Search for Right-Handed Neutrinos and Weak Bosons at CMS with the LHC at $\sqrt{s} = 13\,\mathrm{TeV}$ and an integrated luminosity of $137\,\mathrm{fb}^{-1}$

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

Andrew Christopher Evans

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

Jeremiah Mans

December, 2020

Acknowledgements

This thesis would not be possible without the countless people in and throughout my life who have taught, encouraged, and inspired me. I would like to thank my parents, Bob and Karen, for raising me endlessly curious and passionate about truth. I also want to thank my siblings, Beth, Megan, and Laura, for keeping me humble, grounded, and never lonely. Without a doubt, this journey would be much more drab, and far the lesser without my endlessly faithful wife, Emma. Over the years I've been blessed with many friends, and I want to especially thank all of those in the Pays-de-Gex, particularly those who attended Crossroads Church Thoiry, for being our family and truly making us feel at home. I want to thank my advisor, Jeremy, for giving me all of the tools and experience to become a physicist. I want to thank my colleagues at CMS and UMN for all of the help and camaraderie: Alex, for getting me started with CMS, Seth, for being my de-facto postdoc for awhile, Nicole, for helping get this analysis off the ground, and Mike for seeing the project through. And of course I want to thank Caleb, Josh, Zach, and Tom for all the laughs and memes to keep me going. I would also like to thank Laura Clapper for illustrating my chapter introductions. Lastly, I want to thank my Lord and saviour Jesus Christ, for through him all things were made and in him is life, the light of all mankind.

This dissertation is based upon work supported by the U.S. Department of Energy, Office of Science, Office of High Energy Physics under Award Number DE-SC-0011845.

Dedication

"When I look at your heavens, the work of your fingers, the moon and the stars, which you have set in place, what is man that you are mindful of him, and the son of man that you care for him?"
Psalm 8:3-4 [1]

"So, whether you eat or drink, or whatever you do, do all to the glory of God." 1 Corinthians 10:31[1]

SDG

[1]Scripture quotations are from the ESV® Bible (The Holy Bible, English Standard Version®), copyright © 2001 by Crossway, a publishing ministry of Good News Publishers. Used by permission. All rights reserved.

Abstract

This thesis presents the most expansive limits on the W_R decay to two muons or electrons, and two jets at CMS. Left-right symmetric extensions to the Standard Model are motivated and presented giving rise to the purpose of this search. New techniques in boosted jet reconstruction are employed to expand past analytical approaches to now include much lighter N_R hypotheses. With a larger signal acceptance compared to previous analyses, a new signal selection strategy is developed and explained. Then, standard model backgrounds for this analysis are discussed and estimated using several control regions. Data collected by the CMS collaboration over three years of Run II for a total of $137\,\text{fb}^{-1}$ is combined and limits are set on the W_R mass at $4-5$ TeV and $4-5.5$ TeV for the electron and muon flavor search respectively for N_R masses ranging between $100-3000$ GeV.

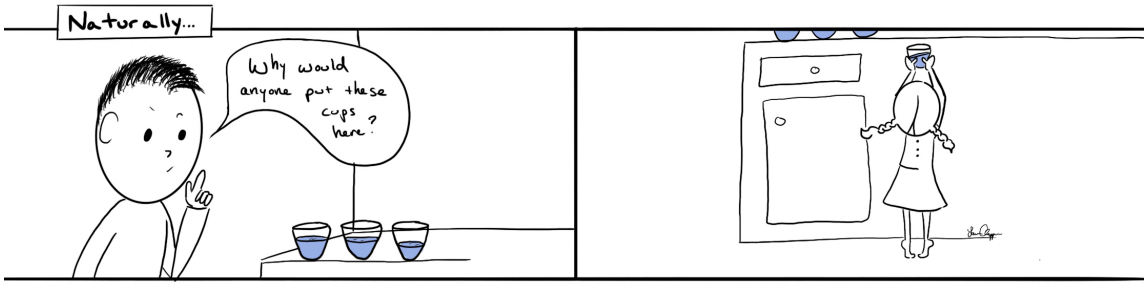
Contents

List of Tables

List of Figures

Chapter 1

Introduction



The Standard Model of particle physics provides a broad and precise description of physical phenomena at the most fundamental and minute level. While the performance of the theory is, quite frankly, stunning, the Standard Model does not explain everything. From the beginning, a core component of the Standard Model, the Weak Force boson (W), has only interacted with left-handed particles. While this behaviour was motivated by experiment, beginning with [1], why would the spin-momentum relationship (handedness) lead to this? All other fundamental interactions in the Standard Model proceed equally left-handed and right-handed. From this, one can ask the question, what if the weak force does interact in a symmetric way, but, the right-handed interaction is suppressed until sufficiently high energies are reached. The LHC, achieving approximately five times the collision energy of its predecessor, could create the high energy environment needed to solve this puzzle.

This thesis details an analysis performed with the CMS collaboration at CERN searching for a "left-right symmetric" addition to the Standard Model evidenced by a new right-handed neutrino and right-handed intermediating W_R boson.

The most effective method for finding a collision event of interest is typically to look for

the generation of particle kinds not present in the initial colliding particles. At the LHC, protons collide and quarks and gluons, interacting via QCD, abound. Leptons, however, can only be produced through an exchange of a different interaction: Weak or Electromagnetic. The W_R decay to a lepton and right-handed neutrino is then the logical interaction to search for. As the N_R is likely heavy and also decays, another lepton and two quarks are most likely produced in this process. The Feynman diagram for the W_R production from quarks and subsequent decay to a lepton and a N_R is shown below. This analysis combines the

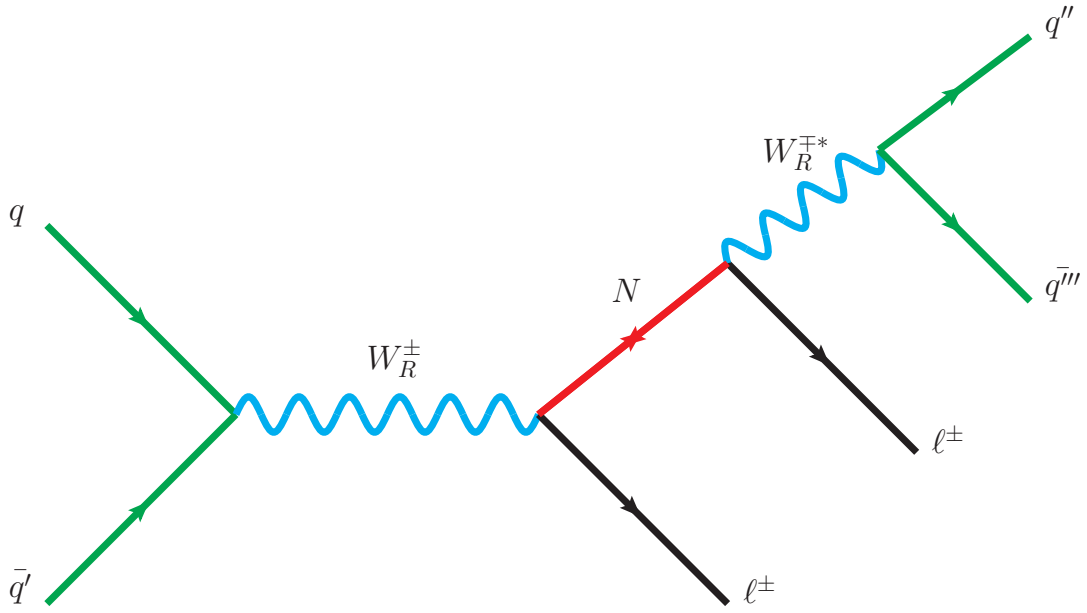


Figure 1.1: The principle Feynman diagram for this analysis. Two quarks annihilate into a right-handed W boson. This decays to a lepton and a right-handed neutrino. This neutrino decays by a virtual right-handed W and another lepton. This virtual boson decays to quarks, which produce jets.

technique of past CMS searches of identifying all four decay objects with a new technique relying on the resolution of just three. In addition, 2016, 2017 and, 2018 data sets are used, more than tripling the integrated luminosity of past searches.

Previous efforts have used lower collision energies and fewer events. They have been performed at $\sqrt{s} = 7\,\text{TeV}$, $\sqrt{s} = 8\,\text{TeV}$, and $\sqrt{s} = 13\,\text{TeV}$ [2][3][4] at CMS searching for the two lepton and two jet signature. While these analyses have excluded certain ranges of right-handed neutrino and W_R mass, there remains a possibility that the four particle

decay signature is not completely visible because the right-handed neutrino is, while heavy, much lighter than the W_R. In this case, the first muon and the right-handed neutrino decay relatively opposite of each other. This neutrino, carrying significant momentum, results in the next particles in the decay being Lorentz boosted together in the lab frame. As a result, these particles can be indistinguishable, or poorly reconstructed.

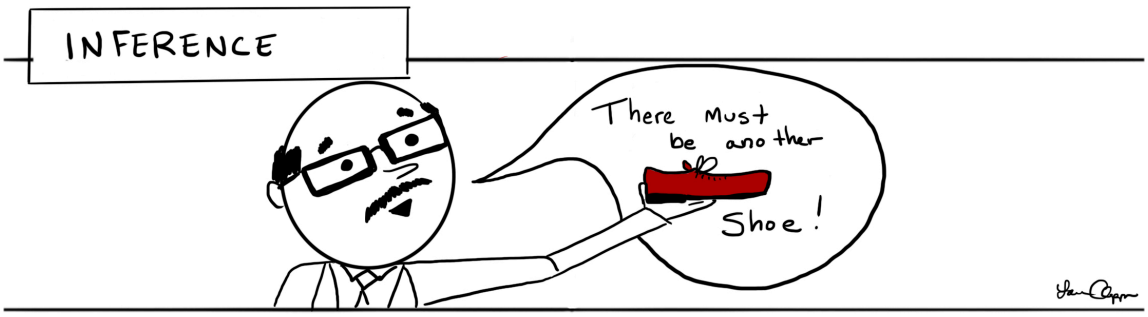
The event selection in this analysis is divided into two separate regions: boosted and resolved. The boosted event selection reconstructs the event with a high energy muon and large jet containing a muon with a novel isolation. The mass of the summed Lorentz vectors of these two objects is searched for excesses. In the resolved region, two jets and two muons, all well physically separated, are selected. These four objects are summed and the mass of the result is searched. Both of these selections are designed to reconstruct a signal well, control for backgrounds, and remain orthogonal to each other, allowing for a statistical recombination of the two selection regions.

The complete analysis represents the collaborative effort of several people. The search for W_R events decaying to muons, particularly the boosted topology, and its integration with the broader analysis, has been the focus of the author. The search for signal decaying to electrons is presented for completeness.

This thesis begins with an introduction to the Standard Model and its history in Chapter 2. An introduction to the CMS detector and the author's work on the HCAL phase I upgrade are in Chapter 3. Details on new boosted-jet techniques are covered in Chapter 4. The focus then shifts to the strategy of the analysis in Chapter 5 and the reconstruction and identification of particles in the analysis in Chapter 6. The estimation of backgrounds is discussed in Chapter 7 and the upper confidence limits on the W_R mass are provided in Chapter 8.

Chapter 2

Physics of the Standard Model and Left-Right Symmetric Extensions



2.1 Development of the Standard Model

The standard model of particle physics, or simply the standard model, encapsulates discoveries in physics stretching well over a century and was chiefly formulated by Weinberg, Glashow, and Salam [5][6]. The standard model combines electromagnetism, the nuclear weak force, and the nuclear strong force to describe the properties and interactions of all matter, excluding gravity.

At the beginning of the 20th century, Einstein and Lorentz developed special relativity. Maxwell's early work developing the four equations of electricity and magnetism was shown to be fundamentally Lorentz invariant, valid in all reference frames [7][8]. In the 1920s Dirac created the first quantum field-theory, recreating Maxwell's equations for relativistic and quantum cases [9]. With this reformulation, the first new particles, anti-particles, were

predicted and the fundamental spin of particles was explained. Dirac's predictions were confirmed by Anderson in 1932 when cosmic-rays were observed decaying to oppositely charged electrons (anti-electrons, or positrons) [10]. Then in the 1940s Feynman developed Quantum Electrodynamics (QED) and with the work of others created a completely-consistent theory of all electromagnetic interactions [11].

The weak nuclear force was first motivated by beta decay. Pauli noticed a lack of conservation of momentum in the decay, necessitating an extremely light and unseen particle, which was called "little neutral one", or neutrino. This allowed for momentum conservation and Fermi proposed a contact interaction between a proton, electron, and the neutrino. This interaction between four-fermions successfully predicted the spectrum of beta decay [12], though it has since been updated to include the weak boson, only directly observable at much higher energies than nuclear decays.

Into the 1950s further study of the weak force lead to a puzzle regarding the conservation of parity. Two particles were discovered, τ and θ, which were identical except for their parity [13]. Lee and Yang proposed that these were the same particles, which instead violated parity symmetry in its decay [14]. The measurement of the beta decay of polarized Co_{60} by Wu confirmed the parity-violating nature of the weak force [1]. Meson decays in a storage ring were also studied, revealing more evidence for parity violation [15]. Sudarshan, Marshak, Feynman, and Gell-Mann all worked to develop a new model for the weak force which would include parity-violation. This was the $V-A$ model, vector minus axial vector. Axial vectors do not transform the same under parity as regular vectors [16][17].

Parallel to the work developing the $V-A$ formulation, Yang and Mills created a non-Abelian gauge theory describing the weak force in 1954. This replaced the contact interaction, first theorized by Fermi, with a charge 1, spin 1 particle as an intermediary [18]. This theory was physically impossible for the weak force, as the weak boson proposed was massless, requiring the weak force to be of infinite range, like electromagnetism. Glashow modified the proposed theory in 1960. He added the $V-A$ model of Sudarshan and Marshak and combined Feynman's QED, describing it all in a single theory with the gauge group $\mathrm{SU(2)_L \times U(1)}$ [19] and predicting an additional heavy boson mixing with the photon. The weak $\mathrm{SU(2)_L}$ group only couples to left-handed chiral states. The last piece of the puzzle was completed by Weinberg and Salam [5][6]. They added the Brout, Englert, and Higgs mechanism [20][21] —which gives mass to vector bosons in the Yang-Mills theory—and predicted an additional boson, the Higgs, not to be discovered until 2012 [22][23].

At the same time as the development of the so called "electroweak sector" of the standard

model, cosmic rays and nuclear interactions were providing evidence for another fundamental force of a different sort, the strong nuclear force. Starting in 1947, cosmic ray events studied in cloud chambers led to the discovery of particles such as kaons [24]. These particles were strange, as they seemed to have an additional quantum number no other particles did. Gell-Mann and Zweig proposed a symmetry group $SU(3)_C$ to explain this [25] [26]. This quark model theorized that all strongly interacting particles were composed of fundamental quarks, three observed and one theorized. The extra property of kaons and other similar mesons, was that they were made partly of "strange quarks". More typical mesons, as well as protons and neutrons, were made of "up" and "down" quarks. The theory predicted one more quark, which was discovered and named "charm". The theory was later extended to include two more, the "bottom" and "top" quarks. The interactions in the quark model are based on color charge and only color neutral particles can be freely propagating. Adding the $SU(3)_C$ group to the electroweak sector gives the complete standard model picture, a $SU(3)_C \times SU(2)_L \times U(1)$ symmetry group.

Every particle predicted in the standard model has now been observed. The first indirect observation of the Z was in 1973 at CERN [27]. The charm quark, confirming quark theory, was discovered at BNL and SLAC and then the bottom quark was soon discovered by the E288 experiment [28][29][30]. In 1983 the weak bosons were directly discovered by the UA1 and UA2 experiments at CERN [31][32][33][34]. The top quark, the heaviest particle in the standard model, was discovered in 1995 at FNAL by the D0 and CDF groups using the Tevatron [35][36]. Finally, the Higgs boson was first discovered by CMS and ATLAS at CERN in 2012 with the LHC. In addition to the successful discovery of all the particles in the standard model, extremely precise verification of its physical constants have been done. As an example, the gyromagnetic ratio of the muon has been calculated to within one part in ten billion, and agrees with measurements to better than one part in one billion [37].

2.2 Components of the Standard Model

Having followed the historical development of the standard model, let us now take a systematic view of its parts. The standard model of particle physics combines three of the four discovered fundamental forces in the universe: the electromagnetic force, the weak force, and the strong force. There has yet to be a successful quantum field theoretic description of gravity, so it is not included.

The standard model describes the fundamental particles and forces in nature with each

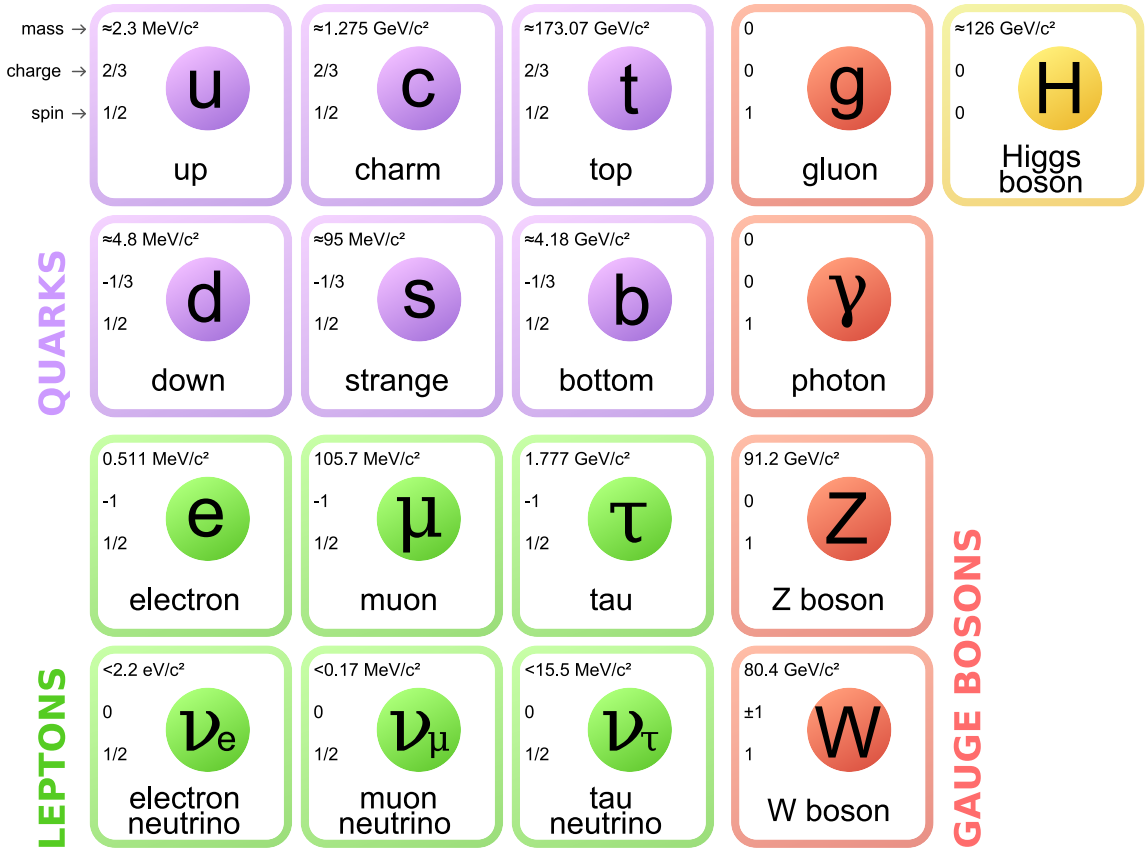


Figure 2.1: The particles of the Standard Model. The quarks are shown in purple, the leptons in green, the exchange particles in red, and the Higgs is shown in yellow. Each particle's name, mass, charge, and spin is also shown.

fundamental force corresponding to a preservation of certain quanta. Each fundamental force is associated with particles which carry the force in exchange between interacting matter, which are charged under the force in question. These interactions in the standard model can be described using Lie algebra, allowing the model to be described as combinations of groups.

A diagram of the standard model and each particle's properties is shown in 2.1. There are three kinds of particles in the standard model: quarks, leptons, and gauge bosons. Quarks and leptons make up all of matter and are fermions, while the gauge bosons define the fundamental force interactions and are bosons. There are three generations of quarks and leptons, each heavier than the last. The lightest generation is the up and down quarks and the electron (and electron neutrino). These particles make up all stable matter in the universe. Heavier generations have the same charges as these, though they all can decay

via the weak force into lighter generations.

The quarks interact with all of the forces described in the standard model. They have $\mp\frac{1}{3}$ or $\pm\frac{2}{3}$ electric charge and a spin of $\frac{1}{2}$. The leptons are the electron, muon, and tau, each with a corresponding neutrino. Each lepton has a charge of -1 and a spin of $\frac{1}{2}$. The neutrinos are not charged however.

At this point it is important to note the unexpected mass distribution in the standard model, which can be read from Fig. 2.1. In the first generation of fermions, the up quark and down quark have very similar mass. From there, the electron has roughly one-quarter of the mass of the down quark. The last particle in the first generation, the electron neutrino, however, has at most one-one-hundred-thousandth of the mass of the others. This significantly smaller mass is difficult to explain naturally and potential sources of the neutrino mass are explained further in section 2.3.1.

The five bosons of the standard model emerge from the quantum field theories that define it. The gluon solely interacts with quarks and gluons and is part of the $SU(3)_C$ group. The photon (massless) and Z boson (massive) are part of a superposition of weak-flavor neutral fermion-antifermion interactions in the electroweak sector, and the W boson comes from the axial vector, symmetry-breaking part of the electroweak sector. Its interactions with matter involve weak-flavor transitions, most commonly seen on earth as β-decay. The Higgs boson, from the Higgs field, gives the weak bosons, W and Z mass, and additionally provides a mechanism for fermion mass. These interactions will be discussed further in the following subsections.

2.2.1 The Strong Force

The strong nuclear force contained in the standard model describes a complicated three flavor interaction $SU(3)_C$, described by Quantum Chromodynamics (QCD). Each flavor is described as a "color" or "color charge": red, green, and blue. Each quark may have one of these colors; anti-quarks have anti-colors. The intermediating particle, the gluon, carries mixtures of color/anti-color. QCD requires the conservation of this color in all interactions. In addition to the conservation of color, free particles must be color neutral (white or black). This feature, called confinement, arises from the fact that as quarks are separated from each other, additional quarks can form from the potential energy in the strong field. This prevents any bare quarks or gluons being visible. Any quark or gluon with sufficient energy to travel away from an interaction will form more colored particles in between it and its past partner. This spray of hadrons is called a "jet". Color neutral quark combinations

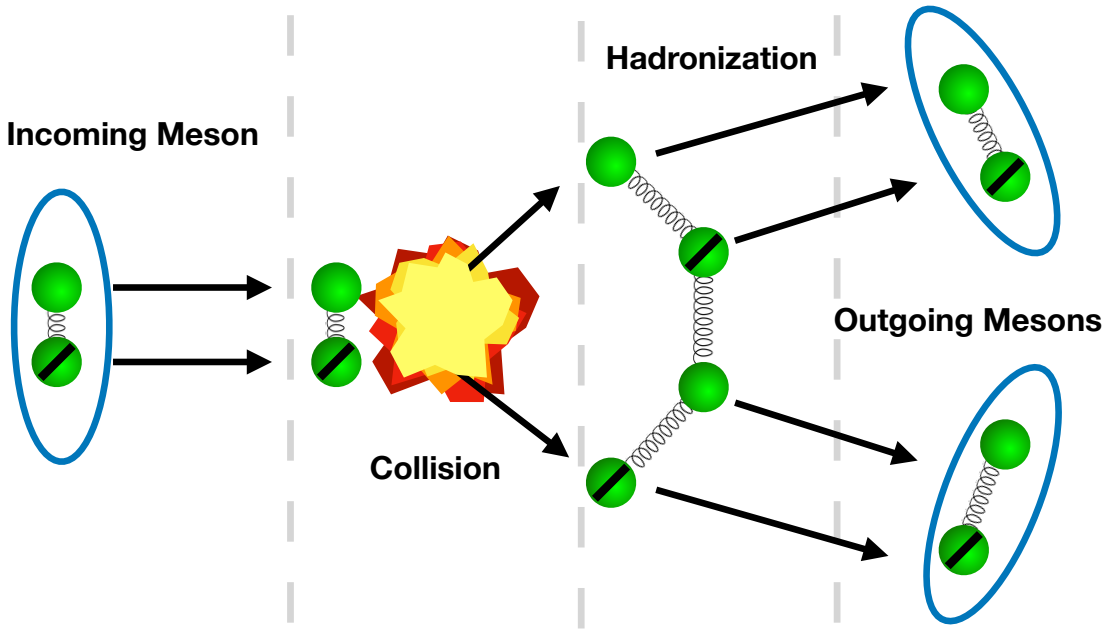


Figure 2.2: A simplified diagram of hadronization. In the collision, the quark and anti-quark of the meson acquire large kinetic energies relative to each other. As the meson's quarks begin to separate, the strong field strength between them increases. At some point, some of this energy is converted to make more quarks, splitting the meson in two. This process continues until the kinetic energy of the originally $q\bar{q}$ pair has been transferred to the rest mass and kinetic energy of multiple hadrons.

are generally mesons, which are color, anti-color quark pairs, or baryons which are three quark, red-green-blue combinations, all are called hadrons.

2.2.2 The Electroweak Interaction

The electroweak interaction comes from the combination of two groups $\mathrm{SU(2)_L} \times \mathrm{U(1)}$. This gives four vector bosons. Three are from the $\mathrm{SU(2)_L}$ group: A^i with $i \in \{1, 2, 3\}$. The last is from the U(1) group: B. The physical states observed, the W, Z and γ are linear combinations of A^i and B. These combinations define the W boson as:

$$W_\mu^\pm \equiv \frac{1}{\sqrt{2}} \left(A_\mu^1 \pm i A_\mu^2 \right). \tag{2.1}$$

The Z and photon combinations are:

$$Z_\mu \equiv \sin\theta_W B_\mu - \cos\theta_W A_\mu^3, \tag{2.2}$$

$$A_\mu \equiv \cos\theta_W B_\mu - \sin\theta_W A^3_\mu. \tag{2.3}$$

The $V-A$ composition of the standard model weak force is parity-violating, and removes the possibility of right-handed weak interactions. Thus, all neutrinos and all interactions with the W boson are only with left-handed particles. From the moment this theory was developed, physicists have worked at ways to restore parity symmetry with a right-handed weak interaction [38][39].

2.2.3 The Brout-Englert-Higgs Mechanism

The Higgs mechanism takes a central role in the standard model. The Higgs field couples to every particle with mass in the standard model. These interactions not only explain the mass of the electroweak bosons, which the Higgs mechanism is essential for, it also couples to fermions. The fermion mass term, in Weyl basis looks like this:

$$\mathcal{L}_D = -m\bar{f}f = -m\chi^\dagger_L\chi_R - -m\chi^\dagger_R\chi_L, \tag{2.4}$$

where χ_L and χ_R are left and right-handed component spinors of the fermion, f. This mass term is called the Dirac mass and is used for all the massive fermions in the standard model. Neutrinos do not have a mass term in the standard model, and as they are exclusively left-handed, a mass term cannot simply be "pencilled in". The evidence for neutrino mass is extensive [40][41] and requires new mass terms in the standard model. Left and right-handed neutrino masses are discussed further in section 2.3.1.

2.3 Left-Right Symmetric Extensions to the Standard Model

The left-handed nature of the weak force $\mathrm{SU(2)_L}$, and the left-right symmetry of every other group, leads to a desire to complete the symmetry of the standard model by adding a right-handed component to the weak force. Completing the standard model has additional functional benefit in allowing for neutrinos to have a very light mass in a natural manner. By the see-saw method, the currently observed left-handed neutrinos can be very light, countered by very heavy right-handed neutrinos. The symmetry would have to be broken at some energy level, allowing for the currently observed asymmetry at low energies.

Left-right symmetric extensions (LRS) models would require adding an additional symmetry group to the standard model or simply making the standard model a part of a larger

symmetry like SU(5) or O(10). To simplify discussing the theory, we'll focus on the simplest extension. Here, the electroweak sector simply gains another SU(2) group, make it $\mathrm{SU(2)_R} \times \mathrm{SU(2)_L} \times \mathrm{U(1)}$. Adding this extension creates three additional particles, the W_R and N_R, which are discussed in later in this chapter, and a partner to the Z, the Z'. We will assume that left and right-handed interactions have the same coupling strength. While this is not required in a more complicated LRS model, it provides a benchmark for the search. The right-handed extension additionally complicates the Higgs sector, requiring at least a Higgs doublet to couple to the right and left-handed weak bosons independently.

2.3.1 See-saw Mechanism

Left-right symmetric models provide a straightforward and natural mechanism for giving right and left-handed neutrinos mass. This mechanism is called the see-saw where one neutrino is very heavy and the other very light [42]. The see-saw Lagrangian combines Dirac and Majorana mass terms; Majorana mass terms only being possible for fermions without electric charge [43]:

$$\mathcal{L} = \frac{1}{2} \begin{pmatrix} \bar{\nu}_{Li} & \bar{\nu}_{Ri} \end{pmatrix} \begin{pmatrix} B_i' & M_i \\ M_i & B_i \end{pmatrix} \begin{pmatrix} \nu_{Li} \\ \nu_{Ri} \end{pmatrix}. \tag{2.5}$$

Each generation of neutrino is denoted by i. The terms ν_R and ν_L are pure left and right-handed spinors. The M is a Dirac mass component and gets its value from the vacuum expectation value of a LRS-Higgs field. The B' and B masses are Majorana components and receive their mass from the individual left and right-handed LRS-Higgs doublets. With negligible mixing between generations, the mass eigenvalues are:

$$\lambda_+ \simeq B,\ \lambda_- \simeq -\frac{M^2}{B}. \tag{2.6}$$

With $B \gg M$, which comes from that observation that $M_{W_R} \gg M_{W_L}$—if it exists. Now we can write the mass terms for the two observable neutrinos. First with the mass eigenvalue of λ_- we have the currently observed neutrinos:

$$\nu \simeq \frac{1}{\sqrt{M^2 + B^2}} \left(B\nu_L - M\nu_R \right) \simeq \nu_L - \frac{M}{B}\nu_R. \tag{2.7}$$

It can be seen in equation 2.7 that the right-handed spinor is a small component consistent with current observations. The heavy primarily right-handed neutrino state is then given

as:

$$N \simeq \frac{1}{\sqrt{M^2 + B^2}} \left(M\nu_L + B\nu_R\right) \simeq \nu_R + \frac{M}{B}\nu_L. \tag{2.8}$$

2.3.2 Experimental Study of LRS Models

While right-handed extensions to the weak force have been studied and sought for a long time, no specific experimental evidence exists for them. In addition, the models themselves have too many free parameters to make any exact prediction. Several different strategies can be employed. Direct searches, where a new particle is produced on-shell and its decay products are seen, can probe for any of the additional particles predicted in LRS models. Indirect searches, which look for subtle changes in known processes from additional or higher-order interactions involving a right-handed weak force, can also be done. The direct search for the W_R and N_R is the focus of this analysis. For a more detailed discussion of phenomenological constraints on the masses of right-handed weak force particles and their mixing see [44].

Direct Searches

Any LRS model will include a new neutral-current boson, the Z'. Perhaps the clearest signature of the existence of LRS interactions at the LHC would be the Z' production and decay to leptons, Z' $\to \ell\bar{\ell}$. The heaviest standard model di-lepton resonance corresponds to the Z. As possible Z' masses are much higher than the Z mass, this mass region would be relatively background free. However, in a LRS model largely following the standard-model left-handed interaction, the Z' will be heavier than the W_R. The Z' could be too heavy to produce at the LHC, while the W_R could be quite visible. Lower limits on the mass of the Z' have been set with this channel and are currently up to 4.5 TeV [45].

Direct searches for the W_R can be performed in a few different ways. One option would be focusing on the $W_R \to \ell\overline{\nu}$ decay. This search signature could be quite similar to a standard model W, assuming that the right-handed neutrino is relatively light and stable. However, right-handed neutrinos light enough to decay from a Z resonance would have played a role in the measurement of the well-constrained Z decay width. So, searches for this decay mode assume a heavy and decaying N_R, as this analysis does.

The hadronic decay of the W_R, which would mirror the production mechanism we rely on, $W_R \to \mathrm{q}\overline{\mathrm{q}'}$, can also be searched for. It has a high branching fraction of the W_R decay and has no sensitivity to N_R mass, which can, as will be seen in this analysis, change the W_R decay topology significantly. However, the $W_R \to \mathrm{q}\overline{\mathrm{q}'}$ channel has significant QCD

backgrounds, even in the highest masses reached at the LHC, presenting a challenge for the search. Selecting the quark flavor in the decay and searching for $W_R \to t\bar{b}$ can be done, and would reduce the amount of QCD. The top produced in the W_R decay can decay producing leptons which would give a better handle on the study than a generic jet search, though with a cost in branching ratio. Searches for heavy di-jet resonances have been performed at CMS with the most recent search using the same years as this analysis. These searches can exclude many different signal models, including the W_R. Masses of a di-jet resonance similar to the W_R were excluded up to 3.6 TeV [46]. A third W_R signature to search for relies on a possible beyond-the-standard-model behavior where some LRS models create a doubly charged W_R. This decay, which produces two leptons of the same sign, is very unique from the standard model and can be also be searched for. The most recent search was performed by the ATLAS collaboration, setting a limit at 4.7 TeV [47].

Indirect Searches

Direct searches provide an obvious way of finding new particles. However, any LRS extensions would also affect standard model processes and constants through additional virtual interactions with the new particles. There are many examples of this: exotic behaviors like neutrino-less double beta decay, left-right mixing, or even lepton-flavor violation could occur. The lack of observation for these behaviors, at present, would need to be explained by their having an exceptionally low rate, which can be achieved by the new particles having masses $\gtrapprox 2$ TeV. An LRS model could also affect currently understood behaviors, but to a degree too small to be yet measured. The electron electric dipole and neutron meson mixing and mass difference are examples of this. Measurements of all of these, or upper limits on possible rates of processes not in the standard model, provide additional limits which can constrain the search area for direct searches. The best interaction to study for an indirect search, or limit, on LRS models is with neutral mesons. The mixing of neutral meson states—and their mass—are affected by CP-violation.

Neutral Kaons are mesons, composites of two quarks, made of one strange and one down quark. One or the other will be an anti-particle, with the particle being the down and anti-strange combination. Physically-propagating neutral Kaon states are neither, but two different combinations of the Kaon and anti-Kaon. These are called K short and K long, based on the striking observation that one (the K long) has a much longer lifetime than the K short, and can be observed to travel in the lab frame. This mixing of K_0 and $\bar{K_0}$ comes from CP-violation. This mixing involves a diagram with the exchange of two

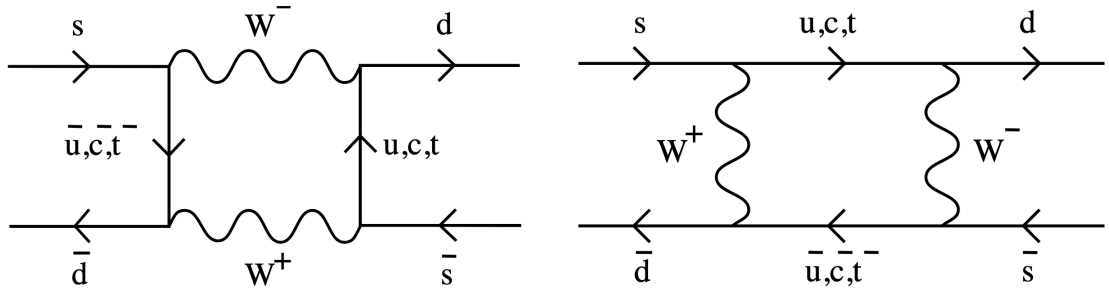


Figure 2.3: The two leading kaon mixing diagrams are shown here. As quarks occur naturally left and right-handed, a virtual W_R would be contribute identical diagrams to this process [48].

W bosons. If there is a W_R, this diagram exists not just for the standard model W but also the right-handed W boson, as quarks can be right and left-handed. This increases the mixing, but is also heavily dependent on the mass of the W_R. The heavier the W_R, the less significant its contribution. Precision measurements of Kaon mixing allow for a lower limit on the W_R mass to be set at 2.5 TeV [49].

2.4 New Particle Production at the LHC and PDF Effects

As the highest energy particle collider created to date, the LHC has the possibility to reach collision energy levels where new, previously unobserved particles could be created. Descriptions of the LHC and CMS are in Chapter 3. As this analysis is a direct search for particles beyond the standard model, its important to review the way in which the LHC generates high mass particles. An overview of the rates of different standard model interactions, and hypothesized beyond the stand model interactions are shown in Fig. 2.4. The cross-sections searched in this analysis vary, but in the upper-most mass limit are on the order of a femtobarn, which as can be seen in the figure, occur no more than several times a year.

