_2^0 \end{pmatrix}, \qquad \langle\Phi\rangle = \begin{pmatrix} v_1 & 0 \\ 0 & v_2 \end{pmatrix}. \tag{2.22}$$

This additional multiplet will provide the analog of the SM Higgs and break the $\mathrm{SU}(2)_L$ symmetry. This leads to the following masses for the vector bosons

$$m_{W_L} = \frac{1}{2}gv,\ m_{W_R} \simeq \frac{1}{2}gU_R,\ m_Z = \frac{1}{2}v\bar{g},\ m_{Z'} \simeq \frac{1}{2}U_R\bar{g}, \qquad v^2 \equiv v_1^2 + v_2^2,\ \bar{g}^2 \equiv g^2 + g'^2 \tag{2.23}$$

where we have assumed $U_R \gg v$. We are also neglecting possible left-right mixing which must be small as we will discuss later. This addition leads to a model in which both $\mathrm{SU}(2)_L$ and $\mathrm{SU}(2)_R$ are broken, and the right-handed more strongly than the left-handed gauge group. Under the assumptions of strict LR symmetry made here it is also clear from 2.23 that the right-handed vector bosons would obey the same mass relation as their left-handed counterparts, $m_{W_R} = \cos\theta_R m_{Z'}$, where $\cos\theta_R \equiv g/\bar{g}$ is the equivalent to the Weinberg angle in the SM.

The final piece the LRS models have to offer is an explanation for the non-vanishing but small masses of the left-handed neutrinos. As discussed in the previous section, the lack of right-handed neutrinos precludes a Dirac-type mass term for the neutrinos. This leads to certain theoretical challenges explaining the mass of neutrinos without raising concerns over fine tuning of the mass difference and explaining why the neutrinos are the only fermions without a Dirac mass component. These issues are both solved with the additional right-handed neutrinos introduced in LRS models through the see-saw mechanism [50, 6].

The generic see-saw mechanism involves a Lagrangian containing both Dirac and Majorana mass terms of the form

$$\mathcal{L} = \frac{1}{2}\begin{pmatrix} \bar{\nu}_{Li} & \bar{\nu}_{Ri} \end{pmatrix}\begin{pmatrix} B_i' & M_i \\ M_i & B_i \end{pmatrix}\begin{pmatrix} \nu_{Li} \\ \nu_{Ri} \end{pmatrix} \tag{2.24}$$

where i denotes the lepton generation and ν_L and ν_R are the pure left and right-handed neutrino spinors. In the LRS model considered above, M terms, the Dirac mass, would arise from the VEV of the field Φ, and B' and B, the Majorana masses, would come from the VEV of χ_L and χ_R respectively. This leads to $B' \sim 0$, $B \sim U_R$, and $M \sim v$ which give the condition $B \gg M$ (as we know $m_{W_R} \gg m_{W_L}$). Using these assumptions we can diagonalize the mass matrix in equation (2.24) to find the mass eigenvalues, ignoring mixing between generations, of

$$\lambda_+ \simeq B, \qquad \lambda_- \simeq -\frac{M^2}{B} \tag{2.25}$$

This gives a very heavy mass state λ_+ and a very light state λ_-. The measurable mass states can then be formed as follows

$$\nu \simeq \frac{1}{\sqrt{M^2+B^2}}\left(B\nu_L - M\nu_R\right) \simeq \nu_L - \frac{M}{B}\nu_R \tag{2.26}$$

with mass eigenvalue λ_- and

$$N \simeq \frac{1}{\sqrt{M^2+B^2}}\left(M\nu_L + B\nu_R\right) \simeq \nu_R + \frac{M}{B}\nu_L \tag{2.27}$$

with mass eigenvalue λ_+. This naturally gives rise to a very light left-handed neutrino and a very heavy right-handed neutrino while requiring no fine tuning of the scalar couplings compared to the changed fermions. Furthermore, the resulting mass states are nearly pure states in the left-right parity space as well. This will lead to very small left-right mixing which is well supported by experiment. Limits on left-right mixing are presented in section 2.5.

The LRS model discussed above is simply an illustration of the possible LRS extensions to the SM and several properties of this model were chosen to show the interesting features in a simplified environment. For instance, we choose to assume that $g_L = g_R$ as a general rule, but this need be true only above the left-right breaking scale. As each coupling constant evolves down from the breaking scale, they are only related up to an additive constant [48]. Although this assumption is made to simplify the calculation here, it will be used to set limit contours as some assumption must be made about signal strength. This model is also too naive to explain the matter/antimatter asymmetry. This can be accomplished with more complex models such as that presented in [5]. Furthermore, LRS models can also be cast in terms of charge parity as the spontaneously broken symmetry, in place of spatial parity. The choice of spatial parity is motivated from the historical perspective and because it leads to a slightly simpler model. Use of charge parity will not affect any of the qualitative features of the model, but it can change some numerical assumptions when setting indirect limits. The choice of charge parity as the spontaneously broken symmetry also lends itself to being embedded into larger symmetry groups more readily than spatial parity [51].

2.5 Phenomenology

Although LRS models form an elegant theoretical framework, they are not yet predictive. They contain no theoretical constraints on which mass range we will find the particles of the right-handed sector. Furthermore, we are missing many of the ingredients required to perform detailed calculations in an LRS model. One of these is the true value of g_R. We also have no hint about the contents of the right-handed PMNS matrix and only weak constraints on the right-handed CKM matrix. This also means that indirect search experiments can at best guide the direct search experiments. To gain quantitative predictive power in these models will require direct measurements of the model parameters [51].

LRS models present several signatures for potential direct detection experiments. The heavy vector bosons can be detected directly, or the neutrinos may also be detected. The clearest signature which would suggest the possibility of a LRS model's reality is the decay $Z' \to l\bar{l}$. This decay gives a very clean signature with two high energy leptons which can be reconstructed to give a resonant peak at the Z' mass. This channel suffers though because the larger mass of the Z' compared to the W_R will mean it may be undetectable while a W_R is within reach at LHC energies. Currently the CMS collaboration has placed limits on a SM like Z' to be heavier than 2.96 TeV [52].

Direct searches for W_R are possible in both the hadronic and leptonic channels, each with their own challenges. The "single lepton" decay channel, $W' \to l\nu$, is a powerful search tool, but inherently assumes a predominantly light, stable ν_R and is therefore not compatible with this search. The hadronic decay channel, $W_R \to q\bar{q}'$, can be used to set limits on LRS models and is a valuable channel due to having minimal dependence on heavy neutrino masses, although its mass reach is limited by very high jet and QCD backgrounds. The large background can be partially alleviated by using the sub-channel $W_R \to t\bar{b}$ where the t quark subsequently decays to a leptonic final state, but is still limited in mass reach depending on the heavy neutrino masses. A CMS analysis has used this channel to set a lower bound of $M_{W_R} > 2.03\,\text{TeV}$ [53, 54]. It is also possible to do a direct detection search for the heavy right-handed Higgs boson which in realistic models can have a doubly charged component which would give a very noticeable decay mode to two same sign leptons, but generally the mass of this Higgs would be too heavy

for direct detection, even if W_R or Z' is within the LHC energy reach [55].

Each of the previously discussed channels can provide evidence for an LRS model, but these search channels do not provide conclusive evidence given that a large number of models predict new heavy spin-1 vector bosons. Another tactic is to search directly for the heavy neutrino. This can be accomplished through production and subsequent decay of a Z' or W_R into a right-handed neutrino. As before, the decay $Z' \to N\bar{N}$ will suffer from smaller production cross-section of Z' over W_R as well as a reduced branching fraction to neutrinos if they are a significant fraction of the mass of Z'. This leads to the most effective search channel, when $M_N < M_{W_R}$, being $W_R \to l\bar{N}_l$. In this decay the neutrino then decays through a virtual W_R^* as $N_l \to lq\bar{q}'$, shown in figure 2.2. This leads to a final state which includes two same flavor leptons and two quarks which will hadronize into jets. This final state contains two high energy leptons and is therefore easy to identify and it allows for complete reconstruction of the invariant mass of the W_R. The search involving an intermediate state of a W_R, instead of a Z', does not suffer the same phase space suppression in the branching fraction when the mass of N approaches half the mass of W_R because only one heavy neutrino need be produced.

In addition to the possible direct search paths, there are many indirect search paths to look for an LRS model. These include neutrinoless double beta decay, lepton flavor violation, left-right mixing, electron electric dipole moment, anomalous mixing of neutral mesons, and neutral meson mass differences [47, 55]. Each of these indirect measurements can give constraints on the model, but none is sufficient for proof of discovery. The most stringent limits on the mass of W_R come from mass and mixing measurements in neutral meson systems. In most cases, the constituents of an LRS model would provide additional CP violating terms.

In the case of $K_S - K_L$ mass difference, CP violation arises from terms allowing $K_0 \leftrightarrow \bar{K}_0$ mixing. The mechanism for $K_0 \leftrightarrow \bar{K}_0$ mixing is the box diagrams, shown in figure 2.3 [56], whose amplitude will involve the CP violating phase in the CKM matrix. In an LRS model it is possible to replace one or both W_L bosons with a W_R boson which will add additional terms and increase the rate of CP violation. The box diagram which contains both a W_R and a W_L is in fact highly enhanced over the SM term on the order of 1000-1500 times but supressed by the mass of the W_R. Using the measurements of this process it is possible to set limits of $M_{W_R} > 2.5\,\text{TeV}$ in the case

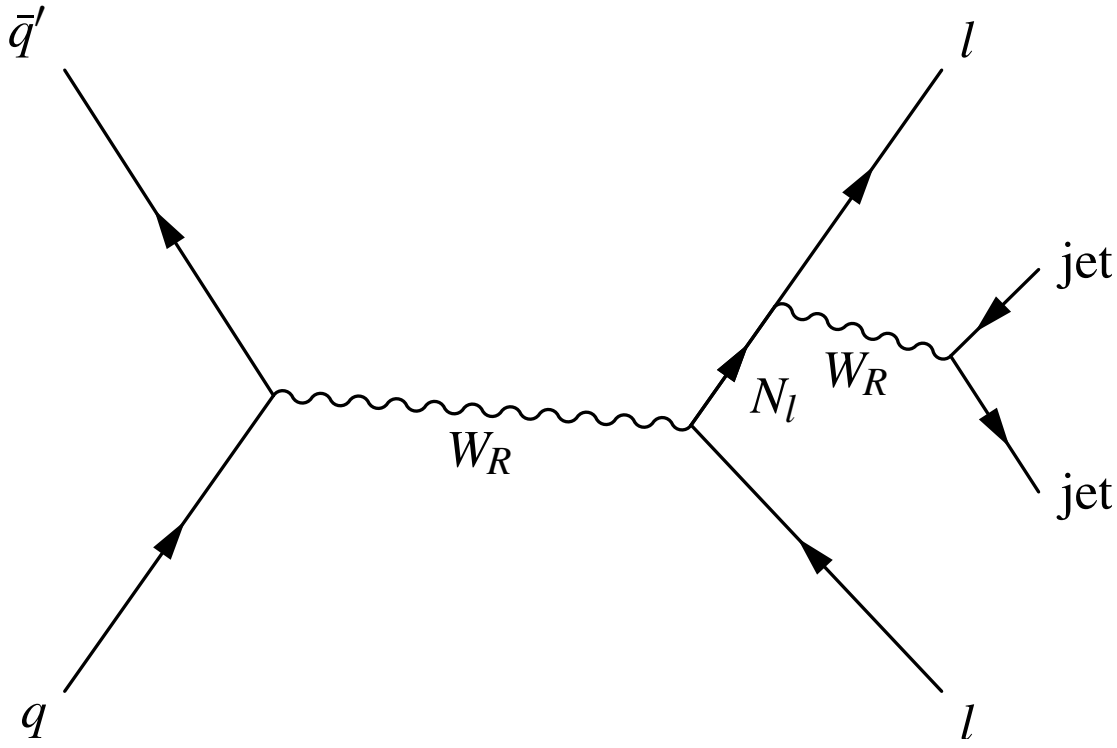


Figure 2.2: The primary production for a W_R boson is quark anti-quark fusion. The W_R decay of interest is then through the right-handed heavy neutrino, N, which results in a final state with two same flavor leptons and two jets.

of spatial parity breaking and $M_{W_R} > 2.3\,\text{TeV}$ for charge parity [55].

Similarly, limits can be set based on the direct weak CP violation in kaon systems. The direct CP violation in $\Delta S = 1$ strangeness violating transitions can be parameterized as

$$\epsilon' = \frac{i}{\sqrt{2}} \frac{A_2}{A_0} \left(\frac{\text{Im}A_2}{\text{Re}A_2} - \frac{\text{Im}A_0}{\text{Re}A_0} \right) \frac{q}{p} e^{i(\delta_2 - \delta_0)} \tag{2.28}$$

where p and q are the $K_0 \leftrightarrow \bar{K}_0$ mixing parameters, A_I are the isospin 0 and 2 decay amplitudes for $K \to \pi\pi$, and δ_I and the strong phases of of $\pi\pi$ scattering. Although the SM contributions to this amplitude are difficult loop corrections, LRS models give tree level contributions to this term. In the case of spatial parity as the symmetry of choice a limit of $M_{W_R} > 3.1\,\text{TeV}$ can be set. In the case of charge parity on the other hand the freedom to choose additional phases in the right-handed CKM matrices allow no limit to be set [55].

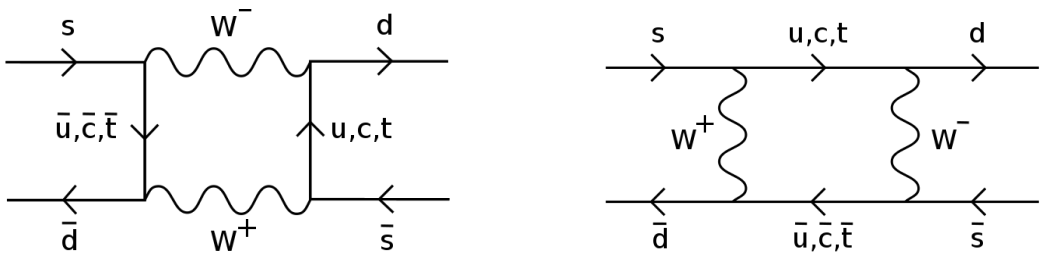


Figure 2.3: Neutral kaon mixing proceeds through these two box diagrams in the SM.

It is not clear a-priori whether spatial or charge parity is the preferred symmetry for LRS models, therefore the limit which is valid for a search is the less stringent of the two possible parity choices. This gives a final limit of $M_{W_R} > 2.5\,\text{TeV}$. This limit itself is subject to large theoretical error, but it is valid to support a W_R consistent with an LRS model above the 2 to 3 TeV range.

It is also possible to use the B_d and B_s systems in a similar way to constrain possible LRS model parameters. In these systems though, there was a slightly more interesting case because of the $\sim 3\sigma$ discrepancy in the B sector forward-back asymmetry CP violation measured by the D0 Collaboration [57]. An LRS model can possibly explain this discrepancy. The authors of [55] show that in the case of charge parity there is a region of space in M_{W_R} and M_{H_R} (where M_{H_R} is the right-handed Higgs mass) which would resolve this discrepancy with M_{W_R} in the reach of the LHC, with a mass of roughly 2 to 4 TeV. Eventually this constraint imposes a weak upper bound on M_{W_R} as it would be forced above M_{H_R} and the LRS models become less appealing than other models which may also offer solutions to this discrepancy. However, more recent measurements by the LHCb experiment have found no discrepancy between B sector CP violation and SM predictions [58]. Furthermore, the latest LHCb data combined with new theoretical consideration of the roll of the flavor changing neutral Higgs has increased the lower limit from B meson mixing to $M_{W_R} > 3\,\text{TeV}$ [59].

Mass constraints can also be placed on the right handed neutrinos themselves. The best direct limit is placed by measurements taken at LEP by the L3 Collaboration. They conducted a search for pair-production of neutral leptons which would in turn

decay to a charged lepton (e or μ) and a W. The W was allowed to decay to hadrons. They placed a 95% CL lower bound of 89.5 GeV on the mass of any unstable e or μ flavor neutral lepton [60].

The primary production mechanism for W_R at the LHC is $q\bar{q}' \to W_R$. The production and decay chain is shown in figure 2.2. A Majorana mass for N is not assumed a-priori. To study this production mode and the following decay through a right-handed neutrino Monte Carlo (MC) was generated using the PYTHIA event generator [61]. These samples were generated with several assumptions, but were kept as general as possible. The events were generated assuming strict left-right symmetry. This included the assumption that $g_L = g_R$, the CKM matrices are equal, right-handed neutrino mixing is negligible, and left-right mixing angles are small. The assumption that the left-right mixing is negligible is a good assumption given the measured constraints. The $W_R \leftrightarrow W_L$ mixing angle is constrained by measurements of neutrinoless double beta decay to be $\tan\zeta < 4.7 \times 10^{-3} \left(\left\langle m_N^V \right\rangle / (1\,\text{TeV})\right)^{1/2}$ [62]. Here $\left\langle m_N^V \right\rangle^{-1} = \sum_j V_{ej}^2 m_j^{-1}$ is the effective mass of the right-handed electron-flavor neutrino, where V_{ej} are the electron flavor terms in the right-handed PMNS matrix and m_j represent the right-handed mass eigenvalues.

The remaining assumptions are not as easy to justify, and in general may change the overall structure of the model. In the case of W_R production and subsequent decay to $q\bar{q}' \to llq\bar{q}'$, the overall cross-section and branching ratios will be affected by the scale of g_R and the right-handed CKM matrix. These theoretical uncertainties affect the overall rate directly and thus can be accounted for with a scale factor on the theoretical expectation. Possible differences between the left and right-handed PMNS matrices can cause larger complications. If inter-generational mixing between the right-handed neutrinos is small (i.e. the right-handed PMNS matrix is approximately diagonal) these effects can be ignored. If the neutrino is produced close to rest, then the mixing probability will be modulated by $\sin^2(\Delta m/2\Gamma)$ where Δm is the mass difference between neutrinos and Γ is the full width of the decaying neutrino. If Δm is small compared to Γ then the mixing will be negligible, however if the mass difference is very large ($\Delta m/2\Gamma \gg \pi/2$) then the mixing will be maximal. If the PMNS mixing elements are large between generations and the mass difference between the neutrinos is sufficiently large, this would cause the signal to be spread between multiple final states including states with different flavor

	Electroweak Standard Model	Left-Right Symmetric Model
Gauge Group	$\mathrm{SU(2)}_R \times \mathrm{U(1)}$	$\mathrm{SU(2)}_R \times \mathrm{SU(2)}_L \times \mathrm{U(1)}$
Fermions	$Q_L = \left(u^i, d^i\right)_L$, $L_L = \left(l^i, \nu^i\right)_L$ $Q_R = \left(u^i, d^i\right)_R$, $L_R = l^i_R$	$Q_L = \left(u^i, d^i\right)_L$, $L_L = \left(l^i, \nu^i\right)_L$ $Q_R = \left(u^i, d^i\right)_R$, $L_R = \left(l^i, N^i\right)_R$
Gauge Bosons	$W_L^{\pm}$, Z, γ	$W_L^{\pm}$, $W_R^{\pm}$, Z, Z', γ
Higgs Sector	$\Phi = \begin{pmatrix} \phi^0 \\ \phi^- \end{pmatrix}$	$\Phi = \begin{pmatrix} \phi_1^0 & \phi_2^+ \\ \phi_1^- & \phi_2^0 \end{pmatrix}$, $\chi_{L,R} = \begin{pmatrix} \chi_{L,R}^+ \\ \chi_{L,R}^0 \end{pmatrix}$

Table 2.1: Summary of electroweak portion of the SM and the basic LRS model.

leptons. Furthermore, for some neutrino masses, the oscillation rate may be energy dependent and would modify the event kinematics. Although the results presented in this thesis are applicable for a wide range of mixing parameters, the background estimation may loose their validity for the case of strong mixing. Particularly, this is true for the top pair production background estimated from data.

In addition to being the easiest channel to directly produce right-handed neutrinos, this channel provides many useful pieces of information in the case of a discovery. The invariant mass of both N and W_R can be reconstructed completely without any missing energy. Furthermore, if a significant detection is made, the nature of the neutrino mass will be easy to determine by comparing the same charge lepton rate to the opposite charge lepton rate. Additionally, measurements of several right-handed CKM and PMNS matrix elements could be made.

2.6 Left-Right Symmetric Model Summary

Table 2.1 summarizes the differences between the electroweak portion of the SM and the basic LRS model based on the $\mathrm{SU(2)}_R \times \mathrm{SU(2)}_L \times \mathrm{U(1)}$ symmetry group.

Chapter 3

The CMS Experiment

This chapter describes the experiment, the CMS detector, and the facility at which it operates, the LHC, at CERN. The summary that follows in this chapter describes in detail those features of the CMS detector crucial to the search for W_R presented herein, a full description of the CMS Detector may be found in [63].

Making use of the proton-proton collisions at the LHC's four interaction points are the four main experiments: ATLAS, CMS, LHCb, and ALICE. ATLAS and CMS are large scale general purpose detectors which are designed to carry out a wide program of physics searches and measurements. LHCb is a specialized detector designed to study b quark physics, while ALICE is primarily designed to study heavy ion collisions.

3.1 The Large Hadron Collider

The LHC is a hadron collider built near Geneva, Switzerland spanning the French and Swiss border. In its most recent data taking period in 2012, the LHC ran with a center of mass energy of 8 TeV, with a planned upgrade to 13 TeV in 2015. This makes it the most powerful particle accelerator built to date. The collider itself is a ring with a 27 km circumference [64].

The journey of a proton to the LHC is not a simple one. The entire CERN accelerator complex is summarized in figure 3.1 [65]. Bunches of protons start by being accelerated by a linear accelerator, LINAC 2, and are then injected into the Proton Synchrotron (PS). The PS boosts the protons to an energy of 26 GeV and then injects the protons

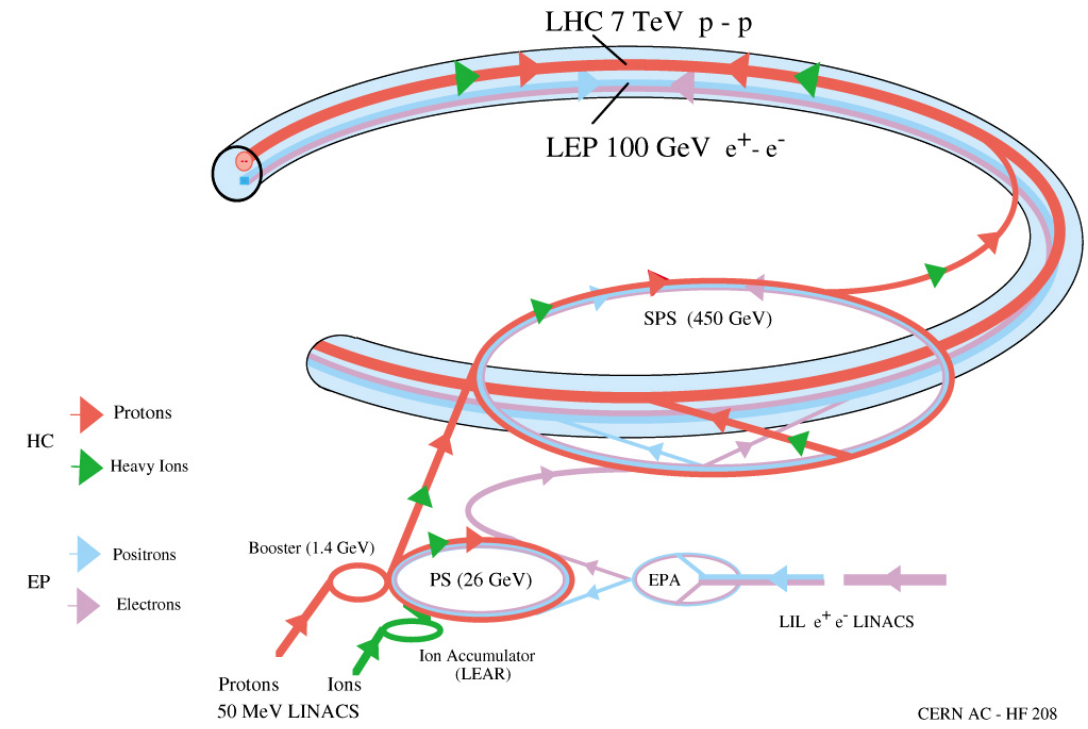


Figure 3.1: The LHC and its injector chain.

into the Super Proton Synchrotron (SPS). The SPS collects the proton bunches from the PS and further boosts their energy to 450 GeV. At this point, the protons are injected into the LHC itself. Once the desired number of proton bunches have been injected into the LHC, the final acceleration to 4 TeV per beam and focusing of the beams occurs before collisions take place. Once the protons have reached the desired energy, they are then steered such that the two beams collide at four points around the LHC ring for each of the four experiments.

The LHC guides the proton bunches around the ring with the use of dipole magnets. At 8 TeV center-of-mass energy, the dipole magnets operate at a field strength of ~7.5 T. In addition to the dipole magnets, quadrupoles and higher order magnets are used to keep the beams focused and the proton bunches together. The protons are accelerated by 16 superconducting RF cavities (eight per beam) which operate at a frequency of

400 MHz. The RF cavities serve to slow fast-moving and accelerate lagging protons to keep the bunches tightly packed in physical and momentum space. The LHC is designed to operate with a proton bunch being accelerated every ten cycles of the RF frequency. This means for the 27 km ring there is space for ~3600 bunches. However, after accounting for additional spacing needed to inject and extract the beams from the machine, it can accommodate 2808 bunches per beam.

The LHC was designed to not only have the highest center of mass energy of any collider built to date, but also to have the highest instantaneous luminosity for a hadron collider. Luminosity, related linearly to the number of particle interactions in a given time, can be expressed as

$$\mathcal{L} = \frac{fnN^2\gamma}{4\pi\epsilon_n\beta^*}F. \tag{3.1}$$

Here f is the frequency of interaction for a particular bunch, n the number of bunches per beam, N the number of protons per bunch, γ is the Lorentz boost factor, ϵ_n is the beam emittance, β^* is the beta function, and F is a geometric factor based on the angle at which the beams cross. The mean emittance and beta function describe how focused the beams are at the interaction point. Their product is equal to σ^2, where σ is the cross-sectional area of each beam at the interaction point. These parameters along with other relevant information can be found in table 3.1 [66]. In this formula f is fixed by the circumference of the machine, but n, N, and σ can be manipulated to optimize the luminosity. In addition to monetary and technological design constraints, which limit the luminosity, there are also a balance of physics goals which serve to limit the luminosity. The luminosity increases as the bunches have more protons added to them (higher N), and as they are squeezed more more tightly (smaller σ), but this eventually leads to many proton-proton scattering events per bunch crossing which causes additional background and makes data analysis more difficult. Certain measurements may even become impossible. This problem is avoided by increasing n, but more bunches means collisions are close together in time, and this leads to its own challenges. For the data taken in 2012, the solution to maximize the luminosity was to squeeze the beams as much as possible, but only fill every other possible bunch in each beam. This leads to bunches spaced 50 ns apart as opposed to the design spacing of 25 ns. The primary factor behind this decision was technological as the overall luminosity would have been lower if every bunch was filled due to difficulty with injecting sufficiently high population

Parameter	Value in 2012	Design value
Beam energy [TeV]	4	7
β^* [m]	0.6	0.55
Bunch spacing [ns]	50	25
Number of bunches	1374	2808
protons per bunch	1.6 - 1.7$\times 10^{11}$	1.15$\times 10^{11}$
Normalized emittance at start of fill [mm.mrad]	2.5	3.75
Peak luminosity [cm^{-2}s^{-1}]	7.7$\times 10^{33}$	1$\times 10^{34}$
Max. mean number of events per bunch crossing	$\approx$40	19
Stored beam energy [MJ]	$\approx$140	362

Table 3.1: LHC operation parameters for 2012 compared to the original design parameters.

bunches at LINAC 2.

3.2 CMS Overview

The CMS detector is a general purpose detector which is designed to make a wide range of measurements. The design goal of the CMS detector is to allow for the measurement and identification of as many particles with the highest efficiency and acceptance which is feasible with the given budget and technological constraints. In this context, acceptance is defined as the active volume of the detector, both its coverage in physical space as well as its limits in momentum or energy space. Efficiency is then defined as the probability that a particle falling into the acceptance region of the detector will be properly identified and reconstructed. These constraints lead to detectors for collider physics being constructed in the shape of finite nested cylinders with each part, or sub-detector, forming a different layer. The proton beams enter the detector through gaps in the center of the ends of the cylinder. To ensure maximum coverage for particles leaving the proton-proton interaction region, the end plates of the cylinders, called endcaps, and the central portion of the cylinder, called the barrel, overlap. A cutaway schematic of CMS is shown in figure 3.2 [67].

Before going into the details of the CMS detector it is important to first understand the coordinate system adopted by CMS and hadron collider experiments in general. The x direction is defined to point toward the center of the LHC ring, the y axis up, and

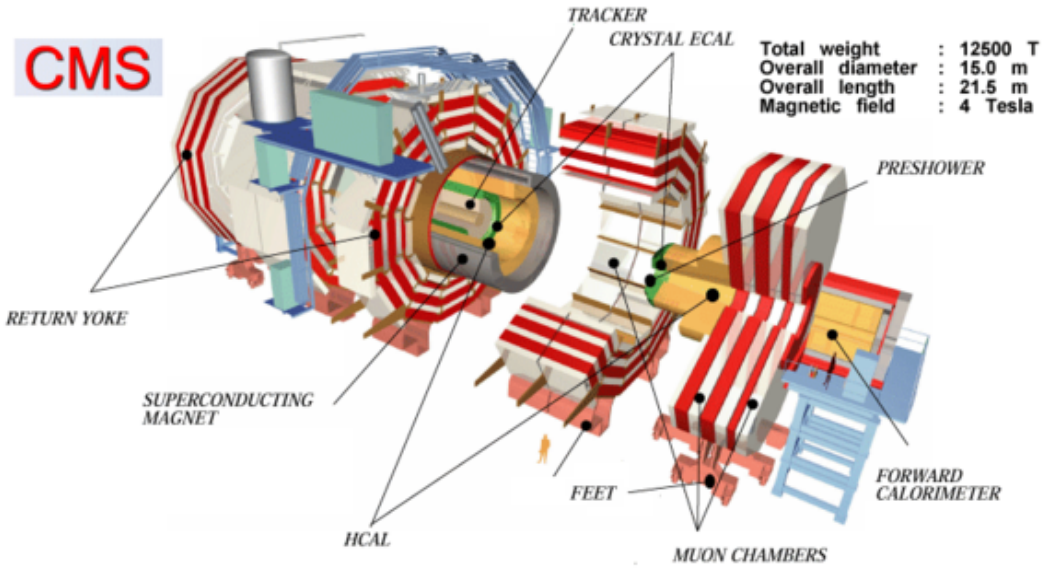


Figure 3.2: A cut-away view of CMS.

the z axis counter-clockwise along the beamline such that a right handed coordinate system is created. The origin is designated to be the center of the detector at the nominal interaction point for proton-proton collisions. CMS uses a cylindrical coordinate system with the coordinates (η, ϕ). Here ϕ is the azimuthal angle in the $x - y$ plane and η is the pseudorapidity. Pseudorapidity is defined as

$$\eta \equiv -\ln \tan \frac{\theta}{2} \tag{3.2}$$

where θ is the polar angle with respect to the z axis.

Pseudorapidity is related to the rapidity, Y, and is preferred over the polar angle for several reasons. The rapidity as defined for use in particle physics is

$$Y \equiv \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) \tag{3.3}$$

where E is the energy of a particle and p_z is its momentum along the beam direction (note this differs from the usual relativistic definition). It is straightforward to show that if two particles come from the same interaction, the difference in their rapidity is invariant of the boost of the initial state. This is a very desirable feature as the boost along the beam direction is not well defined in a hadron collider because it is

impossible to know the exact momentum of the interacting partons. Pseudorapidity is then the rapidity assuming a massless particle. This not only simplifies the calculation, as only the polar angle is required, but at the large energy scales of the LHC it is a good approximation that all decay products reaching the detector are massless.

Similar to the preference for η over θ, it is uncommon to discuss energy and momentum directly. It is more common, and more useful for hadron physics, to use transverse momentum $p_{\mathrm{T}} \equiv |p| \,/\, \cosh \eta$, defined as the component of the momentum perpendicular to the incoming proton beams. The transverse energy is then defined as $E_{\mathrm{T}} \equiv E / \cosh \eta$. These quantities remove the dependence on the unknown momentum along the beam line in the initial state of the parton collision. The quantity $\Delta R \equiv \sqrt{\Delta\eta^2 + \Delta\phi^2}$ is also heavily used as a measure of the distance between two objects in angular space.

The CMS detector is comprised of several major components, called sub-detectors, including the tracker, electromagnetic calorimeter (ECAL), hadronic calorimeter (HCAL), muon chambers, and trigger. The tracker is used to detect the path and momentum of charged particles coming from the proton collisions. The ECAL and HCAL are designed to measure the energy of electrons, positrons, photons, and strongly interacting particles respectively. The tracker, ECAL, and HCAL are situated inside the solenoid magnet which provides a 3.8 T magnetic field parallel to the beam line. The muon chambers are designed to detect the path and momentum of muons which are the only charged particles to pass through HCAL and the magnet. The muon chambers are situated outside the magnet in between layers of iron return yoke which gathers the return field of the solenoid. The final central component of CMS is the trigger which is responsible for quickly looking at data from each collision and deciding if that event should be recorded or discarded.

3.3 The Magnet

Central to the design and construction of CMS is the solenoid magnet. A key goal of CMS is the accurate reconstruction of the p_{T} of charged particles. While the energy of electrons can be measured with calorimetric techniques, muons are minimum ionizing particles at the energies encountered at the LHC and will generally pass through the entire detector with minimal energy deposition. This necessitates the use of a strong

magnetic field to deflect charged particles and measure their momentum through the curvature of their paths. A magnetic field also allows the sign of the charge of a particle to be determined using the direction of the deflection. For both applications, the larger the field, the better the determination can be made.

The CMS magnetic field is provided by a single solenoid 12.9 m long and 5.9 m in diameter. The solenoid provides a uniform 3.8 T magnetic field for the detector components inside it. The fringe field outside the magnet is then gathered by the iron return yoke where it is used to further bend the muons as they pass between the muon chambers without requiring a second magnet for this purpose. The magnet is constructed of a coil of four superconducting wires, carrying 20 kA, which are encased in aluminum to spread heat and add strength. This assembly is encased in a stainless steel vessel which serves as a cryostat and as structural support. The cryostat must be very strong as it must support both the huge magnetic forces as well as the weight of all the detector components inside the magnet.

The solenoid is large enough to contain not only the tracker, but the calorimeters as well. This design was chosen to minimize the un-instrumented mass between the interaction point and the calorimeters where particles can lose energy. Additionally, this design helps to optimize the bending power on the muons. Placing the calorimeters inside the magnetic field introduces several design challenges and benefits which will be discussed in more detail in the following sections.

3.4 The Tracker

The primary purposes of the tracker in CMS are to measure the momentum of charged particles and to identify the locations of particle interaction vertices. The latter task is accomplished primarily by the silicon pixel detector while the momentum determination is handled primarily by the silicon strip tracker. Together these components form the central tracking system for CMS. Both detectors use silicon measurement technologies to measure the ionization of charged particles passing through them. The tracker is designed to have coverage for all charged particles with $|\eta| < 2.5$.

The silicon detectors operate by collecting the ionization charge deposited as charged particles pass through the active silicon component of the tracker. The ionization charge

is collected using a modest (~100 V) electric field. The charge is collected by anodes on the surface of the silicon and measured with charge collection devices which are attached to the active silicon region of the detector. Silicon detectors were chosen for the CMS tracker because they are capable of high granularity, are radiation hard, and have a fast recovery time. All three considerations are particularly important in the tracker as it is the closest sub-detector to the interactions and therefore receives the highest density of particles.

The pixel tracker is the innermost portion of CMS. It is located as close to the beam pipe as possible, the inner layer being only 4.4 cm from the nominal interaction point. The pixel tracker has two additional layers located 7.3 cm and 10.2 cm from the interaction point. These layers, called the barrel, are shaped as nested cylindrical shells 53 cm long. In addition to the barrel, there are also two disk shaped sections on either end of the barrel in order to cover higher η. The layout of the pixel detector is shown in figure 3.3.

The pixel detector is constructed of individual rectangular pixel modules which contain a grid of pixels. The modules are then assembled into the cylinders of the barrel and the endcaps. In the barrel, they are arrayed in a rectangular geometry around the outside of the cylinder with the face of the detectors aligned toward the interaction point. The endcap disks are constructed of wedges constructed out of the same rectangular modules. The wedges are arrayed like a fan with the modules pointing 20° off center from the interaction point.

The individual pixels are $100 \times 150\,\mu$m, but they are capable of much better resolution than this size. If all the charge is deposited in a single pixel then the best resolution one could hope to achieve is the pixel size over $\sqrt{12}$. If the signal is shared between two or more channels, then the ratio of charge in each channel can allow for a more precise weighted-average to be made. Because the magnetic field runs parallel to the face of the pixel modules in the barrel, the charge is spread across multiple pixels further increasing the charge-sharing. The tilt of the endcap wedges creates a magnetic field parallel to the long axis of the pixel modules which amplifies the effect of charge sharing in the endcaps as well.

The silicon strip tracker is located outside the pixel tracker. The primary purpose of the strip tracker is to measure the momentum of charged particles as they leave the

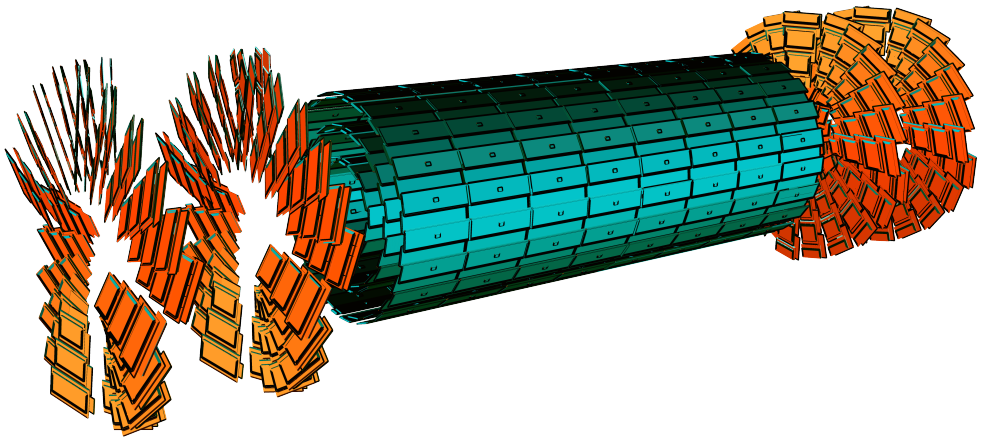


Figure 3.3: A view of the CMS pixel Tracker.

detector, particularly muons, which pass through the calorimeters as minimum ionizing particles. The momentum resolution of a tracking detector is inversely proportional to the magnetic field and the lever arm of the measurement. To optimize the momentum resolution, the longest track measurement in the largest magnetic field possible are used. The CMS strip tracker has ten barrel sections extending from just outside the pixel tracker to 1.1 m from the interaction point. To augment the barrel region there are also twelve endcap discs on either end of the strip tracker. The strip tracker geometry is shown in figure 3.4.

The strip tracker covers 200 times the area of the pixel tracker so it was not feasible economically or practically to implement it with pixel detectors. The solution implemented is to use larger silicon modules which are segmented into strips along the full length of the module and vary from 80 to 180 μm wide from inner to outer layers. The

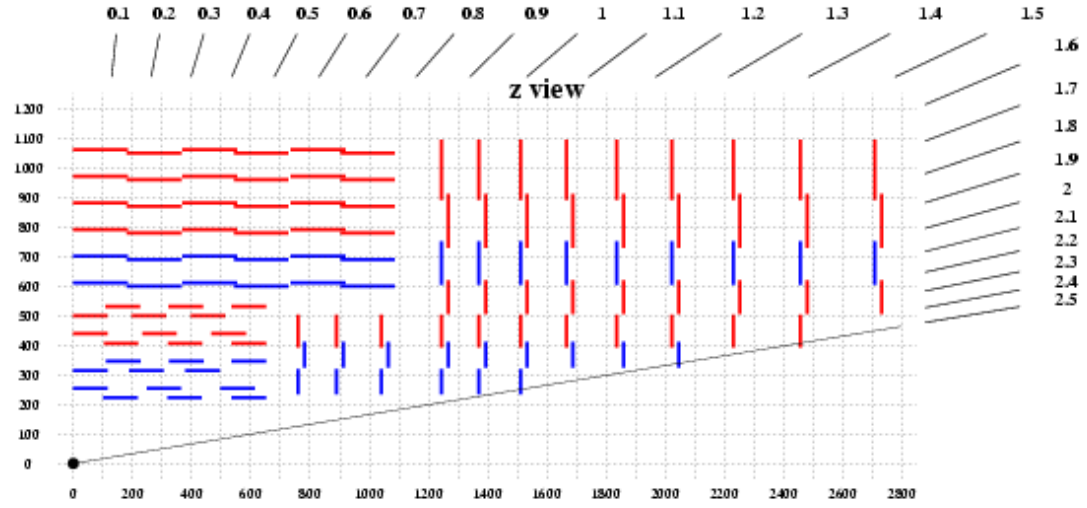


Figure 3.4: A quarter cross-section view of the CMS strip tracker. On the top and right side of the figure η is labeled. On the bottom and left the linear distance from the interaction point is mm is indicated.

long dimension of each strip is aligned with the magnetic field. This much coarser channel geometry is acceptable due to much lower particle fluxes in the strip detector and because for measurements of momentum the crucial direction is in the $r-\phi$ plane. The two innermost layers, and the two middle layers are constructed with a double layers of modules which are offset by an angle of 100 mrad. This angle allows for an accurate 3D position measurement in these layers.

In order for the tracker to measure particle tracks to the best possible precision, the tracker position and alignment must be known better than the anticipated resolution. The only way to achieve the precision required is an in situ calibration and alignment. The primary calibration was done before the start of beam collisions using cosmic ray muons with the magnet field off and on. After the start of proton-proton collisions in 2010, collision data was also used to calibrate the tracker alignment. In addition to an alignment better than 100 μm, the alignment must be constantly checked over time using collision data and additional cosmic data runs. The momentum response for the tracker is further calibrated using measurements of the Z boson. By reconstructing the invariant mass of $Z \to \mu\mu$ events as a function of the positive muon ϕ, distortions of

the tracker in the $r - \phi$ plane can be studied and correction factors derived [68].

After calibration and alignment the tracker is capable of very accurate reconstruction of tracks and primary vertices. The reconstruction is capable of finding primary vertices with a resolution of better than $100\,\mu$m in the $r - \phi$ plane and better than $150\,\mu$m in the z direction. This is crucial to allow for accurate identification of particles which come from extra, uninteresting proton-proton interactions in the same bunch crossing, called pileup collisions [68]. The momentum of a 100 GeV muon can be measured with a resolution of about 1% in the barrel and about 3-6% in the endcaps.

3.5 The Electromagnetic Calorimeter

The Electromagnetic Calorimeter (ECAL) is designed to measure the energy of electrons and photons coming from the interaction point. The primary design concern for ECAL was the energy resolution for photons coming from $H \to \gamma\gamma$ events. With a calorimeter, the best energy resolution is achieved when the full energy of the particle can be contained in the active medium of the detector. This means that for the optimal resolution, a detector should have a high density of large Z material with many radiation lengths of material. For CMS there is the additional requirement that the chosen material be radiation tolerant. Finally, to allow for precise di-photon mass resolution, the detector must have a high granularity to infer the direction of individual photons which leave no track in the tracker.

ECAL is constructed primarily of lead-tungstate crystals. Lead-tungstate is a clear inorganic scintillator with a very high density. The ECAL crystals have a radiation length $X_0 = 0.89\,$cm and a Moliere radius of 2.2 cm. This allows for a fairly compact detector which can fit inside the solenoid but still contain $\sim 25 X_0$ of material. Additionally, the small Moliere radius means that showers will remain compact in $\eta - \phi$ space which helps resolve individual particles even in events with very high detector occupancy. In addition to compact spatial resolution, the scintillation mechanism of lead-tungstate is very fast. This prompt and brief signal allows for accurate timing measurements as well as little dead time from one bunch crossing to the next. Lead-tungstate is fairly radiation hard. The primary mechanism of radiation damage is the formation of color centers which darken the crystal and reduce light yield. This change

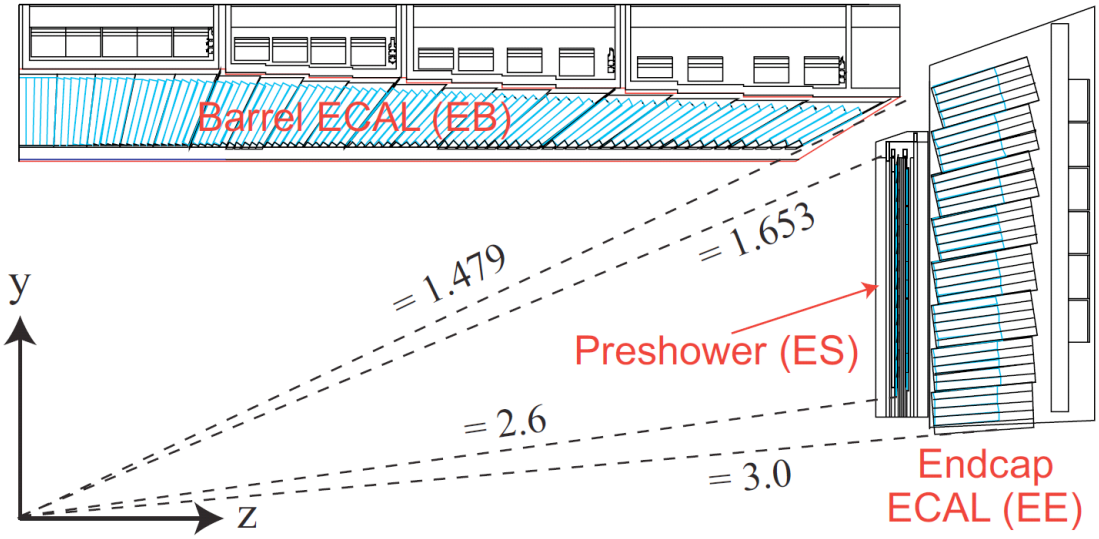


Figure 3.5: A quarter cross-section view of the CMS ECAL.

in transparency is constantly monitored with laser pulses and corrected during physics runs.

ECAL, like the tracker, is split into barrel and endcap segments. The barrel segment is constructed of crystals with a cross section of $2.2 \times 2.2\,\mathrm{cm}^2$ and 23 cm long, though the exact geometry of each crystal is dependent on its location in the barrel. The barrel is assembled so that 5×5 groups of crystals form a rectangular geometry in $\eta - \phi$ space. The crystals are aligned so that their long axis points 3° off center from the nominal interaction point. This small offset is designed to hide the small gaps between crystals from particles leaving the nominal interaction region. The endcap segments are constructed of slightly larger crystals with a front face of $2.86 \times 2.86\,\mathrm{cm}^2$ and are 22 cm long and are all identical in geometry. Unlike the barrel, 5×5 groups of crystals in the endcaps are arranged in a grid in the $x - y$ plane. The ECAL crystal geometry is shown in figure 3.5. In total, there are 61200 crystals in the barrel and 7324 crystals in each endcap.

In addition to the barrel and endcap, there is also a portion of ECAL called the preshower which is situated in front of the endcaps. The preshower is a thin detector constructed of lead and silicon which has a higher granularity than the lead-tungstate

crystals. The purpose of the preshower is to give a high spatial resolution of the start of electromagnetic showers in the endcap. It was originally thought this would help to separate single photons from highly boosted $\pi^0 \to \gamma\gamma$ decays, but was found to be ineffective for this purpose.

Although lead-tungstate is a compact and radiation hard scintillator, its light yield is very low, about 30 photons/ MeV. This requires the use of high gain photo-sensors to make up for the small signal. Two photo-sensors are used in ECAL. In the barrel, avalanche photo-diodes (APDs) are used. In the endcaps, vacuum photo-triodes (VPTs) are used. These sensors were chosen due to their high signal gain and their relative insensitivity to the magnetic field in which they were required to operate. The signal from the photo-detectors is then digitized and the signal is stored on the front-end electronics until the event is accepted by the level-1 trigger. At this point the event is transmitted onward over an optical data link.

ECAL was initially calibrated before being installed using radioactive sources and dedicated electron beams. These calibrations were then refined with various techniques. The most common technique used is the technique of ϕ symmetry. This method asserts that, due to complete symmetry in ϕ the energy response should be independent of ϕ for any particular ring in η and therefore the crystals are calibrated so that any asymmetries in ϕ are removed. Absolute calibrations are also performed using the reconstructed mass of the Z boson which is well-known and relatively background free near the resonance peak and therefore provides a uniform point of reference at energies relevant to LHC physics. The resolution of ECAL can be parameterized as follows

$$\frac{\sigma_E}{E} = \frac{2.8\%}{\sqrt{E/(1\,\mathrm{GeV})}} \oplus \frac{0.128\,\mathrm{GeV}}{E} \oplus 0.3\%, \tag{3.4}$$

where the first term corresponds to the statistical fluctuations in the measurement, and the second and third terms are noise terms. An analysis of ECAL performance, in particular how it affected the discovery of the Higgs Boson, can be found in [69].

3.6 The Hadronic Calorimeter

The Hadronic Calorimeter (HCAL) is designed to measure the energy of strongly interacting particles which cannot be fully contained within ECAL. The measurement of

hadronic particles is a much more complicated task than measuring light electromagnetically interacting particles such as electrons and photons. Unlike electrons and photons, quarks are confined by the strong force to color neutral particles. Because of this any "bare" quarks which are created in a proton-proton collision are hadronized. The strong force causes new quark-antiquark pairs to be created until stable or metastable states, almost exclusively pions, kaons, and protons, are created. So instead of a simple single particle interacting and showering in the detector, there is a spray of particles, called a jet.

The difficulties of hadronic calorimetry are further complicated by the ways in which hadrons interact with the calorimeter. To good approximation, a jet can be considered to contain only pions. The neutral pion component will decay promptly into photons which can be detected in ECAL. Electromagnetically, the charged pions look like minimum-ionizing particles, but they also undergo strong interactions with the nuclei in the detector. As nuclei are sparsely distributed and the strong force has a short range due to confinement, the overall interaction cross-section for nuclear interactions is lower than for electromagnetic interactions. This leads to a very dispersed shower structure and a long interaction length, λ_I. A pion will travel through some amount of detector before interacting, creating a localized shower, and perhaps continue on, to cause another shower later in the detector. A cartoon of a hadronic shower is shown in figure 3.6 [70].

The energy resolution of a hadronic calorimeter is also degraded due to statistical fluctuations between the hadronic component of showers (from charged pions) and the electromagnetic component (from neutral pions). Hadronic showers are generally started by a small numbers of particles, $\mathcal{O}$(tens), therefore the statistical fluctuation in the number of charged versus neutral hadrons on an event-by-event basis is significant. This would not be an issue if the detector response for hadronic and electromagnetic energy deposition were the same, but this is not the case in real calorimeters. This means that the detector response must be corrected on average and the fluctuations absorbed into the energy resolution of the calorimeter. This correction factor is commonly referred to as the π/e, or hadronic over electromagnetic correction.

These concerns lead to the ideal design for a hadron calorimeter to contain many interaction lengths, and be able to distinguish between electromagnetic and hadronic

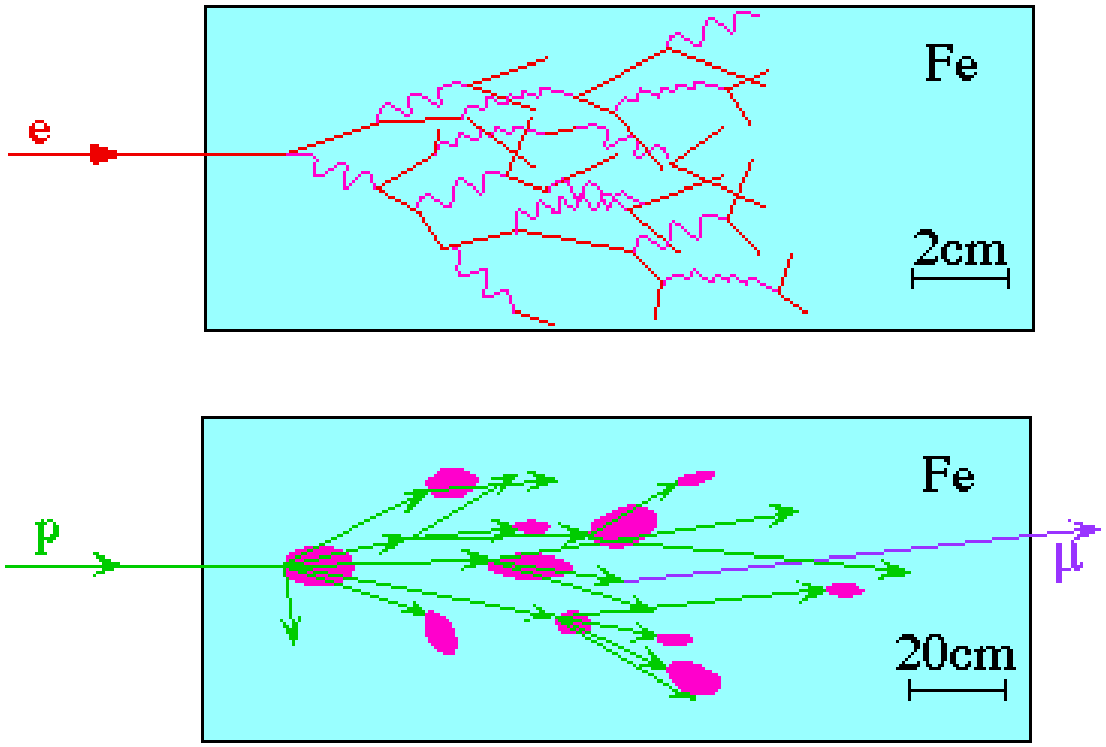


Figure 3.6: Cartoon of shower formation. The cartoon on top shows an electromagnetic shower while the bottom shows a hadronic shower. Each purple ellipse in the hadronic shower represents an electromagnetic shower.

interactions. To implement such an ideal detector would be costly both in space and money. The common solution is to accept that the hadronic and electromagnetic fluctuations are the limiting factor and to construct high-density, low-cost sampling hadronic calorimeters. In CMS this choice is even more appealing due to the size constraints imposed by keeping the calorimeter inside the magnet. A sampling calorimeter is constructed of alternating layers of active detector with layers of passive absorber material. This design will inherently lead to worse resolutions than one which allows for measurement of the entire shower, but has significantly lower cost per interaction length.

The HCAL for CMS is constructed with brass as the main absorber material and plastic scintillator as the active detector. The 17 plastic scintillator layers are all 3.7 mm thick, with the exception of the first and last which are 9 mm thick. In between each scintillator panel is a layer of brass, except the first and last absorber layers which are

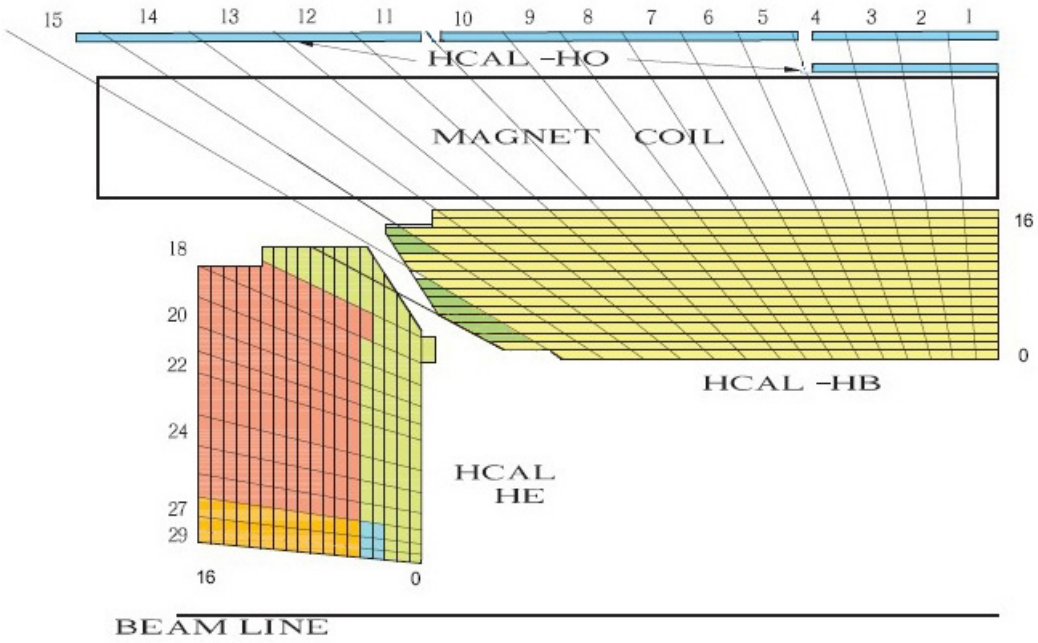


Figure 3.7: A quarter cross-section view of the CMS HCAL.

made of stainless steel. The first stainless steel layer is 61 mm. The next eight layers of brass are 50.5 mm. The last six brass layers are 56.5 mm. The final stainless steel absorber plate is 75 mm thick. The first scintillator layer is thicker to catch any shower escaping ECAL or forming in the support structure between ECAL and HCAL. The last layer is thicker to catch the tails of any shower escaping HCAL. The absorber thicknesses are adjusted with depth to maximize sampling early in the shower. Additionally, there is a layer of scintillator outside the magnet which was also designed to catch any showers that made it through the magnet, but this was not active for LHC Run-I. ECAL and HCAL together vary between 7 to 10 interaction lengths depending on η. The HCAL layer geometry can be seen in figure 3.7.

Like the other sub-detectors, HCAL is also constructed in a barrel and two endcaps. The barrel covers $-1.4 < \eta < 1.4$. The scintillator panels in the barrel are divided into sections which each cover $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$. Each barrel channel in HCAL corresponds to one 5×5 grouping of ECAL crystals. Unlike the ECAL endcaps, the

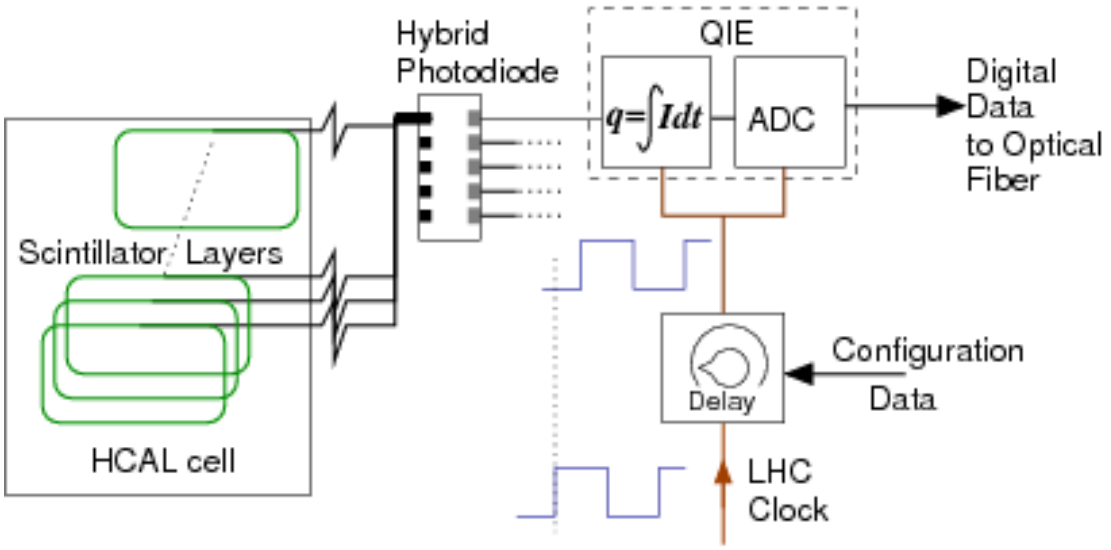


Figure 3.8: HCAL front end electronics

HCAL endcaps are constructed to maintain a rectangular geometry in $\eta - \phi$ space. At the lowest η the channel size approximately matches the barrel and they grow in η and ϕ as space becomes limited at the highest η of 3.0.

Both the barrel and endcaps use the same readout technique shown in figure 3.8. Each scintillator panel has a single wavelength shifting fiber embedded in it. This fiber is then bonded to a traditional optical fiber which carries the signal out of HCAL to a readout box (RBX) which sits behind HCAL but still inside the magnet. Each RBX houses the photo-sensors, digitization electronics, control module, and optical data transmitters for 72 channels. The whole of HCAL uses 72 RBXs. The optical fibers are passed into a "splitter" box which optically sums the signals from multiple layers in each tower of HCAL and distributes them to the photo-detectors. The summed layers are color coded in figure 3.7. In some sections near the barrel-endcap junction and in the endcap there are multiple depths which are read out independently.

The photo-detectors chosen for HCAL are hybrid photo-diodes (HPDs). HPDs are single stage, high-gain, multi-channel photo-sensors. They are semiconductor devices which utilize a PIN diode array behind a photo-cathode. There is a very high negative bias voltage of ~8 kV between the photo-cathode and the PIN diode array. This allows for a large gain of ~2000 which helps to make up for lower light yield of the sampling

calorimeter design. The single-stage amplification makes the devices compact and resistant to high magnetic fields present inside the magnet. The magnetic field actually enhances the performance of the HPDs. By aligning the magnetic field to be normal to the photo-cathode the cross-talk between channels is greatly reduced. Furthermore, the intrinsic noise of the devices is reduced in a strong magnetic field. These effects, or lack thereof, are the reason that the barrel section of HCAL which sits outside the magnet was inoperable for Run-I. With increased cross-talk and noise in this portion of the detector the system proved too noisy to successfully extract physics results.

The RBX also houses the basic front-end control module, called the Clock and Control Module (CCM). The CCM handles all communication with the RBX excluding data transmission, which is handled by a dedicated gigabit optical link (GOL). The CCM receives the clock which controls the synchronization of data taking. It then distributes the clock to the four readout modules (RMs) which each control 18 channels. The CCM also receives and distributes basic configuration information for each channel. These include pedestal and clock delay settings for each channel, LED settings for test pulsing channels, and GOL settings. Finally the CCM monitors basic settings of the RBX such as temperature and low voltage settings which are sent back to the monitoring software. All settings are stored in a software database and uploaded to each RBX at the beginning of every cycle of data taking.

The HCAL readout electronics after the HPDs are summarized in figure 3.8. The signals from the HPDs are read out by a charge integration encoder (QIE). The QIEs work on a 40 MHz clock. The current from the HPD is split equally to four capacitors of different sizes, for different gains, which collect the charge for 25 ns. The voltage is then digitized from the channel with the highest non-saturated output. The digitized voltages are then packed into a byte stream along with additional information such as the chosen gain and are transmitted by the GOL to the back-end electronics which are located off detector.

The data arrives at the HCAL Trigger and Readout (HTR) card where it is unpacked. The HTR prepares the trigger information and sends this off to the trigger electronics. The data is then buffered for up to 3.2 μs while waiting on a decision from the trigger electronics. If an accept is received from the trigger, the data is transmitted on to the high level trigger which will be discussed further in following sections.

Calibration of the hadronic calorimeter began by characterizing the detector response. The response and light yield of the scintillator is determined both before and after installation into the absorber using radioactive sources. From this, a basic energy response of the scintillator can be derived. The fully assembled detector was also exposed to pure beams of electrons and charged pions of known energy. This allows for the detector response to hadronic and electromagnetic showers to be measured independently and the π/e correction to be derived. For these tests it was important to have both ECAL and HCAL in the particle beam in their realistic detector geometry as they will have different π/e corrections. This allows the non-linearity in response to be measured and corrected. After these corrections are applied the resolution of ECAL plus HCAL for hadronic particles is measured to be

$$\frac{\sigma_E}{E} = \frac{84.7\%}{\sqrt{E/(1\,\mathrm{GeV})}} \oplus 7.4\%. \tag{3.5}$$

A full description of the calibration and performance of HCAL can be found in [71, 72].

In addition to the barrel and endcap HCAL there is an additional detector which is designed to cover the forward region, $3 < |\eta| < 5$. This detector is situated in one of the highest radiation environments in the detector and must be highly radiation resistant. This was accomplished with a detector designed by using quartz fibers embedded in a steel absorber which are both very resilient to radiation. The quartz fibers run parallel to the beam at 5 mm spacing in the steel. The fibers are read out in pairs, with one fiber extending the whole length from front to back of the detector while the other fiber stops about one interaction length from the front of the calorimeter. As there is no ECAL in front of this detector, this fiber geometry allows electromagnetic showers to be separated from hadronic showers. As the detector is outside the magnet, 11 m from the interaction point, it is able to use traditional photo-multiplier tubes for its light detection. Otherwise the same readout electronics are used as in the rest of HCAL. Although I was responsible for some upkeep work on this detector, it is not used for the search for heavy neutrinos presented in this thesis.

3.7 The Muon Chambers

The purpose of the muon chambers in CMS is to supply information for the trigger system and assist in muon identification. Without the muon chambers it would be impossible to trigger on muons because the tracker information cannot be used to make early trigger decisions. The tracker provides the primary measurement of muon p_{T}, but for $p_{\mathrm{T}} > 200\,\mathrm{GeV}$ the information from the muon system can be combined to improve the p_{T} determination. The muon chambers are located outside of the magnet and are interleaved with the iron return yoke which contains the flux to reduce fringe fiends. Due to their location far from the interaction point, the muon chambers are required to cover a very large surface are of $25{,}000\,\mathrm{m}^2$. To achieve this level of coverage at a reasonable cost, gas based detectors were chosen. A combination of technologies is used depending on the location in the detector.

In the barrel section of the muon chambers the technologies used are drift tubes (DT) and resistive plate chambers (RPC). The layout of the CMS muon system can be seen in figure 3.9. DTs are gas ionization detectors which collect the ionization charge on an anode wire strung down the length of a tube whose body acts as the cathode. By exploiting the timing and pulse shape it is possible to locate the position of an ionizing particle passing through a DT to $\sim$200 μm. The muon barrel system contains four layers of DT stations separated by return yoke iron which each contain 12 or 18 layers of DTs. The first three layers of DT stations have tubes aligned to measure both the $r-\phi$ and z coordinates. The fourth layer makes a measurement only in the $r-\phi$ coordinate. Together each station is capable of reconstructing a track segment with both position and angle of the track and every muon passes through at least three muon stations. Each DT station is capable of making a muon track measurement with around 100 μm precision in position and 1 mrad in angle.

To augment the DTs in the barrel, RPCs are installed along with each DT station in the barrel. RPCs are constructed of two parallel high-voltage resistive plates separated by a narrow gas gap. Any ionizing particles create an electron cascade that is collected on the anode which is divided into strips. The RPCs have worse position resolution but complement the DTs with a smaller drift distance which gives them much faster timing response, $\sim$1 ns. This fast response time provides prompt information for use

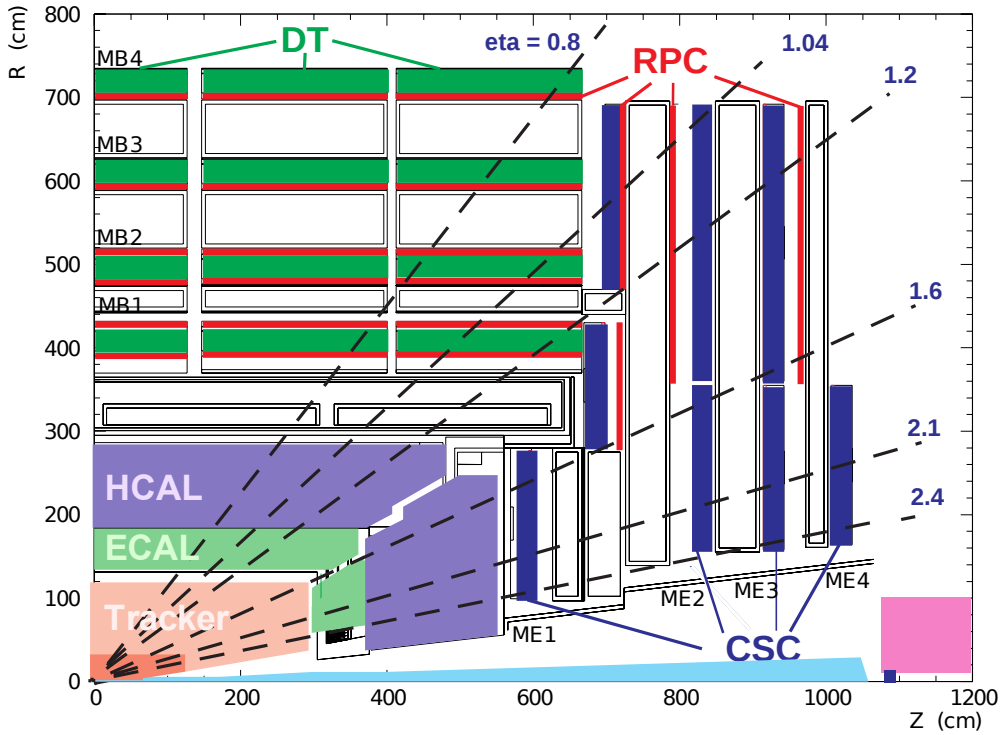


Figure 3.9: A quarter cross-section view of the CMS muon chambers.

in the muon trigger system as well as an accurate measurement of the timing of muon interactions. This is important to help separate which bunch-crossing each muon came from.

In the endcap section of the muon system, the higher magnetic field and particle flux makes DTs ineffective. Instead cathode strip chambers (CSC) are used. A CSC is constructed of a flat gas chamber with a series of anode wires on one face and an orthogonal set of cathode strips on the opposite face with a small gas gap in between. The charge signal on the anode wires provides a prompt signal in time, while a weighted average of charge on the cathode strips provides a precise measurement of particle position. The prompt anode signal is used for the muon trigger system, while the slower cathode information is used for final reconstruction.

Each endcap is constructed of three layers CSCs and RPCs separated by iron return yoke up to a pseudorapidity of $|\eta| < 1.6$. Like the DTs in the barrel, each layer of CSCs

contains a stack of six CSC modules which allow for each layer to make a measurement of muon position and direction to a precision of 200 μm and 10 mrad respectively. Beyond $|\eta| > 1.6$ the endcaps consist of only three layers of CSCs with no RPCs. The muon chamber endcap geometry is shown in figure 3.9.

3.8 The Trigger

The trigger is responsible for selecting which events are interesting, and therefore will be saved, and which are not. The CMS trigger system is divided into two parts: the level-one (L1) trigger which consists of a fast hardware trigger to take the 20 MHz of collision events down to about 100 kHz and the high level trigger (HLT) which uses a computer farm running a software trigger to reduce the data rate further to about 1 kHz for storage on disk.

The L1 trigger is designed to operate with high speed and low latency. The L1 trigger is required to operate at the full LHC collision rate of 40 MHz (though the collision rate in 2012 was only 20 MHz) and make its decision to keep or reject each event within 3.2 μs. Information from the calorimeter and muon sub-systems are used in the L1 trigger decisions, but tracker information is unavailable. Each sub-detector collects the information for each event and uses a course-grained version of the information to calculate basic trigger information to be sent to the L1 trigger. This information includes estimates of the E_{T}, p_{T}, η, and ϕ of electrons, muons, and jets as well as basic combinations such as a sum of all E_{T}. The L1 trigger hardware then selects events to pass based on predefined selection criterion which are selected to balance the rate of selected events of interesting topology to fill the 100 kHz allotted rate.

While the L1 trigger is processing an event, the associated information from each sub-detector is stored in a hardware pipeline. If an event is selected then the data is moved from the pipeline and is packaged and sent to the central DAQ system. It is held there until all the required information is collected from each sub-detector and the event can be assembled and sent on to the HLT.

The HLT uses a software based system to filter events. The HLT currently contains more than 10000 processor cores. Events selected by the L1 trigger are passed in their entirety (about 2.5 MB each) to a processor in the HLT processor farm. This allows the

HLT to have access to the full event information from every sub-detector. The selection algorithms start with very simple trigger paths, such as selecting events with a muon with $p_{\mathrm{T}} > 40\,\mathrm{GeV}$ and work their way more complicated algorithms for selecting events. The software which runs on the HLT computers is the same software which is used for the final reconstruction data, so the trigger algorithms have access to any algorithm available to full reconstruction limited only to the maximum computing time allotted. This approach allows for a very flexible trigger system which can be easily reconfigured and allows for advanced reconstruction techniques to maximize signal efficiency and background rejection.

Chapter 4

HCAL Upgrade

This chapter takes a brief interlude from the search for heavy neutrinos and discusses work performed to upgrade the readout electronics of HCAL. Several factors contribute to the need to upgrade HCAL and all major sub-detectors in CMS. As the LHC operation continues, the instantaneous luminosity will climb as well. Although this allows CMS to collect data at a faster rate, it also brings a more challenging environment with added pileup. The average peak number of pileup collisions during the $\sqrt{s} = 8\,\text{TeV}$ portion of LHC run-I was ~21. In the next fifteen years, the LHC intends to deliver $3000\,\text{fb}^{-1}$ of $\sqrt{s} = 13\,\text{TeV}$ proton-proton collisions to CMS and ATLAS [73]. At the highest instantaneous luminosity this will mean picking interesting events out of over 100 average pileup interactions. In order to maintain the same level of physics performance as CMS currently enjoys, many detector upgrades will be needed. Particularly, it is important to increase the density of readout channels to deal with pileup.

Another motivation for detector upgrades is radiation damage. The active components of CMS, especially in the forward (high $|\eta|$) region of the detector are subjected to very high radiation doses from the proton-proton interactions. In silicon detectors this causes more noise and higher operating power. In a scintillating detector such as HCAL the primary effect of radiation damage is darkening of the scintillator leading to reduced light yield [74]. In the short term, light yield loss can be corrected using dedicated measurements of scintilator transparency throughout the course of a physics run, but eventually the effect will lead to serious degradation of performance. The long-term solution to this problem is to replace the active elements of the detector, particularly in

the forward region, with more radiation-tolerant materials.

A final motivation for detector upgrades is the incorporation of new advanced technologies which were not available when CMS was originally designed. Some upgrades, such as those to the readout electrons, are designed to increase the total bandwidth and flexibility of the readout electronics well beyond the initial CMS design goals. Other upgrades, such as the installation of new photo-detectors for the portion of the HCAL barrel outside the magnet, allow portions of the detector with lackluster performance to meet their initial design goals.

This chapter will focus on the HCAL phase I upgrade to the readout electronics [75]. Long term upgrades to CMS, called the phase II upgrades, including upgraded active components of the detector, are not scheduled to be installed until at least the year 2022. However, the phase I upgrade is already under way with an expected completion by 2018.

4.1 HCAL Phase I Upgrade Overview

The HCAL phase I upgrade will replace front and back end readout electronics for HCAL. The front end electronics include the photo-detectors, readout modules, CCM, and GOL. The back end electronics include the HTR and DCC cards. The existing HCAL readout integrates the light from all but the first layer as shown in figure 3.7 for most values of η. The upgraded readout will take better advantage of the detector by increasing the number of depth segments to four in the barrel and five in the endcaps as shown in figure 4.1. The largest benefit from the increased segmentation is the ability to apply independent detector response corrections for various depths.

Originally the number of channels in HCAL was limited by the large size of the HPDs used to read out the detector. The greater than doubling of HCAL channel count is accommodated by using a relatively new form of photo-detector called a silicon photo-multiplier (SiPM). SiPMs are small, have a low operating voltage (~80 V), are not affected by magnetic fields, and are resilient to the effects of radiation up to the dose expected at the LHC. SiPMs are constructed from arrays of tiny avalanche photo diodes which are operated in Geiger mode. The signal from the array is then read out as a single current pulse proportional to the number of pixels which received photons. The SiPMs

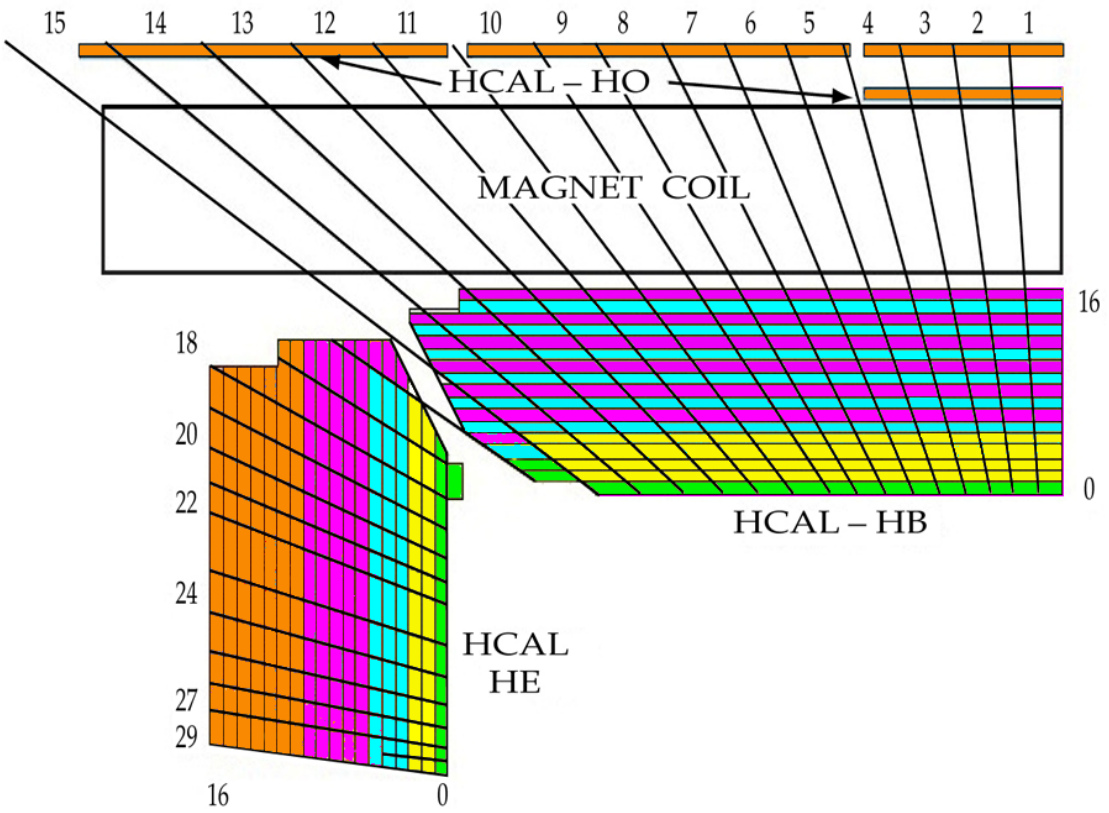


Figure 4.1: New segmentation of HCAL after phase I upgrade.

will be read out with a new eight-bit charge-to-digital converter. Additionally, a time to digital converter will give an accurate time measurement which can be used by the L1 trigger. After digitization, the signals are bit-packed by an FPGA and transmitted to the back end electronics via 4.8 Gbps optical link. The link can be operated at speeds as high as 6.4 Gbps if required by future upgrades. The new readout electronics are summarized in figure 4.2.