2.4.1 Parton Density Functions and Proton-proton Collisions

Collisions at the LHC take place between protons. Each proton carries with it three quarks and a large amount of QCD potential energy, which manifests as a combination of gluons and short-lived quark and anti-quak pairs of all flavors. Actual interactions occur between

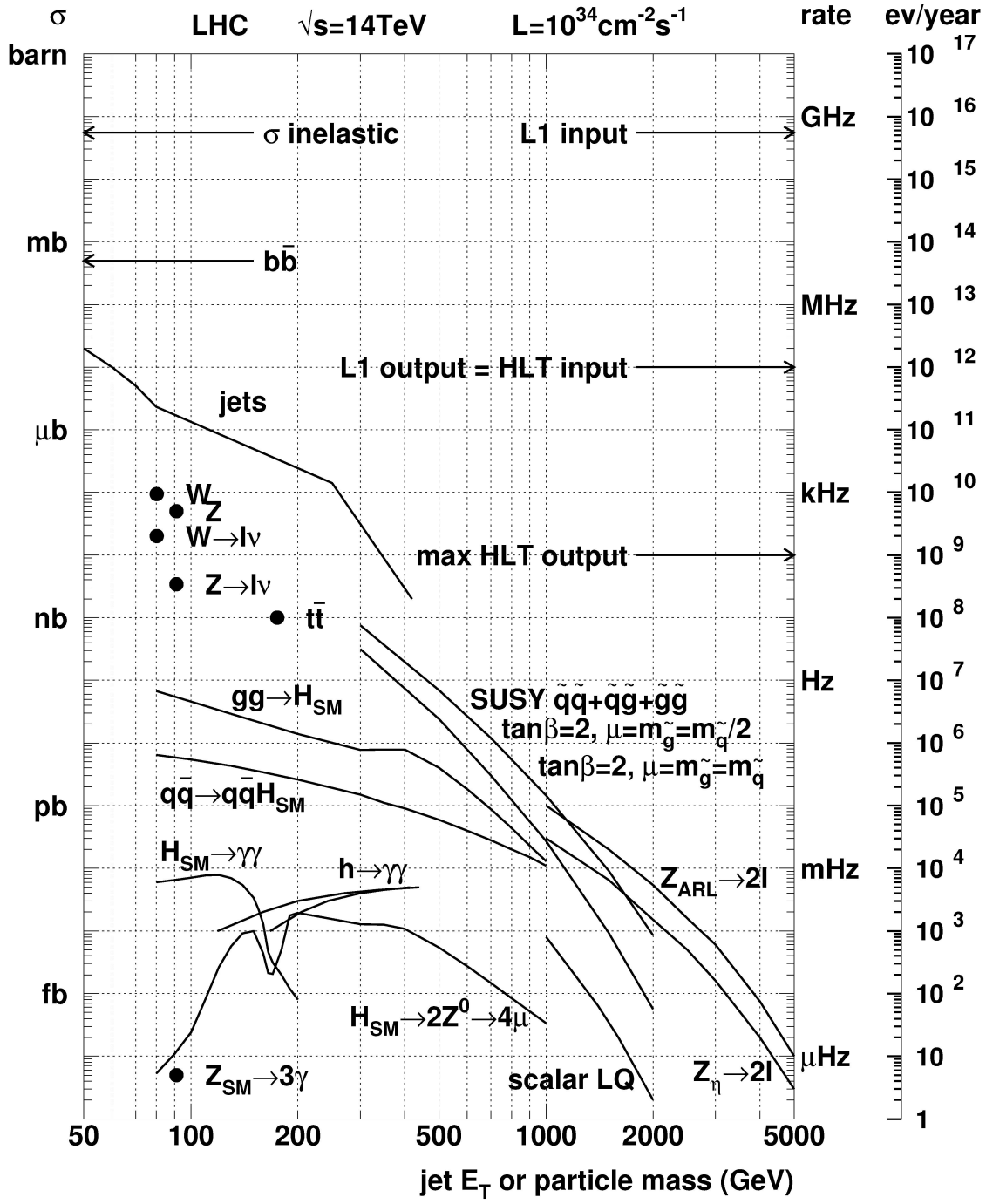


Figure 2.4: Interaction rates for various standard model and theorized beyond the standard model processes are shown. From top to bottom interactions occur at lower and lower rates until, below the inverse-femtobarn level, they would not be expected to occur at the LHC [50].

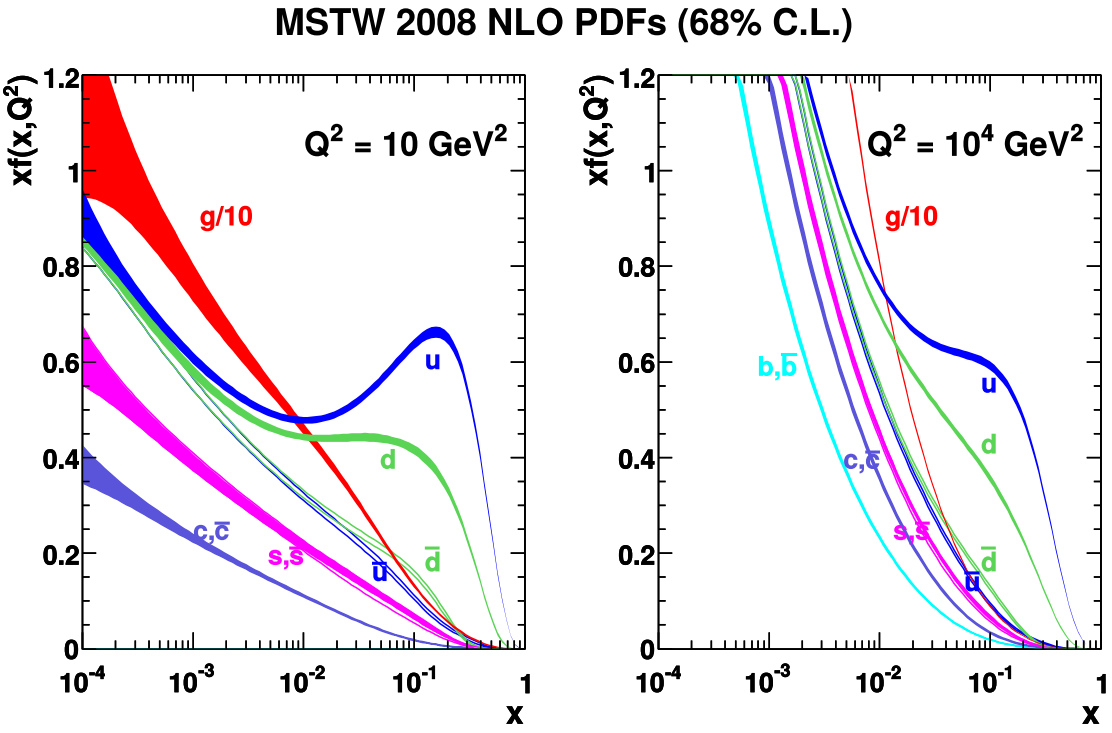


Figure 2.5: Two sets of parton density functions (PDFs) are presented here. A collision at lower energies is shown on the left, and a higher energy collision, more relevant to W_R production is shown on the right. While the variables xf, x, and Q^2 are beyond the scope of this chapter, the y-axis can be interpreted as the rough probability of interaction, and the x-axis the momentum fraction exchanged in collision. At high momentum fractions (toward the left and as required by a heavy W_R) the shell quarks of a protons are seen to dominate. At lower momentum fractions (as would occur with standard model particle production at the LHC) gluons dominate [51]

.

these parts of the proton, called "partons". In theory, any quark or gluon can interact in the proton-proton collision. The amount of momentum carried by an interacting parton varies according the amount of momentum exchanged in the collision, always less than the total collision energy of $\sqrt{s} = 13\,\text{TeV}$. Thus, any particle produced at the LHC must have a mass substantially smaller than 13 TeV. Figure 2.5 shows the relationship between parton interaction probabilities and momentum exchanged in collision for varying collision energies. The extremely heavy W_R would mostly be produced in a collision between a valence quark, and a sea anti-quark. However, many of the most frequent standard model boson production modes involve gluons at LHC energies.

In every high energy particle collision, the energy of the collision available to produce new particle mass is defined as:

$$\sqrt{s} = \sqrt{\left(\sum_{i=1}^{2} E_i\right)^2 - \left(\sum_{i=1}^{2} p_i\right)^2}. \tag{2.9}$$

All particle creation at the LHC must obey the relation in Eq. 2.9, however, particles lighter than this relationship receive an additional bonus from having additional momentum phase-space to be produced in. Extremely heavy particles, like the W_R will generally be produced close to at rest in the lab frame as a result.

As a result of the parton momentum exchange probabilities, this analysis only searches for particles up to a mass of 7 TeV, and is only successful in setting a limit at less than 6 TeV. Heavier particles (even particles with a mass higher than 13 TeV) can be searched for at the LHC, but they will be increasingly virtual in interaction at the LHC as their resonant mass increases. These are very challenging to search for, as they present as a broad excess, rather than a peaking, resonant, mass distribution.

W_R–N_R mass effects at the LHC

While this analysis can be colloquially referred to as a "bump hunt", the shape of the W_R mass spectrum in our selection can stretch beyond a typical resonant peak. This analysis exclusively seeks W_R events which decay to on-shell right-handed neutrinos. This necessarily truncates the W_R mass spectrum to be produced above the N_R mass generated. In the lower N_R mass boosted regime, the W_R mass spectrum is less and less truncated, opening up the possibility of low-mass offshell W_R events to be produced. These events are enhanced by the steeply falling probability of collisions at higher mass. The full shape of the generated W_R mass is shown in Fig. 2.6 at one of the W_R–N_R mass points where the off-shell contribution is most notable.

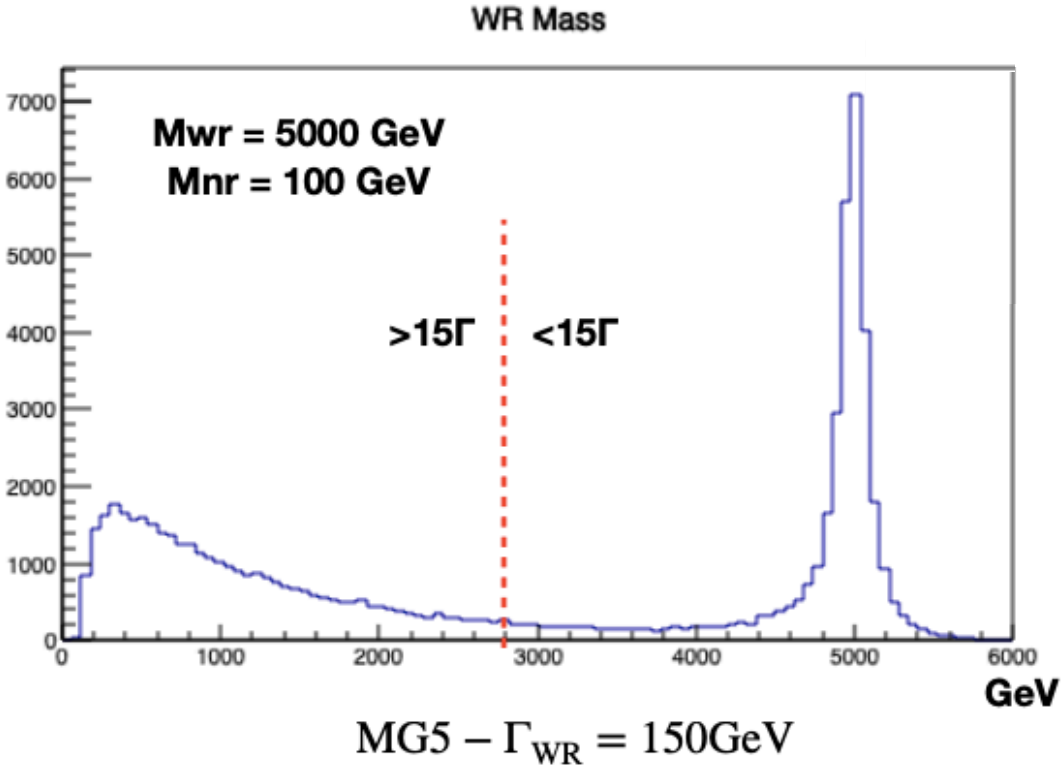
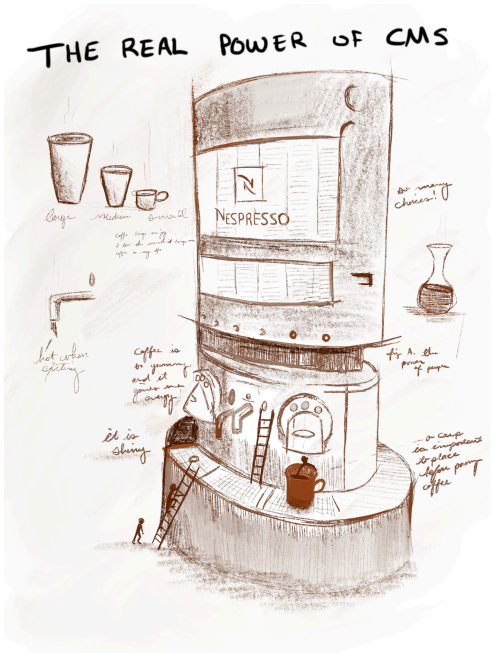


Figure 2.6: The generated mass spectrum for a W_R–N_R mass of 5000-100. The on-shell/off-shell cutoff configured in the Monte-Carlo generator is shown as well and corresponds to 15 times the natural width of the W_R. It can be seen that nearly half of the generated W_R events fall below the cutoff.

.

Chapter 3

The CMS Experiment



This chapter details the CMS experiment as well as the "crown jewel" of the CERN accelerator complex, the Large Hadron Collider (LHC), which is used by CMS. Specific attention is paid to the components of CMS and the LHC most relevant to the W_R search. The LHC is the largest and highest-energy particle collider ever created. It is a 27 kilometer circumference synchrotron accelerator with two beam pipes each designed to carry approximately 2700 bunches of protons which can be brought into collision at four interaction points around the ring. The four experiments around the ring, clockwise looking down, are ALICE, CMS, LHCb, and ATLAS. It was first operated with $\sqrt{s} = 7\,\text{TeV}$, then

$\sqrt{s} = 8\,\text{TeV}$, and now $\sqrt{s} = 13\,\text{TeV}$.

3.1 The Large Hadron Collider

The LHC was constructed outside Geneva, Switzerland. It was first operated in the Fall of 2008, and after an electrical failure in the magnet power system, was repaired and recommenced operation in 2010. The LHC is part of an accelerator system as illustrated in Fig. 3.1. Each accelerator adds a portion of energy to the beam, progressively increasing to the LHC injection energy. The start of the LHC proton beam is a humble gas bottle of hydrogen. From there, the gas is ionized and accelerated with a linear accelerator (LINAC 2) which accelerates protons directly into the Proton Synchrotron (PS). The PS accelerates the protons to an energy of 26 GeV before injecting them into the Super Proton Synchrotron (SPS). The SPS accelerates proton bunches up to 450 GeV. Once proton bunches have reached this energy, they can be accelerated into the LHC. These bunches are then accelerated to 6500 GeV. The LHC was originally designed for operation at up to $\sqrt{s} = 14\,\text{TeV}$ but has not done so yet. In the last data taking period, which concerns this analysis, the LHC ran at $\sqrt{s} = 13\,\text{TeV}$.

Protons are accelerated in each of the synchrotrons with radio-frequency (RF) cavities. The RF cavities fill the beam path with standing electro-magnetic waves. Protons passing through the standing waves are accelerated and kept in bunches. Slow (and trailing) protons are accelerated more than faster (and leading) protons in each bunch. This ensures that each bunch of protons is well-formed in momentum and physical space. At the LHC, there are 8 RF cavities for each beam and they are operated at 400 MHz. Bunches collide at intervals of 25 ns. This is part of the design of the RF cavities, as a bunch is accelerated every 10 cycles of the radio frequency. This gives a maximum number of bunches at the LHC of roughly 3600. Each step in the synchrotron acceleration chain requires additional clear space so that injection (and dumping) magnets can be brought to full energy in between bunches. This gives an actual total of 2808 bunches.

While the LHC holds the record for the highest center-of-mass energy-per-nucleon of any hadron collider built to date, allowing it to probe previously unreached mass regions. It is also the most luminous hadron collider. The record instantaneous luminosity and a high-duty cycle has allowed the CMS experiment to record an unprecedented number of collisions. The luminosity, shown in Eq. 3.1, is proportional to a number of factors:

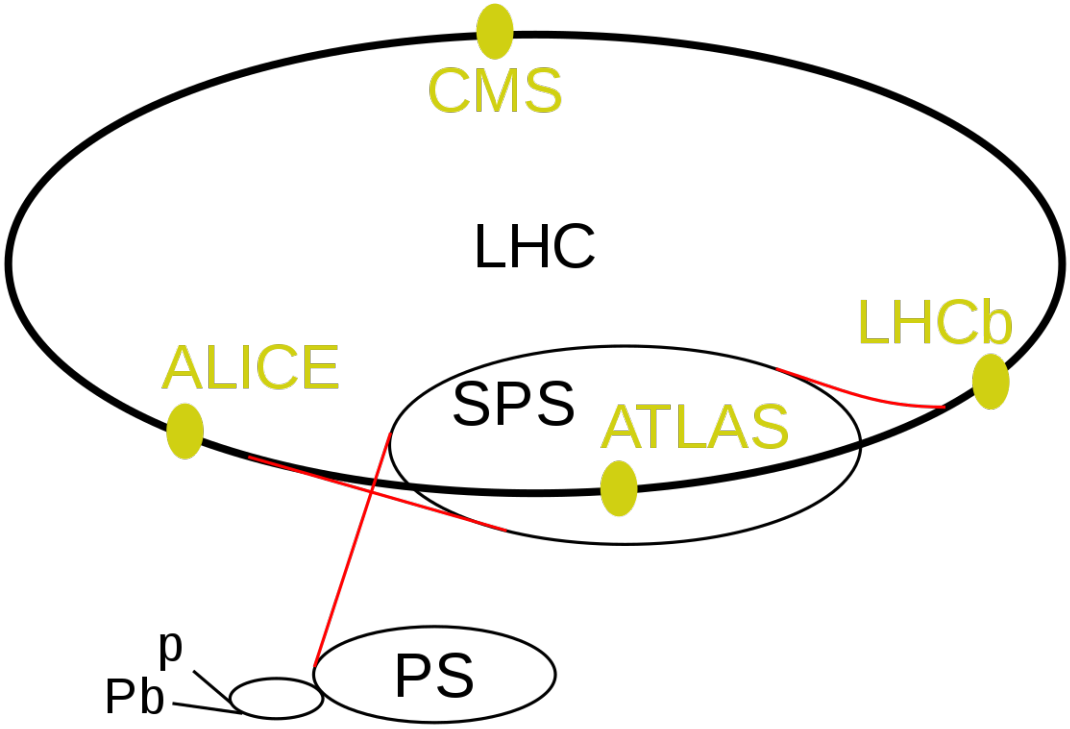


Figure 3.1: This diagram shows the accelerator complex feeding into the LHC. Starting with a linear accelerator, the protons are accelerated into the Proton Synchrotron (PS) and then Super Proton Synchrotron (SPS) before injecting into the LHC.

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

Orbital frequency, f, is the frequency a given bunch interacts (roughly 11 kHz), n is the number of bunches of protons in each beam, N is the number of protons in each bunch, γ is the Lorentz boost of the protons, ϵ_n is the beam emittance, β^* is the beta function, and F is based on the shape of the beam crossing. Emittance represents the total "size" of the beam in momentum-position phase-space, and the beta function describes the focusing of the beams as they cross by high-powered quadrupole magnets. Multiplying them together gives the cross-section area of the beams at the crossing point. At the LHC the number of protons per bunch, the number of bunches and the beam cross-section at collision can be changed. It can be seen from Eq. 3.1 how these parameters affect the luminosity. Increasing the number of bunches and the protons per bunch increases luminosity. Likewise, focusing the beams to a small cross-section increases luminosity.

Each of these variables have their limits and challenges associated with changing them. The total number of bunches the LHC can hold is predefined by the frequency of the RF cavities as well as the need for injection and ejection of the beams in the accelerator chain. The number of protons per bunch is limited for detector safety. Higher proton density bunches are more difficult to control and can become unstable, potentially damaging the accelerator. Higher proton density in bunches also leads to an increase in additional proton interactions in a given crossing. Increasing proton density is the fastest way to boost instantaneous luminosity. This poses a challenge to the detectors to sufficiently determine the properties of the interaction of interest and increases the rate of radiation damage in the detector.

The first year of operation saw $44.96\,\text{pb}^{-1}$ delivered luminosity at $\sqrt{s} = 7\,\text{TeV}$. The accelerator continued operation at this energy for 2011 delivering $6.10\,\text{fb}^{-1}$. In 2012 the energy increased to $\sqrt{s} = 8\,\text{TeV}$ and $23.3\,\text{fb}^{-1}$ were delivered. After a longer shutdown for three years and significant upgrade work, the LHC resumed operation at $\sqrt{s} = 13\,\text{TeV}$ and ran for 2015, 2016, 2017, 2018. A total of $162.85\,\text{fb}^{-1}$ was delivered over those four years. During Run II, 2015-2018, the LHC was operated at its tightest bunch spacing of 25 ns. The first year of Run II (2015) was considered a commissioning year and collected significantly less data and is not considered in this analysis. From 2016-2018 the LHC continued to outperform itself, increasing the number of protons per bunch beyond design specifications and ultimately delivering $162.85\,\text{fb}^{-1}$ collisions. Some of the relevant parameters for the three years covered in this analysis are summarized in Table 3.1. A chart of the delivered

Parameters	2016	2017	2018
Beam Energy	6.5 TeV	6.5 TeV	6.5 TeV
Bunch Spacing	25 ns	25 ns	25 ns
Number of Bunches	2808	2808	2808
Peak Luminosity Per Day	15.3 Hz/nb	20.7 Hz/nb	21.4 Hz/nb
Mean Number of Events per Crossing	27	38	37
Delivered Integrated Luminosity	40.99 fb^{-1}	49.79 fb^{-1}	67.86 fb^{-1}

Table 3.1: Operation parameters of the LHC during Run II, 2016-2018.

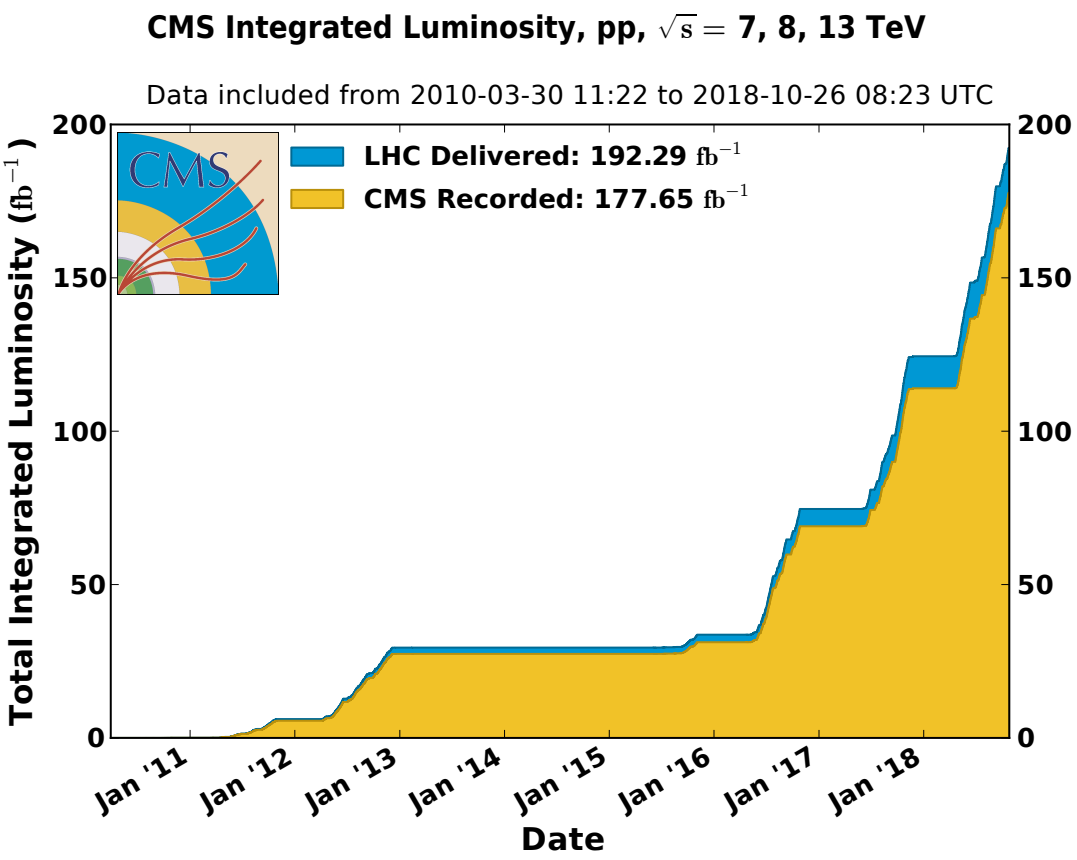


Figure 3.2: This graph shows the integrated luminosity of the LHC over all the years of operation. It can be seen that recent years have seen a dramatic increase in instantaneous luminosity, with the last year, 2018, contributing approximately one third of the total.

luminosity of the LHC from the beginning is shown in Fig. 3.2.

3.2 The Compact Muon Solenoid

The Compact Muon Solenoid (CMS) at CERN is one of the general purpose detectors at the LHC and the one used for this analysis. Each part of the detector is built around the previous moving outward from the interaction point in a cylindrical shape. Through the long axis of CMS runs a single beam-pipe containing both proton beams which are brought into collision in the very center. CMS is designed to accept as many events as possible and with as much detail as possible. It covers a large solid angle to allow for as many of the post-collision particles as possible to encounter an active part of the detector. CMS also has extremely fast electronic readout to allow it to record events as often as possible and as accurately as possible.

The CMS detector has several major detector components, sub-detectors, as well as a very large solenoid magnet. A drawing of CMS and its different components is shown in [52]. The support for these components is integrated as part of the iron magnet return yoke. Starting at the interaction point and moving outward transversely, the CMS detector is made up of a silicon-tracker, an electromagnetic calorimeter (ECAL), a hadronic calorimeter (HCAL), the magnet, and then the muon chambers with the iron magnet return yoke interspersed. Each of these parts will be discussed further shortly. Additionally the detector changes on the two ends of the cylinder, called endcaps, this is to account for the differing particle fluence at different collision angles.

As the CMS detector is discussed, it is first helpful to describe the coordinate system used. This coordinate system is standard for hadron collider experiments. The Cartesian coordinates are defined as follows: x points to the center of the collider, the y points up, and the z coordinate points in the counter-clockwise, looking down, direction along the beam. This creates a right-handed coordinate system. The origin of these coordinates lies at the center of collisions within the detector. An angular coordinate system is generally used to better reflect the geometry of the particle collisions and the detector itself. This coordinate system is defined with η, ϕ. Here, ϕ is simply the azimuthal angle in the $x-y$ plane and η is defined as pseudo-rapidity:

$$\eta \equiv -\ln \tan\left(\frac{\theta}{2}\right). \tag{3.2}$$

In this equation, θ is the the polar angle along the z axis.

Pseudo-rapidity is used as a simpler alternative to particle rapidity defined with the

particle momentum along the beam direction and its energy:

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

When a particle is massless, the Y simplifies to η, which is a good approximation for most decay particles, considering the energies of particles at the LHC.

At LHC energies, proton collisions actually involve the collision of individual components of the protons, quarks and gluons. The momentum of these partons is not known during the collision. The result is that each collision has a different amount of boost along the z axis. However, the rapidity of a particle does not change when the particle is boosted along the z axis. As such, defining a coordinate system with η (which approximates Y) allows for post-collision particle directions to be discussed in a way roughly invariant of their initial momentum in the z direction. Given the large inelastic cross-section at the LHC, this also guarantees that any given sweep of $\Delta\eta$ will have roughly the same particle fluence.

As a result of the unknown longitudinal momentum, a particle's momentum and energy is generally only considered in the transverse direction. The transverse components are calculated from η. The transverse momentum and energy of a particle is defined as:

$$p_T \equiv |p| / \cosh \eta, \tag{3.4}$$

$$E_T \equiv E / \cosh \eta. \tag{3.5}$$

Energies and momenta determined from the detector can then be re-expressed independent of the particle's longitudinal boost. The η and ϕ are also used in conjunction to calculate particle's separation independent of their longitudinal momentum as well. This is defined as ΔR:

$$\Delta R \equiv \sqrt{\Delta\eta^2 + \Delta\phi^2}. \tag{3.6}$$

3.2.1 Particle Flow

At CMS every candidate particle is reconstructed using as many subdetectors as may contain information about the particle. A diagram of what this process looks like is shown in Fig. 3.3.

This analysis uses jets and muons to find candidate W_R events. Each of these requires several subsystems to be well-measured. For muons, the combination of hits in the tracker

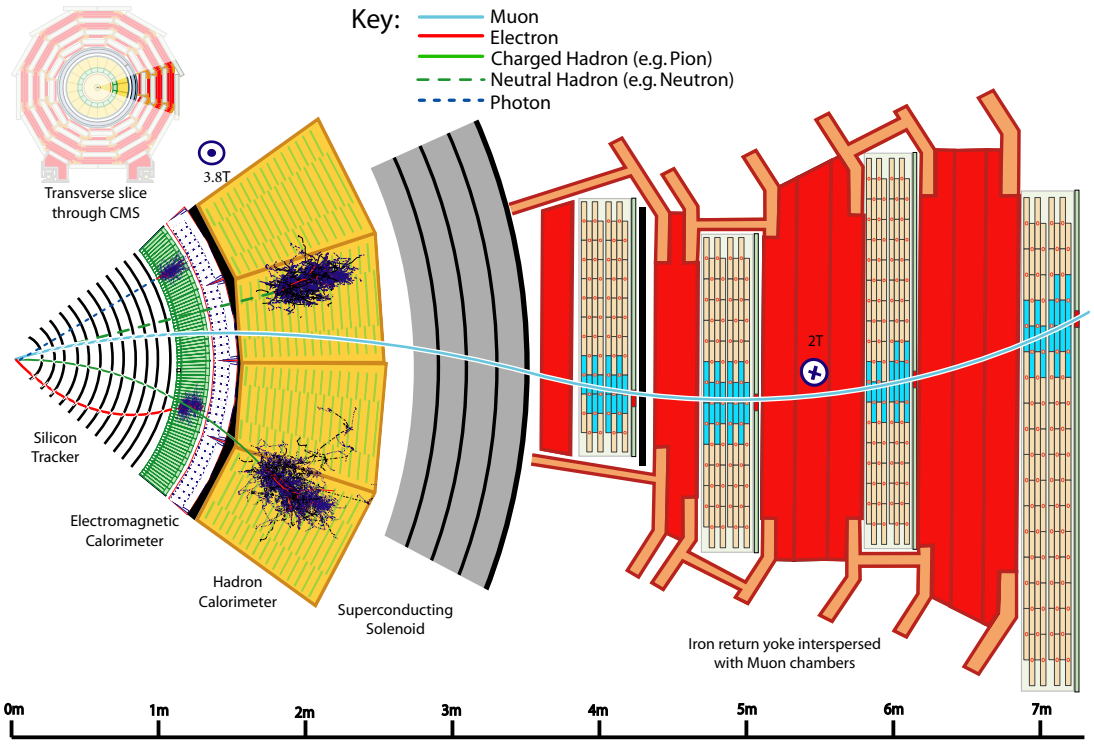


Figure 3.3: Particle flow for some commonly observed particles in CMS. The path of each particle and its detector response are shown diagrammatically here. By tying together information about each particle from the different subdetectors it travels through, an accurate calculation of its properties can be made.

and the muon chambers is used to reconstruct the particle momentum. For jets, charged particles in the jets leave tracks, and all its components deposit energy in the ECAL and the HCAL. The combination of energies deposited in the two calorimeters are combined to estimate the jet energy.

With several proton-proton collisions occurring at the same moment, great care is taken to ensure that the components of jets all come from the same source. The tracker is used to identify which specific location each jet component comes from in the collision, called a vertex. The PUPPI algorithm is used to prune jets based on this. PUPPI is discussed further in Section 6.1.1. Muons used in the analysis are also required to originate from the same vertex.

3.2.2 Tracker

The tracker is at the center of CMS and enables the reconstruction of charged particles passing through the detector from the collision point. All particles travelling into CMS with a $|\eta| < 2.5$ from the interaction point pass through the tracker. There are two main parts of the tracker: the strip tracker and the pixel detector, each made with silicon. The strip tracker is used primarily for measuring momentum, and the pixel detector for position. Silicon detectors operate by collecting ionized charge in material. Each active layer in the detector gives an interaction location for the particle, which is related to the particle's momentum and charge.

As the pixel tracker is closest to the beam line of all sub-detectors, just 4.4 cm from the interaction point, it must also have the highest position resolution. Silicon was chosen to satisfy the stringent requirements of the detector. It has a very fast refresh time (faster than the time between interactions of 25 ns), is radiation sufficiently hard and pixels can be made small enough for the needed resolution.

The strip tracker sits around and outside the pixel tracker. The strip tracker allows for the determination of charged particles' momenta, by measuring their track curvature. With the high magnetic field of CMS and a longer distance from the interaction point, the measurement of momentum is improved over just the pixel detector. While a larger pixel detector could be built in the place of the strip tracker, this was not practical. Each strip runs the length of a tracker module and is between 80 and 180 μm wide depending on its distance from the interaction. A cross-sectional view of the strip tracker can be seen in 3.4.

The precision of the tracker is ultimately key for identifying the primary interaction vertex in the event. This allows for the rejection of particles stemming from an uninteresting

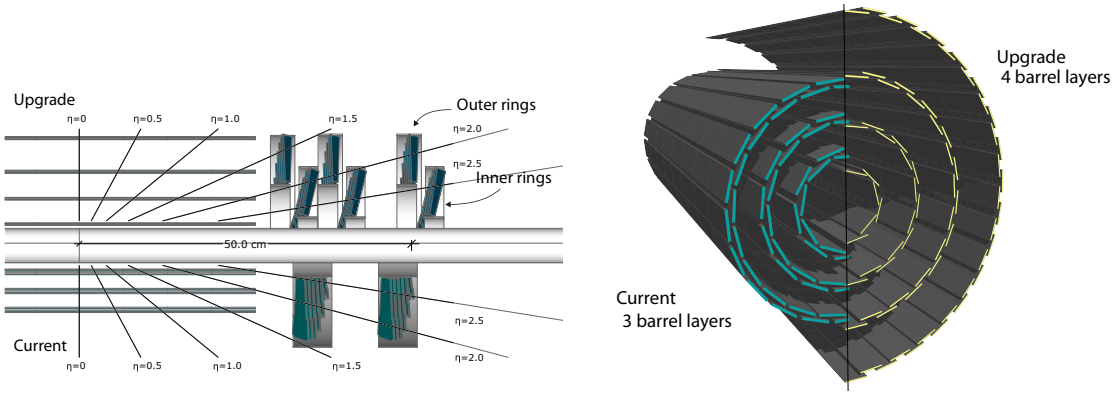


Figure 3.4: A cross-section of a part of the CMS tracker is shown. On the left, a conceptual layout comparing the different layers and disks in the current and upgrade pixel detectors is shown. On the right, a transverse-oblique view comparing the pixel barrel layers in the two detectors can be seen. [53]

vertex (pileup). In the transverse plane its resolution is on the order of 100 μm and longitudinally, 150 μm. The resolution of a muon with 100 GeV of momentum is approximately 1% in the barrel region and $3-6\%$ in the endcap region. For more information about the tracker see [53].

3.2.3 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) absorbs and measures the energy of charged particles passing through it. Electrons and photons are typically completely stopped in the ECAL, while quark-matter particles will leave some energy and continue through to the HCAL. The ECAL is a total-absorption calorimeter made out lead-tungstate crystals. Each crystal absorbs the energy of particles moving through by electromagnetic interactions, a fraction of the energy is then converted to scintillation light. The amount of light produced is then measured.

The ECAL is divided into three sections covering all of ϕ and out to an η of 3.0: the barrel (EB), end-cap (EE) and pre-shower sections (ES). The barrel and end-cap have the same design, but with a different shape of the detector, though the crystals all still point near the interaction point at the center of CMS. The ECAL has a total of 61200 crystals in the barrel and 7324 crystals in the endcap.