The data from the front end is received by the micro HCAL trigger and readout (μHTR) card. The micro serves to distinguish these cards, based on the μTCA standard, from the old HTR cards, based on the VME standard. A simplified design of the μHTR card is shown in figure 4.3. The data flow is handled by two FPGAs. The first FPGA receives the data from the optical link and unpacks the data. It then constructs the L1 trigger information from the raw data and sends this through another optical link

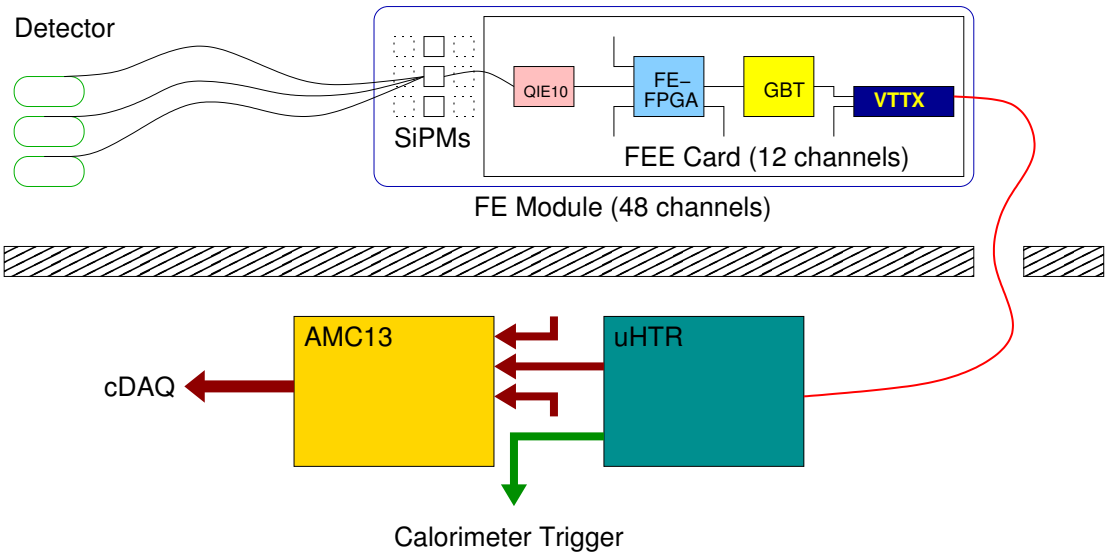


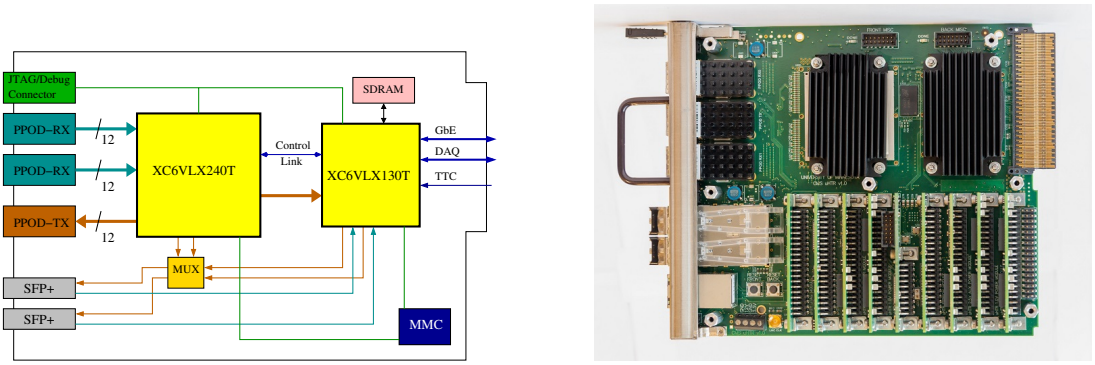
Figure 4.2: HCAL phase I upgrade readout chain.

directly to the trigger electronics. The data is then buffered until a response from the trigger is returned for each event. The events which pass the trigger are then moved into the second FPGA where they are held until all portions of the event from every μHTR is ready. At this point the event is transferred to the central data acquisition system through another card called the AMC13 (replacing the DCC in the existing back end).

4.2 uHTR Mezzanine Testing

The μHTR card is one of the main hardware contributions of the University of Minnesota to the HCAL phase I upgrade. In addition to the design of the main baseboard for the μHTR, several small supporting boards called mezzanines were also designed for the μHTR. The mezzanines host support circuitry including DC-DC converters, linear regulators, flash memory modules, and control CPLDs. These mezzanines, particularly those hosting the DC-DC converters which are soldered directly to the baseboard, need to be thoroughly tested before being installed.

The voltage conversion mezzanines convert the input 12 V from the crate into other voltages needed on the μHTR. They come in two varieties. The first, called power mezzanines (PM), host two DC-DC converter modules which work together to output a



(a) (b)

Figure 4.3: A block diagram of the μHTR (left) and a photo of a completed μHTR (right).

single voltage, either 1.0 V or 3.3 V. The second type, called auxiliary power mezzanines (APM), host a single DC-DC converter which supplies either 1.6 V or 2.5 V and a series of linear regulators which supply 1.0 V and 1.2 V. Each μHTR hosts a 3.3 V and two 1.0 V PMs and one each of the two APMs.

The PMs and APMs are soldered to the μHTR baseboard because of the high current they handle. This requires that they be tested thoroughly with a dedicated test board before they are installed in a μHTR. The design of the board, shown in figure 4.4, is intended to supply a known load equal or in excess of the load which will be experienced by each mezzanine in final operation. The test board is capable of holding a full set of five PMs plus APMs required for a single μHTR. The test board is also capable of enabling and disabling the mezzanines themselves and MOSFETs controlling the load resistors. The board can monitor the input voltage, input current, output voltages, temperature, and reading or writing to an EEPROM for each mezzanine. Additionally, the test fixture can adjust the output voltages up or down by 5%, called margin control.

The interface to the test board is made using the I^2C standard. I^2C is a two wire (clock and data) communication standard which allows bi-directional communication with multiple integrated circuits (IC) on the same communication bus based on their preassigned addresses. The I^2C interface is a good choice here because it is easy to use and there is a wide array of ICs covering all the functionality needed for the test fixture

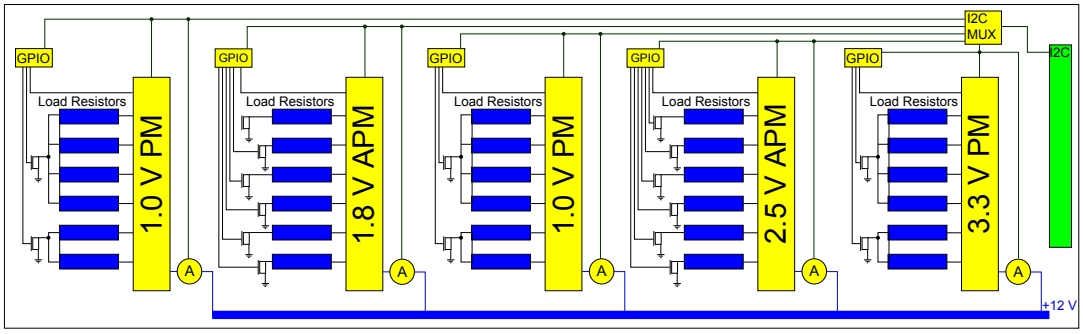
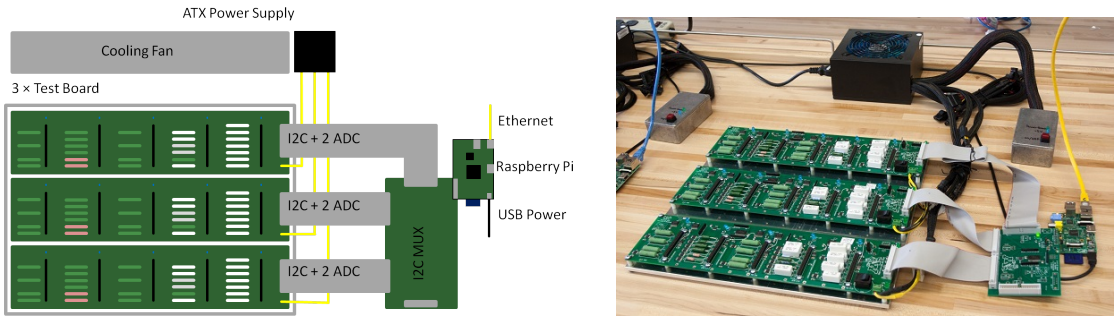


Figure 4.4: PM and APM test fixture design.



(a) (b)

Figure 4.5: A block diagram of a PM and APM test station (left) and a photo of a completed test station (right).

already existing for commercial purchase.

Test stations were assembled with three test boards. Each station's test boards are powered using a commercial ATX power supply. To dissipate the large amount of heat generated from the load resistors, each test station is actively cooled with a cooling fan. The layout of each test station is shown in figure 4.5. There are three test stations for a total of nine test boards.

The interface to each test station is provided by Raspberry-Pi. The Raspberry-Pi is a small, inexpensive system on chip computer based on an ARM processor. It is capable of running a full version of Linux which gives it great flexibility. Additionally, the Raspberry-Pi natively supports several communication standards including I^2C and Ethernet. To allow one Raspberry-Pi to control three test boards in each station, a

simple multiplexer board was designed. The multiplexer board allows the Raspberry-Pi to select between up to four test boards and adds additional ADCs which monitor the input voltage to the test board.

Each test station is controlled by an interface program running on a central computer. The central computer (not a Raspberry-Pi) is responsible for starting and stopping tests, monitoring the mezzanines during tests, and assigning serial numbers to the mezzanines. Each PM and APM has an on board EEPROM memory which stores a MAC address and a block of user configurable memory. At the start of testing, the test software assigns a unique serial number to each mezzanine along with some additional information including the mezzanine type and date of the test. The serial number and MAC address can both be used to track the provenance of the mezzanine later.

The test software is capable of running tests with margin up, margin down, nominal voltage, and a high load test with extra load resistors active. The standard test procedure starts with a short two hour test which runs each mode for 30 min. If the mezzanines show no obvious failures after this two hour test, then they are each tested for seven days at nominal load and voltage. While running both the short test and long test the software records the input voltage and current, the output voltages, and the mezzanine temperatures every ten seconds. The test is aborted if at any point the input currents or mezzanine temperature get to high. After the test is complete each mezzanine's performance can be evaluated based on the log of voltages, currents, and temperatures produced. During a test the interface software allows for easy monitoring of the status of all active tests as well.

The test software executes I^2C instructions on the test setups through the Raspberry-Pis. The central test computer runs the tests for each board in parallel. Each Raspberry-Pi runs a TCP/IP server which will accept an open connection from at most one test client at a time. When each test client desires to communicate with its test board it opens a TCP/IP socket and forwards its I^2C command to the appropriate Raspberry-Pi. The Raspberry-Pi then executes the I^2C command received on the requested test board. The result of the communication is then sent back to the test client and the socket is closed. The server on the Raspberry-Pi then processes the next socket request. Meanwhile the test client receives the response to its I^2C query and moves on to its next task. In this way there is never conflict between test clients over the I^2C bus. A

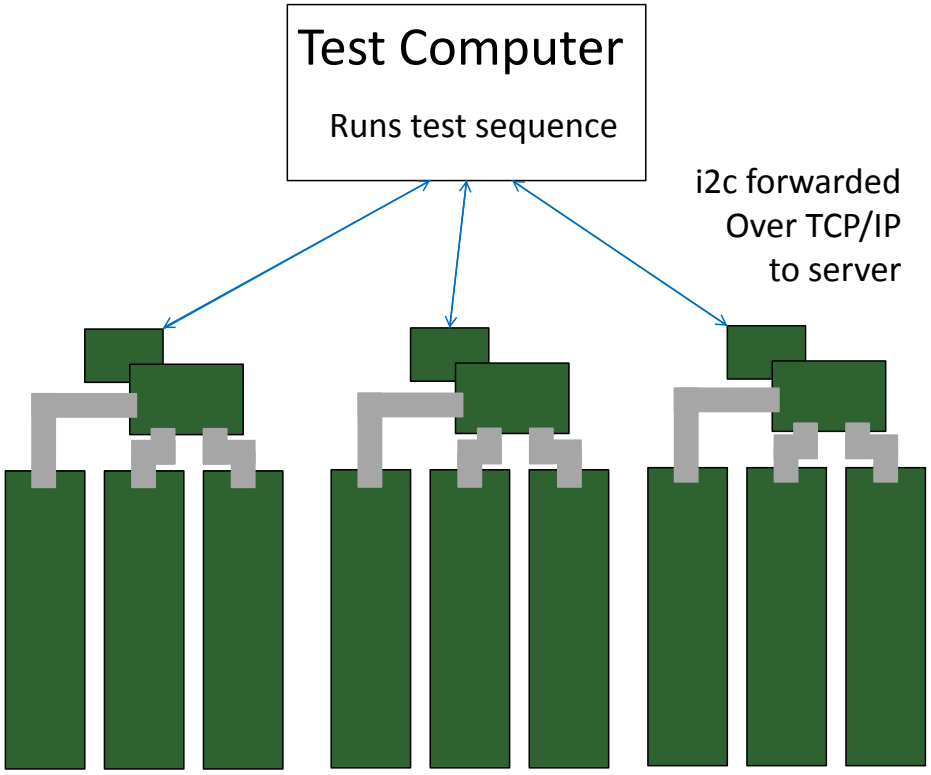
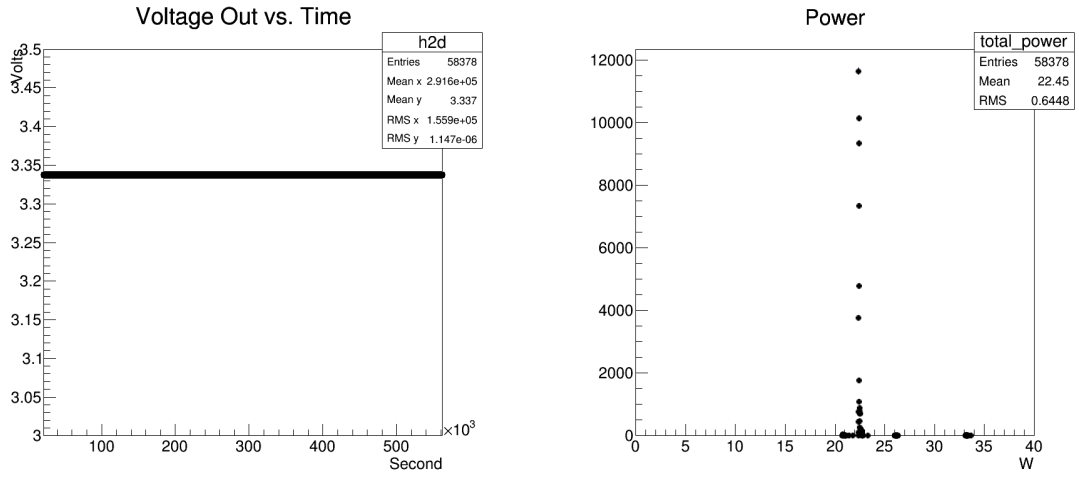


Figure 4.6: Mezzanine tester configuration.

diagram of the entire test setup is shown in figure 4.6.

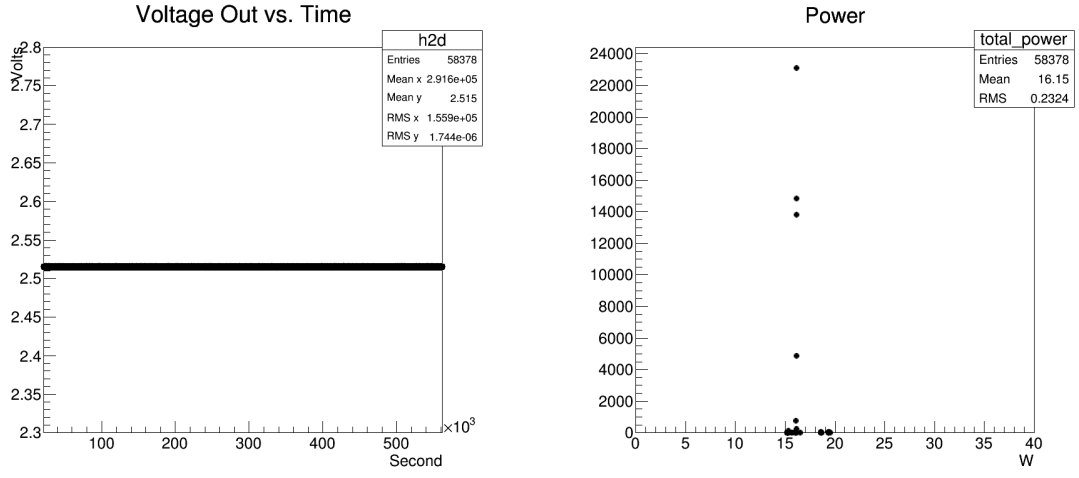
This system gives a flexible and user-friendly way to efficiently perform burn-in testing of the PMs and APMs. The testing is now under way and progressing smoothly. With 160 sets of five mezzanines to be tested the total process will take four months to complete, but the test setup has proven to be reliable and robust. As of the writing of this chapter, 36 sets of five mezzanines have been tested.

Example results of a 3.3 V PM is shown in figure 4.7 and a 2.5 V APM in figure 4.8. On the left of each figure the output voltage of the mezzanine is shown as a function of time. Each of the mezzanines chosen shows a very stable output voltage of a properly functional DC-DC converter. On the right of each figure, a histogram of the total power draw by the mezzanine is shown. The small peaks to either side of the main peak in the power output are from the voltage margin and high load tests.



(a) (b)

Figure 4.7: 3.3 V PM test results showing voltage versus time (left) and histogram of the total power (right).



(a) (b)

Figure 4.8: 2.5 V PM test results showing voltage versus time (left) and total power (right).

Chapter 5

Event Selection

This chapter details the selection of events and reconstruction of individual objects used for the search for $W_R \to lN_l \to lljj$ decays. It further explains how a W_R candidate is constructed and the selection requirements placed on it. The search is conducted using the final states $eejj$ and $\mu\mu jj$ for the electron and muon channels respectively. The final state containing $e\mu jj$ is used for background estimations. After all selection the four-object invariant mass of the two leptons and two jets is used to search for possible W_R signals.

5.1 Data and Monte Carlo

5.1.1 Data

The data used for this analysis comes from the full dataset collected at the center of mass energy $\sqrt{s}$ =8 TeV by the CMS experiment between April 2012 and December 2012. The LHC performance during this time was exceptional, delivering more than 23 fb^{-1} of integrated luminosity as seen in figure 5.1. The data is split into four run eras called 2012A, B, C, and D which can be distinguished in figure 5.1 [76] by the plateaus in the integrated luminosity. Each run era represents a period of reasonably consistent running conditions for the LHC which are separated by periods where the machine was being maintained and minor operational upgrades were applied to increase the instantaneous luminosity. After accounting for beam quality and detector down

time, there are $19.7\,\mathrm{fb}^{-1}$ of data which are used for physics analysis.

The entire dataset for CMS is too large to be practical to work with (many petabytes of data). Therefore, the data is split into smaller datasets which are arranged so that any particular analysis should need only one such dataset. Each of the smaller datasets includes data for a particular type or category of events such as those containing muons, or electrons, or high p_{T} jets. The sorting for each event is performed by the HLT and each event may end up in multiple datasets depending on which triggers are passed. As the luminosity of the LHC increased over time, many datasets were split into several smaller datasets to ensure the size of each remained small ($\mathcal{O}(10\,\mathrm{TB})$). Data used in this analysis were collected in the "SingleMu", "Photon", and "DoublePhotonHighPt" datsets. The datasets are summarized in table 5.1. The datasets used for this analysis were reconstructed, the process of turning raw detector response into physics objects, starting in January of 2013 to include the latest calibrations derived directly from 2012 data.

The "SingleMu" dataset, containing events that pass single muon triggers, is used to collect data for events with the final state $\mu\mu jj$ or $e\mu jj$. Data with the final state $\mu\mu jj$ is used to search for muon flavor heavy neutrinos. The $e\mu jj$ data is used for background estimation. In each event at least one of the selected muons must pass the HLT path "HLT_Mu40_eta2p1." This trigger requires the muon to have $p_{\mathrm{T}} > 40\,\mathrm{GeV}$ and $|\eta| < 2.1$. This HLT path was selected as it had the lowest p_{T} threshold of any un-prescaled, non-isolated muon trigger. Un-prescaled means that at no point during the data taking was the rate of this trigger artificially reduced by throwing away some fraction of the passing events. The non-isolated requirement means that muons which have other nearby activity in the detector are allowed to pass the trigger which is important to estimate the QCD fake rate for muons (this is discussed further in section 6.4). This trigger also has a higher overall efficiency for $\mu\mu jj$ events than a trigger

Run Era	$\mathcal{L}_{int}$ (pb^{-1})	Electron Dataset	Muon Dataset
Run2012A	876	Photon	SingleMu
Run2012B	4412	DoublePhotonHighPt	SingleMu
Run2012C	7055	DoublePhotonHighPt	SingleMu
Run2012D	7369	DoublePhotonHighPt	SingleMu

Table 5.1: Relevant information regarding the datasets used in this analysis.

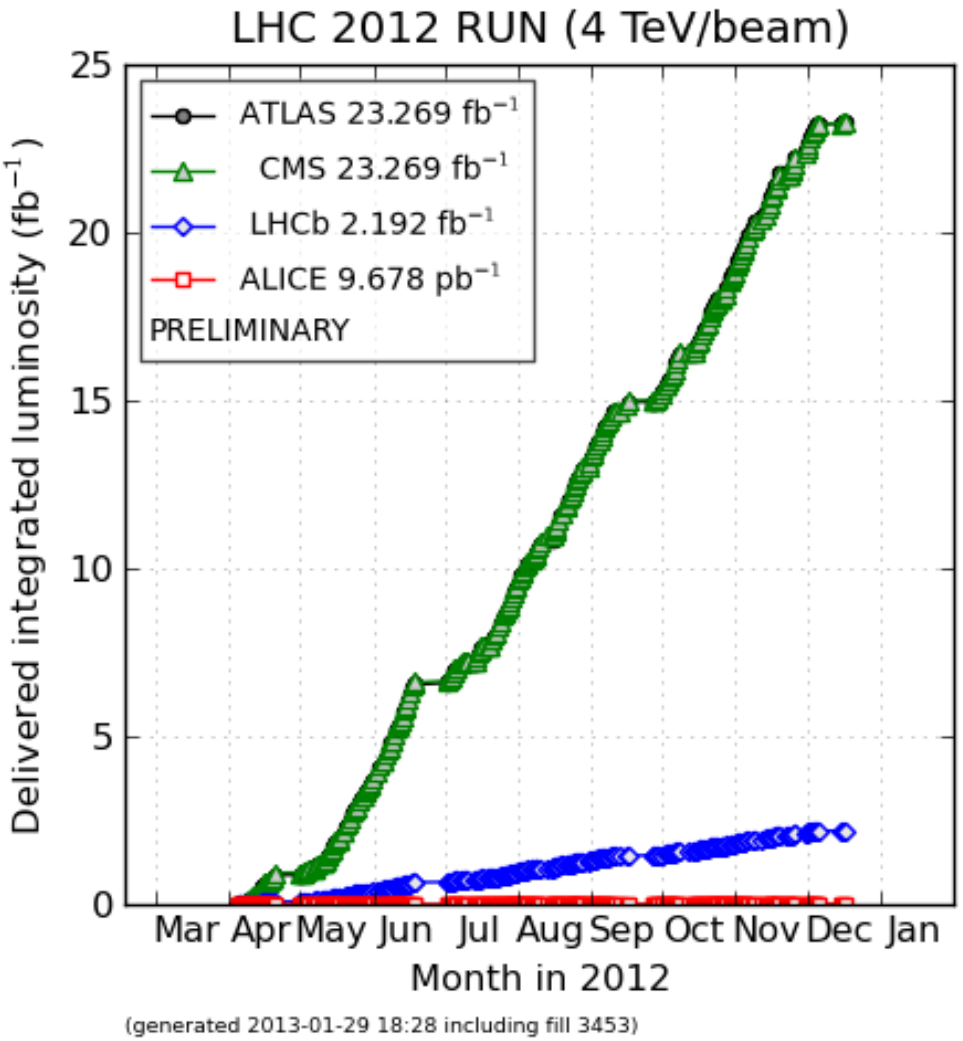


Figure 5.1: The LHC integrated luminosity for the 8 TeV dataset.

which directly requires two muons.

The "Photon" and "DoublePhotonHighPt" datsets, primarily requiring a single photon like energy deposit and two high-p_{T} photon-like energy deposits respectively, are used to collect data with the final state $eejj$. This data is used to search for electron flavor heavy neutrinos. Two datasets are used for this purpose as the "Photon" dataset was split into several smaller datasets at the end of run 2012A. To ensure that a consistent data sample is used between the two datasets, both of the selected electrons in each event are required to have passed the HLT path "HLT_DoubleEle33_CaloIdL_GsfTrkIdVL." This trigger requires that the event have at least two ECAL clusters with $E_{\mathrm{T}} > 33\,\mathrm{GeV}$ with associated tracks. This trigger was chosen over a similar trigger requiring only a single electron because the E_{T} threshold for single electron triggers is much higher than their muon counterparts. This necessitated the use of the double electron trigger to allow the electron and muon channels to keep the same E_{T} requirements.

In addition to the selection requirements above, we apply several cleaning filters to the data which remove any events which might contain possible detector issues. Events which do not contain any properly reconstructed primary verticies are removed. We also remove events which have anomalous noise or improper reconstruction in the tracker, calorimeters, or muon systems. Finally, we remove events which are flagged as having come from single-beam backgrounds such as the proton beam interacting with a stray gas particle.

5.1.2 Monte Carlo

Monte Carlo (MC) simulation is used in several places in this analysis to estimate backgrounds, derive scale factors, estimate the signal shape, and make corrections to limits. The MC used to estimate backgrounds and scale factors were all generated centrally by the CMS MC production group. A set of Drell-Yan (DY)+jets, W+jets, $t\bar{t}$+jets, and WW scattering MC generated using MadGraph [77] is used to estimate the backgrounds and calculate scale factors. A DY+jets sample generated using SHERPA [78] is used for a cross-check. Samples of tW and $\bar{t}W$ generated using POWHEG [79, 80, 81] and di-boson (ZZ, ZW, WW) samples generated with PYTHIA [61] are used to estimate their respective background contributions. The MC samples and the generators used to create them are summarized in table 5.2 along with the cross-section, number

Process	Generator	Events ($\times 10^6$)	σ, pb	$\mathcal{L}_{int}$ (fb^{-1})
$t\bar{t}$ +jets, $t \to bl^+\nu$	MadGraph	4.25	23.64	180
Z+jets, $Z \to l^+l^-$	MadGraph	30.5	3503.7 (NNLO)	8.69
Z+1 jet, $Z \to l^+l^-$	MadGraph	24.0	561.0 (LO)	42.9
Z+2 jets, $Z \to l^+l^-$	MadGraph	21.9	181.0 (LO)	121
Z+3 jets, $Z \to l^+l^-$	MadGraph	11.0	51.1 (LO)	216
Z+4$^+$ jets, $Z \to l^+l^-$	MadGraph	6.40	23.0 (LO)	278
WW	pythia	10.0	55.5	180
WZ	pythia	10.0	33.6	298
ZZ	pythia	9.80	17.7	554
W^-t	powheg	0.498	11.1	45
$W^+\bar{t}$	powheg	0.498	11.1	45
W^+W^+jj	MadGraph	0.100	0.248	403
W^-W^-jj	MadGraph	0.096	0.089	1085
W+jets, $W \to l\nu$	MadGraph	56.5	37509 (NNLO)	1.5
$W_R \to lN_l$	pythia	0.1 each	$4 \times 10^{-4} - 3.2$	–

Table 5.2: Summary information for the Monte Carlo samples used in this analysis. Cross-sections are calculated to next-to-leading order (NLO) unless noted otherwise.

of events in the sample and effective integrated luminosity of the sample.

In addition to the MC used in background studies, MC simulations of various signal hypotheses were centrally generated by CMS as well. The signal MC was generated using the LRS model built into pythia 6 with the assumptions listed in the end of section 2.5. Signal MC samples were generated for M_{W_R} in 100 GeV increments from 700 to 4000 GeV. Each sample is generated with the heavy neutrino mass assumed to be half of the W_R mass. A sparse grid of signal points with $M_{W_R} = 1, 1.5, 2, 3$ TeV and $M_N = \frac{1}{16}, \frac{1}{8}, \frac{3}{16}, \frac{1}{4}, \frac{1}{2}, \frac{5}{6}$ M_{W_R} was also produced to facilitate limit mapping which will be discussed in section 8.4.

Hadronization, the process of turning bare quarks produced in the generation of the sample to jets, is handled by pythia 6 using the CTEQ6L1 parton distribution functions [82] for all MC samples used in this analysis. The MC is then passed through the full CMS detector simulation which uses geant 4 to mimic the detector's response to each event [83]. Finally, the simulated detector response is run through the full CMS reconstruction algorithm which is run on data to produce an output format which is the same for MC simulation and data.

The centrally produced MC is mixed with so-called "minimum bias" events. Minimum bias events are events which pass a very minimum set of requirements to make in into a data set and are used to represent a typical proton-proton collision. Additional proton-proton collisions simulating these typical interactions are added to the MC to account for pileup interactions which take place during each bunch crossing along with the interesting interaction which triggered the event. The number of extra interactions which are added to each event is pulled from a distribution which is expected to match data, but because the pileup distribution changes from run to run, the MC cannot match all possible data distributions. This is accounted for by a reweighting technique. The distribution of the number of pileup events for the data range of interest is computed. This distribution is then compared to the distribution used to generate the MC sample and a weight factor is derived for each event, biased on the number of pileup interactions in that event, so that the MC and data pileup distributions are brought into agreement.

In addition to the centrally produced MC, privately produced signal samples are used for limit corrections and systematic studies. These samples are produced with the same PYTHIA configuration as used in the central production, but we omit the simulation and reconstruction steps. These samples will be discussed in more detail in the appropriate sections.

5.2 Object Selection

5.2.1 Jet Selection

The recommended CMS jet reconstruction, described below, is to cluster particles reconstructed by the particle-flow technique [84] using the anti-k_T clustering algorithm [85]. To account for particles which originate from pileup interactions the charged tracks which do not match to the vertex with the highest sum p_{T} of all particles are removed from the clustering algorithm. The contributions from neutral particles from pileup interactions are removed by subtraction of their average contribution to each jet based on the area of the jet in $\eta - \phi$ space [86]. Each jet has a series of energy corrections to account for any non-linearity in the energy response of the detector. Finally, basic jet identification requirements are applied to ensure each selected jet passes basic quality requirements.

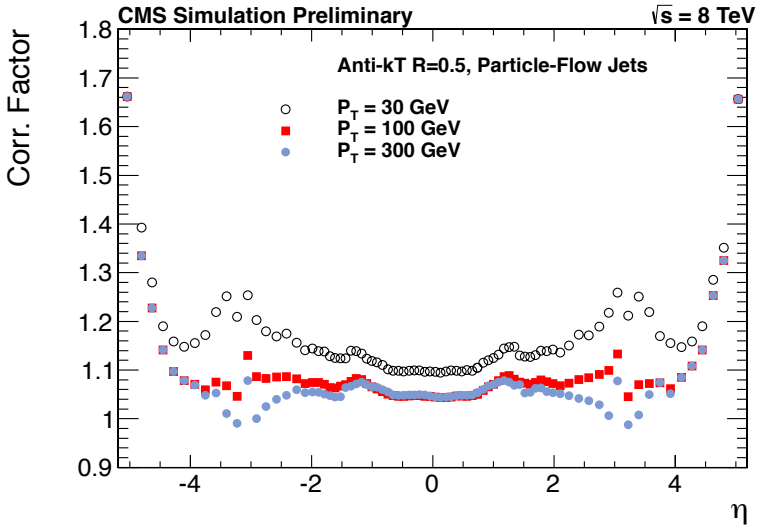
The particle flow reconstruction technique makes use of information from every sub-detector when reconstructing particles [84]. The basic principle is to reconstruct individual particles by matching their signatures in each detector. To keep the algorithm from becoming too complex, the particles are reconstructed in four categories: photons, electrons, muons, and charged pions. Any particle with only an ECAL energy deposit is considered a photon. An ECAL deposit with a matching track is considered an electron. Any track in the central tracker which matches to a track in the muon system is a muon. A cluster of energy in HCAL with at least one corresponding track and ECAL cluster is considered a pion. For each object the p_{T} measurement is taken as a weighted combination of that from the sub-detectors involved in its measurement, increasing the measurement accuracy over the full range of possible p_{T}.

The anti-k_T jet clustering algorithm [85] forms jets by combining particles based on a distance parameter between the ith and jth particles defined as

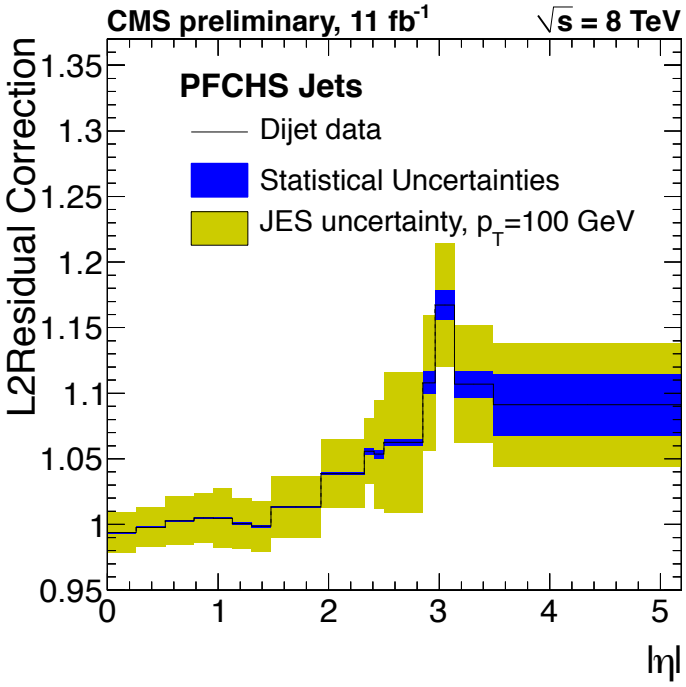
$$d_{ij} = \min(k_{Ti}^{-2}, k_{Tj}^{-2}) \frac{\Delta_{ij}^2}{R^2} \tag{5.1}$$

where $\Delta_{ij}^2 = \Delta Y_{ij}^2 + \Delta\phi_{ij}^2$ is the angular separation between particles i and j, k_{Ti} is the p_{T} of the ith particle, and R is a distance parameter which defines a characteristic size of the jet. Particles are clustered to form jets by combining objects with the smallest distance parameters. A object is flagged as a jet and removed from consideration when $d_{ij} > k_{Ti}^{-2}$ for all j. The algorithm continues to iterate, combining objects and removing "completed" jets, until no particles remain. This algorithm has many theoretical advantages which are detailed in [85]. Additionally, it is computationally faster than equally robust cone-based algorithms. For this analysis, a size parameter of $R = 0.5$ is used because as the standard jet size in CMS, these are the most well understood jets.

The energies of the jets are then corrected using a multistage approach. Jets are corrected for the effects of neutral particles from pileup interactions. Corrections based on the η and p_{T} of the jets are then applied. These corrections are derived from MC studies and are applied to jets in both MC and data. A second set of corrections is then applied only to data, based on studies of dijet, Z+jet, and γ+jet events, to correct for any effects not well modeled in the MC. Jet energy corrections from MC and data studies are shown in figure 5.2 [87]. More details on jet energy corrections can be found in [88]. The average jet energy resolution on 8 TeV data is shown in figure 5.3 [89].



(a)



(b)

Figure 5.2: Plot of jet energy correction factors derived from MC for several example jet p_{T} values (top). Plot of dijet asymmetry corrections derived using data (bottom).

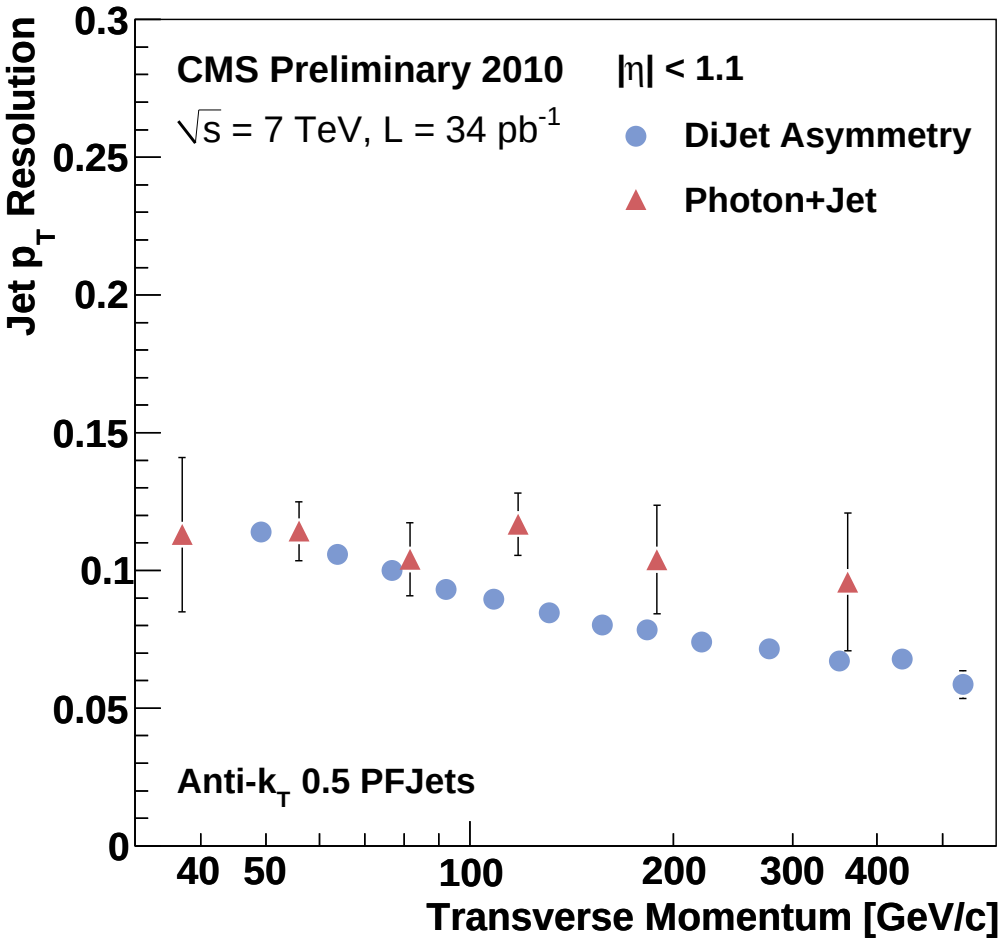


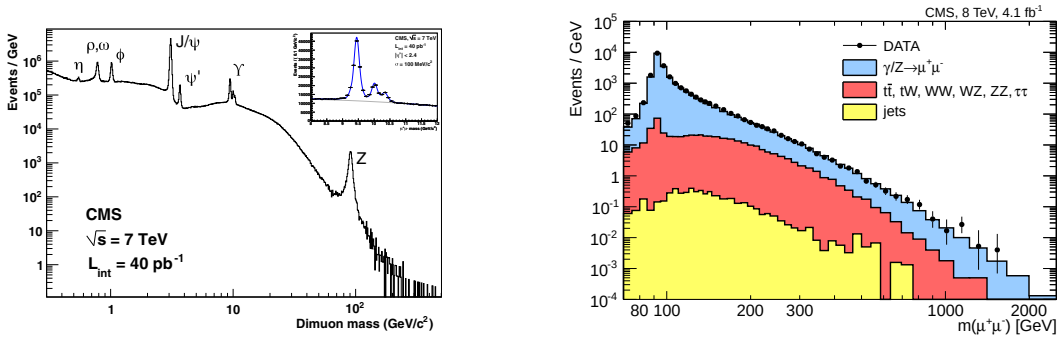
Figure 5.3: The resolution for particle flow jet p_{T} in the central region of CMS.

The final stage of jet selection is the cleaning, identification, and p_{T} requirements. The full particle flow jet collections will contain not only hadronic jets, but leptons as well. Any object in the collection of jets which is identified as a good quality electron or muon is removed from the collection of jets. The remaining jets are required to pass "loose" jet identification which require the jet to have more than one constituent object, the charged and neutral electromagnetic fraction to be less than 0.99, the neutral hadronic fraction to be less than 0.99, and at least one charged hadronic object to have been used in the construction of the jet. Finally, all jets are required to have at least 40 GeV of p_{T}.

5.2.2 Muon Selection

Muons from a W_R decay are expected to have a very large p_{T} due to the large mass of the W_R. This analysis therefore uses muon selection which is optimized for high-p_{T} muons. The CMS recommendation for muon momentum determination with high-p_{T} is the "TuneP" algorithm which combines several reconstruction techniques to yield the best p_{T} resolution [90]. After reconstruction, muons are required to pass tight identification requirements which are also optimized for high-p_{T} muons.

Muon reconstruction starts by identifying possible muon candidates based on a track in either the tracker or the muon chambers. The first algorithm starts with a track, the seed, in the muon chambers and extrapolates it back to a common plane where its position on this common plane is compared to all tracks in the central tracker which are consistent with a muon. If a match is found, the combined object is considered a muon. A second algorithm starts with a seed track in the central tracker and performs the same matching to a track in at least one muon chamber. In either case, if there are multiple matching tracks, the track which is the best match is chosen. The two algorithms are designed to be complementary with the central tracker seed algorithm being more efficient for low p_{T} muons, which may not penetrate the iron of the return yoke to the second muon station, and the muon chamber seeded algorithm providing an efficient and low background method of finding high-p_{T} muons. Plots of the di-muon invariant mass showing low and high p_{T} selections is shown in figure 5.4 [91]. These plots demonstrate the excellent muon resolution of CMS, particularly at low momentum, and show the good agreement between muons reconstructed in data and MC.



(a) (b)

Figure 5.4: Di-muon invariant mass for early 7 TeV data (left) and high mass candidates from 8 TeV data (right).

Once the reconstruction algorithm has picked muons from the possible tracks, the momentum of the muons is reconstructed with the "TuneP" algorithm [90]. This algorithm is designed to be used for any analysis which has a sizable fraction of muons above 200 GeV where the muon chamber information can improve the p_{T} determination. "TuneP" uses three methods of fitting the muon track, selecting the fit with the best goodness of fit, to determine their p_{T}. The first fit uses the central tracker along with the muon stations but removes any muon station hits which are consistent with showering initgiated by the muon in the return yoke. The second fit uses the central tracker with only the first layer of muon chambers to account for the possibility of a hard scatter significantly changing the muon trajectory in the yoke. The third fit uses only the central tracker. This technique significantly improves the muon p_{T} resolution above 200 GeV by reducing incorrect p_{T} assignment due to muon showering at high p_{T}.

Muons used in this analysis are required to pass the "tight high-p_{T}" identification criteria. These selection criteria are designed to reduce muon fakes from jets and non-prompt muon sources and are optimized for good efficiency of high-p_{T} muons. The requirements include that at least one muon chamber be used in the muon track fit, at least six tracker hits be included in the fit, of which one is in the pixel tracker, $\Delta p_{\mathrm{T}}/p_{\mathrm{T}} < 0.3$ for the track used to compute the p_{T}, and the muon track must pass within 2 mm of the beam interaction point in the transverse plane. A full list of requirements can be

	$\|\eta\| < 0.9$	$0.9 < \|\eta\| < 1.2$	$1.2 < \|\eta\| < 2.1$	$2.1 < \|\eta\| < 2.4$
Trigger	0.928 ± 0.0002	0.8302 ± 0.0006	0.8018 ± 0.0005	N/A
High-p_{T} ID	0.9900 ± 0.0003	0.9923 ± 0.0006	0.9949 ± 0.0004	0.9923 ± 0.0012
Isolation	0.9996 ± 0.00001	0.9994 ± 0.0001	0.9997 ± 0.0001	0.9997 ± 0.0001

Table 5.3: Summary of muon ID and isolation data/MC efficiency scale factors, and the trigger efficiency from data, as determined by the muon physics object group using 2012 Monte Carlo and data.

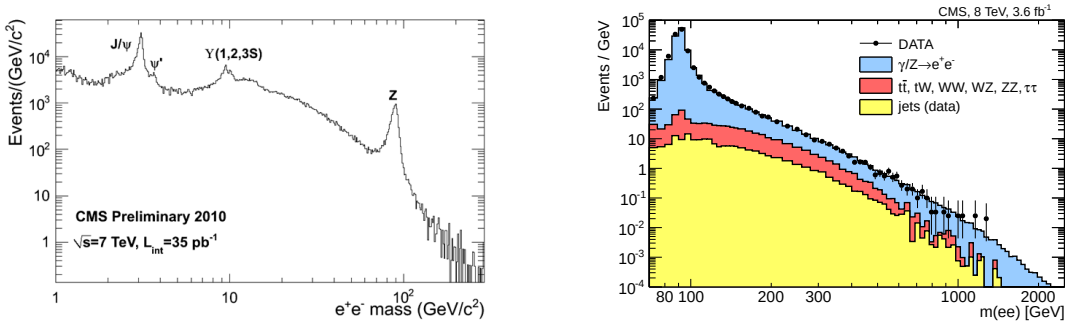
found in ref [92]. Additionally we require that the muon be well isolated in the detector by imposing that the sum of track p_{T} in a cone of $\Delta R < 0.3$ around the muon (not including the muon itself) be less than 10% of the muon p_{T}. These requirements are applied to muons in both data and MC.

In MC, additional corrections are applied to muons to better match the data. These corrections are applied to account for the muon trigger efficiency and differences in identification and isolation efficiency in MC and data. The muon trigger efficiency, identification, and isolation scale factors for data and MC are calculated by the muon physics object working group in CMS [93]. Their results are summarized in table 5.3. The overall trigger efficiency is better than 98% because either muon may pass the single muon trigger. In the case of the trigger efficiency, each event in MC is adjusted by a weight factor dependent on its individual trigger efficiency. The identification and isolation scale factors are derived as a ratio of data to MC efficiency and these scale factors are applied as weight factors to each event.

5.2.3 Electron Selection

Like the muons from W_R decays, electrons are also expected to have a very large E_{T}. The electrons from this analysis are reconstructed based on ECAL energy deposits with additional information from the central tracker [94]. Reconstructed electrons are then subjected to identification and isolation requirements which are optimized for high-E_{T} electrons.

Reconstruction of high-E_{T} electrons starts with identification of an electron-like cluster in ECAL. The ECAL crystals are clustered by a dedicated algorithm which combines them to create possible electron candidates with a narrow structure in η but an elongated structure in ϕ. This cluster shape is caused by bremsstrahlung radiation as



(a) (b)

Figure 5.5: Di-electron invariant mass for early 7 TeV data (left) and high mass candidates from 8 TeV data (right).

an electron passes through the tracker and is bent along the ϕ direction by the magnetic field. Possible electron clusters are matched to either one or three tracks in the outer layers of the tracker. Three tracks are possible when a hard scattering in the tracker causes pair production of another pair of significant E_{T} electrons. The tracker consists of about one radiation length of material so the tracks are reconstructed using a Gaussian-sum filter, a modified approach of the Kalman filter described if ref [95], which accounts for the significant energy loss possible as the electrons traverse the tracker. Traditional electron E_{T} determination uses a weighted sum of track p_{T} and ECAL E_{T}, but this only benefits electrons with $E_{\mathrm{T}} < 20\,\mathrm{GeV}$. To reduce possible issues with bad track fit quality distorting electron energy, the E_{T} of high-E_{T} electrons is taken only from ECAL. However, the η and ϕ of the electron is still taken from the track fit. Plots of the di-electron invariant mass showing low and high p_{T} selections is shown in figure 5.5 [91]. When compared to the di-muon mass plots in figure 5.4, muons have better resolution at low p_{T}, but electrons have better E_{T} resolution at high E_{T}.

The electrons used in this analysis are required to pass high energy electron positron (HEEP) identification and isolation requirements designed to maximize efficiency for high-E_{T} electrons. This selection requires that the electron be within $|\eta| < 2.5$ but excludes the interface region between the ECAL barrel and endcap, $1.442 < |\eta| < 1.56$. The ECAL clusters are subjected to basic shape requirements to ensure they are electron-like and the HCAL over ECAL energy ratio is required to be less than 0.05.

The track is required to pass within 2 mm of the beam interaction point in the transverse plane and no more than one layer of the inner tracker may be missing from the electron track. Isolation requirements are imposed with both tracker and calorimeters [96].

Correction factors are applied to the MC to account for minor differences between data and MC for electrons. The HEEP identification and isolation efficiencies, as calculated by the CMS HEEP group, are in good agreement between MC and data, but a small correction factor of 0.998 in the barrel and 0.979 for endcap electrons is applied as a weight factor on individual events. The trigger efficiency of the double electron trigger used is measured by the $Z' \to ee$ analysis group in CMS to be flat in di-electron mass and electron E_{T} and have an efficiency above 99% for events with two electrons having $E_{\mathrm{T}} > 37\,\mathrm{GeV}$ [91]. No corrections are applied to MC for the trigger efficiency given its high efficiency and flat response.

5.3 W_R Candidate Selection

After defining a set of events and the individual objects, W_R candidates are selected. In each event, a W_R candidate is constructed from the highest p_{T} (E_{T}) muons (electrons) and the two highest p_{T} jets. In the case of the $e\mu jj$ final state used to estimate the $t\bar{t}$ background, the highest p_{T} muon and highest E_{T} electron from the event are chosen. Leptons selected to be part of a W_R candidate must have $p_{\mathrm{T}} > 40\,\mathrm{GeV}$ so that the leptons are above the p_{T} threshold of their respective triggers. The choice to always use the highest p_{T} objects is motivated because M_{W_R} is unknown, but it is heavy enough that its decay products will be the largest p_{T} objects in any given event. An event display showing a good quality event with two muons and two jets can be seen in figure 5.6.

If either of the leptons selected for a W_R candidate are found within $\Delta R < 0.5$ of the each other or the jets selected for a W_R candidate, the event is rejected. This requirement removes any poor quality leptons which may otherwise pass the identification and isolation cuts as good quality leptons candidates are already removed from the jet list.

Two additional requirements are placed on W_R candidates. The highest p_{T} lepton is required to have $p_{\mathrm{T}}(l_1) > 60\,\mathrm{GeV}$ and the di-lepton invariant mass, M_{ll}, is required to be greater than 200 GeV. The $p_{\mathrm{T}}(l_1) > 60\,\mathrm{GeV}$ requirement is chosen because the

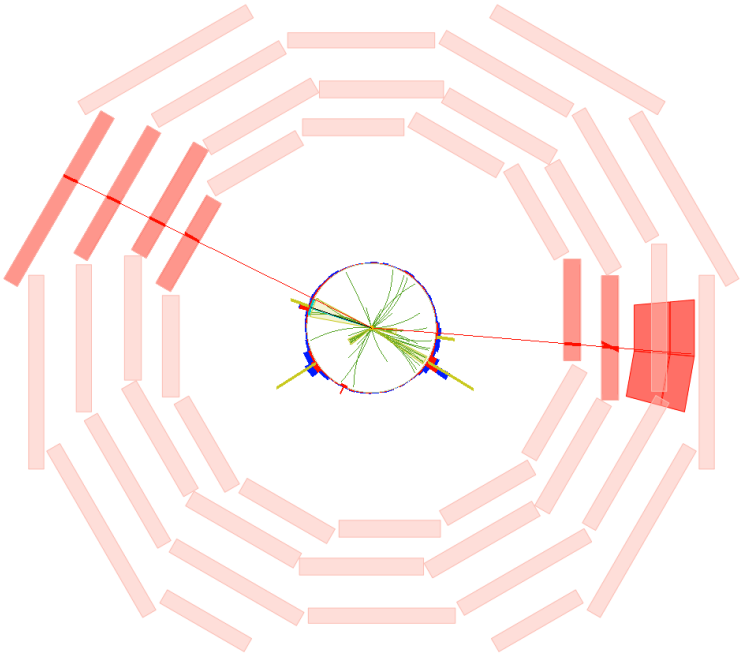


Figure 5.6: An event display showing two good quality muons and two jets satisfying all selection requirements. Muons are represented by the red lines while jets by yellow lines. The green lines are tracks reconstructed in the silicon tracker. The red and blue columns represent the ECAL and HCAL energy deposits respectively. Two additional jets are present in the event, but both are well separated from the selected objects.

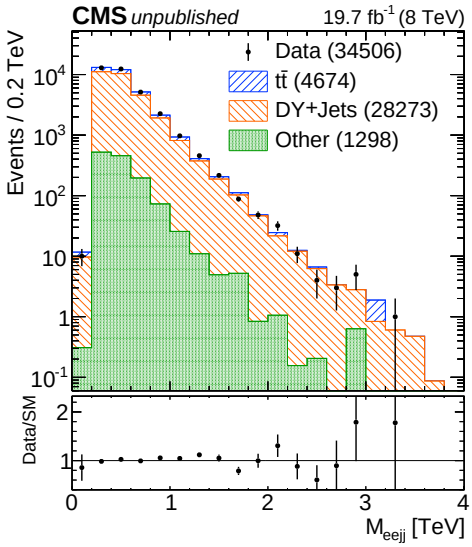
lepton from the initial back to back decay of the W_R into a charged lepton and an N_l will cause the first lepton to have a very large p_{T} for most values of M_N. The data are shown in figures 5.7 and 5.8 for the electron and muon channels respectively after the $p_{\mathrm{T}}(l_1)$ requirement but before the M_{ll} requirement. The distributions show reasonable agreement between the data and the SM predictions discussed in more detail in the next chapter.

The selection requirements for this analysis are designed to maximize the signal versus background while keeping the selection simple to allow for possible reinterpretation of the result with other theoretical frameworks. The efficiency of the above selection for signal like events is estimated using signal MC generated as described in section 5.1.2. The efficiency times acceptance found by applying the full event selection is shown in figure 5.9. As M_{W_R} increases so does the efficiency because the decay products of lighter

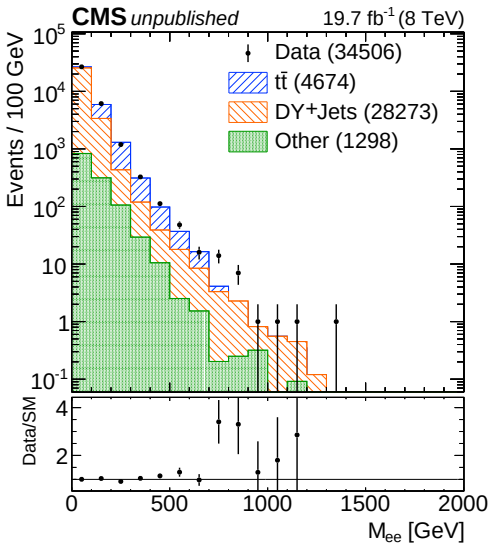
Electron Channel				
M_{W_R}, M_N (TeV)	$eejj$	$p_{\mathrm{T}}(e_1)$	M_{ee}	$M_{eejj} > 600\,\mathrm{GeV}$
2.0, 1.67	56.81(56.81)	56.81(99.99)	53.61(94.37)	53.58(99.95)
2.5, 1.25	57.80(57.80)	57.80(100.00)	56.91(98.46)	56.90(100.00)
3.0, 0.75	46.72(46.72)	46.72(100.00)	45.83(98.09)	45.77(99.87)
Muon Channel				
M_{W_R}, M_N (TeV)	$\mu\mu jj$	$p_{\mathrm{T}}(\mu_1)$	$M_{\mu\mu}$	$M_{\mu\mu jj} > 600\,\mathrm{GeV}$
2.0, 1.67	69.58(69.58)	68.80(100.00)	64.96(94.41)	64.94(99.97)
2.5, 1.25	69.73(69.73)	68.92(100.00)	67.86(98.47)	67.86(100.00)
3.0, 0.75	56.87(56.87)	56.37(100.00)	55.39(98.26)	55.37(99.96)

Table 5.4: Absolute(relative) efficiency for each selection stage for selected signal points. At the $lljj$ selection stage the four objects with minimum p_{T} and ΔR requirements are required. The $p_{\mathrm{T}}(l_1)$ and M_{ll} stages apply the $p_{\mathrm{T}}(l_1) > 60\,\mathrm{GeV}$ and $M_{ll} > 200\,\mathrm{GeV}$ cuts in succession. Finally, the signal efficiency for the region $M_{lljj} > 600\,\mathrm{GeV}$ is shown.

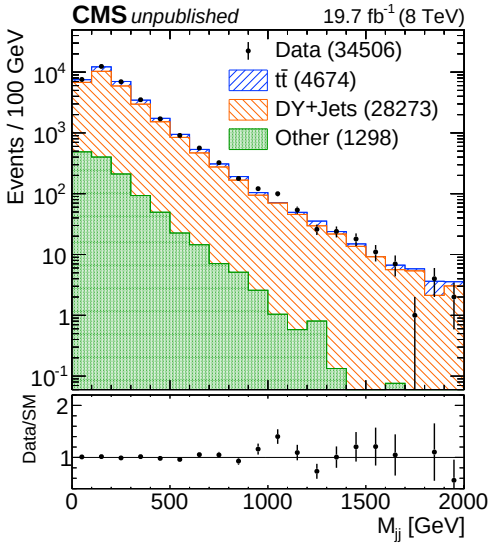
W_R candidates may miss the p_{T} requirements on the individual objects. The general decrease in acceptance when $M_N \ll M_{W_R}$ is due to N being highly boosted, so its decay products fail the isolation or ΔR requirements. The overall efficiency for muons is higher because the efficiency of muon reconstruction and identification is higher than for electrons. The relative and overall efficiencies are shown for each stage of cuts for three representative signal points in table 5.4. The majority of the signal loss occurs during the initial object selection, with the remaining cuts being more than 95% efficient for most M_{W_R} and M_N assumptions. Sample signal shapes for the primary selection variables as well as the ljj masses are shown in figures 5.10 and 5.11. Depending on the mass of the heavy neutrino, either the ljj mass constructed with the leading or sub-leading lepton will show a resonant peak. The M_{lljj} mass distributions are shown in figure 5.12. As these show a clear resonance structure for all M_{W_R} and M_N assumptions the M_{lljj} distributions are used as the final variable of merit to search for possible W_R production and in limit setting.



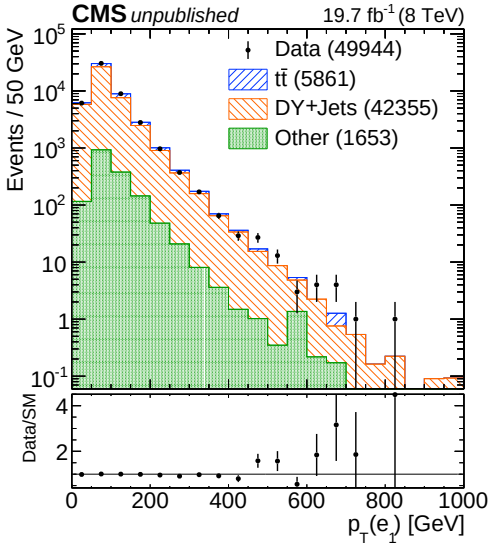
(a) M_{eejj}



(b) M_{ee}

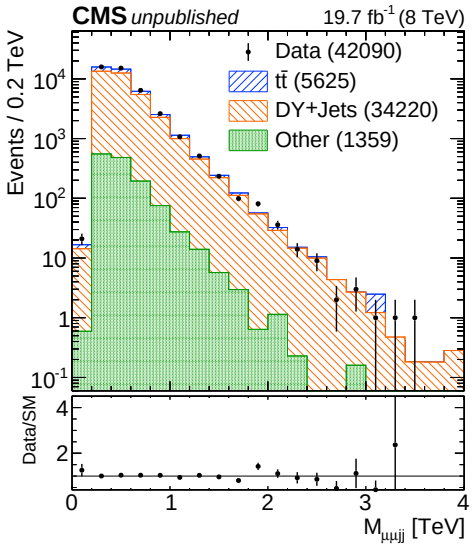


(c) M_{jj}

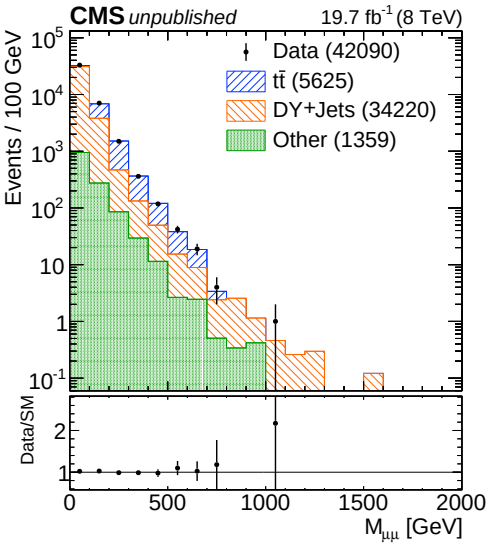


(d) $p_{\mathrm{T}}(e_1)$

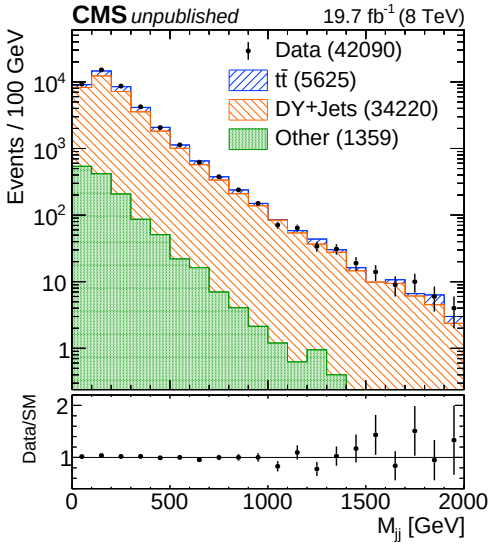
Figure 5.7: Plots showing electron channel data before the $M_{ll} > 200\,\mathrm{GeV}$ cut is applied.



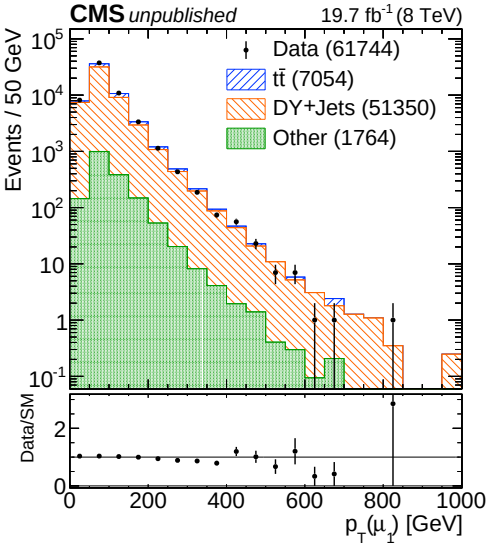
(a) $M_{\mu\mu jj}$



(b) $M_{\mu\mu}$

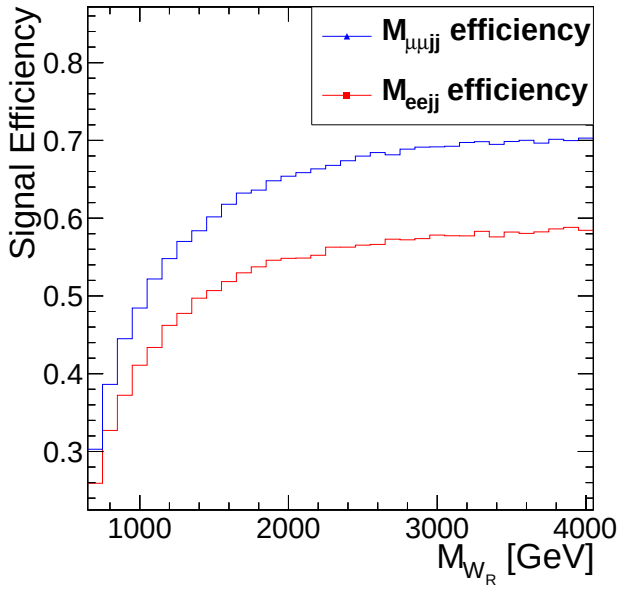


(c) M_{jj}

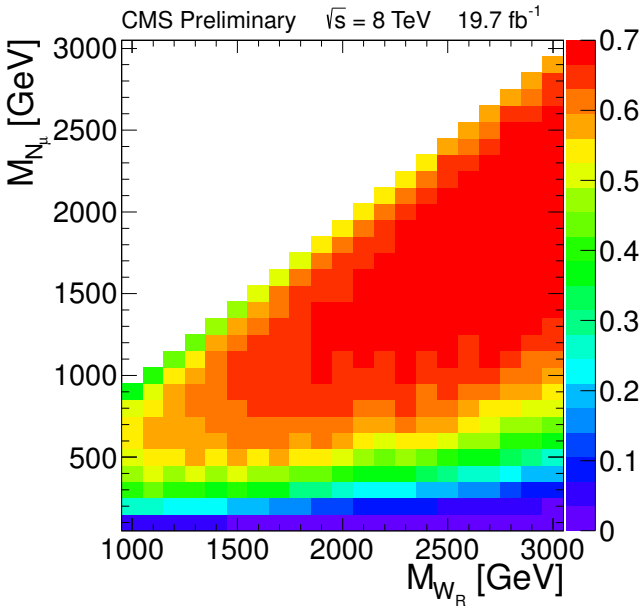


(d) $pt(\mu_1)$

Figure 5.8: Plots showing muon channel data before the $M_{ll} > 200\,\text{GeV}$ cut is applied.

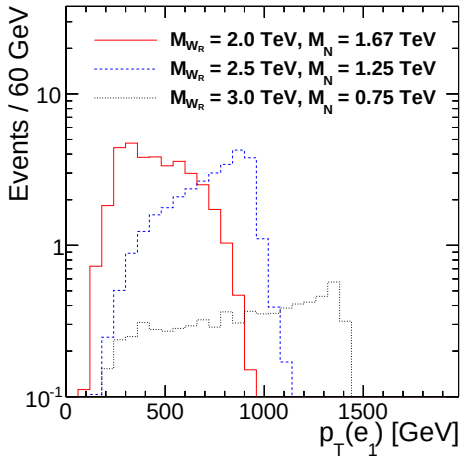


(a)

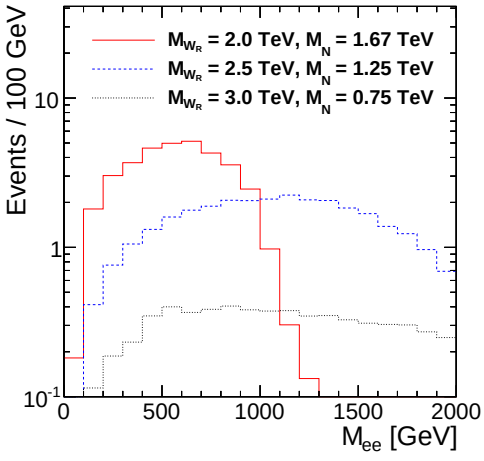


(b)

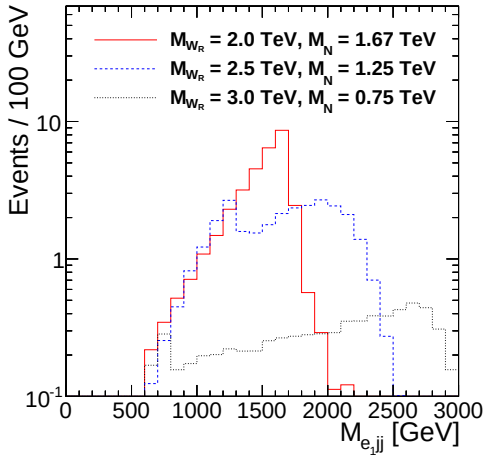
Figure 5.9: Signal efficiency for final selection as a function of the chosen M_{W_R} (top) and as a function of both M_{W_R} and M_N (bottom).



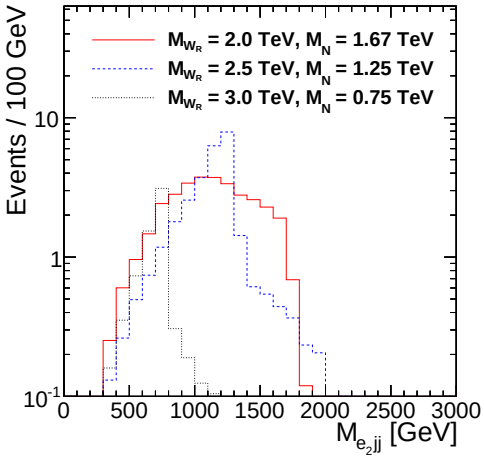
(a) $p_{\mathrm{T}}(e_1)$



(b) M_{ee}

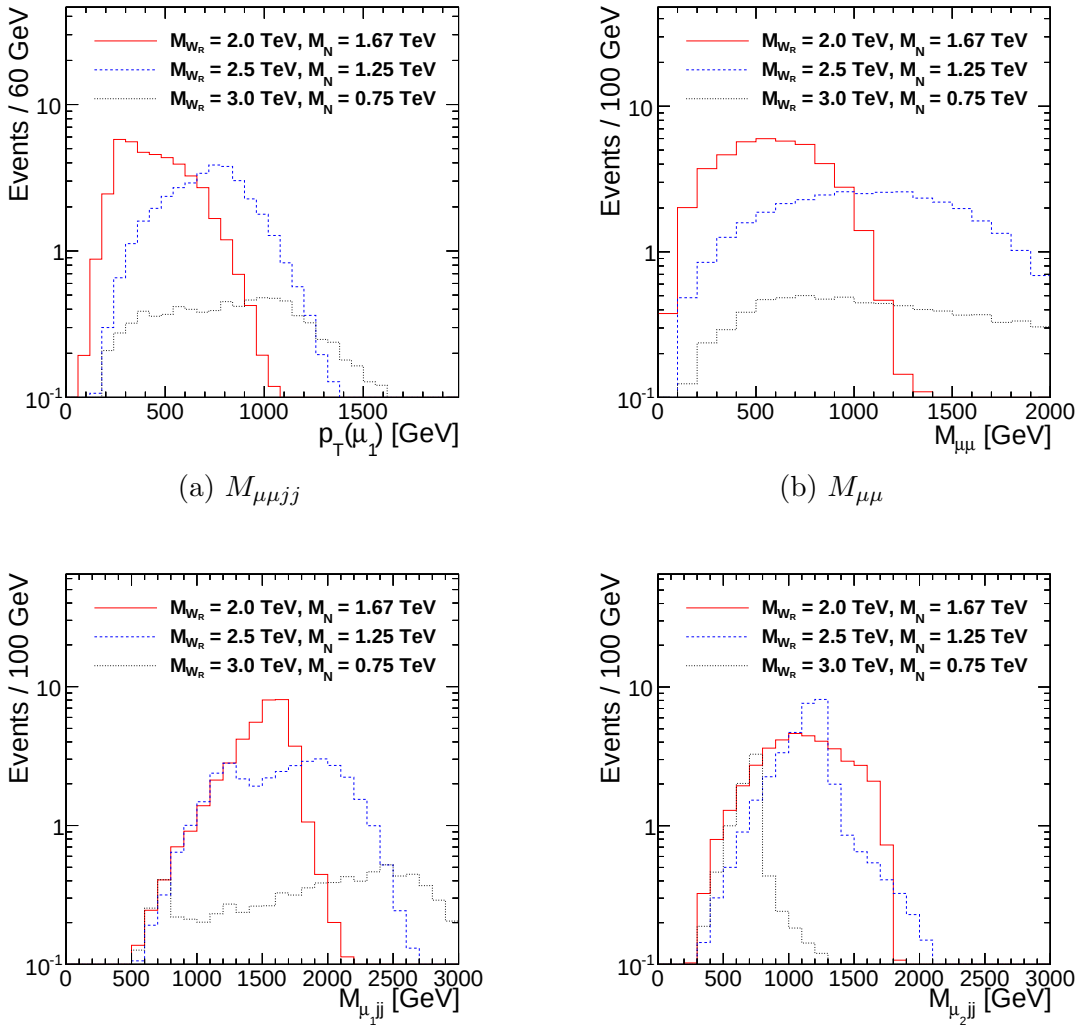


(c) $M_{ejj}(e_1)$



(d) $M_{ejj}(e_2)$

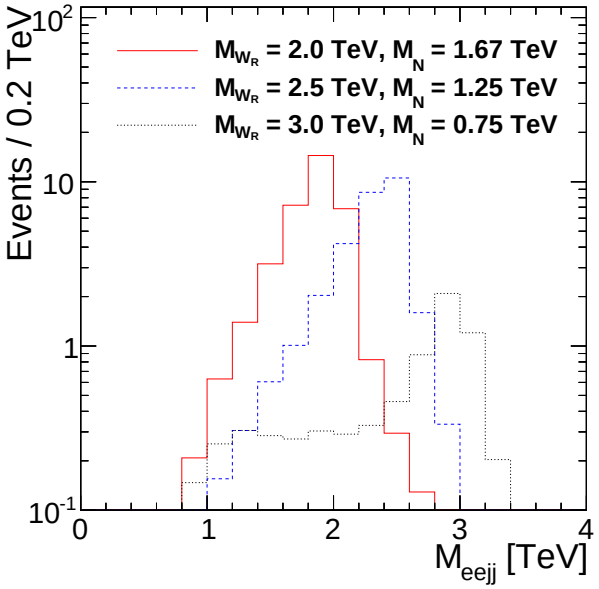
Figure 5.10: Plots showing sample signal distributions for the electron channel. Plot (a) shows the p_{T} distribution of the leading electron. Plot (b) shows the M_{ee} mass after the $p_{\mathrm{T}}(e_1) > 60\,\mathrm{GeV}$ cut is applied. Plots (c) and (d) show the ejj masses after both the $p_{\mathrm{T}}(e_1)$ and M_{ee} cuts have been applied.



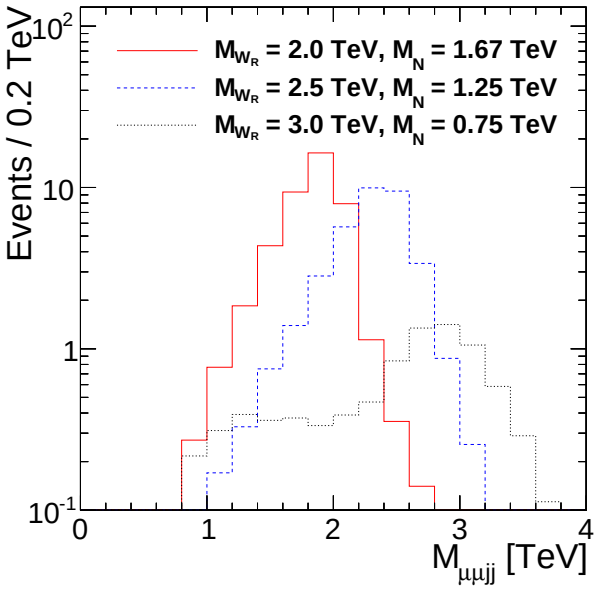
(a) $M_{\mu\mu jj}$ (b) $M_{\mu\mu}$

(c) $M_{\mu jj}(\mu_1)$ (d) $M_{\mu jj}(\mu_2)$

Figure 5.11: Plots showing sample signal distributions for the muon channel. Plot (a) shows the p_{T} distribution of the leading muon. Plot (b) shows the $M_{\mu\mu}$ mass after the $p_{\mathrm{T}}(\mu_1) > 60\,\mathrm{GeV}$ cut is applied. Plots (c) and (d) show the μjj masses after both the $p_{\mathrm{T}}(\mu_1)$ and $M_{\mu\mu}$ cuts have been applied.



(a)



(b)

Figure 5.12: Sample signal shapes for variable of merit M_{lljj} for electron (top) and muon (bottom) channels.

Chapter 6

Background Estimation

After selecting the events and defining the signal region in chapter 5, the background contamination of the signal region is discussed here. All important sources of background contain at least two real leptons due to the two high-p_{T} leptons required in the final state. The dominant sources of background are pair-production of top quarks, referred to as $t\bar{t}$ production, and Z and γ^* production in association with jets, referred to as DY+jets production. The remaining background contributions come from pair production of vector bosons, single top production with a W boson, and W boson scattering. Possible contributions from QCD multijet backgrounds are found to be negligible. The following chapter details how each of these backgrounds is estimated and any cross-checks performed.

6.1 $t\bar{t}$ Background Estimation

The overall dominant background for the $eejj$ and $\mu\mu jj$ final states comes from pair-production of top quarks. The dominant decay mechanism of top quarks is $t \to Wb$. The purely hadronic decays of the top to three jets is an insignificant source of background, but the leptonic decays of the t will produce a real charged high-p_{T} lepton. This leads to the final state of $t\bar{t} \to l\bar{l}'b\bar{b}\nu\bar{\nu}'$. The neutrinos will be undetectable in CMS and therefore the effective final state is the same as in this analysis. The $t\bar{t}$ decay to leptons is shown in figure 6.1.