Each crystal of lead-tungstate in the ECAL in the barrel region is $2.2\times2.2\,\mathrm{cm}^2$ and 23 cm long. In the endcap, they are slightly wider and shorter at $2.86\times2.86\,\mathrm{cm}^2$ and 22 cm long. These crystals are arranged in 5x5 groups arranged to be rectangular in $\eta-\phi$. The long

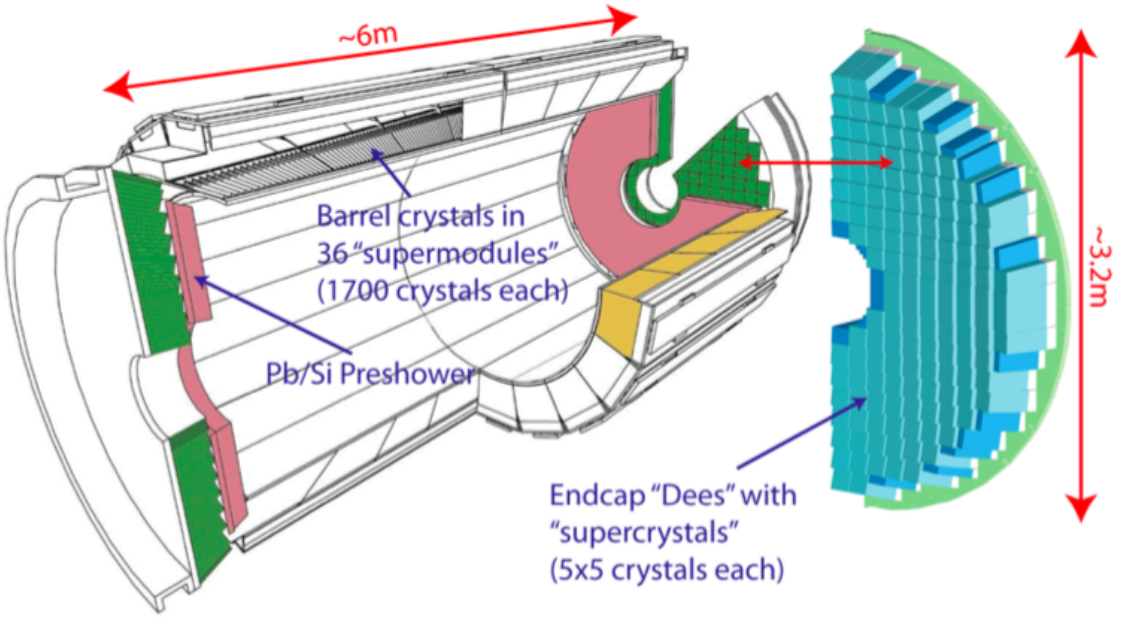


Figure 3.5: This cut-away of the ECAL shows the barrel and endcap regions, as well as the pre-shower. Individual crystals and their alignment towards the nominal interaction point can be seen. [54]

axis of the crystals are misaligned by 5 ° from the center of the detector. This reduces the chance that a particle travelling from the interaction point proceeds along the gap between crystals.

The ECAL is the first usage of lead-tungstate, $PbWO_4$, in a collider calorimeter. Lead-tungstate (as crystallized for CMS) is clear, and its high density gives it a radiation length of just $X_0 = 0.89\,\text{cm}$ and a Moliere radius of 2.2 cm. The Moliere radius is proportional to the volume over which the shower of energy spreads. The short radiation length of the crystal allows the ECAL to be smaller, and the small Moliere radius allows for higher spatial resolution in the detector.

The disadvantage of lead-tungstate is its low light yield with only ~ 30 photons/MeV. High-gain photo-sensors are used to amplify this signal. The barrel region uses avalanche photo-diodes (APDs) and the endcaps use vacuum photo-triodes (VPTs). Each has a high gain, while remaining relatively insensitive to the high magnetic field environment they operate in. Signals coming from each crystal are digitized and stored in electronics on the detector, and are only readout on request by the CMS trigger system. This reduces the data rate for ECAL, as sending all of the information for all crystals each collision was not

be feasible in 2006, when the detector was constructed.

The best performance of the ECAL requires near-constant calibration. As with all detector components of CMS, lead-tungstate is radiation resistant. However, the crystal structure is slowly damaged by radiation. Radiation creates imperfections in the crystal, reducing the light produced from an interaction. To maintain the highest possible precision, the ECAL is continuously recalibrated with a built-in laser system, which can operate during the abort-gap of the LHC while collisions are underway. The ECAL was first calibrated with dedicated beams and radioactive sources prior to being installed in CMS. With a known source, the whole detector was calibrated such that there is no asymmetry in the response based on $\eta-\phi$. The absolute calibration of the detector is also done with physics collisions in CMS using the Z boson. Its mass is very well understood from independent measurements and the signal is visible strongly over background. The resolution is a function of energy and is parameterized as:

$$\frac{\sigma_E}{E} = \frac{2.8\,\%}{\sqrt{E/\,(1\,\mathrm{GeV})}} \oplus \frac{0.128\,\mathrm{GeV}}{E} \oplus 0.3\,\%. \tag{3.7}$$

More details on the performance of the ECAL can be found in [54].

3.2.4 Hadron Calorimeter

The hadron calorimeter (HCAL) sits outside the ECAL and samples the energy from hadrons produced in collisions. The HCAL also stops most of the remaining particles, muons and neutrinos being the most common exceptions, and provides energy measurements of neutral hadrons, commonly neutrons and K_L particles, produced in collisions. The HCAL sub-detector also consumed two and a half years of the author's efforts and receives special focus in this section and section 3.3.

There are four sub detectors in the HCAL: HCAL Barrel (HB), HCAL Endcap (HE), HCAL Forward (HF), HCAL Outer (HO). These correspond to the barrel, endcap, forward and outer regions. HB, HE and HO follow the same overall design. Each is made out of layers of brass and plastic scintillator. The brass is used to cause hadrons passing through to shower and it absorbs some of their energy. The energy of these showers is then measured in the subsequent plastic scintillator layer. Showers often pass through several of the HCAL layers before being totally absorbed. The energy in each of these depths can then be summed for a total energy measurement, or compared to give more information on the shower evolution. A cross-section of the HCAL showing the regions HB, HE, and HO cover

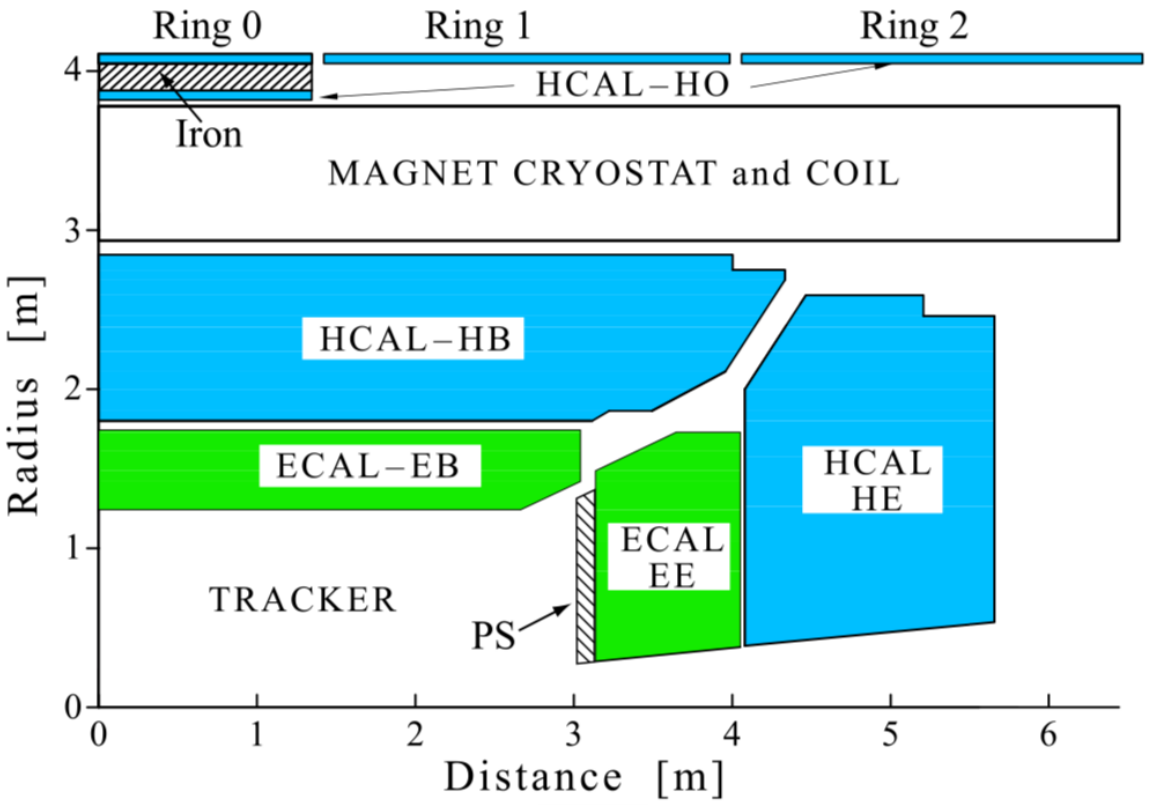


Figure 3.6: This cut-away of the HCAL shows the barrel and endcap regions, and it's arrangement with adjacent CMS components. The barrel and endcap regions are within the magnet. The gap between them allows for cabling to the inner detector. HF is not pictured as it sits separately in a higher η. [55]

is shown in Fig. 3.6.

HF is the most forward part of the HCAL and the most forward detector used in event reconstruction. It covers $|\eta|$ of up to 5.0. As a result of the significantly higher particle energy density in the forward regions, HF has a special, exceptionally radiation hard design. The bulk of HF is comprised of steel, with quartz fibers running horizontally through it. These quartz fibers measure the number of shower particles passing through them by the direct Cerenkov radiation produced in them. Each channel in HF comprises two bundles of such fibers, one short and one long. The long fibers run the length of HF, whereas the short fibers end farther from the inward side. The difference between the energy deposited in the fibers can be used to determine how deeply the transiting particle showered, which is related to whether the particle is a photon or electron as opposed to a hadron, which will shower deeper into the detector.

The HCAL has been specifically designed to measure the energy of hadrons. The hadrons detected by the HCAL are mostly charged pions at higher jet energies, and mostly neutrons and protons at lower jet energies, which are light quark anti-quark pairs. Neutral pions decay quickly to photons, which are well absorbed in the ECAL. Charged pions, however, decay more slowly, via the weak-force, or can be captured and absorbed by a nucleus. These, as well as neutrons and protons, pass through the ECAL. Any electrically charged hadron will interact with the tracker and the ECAL, leaving only some energy, as they are much heavier than electrons, with the same magnitude of charge. Hadrons will, however, interact via the strong force with nuclei in the detector. This happens in the ECAL as well as the HCAL. Each interaction can create more hadrons with energy sufficient to continue travelling through the detector. The energy of a hadron initially interacting with the detector is then distributed throughout the ECAL and the HCAL and requires a more careful reconstruction.

The specific ratio of neutral and charged pions involved in a shower is inherently random. Each shower will contain different amounts. Since each shower is started by only tens of hadrons, these fluctuations can be significant. The calorimeters do not have the same response to the electromagnetic interactions from neutral pion decays and the nuclear interactions of charged pions. This results in a degraded resolution of the HCAL from what might be theoretically expected. The behavior of the detector response with respect to different hadrons, and the distributions of them in a shower have to be understood to make accurate measurements.

Signal amplification from the plastic scintillators was originally done with hybrid photodiodes. These are in the process of being replaced, however, with silicon photo-multipliers (SiPMs). SiPMs offer improved photodetector efficiency and lower effective noise, improving detector performance despite radiation degradation of the scintillator. The forward part of the HCAL, HF, is much more removed from the magnetic field environment and continues to use photo-multiplier tubes.

Before installation, parts of the HCAL were fully assembled along with the ECAL and tested with pure electron and pion beams. These tightly controlled beams allowed for the π/e correction to be derived and the detector response calibration. The original resolution of the CMS ECAL and HCAL combined for hadronic particles is:

$$\frac{\sigma_E}{E} = \frac{84.7\,\%}{\sqrt{E/\,(1\,\mathrm{GeV})}} \oplus 7.4\,\%. \tag{3.8}$$

As can be seen in Eq. 3.8, hadronic particle energy reconstruction resolution has two significant components. On the left, the stochastic term, relating to shower-particle counting-statistics is very large until high energies. The right hand term represents the calibration uncertainty in the detector. It can be seen that hadrons do not have near the resolution as electrons and photons, particularly at lower energies. As with all sections of CMS, radiation damage is problematic and requires continual study. The HCAL does not require the same level of calibration as the ECAL, but still uses a laser system, operating within and outside of physics data taking to monitor detector performance.

3.2.5 Muon Chambers

The muon chambers detect muons passing through and out of the detector. The muon detectors, in conjunction with the tracker, allow for the momentum of the muons to be measured very accurately. As the radius of curvature of muon tracks is proportional to their momentum, with sufficiently low momentum muons (generally $p_T < 200\,\text{GeV}$) the tracker is sufficient for determining the momentum. However, the straighter tracks of higher momentum muons benefit greatly from the additional information from the muon system. Sitting outside the CMS magnet, all particles, other than neutrinos, are absorbed prior to reaching the muon detectors. By connecting tracks with the muon detectors, muon identification can be performed in a straightforward manner.

The muon chambers are further from the interaction point than any other endcap or barrel detector. They cover an area of roughly $25,000\ \text{m}^2$. Depending on the specific region of the detector, one of three technologies is chosen. In the barrel section, the detectors are drift tubes (DT) and resistive plate chambers (RPC). The RPCs are used additionally in the endcap along with cathode strip chambers (CSC). CSCs continue to the highest rapidity region of the muon detector. A diagram of the layout of the muon chambers is shown in Fig. 3.7.

All of the muon chambers are a type of gas ionization detector. Muons passing through gas ionize molecules, these ions are collected by an anode wire running down the center of the tube, the body of which functions as a cathode. The timing and shape of the current in anode as it collects charge allows for a determination of the position of the muon as it passed through. The barrel part of the muon system contains four layers of DTs. These layers are separated by the iron return yoke of the magnet and each contains either 12 or 18 layers. The alignment of the tubes in the first of the three sets is such that position in the r–ϕ plane and along the z direction can be measured. The outer layer just measures in

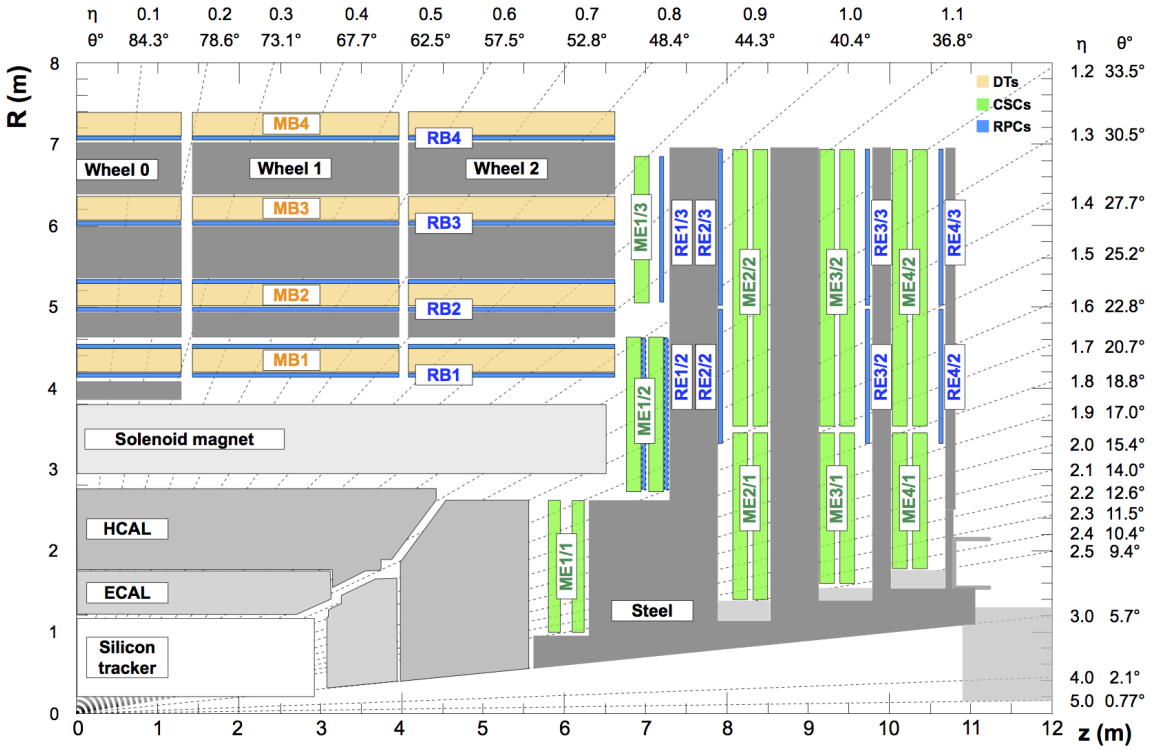


Figure 3.7: This cut-away of the muon system shows the barrel and endcap regions, and its arrangement with adjacent CMS components.[56]

the r–ϕ plane.

In each of the four layers of DTs, RPCs are installed for improved performance. RPCs have worse position resolution than DTs, but a faster response time, as the ionized particles drift a smaller distance. RPCs contain two parallel plates of resistive material. A high-voltage is setup across the plates and gas fills the gap. Muons passing through this gap create ionizing particles. The electrons from this ionization drift to the strips of anodes in each of the RPCs. As the muons travel from the interaction point at roughly one foot per nanosecond, they may not exit the detector before the next interaction has begun. The timing of the muon interactions as it passes through must be known precisely. The response time of the RPCs is 1 ns, much smaller than the time between a bunch-crossing.

In the endcap region, the magnetic field can be higher strength, and have a more complicated shape as the field lines connect through the solenoid. This, along with the higher particle density in this region, creates challenges necessitating another design for muon measurement. CSCs are used instead of DTs. A CSC is a flat gas chamber with anode and cathode wires running orthogonal and on opposite faces. The signal time at the anode is relatively prompt and can be used for the CMS trigger system. The charge on the cathode strips gives a more precise particle position, but is too slow for triggering. RPCs are also in the endcap for further performance. The endcaps have three layers of CSCs and RPCs with iron return yoke in between. Each CSC layer has six slices, giving excellent muon position and direction measurements. Beyond $|\eta| > 1.6$ the endcap only has CSCs. A complete discussion of the muon detector system performance can be found here [56].

3.2.6 The Magnet

The central solenoid magnet of CMS is a critical component of the detector. It produces a magnetic field of 3.8 T throughout the barrel regions of the tracker, ECAL, and HCAL, with an internal diameter of roughly 7 m. Any larger, and the magnet could not be delivered to the CMS site upon assembly. It also produces a 2 T field within the iron return yoke surrounding and supporting the muon chambers. The superconducting material is niobium-titanium, embedded in aluminum, with 18,000 A of current running through it. This results in a stored energy of approximately 2.66 GJ. The magnetic field exerts a force on charged particles according to their momentum perpendicular the magnetic field, which bends their paths. This allows for particle charge to be distinguished, as well as momentum measurement based on the radius of the track. Once constructed, the cosmic ray measurements with the muon system allowed for precision understanding of the magnetic field throughout

the detector. [57].

3.2.7 The Trigger

The LHC delivers collisions at 40 MHz. Only 1 in 400,000 can be written to disk. The complicated task of deciding whether an event is worth keeping is given to the trigger system. The CMS trigger system is comprised of two layers: a hardware level trigger designed to handle the full 40 MHz of events and filter it down to 100 kHz called the level one (L1) trigger, and the high level trigger (HLT) which is operated by a computer farm which further reduces the rate to 100 Hz.

The L1 trigger is designed to be able to make trigger decisions very quickly, as the new events continue to arrive, and to have as little latency as possible, as the more latency, the more information must be cached in the hardware prior to readout. As a result of these constraints, the L1 trigger can only perform simple and regional calculations to determine an event's characteristics. The ECAL and HCAL calculate basic information about the detector signals and give this information to the global trigger (GT). At the same time, the muon sub-systems calculate trigger information to pass to the GT as well. With this combined information events are saved based on the total information available. From here, a signal, Level One Accept (L1A), is sent to all the relevant detector hardware. The total event information, stored in a hardware pipeline, including the tracker systems, is sent to the HLT for further decisions. This happens on the order of microseconds.

The HLT software is run by roughly 10,000 CPU cores. Each event, which is several megabytes, with its full information available to the HLT, must be understood to winnow and keep just 1 in 1000. To do this, simple trigger paths, which are categories an event may fall into which would make it worth keeping, are calculated first. After simpler paths are exhausted, more complicated calculations have time to be made. Once an event passes an HLT trigger path it is immediately saved and the next can be analyzed. The software running on the HLT computers is the same as that used by the final reconstruction software. This allows full access to any algorithms used in the software.

The trigger system at CMS is able to be reconfigured easily, and is continually studied and improved to maximize its efficiency in selecting worthwhile events, and to prevent unintentional biasing of the saved data. With the total delivered luminosity of $162.85\,\mathrm{fb}^{-1}$, CMS has saved petabytes of collision data from the second run.

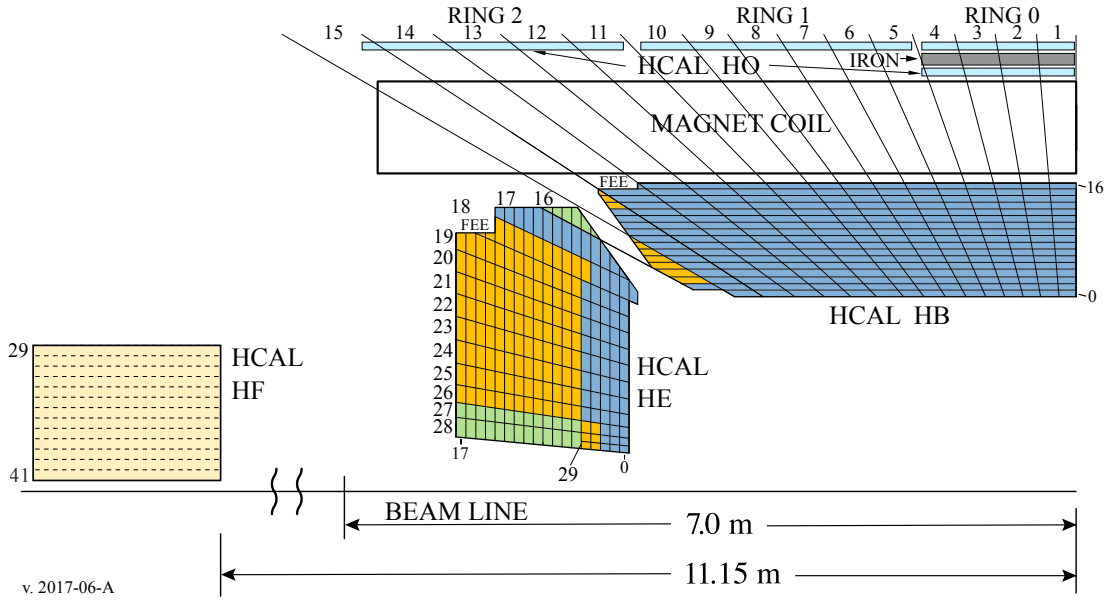


Figure 3.8: The Run I depth segmentation used by the HCAL. It can be seen that the majority of the HCAL is summed into one or two depths.

3.3 The HCAL Phase I Upgrade

3.3.1 Overview

The Hadronic Calorimeter for CMS as discussed in section 3.2.4, is critical in the measurement of hadronic activity in the detector and electron and photon identification. As the detector has aged, primarily from radiation during beam collisions, the performance of the plastic scintillator has decreased. In addition, the original HCAL design had limited readout bandwidth with the result that the energies recorded in several layers of plastic scintillator were summed to make depths. With more bandwidth, more layers can be recorded individually. Upgrading the light-gathering sensitivity of the detector, as well as increasing the bandwidth of information moving off-detector, gives the HCAL the ability to increase its performance for low-energy events even with the lower light yields of the damaged scintillator. The detector also has more depth information, by summing layers of plastic scintillator into more regions. With additional depth information, each layer of the HCAL can be individually calibrated. This improves performance especially as parts of the detector closer to collisions are radiation damaged more quickly. The number of depths in the HCAL before and after the Phase I upgrade is shown in Fig. 3.8 and 3.9 respectively.

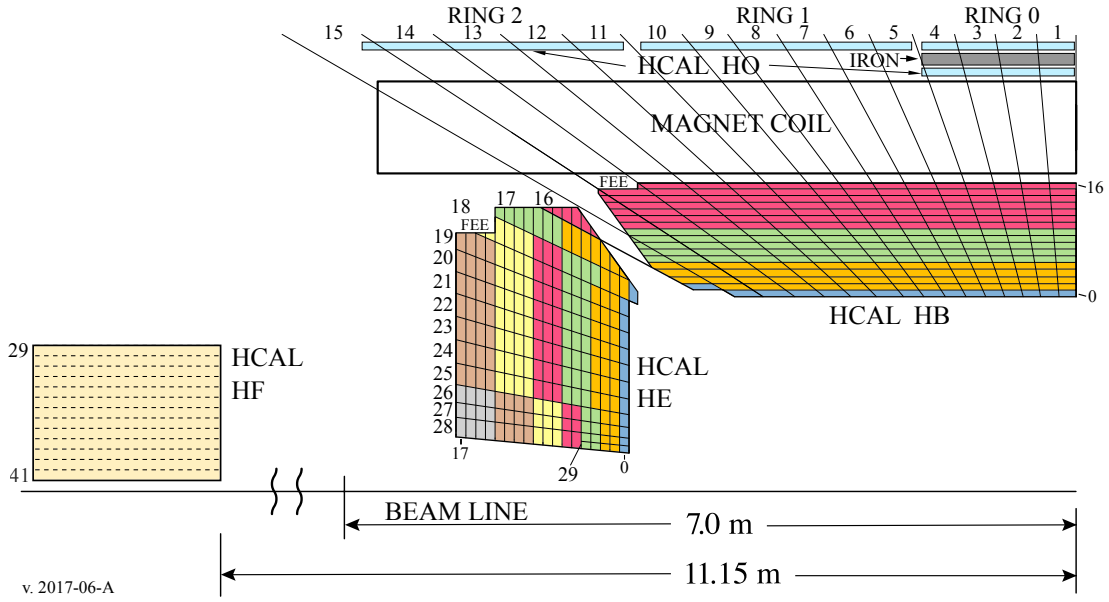


Figure 3.9: The complete phase I upgrade HCAL detector depth segmentation. Several more depths are available throughout the HCAL.

3.3.2 Front-End Control

With the upgrade of the HCAL detector readout capabilities came a desire to redesign the on-detector electronics control system. We'll give an overview of the hardware involved in the front-end control system, and then discuss the software elements involved in each piece. The goal of the new system was to manage the large amount of information needed to monitor and configure the detector hardware in a robust way. Here the author will make a brief detour to discuss these changes and one issue, important to the analysis, which occurred as a result.

Hardware Layout

A natural way to organize the hardware involved in the HCAL front-end control system is to distinguish between components on the detector, and components off the detector. CMS has two caverns, called the service cavern (USC) and the experimental cavern (UXC). The service cavern is protected from the radiation and magnetic field environment of CMS and contains the majority of the controlling hardware, from PCs in racks to power-supplies and custom detector readout electronics. UXC is almost entirely filled with the actual CMS detector and includes some cooling and power supplies as well as any electronics

required to amplify and digitize any detector signals. By necessity, these electronics are specially selected for each region of the detector to fit the constraints of not just enduring the high magnetic field and ionizing radiation, but also a compact size and reasonable power consumption.

In the front-end, the HCAL electronics are divided into readout boxes (RBXs). Each RBX contains the electronics necessary to digitally process the signals from scintillator. In parts of the HCAL with SiPMs, these are also contained in the RBX. For HF, the RBXs are located in racks adjacent to the detector, and process signals from the PMTs. For the HCAL phase I upgrade, each RBX contains a next-generation Clock and Control Module (ngCCM). The ngCCM receives and sends all fast and slow control through each RBX over a backplane. The ngCCM communicates with the USC side electronics through the next-generation Front-End Control Card (ngFEC). Each RBX contains several charge integration and digitization chips (QIEs) as well as a calibration unit (CU). Each ngFEC sits in a uTCA crate in USC. From here, it receives fast control messages from the trigger and clock distribution system (TCDS) and communicates with the ngCCM server, which will be discussed in 3.3.2 via IPbus. Data from the front-end is sent to back-end crates to FPGA based models called uHTRs. A diagram for the next-generation front-end system is shown in Fig. 3.10.

Software Control

The software control system is designed around a single application forming the hardware to software interface. This is called the `ngCCM server`. Other specific applications interface with the server, and reside in the HCAL application framework for handling specific tasks. The `ngCCM server` communicates with each ngFEC over IPbus. Its communications are either directly with the ngFEC, or are forwarded to an ngCCM over one of the two I2C links. The server has three client interfaces. The first is over websockets which sends and receives text strings in communication with the server, this link allows communication with the HCAL online software applications, and it is extensively used by an application called `ngRBXmanager`. The second client interface is used by the detector control and safety (DCS) system. DCS is responsible for managing power, cooling, and any other physical needs, as well as monitoring for any hazards. The third client interface is a command line interface program which is primarily used by hardware and software experts. A control scheme showing the various applications can be seen in Fig. 3.11.

The HCAL online software applications communicate with three primary purposes.

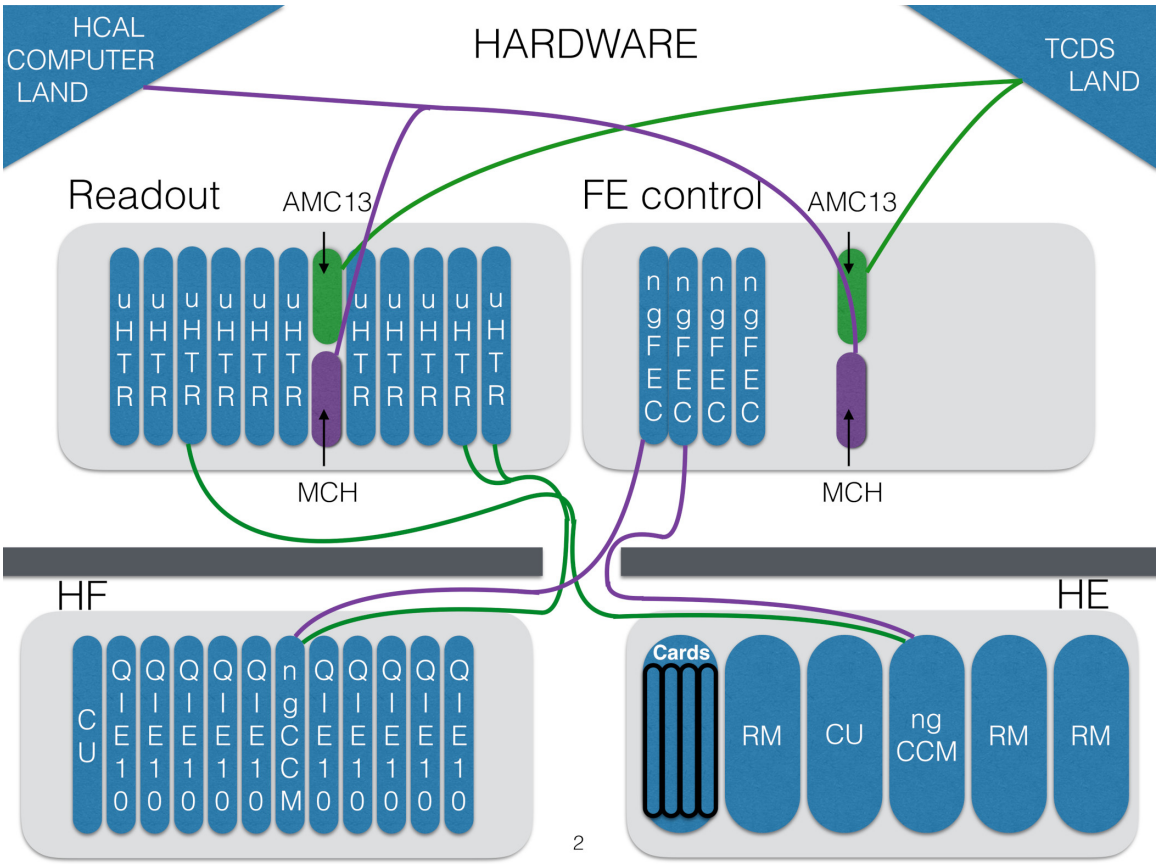


Figure 3.10: In the upper part of the diagram, the HCAL and computer system and TCDS are shown. These connect with back-end readout crates, containing uHTRs, and frontend control crates containing ngFECs. These crates sit in USC. The AMC13 and MCH are one to a crate and handle communication with various components over the crate backplane. From USC to UXC fiber optic cables for fast and slow control and detector readout sit. In the front-end, there are different RBXs layouts for the different the HCAL subdetectors. The QIEs, and there bundling in RMs in HE are shown, along with the CU and ngCCM.

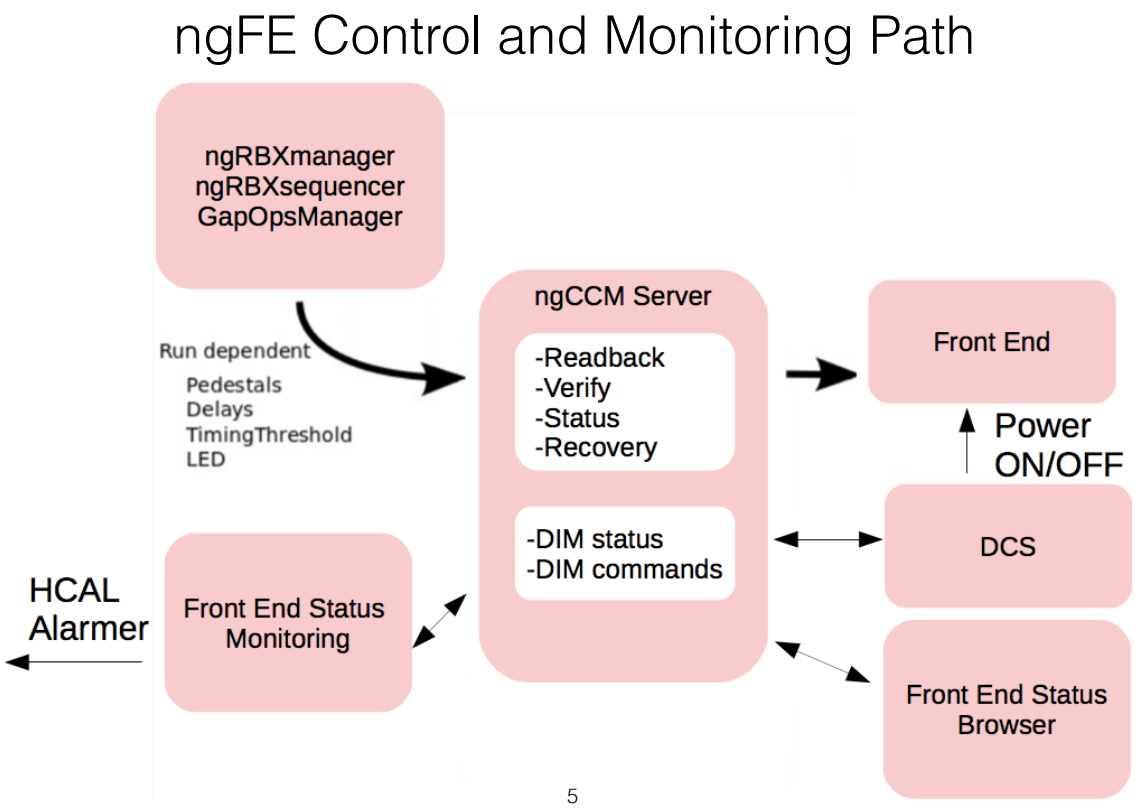


Figure 3.11: A layout of the HCAL front-end control software for the phase I upgrade. The ngCCM server is the central piece of software which receives communication from the HCAL applications and DCS.

`ngRBXmanager` is responsible for taking the HCAL configuration files and generating `ngCCM server` websocket client commands. These are sent when the detector is configured for data taking. The `ngRBXmanager` also ensures that the `ngCCM server` is properly working and its configuration of the detector was successful. The `ngRBXmonitor` communicates with the `ngCCM server` in the same fashion as `ngRBXmanager`, but does so at a regular interval. The focus of the monitor is to query the status of the server and hardware registers throughout the system. This information is available on request through the web, compiled into log tables, and digested by the HCAL alarmer, which can send automatic communications to experts if the system drifts from an acceptable state. The last job of the HCAL online software is done by the `ngRBXsequencer` and `GapOpsManager`, which each fill roles for special hardware configurations. These setup the detector for calibration and validation outside of physics events.

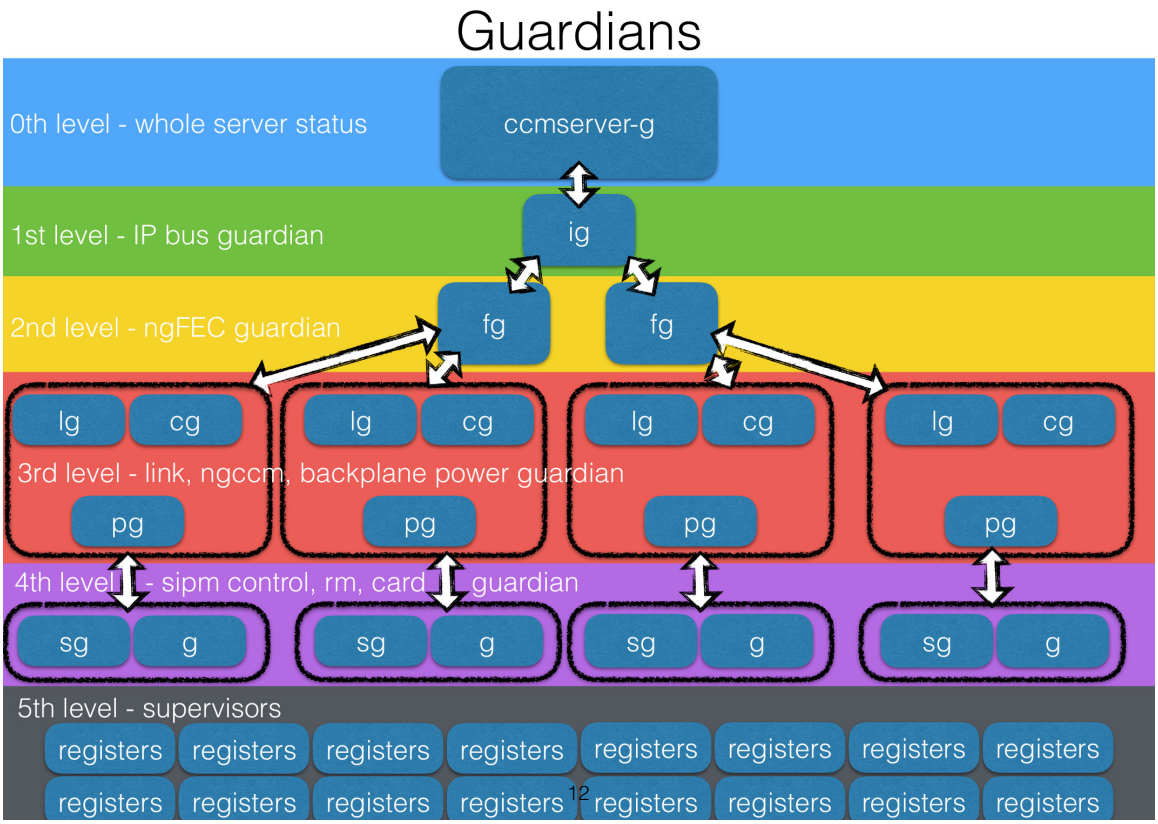


Figure 3.12: A tree of the `ngCCM server` guardians is shown. The trunk level being a purely software guardian for the server, this then branches out following the physical structure of the front-end system.

The `ngCCM server` has a unique structure, which, though complicated, has served to put several different needed software tasks into one program. The server is built with a "devActor" framework, where the server is made up of many individual actors, called "guardians". The guardians are configured when the server starts, each is responsible for a group of components of the front-end control, and any components which are "downstream" in communication. A guardian is created for each ngFEC the server communicates with, and a guardian for each RBX an ngFEC communicates with, these sitting under the ngFEC guardian. Commands propagate down the tree, and status information propagates up. At the base level, hardware registers have supervisors, the leaves on the tree. There are supervisors at every level with hardware registers, but majority of supervisors are for registers within RBXs. A diagram of this can be seen in Fig. 3.12

3.3.3 HEM Readout Box Failure

On June 30th, 2018 the new endcap electronics were stress-tested in an unexpected way. The smoke alarms on the surface in the room containing the main power supplies for the cavern were triggered. The result was an almost-immediate power cut to all of CMS (including the HCAL) electronics in USC. Following the power cut and recovery, two sectors on the negative η side of the detector (labeled HEM15 and HEM16, which make up 40 degrees of the endcap on their side) failed to power on. The power supply responsible for sending low voltage power to the control units for these sectors reported an over voltage limit error "OvHVMax". The power supplies responsible for each were replaced. Power-up was successful, but communications failed. The current consumption of the readout boxes also appeared strange. Another replacement of the power supplies was attempted, and again communication was unsuccessful. A graph of the current draw of the RBXs is shown in Fig. 3.13.

After extensive diagnostic tests and attempts to reproduce the event, it was determined that the power supplies had, in error, sent a voltage pulse of roughly 20 V. Given the maximum recommended voltage of 12 V for the module, this was sufficient to completely destroy the hardware. Due to the tight environment in which the HE RBXs are installed, it was not feasible to replace them in the middle of the 2018 data collection period. The low voltage power supplies were then taken for testing both at CERN and with the manufacturer, CAEN. After considerable testing, it was determined that under certain power-up conditions the power supply could send a brief higher voltage pulse on the low voltage line. Fearing a repetition of the issue as it was being understood, a very simple Zener diode protection circuit was added, which would prevent the low voltage line from exceeding the Zener breakdown, allowing for a cheap and simple solution to prevent damage to the RBXs that remained functioning, but had potentially flawed power supplies. The rest of the year's data was collected with HEM15 and HEM16 powered off. With 40 degrees of one endcap of the HCAL off, the physics performance of the detector is significantly affected, but not fully lost. CMS uses information from as many sub-detectors as possible to create reconstructed "particles". Any hadronic activity entering this region then suffers reduced precision, but is not completely lost. The amount of energy that would be expected in the dark region of the HCAL must be extrapolated. Electron identification also suffers. While there are many different properties that distinguish electrons in the detector, one key way they are distinguished from charged pions is that charged pions shower in the ECAL and HCAL, electrons do not, as they are too light to travel this far. The ratio of energy deposited in the

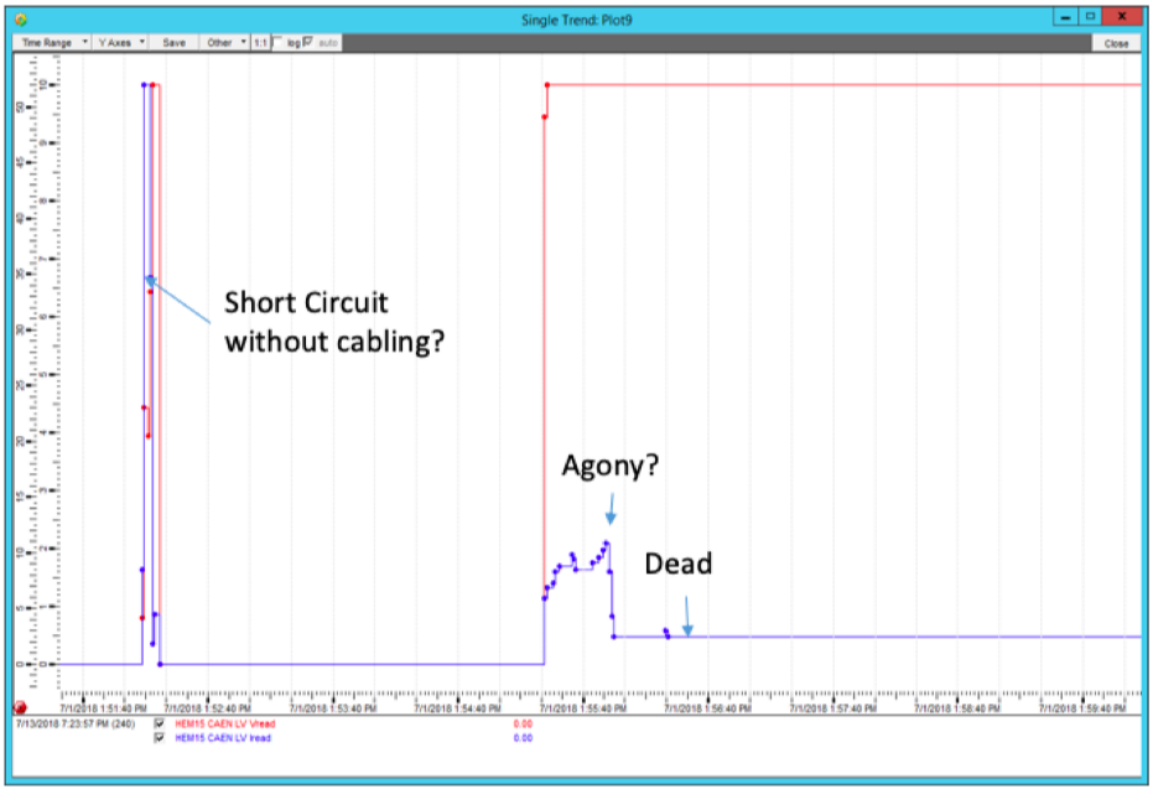
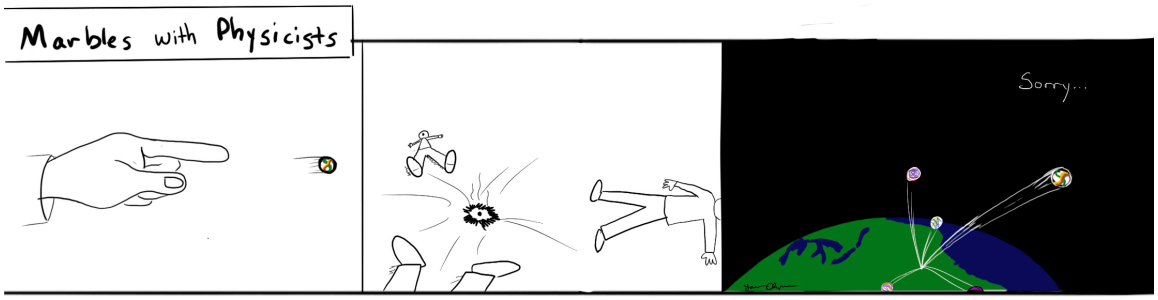


Figure 3.13: The voltage (red) across and current through (blue) one of the failed readout boxes, HEM15, is shown. The second jump in voltage corresponds to a current rise, fluctuation and then failure. The module was already broken at this point, and simply finishing "frying".

ECAL and HCAL behind allows for a discrimination which is impossible with the HCAL off. The specific effects on this analysis was studied and is discussed later in Section 6.4.

Chapter 4

Boosted Physics as an Experimental Tool



Having covered the theory of the standard model, LRS extensions, and the CMS experiment at the LHC in broad strokes, it is now important to focus in on cutting-edge techniques in boosted physics that make parts of this analysis possible.

LRS models do not give any suggestions for the N_R and W_R mass relationship. In the lighter N_R phase-space, the neutrinos will be produced with large transverse momentum, called "boost". Therefore, searches where the N_R is much lighter than the W_R, though heavy in standard model comparison, must be performed in tandem with past searches where N_R is assumed of similar mass to the W_R. A boosted N_R search is challenging, as the three decay products of the N_R are no longer likely to be individually isolated, as traditional W_R searches assumed. The specific ratios of the mass of the W_R and N_R where this object overlap from boost becomes significant is shown in Fig. 4.1. A more detailed discussion on the physics of N_R "jets" can be found in [58].

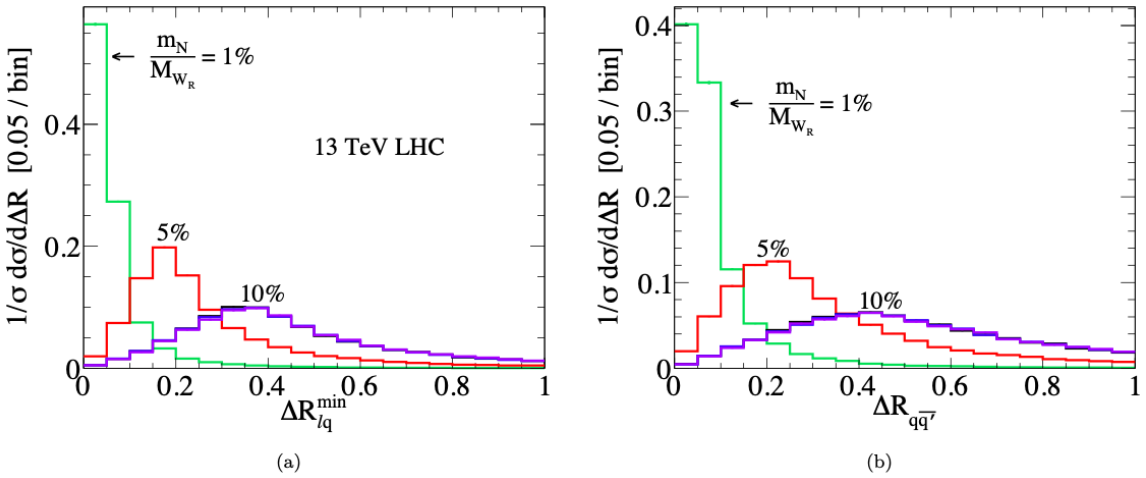


Figure 4.1: Minimum angular separation of the lepton and a quark from N_R decays (a) and the separation of the two quarks in the N_R decay (b) are shown. Three different mass ratios of the N_R and W_R are shown in each. For jets reconstructed with $\Delta R = 0.8$, virtually all of these N_R decays would have their quarks and lepton found within the jet[58]. As the N_R mass decreases, the decay products begin to fit within the $\Delta R = 0.4$, the smaller jets used in this analysis.

The substructure of jets made from boosted decay objects is the subject of ongoing study at the CMS and the LHC. The ability to distinguish multiple objects within jets allowed CMS to publish the groundbreaking observation of the Higgs to bb decay in 2018 [59]. This thesis adapts some of these techniques to expand the mass limits on the W_R and N_R where the N_R forms a merged jet. These are discussed in the following sections.

4.1 Boost

Before delving too deeply into the algorithms used to study jet substructure, it's important to take a step back and cover the basics of boost as a special relativity concept and one of the more simple examples of it at CMS, the decay of a boosted particle into two parts.

4.1.1 Special Relativity

Boost is a concept coming from special relativity. A basic premise of special relativity is that the speed of light should always be observed as constant, regardless of the relative motion of the observer with the respect to the emitter. The consequences of this, time dilation and length contraction, are observed as particles generally travel close to the speed

of light. Distances parallel to the motion of the particles are measured shorter in the lab frame than they would be measured with the particles at rest. Both time dilation and length contraction are proportional to γ, which is called "the boost". The boost can be written as a function of the speed of an object and the speed of light, or in natural units is the ratio of the energy over the rest mass of a particle:

$$\gamma = E/m_0. \tag{4.1}$$

Using γ, the length contraction and time dilation can be defined as

$$L = L_0/\gamma,\ t = \gamma t_0. \tag{4.2}$$

A measured proper length L_0, or the length measured when at rest, is divided by γ and becomes equal to lab measured value. The proper time t_0 is dilated by the factor γ. However, distances perpendicular to the motion are not affected.

4.1.2 Boosted Two-Body Decays

The effect of special relativity contracting only the parallel coordinates can have significant effects on a process as viewed in the lab frame. As an example, consider a process like $H \rightarrow b\bar{b}$. Considered at rest, the two decay quarks will travel back-to-back to conserve momentum. As can often happen at the LHC, the Higgs boson may be produced with significant momentum, and while its decay particles are produced back-to-back in the Higgs frame of reference, they are additionally travelling in the direction of the Higgs in the lab frame. This results in two quarks which travel in the same direction in the lab frame.

When the momentum of the particle undergoing decay is larger than its rest mass, the momentum the decay particles get from the boost will exceed the momentum received from the decay energy itself. In this case, and as is shown in Fig. 4.2, the decay particles will both travel in the direction of the initial particle, and in the lab frame appear closer and closer to each other as the boost continues to increase. At CMS, a rule of thumb in a two body decay is that the angular separation between the particles proportional to the mass and transverse momentum of the initial particle:

$$\Delta R \approx 2m/p_{\mathrm{T}}. \tag{4.3}$$

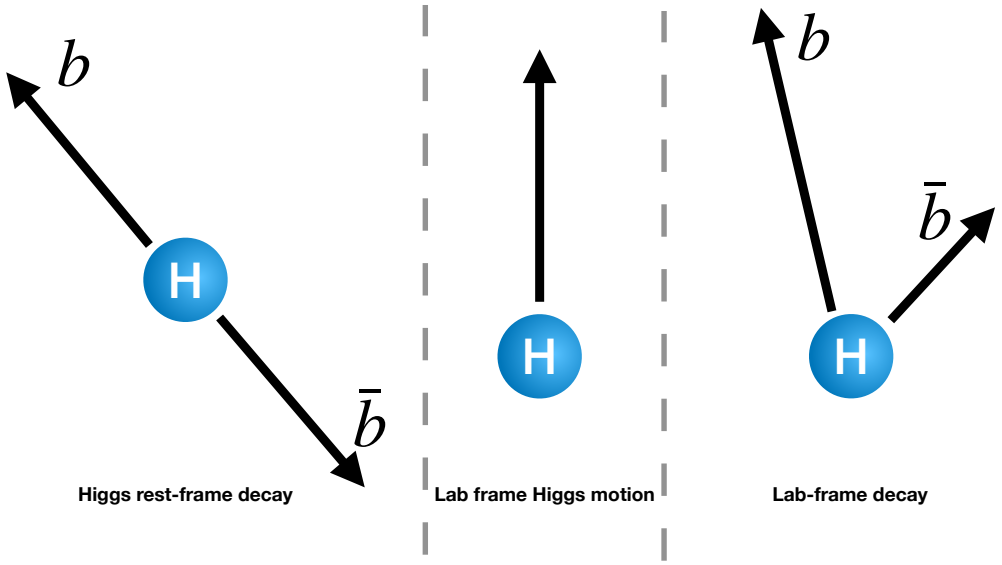


Figure 4.2: On the left, a back-to-back decay of the Higgs to two b quarks is shown. With significant momentum of the Higgs, this process is observed differently in the lab frame of reference, shown on the right.

4.2 The "impossible" Higgs decay

The search for the Higgs boson decaying to two b-quarks in CMS is a landmark study for boosted physics and provides a perfect frame-of-reference to discuss boosted techniques in.

The standard model Higgs boson couples to fermions according to their mass. With the heaviest fermion available for Higgs decay being the b-quark, the 58% of Higgs will decay to pairs of these quarks. The two leading discovery channels for the Higgs, however, were the $H \to ZZ \to 4\ell$ and $H \to \gamma\gamma$ decay channels. These make up less than half a percent of Higgs decays, far less frequent than a decay to b quarks [60]. From the beginning of the design of CMS and the LHC, it was not expected to be possible to even observe $H \to b\bar{b}$ as the rate at which everyday QCD processes can produce pairs of b quarks is vastly higher than the Higgs process.

As the collision energy and luminosity of the LHC increased in Run II, a new technique for analyzing Higgs decays emerged. Some small fraction of Higgs are produced in the LHC with significant transverse momentum. This Higgs momentum, or "boost", collimates the decay products of the Higgs in the lab frame making the signature distinct from a QCD process. A QCD process is very unlikely to produce two relatively collinear heavy quark

pairs with significant momentum. While the chance of a boosted Higgs is also low, the ratio of signal to background rises dramatically with the requirement of a boosted object. The new challenge then, is distinguishing two individual b quarks pushed together into a single jet by the boost of the Higgs. For full details on the boosted H $\rightarrow b\bar{b}$ analysis see [59].

4.3 Jet Algorithms

As discussed in Section 2.2.1, quarks and gluons exiting a hard collision at CMS produce a multitude of soft collinear QCD radiations. The hadronization process occurs extremely rapidly, and by the time the hadronized particles have travelled outside the beamline, they have formed a "jet": several energetic hadronic particles travelling roughly collinearly. As the jet of hadrons continues through the detector, each is measured in different ways according to its specific quark composition. In order to capture the kinematics of the original quark or gluon involved in the hard process, all of the QCD radiations must be associated and summed until all of the parts of the hadronization have been combined.

CMS generally uses a variant of the Cambridge-Aachen jet algorithm called anti-k_{T}. Each type of jet algorithm tries to organize all the particles observed in the detector into a set of jets, and has advantages and disadvantages dependent on a large variety of situations.

4.3.1 The anti-k_{T} Algorithm

Particle flow (PF) jets are created using a jet algorithm to associate individual PF particles. The two key relations in this algorithm are:

$$d_{ij} = \min\left(k_{ti}^{-2}, k_{tj}^{-2}\right)\frac{\Delta_{ij}^2}{R^2}, \tag{4.4}$$

$$d_{iB} = k_{ti}^{-2}, \tag{4.5}$$

where:

$$\Delta_{ij}^2 = (y_i y_j)^2 + (\phi_i \phi_j)^2 . \tag{4.6}$$

The first distance relation d_{ij} is computed between every particle or proto-jet in the event. The second distance d_{iB} is calculated between each proto-jet and the beam. The reconstruction begins by clustering together the closest pair, and continues until every object is included. For instances where d_{iB} is the closest, the jet is removed from consideration for further clustering.

As an example to understand the behaviour of the anti-k_{T} algorithm, consider an event with a few well-separated hard particles with transverse momenta $k_{t1}, k_{t2}, ...$ and several softer particles. The $d_{1i} = \min\left(1/k^2, 1/k^2\right) \Delta_{1i}^2/R^2$ between the first hard particle and one of the soft particles, i, is determined by the transverse momentum of the hard particle and the Δ_{1i} separation. The d_{ij} between similarly separated soft particles will be much larger than d_{1i}. As a result, soft particles will tend to cluster with the hard particles in the event instead of forming clusters with each other. If a hard particle has no hard neighbours within a distance $2R$, then it will simply accumulate all the soft particles within a circle of radius R, and form a conical jet.

In the case where a second hard particle is present and $R < \Delta_{12} < 2R$ two hard jets will form, but it will not be possible for both to be perfectly conical. If one of the jets has significantly more k_t, that jet will be conical and the other jet will only be partly conical, missing the soft radiation overlapping with the higher k_t. If the two jets are roughly equivalent in k_t neither will be completely conical and the overlap region will be divided between the two. This boundary region is defined by $\Delta R_{1b}/k_{t1} = \Delta R_{2b}/k_{t2}$ [61].

4.3.2 Jet Boost

Looking at Cambridge-Aachen algorithms broadly, one algorithmic assumption is important to revisit when studying boosted objects. The anti-k_{T} searches to combine particles within a defined maximum size cone in $\eta - \phi$ space. Particles close in momentum and position are iteratively combined together within the cone. These particle pairs are then combined with each other and so on, until all well associated particles have been clusters (and those clusters clustered) to make a single jet.

This process works well when the original partons come from an interaction approximately at rest transversely in the detector. Significant transverse boost of the production frame additionally collimates the hadronization in the lab frame, resulting in overlapping jets. To combat this, boosted physics searches use larger cone size jets to collect all the particles in a region of the detector which could be associated, for example, with the $\mathrm{H} \to b\bar{b}$. These "fat" jets are then reclustered into small jets, often two or three, to distinguish sub-components of the jet more tightly associated with each other than the rest of the jet.

4.4 Techniques in Boosted Physics at Hadron Colliders

To leverage the kinematic uniqueness of boosted systems significant advancements in techniques have arrived in recent years for the study of jets formed from particles seen as merged together as a result of their boost.

4.4.1 Jet Substructure

Several different techniques are used to study the constituents of "fat" jets in boosted physics. N-subjettiness is used to determine how many tightly associated subclusters a fat jet might have, key for identification of $\mathrm{b\bar{b}}$ jets in the boosted Higgs analysis. The techniques used in this analysis are lepton subjet fraction (LSF) and soft-drop mass and are discussed further below.

N-subjettiness

N-subjettiness is based on reclustering the components of a large jet into multiple smaller jets and then evaluating whether the subcluster splittings are natural or forced. Heavy QCD jets, which can spoof interesting physics, like a hadronically-decaying electro-weak boson, will have many energetic components all travelling roughly collinearly. A hadronic decay of a Higgs, however, will have two adjacent b-quark jets and little else.

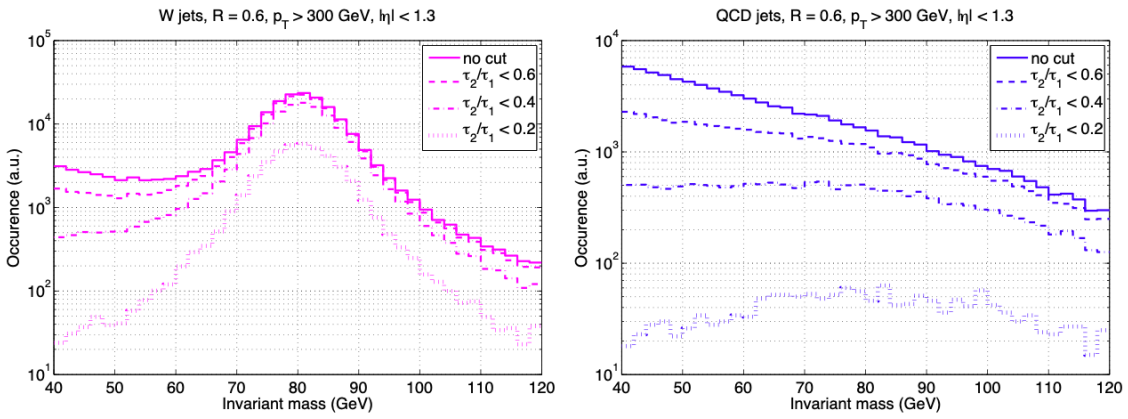
In order to distinguish between a jet of many approximately equally energetic components which can all be associated, and a jet with two distinct clusters of particles, N-subjettiness defines a variable τ_N which can be calculated for different numbers of subjets, N. The variable τ_N can be thought of as a measure of the "cost" to associate subclusters, measured in position and momentum space. Generically, τ_N is calculated as:

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min \{\Delta R_{1,k}, \Delta R_{2,k}, \cdots, \Delta R_{N,k}\}. \tag{4.7}$$

A sum over all particles (k) in the larger jet is performed, comparing each to the kinematic centers of a chosen number of subclusters. For each particle, its momentum is multiplied by the ΔR of the closest subcluster center. For only one jet cluster, τ_N is simple to calculate:

$$\tau_1 = \frac{1}{d_0} \sum_k p_{T,k} \Delta R_{1,k}. \tag{4.8}$$

For the case where two subclusters are considered, the closest of the ΔR between each



(a) (b)

Figure 4.3: N-subjettiness performance can be seen in (a) with W jets and (b) QCD jets. Tighter cuts on the τ_2/τ_1 ratio yield better W mass peaks [62].

particle and the two subcluster kinematic centers is taken.

$$\tau_2 = \frac{1}{d_0} \sum_k p_{T,k} \min \{\Delta R_{1,k}, \Delta R_{2,k}\} . \tag{4.9}$$

The ratio of τ_N for different values of N is taken to compare the relative "cost" of different subcluster arrangements. In the H $\rightarrow b\bar{b}$ study, the most important ratio is $\tau_2/\ \tau_1$, called τ_{21}. A number closer to zero points a more natural di-cluster than a single jet clustering. The discriminating power of τ_{21} to distinguish between an electroweak boson's hadronic decay, and a QCD jet is shown in Fig. 4.3.

This analysis studied the benefits of various N-subjettiness ratios, and did not find any to be particularly helpful for discriminating signal and background. Which suggests that the leading backgrounds for the analysis have a similar topology to the boosted neutrino decay. The QCD background which is important in H $\rightarrow b\bar{b}$ is suppressed in this analysis by requiring a lepton within the fat jet cluster. We therefore move to an important tool for lepton identification in the boosted regime. A complete discussion of N-subjettiness is found in [62].

Lepton Subjet Fraction

Lepton signals at hadron colliders like the LHC serve as clear indicators of interesting physics. While quarks and gluons play large roles in interesting particle physics, hadron

colliders produce them in over abundance. The events of quarks and gluons in the final state, hadronizing to jets, make up the vast majority of the interactions at the LHC and make distinguishing a specific process of interest challenging. Leptons, on the other hand, indicate an electroweak interaction occured. If that interaction occurred early in the event (rather than as part of a meson decay), it is often a signature of an interesting event.

To select away from meson decay, leptons used in searches for new physics are required to be isolated from other detector signals. A hard electroweak interaction typically kicks its lepton far away from the rest of the interaction, well isolating it. This scenario holds for the resolved version of this analysis, where the N_R and W_R are not moving before decaying, and each particle travels far from the others. When a meson decays, however, the muon will share that meson's momentum, and travel roughly collinear with the hadronic activity associated with the meson, which is part of a jet, The jet, therefore, will often contain leptons from the decay of its charged mesons. The situation where a lepton is travelling along with other jet constituents is similar to when the N_R is boosted. A boosted N_R will produce a muon physically close to the quarks in the N_R decay similarly to the meson decay picture.

In the typical non-boosted analysis, relative momentum isolation is used to remove decay-in-flight leptons. The p_T of all particles within a certain size cone are summed and compared with that of the lepton. This is defined as:

$$\mathcal{R}^{\ell}_{\text{iso}} = \frac{\sum_i p_{T,i}}{p_{T,\ell}}. \tag{4.10}$$

The cone size is typically 0.3, and the value of $\mathcal{R}^{\ell}_{\text{iso}}$ can be tuned for a particular analysis and often falls within the range 0.1-0.2. It can be readily seen that the fixed-size cone fails again for boosted objects. In the specific case of the lepton in a N_R jet, it will be momentum-isolated from the two quarks in the N_R decay, but the relative isolation defined in Eq. 4.10 will fail to identify this.

To measure the isolation of leptons in boosted jets, a technique called Lepton Subject Fraction (LSF) is used. In this analysis, the selected "fat" jet is clustered into three subjets corresponding to the expected number of original constituents (a lepton, and two quarks). Any lepton within the main jet will be clustered into one of the 3 subjets. The relative transverse momentum of the subjet and the transverse momentum of the lepton defines LSF. In the re-clustering, the muon should be too far away in momentum space to be clustered with most of the two quarks' QCD radiation. Therefore, the lepton will be mostly

alone in its subjet and the ratio of its momentum to the momentum of the entire subcluster will be close to one:

$$\text{LSF} = \frac{p_{T,\ell}}{p_{T,sj}}. \tag{4.11}$$

The performance of LSF versus relative isolation is shown in Fig. 4.4 [63].

Soft Drop Mass

At the current LHC luminosities, multiple events interact and are recorded simultaneously in a bunch crossing. This "pileup" of events is problematic for jet measurements. As each jet is constructed, QCD radiation from the pileup and the underlying event will additionally be clustered with the hadronization of interest, or in many cases be made purely of QCD. The soft-drop procedure removes soft, wide-angle radiation in jets. The mass of a jet from QCD contributions tends toward zero after this process while the mass coming from hard radiation remains.

In soft drop, jet constituents are clustered in a pair-based clustering scheme with angular ordering (the Cambridge-Aachen jet algorithm[61]). The clustering is brought to where there are just two subjets, we can label them as j_1 and j_2. These are then compared with:

$$\frac{min\left(p_{T1}, p_{T2}\right)}{p_{T1} + p_{T2}} > z_{cut}\left(\frac{\Delta R_{12}}{R_0}\right)^{\beta}. \tag{4.12}$$

On the left, the p_{T} fraction of the lowest subcluster is compared to the angular separation based variable $\left(\frac{\Delta R_{12}}{R_0}\right)^{\beta}$ multiplied by a scale, z_{cut}. The behaviour is then tune-able with two variables: z_{cut} and β, depending on the desired response. At CMS typically $\beta = 0$ and $z_{cut} = 0.1$. This simplifies the relationship to:

$$\frac{min\left(p_{T1}, p_{T2}\right)}{p_{T1} + p_{T2}} > 0.1. \tag{4.13}$$

If this condition is true, neither subcluster is removed from the jet, as they are each considered sufficiently hard. If the condition is false, the softer subcluster is removed, and the jet is redefined to be the subjet with largest p_{T}. The two subclusters of this new jet are evaluated and the procedure continues until the two compared subclusters are comparably hard. A cartoon of this process on a QCD jet and a N_R jet is shown in Fig. 4.5. The soft drop mass procedure is detailed in this paper [64].

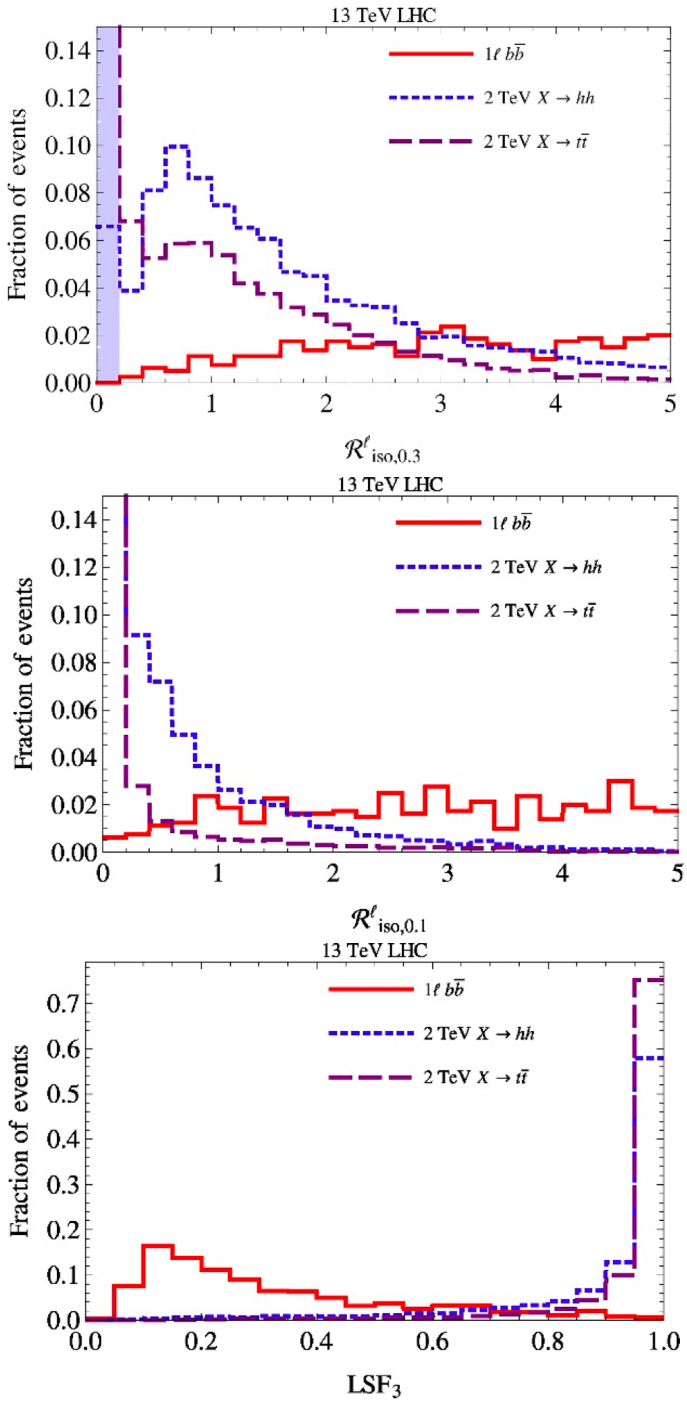


Figure 4.4: From top to bottom, these plots show the performance of a relative isolation cut of 0.3, 0.1, and LSF_3. Each shows the behavior for a semi-leptonic $b\bar{b}$ and a theoretical 2 TeV object decaying to two Higgs or a $t\bar{t}$ pair.