The estimation of the $t\bar{t}$ contribution to the signal region is based primarily on

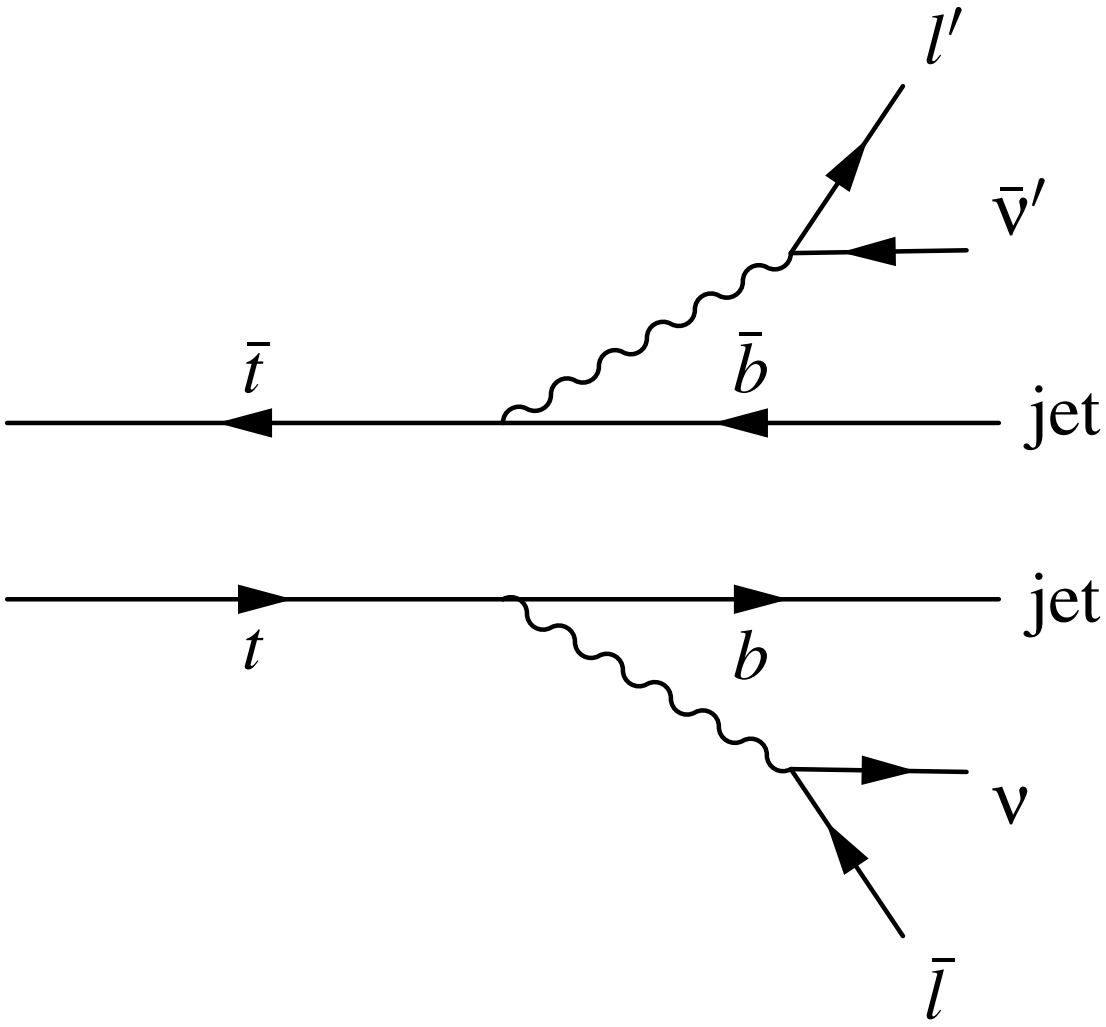


Figure 6.1: Decay of $t\bar{t}$ including two leptons and two jets.

data with a final state of $e\mu jj$ selected as described in chapter 5. This control region is expected to be highly enriched in $t\bar{t}$ events because they are the only SM process with a high production rate for two different flavor leptons in association with two jets. Furthermore, this final state is expected to be signal free because the W_R decays produce two same flavor leptons. A possible exception to this is a tau flavor heavy neutrino where the two taus decay to an electron and a muon, but this can contribute at most at the level of 3 - 5% of the total $t\bar{t}$ background in the signal region. The tau flavor heavy neutrino decays will also contribute at the same level in all three possible final states and although this would make the measurement insensitive to a tau flavor heavy neutrino, it would not interfere with the electron or muon flavor search results.

The shape of the $t\bar{t}$ distribution is estimated by applying the full event selection to the $e\mu jj$ channel data. Although this consists largely of $t\bar{t}$ events, some other backgrounds, primarily vector boson pair production and tW production, contribute to this final state at a small level. To account for this small contamination, the contributions from these backgrounds are subtracted using MC simulation to which the full $e\mu jj$ selection has been applied. The $e\mu jj$ distribution with background included is shown in figure 6.2. Also included in the figure is an exponential fit to the distribution starting at 800 GeV. In the high statistics region of this distribution, the distribution is expected to be exponential as seen in $t\bar{t}$ MC. The fit is generally in good agreement with the data, but in the bin from 1.0 to 1.2 TeV a significant downward deviation of $\sim 1.85\sigma$ is found. This deviation is not present in MC and so we chose to replace this bin with the value derived from the exponential fit.

Naively the background estimation from the $e\mu jj$ channel would need to be normalized by a factor of 0.5 to account for the combinatoric factor leading to twice as many cross flavor decays of $t\bar{t}$ as same flavor decays. In reality, this factor is different for both electron and muon channels due to differences in identification, isolation, and trigger efficiency for each lepton. The normalization factors are calculated using three different techniques. The first technique, used for the final normalization factors, uses the ratio of events in $t\bar{t}$ MC passing the final selection with $M_{W_R} > 600$ GeV in each final state, $N_{eejj}/N_{e\mu jj}$ and $N_{\mu\mu jj}/N_{e\mu jj}$, to estimate the scale factors. Two estimation techniques based on data are also used as cross checks. The results of all three methods are summarized in table 6.1.

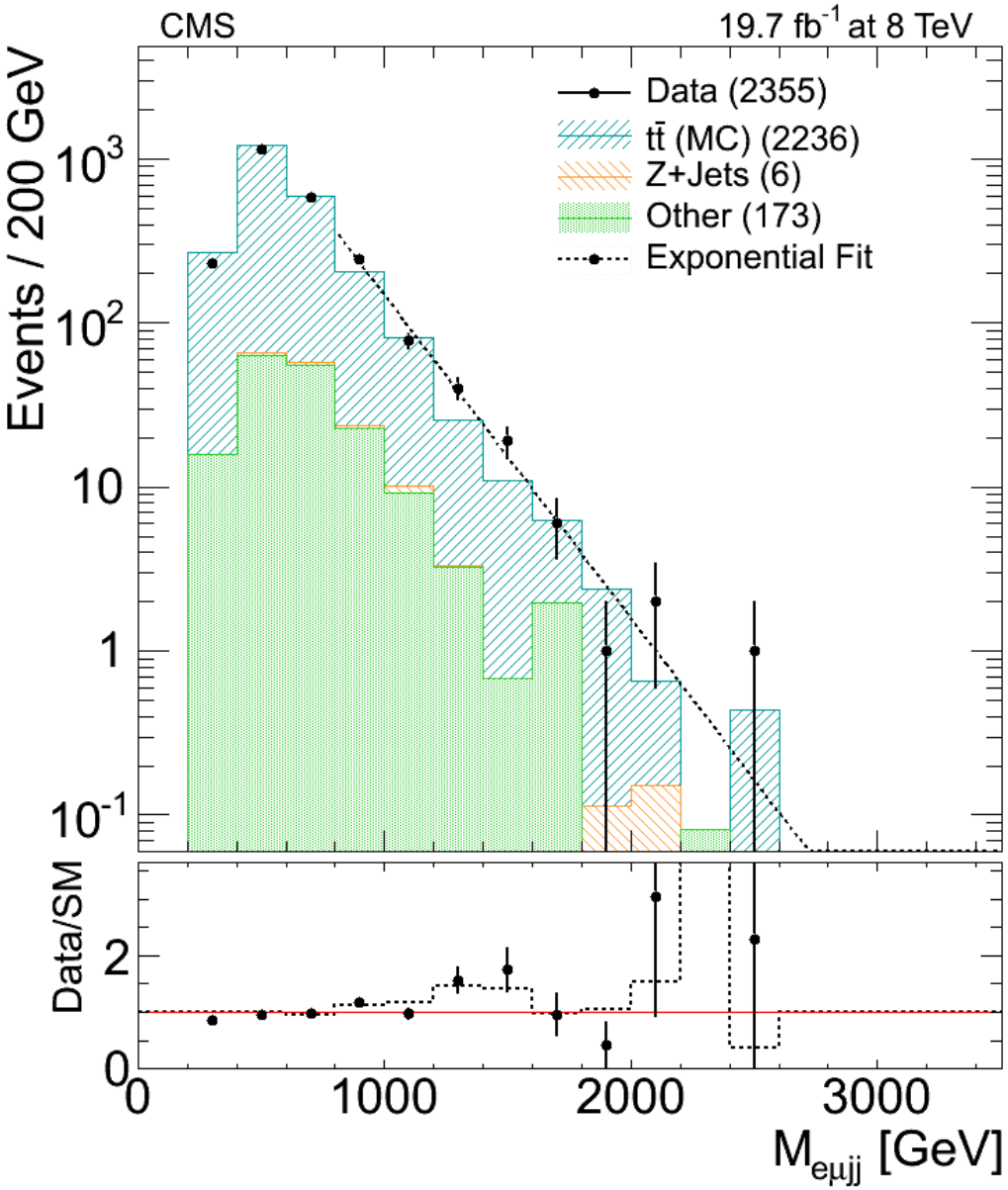


Figure 6.2: $e\mu jj$ channel data compared to MC prediction and an exponential fit.

Method	$ee jj / e\mu jj$	$\mu\mu jj / e\mu jj$
$t\bar{t}$ MC	0.524 ± 0.007	0.632 ± 0.008
efficiency	0.517 ± 0.010	0.626 ± 0.012
b-tagging	0.53 ± 0.02	0.65 ± 0.03

Table 6.1: Normalization factors for $t\bar{t}$ background estimate.

The first alternate estimation approach is based on the relative efficiency of each final state. The event yield in each final state can be written (up to a common numerical factor) as

$$N^{t\bar{t}}_{eejj} \propto \epsilon_e^2, \quad N^{t\bar{t}}_{\mu\mu jj} \propto \epsilon_\mu^2 \left(1 - (1 - \epsilon_{trig})^2\right), \quad N^{t\bar{t}}_{e\mu jj} \propto 2\epsilon_e \epsilon_\mu \epsilon_{trig} \tag{6.1}$$

where ϵ_e and ϵ_μ are the overall reconstruction, identification, and isolation efficiencies for electrons and muons respectively, and ϵ_{trig} is the muon trigger efficiency. The electron trigger efficiency is taken to be 100% and therefore ignored for this calculation. The normalization factors for the electron and muon channels can then be written as

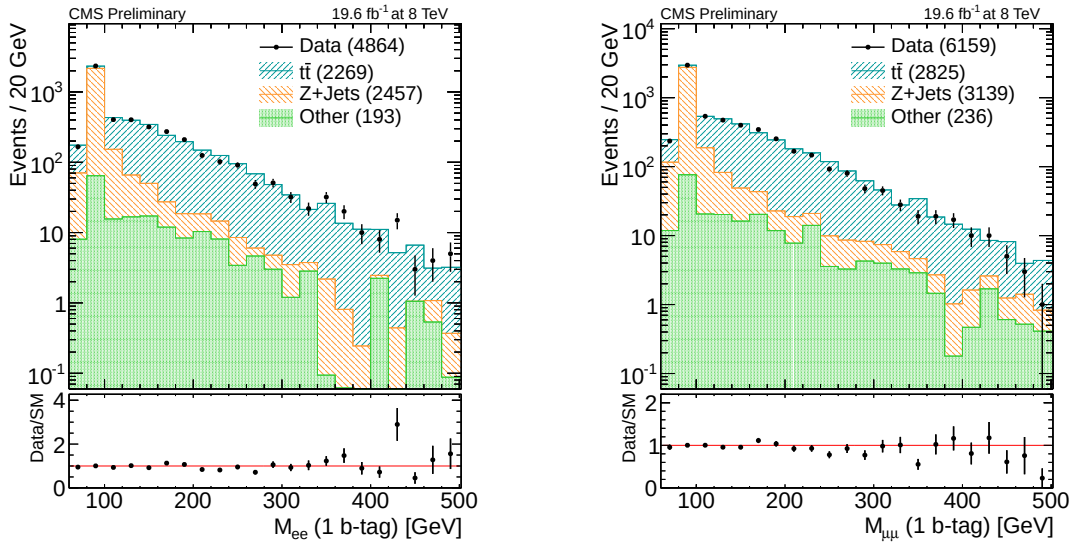
$$k_e \equiv \frac{N^{t\bar{t}}_{eejj}}{N^{t\bar{t}}_{e\mu jj}} = \frac{1}{2}\frac{\epsilon_e}{\epsilon_\mu}\frac{1}{\epsilon_{trig}}, \quad k_\mu \equiv \frac{N^{t\bar{t}}_{\mu\mu jj}}{N^{t\bar{t}}_{e\mu jj}} = \frac{1}{2}\frac{\epsilon_\mu}{\epsilon_e}\left(2 - \epsilon_{trig}\right). \tag{6.2}$$

The average muon trigger efficiency can be calculated from the muon η distribution and the trigger efficiencies in table 5.3 to be 87.2%. The normalization factors in 6.2 depend only on the ratio of efficiencies, ϵ_e/ϵ_μ. This ratio is approximated using the control region $120 < M_{ll} < 200\,\text{GeV}$ as

$$\frac{\epsilon_e}{\epsilon_\mu} = \sqrt{\frac{N_{eejj}(120 < M_{ll} < 200\,\text{GeV})}{N_{\mu\mu jj}(120 < M_{ll} < 200\,\text{GeV})}} \tag{6.3}$$

where $N_{eejj}(120 < M_{ll} < 200\,\text{GeV})$ and $N_{\mu\mu jj}(120 < M_{ll} < 200\,\text{GeV})$ and the number of events in data in the control region for electron and muon events passing all other final selection criteria. This approximation holds because the lepton fake rate is negligible (as will be shown in section 6.4) and the muon trigger efficiency is nearly 100% for $\mu\mu jj$ events as either muon may trigger the event. The scale factors calculated with this method are shown in the second row of table 6.1.

The third technique used to calculate the scale factors uses b-tagging to identify jets which originate from b quarks. Jets originating from b quarks can be distinguished from



(a) (b)

Figure 6.3: Plot of M_{ll} for events with one b-tagged jet for $eejj$ channel (left) and $\mu\mu jj$ channel (right).

the lighter quarks and gluons because B hadrons have a characteristic lifetime that is distinct from both promptly decaying and long lived hadrons containing light quarks. This causes the b-jets to have an offset vertex from the primary vertex of the proton-proton collision on the order of $400\,\mu$m. Other characteristics, such as the number of tracks and the spread of the tracks, can also be used to distinguish b-jets from light quark and gluon induced jets. This analysis uses the "combined secondary vertex" method which combines these techniques, as described in [97], to tag b-jets with an efficiency of 63% and a fake rate much less than a percent. The b-jet selection is then applied to the final selection for all three final states and the scale factors for muon and electron data are derived by the ratio of events in the region $120 < M_{ll} < 200\,\text{GeV}$. Contributions from other SM backgrounds are subtracted using MC. The M_{ll} distributions are shown in figure 6.3 and the normalization factors are shown in the third line of table 6.1. All three methods of determining the normalization factors are consistent within their statistical uncertainties.

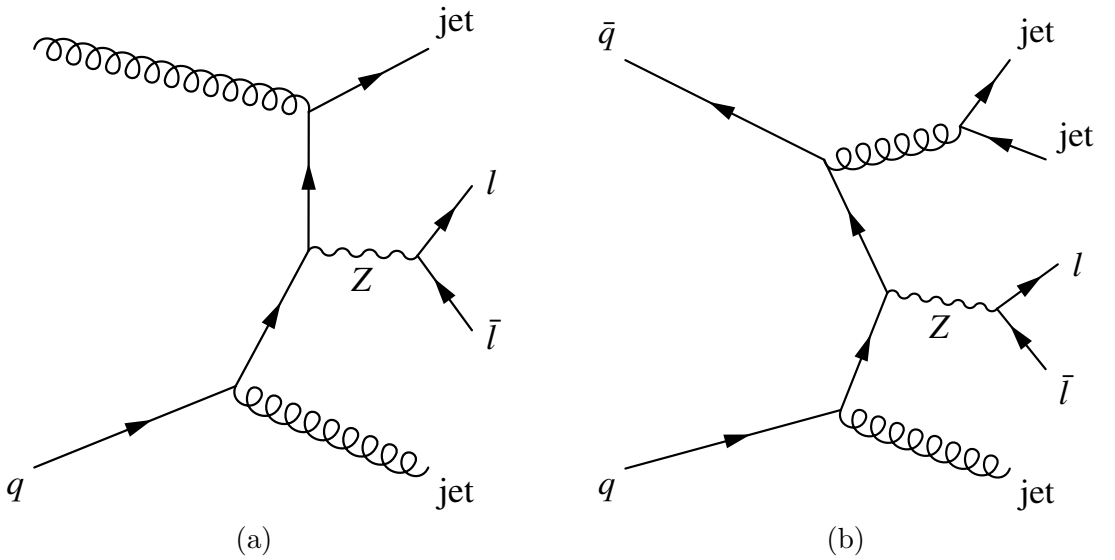


Figure 6.4: Possible DY+jets production diagrams with two jets (left) and three jets (right).

6.2 DY+jets Background Estimation

The second largest background overall and the largest at high M_{lljj} is Z^* and γ^* produced in association with jets, referred to as DY+jets. In this process a Z^* (or γ^*), which decays to a pair of leptons, is produced with one or multiple initial state radiation (ISR) jets. ISR jets happens when the particles which are fusing to create the boson emit a quark or gluon which hadronizes into a jet. ISR may happen multiple times and can lead to multiple high-p_{T} jets being produced along with the vector boson. Examples of processes which may lead to $eejj$ or $\mu\mu jj$ final states are shown in figure 6.4.

The DY+jets background is estimated using MadGraph MC which is generated in multiple samples based on the number of jets produced in association with the boson as well as an inclusive sample used to estimate the normalization and zero jet sample. The individual samples are generated with the assumption of one, two, three, and four hard ISR jets produced with the boson. This is in addition to any jets which may arise from pileup collisions. These individual samples are used in favor of the inclusive sample with all jet multiplicities included because the selection requiring two jets with

$p_{\mathrm{T}} > 40\,\mathrm{GeV}$ heavily biases the analysis to the two and higher jet content events. The individual samples allow for these jet multiplicities to be generated at a much higher population with respect to the inclusive sample. Any possible contribution from the zero jet category is estimated by selecting only the zero jet events from the inclusive DY+jets MC.

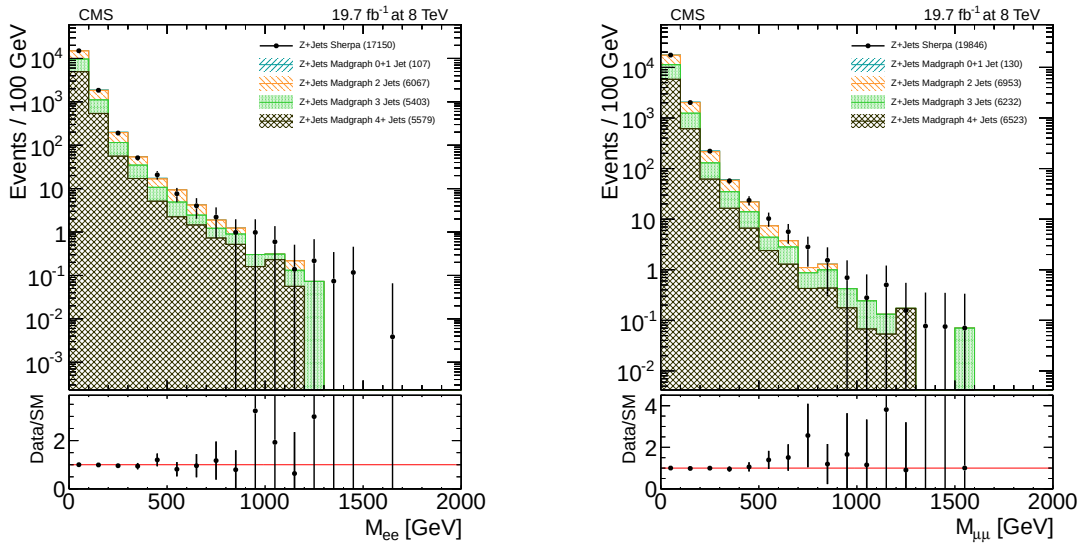
The individual jet multiplicity samples are normalized based on their effective integrated luminosities, $\int\mathcal{L} = N/\sigma$ calculated based on σ, the leading order (LO) cross-section of the sample, and N, the number of events generated. These numbers are shown in table 5.2. Each sample is normalized such that its effective integrated luminosity matches the integrated luminosity of data. All samples are then summed to find the total inclusive DY+jets shape, shown in figure 6.5. As a cross-check we also show the DY+jets estimate made using the SHERPA generator. The SHERPA sample overall normalization is fixed to the MADGRAPH sample in the Z peak region. The majority of events come from the two, three, and four jets samples, this is strong evidence that pileup jets are not a major contribution in this analysis. This is largely due to the soft nature of most pileup jets, particularly after the pileup jet cleaning is applied.

The whole DY+jets distribution is normalized to the next to next to leading order (NNLO) cross-section as calculated with FEWZ [98], a tool for calculating NNLO cross-sections of electroweak processes. The normalization factor is calculated as

$$k_{\mathrm{NNLO}} = \frac{\sigma_{\mathrm{NNLO}}\epsilon_{Inc}}{\sum_i \sigma_i\epsilon_i} \tag{6.4}$$

where σ_{NNLO} is the NNLO cross-section listed for the inclusive DY+jets sample in table 5.2, σ_i are the LO cross-sections of the individual jet multiplicity samples, ϵ_{Inc} is the cut efficiency of the inclusive sample, and ϵ_i are the cut efficiencies of the individual jet samples. The cut efficiencies are calculated as the number of events passing all selections divided by the total number of events in each sample. For both the electron and muon channel background estimate this normalization factor is ~1.2.

Finally, the normalization of the DY+jets background estimation is checked using data in the Z peak region of $60 < M_{ll} < 120\,\mathrm{GeV}$. The intra-normalization of the individual jet multiplicity samples is kept fixed, but the overall normalization is allowed to float in the normalization region. Contributions from other backgrounds are included

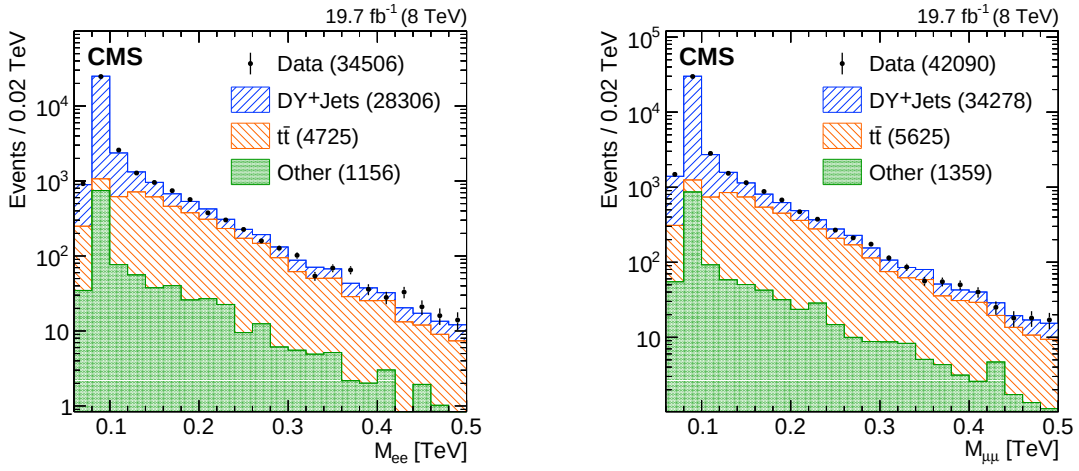


(a) (b)

Figure 6.5: M_{ll} distribution broken down for different jet multiplicity samples for the electron channel (left) and muon channel (right). A SHERPA DY+jets sample is also shown for comparison.

using MC estimates, or data in the case of $t\bar{t}$, with their normalization fixed. A log-likelihood fit is performed comparing the DY+jets background estimate to data in the Z peak region to extract the normalization factors. The result after normalizing the DY+jets contribution can be seen in figure 6.6. The distribution of M_{ll} is in good agreement for both the electron and muon channel data over the entire mass range even though the normalization is only done in the $60 < M_{ll} < 120\,\text{GeV}$ region. The scale factors for the muon and electron channel are 1.027 ± 0.006 and 1.000 ± 0.007 respectively. The uncertainties are based on MC and data statistics.

To negate the contribution of the DY+jets and other Z containing backgrounds the $M_{ll} > 200\,\text{GeV}$ requirement is imposed. From figure 6.6 it can be seen that this cut effectively removes close to 99% of the DY+jets background. It also retains most to all of the signal efficiency due to the very large lepton p_{T} expected in W_R decays.



(a) (b)

Figure 6.6: M_{ll} distribution for electron channel (left) and muon channel (right) after normalization in the Z peak region.

6.3 Other Electroweak Backgrounds

The remaining significant backgrounds all come from electroweak sources. The signals contributing in this category are vector boson pair production, WW, WZ, and ZZ, single top production in association with a W boson, tW, and W boson scattering, $W^{\pm}W^{\pm}jj$. All these backgrounds sources are capable of producing two or more same flavor charged leptons and up to two jets. Processes that do not produce two jets directly, like WW, can pick up additional jets through ISR, and pileup. Example processes which can contribute to $eejj$ and $\mu\mu jj$ final states are shown in figure 6.7. Note that several processes, including WW, $W^{\pm}W^{\pm}jj$, and tW, have small but non-negligible contributions to the $e\mu jj$ final state as well.

These remaining backgrounds are all estimated from MC simulation. The di-vector boson production is estimated from MC generated with PYTHIA, the tW from MC generated with POWHEG, and the $W^{\pm}W^{\pm}jj$ from MC generated with MADGRAPH. Each MC sample is normalized to the integrated luminosity of data using the NLO cross-section for the particular process as listed in table 5.2. The breakdown of the "Other" category of backgrounds is shown in figure 6.8. Contributions from W+jets

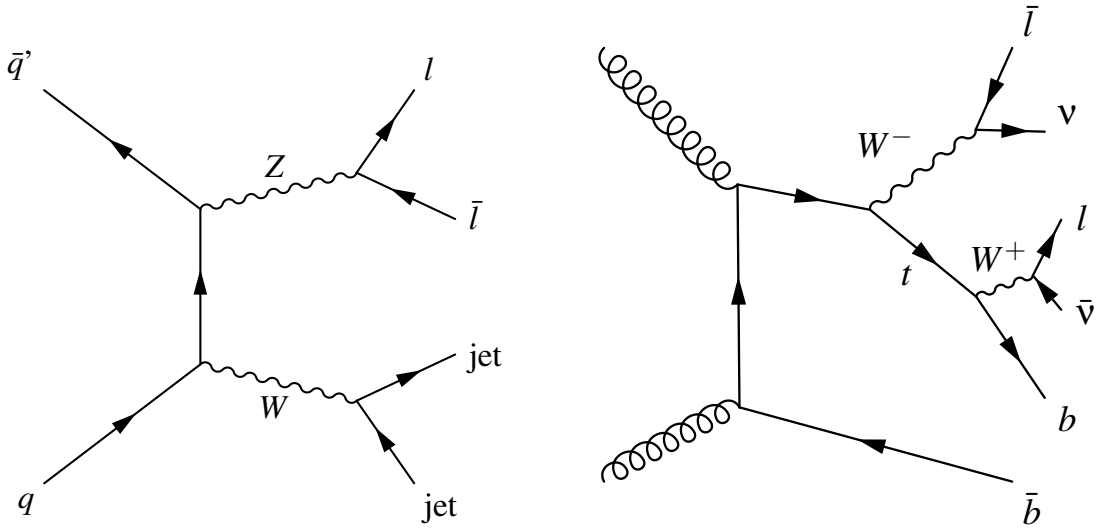


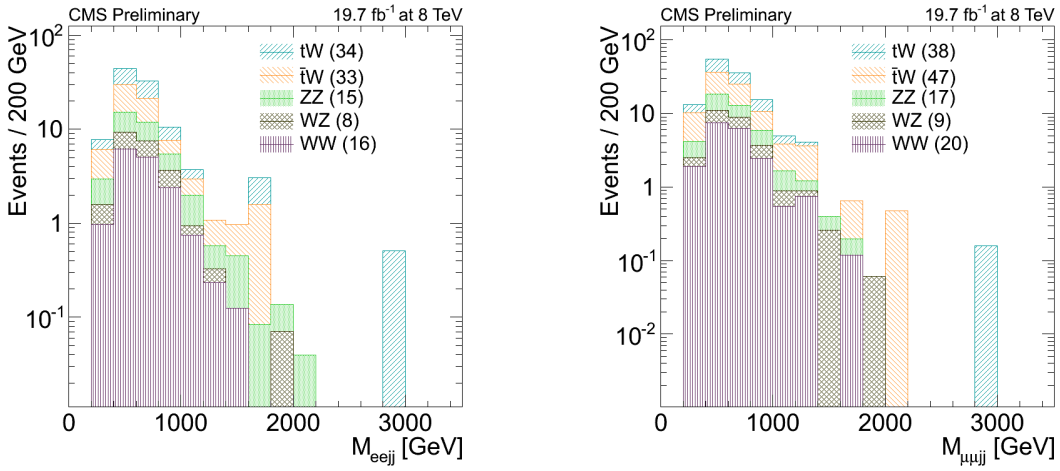
Figure 6.7: Examples of ZW production (left) and tW production (right).

and γ+jets processes are neglected as they contribute at a rate much less than any of the other background in this category due to their inability to produce two real high-p_{T} leptons.

6.4 QCD Background Estimation

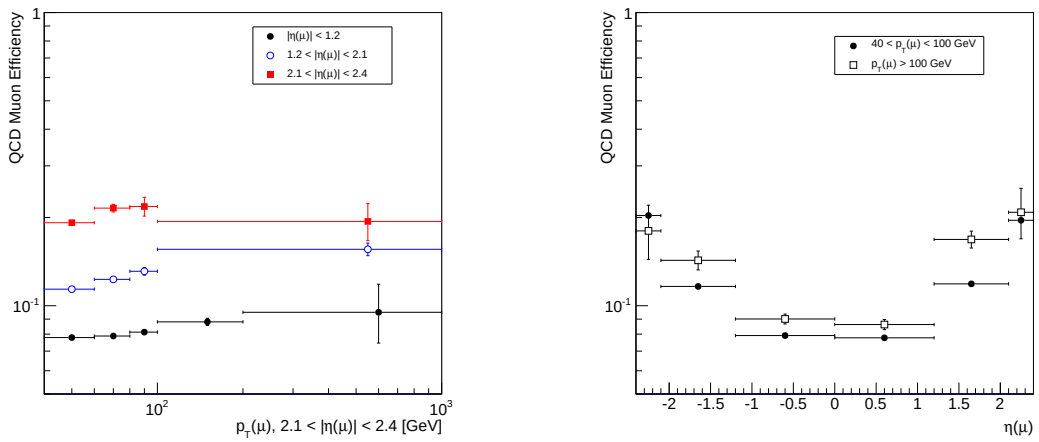
The final background considered in this analysis is possible contributions from quantum chromodynamics (QCD) multijet production. This process is very prevalent in the environment of a hadron collider such as the LHC, but will not make a large contribution to the background as it would require two jets to be mistaken for high-p_{T} leptons. QCD backgrounds are estimated using techniques based primarily on data because simulations of QCD interactions are particularly difficult due to both theoretical limitations and limited computational resources. The QCD background is estimated by first determining the probability that a jet will be mistakenly identified as a lepton. This fake rate is then applied to a multijet control sample to estimate the possible background distribution.

For electrons, the fake rate calculation determined by the CMS $Z' \to ee$ analysis team is used [52]. They estimate the electron fake rate using a data sample collected with



(a) (b)

Figure 6.8: Sum of all backgrounds in "other" category for electron (left) and muon (right) channel.



(a) (b)

Figure 6.9: Muon fake rate as a function of p_{T} (left) and η (right).

pre-scaled single photon triggers which require one and only one ECAL cluster greater than 10 GeV. The reconstructed electrons have minimal selection requirements applied. The fake rate is then defined as the number of electrons passing the full HEEP electron selection divided by the number passing the basic requirements. Contamination of the sample by events containing real leptons, primarily W+jets and γ+jets, are accounted for using MC. The fake rate is then parameterized as a function of E_{T} and η.

The muon fake rate is estimated in a similar way to the electron fake rate. A jet which fakes a muon is only likely to do so if the jet has produced a real muon in the decay cascade during hadronization. To study this a sample of μ+dijet events with low missing E_{T} and no additional high-p_{T} jets is selected. Muons are selected which are within $\Delta R < 0.3$ of one of the two jets and loosened muon identification requirements are applied. The fake rate is defined as the ratio of events passing the full tight high-p_{T} muon identification plus isolation requirements to the events passing the loosened requirements. The muon fake rate as a function of p_{T} and η is shown in figure 6.9.

Once the fake rates are derived, they are applied to two control samples containing one fake lepton and one real lepton and two jets, $lf_{l}jj$, and another with two fake leptons and two jets, $f_{l}f_{l}jj$. Here l represents a lepton satisfying all selection criterion

from sections 5.2.2 and 5.2.3 for muons and electrons respectively, and f_l the reduced selection for fake leptons. The fake leptons are chosen based on the same loosened selection requirements used to determine the fake rate. All other selection requirements are applied as described in chapter 5 with the addition in the muon case that the selected jets may not be the jet chosen for the fake muon. For the single fake sample, lf_ljj, the fake rate is applied as a weight on the event depending on the p_{T} and η of the fake lepton in the detector. For the double fake sample, f_lf_ljj, the fake rate is applied twice, once for each fake lepton dependent on its p_{T} and η.

The single fake lepton estimation method is heavily contaminated by events from electroweak processes due to the single real lepton. This is accounted for by applying the QCD control sample selection to the full set of MC samples used for the background estimation as well as a W+jets sample generated with MadGraph. After background subtraction, the single fake sample is normalized by 0.5 to account for the combinatorial factor due to either lepton possibly being the fake. The double fake lepton sample on the other hand should be dominated by QCD fakes, but will have very low event yield due to the high kinematic cuts applied. Distributions of the QCD background with electroweak contributions added are shown in figure 6.10 with the $M_{ll} > 200\,\mathrm{GeV}$ cut removed. The final background subtracted event yields are shown in table 6.2 for different stages of selection along with an additional requirement that $M_{lljj} > 600\,\mathrm{GeV}$ where each cut is applied in succession. The single and double lepton fake estimates are consistent to within the statistical uncertainty of the estimate. Furthermore, after the $M_{ll} > 200\,\mathrm{GeV}$ cut is applied the QCD background is consistent with zero in both channels and therefore is ignored in this analysis.

		One fake lepton	Two fake leptons
$eejj$	$p_{\mathrm{T}}(e_1) > 60$ GeV	$(172-155)/2 = 8.5$	$4.6 - 0.2 = 4.4$
	$M_{ee} > 200$ GeV	$(27-20)/2 = 3.5$	1.8
	$M_{eejj} > 600$ GeV	$(19-14)/2 = 2.5$	1.7
$\mu\mu jj$	$p_{\mathrm{T}}(\mu_1) > 60$ GeV	$(157-155)/2 = 1$	$3.1 - 0.4 = 2.7$
	$M_{\mu\mu} > 200$ GeV	$(36-33)/2 = 1.5$	$0.4 - 0.4 = 0.0$
	$M_{\mu\mu jj} > 600$ GeV	$(19-25)/2 < 0$	$0.2 - 0.4 < 0$

Table 6.2: Estimated level of QCD multijet background to the M_{lljj} distribution. The number of multijet events is taken as the difference of the data and electroweak MC distributions.

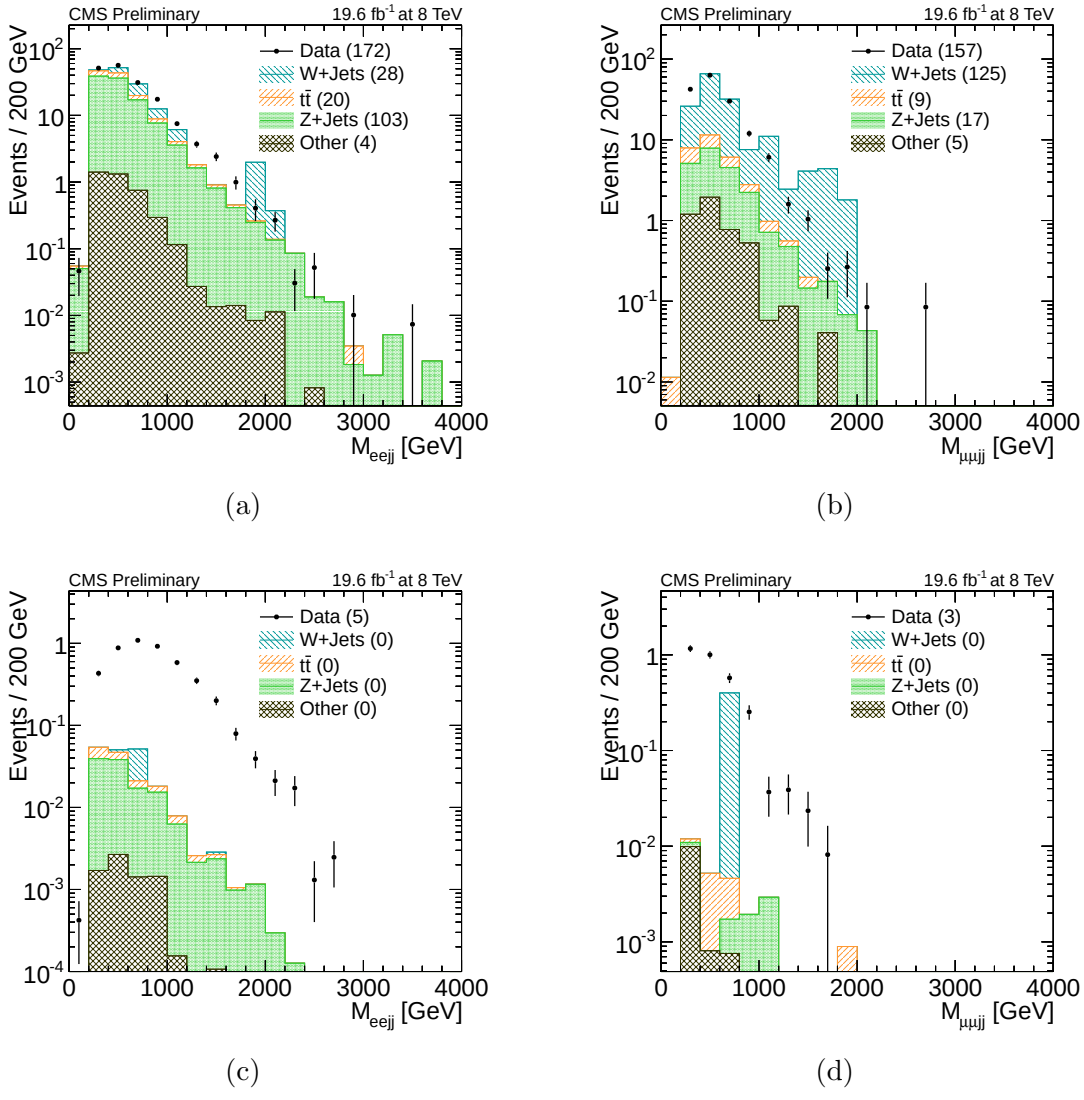


Figure 6.10: Distribution of QCD estimate for electron (left) and moun (right) for one fake (top) and two (bottom) leptons.

	Data	Signal	Total BG	$t\bar{t}$	DY+jets	Other
2 electrons, 2 jets	34506	30	34154	4725	28273	1156
$M_{ee} > 200\,\text{GeV}$	1717	29	1747	1164	475	108
$M_{eejj} > 600\,\text{GeV}$	817	$29 \pm 1 \pm 3$	783 ± 51	476 ± 42	252 ± 24	55 ± 12

	Data	Signal	Total BG	$t\bar{t}$	DY+jets	Other
2 muons, 2 jets	42090	35	41204	5625	34220	1359
$M_{\mu\mu} > 200\,\text{GeV}$	2042	35	2064	1382	549	133
$M_{\mu\mu jj} > 600\,\text{GeV}$	951	$35 \pm 1 \pm 4$	913 ± 58	562 ± 50	287 ± 26	64 ± 12

Table 6.3: The total numbers of events reconstructed in data, and the expected contributions from signal and background (BG) samples, after successive stages of the selection requirements are applied. The "Signal" column indicates the expected contribution for $M_{W_R} = 2.5\,\text{TeV}$.