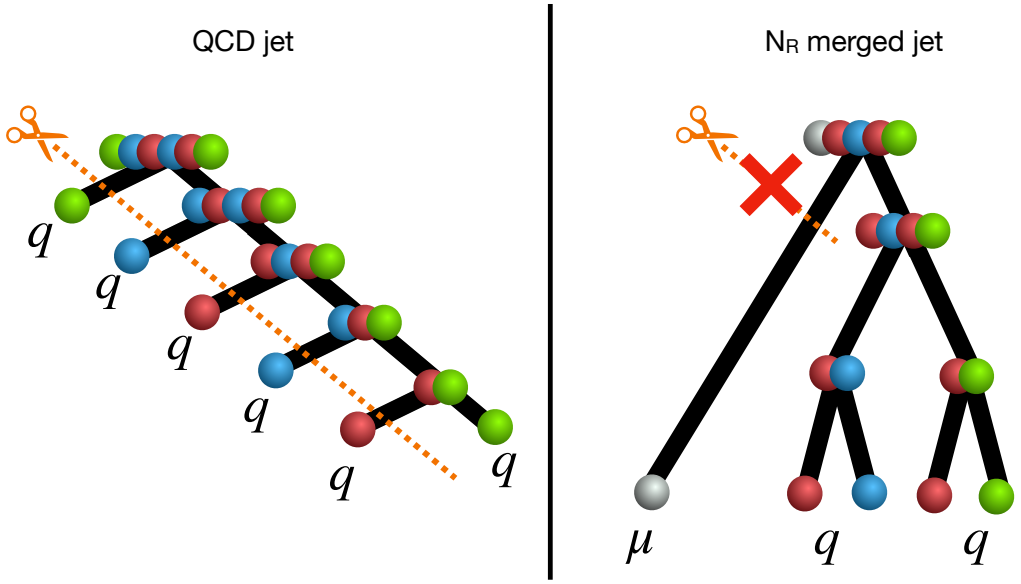
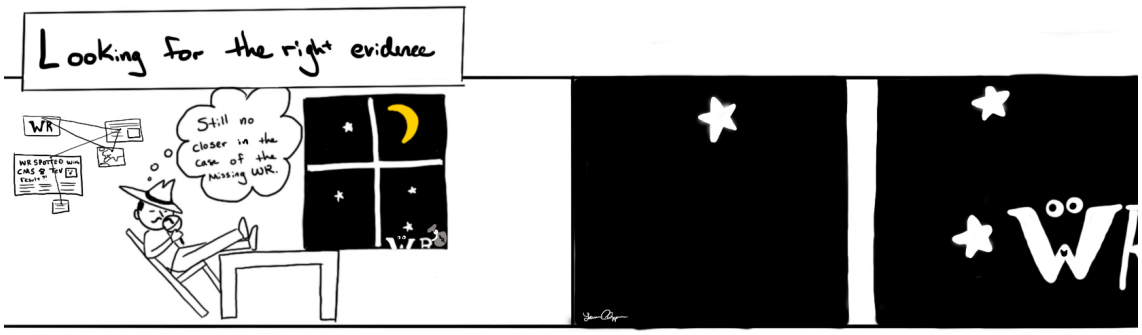


Figure 4.5: A cartoon of the structure of a QCD jet and a N_R jet is shown on the left and right, respectively. The QCD jet is made up of many soft radiations, each often having much less momentum than the subcluster it is joined with. As the soft drop procedure proceeds, most of the jet is trimmed away. The remaining mass of the jet is minimal. In the N_R case, while there is some color-flow between the two quarks, the muon and the quarks are each hard enough that soft drop does not trim any parts of the jet. The remaining jet mass is close to the N_R mass.

Chapter 5

Analysis Strategy



The goal of this chapter is to provide an overview of the strategy adopted by this analysis. This analysis performs a direct search for LRS physics with a higher integrated luminosity than ever before, and uses boosted-object reconstruction techniques in an orthogonal selection region in order to search for LRS physics in an expanded W_R–N_R mass phase-space.

5.1 Overview

At this point, it is valuable to return attention to the Feynman diagram illustrating the W_R–N_R production and decay, Fig. 5.1. While the primary decay mode of a W_R would be a back-to-back quark-antiquark pair, as discussed in Section 2.3, searching for the leptonic decay of the W_R is more statistically powerful. However, the separation of the four decay objects is heavily dependent on the N_R mass. For a N_R mass substantially lighter than the W_R mass, the N_R decay particles collimate in the lab frame, boosted by the momentum of the N_R. The lepton produced in the initial W_R decay process will generally travel opposite

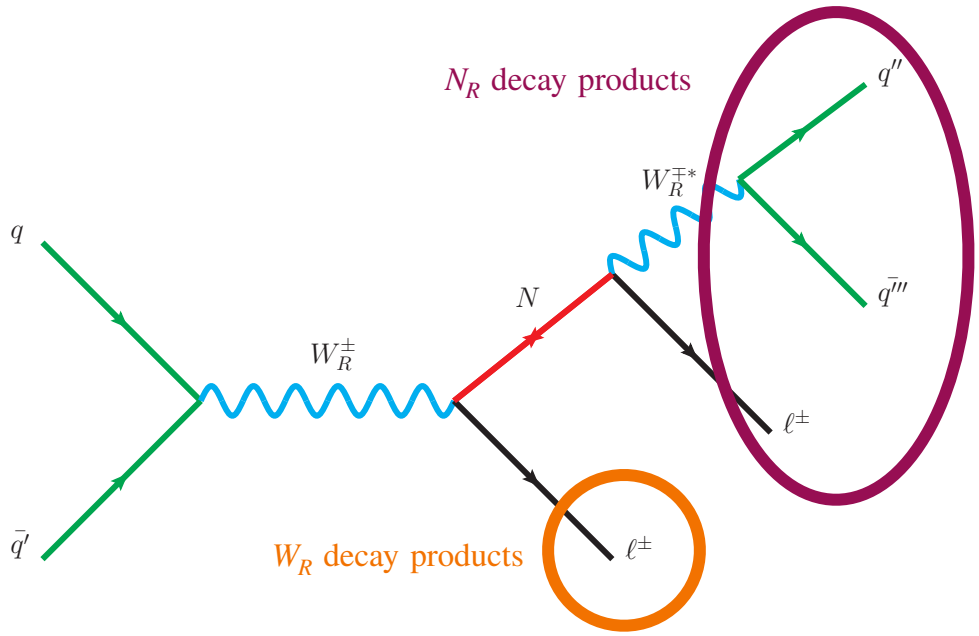


Figure 5.1: The principal Feynman diagram for this analysis. Two quarks annihilate into a W_R boson, which decays to a lepton and a right-handed neutrino. This neutrino decays by a virtual W_R and another lepton. This virtual boson decays to quarks, which produce jets. The final state objects produced from the W_R and N_R decay are highlighted separately. The relative separation of the N_R decay products from the W_R decay products changes in the W_R–N_R mass space.

to the N_R path, as the W_R will typically not be produced with significant momentum.

Past analyses have required all four final-state objects to be well-separated in the lab frame [4], which relies on an increasingly-narrow kinematic phase-space as the N_R becomes lighter relative to the W_R. Some new selection criteria must be integrated into the analysis to pursue the lighter N_R phase-space.

For this analysis, the strategy is to consider every W_R–N_R hypothesis with two orthogonal selections. As no event can pass both selections, the total number of events passing each selection can be combined in the statistical analysis. In addition, the simple kinematic definition of the selection regions allows future theorists to cleanly reinterpret the results for new theories.

5.2 Kinematic Region Selection

The two signal regions will be discussed, starting with the lepton selection which is common to both signal regions. For an event which passes the lepton selection, the jets reconstructed in the event are evaluated with resolved selection. If this fails, the boosted selection is tried. The combination of the resolved and the boosted selections is designed to allow almost all signal events.

5.2.1 Lepton Selection

Both types of signal selection require two well-separated leptons of significant transverse momentum. The leading lepton in the event must satisfy $p_{\mathrm{T}} > 60\,\mathrm{GeV}$ and the sub-leading lepton must satisfy $p_{\mathrm{T}} > 53\,\mathrm{GeV}$. These two leptons must also be separated by $\Delta R > 0.4$. After these requirements, the dominant backgrounds are those with two real leptons, such as $\mathrm{Z} \to \mathrm{e}^{+}\mathrm{e}^{-}$ and $\mathrm{t\bar{t}}$. General QCD processes and inclusive W processes are highly suppressed. In addition, to keep the leptons within the tracker, and muon chamber acceptance, both leptons must satisfy $|\eta| < 2.4$. This requirement reduces signal as well as background to some extent. Depending on the mass of the W_R, between $70 - 85\%$ of signal events pass these requirements at the generator level.

5.2.2 Jet Selection

Resolved Approach

The resolved jet selection begins by requiring at least two anti-k_{T} size 0.4 jets with a transverse momentum higher than 40 GeV. Requiring this momentum from the jets reduces

background and helps guarantee that the jets did in fact come from a hard-process (more about this will be discussed in Chapter 6). The jets are also required to have their center-of-momentum within $|\eta|$ of 2.4. This requirement mimics the η requirement placed on the leptons. These jets also must be more than 0.4 in ΔR apart from each other and the two selected leptons. This multi-object separation requirement ensures that this event can be reconstructed under a resolved paradigm. If events fail to identify the two jets, or they are not sufficiently separated, the boosted selection is attempted.

Boosted Approach

The boosted jet selection begins by requiring events to have at least one anti-k_{T} size 0.8 jet with a transverse momentum higher than 200 GeV. In a signal event, the N_R jet will have a large amount of momentum, and so this requirement is not significantly penalizing. This gives a great opportunity to remove significant amounts of background. For the same reason as the resolved selection, the fat jet must have an $|\eta| < 2.4$.

The next boosted jet requirement is that the sub-leading lepton must fall within the cone of the fat jet. This requirement ensures that the fat jet selected is N_R-jet like. In addition, while CMS produces jets in abundance, even at these high transverse momenta, high transverse momentum leptons are uncommon within these jets. To guarantee that not only is the sub-leading lepton within the jet, but also that the leading lepton travels away from the N_R jet, the $\Delta\phi$ between the fat jet and the leading lepton must be at least 2.0.

5.2.3 Multi-Object Selections

Having selected all of the jets and leptons necessary to reconstruct the event under either the resolved or boosted paradigm, the qualities of combinations of these objects can now be judged.

W_R Mass

This analysis hopes to reconstruct the W_R mass peak. In the resolved case, this is done by summing the 4-vectors of the four selected objects and calculating the mass of that sum; $M_{\ell\ell jj}$. In the boosted paradigm, the W_R mass can be reconstructed from the N_R jet (J) and the leading lepton (ℓ); $M_{\ell J}$. As this analysis focuses on extending the limit of past analysis in the W_R–N_R mass phase-space, the requirement on the W_R mass can be placed quite high to obliterate remaining background rates. In both analyses, the reconstructed

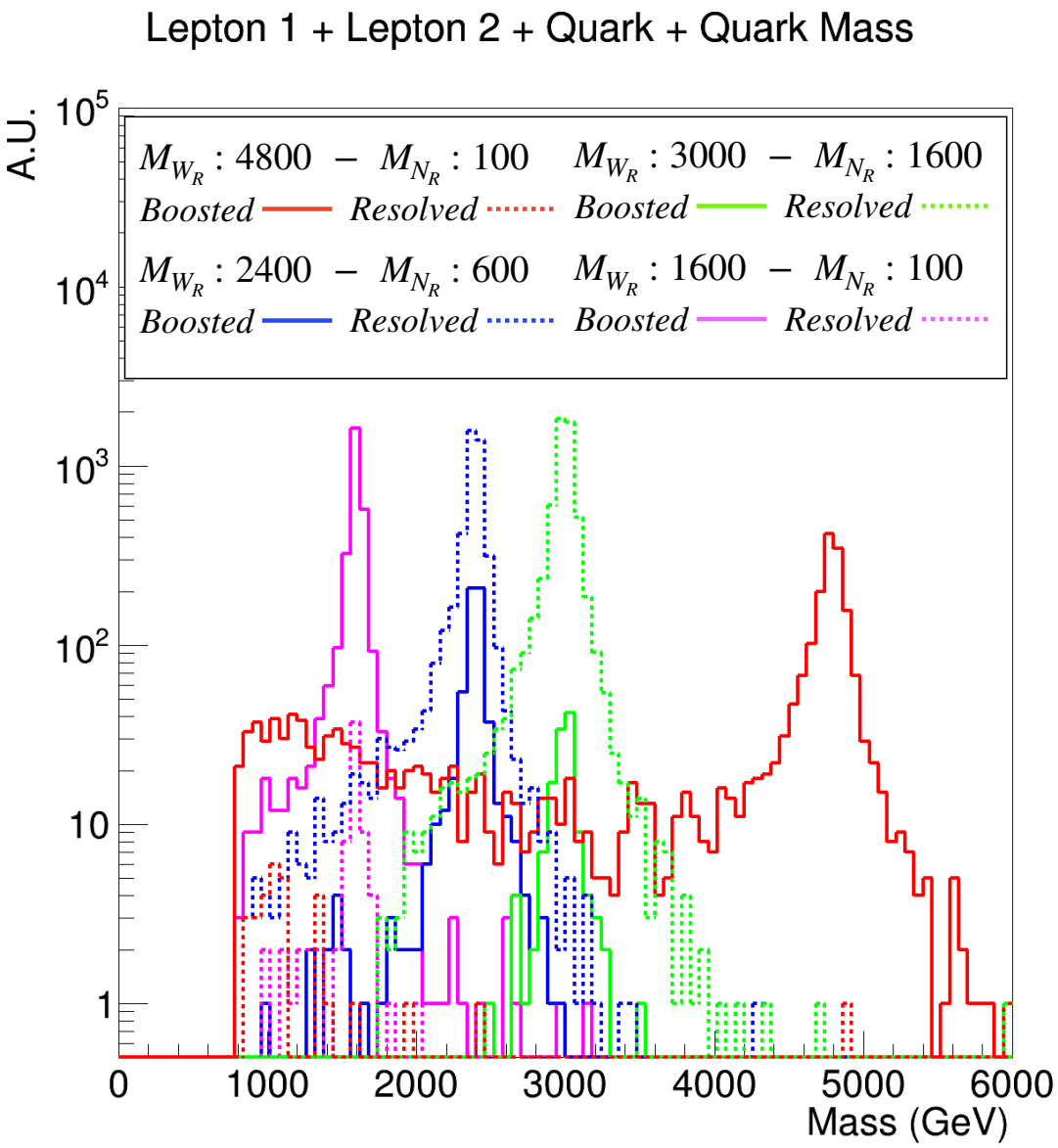


Figure 5.2:
The W_R mass for four different W_R and N_R mass hypotheses. The integral of each histogram is scaled relative to the acceptance of the analysis at that mass point. The resolved selections are shown with dashed lines, and the boosted selections are shown with solid lines. It can be seen that for the most massive and most boosted mass point, the boosted selection far outperforms the resolved selection, which only succeeds in the lower off-shell tail of the W_R spectrum. The shape of this spectrum is discussed in Section 2.4.1.

W_R mass must be at least 800 GeV. The shape of four different W_R masses, spread across the searched region are shown in Fig. 5.2.

Di-lepton Mass

One of the most significant backgrounds in this analysis comes from the Drell-Yan process. Off-shell high-mass Drell-Yan events and high-momentum, mis-reconstructed Drell-Yan both form parts of this background for both boosted and resolved selection regions. To significantly reduce this background, a requirement is placed on the mass of di-object formed by the two selected leptons. This mass must be higher than 400 GeV.

In past analyses, the di-lepton mass requirement was considerably lower, at 200 GeV. While this requirement is certainly high enough to drastically reduce the on-shell Drell-Yan background, it was increased to further reduce all backgrounds for this analysis. A comparison of the expected limits across W_R masses for the new (400 GeV) and old (200 GeV) requirements for the resolved analysis is shown in Fig. 5.3. Generally, the expected limit is improved by $10 - 20\%$ for W_R masses higher than 1000 GeV.

5.2.4 Signal Efficiencies

A scan of the signal efficiencies is shown in Fig. 5.4 and Fig.5.5. In general, signal efficiency improves as the W_R is heavier. For the boosted signal region, it can be that the efficiency also improves as the N_R lightens compared to the W_R up to a point. After this, the electron-within-jet selection becomes less efficient. The efficiency also begins to fall as an increasing fraction of the produced W_R events are lighter and off-shell. This behaviour is discussed in Section 2.4.1.

5.3 Lepton Flavor and Kinematic Sidebands

In this analysis, the behavior of background and signal in the signal region are simulated. When data is studied in the signal region, the question is then posed whether the data behaves more consistently with background or a background + signal hypothesis. In order to understand the behaviour of background and signal without looking directly at the data (which would potentially contaminate the analysis), alternate selection regions are created to validate and correct simulated background behaviour. These regions are commonly called "sidebands". In this analysis, sidebands are created by changing two kinematic requirements, $M_{\ell\ell}$ and $M_{\ell\ell jj}/M_{\ell J}$ and the lepton flavor requirements are changed to create a signal-like (but signal-free) sideband. These requirement changes create, in total, four sidebands.

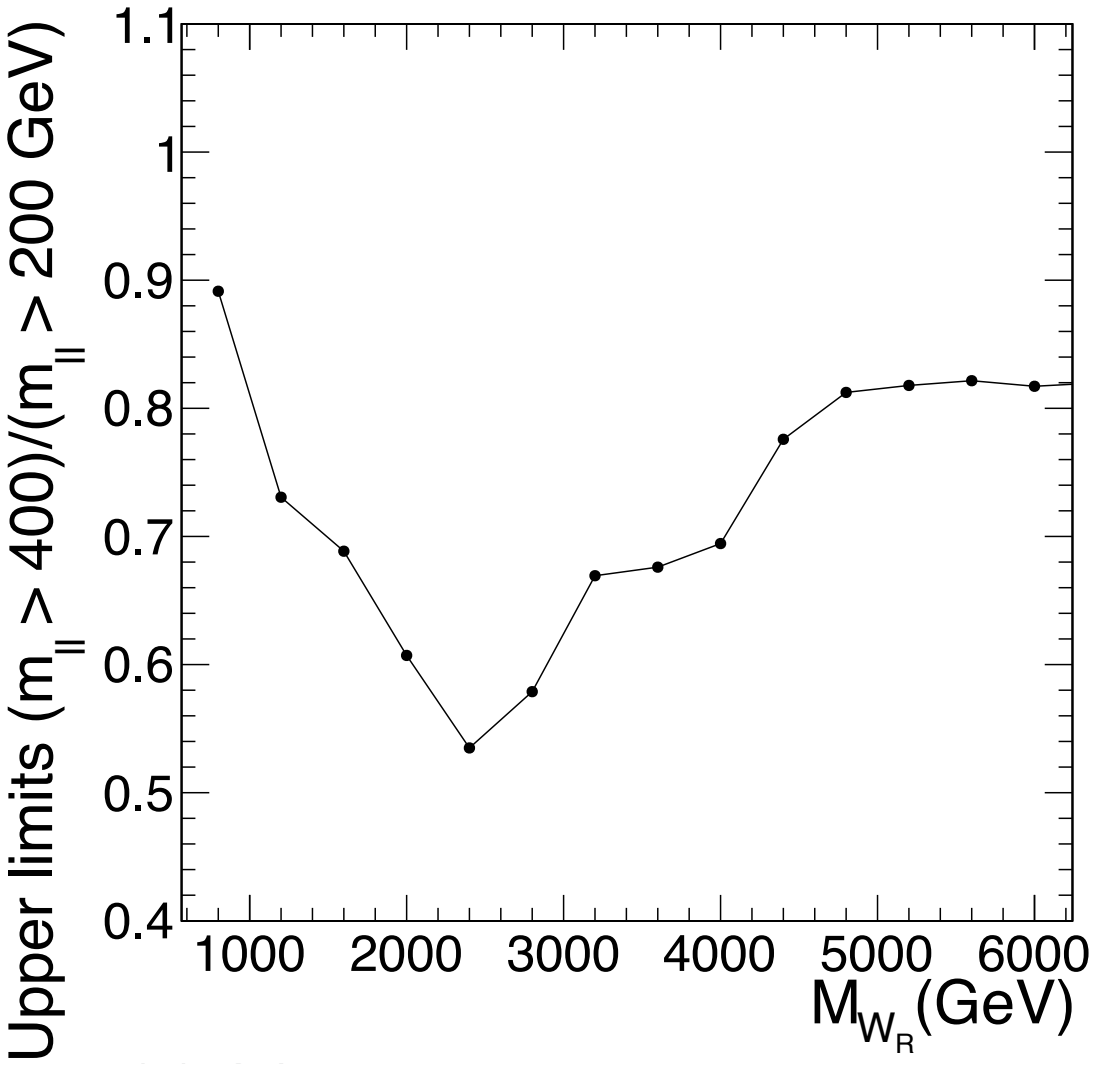


Figure 5.3: The ratio of the expected limits in the muon channel of the resolved analysis with the $M_{\ell\ell} > 400\,\text{GeV}$ and $M_{\ell\ell} > 200\,\text{GeV}$. The expected limits are calculated for signals with $M_{W_R}/M_{N_R} = 2$. The expected limits are stronger for the $M_{\ell\ell} > 400\,\text{GeV}$ selection for all signals with $M_{W_R} > 1000\,\text{GeV}$.

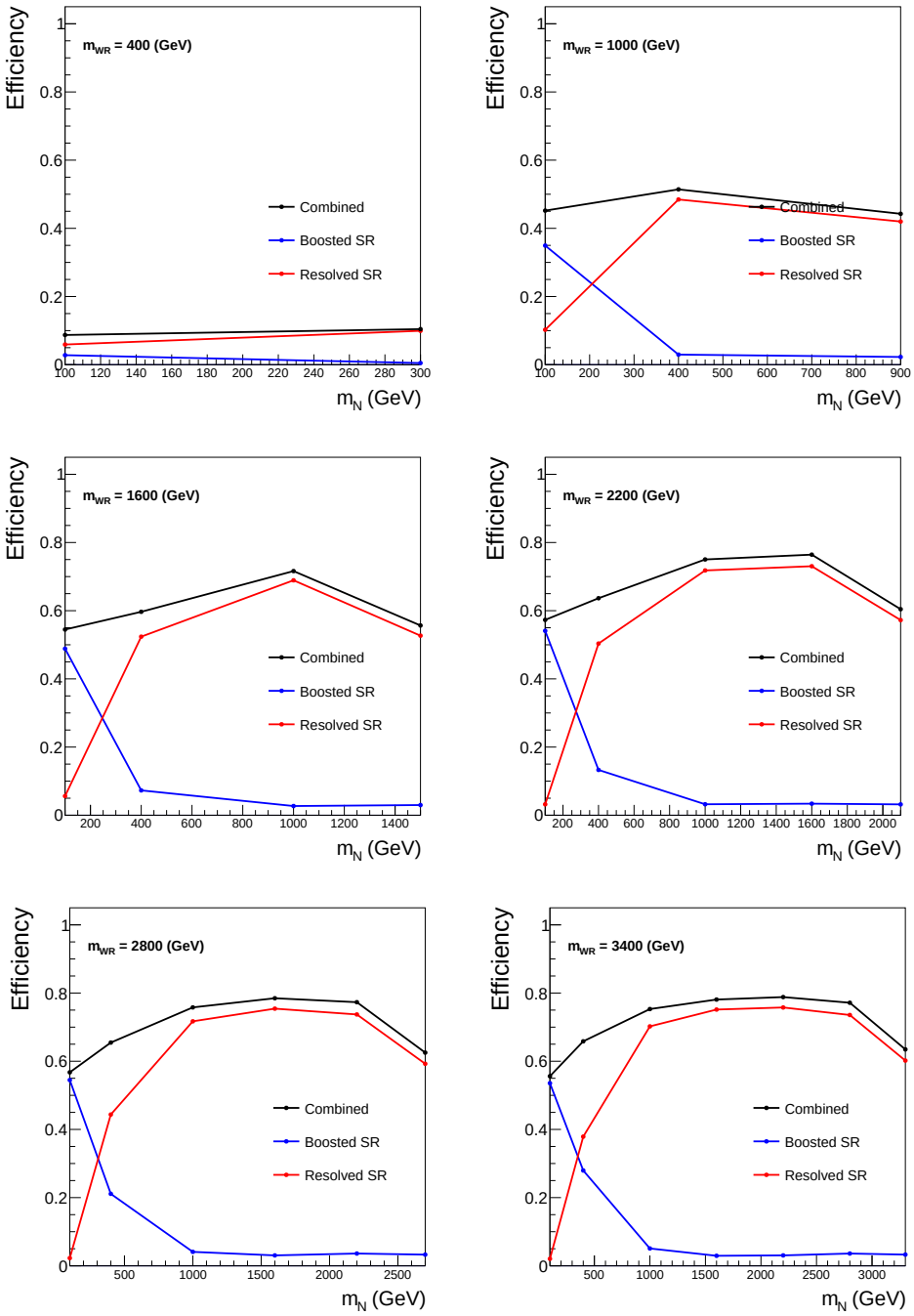


Figure 5.4: The signal efficiencies in the signal regions (SR) for W_R masses of 400 GeV – 3400 GeV. In blue and red, the boosted and resolved signal region efficiency is shown respectively. The combined efficiency of each selection at a given mass point is shown in black. Efficiencies rise as the W_R mass rises into this analysis' selections.

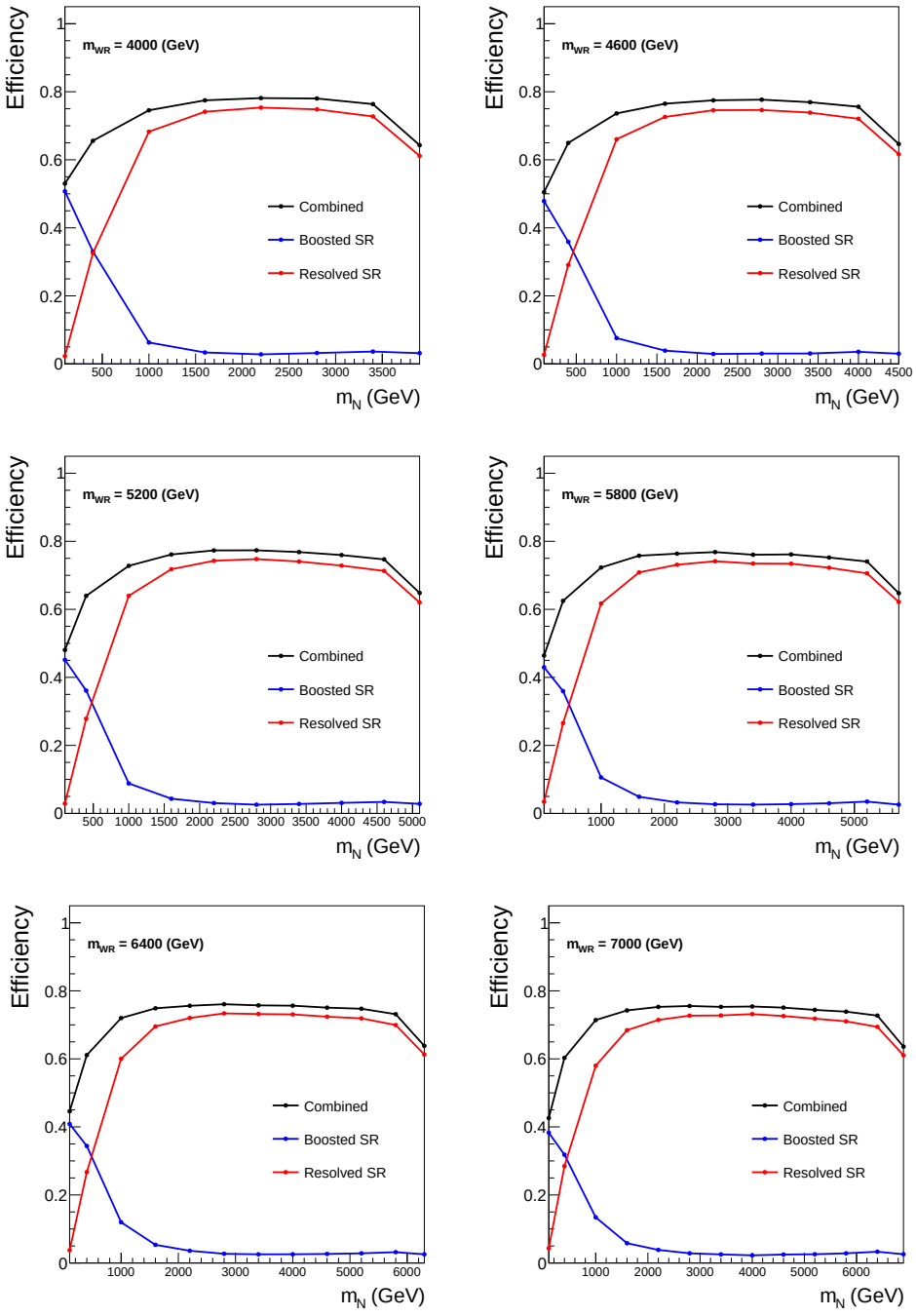


Figure 5.5: The signal efficiencies in the signal regions (SR) for W_R masses of 4000 GeV – 7000 GeV. In blue and red, the boosted and resolved signal region efficiency is shown respectively. The combined efficiency of each selection at a given mass point is shown in black. Efficiencies rise as the W_R mass rises into this analysis' selections.

5.3.1 Kinematic Sidebands

W_R Mass

The shape of $M_{\ell\ell jj}/M_{\ell J}$ for all backgrounds in this analysis is expected to approximately exponentially fall before and through the signal region. Therefore, this background behaviour can be studied in a region with a reconstructed W_R mass of less than 800 GeV. This forms the low-mass sideband of the analysis.

Di-lepton Mass

To form a sideband where the behaviour of Drell-Yan can be studied, the $M_{\ell\ell}$ mass requirement is changed to be less than 200 GeV in the resolved selection and less than 150 GeV in the resolved selection. This allows resonant Drell-Yan events to be studied and their behaviour extrapolated to the signal region.

Forming the low di-lepton mass selection region requires the loosening of additional requirements in the boosted selection. The sub-leading lepton is no longer required to be within the fat-jet. Requiring the second lepton to be within the jet in the low mass region almost entirely depopulates the region, severely limiting what can be studied in it.

5.3.2 Flavor Sideband

This analysis focuses on two lepton flavors, muon and electron. While, generally, any LRS extension would apply to all lepton flavors, the tau flavor is not searched for, as it poses several different, unique challenges [65]. Likewise, LRSM theories do not preference a right-handed neutrino mass-hierarchy. Therefore, there is no reason to expect the right-handed tau neutrino to appear in a similarly searchable mass region as the two flavors searched for in this analysis.

As signal events will produce two same-flavor leptons, this analysis selects background events which contain one electron and one muon as well. This selection produces an almost completely signal-free flavor sideband where lepton-flavor independent backgrounds can be studied, the most significant of which is $t\bar{t}$. While the low $M_{\ell\ell}$ region is dominated by Drell-Yan events, these events are unlikely to appear in the flavor sideband (as the Z boson produces same flavor leptons), unless the Z decays to the tau, and these independently to two different leptons. The Z decay to leptonically-decaying taus occurs at a much lower rate, however, than the second dominant background. A $t\bar{t}$ event has two independent W boson decays in its decay process. This means $t\bar{t}$ can be studied thoroughly in both data and

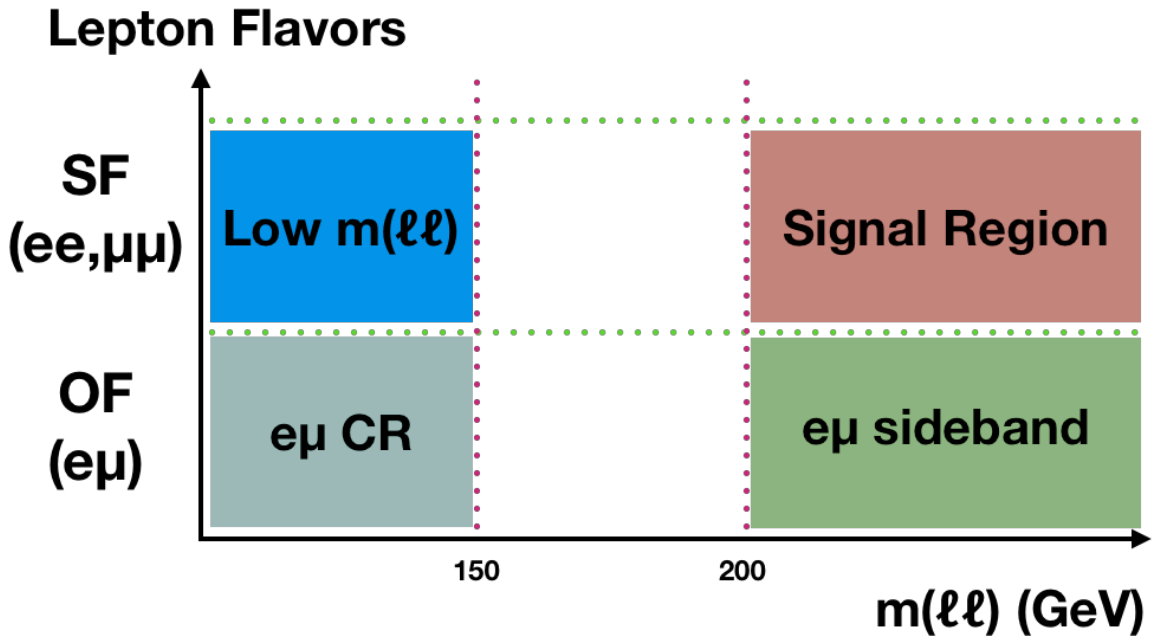


Figure 5.6: The multi-object mass and lepton flavor permutation are used to create four selection regions. This allows for the $t\bar{t}$ behaviour to be studied in both data and MC in the signal mass regime.

Monte-Carlo in this region. The flavor sideband also spans two kinematic selection regions, the low and high $M_{\ell J} - M_{\ell\ell jj}$ mass regions. A diagram of the four regions produced is shown in 5.6.

Chapter 6

Object Identification and Triggers

This chapter focuses on the definitions and identifications of objects used in this analysis. While the previous chapter focuses on defining the backgrounds and signal region studied in this analysis, the focus here is on the details of implementing this search in the physical CMS detector. Selecting events and reconstructing particles in real data, and comparing them to simulated events, pose many unique challenges. Covered in this chapter will be the trigger criteria used to select events for saving to disk at CMS, and the reconstruction, identification, and correction of jet, muon and electron objects used in this analysis. The process of creating these objects, called Particle Flow, is discussed in Section 3.2.1.

6.1 Jets

6.1.1 Jet Reconstruction

In the boosted analysis, a larger jet is used (AK8), with $R = 0.8$, and in the resolved analysis, a smaller jet is used (AK4), with $R = 0.4$. After the jet has been constructed

with PF particles, additional algorithms are used in an attempt to reduce the effects of pileup. For the AK8 jets, the pileup per particle identification (PUPPI) algorithm is used [66]. This weights each PF particle prior to being clustered in the jet based on a calculated probability of it coming from a pileup vertex. This is done based on several parameters, event pileup properties, local energy distribution, and tracking. The AK4 jets use charged hadron subtraction (CHS) where all charged hadrons which are not identified as coming from the primary vertex are removed from the jet.

6.1.2 Identification

This analysis uses the official recommended requirements for jet quality from the CMS JetMET group. The identification is designed to reject low quality jets. Jets that are not of interest to this analysis can be produced in various ways including: the overlap of multiple pileup particles, anomolous noise in subdetectors and non-hadronic lepton seeding. The jet quality requirements significantly reduce the presence of these jets in our selection. For the resolved analysis, we additionally veto jets which may be created by an energetic lepton and random PF particles around it. For AK4 jets the following requirements are used:

- the neutral and charged electromagnetic energy fractions must be less than 90%,
- the neutral hadronic energy fraction must be less than 90% (2016) or 80% (2017,2018),
- there is at least one charged hadron in the jet and the charged, hadronic energy fraction is greater than 0%,
- the muon energy fraction must be less than 80%.

For AK8 jets, largely the same requirements are used, but there is no veto on a jet seeded by a lepton. This increases our efficiency, as an energetic lepton is already required to be in the jet, and these requirements could conflict. The AK8 jet requirements are:

- the neutral and charged electromagnetic energy fractions must be less than 99% (for 2016),
- the neutral hadronic energy fraction must be less than 90%,
- there is at least one charged hadron in the jet and the charged, hadronic energy fraction is greater than 0%.

All AK8 and AK4 jets considered in the analysis must pass these identification requirements. With a quality selection of jets in each event, kinematic requirements can be placed on the jets to identify signal and background events as detailed in Chapter 5.

6.1.3 Corrections

The differences between the jet energy in simulation and data requires additional correction. These corrections are produced by a group within CMS which studies the performance of jet objects in the detector. Typically, the energy corrections are $2-3\%$. Jet energies in data and simulation differ by a scale-factor dependent on the p_{T} and η of the jets. The resolution of the jet energies are additionally different in simulation and data. As such, a per-jet random smearing is applied to simulated jets to adjust their resolutions into agreement with data.

The boosted analysis relies on a somewhat unique jet finding configuration, as the heavy neutrino is reconstructed as an AK8 jet as part of this analysis. Jet energy corrections are studied and calculated using a simulated QCD sample. The generated right-handed neutrino energy was compared to the energy of the AK8 jet reconstructed from it, both with and without jet energy corrections applied. This is shown in Fig. 6.1. It can be seen that the corrections have a minimal effect on the energy. The peak is in substantially the same place, and the relative broadness of the reconstructed energy is expected. The study and corrections applied are based on the 2016 detector conditions, but no significant changes were expected for 2017 and 2018 data, and they were not studied. Additionally, jet corrections can be correlated between years, and this is shown in Table 8.4.

6.2 Muons

6.2.1 Reconstruction

Muons in this analysis are required to be "global". Global muon reconstruction starts with hits recorded in the muon system and then looks for a matching track recorded in the tracker. There are several algorithms that perform this matching and more information on the muon reconstruction algorithms can be found in [67]. Each algorithm is best suited for a specific range of muon η and p_{T} values, so no single algorithm is always best for the high p_{T} muons studied in this analysis. This degeneracy in algorithm performance results from reductions in the tracker and muon detector momentum resolution as muon p_{T} increases. A specific combination of these algorithms, called TuneP, was chosen to determine the p_{T} of high energy muons. TuneP is used to recalculate all of the muons' p_{T} in this analysis.

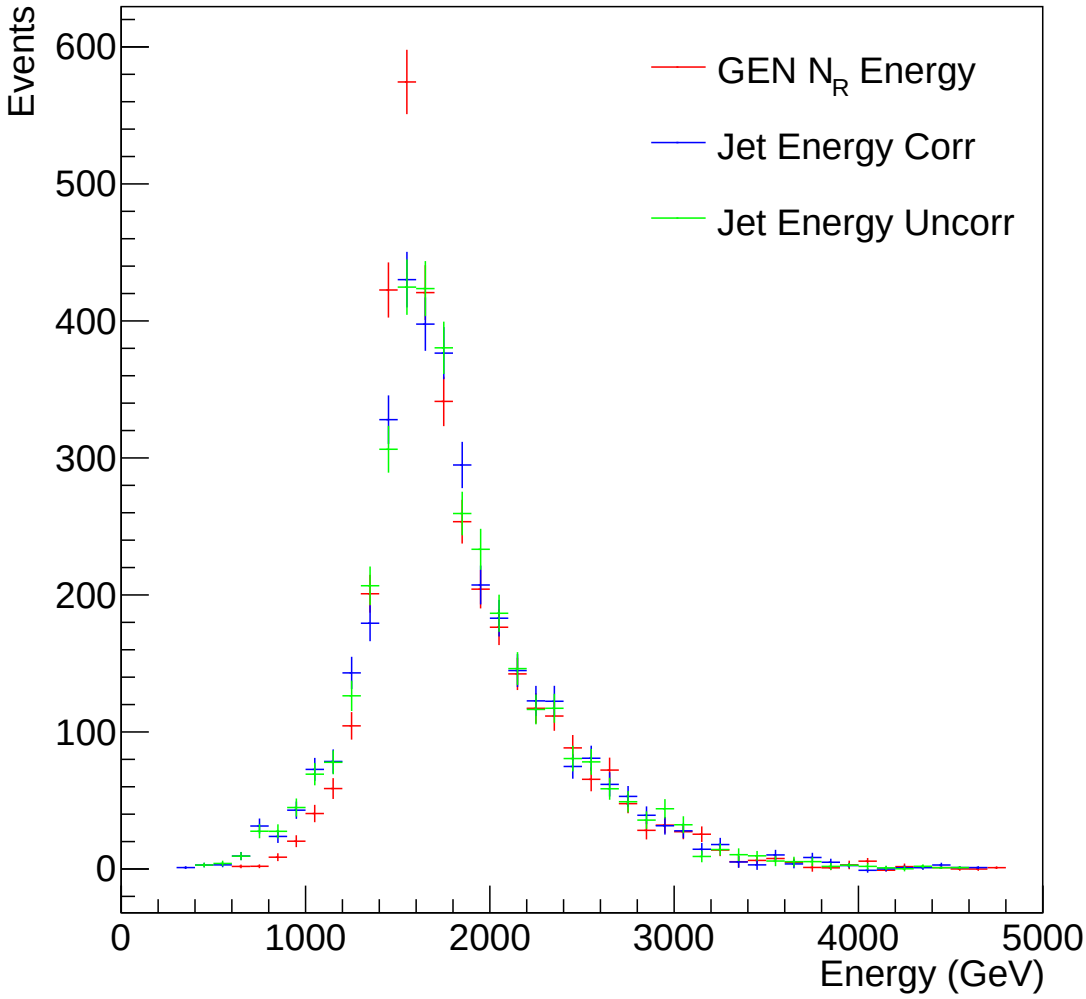


Figure 6.1: The generated right-handed neutrino energy is compared to the corrected and uncorrected jet energy for the AK8 jet reconstructing the right-handed neutrino.

6.2.2 Identification

Two types of muon identifications (ID) were used to select muons in this analysis. Isolated muons used in the resolved analysis, and the single isolated muon in the boosted analysis are all selected the same way. These are called "high-p_{T}" muons. The muon expected to be contained within the N_R jet is reconstructed with looser requirements as its reconstruction is expected to be more challenging. These are called "loose" muons. Loosening these requirements for this muon increases the acceptance rate of signal events without substantially increasing the background event rate. The high-p_{T} muon criteria are:

- The muon is reconstructed as a "global" muon.
- At least one muon-chamber hit is included in the global-muon track fit.
- There are muon segments in at least two muon stations.
- The p_{T} relative error, $\sigma_{p_{\mathrm{T}}}/p_{\mathrm{T}}$, of the muon best track, is less than 0.3.
- To prevent muons from cosmic rays and from decays of long lived particles, the transverse impact parameter must be less than 2mm with respect to the primary vertex. The longitudinal distance of the track must be less than 5 mm from the primary vertex.
- The muon track has at least one pixel hit.
- At least 6 tracker layer hits are required in the reconstruction.

The Loose muon ID criteria are:

- The muon is a particle flow muon.
- The muon is a "global" or "tracker" muon.

To reject muons from jets, each muon, other than the loose muon in boosted events, must be isolated from other tracks in the tracker. The energy of all tracks in a cone of $R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 0.3$ around the muon, excluding the muon track, must be less than 10% of the muon p_{T}. After applying these muon identification and isolation requirements, simulated events were re-weighted according to CMS prescriptions to account for differences in the ID and isolation efficiencies between data and simulations [68, 69].

6.3 Electrons

6.3.1 Reconstruction

Electrons are reconstructed from ECAL clusters matched to tracks made with the global-sum-fit (GSF) algorithm. The clusters in the ECAL are required to have a shower shape consistent with an electromagnetic interaction. The energy measurement of electrons in CMS is not perfect for various reasons. The measurements, however, can be corrected using the very precise measurement of the Z boson mass from other experiments. DY $\rightarrow e^+e^-$ events in CMS are studied to generate these corrections. More details on the electron reconstruction can be found in [70][71]. Monte-Carlo events must also be corrected to account for differences between simulations and data in the energy resolution of electrons. Electrons in MC events are smeared to account for their artificially high resolution. In total, electrons in MC are smeared, scaled, and the reconstruction efficiency is scaled based on the recommendations of the electron and photon (EGamma) POG [72, 73, 71].

6.3.2 Identification

Reconstructed electrons are required to pass the high-energy electron-photon (HEEP) identification [74]. This identification requires a reconstructed electron to contain a high quality, isolated track spatially linked to an isolated ECAL energy deposit. In addition, the shower shape of the ECAL energy deposit must be consistent with a true electromagnetic shower. The requirements of HEEP are specifically tuned for high energy electrons, typically with $E > 200\,\text{GeV}$. The electrons have two selection categories (tight and loose), chosen such that resolved and boosted signal events are optimally reconstructed. The "Tight" electron requirements are used to reconstruct both electrons for resolved events and the first electron in the boosted events.

Differences in electron ID efficiencies between data and simulation are taken into account by applying a scale factor provided by the EGamma POG [71]. Discrepancies in energy scale and resolution between data and simulation are corrected following the EGamma POG prescriptions for scales and smearings [72]. The electron energy scale was corrected in data, by a multiplicative factor dependent on both the η and R_9 of the electron. Where R_9 is the ratio of the middle and surrounding 3×3 ECAL crystals that the electron showers in. The electron energy in simulated events was smeared to take into account the effective resolution in data. A Gaussian smearing which depends on η and R_9 was applied.

The "Loose" electrons are used in boosted selection events to identify the electron lying

within the AK8 jet. As the electron reconstruction is more challenging with the surrounding jet constituents, the electron identification requirements are loosened to keep signal acceptance high. The requirements for electrons are summarized in Table 6.1.

Requirement	Loose	Tight
ID	Cut Based Loose without relIsoWithEA	HEEPV70

Table 6.1: Electron selection requirements

6.4 HEM failure

CMS relies on particle flow to leverage information from as many sub-detectors as possible. As such, the loss of sectors in the HCAL as result of the HEM failure can be mitigated to some extent. The momentum and direction of particles can be determined from the tracker, as well as the sign of the charge. Charged particles and photons interact in the ECAL and long-lived charged and neutral hadrons interact with the HCAL. Muon reconstruction typically relies on just the tracker and muon chambers, and are thus unaffected by the missing section of the HCAL. Hadronic activity, however, suffers reduced precision as a unknown fraction of the hadronized particle's energy will be unmeasured. While the process of showering in the detector is relatively well understood, the amount of energy that would be expected in the dark region of the HCAL must be extrapolated. Electron identification also suffers. There are many different properties that distinguish electrons in the detector and one key way they are distinguished from charged pions is that charged pions shower in the ECAL and HCAL. Electrons do not shower in the HCAL, as they are too light to penetrate this far. The ratio of energy in the ECAL with the HCAL sectors behind it is important to know, and impossible to extrapolate with the HCAL off.

To study the effects of the HEM failure on this analysis, the event rate in the low mass boosted signal region was studied. Both muon and electron flavor analyses were checked. The low mass signal region is the most similar selection to our actual signal region. Effects on this analysis were estimated to be related to the total number of events passing the analysis requirements as opposed to a change in reconstruction in the number of events. While it is probable that there is some level of effect from the HEM failure, as this analysis has very few events, any effect was determined to be smaller than the statistical precision available.

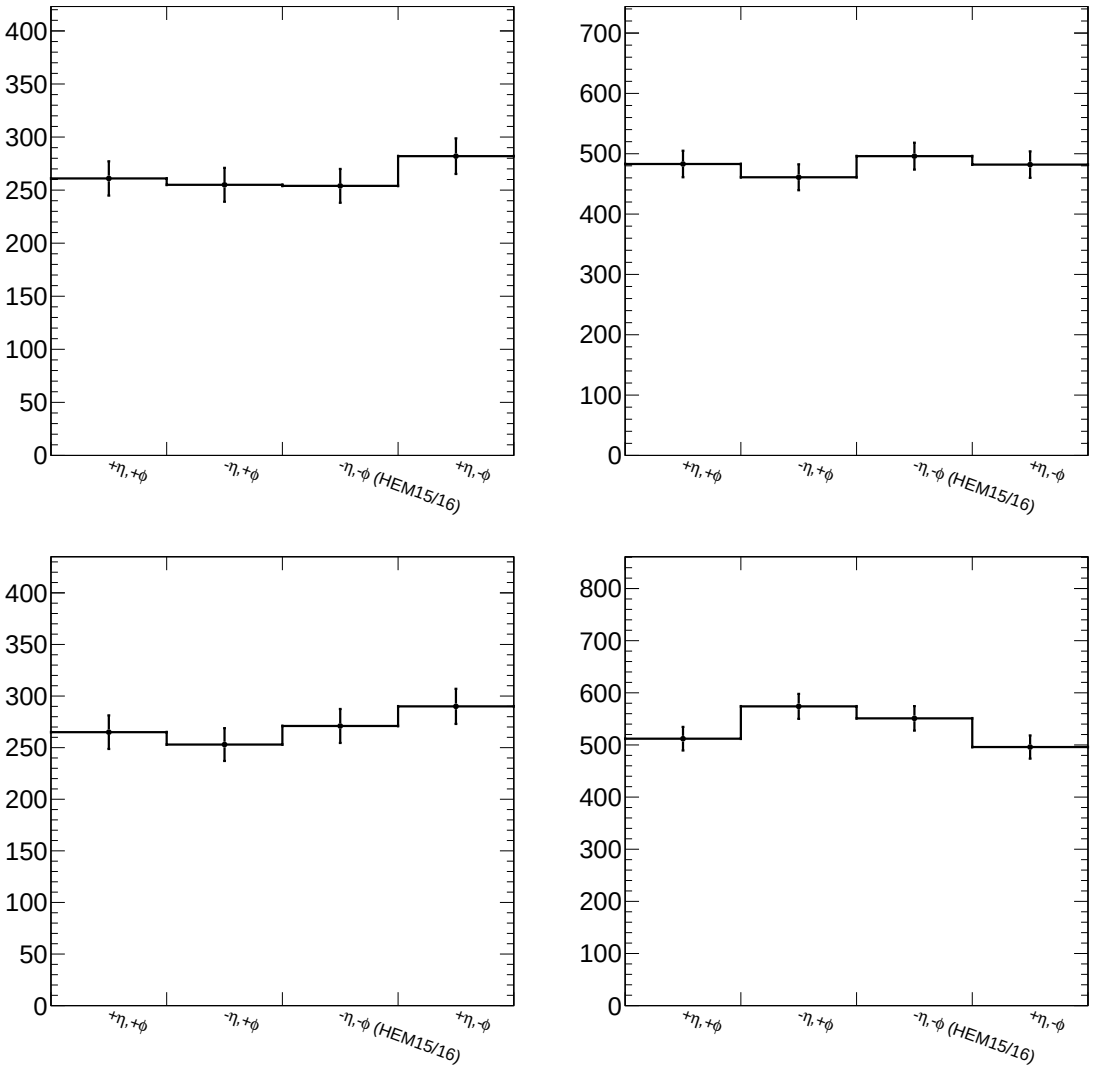


Figure 6.2: The number of events passing the low mass signal selection with a particle within one of the four endcap quadrants is shown. In both rows, the data taking period prior to the the failure (eras A-B) are shown on the left. On the right, (eras C-D) rates of events after the failure are shown. The muon flavor analysis is on top and the electron flavor analysis is on bottom. In all of these regions the HEM failure region does not have a statistically significant difference.

6.5 Triggers

The overarching goal of our trigger selection is simplicity. CMS defines hundreds of different requirement sets (triggers) to meet many needs. A trigger which only passes events with a very high muon p_{T}, for example, significantly reduces the amount of events needed to be sifted through. Trigger selections must be careful not to be so strict as to reduce signal acceptance, while too-accepting triggers are overall problematic for the detector readout, as only so many events can be saved in a given amount of time. Triggers accepting many events are generally pre-scaled. Pre-scaling introduces a random selection designed to lower the rate at which a trigger passes events without biasing the trigger.

Trigger requirements can be thought of as coming in two forms. The first kind of requirement focuses on the quality of the object being triggered on. Requirements can be imposed to filter out objects that may come from detector noise, pileup, or are difficult to reconstruct for some reason (e.g. a muon that travels through a gap in the detector). These requirements lower the trigger rate and help ensure little of CMS bandwidth is taken up by events that will eventually be rejected. On the other hand, these quality requirements will always reject "good" candidate objects at some rate and some quantity of truly interesting events will be lost to the bit bucket.

Another sort of object requirement, like transverse momentum, does not directly require any object quality. However, high transverse momentum objects are reliable indicators of a particularly hard interaction in the detector. As the momentum requirement is raised, this requirement serves as an increasingly reliable proxy for object quality. As an example, the probability of a fake muon having several hundred GeV of momentum is extremely small. The advantage of a momentum trigger requirement is that as the momentum increases you can reasonably expect the efficiency to asymptotically approach the maximum possible value, given detector coverage.

A balance, then, is needed between object quality and transverse momentum. Thus, the path forward is to choose the most accepting triggers which are not yet so accepting to be pre-scaled. Additional higher momentum (with less stringent identification) requiring triggers are then added to increase efficiency where a low momentum trigger may be lacking. This analysis uses single lepton triggers of corresponding flavor to the search region. Using a single lepton trigger is especially advantageous when multiple energetic leptons are expected in an event, as the combined probability of rejecting the event at the trigger level is much lower.

6.5.1 Muon Triggers

For simplicity, the boosted and resolved selection muon triggers are the same. For each of the years the triggers are only slightly different.

Table 6.2: Muon Selection Triggers

Year	Triggers
2016	HLT_Mu50_v* OR HLT_TkMu50_v*
2017	HLT_Mu50_v*
2018	HLT_Mu50_v* OR HLT_OldMu100_v* OR HLT_TK_Mu100_v*

The HLT_Mu50 triggers require a muon with at least 50 GeV to be in an event. There were some inefficiencies with the detailed requirements of this trigger in 2016 and 2018, leading to the addition of triggers to complement the selection. Each of these selections are the recommended triggers specified by the muon physics object group (POG) at CMS.

The trigger efficiencies used in our dimuon regions were officially measured by the muon POG as a function of the p_{T} and η of a muon passing the HighpT ID [75, 76, 77]. The efficiencies were measured in data and MC, and the ratio is used as a correction factor that is applied to MC. The η averaged p_{T} behavior of these triggers in MC and data, as well as the ratio of data over MC can be seen in Figure 6.3.

6.5.2 Electron Triggers

Electron trigger choice was more challenging than for muons. The lowest-p_{T} threshold unprescaled trigger does not have as high an efficiency at higher p_{T} as would be expected. Two additional triggers were added, one higher p_{T} electron trigger and one higher p_{T} photon trigger. Electrons and photons can be differentiated later with parameters not used by the trigger. In addition, part of the 2017 data was not taken with the HLT_Ele115_CaloIdVT_GsfTrkIdT trigger enabled, and so it is not included for the year. Omitting this trigger does make a small but distinguishable reduction in the overall trigger efficiency for a small mass region. Electrons falling in this region have less transverse momentum than a typical high $M_{\ell\ell jj}/M_{\ell J}$ event electron would have. As a result, the trigger change causes no perceptible change in the analysis performance.

Electron Trigger Efficiencies

The electron trigger combinations used in this analysis had not been officially studied. As it is impossible to guarantee that the behaviour of a trigger applied to MC and data will

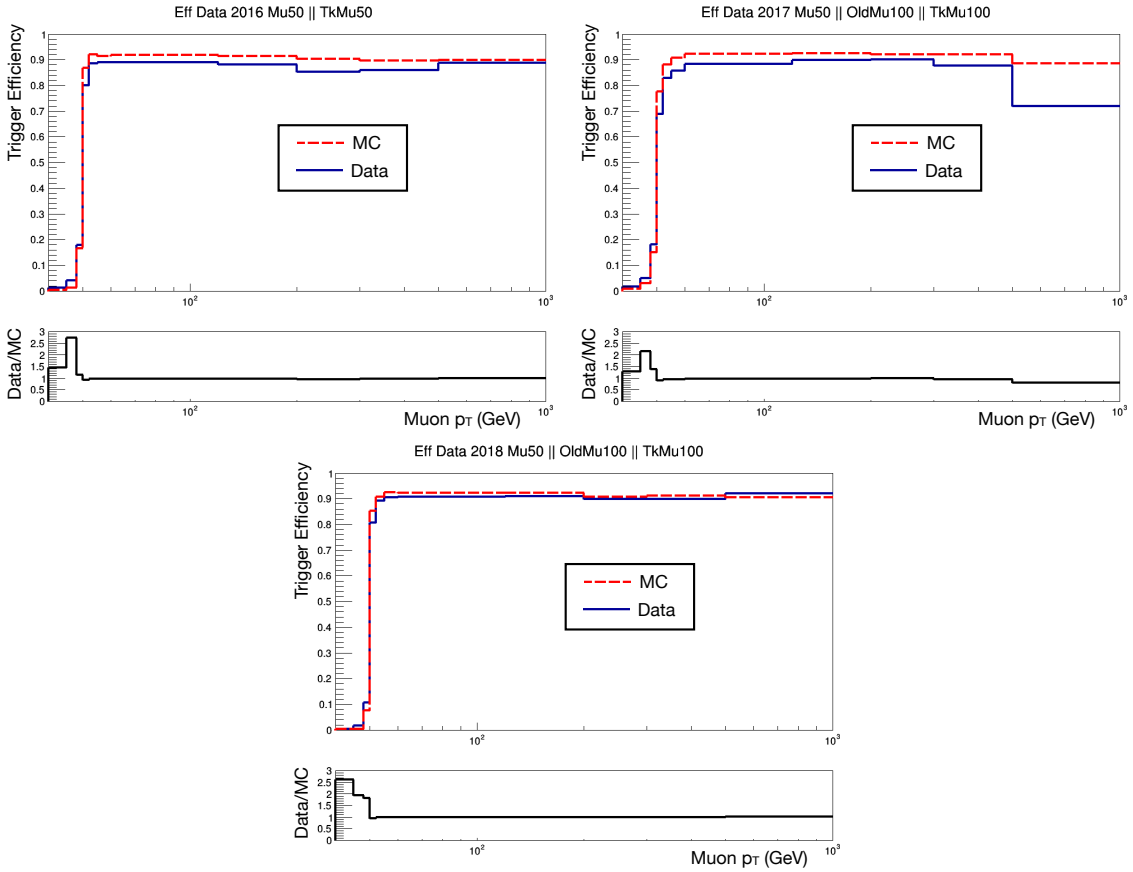


Figure 6.3: Muon trigger efficiencies are shown for the 2016 (top left), 2017 (top right), and 2018 (bottom) years. The trigger combinations do rise quite high, but never become fully efficient at the highest muon momenta. Muon momenta becomes increasingly difficult to measure as higher p_{T} as its path straightens in the detector.

be the same, all triggers, and trigger combinations, have to be studied. The efficiency of the trigger combinations was measured in data and MC and compared to produce a scale factor as a function of electron p_{T} and η. Graphs showing the comparison of data and MC in each of the years are shown in Fig 6.4. The η dependence has to do with whether an electron lands in the barrel region or the endcap region and so these two regions are shown in black and red respectively.

Level 1 Pre-Firing Trigger Inefficiency

An issue with the ECAL trigger system occurred during 2016 and 2017 due to the ECAL trigger primitives. Information packaged by a subdetector and sent to the Level-1 (L1)

Table 6.3: Electron Selection Triggers

Year	Trigger
2016	HLT_Ele27_WPTight_Gsf_v* OR HLT_Ele115_CaloIdVT_GsfTrkIdT OR HLT_Photon175_v*
2017	HLT_Ele35_WPTight_Gsf_v* OR HLT_Photon200_v*
2018	HLT_Ele32_WPTight_Gsf_v* OR HLT_Ele115_CaloIdVT_GsfTrkIdT OR HLT_Photon200_v*

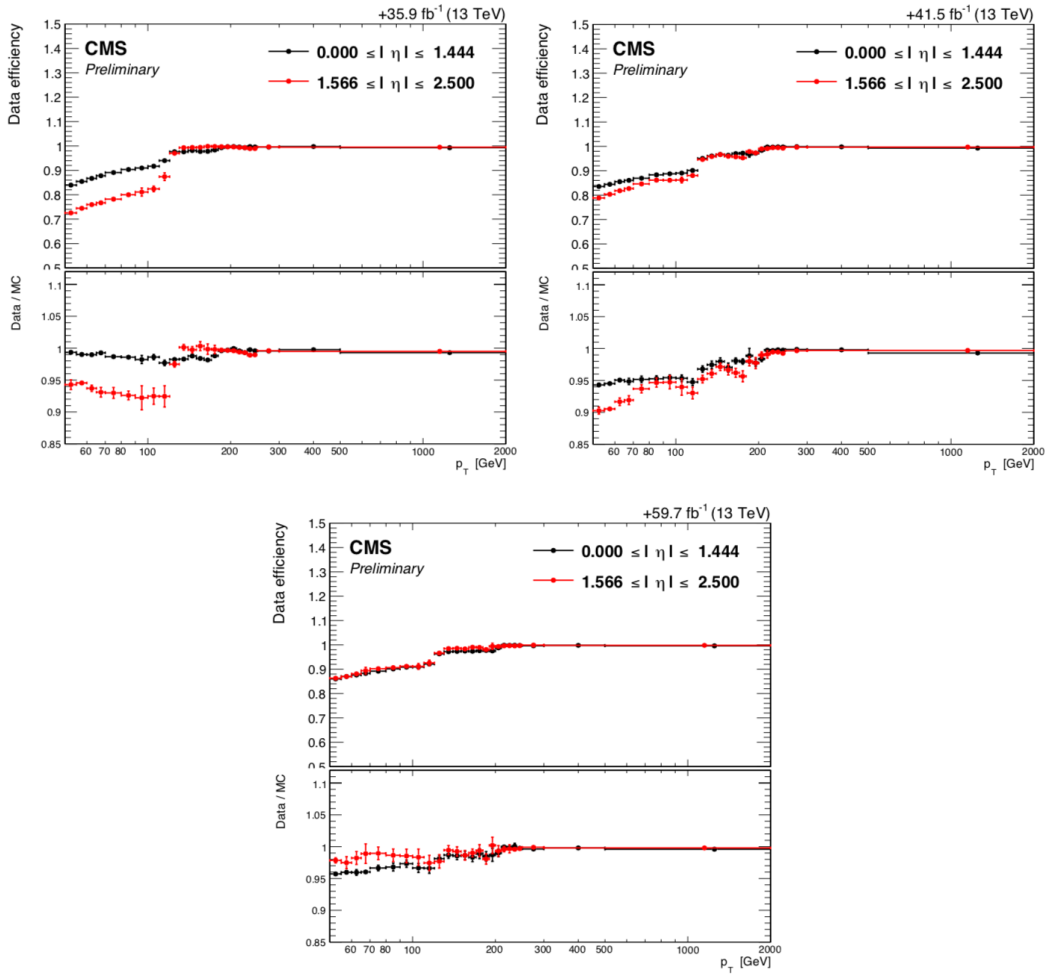


Figure 6.4: HLT trigger path comparison of data and MC. Top left is 2016, 2017 is top right, and 2018 is on bottom.

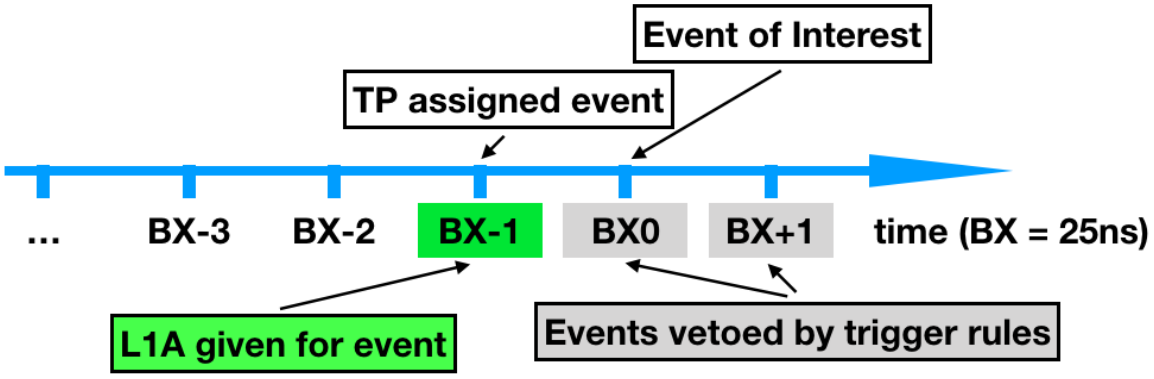


Figure 6.5: The event of interest is shown as BX0. Because of the timing shift, the TP is identified as being with BX-1, the L1A is issued for the wrong event, and the next two BXs are disallowed by trigger rules

trigger system are called trigger primitives. These form the view of the CMS detector at the L1 level. In the inner-most rings of the ECAL endcap, the timing of the detector began to drift, which increased the L1 pre-firing rate for all of the L1 triggers based on the calorimeters.

L1 pre-firing occurs when a L1 trigger fires based on the information from a trigger primitive (TP) which does not correspond to correct bunch crossing (BX). The trigger primitive produced by the ECAL would have come from a certain BX, but because of a timing error, the BX prior to the trigger event is actually read out and sent to the HLT. The CMS trigger rules, which are designed to prevent buffer overflows vetoes more than one L1 trigger acceptance (L1A) signal in three consecutive BXs. The L1A from the incorrect BX then prevents the correct BX from being read out, even if it could have passed for other reasons. A diagram of this is shown in Figure 6.5. While the timing drift causing the pre-firing is understood, there is no way to know on an event-by-event level if it occurred. There is, however, a way to guarantee that pre-firing did *not* occur for a give event, by leveraging the trigger rules. Events which come only 3 BX after a previous L1A cannot be affected by the pre-firing issue. This is shown in Figure 6.6.

Each event that is saved records not just the L1 trigger information for the event, but also for the two BXs before and after. This means the probability that an L1 pre-fire could have occurred, if it were not for the trigger rules, can be calculated for these un-pre-fireable events. Using a tag-and-probe technique, the probability that an ECAL interacting object could have its TP energy placed in the wrong BX can be calculated. This effect must studied for every analysis done involving the ECAL. As each analysis will have different

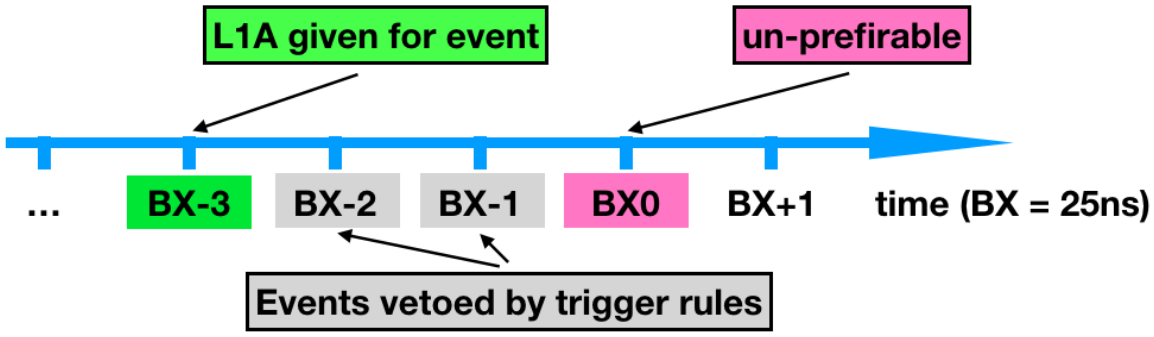


Figure 6.6: The event of interest is shown as BX0. Because of the trigger rules, if BX-3 has an L1A generated, BX-2, and BX-1 are ignored. This means that BX0 cannot be affected by the pre-firing issue.

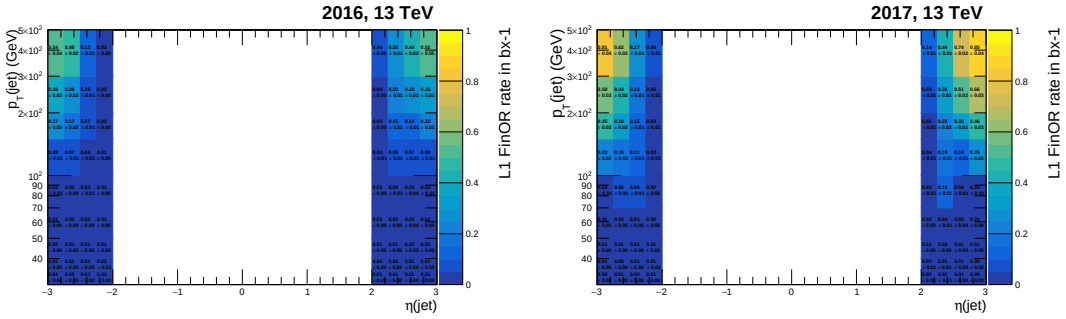


Figure 6.7: The jet pre-firing probability for 2016 (left) and 2017 (right).

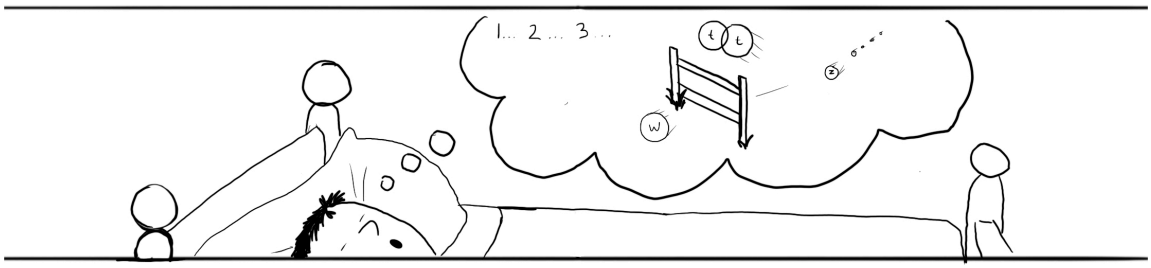
event requirements, the way the pre-firing issue affects it changes as well.

The pre-firing probabilities for this analysis are measured as a function of the p_{T} and η for jets, which deposit significant energy in the ECAL. The events studied must have exactly one muon which corresponds to a passing decision from the HLT IsoMu24/27. Electrons are vetoed. The probed jet is required to pass tight identification requirements and be isolated with $p_{\mathrm{T}} > 40\,\mathrm{GeV}$ as well as $1.75 < |\eta| < 3.5$, placing it in the endcap, where pre-firing can happen. The pre-firing probabilities for this are shown in Fig. 6.7.

These pre-firing probabilities are used to adjust the total jet rate in events taken in 2016 and 2017. As simulated events do not have the pre-firing issue, there is a mismatch of rates between data and MC.

Chapter 7

Background Estimations and Uncertainties



A discussion of the dominant backgrounds, their uncertainties and their estimations are inseparably linked to the complicated statistical calculations performed to estimate a limit on the cross-section of a possible W_R. This chapter serves as the first of two discussing the approaches of this analysis therein. Here, dominant backgrounds and their leading uncertainties are discussed in detail and the remaining uncertainties briefly explained. Section 8 delves into statistical techniques and calculations used in this chapter and presents the resulting limit. A complete list of all of the background simulations used in each of the three years is shown in Appendix B.

7.1 Background Estimations

The $t\bar{t}$, tW and Z/γ^* (Drell-Yan) processes can have the same final state as the signal and they are the main sources of background in both the resolved and boosted regions of this

analysis. In order to better-estimate the two dominant backgrounds, $t\bar{t}$ + tW and Z/γ^* simulation results are compared to and corrected using the two control regions defined in Chapter 5. There are several other backgrounds also accounted for in this analysis, however they occur at a low enough rate to both have a much smaller overall contribution and prove to difficult to directly study from collision data. These backgrounds are estimated directly from Monte-Carlo simulations.

As each year's Monte-Carlo has its own corrections and is partially correlated with the other years, each year's distribution for each background is handled separately for the purposes of limit-setting. However, where possible each year is shown in figures stacked with the others for simplicity. For a complete view of these distributions with each year separated, see Appendix A.

7.1.1 Drell-Yan

To estimate the background from high-mass or high-p_{T} Drell-Yan lepton pairs produced in association with additional jets, a leading order (LO) MC simulation is used. While this simulation produced ample events to make a statistically-precise estimate of the Drell-Yan background in our signal region, there are a few ways in which the simulation has been observed to be inconsistent with both high-order simulations and data. We discuss below the strategy used to correct the LO prediction with next-to-leading-order (NLO) simulation and data.

The $m(\ell\ell jj)$ and $m(\ell J)$ spectrum of all Monte-Carlo events in the Drell-Yan control region (described in Section 5.2.3) is shown in Fig. 8.4. In the Drell-Yan control region, the simulated p_{T} of the Z boson deviates from data. To correct this issue, a p_{T} correction was derived by reweighting Drell-Yan simulation based on NLO Drell-Yan simulation. While it is certain that the NLO simulation better agrees with measurements in the Drell-Yan control region, it is assumed that this relationship holds true in the signal region as well. NLO Drell-Yan simulations are not available with enough generated events to give a statistically-precise shape in our signal region. These corrections are shown in Fig. 7.1.

After these corrections an observed difference remains between the Drell-Yan and data in the Z-mass sideband in the resolved signal mass spectrum. The LO samples used in this analysis over-predict the data distribution noticeably. The difference between data and Monte-Carlo for both lepton flavors is shown in Fig. 7.2. As LO and NLO Monte-Carlo simulate a different number of partons, it is important to consider the NLO spectrum as well. The NLO and LO spectra (with and without Z-p_{T} reweighting) are shown in Fig. 7.3.