6.5 Final Object Distributions

The total background is estimated from the sum of the $t\bar{t}$, DY+jets, and "other" background categories. A table showing the number of events in data and each background category for different levels of selection is shown in table 6.3. The three cut stages shown are applied sequentially. The M_{lljj} distributions are shown in figures 6.11 and 6.12 for the electron and muon channels respectively. The bottom portion of the plot shows the ratio of the data to the SM background prediction. A sample signal distribution is also shown for reference. The data and background prediction are in general agreement in the electron and muon channel with the possible exception near $M_{eejj} = 2\,\text{TeV}$ in the electron distribution. Here a noticeable difference in the data and SM prediction is observed. The significance of this difference will be address in detail in section 8.3. The reduced χ^2 calculated for the electron and muon channels are 1.4 and 0.9 respectively.

Additional kinematic distributions are shown for electron channel data in figure 6.13 and muon channel data in figure 6.14. In each figure the top plots show the invariant mass of each combination of ljj invariant mass. A resonant excess in these distributions would give the mass of a possible heavy neutrino. The bottom two plots in both figures shows the p_{T} of the leading lepton and the di-lepton invariant mass. Each distribution would show a high p_{T}/mass excess in the presence of a W_R signal.

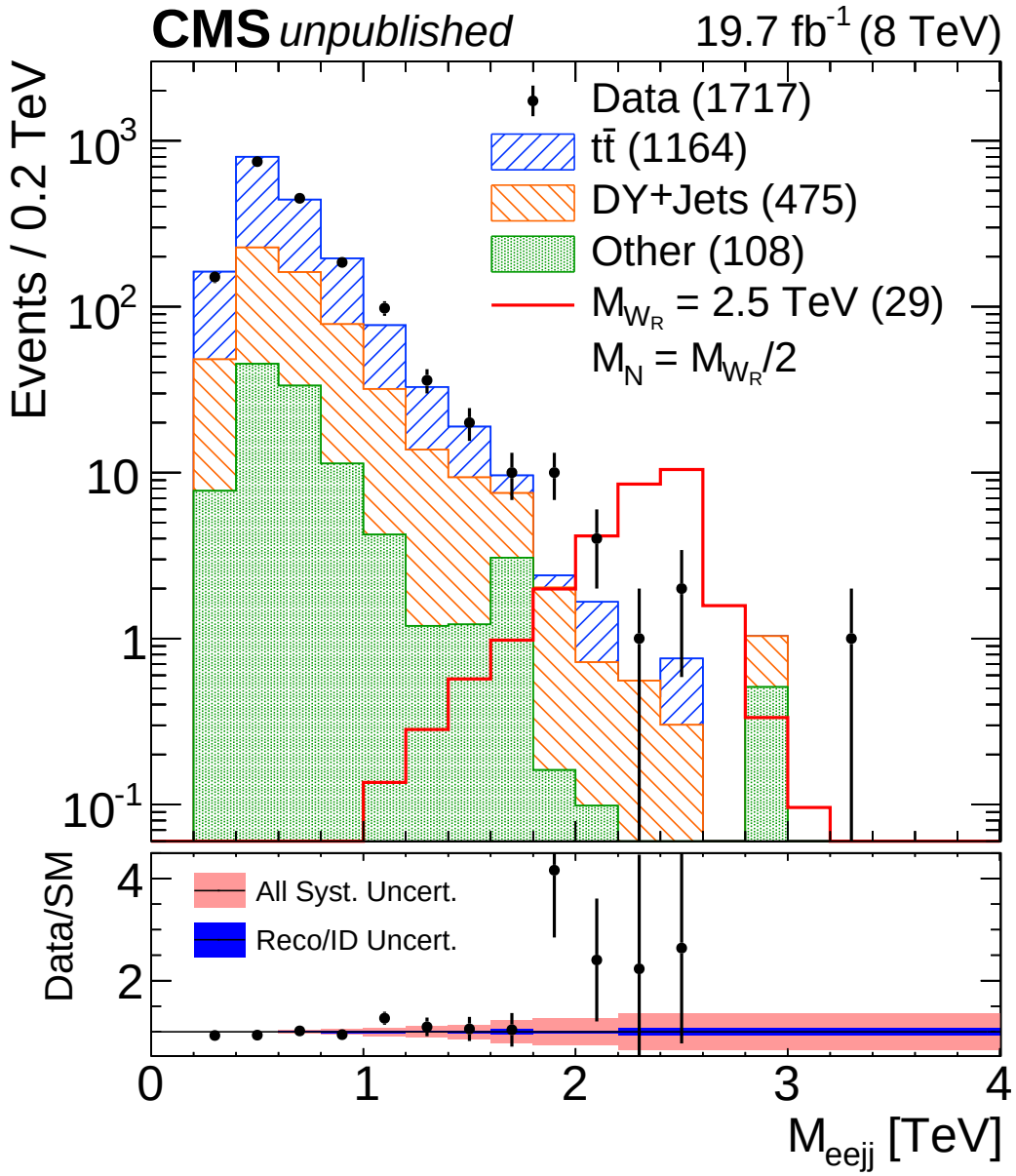


Figure 6.11: Final M_{lljj} distribution for electron channel.

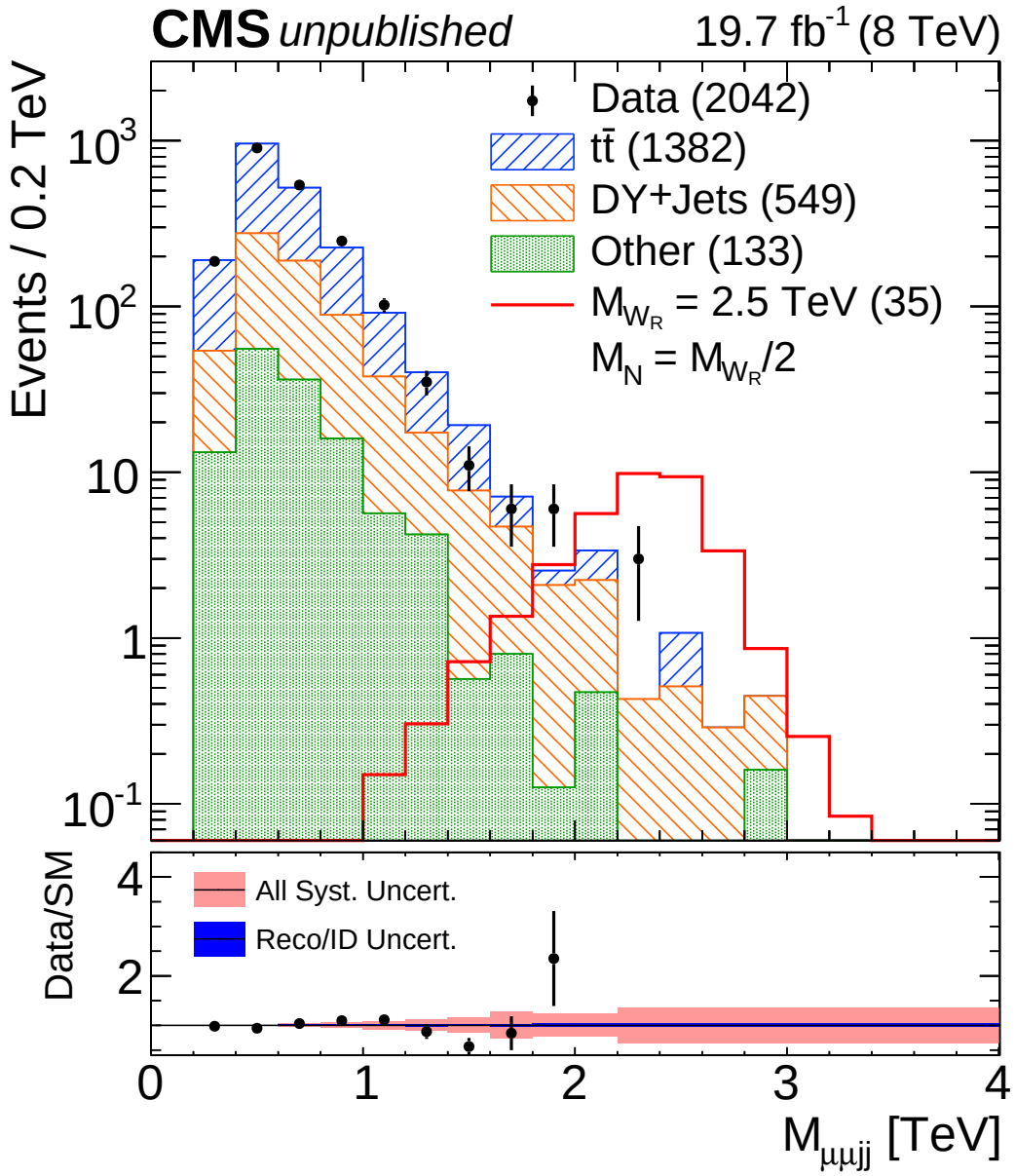
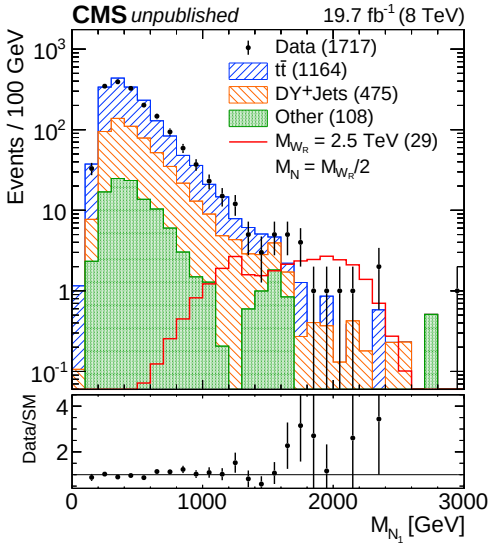
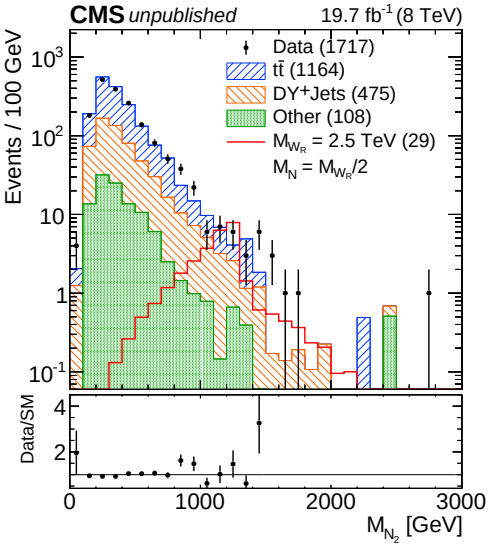


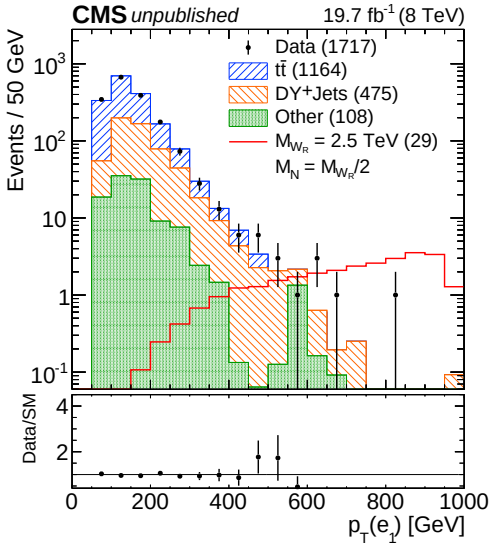
Figure 6.12: Final M_{lljj} distribution for muon channel.



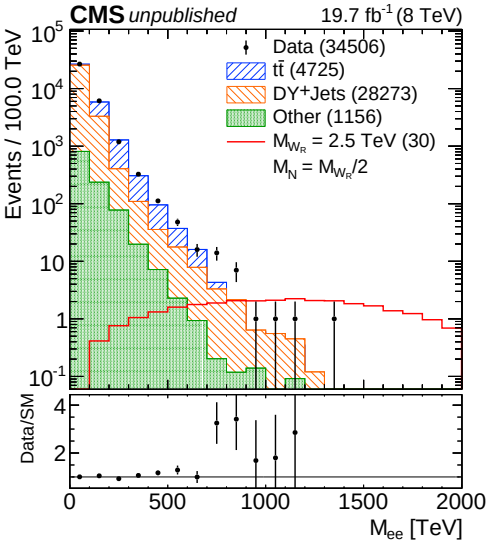
(a) $M_{ejj}(e_1)$



(b) $M_{ejj}(e_2)$

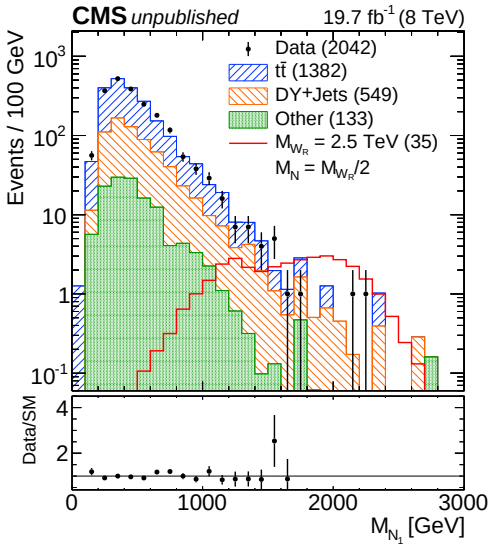


(c) $p_{\mathrm{T}}(e_1)$

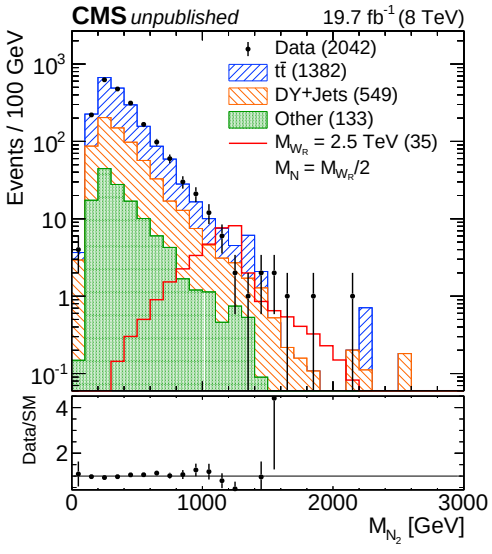


(d) M_{ee}

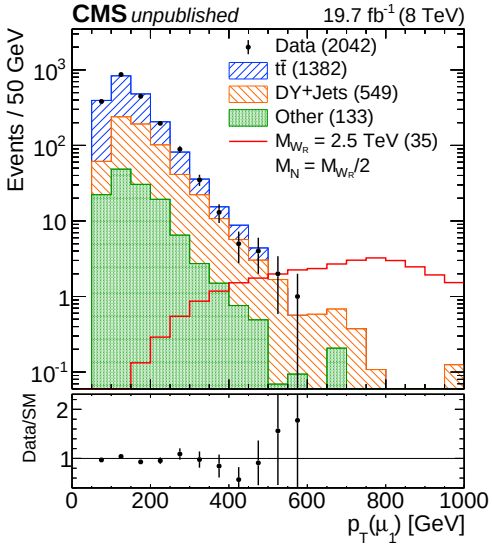
Figure 6.13: Plots of interesting kinematic quantities for electron channel.



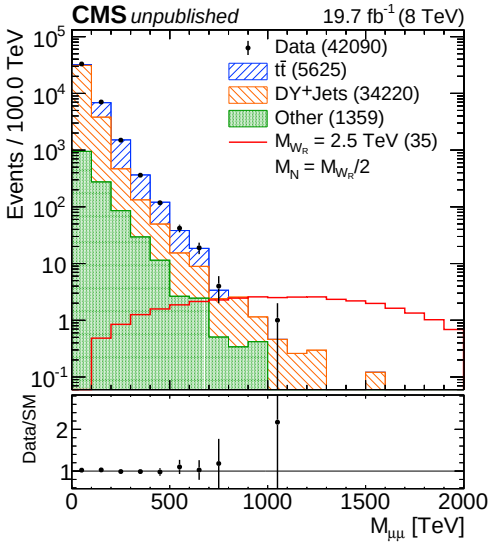
(a) $M_{\mu jj}(\mu_1)$



(b) $M_{\mu jj}(\mu_2)$



(c) $p_{\mathrm{T}}(\mu_1)$



(d) $M_{\mu\mu}$

Figure 6.14: Plots of interesting kinematic quantities for muon channel.

Chapter 7

Systematic Uncertainties

This chapter details the relevant systematic uncertainties considered in this analysis. The dominant systematic uncertainties come from the limited statistics of the background determination and uncertainties in the background modeling. Systematic uncertainties from individual object uncertainties, pileup, signal statistics, and theoretical modeling uncertainties are considered as well. All systematic uncertainties, except a few exceptions noted below, are included as nuisance parameters in the limit calculation. The gamma probability distribution function is used to estimate the uncertainties in background statistics and the log-normal distribution function is used for all other systematic uncertainties.

Systematic uncertainties are calculated individually in each bin used by the limit setting code. Below $M_{lljj} < 1800\,\text{GeV}$, 200 GeV bins are used. Above this point background statistics become limited and a bin from 1800 to 2200 GeV and a single bin above 2200 GeV are used. If an individual bin has insufficient statistics to make a significant determination of a particular uncertainty, it is combined with nearby bins until a statistically significant determination can be made. Unless otherwise noted, each uncertainty is calculated separately for the electron and muon channels.

7.1 Dominant Systematic Uncertainties

7.1.1 Background Statistical Uncertainties

The largest source of systematic uncertainty comes from the limited background statistics. The background statistics are a significant source of uncertainty in the $lljj$ final state due to the requirement of two real high-p_{T} jets in the final state. Generation of MC with additional jets is difficult and often requires dedicated samples such as the exclusive jet multiplicity DY+jets samples to enrich the statistics of the interesting parts of the sample.

The statistical uncertainties for background estimates are not simple Poisson errors because the backgrounds are summations of weighted events. To account for the weight of each event, suppose that N statistically independent measurements of a Poisson distributed quantity X are made. Each measurement of X will yield a result x_i. Now imagine that these N quantities are distributed in some way along a continuous variable W so that the interval between one measurement and the next is represented by w_i. Traditionally, the problem would be phrased such that W represented time between events and therefore w_i would be the time between each individual Poisson event. In the case relevant here, however, W is interpreted as the weight associated with each individual event.

Now construct the following variable

$$n = \sum_i w_i x_i. \tag{7.1}$$

In the case where W represents time, then n represents the total time needed to make all N measurements. If, as is more relevant here, W represents the weight of each event, then n represents the total number of weighted events. The variance of n can be found, knowing that X is Poisson distributed, as

$$\sigma_n^2 = \sum_i w_i^2 x_i, \tag{7.2}$$

where $\sigma_x = \sqrt{x}$ was used for the variance of each x_i and the error on each w_i is neglected. Although this equation is correct, it can be made simpler with no loss of generality if

each weight, w_i, is chosen such that all $x_i = 1$. This makes equations 7.1 and 7.2

$$n = \sum_i w_i, \quad \sigma_n^2 = \sum_i w_i^2. \tag{7.3}$$

Following the argument found here [99], if m is the weighted number of events for which M Poisson events arrive, then the cumulative probability density is

$$P(m \leq n) = 1 - P(m > n) = 1 - \sum_{k=0}^{M-1} \frac{(n/\bar{w})^k e^{-n/\bar{w}}}{k!}. \tag{7.4}$$

Here $\bar{w}$ is the average weight for a dataset defined such that $n = \bar{w}M$.

The probability distribution function of n is found by taking the derivative of 7.4. This yields

$$f(n) = \frac{dP(n \leq m)}{dn} = e^{-n/\bar{w}} \left[\frac{1}{\bar{w}} \sum_{k=0}^{M-1} \frac{(n/\bar{w})^k}{k!} - \sum_{k=0}^{M-1} \frac{k(n/\bar{w})^{k-1}}{k!} \right]. \tag{7.5}$$

It can be shown that 7.5 simplifies to

$$f(n) = \frac{n^{M-1}}{\bar{w}^M \Gamma(M)} e^{-n/\bar{w}}. \tag{7.6}$$

In this form it is clear that $f(n)$ is simply the gamma distribution with a mean of $\mu = n = \bar{w}M$ and a variance of $\sigma^2 = \bar{w}^2 M = \bar{w}n$. Combining this with 7.3 we find

$$\sigma^2 = \sum_i w_i^2 = \bar{w}^2 M = \bar{w} \sum_i w_i \tag{7.7}$$

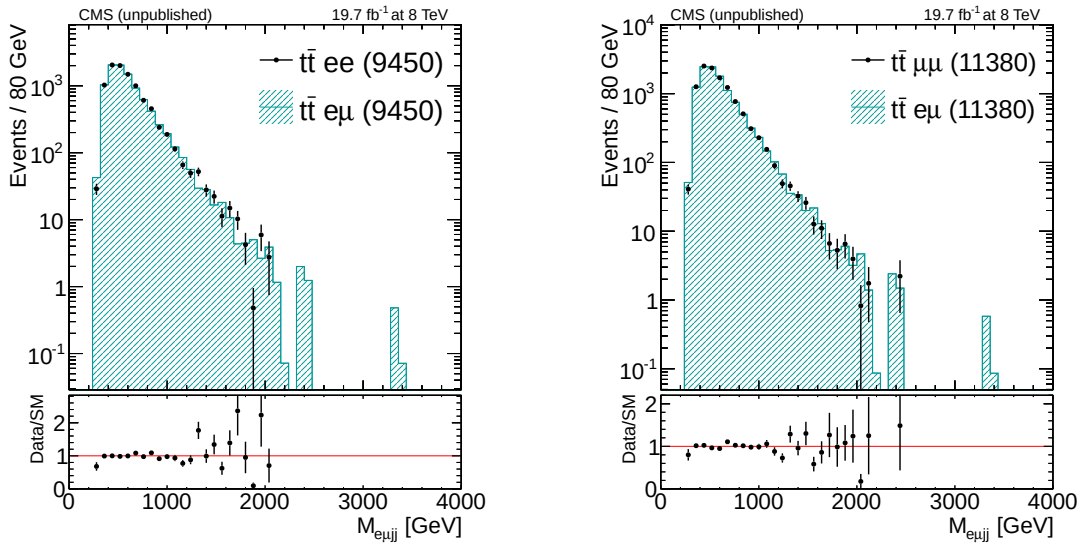
Solving the two equations in 7.7 we find

$$\bar{w} = \frac{\sum_{i=1}^N w_i^2}{\sum_{i=1}^N w_i} \tag{7.8}$$

and

$$M = \frac{\left(\sum_{i=1}^N w_i\right)^2}{\sum_{i=1}^N w_i^2} \tag{7.9}$$

for the average weight $\bar{w}$ and effective number of events M. Note that M and the original number of Poisson events N need not be the same. In the final limit calculation the uncertainties in the backgrounds are estimated on a bin-by-bin basis using the gamma distribution with mean $\bar{w}^2 M$ and average weight $\bar{w}$ calculated from all events in each bin. The relative uncertainty in each bin is shown in the first line for each background table in 7.1 and 7.2.



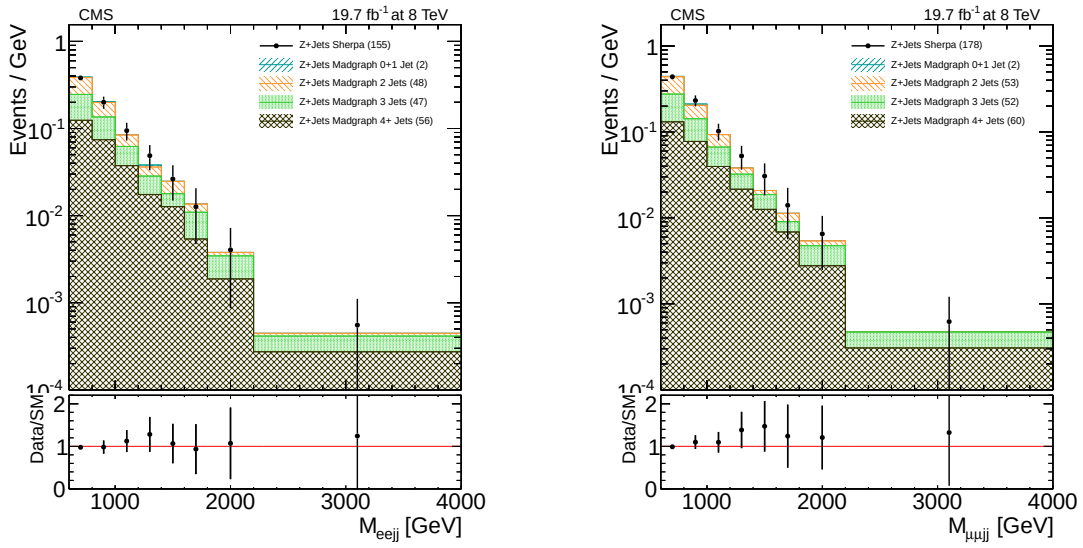
(a) (b)

Figure 7.1: M_{lljj} distribution vs $M_{e\mu jj}$ distribution in $t\bar{t}$ MC for electron (left) and muon (right) channels.

7.1.2 Background Shape Uncertainties

In addition to the statistical uncertainty on each background, an uncertainty is assigned to account for any possible mis-modeling of the backgrounds. For the $t\bar{t}$ background estimation the possibility of differences between the $e\mu jj$ and either $eejj$ and $\mu\mu jj$ channels is studied using MC. After normalization, the electron and muon channels from $t\bar{t}$ MC are compared to the $e\mu jj$ channel MC as shown in figure 7.1. In both channels the distributions are in good agreement between the same flavor lepton distributions and the $e\mu jj$ distribution. Due to this agreement no uncertainty is deemed necessary. An uncertainty is applied to the 1.0 to 1.2 TeV bin based on the difference between the fit point used and the uncorrected distribution.

The uncertainty on the DY+jets background estimation is determined using an independent MC sample generated with SHERPA. The comparison to this sample is used to bound any modeling uncertainties including scale and ISR/FSR uncertainties in the DY+jets MC sample. This sample is also generated with an enriched two, three, and



(a) (b)

Figure 7.2: M_{lljj} distribution prediction from SHERPA and MADGRAPH for electron channel (left) and muon channel (right).

four jet content and provides similar statistics to the exclusive jet multiplicity MADGRAPH samples for the final state containing two leptons and two jets. The distributions of M_{lljj}, with all selections applied, are compared in figure 7.2. In both cases the distributions are statistically consistent, but there are consistent systematic discrepancies between the two samples on the order of 10 to 20% for a few bins. As these differences are on the order of the statistical uncertainty of the samples the electron and muon channels are combined to determine the final systematic uncertainty shown in figure 7.3. Below $M_{lljj} < 1000\,\text{GeV}$ the larger of the statistical uncertainty or the difference is chosen as the systematic uncertainty and above 1000 GeV a weighted average is taken as the uncertainty in each bin.

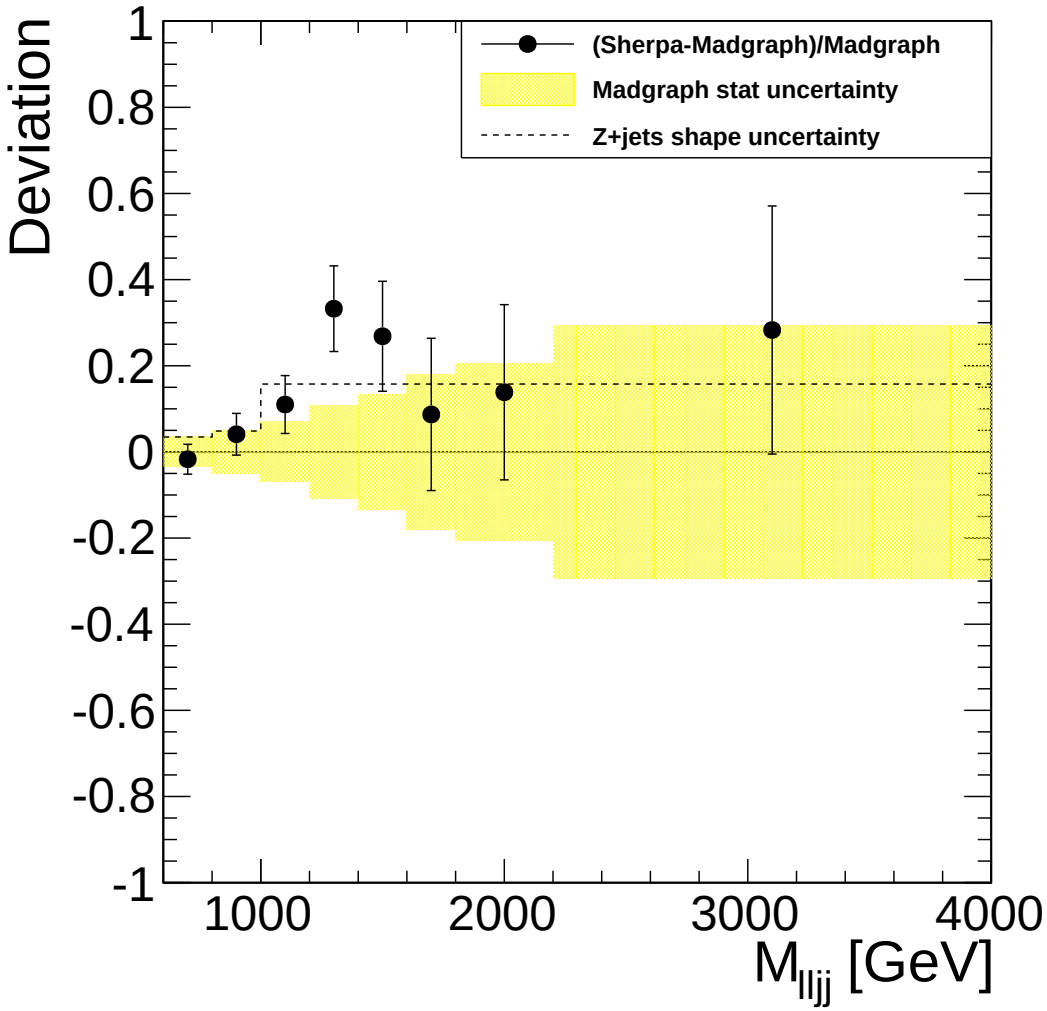


Figure 7.3: Shape systematic for DY+jets background.

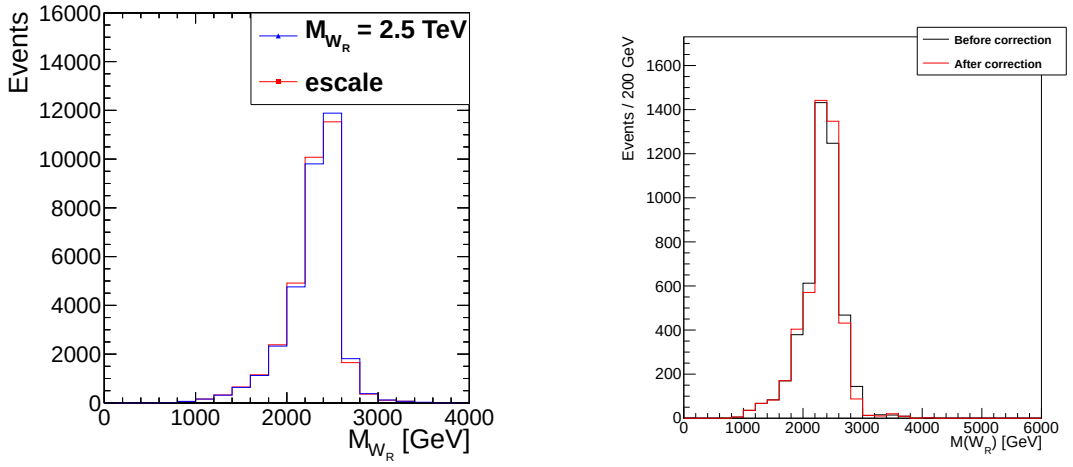
7.2 Sub-dominant Systematic Uncertainties

7.2.1 Object Uncertainties

The largest single object uncertainty is the uncertainty on the jet energy corrections and jet energy resolution. The jet energy correction and resolution uncertainties are determined by the jetMET working group in CMS. Their recipe outlined in [100, 88] is used in this analysis. In the case of the jet energy correction uncertainties, the energy of each jet in MC is adjusted up by an uncertainty factor dependent on the p_{T} and η of the jet. The full set of event selection criteria are then reapplied using the new adjusted jets in each MC sample. The process is repeated for the downward fluctuations in jet energy as well. The relative uncertainty is then parameterized in each bin of the M_{lljj} distribution as the ratio of events passing the final selections with the jet energy corrections perturbed versus the nominal jets for both the upward and downward fluctuations. Before the DY+jets MC uncertainty is calculated, the varied sample is normalized to the unperturbed sample in the region $60 < M_{ll} < 120\,\mathrm{GeV}$ to account for the effects of the DY+jets normalization to data.

The jet energy resolution uncertainty is calculated in a similar way to the jet energy correction uncertainty. The full selection is rerun with the jet energy scaled up and down by factors based on the jet η, which represent the one σ up an downward fluctuations of the jet p_{T} resolution. The relative uncertainty is then derived in each bin in the M_{lljj} distribution.

The electron energy scale and resolution uncertainties are also calculated in a similar fashion to those for the jets. The electron scale factors for each uncertainty are calculated by the $Z' \to ee$ analysis team using data from the Z peak to be 0.6% and 1 to 2% for energy scale and resolution respectively [52]. These scale factors are applied to the electrons, an upward fluctuation for the energy scale and up and downward fluctuations for the resolution, and the bin-by-bin uncertainties are derived for each MC sample. Although these uncertainties are small, they can have a large effect on the signal MC because at high M_{ll} the ~1% shift in energy can push many events from a high population bin to a small population bin. This can lead to several-percent uncertainties for high M_{W_R} signal points. An example of the M_{lljj} distribution with and without the energy scale uncertainty applied is shown to the left in figure 7.4.

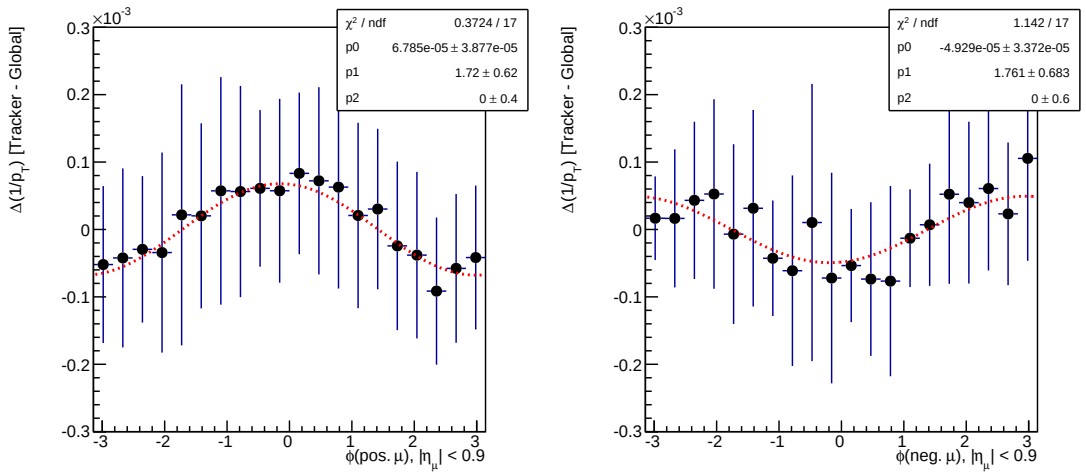


(a) (b)

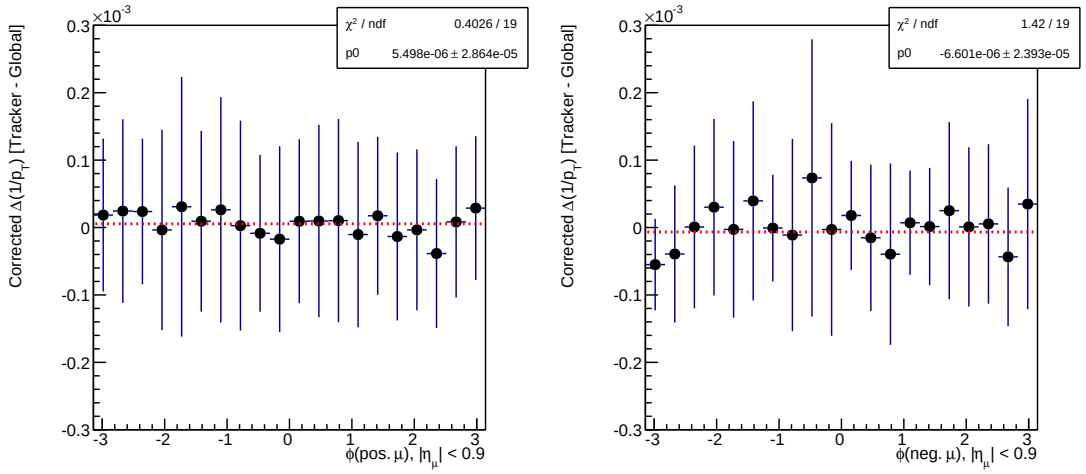
Figure 7.4: $M_{W_R} = 2.5\,\text{TeV}$ signal MC with electron energy scale correction applied (left). $M_{W_R} = 2.5\,\text{TeV}$ signal MC with muon energy resolution correction applied (right).

The muon resolution uncertainty is dominated by an overly large misalignment of the muon chambers in MC. This effect causes poorer performance for muons with $p_{\text{T}} > 200\,\text{GeV}$ in MC. This effect can be seen in the difference in $1/p_{\text{T}}$ of the tracker and tracker+muon system measurements and the effect is shown in the top portion of figure 7.5. This effect can be corrected for on a statistical basis for muons with $p_{\text{T}} > 200\,\text{GeV}$ as shown in the bottom of figure 7.5. The signal shape for a W_R with $M_{W_R} = 2.5\,\text{TeV}$ is shown on the right in figure 7.4 before and after this correction is applied. For background MC and low M_{W_R} signal MC the vast majority of muons have $p_{\text{T}} < 200\,\text{GeV}$ and therefore the effect of this correction is minimal. For signal MC with $M_{W_R} > 2\,\text{TeV}$ this effect can be on the order of ~3%, but most events will fall into the last and largest M_{lljj} bins which will eliminate the effects of bin migration. Therefore, the correction is not applied to the nominal analysis, but is instead applied as a systematic uncertainty.

The final category of object specific uncertainties is based on the lepton identification and reconstruction efficiencies. These uncertainties are calculated by adjusting the event weight for leptons by ±0.5% for muons and ±2%(4%) for electrons in the barrel(endcap). These fluctuations correspond to the uncertainty on the muon and electron



(a)



(b)

Figure 7.5: Difference in $1/p_{\mathrm{T}}$ for tracker only and tracker plus muon system for signal MC with $M_{W_R} = 2.0\,\mathrm{TeV}$ before correction (top) and after correction (bottom).

ID efficiencies. The full selections are then reapplied and the uncertainties extracted from the M_{lljj} distribution as for other uncertainties.

7.2.2 Event and Sample Based Uncertainties

Uncertainties in the muon trigger efficiency are calculated by the muon physics object group in CMS. These uncertainties are less than one percent. The minimum p_{T} threshold for the trigger used by this analysis is recommended to be 45 GeV because the efficiency curve has not plateaued by 40 GeV. In order to maintain consistency between the electron and muon channels, the p_{T} threshold for muons in this analysis was placed at 40 GeV. This choice is justified because the leading muon is required to have $p_{\mathrm{T}} > 60$ GeV. Less than 5% of events contain a muon below 45 GeV and of those the leading p_{T} muon will trigger the event 87% of the time. This means this choice will affect less than 1% of all muon channel events. To account for any added uncertainty this may cause we increase the uncertainty on events where muons with $p_{\mathrm{T}} < 45$ GeV trigger to 10% to account for the maximum possible efficiency loss in this range due to the shoulder of the turn-on curve. The uncertainty is derived as a function of M_{lljj}.

The electron trigger efficiency is nearly 100% and any inefficiency is found to be independent of M_{ll}. Due to this, any effect from the electron trigger inefficiency will be removed by the DY+jets sample normalization to data. The electron trigger efficiency uncertainty is therefore combined with the DY+jets normalization uncertainty.

The uncertainty in the normalization of each background and signal is taken into account separately depending on how each sample is normalized. The normalization factors on the $t\bar{t}$ background estimate are found to be consistent within their statistical uncertainties so the uncertainty on the normalization is taken to be statistical uncertainty on the more precise approach based on data, the efficiency based method.

The uncertainty on the DY+jets estimation is taken as the difference between the electron and muon channel scale factors because the factors are not consistent within their statistical uncertainties. This conservative choice is made because this uncertainty also covers any possible electron trigger uncertainty issues.

The remaining MC samples, as well as the signal samples, are normalized directly to their NLO cross-sections. The normalization uncertainties are therefore taken as the

cross-section uncertainty and the total integrated luminosity uncertainty. The cross-section uncertainties are calculated with each NLO cross-section. The integrated luminosity uncertainty is based on the luminosity calculation performed with the silicon pixel tracker and is found to be 2.6%. An additional uncertainty is applied to the signal MC based on the statistics of each sample. This uncertainty is taken as the relative statistical uncertainty on the whole sample after the full set of selections has been applied. This is less than 1% for all signal samples.

The final event based uncertainty is designed to account for pileup effects. This uncertainty is calculated for each MC sample by varying the mean of the number of pileup interactions in the sample up and down by 5%. This is accomplished by reapplying the pileup event weights based upon a new pileup interaction distribution with the mean shift incorporated. The relative uncertainties are then calculated as a function of M_{lljj} based on the ratio of the perturbed samples to the nominal sample.

7.2.3 Theoretical Uncertainties

The theoretical uncertainties come from parton distribution function (PDF) uncertainties in MC production. The PDFs, used to describe the effective content of the protons during a hard scattering, are calculated by several groups of theorists and experimentalists working in tandem to collect the data from a wide array of measurements. These include precision measurements of proton-proton collisions at as many center of mass energies as possible, as well as dedicated measurements of electron-proton collisions. The primary PDF set used for this analysis is CTEQ6L1 [82].

The signal MC can have very large PDF uncertainties due to the very high Q^2 involved in creating multi- TeV objects. The PDF uncertainties for signal are calculated using the prescription outlined in [101]. Dedicated MC samples are generated, but not run through detector simulation or reconstruction. A simplified analysis is run on the full grid of signal points with each PDF set run through its full set of possible variations. The variations are ordered by positive or negative effects and the contributions are added in quadrature. The PDF uncertainties for signal are factorized into two parts. The first part is the effect on the the event acceptance which is evaluated using simplified selection criteria designed to run on the generator information only. The acceptance uncertainty is a small effect. The second part includes effects on the overall production

cross-section and for $M_{W_R} > 1\,\text{TeV}$ is quite large (~30% at $M_{W_R} = 3\,\text{TeV}$). These effects are separated because the acceptance effects are included in the limit calculations while the cross-section effects are included on the expectation of the theoretical model. For background MC samples, the PDF uncertainties are expected to be small and therefore we rely on uncertainties found using this technique on MC generated at 7 TeV to estimate the background PDF uncertainties.

7.3 Systematic Uncertainty Tables

The systematic uncertainties described in this chapter are shown in tables 7.1 and 7.2 for the electron and muon channels respectively. The uncertainties are shown for each background and for three sample signal points. The dominant uncertainties are shown separately while the sub-dominant uncertainties are combined into a single category by summing the individual uncertainties in quadrature. These uncertainties are only meant to give a sense of scale on the size of the uncertainties. The treatment of the uncertainties in the final limit setting is more complicated and will be covered in detail in section 8.1.

	M_{eejj} mass range (TeV)								
$t\bar{t}$ Background	0.6-0.8	0.8-1.0	1.0-1.2	1.2-1.4	1.4-1.6	1.6-1.8	1.8-2.2	> 2.2	Total
Events	276.2	114.9	45.0	19.3	9.6	2.1	1.4	0.5	469
$e\mu jj$ Stats.	4.7	7.2	12.9	17.5	23.9	65.9	63.4	109.2	7.6
BG shape	0.0	0.0	22.0	0.0	0.0	0.0	0.0	0.0	2.1
Other Syst.	1.9	1.9	1.9	1.9	1.9	1.9	1.9	1.9	1.9
Total	5.0	7.5	25.6	17.6	24.0	66.0	63.4	109.2	**8**

	M_{eejj} mass range (TeV)								
Z+jets BG	0.6-0.8	0.8-1.0	1.0-1.2	1.2-1.4	1.4-1.6	1.6-1.8	1.8-2.2	> 2.2	Total
Events	128.5	67.3	27.8	12.6	8.2	4.5	2.5	1.3	253
MC Statis.	3.6	5.0	7.2	11.2	13.3	18.0	22.5	27.7	5.6
BG Shape	3.5	4.8	15.8	15.8	15.8	15.8	15.8	15.8	6.6
Other Syst.	3.6	4.5	4.4	4.4	4.7	4.4	4.0	9.5	4.1
Total	6.2	8.3	17.8	19.8	21.1	24.3	27.7	33.3	**10**

	M_{eejj} mass range (TeV)								
Other BG	0.6-0.8	0.8-1.0	1.0-1.2	1.2-1.4	1.4-1.6	1.6-1.8	1.8-2.2	> 2.2	Total
Events	32.6	10.5	3.7	1.1	1.0	3.1	0.2	0.5	53
MC Stats.	10.2	16.7	25.9	50.5	56.4	43.5	53.4	100.0	17.2
Other Syst.	10.8	12.5	16.0	10.0	10.3	9.1	9.1	9.1	11.4
Total	14.9	20.8	30.4	51.4	57.3	44.4	54.1	100.4	**21**

	M_{eejj} mass range (TeV)								
$M_{W_R} = 1\,\text{TeV}$	0.6-0.8	0.8-1.0	1.0-1.2	1.2-1.4	1.4-1.6	1.6-1.8	1.8-2.2	> 2.2	Total
Efficiency (%)	3.2	22.6	12.6	1.7	0.6	0.2	0.1	0.0	41.0
PDF Cross-sec.	5.5								5.5
Other syst.	7.5	7.5	12.3	11.3	12.1	11.4	11.2	13.6	**9**

	M_{eejj} mass range (TeV)								
$M_{W_R} = 2\,\text{TeV}$	0.6-0.8	0.8-1.0	1.0-1.2	1.2-1.4	1.4-1.6	1.6-1.8	1.8-2.2	> 2.2	Total
Efficiency (%)	0.1	0.2	0.6	1.6	4.3	10.6	35.5	2.2	55.1
PDF Cross-sec.	11.4								11.4
Other syst.	14.1	7.6	5.5	5.6	6.4	6.1	5.6	8.3	**6**

	M_{eejj} mass range (TeV)								
$M_{W_R} = 3\,\text{TeV}$	0.6-0.8	0.8-1.0	1.0-1.2	1.2-1.4	1.4-1.6	1.6-1.8	1.8-2.2	> 2.2	Total
Efficiency (%)	0.0	0.0	0.1	0.3	0.4	1.0	3.8	52.4	58.1
PDF Cross-sec.	26.2								26.2
Other syst.	12.2	18.8	9.9	6.1	5.8	5.9	6.0	5.4	**5**

Table 7.1: Summary of background and signal event populations, efficiencies, and systematic uncertainties (in %) for the electron channel, summarized for the M_{eejj} bins used for limit inputs. The dominant background uncertainties (statistics and shape) are presented separately from the remaining systematic uncertainties. The weighted average uncertainties are presented in the far right column, and the sum-in-quadrature uncertainty is presented in the bottom row.

	$M_{\mu\mu jj}$ mass range (TeV)								
$t\bar{t}$ Background	0.6-0.8	0.8-1.0	1.0-1.2	1.2-1.4	1.4-1.6	1.6-1.8	1.8-2.2	> 2.2	Total
Events	332.6	138.4	54.2	23.2	11.6	2.5	1.7	0.6	565
$e\mu jj$ Stat.	4.7	7.2	12.9	17.5	23.9	65.9	63.6	109.2	7.6
BG shape	0.0	0.0	22.0	0.0	0.0	0.0	0.0	0.0	2.1
Other Syst.	1.9	1.9	1.9	1.9	1.9	1.9	1.9	1.9	1.9
Total	5.0	7.5	25.6	17.6	24.0	66.0	63.7	109.2	**8**

	$M_{\mu\mu jj}$ mass range (TeV)								
Z+jets BG	0.6-0.8	0.8-1.0	1.0-1.2	1.2-1.4	1.4-1.6	1.6-1.8	1.8-2.2	> 2.2	Total
Events	152.8	72.9	32.1	13.1	7.2	3.9	3.7	1.5	287
MC Stats.	3.3	4.8	6.8	10.3	13.5	18.1	18.6	31.2	5.2
BG Shape	3.5	4.8	15.8	15.8	15.8	15.8	15.8	15.8	6.4
Other Syst.	3.1	3.1	4.0	4.0	4.0	4.1	4.1	4.1	3.3
Total	5.7	7.5	17.6	19.2	21.2	24.3	24.7	35.2	**9**

	$M_{\mu\mu jj}$ mass range (TeV)								
Other BG	0.6-0.8	0.8-1.0	1.0-1.2	1.2-1.4	1.4-1.6	1.6-1.8	1.8-2.2	> 2.2	Total
Events	35.5	15.3	5.0	4.1	0.4	0.7	0.5	0.2	62
MC Stats.	9.8	14.7	27.1	29.3	39.7	72.3	89.2	100.0	15.5
Other Syst.	5.9	5.9	13.4	13.1	9.1	7.2	7.2	7.2	7.0
Total	11.4	15.8	30.3	32.1	40.8	72.7	89.5	100.3	**17**

	$M_{\mu\mu jj}$ mass range (TeV)								
$M_{W_R} = 1$ TeV	0.6-0.8	0.8-1.0	1.0-1.2	1.2-1.4	1.4-1.6	1.6-1.8	1.8-2.2	> 2.2	Total
Efficiency (%)	4.4	26.8	14.0	2.0	0.6	0.3	0.2	0.0	48.3
PDF Cross-sec.					5.5				5.5
Other Syst.	4.8	4.9	9.3	13.6	9.3	9.3	9.3	9.3	**7**

	$M_{\mu\mu jj}$ mass range (TeV)								
$M_{W_R} = 2$ TeV	0.6-0.8	0.8-1.0	1.0-1.2	1.2-1.4	1.4-1.6	1.6-1.8	1.8-2.2	> 2.2	Total
Efficiency (%)	0.1	0.3	0.8	2.3	5.6	13.8	38.9	3.7	65.5
PDF Cross-sec.					11.4				11.4
Other Syst.	4.2	4.2	7.4	12.2	3.5	4.6	3.6	11.5	**5**

	$M_{\mu\mu jj}$ mass range (TeV)								
$M_{W_R} = 3$ TeV	0.6-0.8	0.8-1.0	1.0-1.2	1.2-1.4	1.4-1.6	1.6-1.8	1.8-2.2	> 2.2	Total
Efficiency (%)	0.0	0.1	0.1	0.3	0.6	1.2	5.1	61.3	68.6
PDF Cross-sec.					26.2				26.2
Other Syst.	13.6	13.6	3.2	36.9	37.6	3.2	3.5	3.2	**4**

Table 7.2: Summary of background and signal event populations, efficiencies, and systematic uncertainties (in %) for the muon channel, summarized for the $M_{\mu\mu jj}$ bins used for limit inputs. The dominant background uncertainties (statistics and shape) are presented separately from the remaining systematic uncertainties. The weighted average uncertainties are presented in the far right column, and the sum-in-quadrature uncertainty is presented in the bottom row.

Chapter 8

Limit Setting

This chapter deals with the calculation and interpretation of the 95% confidence level (CL) exclusion limits. The limits are calculated based on the modified frequentist CL_S approach [102, 103]. Limits are first calculated based upon the signal efficiency calculated using the fully reconstructed signal MC generated with $M_N = M_{W_R}/2$. The significance of the limit calculations is discussed along with the significance of the small excess in the electron channel around $M_{W_R} = 2\,\text{TeV}$. The extrapolation of the limits to the full $M_N - M_{W_R}$ plane is then discussed. Finally, limits combining the electron and muon channel data, under the assumption that the mass of all three heavy neutrinos are degenerate, are discussed.

The limits are computed using a multi-bin technique as a function of M_{lljj}. The exact binning choice is discussed in the introduction to chapter 7. This approach has several benefits over a traditional single-bin "cut and count" analysis. The first is additional background rejection which comes from treating M_{lljj} in multiple bins. This is because bins with no expected signal do not contribute to the limits. Further advantages arise from considering correlations between bins for the signal shape and systematic uncertainties.

8.1 The CL_S Technique

8.1.1 Conceptual Overview

The CMS recommendation for setting exclusion limits is to use the modified frequentist CL_S technique [102]. Frequentist statistics are based on the philosophy that the future outcome of an experiment can be predicted based solely on the previous outcomes of the same experiment. This is in contrast to the Bayesian approach to statistics where a probability of belief in an outcome is assigned based on the data and an a-priori assumption of the outcome. An important feature of frequentist limits is that they make a statement on how probable the observed data is given a model, and not on how probable the model is.

The basis of limit setting is the concept of hypothesis testing. Say an experiment has two possible hypotheses describing its outcome. The first hypothesis, H_0, often called the null-hypothesis or the background only hypothesis, models the outcome without any new physics. A second hypothesis, H_μ, includes some new effect (e.g. the existence of a W_R). H_μ is often called the signal hypothesis or the alternate hypothesis. Here μ parameterizes the strength of a possible signal. The first step is to construct a variable which holds some power to discern between the two hypotheses, called the test statistic. This might be as simple as the total number of events, the output of a multivariate analysis, or a more complicated construction which can account for systematic uncertainties.

To discern the two hypotheses, H_0 and H_μ, the probability distribution function of the chosen test statistic, $q(X)$ for a measurement X, must be known or estimated for each hypothesis. For the purpose of this thesis $q(X)$ is chosen such that smaller values are more signal like. The most common estimation technique is to use toy MC to create a large number of pseudo-experiments from which the distributions of the test statistic can be constructed. Once the probability distribution function, $f(\mu)$, has been determined, the probability that the result is not caused by a background fluctuation is

$$\mathrm{CL}_b \equiv \int_{q(X)}^{\infty} dq f(0). \tag{8.1}$$

Here CL_b stands for confidence level of the background only hypothesis (H_0). The probability that the null hypothesis describes a disagreement between data and the

expected result at least as large as the observed difference is $1 - \mathrm{CL}_b$. This quantity is referred to as the p-value. Similarly, the confidence level for H_μ can also be calculated as

$$\mathrm{CL}_{s+b} \equiv \int_{q(X)}^{\infty} dq f(\mu) \tag{8.2}$$

where CL_{s+b} stands for the confidence level of the signal plus background hypothesis. Then based upon the values of CL_b and CL_{s+b}, the alternate hypothesis, H_μ, can possibly be discerned from the null-hypothesis, H_0. Generally the thresholds for claiming the alternative hypothesis may be correct are set before the experiment is performed to remove any bias. In high energy physics it is common to require $1 - \mathrm{CL}_b < 2.87 \times 10^{-7}$ which corresponds to the "5σ" discovery threshold.

In any region where there is no clear evidence for discovery, frequentist exclusion limits are set instead. These limits are calculated based on the confidence levels discussed above. It is common in high energy physics to set 95% confidence level limits. Such limits exclude the region of the parameter of interest (μ) where the probability that the the observed data is described by the alternative hypothesis is less than 5% based on the measurements made. More concretely, this is defined as the region for which $\mathrm{CL}_{s+b} < 1 - 0.95$. Although this is a perfectly correct way to calculate the limits in the frequentist interpretation, this formulation of exclusion limits has some undesirable effects. The most alarming of these is that for small signal, more stringent limits can be placed with larger backgrounds. This is avoided by "normalizing" the signal plus background confidence level by the background only confidence level, CL_b. This leads to the condition for 95% confidence limits of

$$\mathrm{CL}_S \equiv \frac{\mathrm{CL}_{s+b}}{\mathrm{CL}_b} < 1 - 0.95. \tag{8.3}$$

This is called the modified frequentist, or the CL_S, approach to limit setting. It should be noted that CL_S is not actually a confidence level, and due to the normalization by CL_b will tend to give conservative limits compared to the traditional approach using CL_{s+b}.

8.1.2 Limit Implementation

The limit setting calculations for this analysis are implemented using the tool developed by the Higgs analysis group in CMS with the RooStats statistical analysis package

[104]. This tool allows for calculation of expected exclusion limits ignoring the observed data, observed exclusion limits, significance of the observed results, and p-value of the null-hypothesis. It also allows for either CL_{s+b} or CL_S limits to be calculated along with a choice of several possible techniques to calculate the test statistic.

The test statistic used for this analysis is called the profiled likelihood ratio defined as

$$\lambda(\mu) = \frac{L(X|\mu, \hat{\hat{\boldsymbol{\theta}}})}{L(X|\hat{\mu}, \hat{\boldsymbol{\theta}})} \tag{8.4}$$

where $L(X|\mu, \boldsymbol{\theta})$ is the likelihood function for a measurement X, signal scale factor μ, and systematic uncertainties represented by $\boldsymbol{\theta}$. Quantities such as systematic uncertainties which affect the outcome of the limits but are not of direct interest are called nuisance parameters. The variables $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ indicate that the likelihood is maximized with respect to these variables, while $\hat{\hat{\boldsymbol{\theta}}}$ indicates that the likelihood is maximized with respect to $\boldsymbol{\theta}$ for a given μ. On the right-hand side of equation 8.4 the denominator represents the global maximum likelihood for the model given the data X while the numerator represents the largest likelihood for a given signal strength assumption μ. If the model with an assumed signal strength μ matches the data well then the ratio $\lambda(\mu)$ will be close to one, but if it is not a good match it will be close to zero. Although $\lambda(\mu)$ satisfies the requirement for a test statistic, it represents how "signal like" the data is, it is often more convenient to work with log likelihood ratio defined as

$$q(\mu) = -2 \ln \lambda(\mu). \tag{8.5}$$

The advantage of the log likelihood ratio is that it can be calculated as a sum instead of a product and is therefore much easier to compute. Unlike $\lambda(\mu)$, $q(\mu)$ will be close to zero for signal like results while larger values are less signal like.

The systematic uncertainties are accounted for via nuisance parameters in the profile likelihood. The likelihood function can be written as

$$L(X|\mu, \boldsymbol{\theta}) = \prod_i \frac{e^{-(\mu s(\boldsymbol{\theta}) + b(\boldsymbol{\theta}))} (\mu s(\boldsymbol{\theta}) + b(\boldsymbol{\theta}))^{x_i}}{x_i!} \prod_j G_{ij}(\boldsymbol{\theta}) \tag{8.6}$$

where s and b are the number of signal and background events predicted and depend on the systematic uncertainties through the nuisance parameters $\boldsymbol{\theta}$ and x_i are the individual

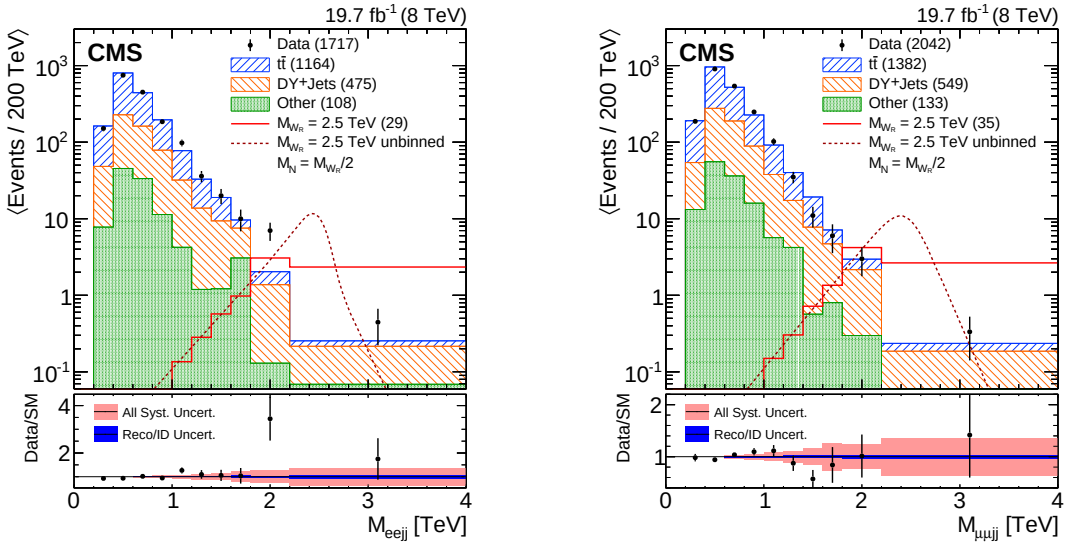
measurement bins of X. The product over i enumerates the bins in M_{lljj} while the product over j enumerates the different sources of systematic uncertainty.

The functions G_{ij} are constraint functions which limit the range over which the nuisance parameters can vary during optimization. As demonstrated in section 7.1.1 the appropriate form of the constraint function for the background statistical uncertainties is the gamma distribution. For the remaining systematic uncertainties, the log-normal distribution is used with a width chosen for each bin and uncertainty. As the likelihood is maximized, the nuisance parameters are moved around their optimal values which will change the signal and background yields. This may in turn raise the first term in 8.6, but the likelihood is penalized by the corresponding constraint functions based on how far from the optimal value each parameter is adjusted. This restricts the fitter from adjusting values of the nuisance parameters significantly farther than the uncertainty on each parameter. The fitter also takes into account appropriate correlations between sources of systematic uncertainty.

8.2 One Dimensional Limits

Using the CL_S technique as described above, limits on the cross-section times branching ratio, $\sigma \times B$, are set in the electron and muon channels under the assumption that only one flavor of heavy neutrino is accessible at LHC energies. Limits are also set using the combined electron and muon channel data under the assumption that the heavy neutrinos are all degenerate in mass and accessible at the LHC. The input data used is selected as described in chapter 5. The background estimations are taken as described in chapter 6. The data and background distributions, binned for limit input, are shown in figure 8.1. In order to simplify the limit calculation, only bins with $M_{lljj} > 600\,\mathrm{GeV}$ are considered. This results in no loss of generality as even the lightest W_R signal considered, with a mass of 1 TeV, has negligible signal contributions with $M_{lljj} < 600\,\mathrm{GeV}$.

The signal inputs to the limit setting tool are taken from the fully simulated and reconstructed signal MC generated with $M_N = M_{W_R}/2$. Each signal point, generated for a particular M_N and M_{W_R}, is required to pass the full selection criteria as described in chapter 5. The efficiency for the signal to pass all selection criteria is then calculated for each limit input bin. This efficiency is then used by the limit setting tool to model



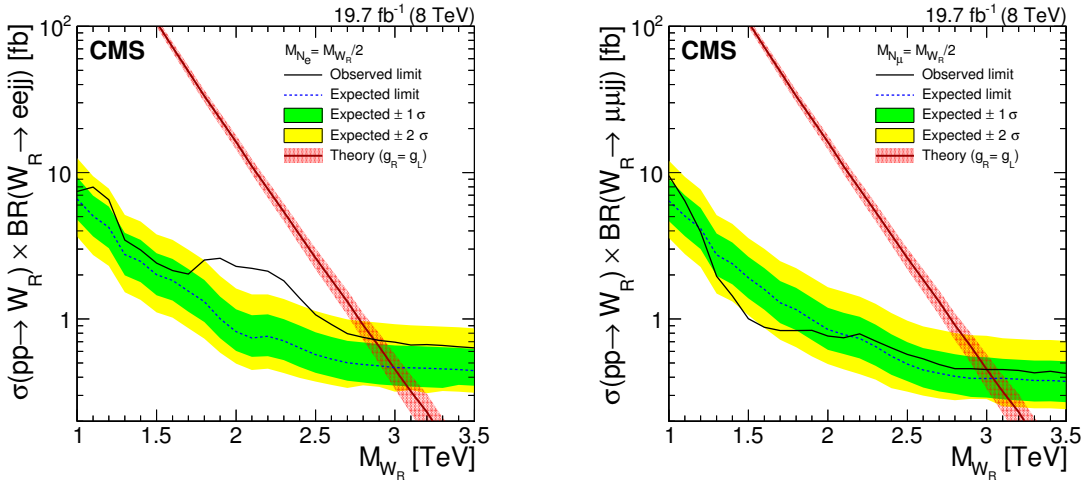
(a) (b)

Figure 8.1: M_{lljj} distribution binned for the limit calculation with systematic uncertainties shown in the lower ratio plot for the electron (left) and muon (right) channels. On the lower portion of the plot the outer light red error band represents all uncertainties while the inner dark blue band represents only the sub-dominant uncertainties.

the signal. A sample signal point with $M_{W_R} = 2.5\,\text{TeV}$ is shown in figure 8.1.

Systematic uncertainties are used as calculated in chapter 7. The background statistics uncertainties are considered completely uncorrelated between bins and between background samples. Other uncertainties are considered correlated between bins, but separate sources of uncertainty are considered uncorrelated.

The limits as calculated for the electron channel and muon channel can be seen in figure 8.2. The region above the solid black line is excluded at the 95% confidence level. The theoretical prediction for $W_R \to lN_l$ production is shown as well. The leading order cross-sections are provided by the MC generation process. The LO cross-section are corrected to NLO cross sections using calculations performed with FEWZ [98]. The expected and observed exclusion limits are in reasonable agreement for the muon channel, but the electron channel shows a significant deviation around $M_{W_R} = 2\,\text{TeV}$. This deviation between the observed and expected limits is caused by the excess of observed events over the background prediction around $M_{eejj} = 2\,\text{TeV}$. The statistical



(a) (b)

Figure 8.2: Single flavor one dimensional limits for the electron (left) and muon (right) channels.

significance of this excess will be discussed briefly in the next section.

Limits combining the electron and muon channel data are shown in figure 8.3. These limits assume correlated systematic uncertainties for any systematic category which would affect both channels in the same way. This includes uncertainties such as the $t\bar{t}$ background estimation uncertainties, jet energy corrections, and luminosity uncertainties. Other uncertainties, such as MC statistical uncertainties and lepton energy resolution, are considered uncorrelated between the electron and muon channels. The combined channel limits show good agreement between observed and expected limits.

8.3 Significance and Excess Discussion

Due to the small excess of events over the background prediction around $M_{eejj} = 2\,\text{TeV}$ in the electron channel, the discovery significance for each signal mass assumption was calculated. The significance is calculated using the same tool and inputs as the limit setting. The plot of significance for the electron and muon channels is shown in figure 8.4(a). The maximum significance calculated in the electron channel of 2.8σ occurs for $M_{W_R} = 2.1\,\text{TeV}$. This significance is large enough to raise interest in this excess,

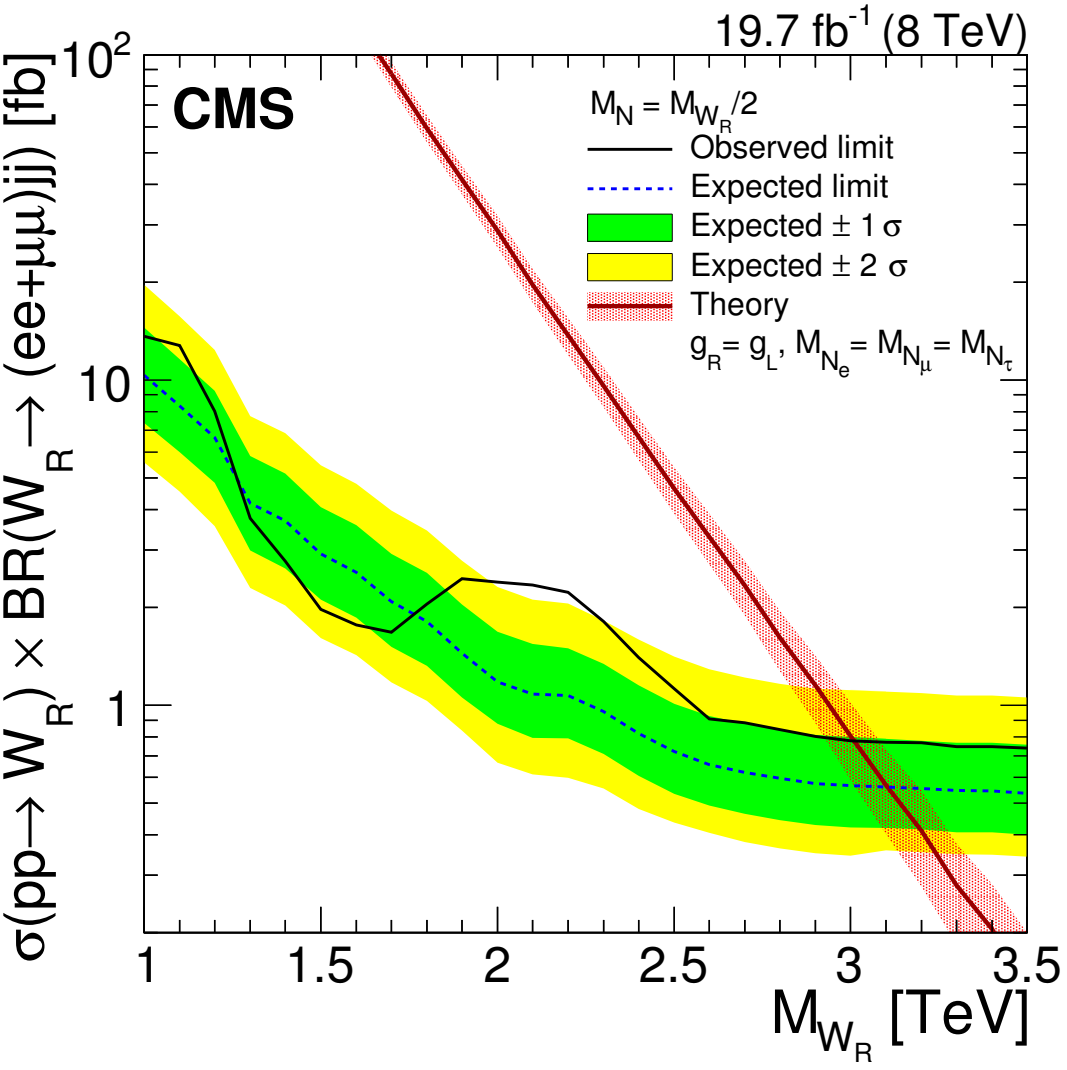
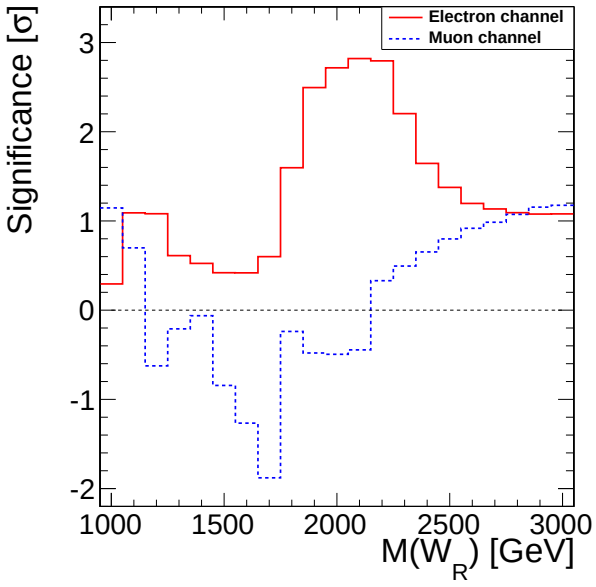


Figure 8.3: Combined electron and muon channel limits under the assumption of degenerate heavy neutrino masses.

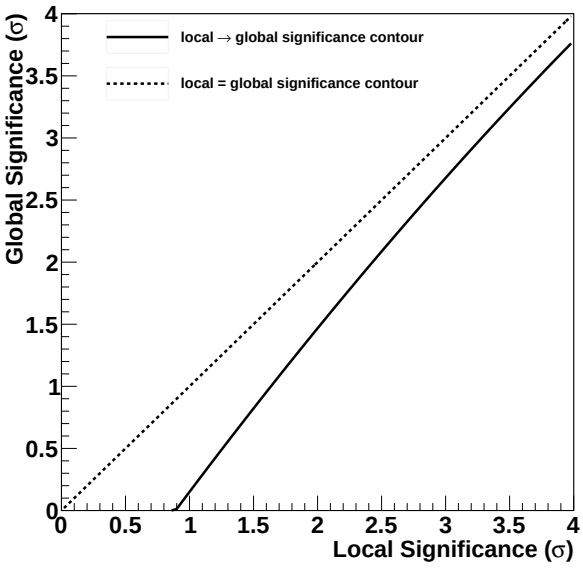
but insufficient to make any certain claims of disagreement with the SM. Additionally, because this is a multi-binned search, the look elsewhere effect must be taken into account.

Qualitatively, the look elsewhere effect is the phenomenon that the probability of a large statistical fluctuation in any one bin is higher than a similar statistical fluctuation occurring in one particular bin. Quantitatively this effect can be accounted for by calculating a conversion from the local significance, defined as the significance of the particular signal assumption, to the global significance which quantifies the probability of the ensemble of measurements made. This correction factor is derived using a series of toy MC based on the total background prediction. For each toy the significance is calculated using a simplified profile likelihood method for each signal mass assumption. The cumulative probability distribution (CPD) is constructed for the maximum significance and a randomly chosen significance from each toy. The global significance is then calculated by matching two CPDs such that the cumulative probability of the local significance on the randomly constructed CPD matches the cumulative probability on the CPD constructed from the maximum significances. The solid black line on figure 8.4(b) shows the mapping of local to global significance. The local significance of 2.8σ corresponds to a global significance of 2.4σ.

In addition to the calculation of significance, the signal like qualities of the electron channel excess have also been studied. A plot of M_{eejj} with a signal point with $M_{W_R} = 2.0\,\text{TeV}$ and $M_N = 1.67\,\text{TeV}$ is shown in figure 8.5 along with several kinematic quantities for events which pass all selection criteria plus the requirement $1.8 < M_{eejj} < 2.2\,\text{TeV}$. This particular signal point is chosen because it is the best match to the excess of the available signal assumptions generated. The signal matches the shape of the excess well in M_{eejj}, but all the remaining plots indicate that this excess is not compatible with a W_R. The M_{ee} distribution of the excess would be expected to extend to larger masses if these events came from the decay of a W_R. Furthermore, there is no distinct mass peak reconstructed for the invariant mass distribution of the e_1jj system where e_1 stands for the sub-leading electron in E_T. Finally, the low peak in the $p_\text{T}(e_1)$ distribution suggests that the heavy neutrino would need to be very heavy ($M_N \gg M_{W_R}/2$) to suppress the p_T of the lepton which comes directly from the W_R decay. However, a very heavy N_e would still produce a large high-E_T tail for the leading electron.

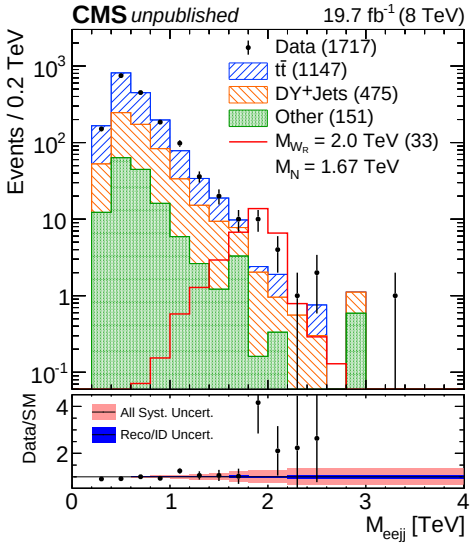


(a)

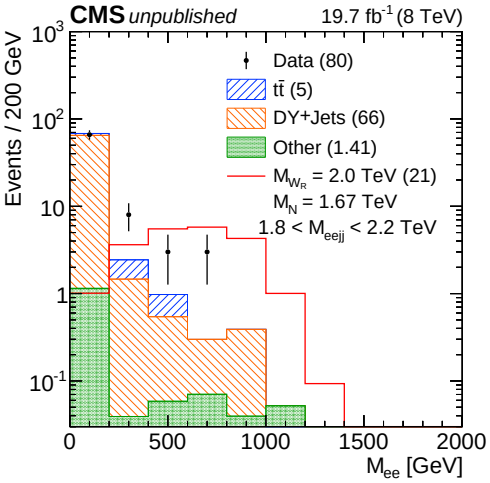


(b)

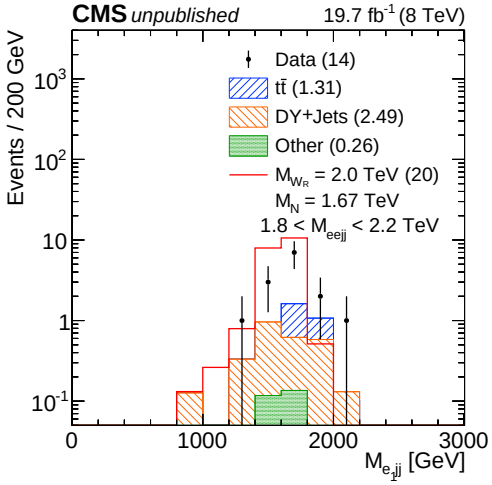
Figure 8.4: Local significance calculations for the electron and muon channels as a function of the W_R mass assumption (top). Conversion chart for local to global significance (bottom). The solid black line shows the conversion from local to global significance.



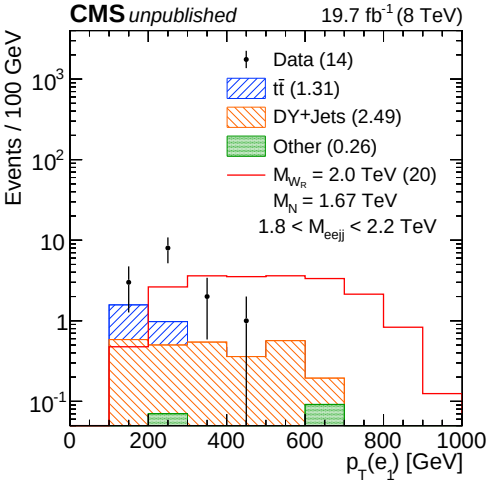
(a) M_{eejj}



(b) M_{ee}



(c) $M_{ejj}(e1)$



(d) $p_{\mathrm{T}}(e1)$

Figure 8.5: Plots showing the electron channel excess along with signal assumption with $M_{W_R} = 2.0\,\mathrm{TeV}$ and $M_N = 1.67\,\mathrm{TeV}$. Plot (a) shows the M_{eejj} distribution. Plots (b), (c), and (d) show M_{ee}, $M_{ejj}(e2)$, $p_{\mathrm{T}}(e1)$ distributions for events from the region $1.8 < M_{W_R} < 2.2\,\mathrm{TeV}$.

8.4 Two Dimensional Limits

After calculating the one dimensional limits along the line of $M_N = M_{W_R}/2$, the limits are extrapolated to find exclusion limits in the full $M_N - M_{W_R}$ plane. It is possible to compute the two dimensional limits in the same fashion as the one dimensional limits, but this requires an order of magnitude more signal MC be generated in a full grid of M_N and M_{W_R} points. In order to reduce the strain of MC production, a much faster technique using the acceptance of the signal as a function of M_N is used to map the limits from $M_N = M_{W_R}/2$ to the full plane. This technique is successful because the effects of efficiency and acceptance can be largely decoupled for a constant M_{W_R}. In the following section the mapping technique will be discussed and the two dimensional limits will be presented.