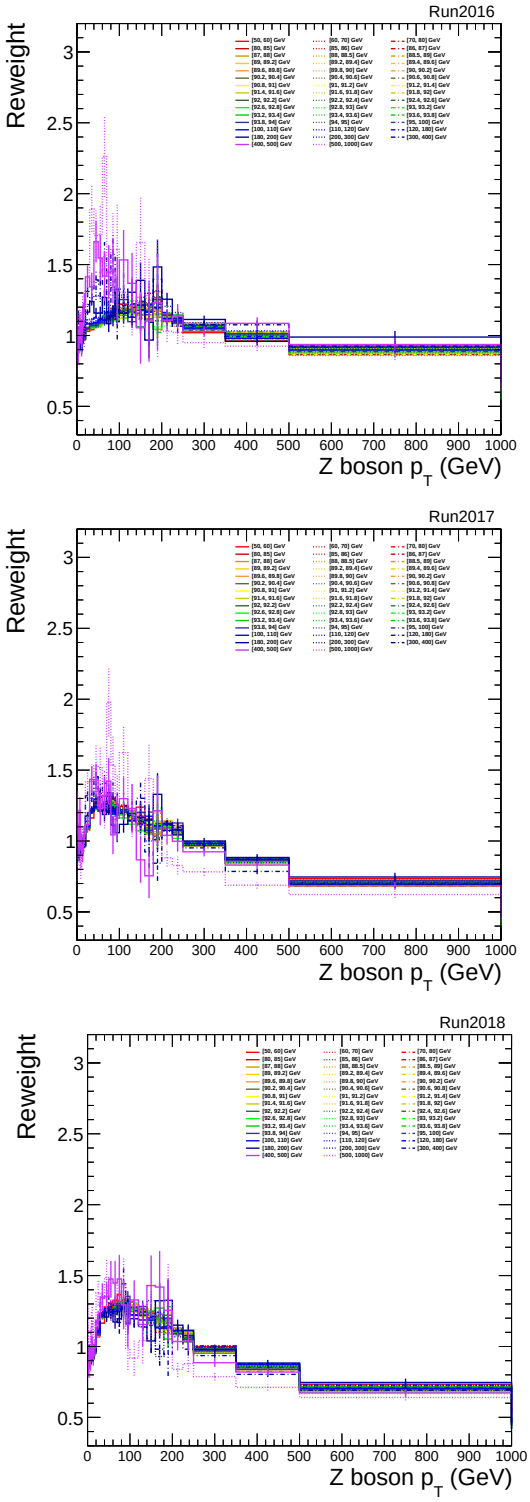


Figure 7.1: The Z-p_{T}-mass correction functions, for 2016 (upper), 2017 (middle) and 2018 (lower) [78].

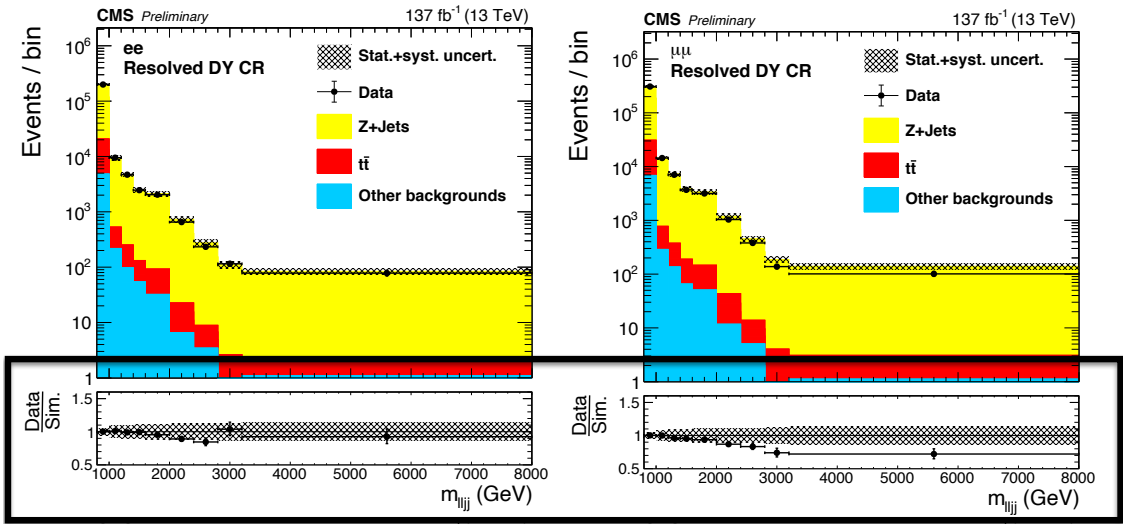


Figure 7.2: Data and Monte-Carlo in the Z-mass sideband are compared in the electron and muon flavors shown on the left and right, respectively. Data over Monte-Carlo is highlighted and shown in the bottom two plots.

The spectrum of the highest p_{T} jet in the selection region is shown alongside the signal mass spectrum. From the jet p_{T} distribution its clear that the LO jet distribution is consistently harder in the high p_{T} tail as might be expected. While there are differences between the shape of this difference between the two lepton flavor selections, these are assumed to be statistical fluctuations.

We hypothesize that the discrepancy in W_R mass observed is due to QCD effects which are independent of the Drell-Yan invariant mass and lepton flavor. A bin-by-bin correction factor (ξ_i) is calculated in this control region and applied to the LO Drell-Yan simulation in all regions. The correction factor is calculated by taking the ratio of Drell-Yan in data and simulation:

$$\xi_i = \frac{\mathrm{Data}_i - \mathrm{MC}_{other,i}}{\mathrm{DY_{LO}}}. \tag{7.1}$$

This ratio is calculated for each reconstructed W_R candidate mass bin (i), in each year, for the boosted and resolved Drell-Yan control regions. The muon and electron flavors are combined, as the jet hardness will be independent of lepton flavor. This correction is then applied to the Drell-Yan in the signal region.

After the shape-wise corrections have been made to each bin of the Drell-Yan simulation, the overall normalization of the Monte-Carlo is corrected to match the data yield. This normalization scale is fit in the control and signal regions simultaneously and at the

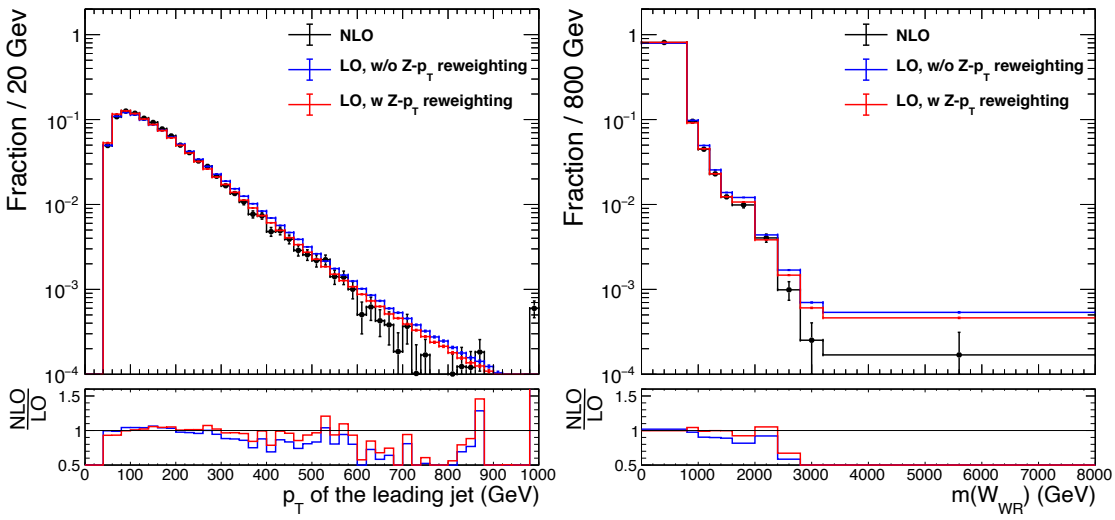


Figure 7.3: The spectra of simulated Drell-Yan are compared at LO with and without reweighting and NLO. The highest-p_{T} jet spectrum is shown on the left, and the signal mass is shown on the right.

same time as the $\mathrm{t\bar{t}}$ and tW simulation yield is fit. These are both discussed further in Section 8.1.2.

7.1.2 Top pair backgrounds estimation

Unlike Drell-Yan, higher-order corrections are not motivated in the $\mathrm{t\bar{t}}$, which can, anyway, be directly compared to the simulation in the flavor sideband. The tW process contributes a significant, though smaller, amount in each region the $\mathrm{t\bar{t}}$ is present and shares many of the same features, it is considered together with $\mathrm{t\bar{t}}$ in their estimation. The agreement between data and Monte-Carlo for them is as good as can be expected given the statistical precision possible. For this reason, only an overall correction is made to the total yield of the $\mathrm{t\bar{t}}$ and tW simulations combined. The background $m(\ell\ell jj)$ and $m(\ell J)$ spectra is shown in the flavor sideband control-regions in Fig. 7.5.

Rate-differences between data and Monte-Carlo are calculated by performing a simultaneous fit of the $\mathrm{t\bar{t}}$ and tW rate in the signal and flavor sideband. This scale factor is calculated independently in each year and for each signal-type hypothesis (muon or electron, boosted or resolved). This is discussed in more detail in Section 8.1.2.

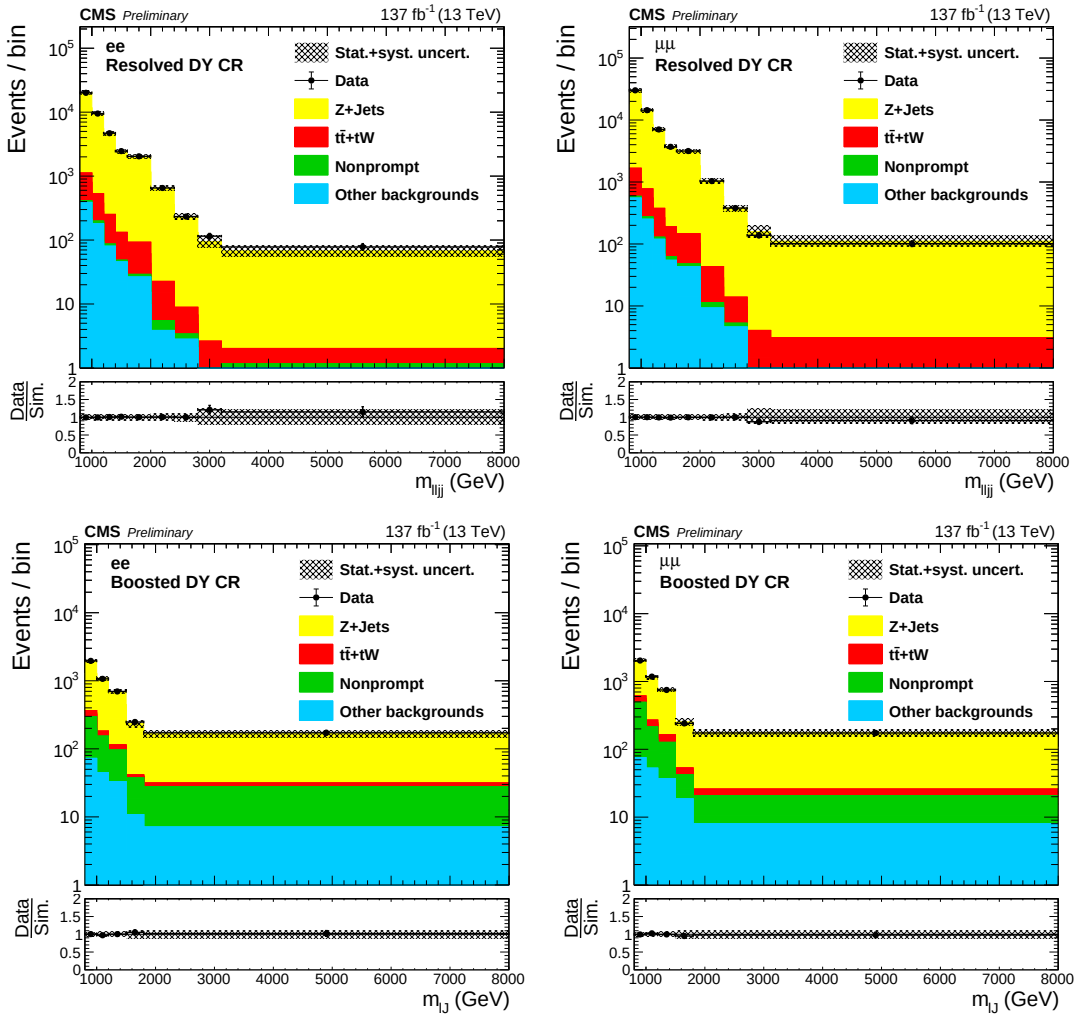


Figure 7.4: The pre-fit $m(\ell\ell jj)$ and $m(\ell J)$ distribution in the low $m_{\ell\ell}$ control regions with three years stacked. Results for the di-electron (di-muon) channel is shown on the left (right), for resolved (upper) and boosted (lower).

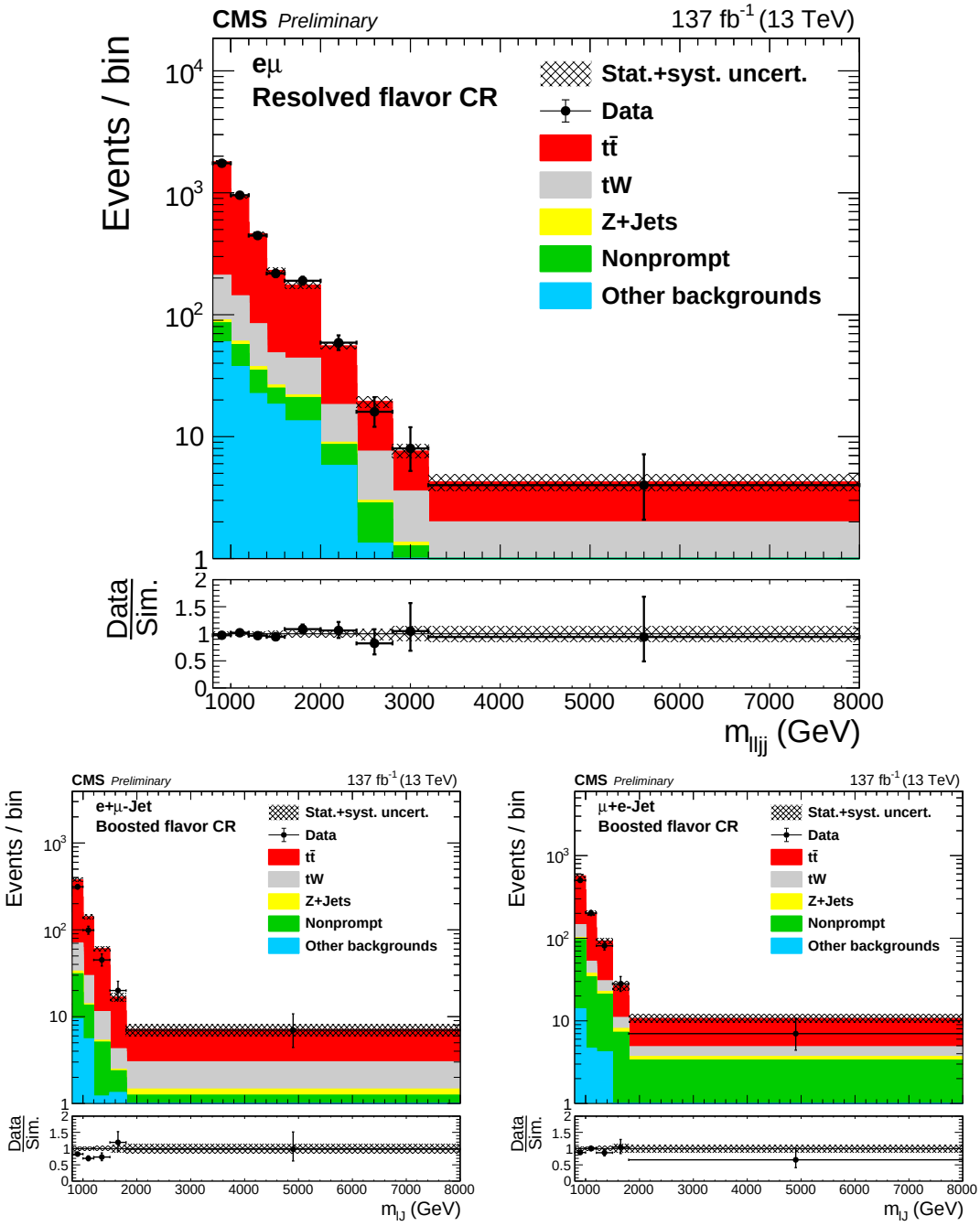


Figure 7.5: The pre-fit reconstructed mass of W_R candidates in the resolved flavor sideband (upper) and the electron(muon)-in-jet boosted flavor sidebands (lower), for all years combined.

7.1.3 Additional Backgrounds

Other, more minor, backgrounds are also estimated by this analysis. These backgrounds all occur at a lower rate than the two primary backgrounds, though the total amount of the other backgrounds can be more significant. Because of their lower rate, and in general, the difficulty of designing control regions for them, the estimates for the shapes of their contributions are taken from Monte-Carlo simulations. In the fitting process described in Chapter 8 the total rates of these backgrounds are fit in two categories, "other" (and always prompt) and "Non-prompt". All of these additional backgrounds are categorized and briefly discussed in this section. A complete list of the simulated backgrounds is in Appendix B. The fraction of each bin made up of a particular background category for each of the signal regions is shown in Fig. 7.6.

"Other" Prompt Backgrounds

A variety of less frequent standard model hard-processes contribute to the background of this analysis. These backgrounds produce the leptons required to pass as a signal candidate occur in the hard-process interaction itself. They are labelled as "Other" as they have one of the smallest contributions to the overall background.

- **tW:** Technically, the most dominant background in this category is the tW interaction. However, this process results in an array of final state particles fairly similar to $t\bar{t}$ production and is shown separately. In the current scheme, tW is combined with the $t\bar{t}$ background in data-versus-simulation comparisons. This background occurs at a similar rate in the flavor sideband and signal regions, but is not significant in the Drell-Yan control region.

- Multi-boson and ttV: At a much lower rate than tW production multiple weak boson can also be produced in collision. Any combination of W and Z bosons can be produced and could decay in fashion passing our signal selections. It is also possible for a vector boson to be produced in conjunction with multiple top quarks, which, likewise, could appear as a W_R candidate. None of these processes occur at a high rate in any of our regions and their combined distributions are referred to as "Other" backgrounds in this chapter.

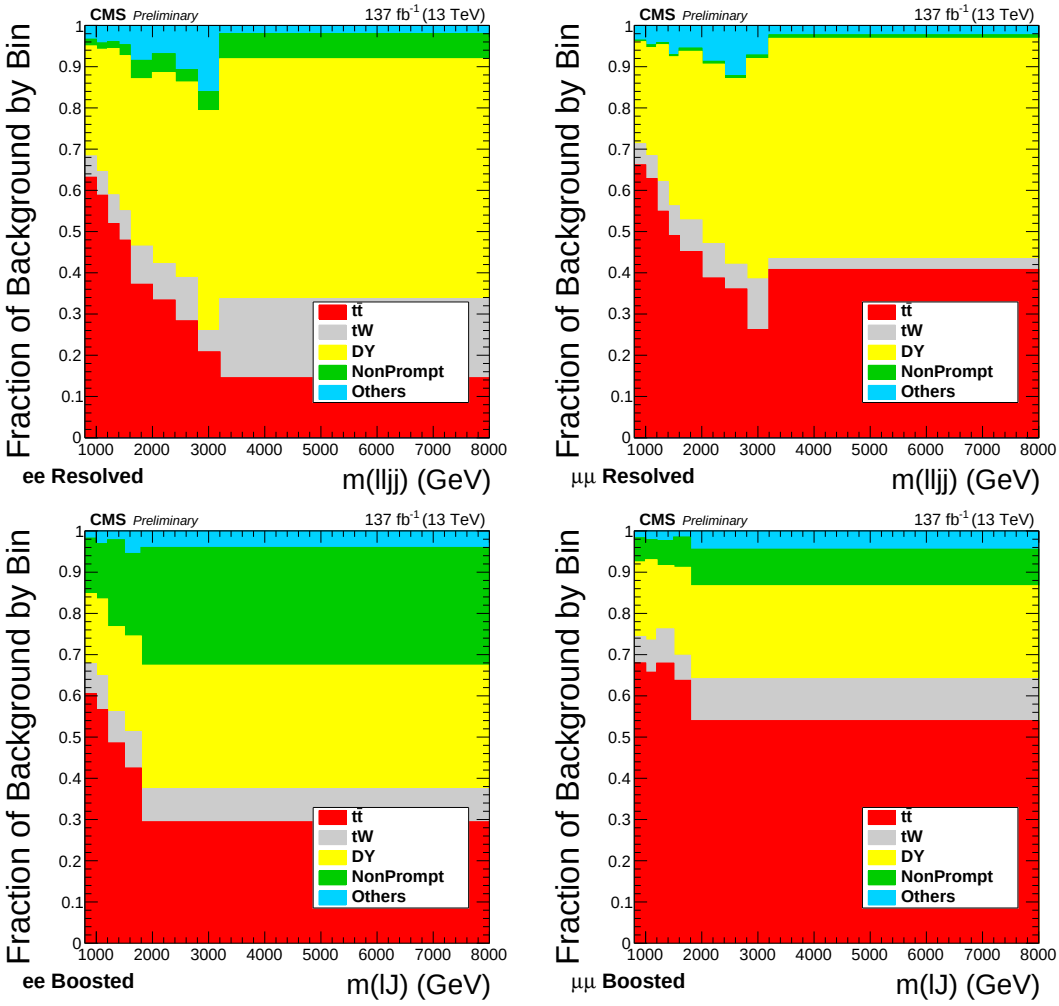


Figure 7.6: These figures show the fractional contribution of each background category in each bin. The electron (muon) analysis is on the left (right) and the resolved (boosted) analysis is on top (bottom).

Non-prompt Backgrounds

In addition to standard model interactions which can genuinely produce a W_R candidate signature, some backgrounds appear from hadron decays-in-flight and other mischaracterizations of an event. The most common of these in this analysis is the identification of an electron from what is likely a charged pion, or a charged pion decay-in-flight to a lepton. Semi-leptonic $t\bar{t}$ decays, single top production, or W boson production produce some of the necessary W_R signature, and can additionally fake an extra lepton.

The largest contributor of this type of background comes from the W boson decay as W boson production occurs at a much higher rate than the other non-prompt backgrounds. This faking typically produces an electron, as a muon interaction in the detector is easier to distinguish as muons are the only particle found in the muon chambers.

The most susceptible analysis regions are the flavor sideband and signal region of boosted electron analysis. This analysis region requires an electron within a jet with relatively loose requirements, providing a path for some rate of fake electrons. The fake-rate may not be reliably estimated in Monte-Carlo simulation, and this could effect the $t\bar{t}$ rate parameter fit. The W-jet process, in particular, could be mis-estimated by 50%. Estimating the fake-rate, and calculating a fake-rate uncertainty has not been attempted for this analysis yet. An eventual understanding of this process may reduce the statistical power of the boosted, electron region of this analysis.

7.2 Systematic Uncertainties

The systematic uncertainties in this analysis are meant to cover a wide range of discrepancies between data and simulation. This section divides the systematic uncertainties into categories. Each systematic uncertainty is ultimately propagated through to the final W_R mass distribution in various ways. As is discussed in Chapter 8, every systematic uncertainty is calculated in every signal region and sideband.

7.2.1 Background Driven Uncertainties

$t\bar{t}$

The dominant uncertainty for the $t\bar{t}$ background is on the normalization of $t\bar{t}$ simulation. This normalization is calculated simultaneously in each region, with the flavor sideband having the strongest influence.

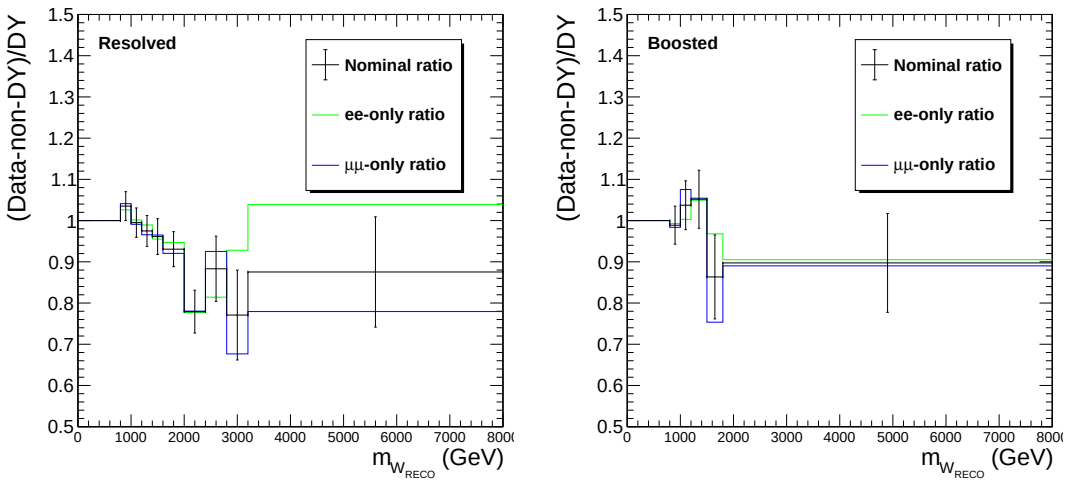


Figure 7.7: The (Data - nonDY simulation)/(DY simulation) obtained in the boosted and resolved Drell-Yan control regions for 2016. The error bars in black solid line indicate the statistical and systematic uncertainties of the data and simulations propagated to the ratios. The green and blue solid lines are the ratios obtained using ee and $\mu\mu$ data, respectively.

Drell-Yan

There are two dominant systematic uncertainties affecting the Drell-Yan simulations. The first is the uncertainty in the Z p_{T}-mass reweighting. This uncertainty varies, but is largest in the highest p_{T} and mass region at 30%. The second dominant uncertainty comes from the residual data and simulation discrepancy noticed in the control region. As mentioned earlier, the uncertainty on the nominal value of the ratio is calculated per bin and includes all of the statistical and other systematic uncertainties. This ratio and the full uncertainty on each bin of the ratio is show in Fig. 7.7.

7.2.2 Object Uncertainties

Systematic differences between simulation and data for the reconstructed jets, electrons, and muons in this analysis are only corrected up to a precision. Beyond this, systematic uncertainties remain for these objects. The event selection is re-run with a parameter (i.e. electron p_{T}) set at its upper and lower 1σ bounds. The final $M_{\ell\ell jj}/M_{\ell J}$ mass distribution is then used to see the effect of each parameter's uncertainty.

- **LSF scale factor**: The most significant object uncertainty in the boosted channel comes from constraining the agreement of the LSF variable between MC and data. Understanding this difference is complicated, as the background and signal produce

the reconstructed N_R jet from different processes. To estimate the effects, we artificially create a 3-prong, signal-like jet by starting with a hadronically decaying, boosted $W \to q\bar{q}$ and injecting either an electron or a muon into the event nearby it. This procedure is performed in data and $t\bar{t}$ simulation and the LSF distributions are compared to derive the scale factor for every year individually. This scale factor is created by taking the ratio of the LSF requirement efficiency on "signal-like" jets in data ϵ_{data} and simulation ϵ_{MC}:

$$\Gamma_{LSF} = \frac{\epsilon_{data}}{\epsilon_{MC}}. \tag{7.2}$$

This factor, Γ_{LSF}, is multiplied by the event weight of each simulated event. The data and simulation distributions of LSF with an injected lepton are shown for both lepton flavors in Fig. 7.8. The difference in efficiency between data and simulation on our LSF selection varies between 0.98 and 1.11 (between all years and lepton flavors) with an uncertainty between 5 and 9 %. The full results are shown in Table 7.1.

Table 7.1: LSF scale factors in each year from injected electron and muon samples.

Year	Injected e	Injected μ
2016	$1.04^{+0.09}_{-0.08}$	$1.01^{+0.06}_{-0.06}$
2017	$1.02^{+0.08}_{-0.08}$	$0.98^{+0.07}_{-0.07}$
2018	$1.11^{+0.08}_{-0.07}$	$1.06^{+0.06}_{-0.05}$

- **Lepton momentum scale and resolution**: The lepton momentum scale uncertainty is computed by varying the momentum of the leptons by their uncertainties.

  Neither of the lepton momentum scale and resolution uncertainties are dominant in their respective channels. Each of these corrections apply to the lepton's momentum value, not the event weight itself. As such, a few percent change in their momenta is very unlikely to move the event between analysis bins let alone cause it to fail event requirements.

  For muons with $p_{\mathrm{T}} < 200\,\mathrm{GeV}$, the Rochester corrections were applied to the muon momentum, which removes bias from detector misalignment or magnetic fields . Systematic uncertainties considered are follows; root-mean-squared (RMS) of pre-generated error sets, difference between results without Z momentum reweighting and variation of profile and fitting mass window, For muons with $p_{\mathrm{T}} \geq 200\,\mathrm{GeV}$, generalized–endpoint (GE) method was applied, and the uncertainties on the muon

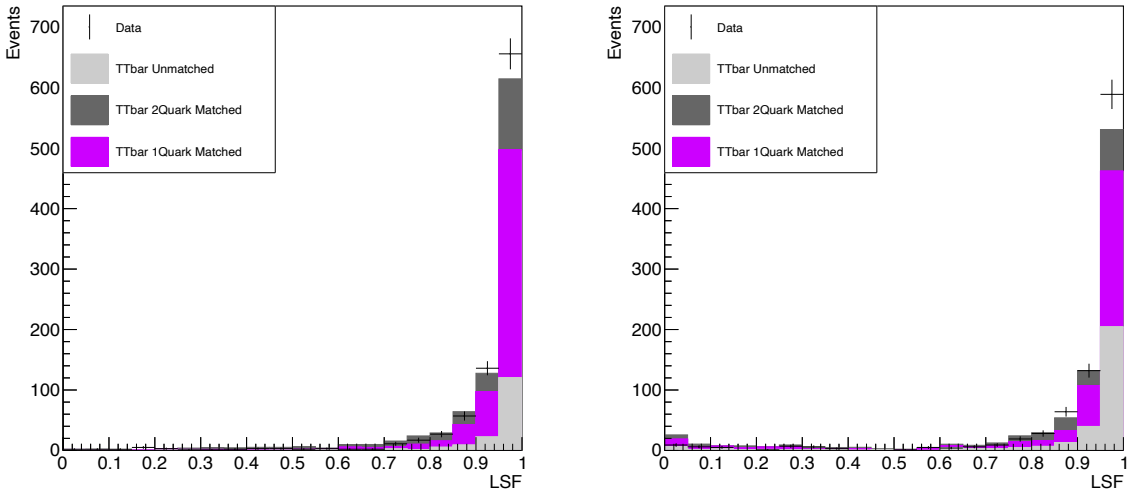


Figure 7.8: LSF is shown for data and simulation in 2018. The $t\bar{t}$ simulation is separated in to three categories dependent on how many of the decay-quarks are within the jet cone of 0.8.

curvature bias are taken from a gaussian distribution. Muon reconstruction and momentum scale give 0.4–1.0 (0–0.4) % and 0.4–2.4 (0.6–3.6) % uncertainties in the background estimation in the resolved (boosted) region.

For electrons, we used the MiniAOD V2 energy corrections [72], and corresponding uncertainties. Electron reconstruction [73], energy resolution, and energy scale [72] give 1.0–1.6 (0–0.5) %, <0.1 (<0.1) %, and 0.5–1.8 (0.5–2.6) % uncertainties in the background estimation in the resolved (boosted) region.

- **Lepton trigger and selection**: Discrepancies in the lepton reconstruction, identification, and isolation efficiencies between data and simulation are corrected by applying a scale factor to all the simulated samples. For the modified loose electron ID, the discrepancy between data and simulation is calculated as part of our LSF scale factor. The scale factors, which depend on the p_{T} and η, are varied by $\pm\sigma$ and the change in the yield in the signal region is taken as the systematic. Electron identificationand trigger give 3.1–3.3 (1.9–2.6) % and 0–0.1 (0.2–0.4) % uncertainties in the background estimation in the resolved (boosted) region. Muon identification, isolation and trigger give 0.2–1.2 (0.3–1.4) %, 0.1–0.2 (0–0.1) %, and 0.1–0.2 (0.6–1.0) % uncertainties in the background estimation in the resolved (boosted) region across all regions/years.

- **Jet energy scale and resolution**: Like with momentum corrections for leptons,

the jet energy scale and resolution apply to the overall jet energy. This does not significantly change the four or two-object mass spectrum. In order have the resolution in the simulation similar to that in the data the momentum of the jets is smeared as:

$$p_{\mathrm{T}} \to \max[0, p_T^{\mathrm{gen}} + c_{\pm 1\sigma} \cdot (p_{\mathrm{T}} = p_T^{\mathrm{gen}})], \tag{7.3}$$

in which $c_{\pm 1\sigma}$ are the data/MC scale factors, which are shifted by $\pm\sigma$.

This results in a systematic uncertainty of less than 1% for all masses.

7.2.3 Event Uncertainties

There are several uncertainties related to event weights applied to background and signal simulations. Their effects are generally small, and vary in application depending on the uncertainty.

- **Integrated luminosity**: The systematic uncertainty on the integrated luminosity are 2.5%, 2.3%, and 2.5% for 2016, 2017 and 2018, respectively [79, 80, 81]. The integrated luminosity is used to scale background and signal simulations for them to be compared to data. Each of the major backgrounds has additional scaling to data in the simultaneous fit making this uncertainty effectively irrelevant for backgrounds. Signal simulations, however, are scaled by luminosity and otherwise uncorrected to create the two-dimensional W_R–N_R mass limits. It remains, however, not a dominant uncertainty in those limits.

- **Pileup**: The number of collision vertices in every event varies. This distribution is modeled in simulation and compared with data for each year. An uncertainty is estimated by varying the minimum bias cross section of pp collisions at 13 TeV. This uncertainty is not very significant.

- **Theoretical uncertainties**: For signal simulation, the uncertainties on the rate and acceptance of the signal are derived from the variation of the QCD scale, the parton distribution functions (PDFs) and α_{s}. The PDF and α_{s} uncertainties for the MadGraph signal samples are estimated from the standard deviation of the weights from the PDF errorsets provided in the NNPDF3.1 parton distribution set. The procedure for estimating the uncertainties associated with the PDF follows the recommendations issued by the PDF4LHC group [82].

- **Pre-firing probabilities**: Followed by the recommendation pre-firing study group in CMS, a 20 % systematic uncertainty is applied in addition to the statistical uncertainty on the pre-firing scale factor. This represents an overall uncertainty of, at most, 1% on the number of background events.

7.2.4 Multi-Year Correlations

For this analysis, the three main data taking years of Run II are analyzed. Each year changes in the detector and LHC conditions occurred, some subtle, and others less so (like the HEM failure in 2018). Collision events between years are generally considered uncorrelated. However, as a conservative default, systematic uncertainties considered in this analysis correlate between years, while the statistical uncertainty of events in different years are completely uncorrelated. A few systematic uncertainties are uncorrelated. A full list of all the uncertainties and the range of their effects on event yields in different signal regions and years is shown in Table 8.4 after the fitting process.

7.3 Pre-fit Signal Region Distributions

The background control regions and the systematic uncertainties have been shown. Now it is time to reveal the most important region, the signal region. Combining all of the information above and the signal region, the model in Chapter 8 calculates the statistical significance of the results. The signal region distributions before the fitting process are shown in Fig. 7.9. The statistical significance of these results and the behaviour of the fitting is discussed in Chapter 8.