8.4.1 Limit Mapping Procedure

The procedure for mapping the $\sigma \times B$ limits to the full $M_N - M_{W_R}$ plane is performed using MC samples generated in a grid with $1.0 < M_{W_R} < 3.4\,\text{TeV}$ and $M_N < M_{W_R}$. The grid is generated in $100\,\text{GeV}$ steps along M_{W_R} and $50\,\text{GeV}$ steps along M_N. To save time, this MC is not run through detector simulation or reconstruction. Instead the leptons and jets produced directly from PYTHIA are used. This is acceptable because the efficiency is basically constant along lines of constant M_{W_R}. This means that this MC only needs to model the acceptance corrections correctly instead of the more difficult efficiency effects. Two different techniques are used for $M_N > M_{W_R}/2$ and for $M_N < M_{W_R}/2$.

The first technique is used for mapping the limits when $M_N > M_{W_R}/2$. This region is the simpler of the two because the corrections are small and the signal shape as a function of M_{lljj} is stable. For each MC mass point generated, the acceptance for the particular mass assumption is estimated using a simplified version of the analysis. The simplified analysis begins by selecting the two highest-p_T leptons and two highest-p_T jets to build a W_R candidate. The selected objects are then subjected to the object p_T requirements, the $\Delta R > 0.5$ requirements between all lepton and jet combinations, and the $M_{ll} > 200\,\text{GeV}$ requirement. The acceptance, $A\left(M_{W_R}, M_N\right)$, is then defined as the number of events passing these requirements with $M_{lljj} > 600\,\text{GeV}$ divided by the total

number of events in the sample. The limits are extrapolated to larger M_N along lines of constant M_{W_R} based upon the ratio

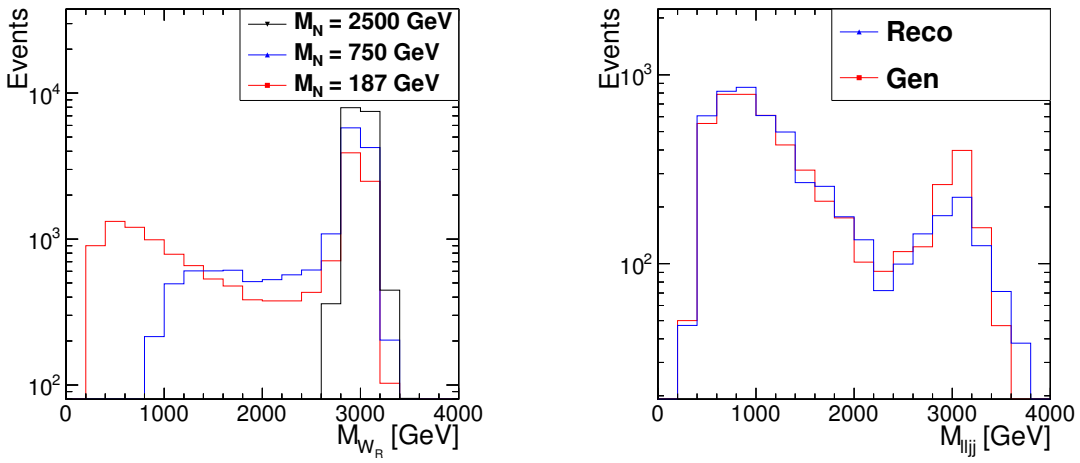
$$\sigma \times B\left(M_{W_R}, M_N\right) = \sigma \times B\left(M_{W_R}, M_{W_R}/2\right) \frac{A\left(M_{W_R}, M_N\right)}{A\left(M_{W_R}, M_{W_R}/2\right)}. \tag{8.7}$$

The second technique is used for the limit mapping when $M_N < M_{W_R}/2$. In this region, the acceptance changes more dramatically due to the decay products of the highly boosted heavy neutrino decay failing the $\Delta R > 0.5$ requirement. The signal shape also changes as a function of M_N in this region. If $M_N \ll M_{W_R}$ then the W_R production can proceed more readily through a light W_R^* than an on-shell W_R due to the large number of low momentum-fraction partons in the proton. These effects require a more sophisticated method than the solution that is applicable above $M_{W_R}/2$. The effect this has on the M_{lljj} distribution is shown on the left in figure 8.6. Again the signal MC without simulation or reconstruction is used with the same basic selection as described above. Instead of evaluating the acceptance, the signal shape is estimated. The right of figure 8.6 shows the M_{lljj} shape for a sample signal point compared to the distribution from fully reconstructed MC which shows good replication of the signal shape with the basic PYTHIA level MC. The shape is then used as input to the nominal limit setting code, but with systematic uncertainties disabled. The results of these limits, $\xi\left(M_{W_R}, M_N\right)$, are then used to derive a correction factor as

$$\sigma \times B\left(M_{W_R}, M_N\right) = \sigma \times B\left(M_{W_R}, M_{W_R}/2\right) \frac{\xi\left(M_{W_R}, M_N\right)}{\xi\left(M_{W_R}, M_{W_R}/2\right)}. \tag{8.8}$$

The theoretical $\sigma \times B$ values are extrapolated using a similar technique. The NLO cross-section as calculated at $M_N = M_{W_R}/2$ are extrapolated to other values of M_N by the ratio of the LO $\sigma \times B$ which is calculated by PYTHIA for each point in the MC grid generated. The theoretical $\sigma \times B$ values are relatively slowly changing until M_N approaches M_{W_R} at which point the branching ratio to heavy neutrinos becomes heavily suppressed.

The success of this limit mapping technique was checked using the sparse grid of fully reconstructed signal MC generated as described in section 5.1.2. The signal efficiency for each mass point in the grid was calculated and the limits computed using the full limit setting tool with systematic uncertainties enabled for both electron and muon channels. These limits are then compared to the limits derived from the limit mapping method.



(a) (b)

Figure 8.6: M_{lljj} distributions for three M_N assumptions with $M_{W_R} = 3.0\,\text{TeV}$ (left). M_{lljj} distributions for fully reconstructed and PYTHIA generator signal with $M_{W_R} = 3.0\,\text{TeV}$ and $M_N = 0.187\,\text{TeV}$ (right).

The ratio of the two limit determination methods, parameterized as a function of the ratio M_N/M_{W_R} is shown in figure 8.7. The ratio is near one except for $M_N/M_{W_R} < 0.2$ the ratio deviates noticeably from one. As a result of this deviation we apply this ratio as a correction factor to the limits. To smooth between points the results are fit to a function of the form

$$f = 1 - e^{a+bM_N/M_{W_R}}. \tag{8.9}$$

The corrections are applied independently for the electron and muon channels to account for any possible acceptance and efficiency differences which are not accounted for by the mapping method. A systematic uncertainty is applied to the correction as the larger of half the spread of the points for any M_N/M_{W_R} value or the statistical uncertainty.

8.4.2 Two Dimensional Limit Results

After applying all correction and scaling factors, the mapped limits can be seen for $M_{W_R} = 2\,\text{TeV}$ in figure 8.8. The limits are reasonably flat for intermediate and large values of M_N, but at small values of M_N the experimental limit jumps drastically due

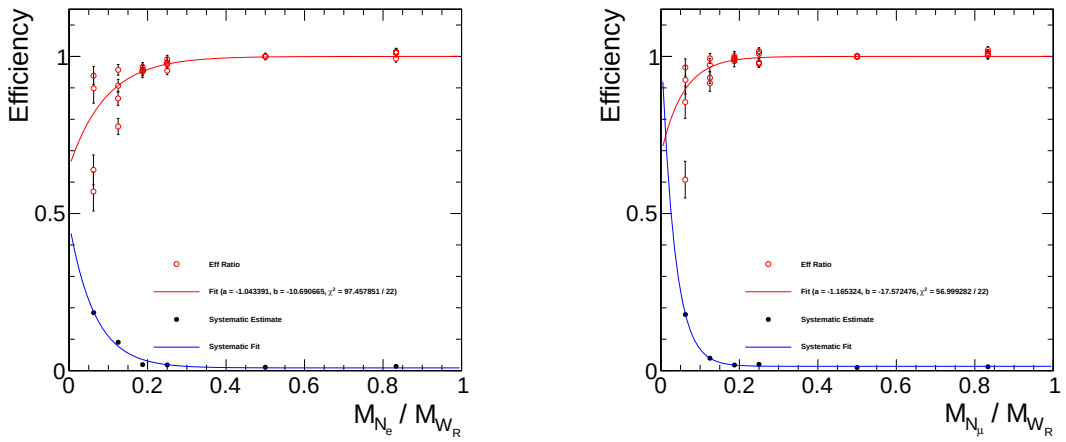
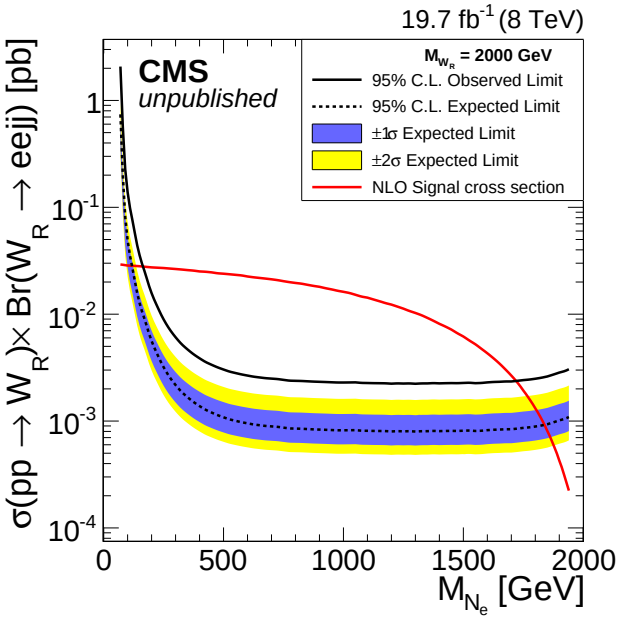
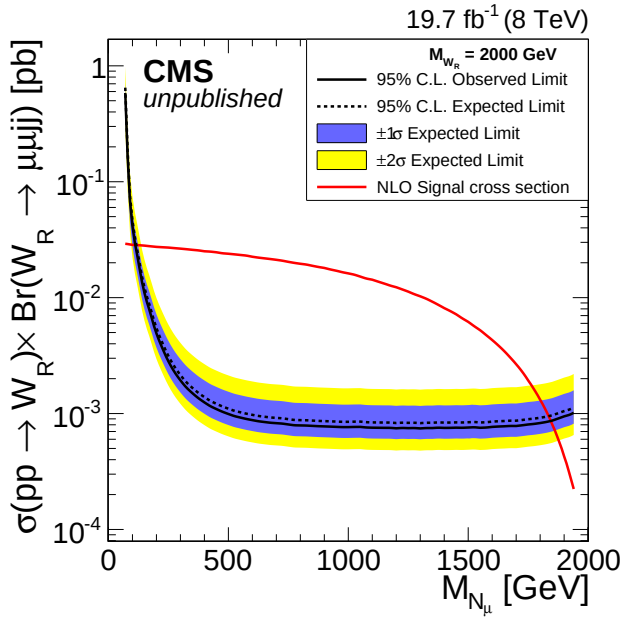


Figure 8.7: Acceptance ratio correction factors for electron (left) and muon (right) channel.

to the acceptance loss from the $\Delta R < 0.5$ separation requirements. At large M_N the steeply falling theoretical expectation restricts the mass limits. The excluded region is defined for each value of M_{W_R} as the region of M_N above the experimental exclusion line, but below the theoretical expectation for $\sigma \times B$. The total excluded region is built by combining all values of M_{W_R}. The exclusion contours are smoothed between values of M_{W_R} using linear interpolation based on the strength of the $\sigma \times B$ limits. The single channel limits are shown in figure 8.9 while the combined channel limits, again under the assumption of degenerate neutrino masses, are shown in figure 8.10. The difference between the expected and observed limit contours in figures 8.8(a), 8.9(a), and 8.10 and due to the excess of events observed in the electron channel.

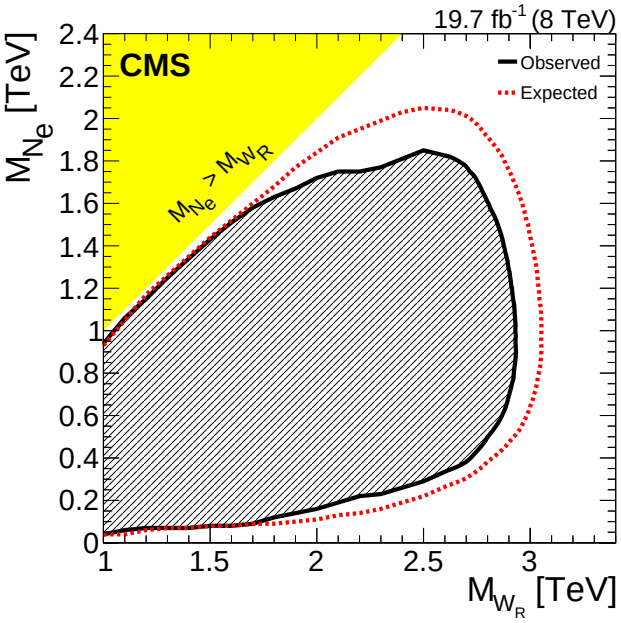


(a)

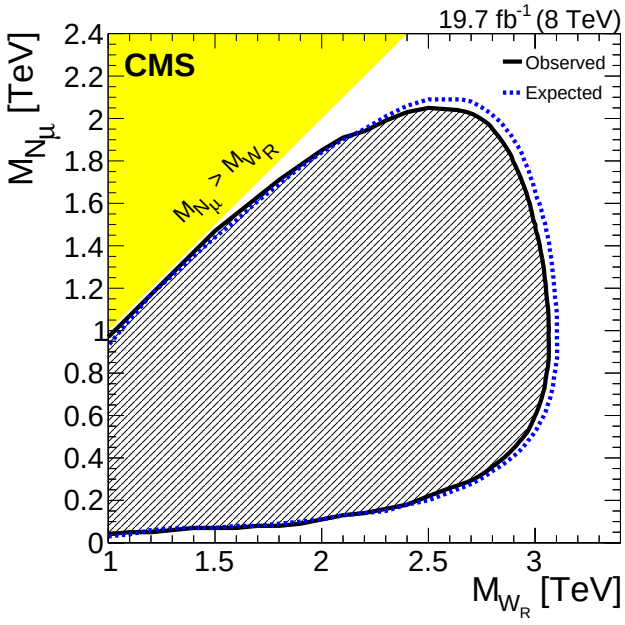


(b)

Figure 8.8: Mapped limits as a function of M_N with $M_{W_R} = 2.0\,\text{TeV}$ for electron (top) and muon (bottom) channel.



(a)



(b)

Figure 8.9: Two dimensional mass exclusion limits for electron (top) and muon (bottom) channel.

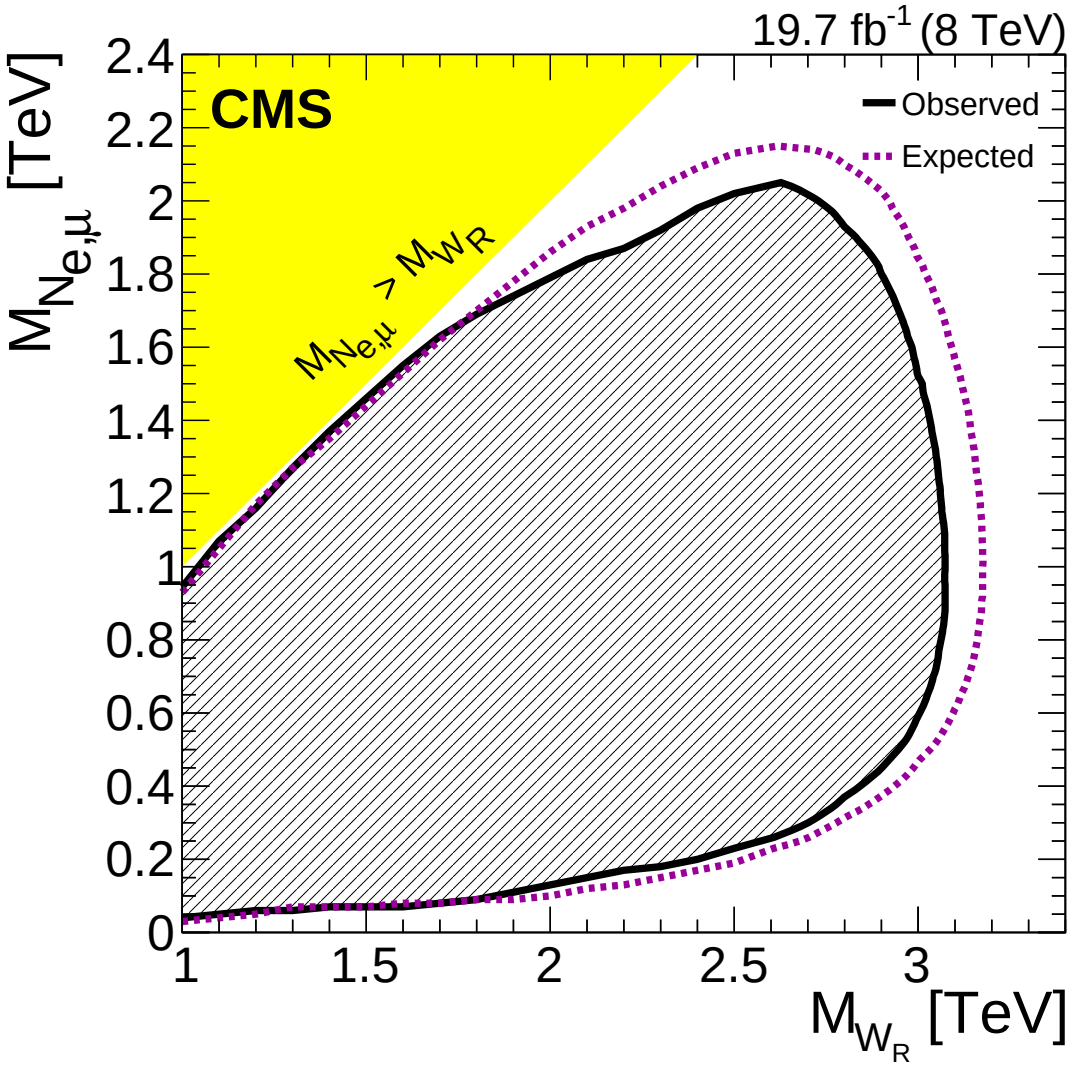


Figure 8.10: Two dimensional mass exclusion limits for combined electron and muon channel.

Chapter 9

Conclusion

The search for the decay $W_R \to lN_l \to lljj$ was performed using the full $19.7\,\text{fb}^{-1}$ of $\sqrt{s} = 8\,\text{TeV}$ center of mass data collected with the CMS detector. This analysis focused on the search for electron and muon flavor heavy neutrinos with the final states $eejj$ and $\mu\mu jj$ respectively. There was no statistically significant excesses of events over the standard model expectation for either channel, but a small excess of events is present in the electron channel with $M_{lljj} \approx 2\,\text{TeV}$. The electron channel excess has a local significance of 2.8σ for $M_{W_R} = 2.1\,\text{TeV}$ which corresponds to a global significance of 2.4σ. Although this suggests that this final state merits continued attention in the next LHC run, this is not sufficient evidence to claim disagreement with the standard model. Furthermore, when studied in detail, this excess was found to be unlikely to be caused by W_R decay.

The data are used to place CL_S 95% upper limits on the allowed cross-section times branching ratio of the process $q\bar{q}' \to W_R \to lN_l$. Limits are set individually for the electron and muon channels assuming that only one heavy neutrino flavor is light enough to be produced at LHC energies. Limits are also set using the combined electron and muon data under the assumption that the heavy neutrinos have degenerate mass. These limits are used to derive an excluded region in $M_{W_R} - M_N$ space based on the assumption of strict left-right symmetry. These limits are the first direct search results to reach to the theoretically interesting region, $M_{W_R} > 3\,\text{TeV}$, allowed by B meson mixing measurements. The excluded region, as a function of M_N, extends as high as $M_{W_R}{\sim}3\,\text{TeV}$ in all three search channels.

References

[1] Laurent Canetti, Marco Drewes, and Mikhail Shaposhnikov. Matter and antimatter in the universe. *New Journal of Physics*, 14(9):095012, 2012.

[2] Steven Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967.

[3] A Salam. Elementary particle physics: Relativistic groups and analyticity, vol. In *Proceedings to the Eighth Nobel Symposium, Almquvist and Wiksell*, 1968.

[4] Jogesh C. Pati and Abdus Salam. Lepton number as the fourth "color". *Phys. Rev. D*, 10:275–289, Jul 1974.

[5] R. N. Mohapatra and X. Zhang. Electroweak baryogenesis in left-right-symmetric models. *Phys. Rev. D*, 46:5331–5336, Dec 1992.

[6] Murray Gell-Mann, Pierre Ramond, and Richard Slansky. Complex Spinors and Unified Theories. *Conf.Proc.*, C790927:315–321, 1979, arXiv:1306.4669.

[7] S. et. al. Chatrchyan. Search for heavy neutrinos and W_R bosons with right-handed couplings in a left-right symmetric model in pp collisions at $\sqrt{s}$=7 TeV. *Phys. Rev. Lett.*, 109:261802, Dec 2012.

[8] Vardan Khachatryan et al. Search for heavy neutrinos and W bosons with right-handed couplings in proton-proton collisions at sqrt(s) = 8 TeV. 2014, arXiv:1407.3683.

[9] J Clerk Maxwell. A dynamical theory of the electromagnetic field. *Proceedings of the Royal Society of London*, 13:531–536, 1863.

[10] Hendrik Lorentz. Simplified Theory of Electrical and Optical Phenomena in Moving Systems. *Proceedings of the Royal Netherlands Academy of Arts and Sciences*, 1:427442, 1899.

[11] Albert Einstein. On the electrodynamics of moving bodies. *Annalen Phys.*, 17:891–921, 1905.

[12] P. A. M. Dirac. The Quantum Theory of the Electron. *Royal Society of London Proceedings Series A*, 117:610–624, February 1928.

[13] Carl D. Anderson. The positive electron. *Phys. Rev.*, 43:491–494, Mar 1933.

[14] R. P. Feynman. Space-time approach to quantum electrodynamics. *Phys. Rev.*, 76:769–789, Sep 1949.

[15] E. Fermi. Versuch einer theorie der β-strahlen. I. *Zeitschrift fr Physik*, 88(3-4):161–177, 1934.

[16] J. Orear, G. Harris, and S. Taylor. Spin and parity analysis of bevatron τ mesons. *Phys. Rev.*, 102:1676–1684, Jun 1956.

[17] T. D. Lee and C. N. Yang. Question of parity conservation in weak interactions. *Phys. Rev.*, 104:254–258, Oct 1956.

[18] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental test of parity conservation in beta decay. *Phys. Rev.*, 105:1413–1415, Feb 1957.

[19] Richard L. Garwin, Leon M. Lederman, and Marcel Weinrich. Observations of the failure of conservation of parity and charge conjugation in meson decays: the magnetic moment of the free muon. *Phys. Rev.*, 105:1415–1417, Feb 1957.

[20] ECG Sudarshan and RE Marshak. The nature of the four-fermion interaction. In *Proceedings of Padua-Venice Conference on Mesons and Newly Discovered Particles*, 1957.

[21] R. P. Feynman and M. Gell-Mann. Theory of the fermi interaction. *Phys. Rev.*, 109:193–198, Jan 1958.

[22] C. N. Yang and R. L. Mills. Conservation of isotopic spin and isotopic gauge invariance. *Phys. Rev.*, 96:191–195, Oct 1954.

[23] Sheldon L. Glashow. Partial-symmetries of weak interactions. *Nuclear Physics*, 22(4):579 – 588, 1961.

[24] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964.

[25] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.

[26] GD Rochester and Clifford Charles Butler. Evidence for the existence of new unstable elementary particles. *Nature*, 160(4077):855, 1947.

[27] M. Gell-Mann. A schematic model of baryons and mesons. *Physics Letters*, 8(3):214 – 215, 1964.

[28] G. Zweig. An SU(3) model for strong interaction symmetry and its breaking. Version 1. CERN-TH-401. 1964.

[29] F.J. Hasert et. al. Search for elastic muon-neutrino electron scattering. *Physics Letters B*, 46(1):121 – 124, 1973.

[30] J. J. Aubert et. al. Experimental Observation of a Heavy Particle J. *Phys. Rev. Lett.*, 33:1404–1406, Dec 1974.

[31] J. E. Augustin et. al. Discovery of a narrow resonance in e+e- annihilation. *Phys. Rev. Lett.*, 33:1406–1408, Dec 1974.

[32] S. W. Herb et. al. Observation of a dimuon resonance at 9.5 GeV in 400-GeV proton-nucleus collisions. *Phys. Rev. Lett.*, 39:252–255, Aug 1977.

[33] G. Arnison et. al. Experimental observation of lepton pairs of invariant mass around 95 GeV/c2 at the CERN SPS collider. *Physics Letters B*, 126(5):398 – 410, 1983.

[34] P. Bagnaia et. al. Evidence for Z0→e+e at the CERN pp collider. *Physics Letters B*, 129(12):130 – 140, 1983.

[35] G. Arnison et. al. Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s}$=540 gev. *Physics Letters B*, 122(1):103 – 116, 1983.

[36] M. Banner et. al. Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN pp collider. *Physics Letters B*, 122(56):476 – 485, 1983.

[37] F. Abe et. al. Evidence for top quark production in $p\bar{p}$ collisions at $\sqrt{}$s =1.8 TeV. *Phys. Rev. D*, 50:2966–3026, Sep 1994.

[38] S. Abachi et. al. Search for high mass top quark production in p$\bar{p}$ collisions at s= 1.8 TeV. *Phys. Rev. Lett.*, 74:2422–2426, Mar 1995.

[39] S. Chatrchyan et. al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716(1):30 – 61, 2012.

[40] G. Aad et. al. Observation of a new particle in the search for the standard model higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):1 – 29, 2012.

[41] J. Beringer et al. (Particle Data Group). *Phys. Rev.*, D86(010001):583 – 585, 2012.

[42] Standard model. `http://en.wikipedia.org/wiki/Standard_Model`, June 2014.

[43] Q. Ho-Kim & Pham X. Y. *Elementary Particles and Their Interaction: Concepts and Phenomena*. Springer.

[44] Raymond Davis, Don S. Harmer, and Kenneth C. Hoffman. Search for neutrinos from the sun. *Phys. Rev. Lett.*, 20:1205–1209, May 1968.

[45] Q. R. Ahmad et. al. Direct evidence for neutrino flavor transformation from neutral-current interactions in the sudbury neutrino observatory. *Phys. Rev. Lett.*, 89:011301, Jun 2002.

[46] Ettore Majorana. Teoria simmetrica dellelettrone e del positrone. *Il Nuovo Cimento*, 14(4):171–184, 1937.

[47] A. G. Akeroyd, Mayumi Aoki, and Yasuhiro Okada. Lepton flavor violating τ decays in the left-right symmetric model. *Phys. Rev. D*, 76:013004, Jul 2007.

[48] G. Senjanović and R. N. Mohapatra. Exact left-right symmetry and spontaneous violation of parity. *Phys. Rev. D*, 12:1502–1505, Sep 1975.

[49] Goran Senjanović. Spontaneous breakdown of parity in a class of gauge theories. *Nuclear Physics B*, 153(0):334 – 364, 1979.

[50] Mohapatra, N. Rabindra, and Goran Senjanović. Neutrino mass and spontaneous parity nonconservation. *Phys. Rev. Lett.*, 44:912–915, Apr 1980.

[51] Goran Senjanović. Seesaw at LHC through leftright symmetry. *International Journal of Modern Physics A*, 26(09):1469–1491, 2011.

[52] CMS Collaboration. Search for resonances in the dilepton mass distribution in pp collisions at sqrt(s) = 8 TeV. (CMS-PAS-EXO-12-061), 2013.

[53] S. Chatrchyan et. al. Search for W→tb decays in the lepton + jets final state in pp collisions at $\sqrt{s} = 8$ TeV. *Journal of High Energy Physics*, 2014(5), 2014.

[54] Search for narrow t + b resonances in the leptonic final state at sqrt(s) = 8 TeV. Technical Report CMS-PAS-B2G-12-010, CERN, Geneva, 2013.

[55] Alessio Maiezza, Miha Nemevšek, Fabrizio Nesti, and Goran Senjanović. Left-right symmetry at lhc. *Phys. Rev. D*, 82:055022, Sep 2010.

[56] CP violation. `http://en.wikipedia.org/wiki/CP_violation`, June 2014.

[57] V. M. et. al. Abazov. Evidence for an anomalous like-sign dimuon charge asymmetry. *Phys. Rev. D*, 82:032001, Aug 2010.

[58] LHCb Collaboration. Measurement of the flavour-specific CP violating asymmetry a^s_{sl} in $\bar{B}^0_s$ decays. CERN-LHCb-CONF-2012-022. 2012.

[59] Stefano Bertolini, Alessio Maiezza, and Fabrizio Nesti. Present and Future K and B Meson Mixing Constraints on TeV Scale Left-Right Symmetry. 2014, arXiv:1403.7112.

[60] P. Achard et. al. Search for heavy neutral and charged leptons in e+e annihilation at LEP. *Physics Letters B*, 517(12):75 – 85, 2001.

[61] Torbjrn Sjstrand, Stephen Mrenna, and Peter Skands. PYTHIA 6.4 physics and manual. *Journal of High Energy Physics*, 2006(05):026, 2006.

[62] M. Hirsch, H.V. Klapdor-Kleingrothaus, and O. Panella. Double beta decay in left-right symmetric models. *Physics Letters B*, 374(13):7 – 12, 1996.

[63] Bayatian, G L et. al. *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*. Technical Design Report CMS. CERN, Geneva, 2006.

[64] Oliver Sim Brning, Paul Collier, P Lebrun, Stephen Myers, Ranko Ostojic, John Poole, and Paul Proudlock. *LHC Design Report*. CERN, Geneva, 2004.

[65] Jean-Luc Caron. The LHC injection complex.. L'ensemble d'injection du LHC. AC Collection. Legacy of AC. Pictures from 1992 to 2002., May 1993.

[66] Mike Lamont. Status of the LHC. *Journal of Physics: Conference Series*, 455(1):012001, 2013.

[67] Compact muon solenoid. `http://it.wikipedia.org/wiki/Compact_Muon_Solenoid`, June 2014.

[68] V. Veszpremi. Operation and performance of the CMS tracker. *JINST*, 9:C03005, 2014, arXiv:1402.0675.

[69] David Barney. The CMS electromagnetic calorimeter its performance and role in the discovery of the Higgs boson and perspectives for the future. Technical Report CMS-CR-2013-410, CERN, Geneva, Nov 2013.

[70] Subatomic physics detectors. `http://www.upscale.utoronto.ca/GeneralInterest/DBailey/SubAtomic/Lectures/LectF05/Lect05.htm`, June 2014.

[71] Abdullin, S. et. al. The CMS barrel calorimeter response to particle beams from 2 to 350 GeV/c. *The European Physical Journal C*, 60(3):359–373, 2009.

[72] S Chatrchyan et al. Performance of CMS Hadron Calorimeter Timing and Synchronization using Test Beam, Cosmic Ray, and LHC Beam Data. *JINST*, 5:T03013, 2010, arXiv:0911.4877.

[73] L Rossi. LHC Upgrade Plans: Options and Strategy. (CERN-ATS-2011-257):6 p, Dec 2011.

[74] S.V. Afanasiev, P. de Barbaro, I.A. Golutvin, I.F. Emeliantchik, A.I. Malakhov, P.V. Moisenz, V.A. Smirnov, and N.M. Shumeiko. Improvement of radiation hardness of the sampling calorimeters based on plastic scintillators. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 717(0):11 – 13, 2013.

[75] J Mans, J Anderson, B Dahmes, P de Barbaro, J Freeman, T Grassi, E Hazen, J Mans, R Ruchti, I Schimdt, T Shaw, C Tully, J Whitmore, and T Yetkin. CMS Technical Design Report for the Phase 1 Upgrade of the Hadron Calorimeter. Technical Report CERN-LHCC-2012-015. CMS-TDR-10, CERN, Geneva, Sep 2012.

[76] LHC luminosity plots for the 2012 proton run. `http://lpc.web.cern.ch/lpc/lumiplots_2012.htm`, June 2014.

[77] Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer. MadGraph 5: going beyond. *Journal of High Energy Physics*, 2011(6), 2011.

[78] T. Gleisberg, S. Hche, F. Krauss, M. Schnherr, S. Schumann, F. Siegert, and J. Winter. Event generation with SHERPA 1.1. *Journal of High Energy Physics*, 2009(02):007, 2009.

[79] Paolo Nason. A new method for combining nlo qcd with shower monte carlo algorithms. *Journal of High Energy Physics*, 2004(11):040, 2004.

[80] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing nlo calculations in shower monte carlo programs: the powheg box. *Journal of High Energy Physics*, 2010(6), 2010.

[81] Emanuele Re. Single-top Wt-channel production matched with parton showers using the POWHEG method. *The European Physical Journal C*, 71(2), 2011.

[82] James Botts, Jorge G. Morfin, Joseph F. Owens, Jianwei Qiu, Wu-Ki Tung, and Harry Weerts. CTEQ parton distributions and flavor dependence of sea quarks. *Physics Letters B*, 304(12):159 – 166, 1993.

[83] S. Agostinelli et. al. Geant4 - a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250 – 303, 2003.

[84] Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV. (CMS-PAS-PFT-10-002), 2010.

[85] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063, 2008.

[86] Matteo Cacciari and Gavin P. Salam. Pileup subtraction using jet areas. *Physics Letters B*, 659(12):119 – 126, 2008.

[87] CMS Collabortaion. JEC performance plots for 2013. `https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsJME2013JEC`, June 2014.

[88] The CMS collaboration. Determination of jet energy calibration and transverse momentum resolution in cms. *Journal of Instrumentation*, 6(11):P11002, 2011.

[89] Jet Energy Resolution in CMS at sqrt(s)=7 TeV. Technical Report CMS-PAS-JME-10-014, CERN, Geneva, 2011.

[90] The CMS collaboration. Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV. *Journal of Instrumentation*, 7(10):P10002, 2012.

[91] S. Chatrchyan. Search for heavy narrow dilepton resonances in pp collisions at sqrt(s) = 7 and 8 tev. *Physics Letters B*, 720(13):63 – 82, 2013.

[92] CMS Collabortaion. Baseline muon selections. `https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMuonId`, June 2014.

[93] Muon ID and Isolation Efficiencies in 2012 RunAB, CERN-CMS-DP-2012-025. 2012.

[94] Electron reconstruction and identification at sqrt(s) = 7 TeV. Technical Report CMS-PAS-EGM-10-004, CERN, Geneva, 2010.

[95] W Adam, R Frhwirth, A Strandlie, and T Todorov. Reconstruction of electrons with the gaussian-sum filter in the cms tracker at the lhc. *Journal of Physics G: Nuclear and Particle Physics*, 31(9):N9, 2005.

[96] CMS Collabortaion. HEEP Electron ID and isolation. `https://twiki.cern.ch/twiki/bin/viewauth/CMS/HEEPElectronID`, June 2014.

[97] The CMS collaboration. Identification of b-quark jets with the cms experiment. *Journal of Instrumentation*, 8(04):P04013, 2013.

[98] Ryan Gavin, Ye Li, Frank Petriello, and Seth Quackenbush. FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order. *Computer Physics Communications*, 182(11):2388 – 2403, 2011.

[99] Nicolas Christou. Poisson, Gamma, and Exponential distributions. `http://www.stat.ucla.edu/~nchristo/statistics100B/stat100b_poisson_gamma.pdf`, January 2007.

[100] CMS Collabortaion. Jet Energy Corrections: Official Software Tools for applying JEC Corrections and Uncertainties. `https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookJetEnergyCorrections`, July 2014.

[101] CMS Collabortaion. Electroweak Utilities for MC corrections and systematics evaluation. `https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideEWKUtilities`, July 2014.

[102] A L Read. Presentation of search results: the CLs technique. *Journal of Physics G: Nuclear and Particle Physics*, 28(10):2693, 2002.

[103] Thomas Junk. Confidence level computation for combining searches with small statistics. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 434(23):435 – 443, 1999.

[104] Lorenzo Moneta, Kevin Belasco, Kyle S. Cranmer, S. Kreiss, Alfio Lazzaro, et al. The RooStats Project. *PoS*, ACAT2010:057, 2010, arXiv:1009.1003.

Appendix A

Glossary and Acronyms

This appendix defines jargon terms in a glossary, and contains a table of acronyms and their meaning.

A.1 Glossary

- **barrel** – The central region of each sub-detector in η is called the barrel.
- **confinement** – Property of the strong force which forces hadrons to be color neutral.
- **endcap** – The portion of each sub-detector which cover the flat face at high $|\eta|$ are called the endcaps.
- **hadronization** – The process by which exposed color charges form hadrons to make themselves color neutral.
- **interaction length** – The characteristic length of a material after which $\frac{1}{e}$ of the energy of a hadronic shower remains.
- **interaction point** – The region at the center of the detector where proton-proton interactions occur.
- **jet** – A spray of strongly interacting particles which originates from the fracturing of hadronic particle.

- **LHC Run-I** – The operational period of the LHC extending from 2010 to 2013. During this period the LHC operated at both 7 and 8 TeV center of mass energy.
- **Moliere radius** – The radius inside which 90% of an electromagnetic shower is contained for a given material.
- **particle flow** – Reconstruction technique which combines information from all sub-detectors.
- **parton** – The individual constituents of a proton including the real valence quarks and virtual quarks, anti-quarks, and gluons.
- **pileup** – Additional proton-proton interactions which occur in the same event.
- **pre-scaled** – Indicates that only some of the events passing a particular trigger were actually saved to disk.
- **primary vertex** – The reconstructed vertex associated with the parton-parton collision.
- **radiation length** – The characteristic length of a material after which $\frac{1}{e}$ of the energy of an electromagnetic shower remains.
- **reconstruction** – The process of transforming raw detector response in to objects useful for physics analysis.
- **simulation** – The step of Monte Carlo production which uses GEANT to simulate the response of the detector to particle interactions.

A.2 Acronyms

Table A.1: Acronyms

Acronym	Meaning
APD	avalanche photo-diode
	Continued on next page

Table A.1 – continued from previous page

Acronym	Meaning
APM	auxiliary power mezzanine
ATLAS	a torroidal LHC apparatus
CCM	clock and control module
CDF	cumulative distribution function
CL	confidence level
CMS	compact muon solenoid
CSC	cathode strip chamber
DAQ	data acquisition
DCC	data concatenation card
DT	drift tube
ECAL	electromagnetic calorimeter
EEPROM	electrically erasable programmable read-only memory
FPGA	field programmable gate array
FSR	final state radiation
GOL	gigabit optical link
HCAL	hadronic calorimeter
HLT	high level trigger
HPD	hybrid photo-diode
HTR	HCAL trigger and readout
ISR	initial state radiation
L1	level one
LHC	large hadron collider
LRS	left-right symmetric
MC	Monte Carlo
PDF	parton distribution function
PM	power mezzanine
PS	proton syncrotron
QCD	quantum chromodynamics
QIE	charge integration encoder

Continued on next page

Table A.1 – continued from previous page

Acronym	Meaning
RBX	readout box
RM	readout module
RPC	resistive plate chamber
SPS	super proton syncrotron
VEV	vacuum expectation value
VPT	vacuum photo-triode