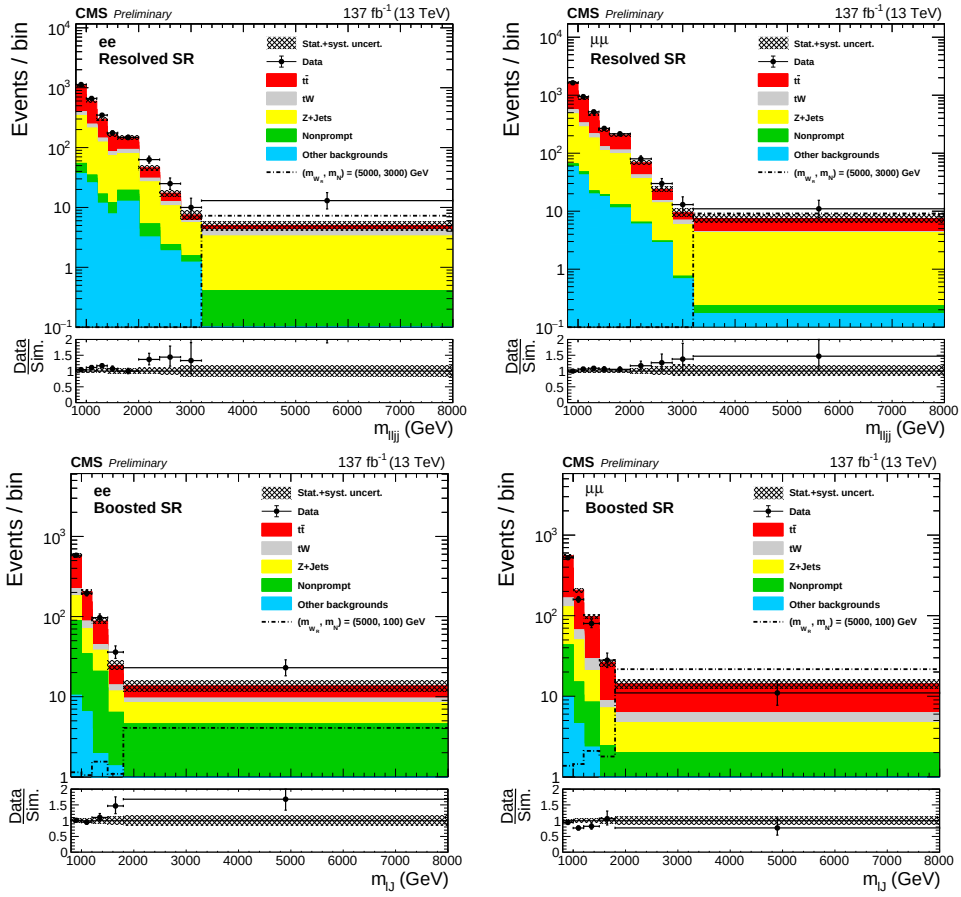
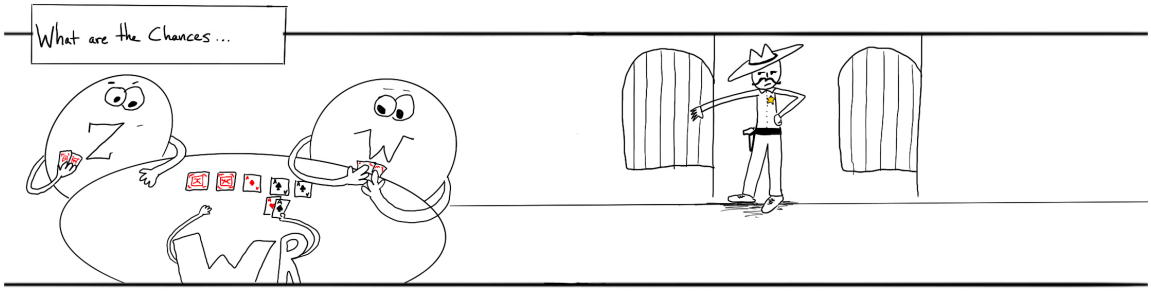


Figure 7.9: The pre-fit reconstructed mass of W_R in the resolved(boosted) signal regions is shown on top(bottom). Results for the di-electron (di-muon) channel are shown on the left (right) for all years combined.

Chapter 8

Hypothesis Testing



The final result of this analysis could be either a measurement of the cross-section for W_R production with subsequent decay for N_R or an upper limit on the cross-section, in the case no signal is observed. By convention, cross-section exclusion limits are set at the 95% confidence level for this analysis. The calculation of these levels and how the results are interpreted are discussed in this chapter. Limits on the cross-section are calculated use the CL_s technique, which is a modified frequentist approach and is the standard technique for limits set by experiments at the LHC [83],[84]. Results for the cross-section are presented one dimensionally as a function of W_R mass for both the boosted and resolved N_R case. Two dimensional limits are also shown with resolved and boosted treated separately and combined, where the model $g_R = g_L$ is assumed.

This analysis considers each bin in the final W_R mass distribution simultaneously. The binning choices for the resolved and boosted analysis are discussed in Chapter 7. Inter-bin correlations improve the ability to distinguish signal from background given the W_R mass's peaking behavior. For the heaviest W_R mass considered, however, the last bin influences the limit far more than any other.

8.1 Simultaneous Fits

All of the histograms and uncertainties which enter into the statistical calculations are shown in Section 7.3. Each lepton-flavor search has three regions, the flavor-sideband, the Drell-Yan control region, and the signal region. The handling of the yield of the Drell-Yan and $t\bar{t}$ + tW backgrounds is discussed in Section 8.1.2.

The fit is performed by minimizing the negative log of the likelihood of each of the rate parameter values simultaneously in each region and year. The agreement of every bin in each of the regions is compared with the measured value and uncertainty for that bin. The systematic uncertainties are modelled as Gaussian distributions, fully correlated between regions. The minimization process is described in detail in Section 8.2.1.

8.1.1 Post-fit Results

After fitting, the data and background-only signal regions are compared and the significance of an excess, or the strength of a limit can be seen. The number of expected and measured events with a W_R candidate mass higher than 3200 GeV (resolved) or 1800 GeV (boosted) as well as the uncertainty on the mean number of expected background events in each signal region is shown in Table 8.1.

Table 8.1: Combined number of events in the three highest mass bins for each channel

Year	Channel	Event type	DY	$t\bar{t}$ + tW	Nonprompt	Others	Total background	Data
2016	ee	Resolved	3.87 ± 0.16	2.36 ± 0.25	1.35 ± 0.73	0.56 ± 0.78	8.14 ± 1.16	10
		Boosted	5.52 ± 0.87	18.70 ± 1.82	7.27 ± 3.83	2.13 ± 0.72	33.62 ± 3.09	39
	$\mu\mu$	Resolved	5.87 ± 0.38	4.02 ± 0.26	0.21 ± 0.38	2.12 ± 1.16	12.23 ± 1.38	9
		Boosted	3.48 ± 1.02	21.77 ± 1.22	1.62 ± 0.35	2.04 ± 0.78	28.91 ± 2.12	27
2017	ee	Resolved	5.11 ± 0.27	3.82 ± 0.44	0.36 ± 0.23	1.32 ± 0.81	10.61 ± 1.19	11
		Boosted	7.37 ± 1.01	20.26 ± 2.44	7.64 ± 4.14	0.52 ± 0.14	35.79 ± 3.64	44
	$\mu\mu$	Resolved	6.66 ± 0.48	4.83 ± 0.33	0.08 ± 0.05	2.90 ± 1.34	14.47 ± 1.63	19
		Boosted	7.23 ± 1.38	27.17 ± 1.69	2.81 ± 0.89	2.02 ± 0.90	39.23 ± 2.62	46
2018	ee	Resolved	8.56 ± 0.83	5.13 ± 0.47	1.13 ± 0.98	1.91 ± 1.20	16.73 ± 1.88	27
		Boosted	15.96 ± 1.86	30.76 ± 3.57	15.92 ± 6.13	1.72 ± 0.64	64.36 ± 5.14	73
	$\mu\mu$	Resolved	9.35 ± 0.54	8.60 ± 1.01	0.13 ± 0.10	1.06 ± 0.57	19.13 ± 1.52	26
		Boosted	9.10 ± 1.11	37.17 ± 2.20	4.71 ± 1.10	1.16 ± 0.28	52.14 ± 2.84	46
Combined	ee	Resolved	17.54 ± 1.17	11.31 ± 1.00	2.84 ± 1.28	3.79 ± 1.83	35.48 ± 2.60	48
		Boosted	28.85 ± 3.23	69.71 ± 5.58	30.84 ± 8.51	4.37 ± 1.63	133.77 ± 7.69	156
	$\mu\mu$	Resolved	21.88 ± 1.17	17.45 ± 1.38	0.43 ± 0.41	6.07 ± 2.39	45.83 ± 3.06	54
		Boosted	19.81 ± 2.58	86.11 ± 3.72	9.14 ± 1.97	5.21 ± 1.94	120.27 ± 4.49	119

The measured and expected distributions post-fit for all years combined and each signal region is shown in Fig. 8.1. For a complete picture of the post-fit distributions in each year see Section A. It can be seen that both the muon flavor signal regions were consistent with the expected distributions within uncertainty. There is a small, approximately 1 σ excess in

the electron-boosted region, and a close to 2 σ excess over the expected number of events in the electron-resolved region. The excess in the resolved signal region warrants further study to exclude non-signal extra events (e.g. a detector issue). A brief discussion characterizing these events is in Section 8.3.4.

While the deviation from the expected number of background events is curious, the excess is not significant enough to perform a cross-section measurement. We instead produce signal upper cross-section limits. The statistical process for this is discussed in Section 8.2.1.

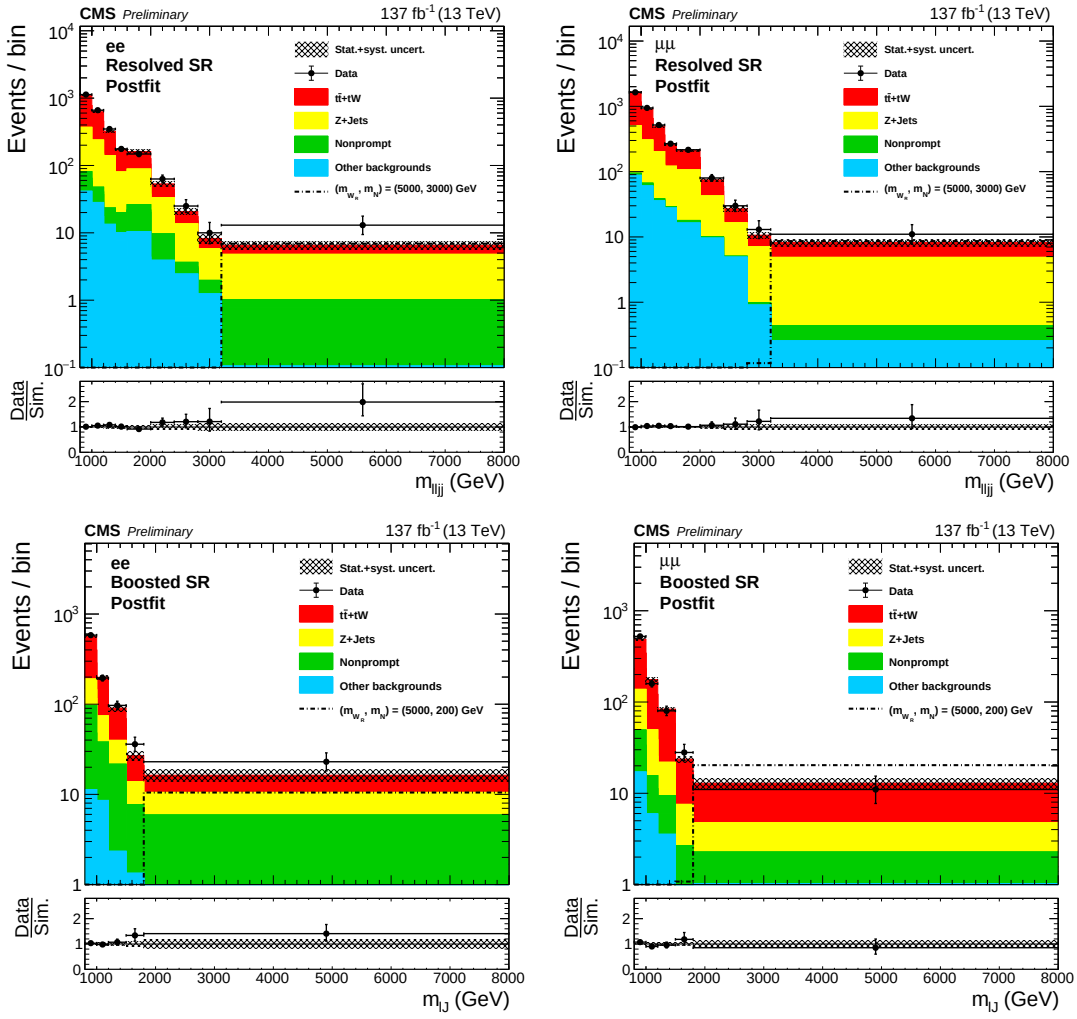


Figure 8.1: The post-fit reconstructed mass of W_R candidates in the signal regions with three year stacked. Results for the di-electron (di-muon) channel are shown on the left (right), for resolved (upper) and boosted (lower).

8.1.2 Simultaneous Normalization Derivation

The simulated yields of the Z/γ^*+Jets (f_{DY}) and $t\bar{t}$ + tW ($f_{t\bar{t}+tW}$) backgrounds as well as the "Nonprompt" (f_{NP}) and "Other" (f_{Other}) background yields and the signal strength (f_{signal}) are simultaneously fit in the control regions and the signal regions. As the control regions for Z/γ^*+Jets and $t\bar{t}$ + tW are relatively pure in that background, these regions constrain their background rates the most.

To calculate the yield scale-factor, the ratio of the total yield between data and Monte-Carlo simulation (f) is calculated by subtracting all other simulated distributions from the data and compared with the relevant simulation alone (Z/γ^*+Jets , $t\bar{t}$ + tW, Nonprompt, Other or signal). The simultaneous fit accounts for all of the statistical and systematic uncertainties affecting the signal and control-region ratios in determining the best fit. A diagram showing the different rate parameters calculated in each of the analysis regions is shown in Fig. 8.2.

Conceptually, two simultaneous fits are performed, one for the di-electron boosted and resolved selection and one for the di-muon boosted and resolved selection. However, the signal strength is also fit for each mass hypothesis. Though all signals are negligible in the control regions, both the di-muon and di-electron fits are performed for every signal mass hypothesis. This does not significantly change the fitted background rates and for simplicity, in both Table 8.2 and Table 8.3, a background only fit result is shown.

$$f = \frac{\text{Data} - \text{MC}_{\text{other}}}{\text{MC}_{\text{fitted}}}. \tag{8.1}$$

The normalization factors for $t\bar{t}$ + tW are shown in Table 8.3 and the normalization factors for Drell-Yan are shown in Table 8.2.

Table 8.2: The background-only fitted rate parameters for the Drell-Yan background.

Year	Event type	Fitted with ee signal regions	Fitted with $\mu\mu$ signal regions
2016	Resolved	1.00 ± 0.04	0.99 ± 0.04
	Boosted	0.99 ± 0.09	0.99 ± 0.08
2017	Resolved	1.00 ± 0.04	0.99 ± 0.04
	Boosted	1.00 ± 0.08	0.96 ± 0.08
2018	Resolved	1.00 ± 0.05	1.00 ± 0.05
	Boosted	0.98 ± 0.08	1.00 ± 0.07

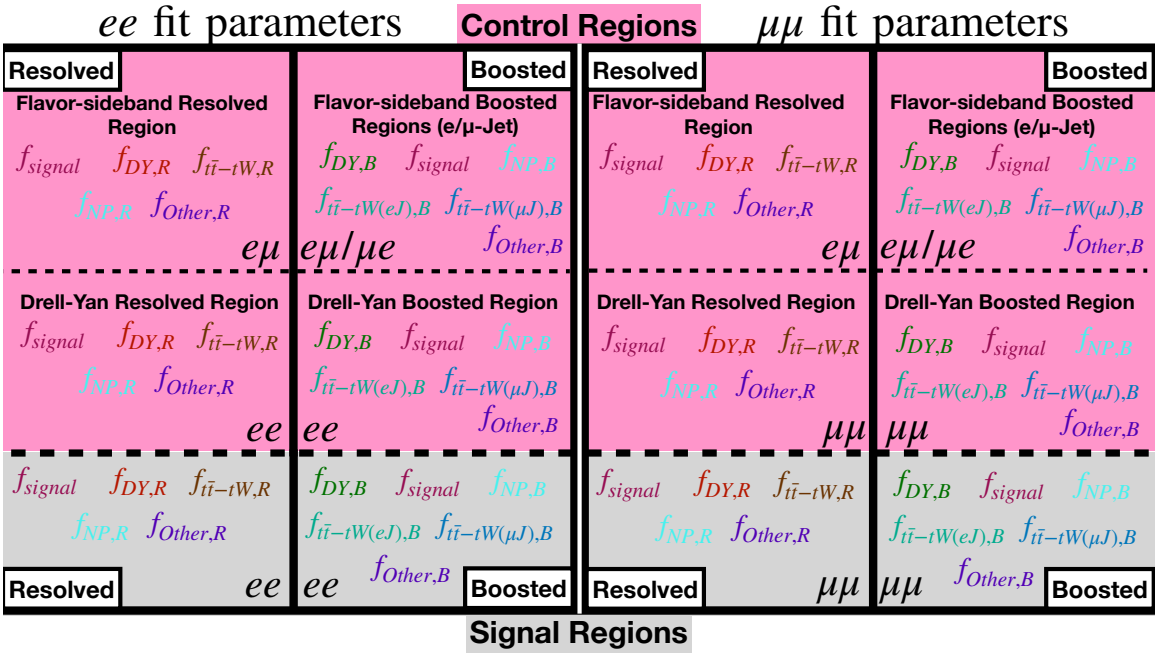


Figure 8.2: A diagram of the rate parameters is shown. Each rate parameter has a unique color and is simultaneously fit in each region it is shown in.

Table 8.3: The background-only fitted rate parameters for the $t\bar{t}$ + tW background

Year	Event type	Fitted with ee SRs	Fitted with $\mu\mu$ SRs
2016	Resolved	0.95 ± 0.05	0.92 ± 0.05
	Boosted with e-Jet	0.86 ± 0.14	0.84 ± 0.14
	Boosted with μ-Jet	0.75 ± 0.09	0.75 ± 0.08
2017	Resolved	1.05 ± 0.05	1.02 ± 0.05
	Boosted with e-Jet	1.04 ± 0.17	0.92 ± 0.17
	Boosted with μ-Jet	0.93 ± 0.12	0.91 ± 0.11
2018	Resolved	0.99 ± 0.05	0.98 ± 0.05
	Boosted with e-Jet	0.87 ± 0.14	0.86 ± 0.14
	Boosted with μ-Jet	0.64 ± 0.08	0.76 ± 0.08

8.1.3 Final Control Region Distributions

This section shows the final background distributions in the flavor sideband and Drell-Yan control regions. The agreement of the Monte-Carlo simulation and data in these distributions demonstrates the overall success of the analytical process this analysis undertakes. The flavor sidebands are shown in Fig. 8.3 and the Drell-Yan control regions are shown in Fig. 8.4.

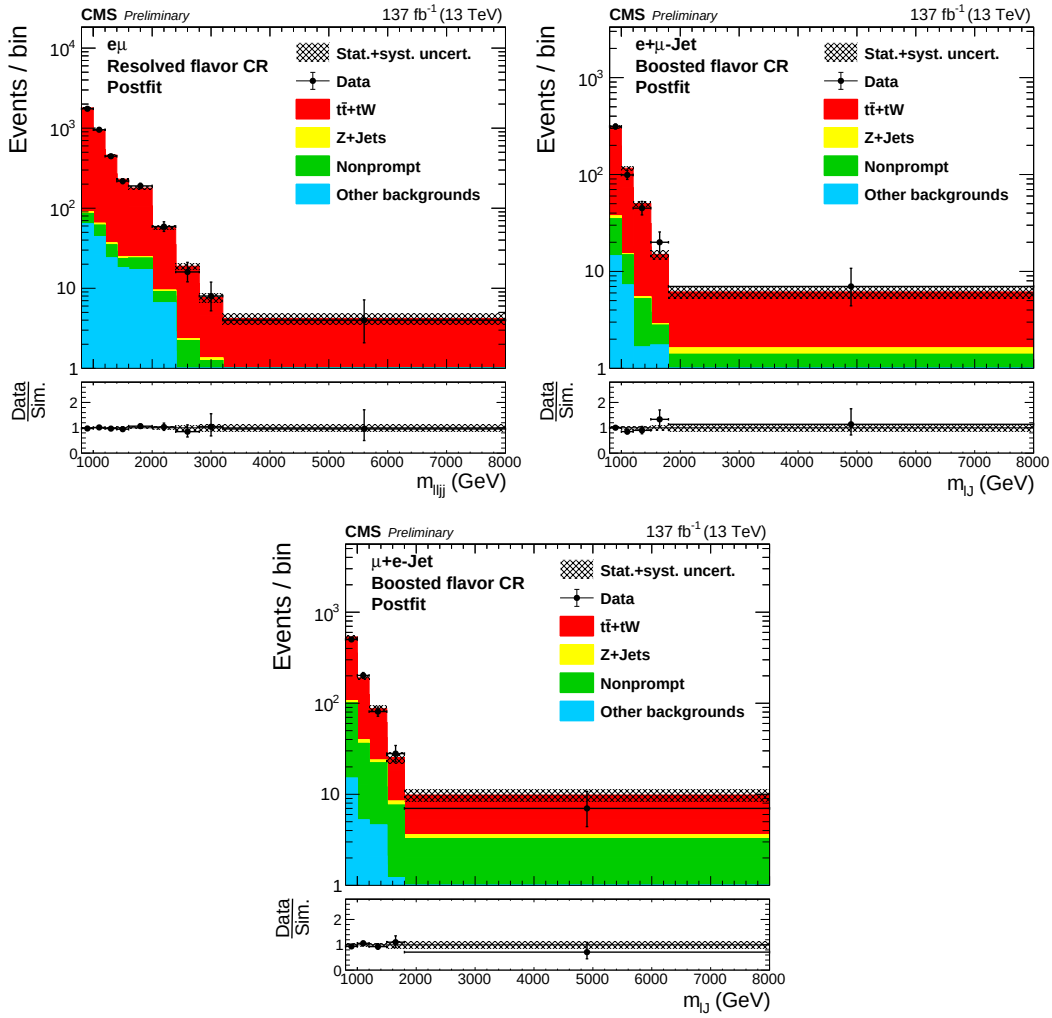


Figure 8.3: The post-fit reconstructed mass of W_R candidates in the resolved flavor sideband (upper) and the electron(muon)-in-jet boosted flavor sidebands (lower), for all years combined.

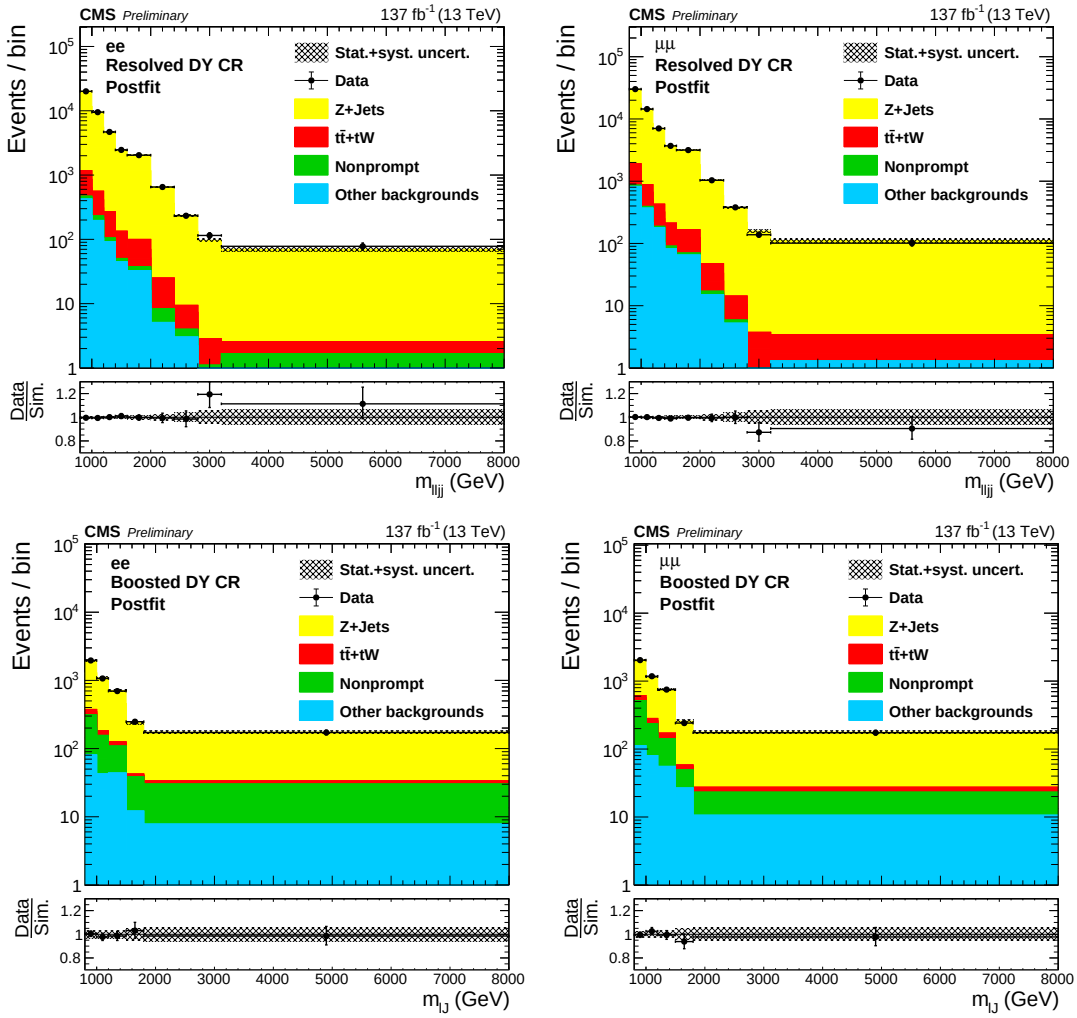


Figure 8.4: The post-fit $m(\ell\ell jj)$ and $m(\ell J)$ distribution in the low $m_{\ell\ell}$ control regions with three years stacked. Results for the di-electron (di-muon) channel is shown on the left (right), for resolved (upper) and boosted (lower).

8.2 The CL_s Technique

8.2.1 Overview

The CL_s technique is a modified frequentist method for setting exclusion limits or establishing observations. The frequentist limits used here answer the question, "How probable is this observation based on a given model". This is in contrast to Bayesian statistics, which calculate a probability of a model given certain results.

Fundamentally, we can consider two hypotheses. One is that the observed data results from standard model background events. This will be called the null hypothesis, H_0. The second hypothesis is that the data results from a combination of standard model background events and a new signal, H_μ. The variable μ will be used to represent the amount of signal strength. For this analysis, the signal would be a W_R boson. Now some variable must be selected to allow us to discern between the two hypotheses. This variable is called the test statistic. The total number of events measured in data after some event requirements is an example of a test statistic and is a simplified version of what this analysis uses.

The probability distribution of the test statistic, defined as $q(X)$, has to be estimated for each of the hypotheses, background only (H_0) and background + signal (H_μ). To estimate the test statistic's distribution for the two hypotheses, toy Monte-Carlo can be used to create many different possible outcomes for $q(X)$ including all of the uncertainties in the analysis. Once the probability distribution is determined, the probability that the result is caused by background only, H_0, is:

$$\mathrm{CL}_b \equiv \int_{q(X)}^{\infty} f(0)\,\mathrm{dq}. \tag{8.2}$$

The confidence level for the background only hypothesis, H_0, is CL_b. The probability that the null hypothesis explains a disagreement at least as large as the disagreement between the measured data and the expected background only result is defined as $1 - \mathrm{CL}_b$. This value is also called the "p-value".

The confidence level for the signal + background hypothesis, CL_{s+b}, can be calculated as:

$$\mathrm{CL}_{s+b} \equiv \int_{q(X)}^{\infty} f(\mu)\,\mathrm{dq}. \tag{8.3}$$

The values of CL_b and CL_{s+b} can be used to discern between the null hypothesis, H_0, and

the signal + background hypothesis, H_μ. The threshold for rejecting the null hypothesis, or the alternative hypothesis, were set by CMS before any analyses commenced. This removes a potential source of bias in the results. The famous "5σ" threshold is defined as $1 - \mathrm{CL}_b < 2.87 \times 10^{-7}$.

Often in searches for new phenomena, no clear evidence for discovery exists in data. Exclusion limits are calculated instead. These limits are calculated based on the confidence interval calculations defined above and are commonly set at the 95th percentile. An exclusion limit is designed to exclude signal by requiring that the probability that the observed data can be described by a background only is less than 5%. This works out to be $\mathrm{CL}_{s+b} < 1 - 0.95$. Calculating limits in this fashion gives problematic results when backgrounds are much larger than expected signal. The $\mathrm{CL_s}$ technique handles this by normalizing the signal + background confidence level with the background only confidence level:

$$\mathrm{CL}_S \equiv \frac{\mathrm{CL}_{s+b}}{\mathrm{CL}_b} < 1 - 0.95. \tag{8.4}$$

This gives the modified frequentist confidence limit, $\mathrm{CL_s}$. $\mathrm{CL_s}$ isn't a true confidence level as it is designed to give values relative to the background confidence level, which are by construction more conservative than CL_{s+b} alone.

8.2.2 Systematic Uncertainties

This section presents the systematic uncertainties in the analysis for all regions and years after the background rate fitting and the likelihood profiling. The uncertainties are summarized in Table 8.4.

8.3 Limits

8.3.1 Expected Limits Computation

Expected limits are calculated and compared with measured limits. The measured and expected limits represent two slightly different things. The measured limit is calculated, as described in Section 8.2.1. This represents the 95% upper confidence level for a signal cross-section. The expected limit, however, represents the theoretical-statistical power of the analysis. The expected limit is computed by randomly generating a simulated "data" sample from the pre-fit distributions in each of the control and signal regions. This "data" is then used in the exact same way as real-measured data in the fitting and hypothesis testing.

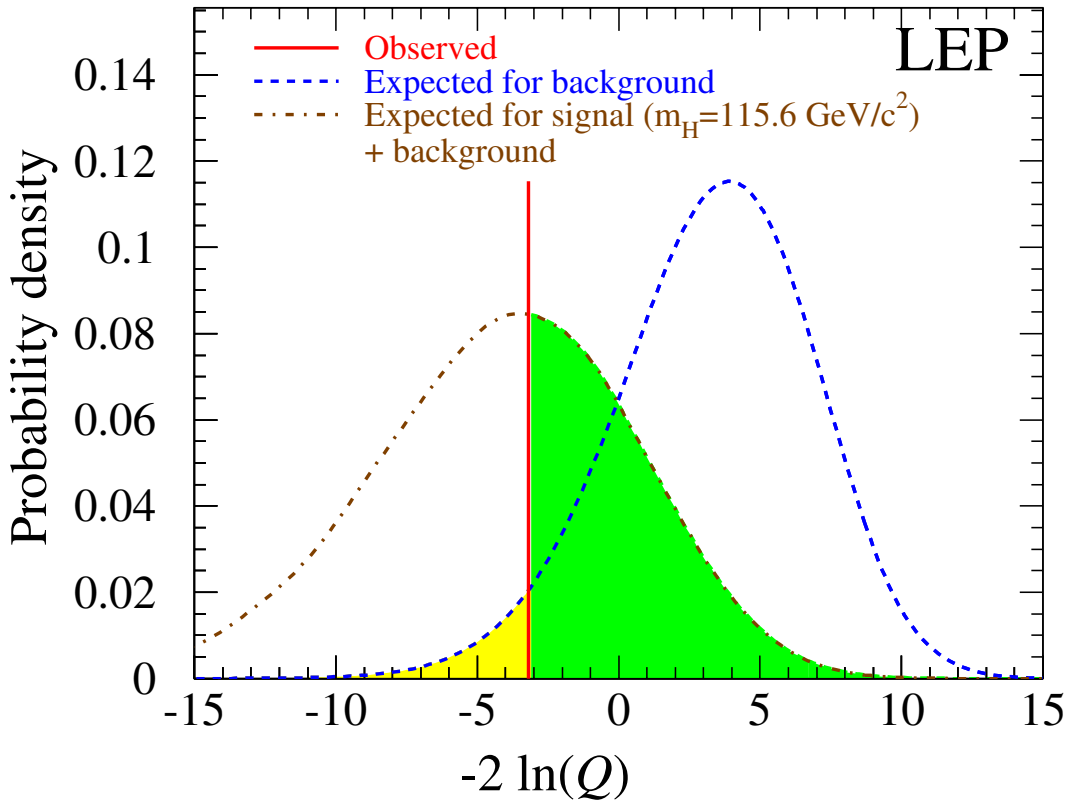


Figure 8.5: An example likelihood profile from a LEP era Higgs search [83]. This plot shows the probability densities for the combined Higgs search at LEP for the background (in blue, on the right) and a signal + background hypotheses (in gold, on the left) for a previously hypothesized Higgs mass of 115.6 GeV/c . The yellow region to the left of the observation is $1 - CL_b$ and the green region to the right of the observation represents CL_{s+b}.

Source	Bkgd./Signal process	Year-to-year treatment	ee bkgd. (%)	ee signal (%)	$\mu\mu$ bkgd. (%)	$\mu\mu$ signal (%)
Integrated luminosity	All bkgd./Signal	Uncorrelated	2.3–2.5 (2.3–2.5)	2.3–2.5 (2.3–2.5)	2.3–2.5 (2.3–2.5)	2.3–2.5 (2.3–2.5)
Jet energy resolution	All bkgd./Signal	Uncorrelated	0.5–1.4 (0.7–1.9)	0–0.3 (0–0.4)	0.2–1.2 (0.2–1.1)	0–0.3 (0–0.3)
Jet energy scale	All bkgd./Signal	Correlated	1.9–4.1 (0.9–2.0)	0–0.2 (0–0.3)	2.1–3.4 (0.6–1.0)	0–0.2 (0–0.4)
Muon reconstruction	All bkgd./Signal	Correlated	NA	NA	0.4–1.0 (0.3–0.7)	4.4–36.8 (5.6–30.7)
Muon momentum scale	All bkgd./Signal	Correlated	NA	NA	0.4–2.5 (0.4–3.6)	0.1–0.2 (0.1–0.3)
Muon identification	All bkgd./Signal	Correlated	NA	NA	0.2–1.2 (0.1–0.6)	0.2–1.1 (0.1–0.5)
Muon isolation	All bkgd./Signal	Correlated	NA	NA	0.1–0.2 (0–0.1)	0.1–0.2 (0–0.1)
Muon trigger	All bkgd./Signal	Uncorrelated	NA	NA	0.1–0.2 (0.1–0.2)	0.7–1.6 (0.5–1.3)
Electron reconstruction	All bkgd./Signal	Correlated	1.0–1.6 (0.5–0.8)	0.8–1.4 (0.4–0.8)	NA	NA
Electron energy resolution	All bkgd./Signal	Correlated	< 0.1 (< 0.1)	< 0.1 (< 0.1)	NA	NA
Electron energy scale	All bkgd./Signal	Correlated	0.5–1.8 (0.5–2.3)	0–0.3 (0–0.5)	NA	NA
Electron identification	All bkgd./Signal	Correlated	3.1–3.2 (1.8–1.9)	4.1–4.4 (2.1–2.4)	NA	NA
Electron trigger	All bkgd./Signal	Uncorrelated	0–0.1 (0.2–0.4)	< 0.1 (0.1–0.2)	NA	NA
LSF scale factor	All bkgd./Signal	Uncorrelated	NA (7.2–8.7)	NA (7.2–8.7)	NA (5.7–7.1)	NA (5.7–7.1)
Pileup modeling	All bkgd./Signal	Correlated	0.2–1.1 (0.5–1.1)	0.1–0.8 (0.2–0.9)	0.3–0.5 (0.3–1.1)	0.1–0.5 (0–0.6)
Prefire reweighting	All bkgd./Signal	Correlated	0–1.4 (0–1.1)	0–0.8 (0–0.9)	0–0.5 (0–0.4)	0–0.4 (0–0.2)
Z p_{T}	Z/γ^*	Correlated	2.6–3.3 (2.7–3.5)	NA	2.7–3.1 (2.8–3.4)	NA
DY reshape	Z/γ^*	Correlated	3.9–4.6 (4.6–5.5)	NA	4.0–4.6 (4.6–5.5)	NA
Nonprompt normalizaion	Nonprompt	Uncorrelated	100 (100)	NA	100 (100)	NA
Rare SM normalizaion	Others	Correlated	50 (50)	NA	50 (50)	NA
PDF error	Signal	Correlated	NA	5.9–11.1 (8.8–39.9)	NA	2.8–6.8 (17.5–40.6)
α_{s}	Signal	Correlated	NA	0–0.2 (0.2–1.3)	NA	0–0.2 (0.2–1.2)
renormalization/factorization scales	Signal	Correlated	NA	0–0.1 (0.3–2.3)	NA	0–0.1 (2.1–2.9)

Table 8.4: A post-fit summary of the relative systematic uncertainties in signal and background. The uncertainties are given for the resolved (boosted) SR. The numbers for signal is obtained for $m_{W_R} = 5\,\mathrm{TeV}$. The range given for each systematic uncertainty source covers the variation across the years.

By generating many "data" simulations, nominal and upper-lower confidence intervals are produced representing how sensitive the analysis is. This represents the expected limit.

8.3.2 One Dimensional Limits

At the 95% confidence level, an upper bound is placed on the production cross-section of the W_R multiplied by the branching fraction of its decay to two leptons and two jets. To construct one dimensional (1D) limits, the relationship of the N_R to the W_R mass is fixed. For the resolved 1D limit, the N_R is assumed to have a mass of half of the W_R mass. For the boosted limit, the N_R mass is fixed at 100 GeV. This is the lowest mass N_R considered at each W_R mass point, and represents the most boosted N_R case for each W_R mass. The expected limit contours at 68% (green) and 95% (yellow) are shown along with the measured limit, and the theoretical cross-section times branching fraction for the $g_{\mathrm{L}} = g_{\mathrm{R}}$ hypothesis. As the W_R mass increases, the background falls drastically, and the cross-section limit improves significantly. At approximately 3 TeV, all signal regions have one bin that continues to infinity. Any W_R hypothesis with its signal peak falling within this last bin would represent the lowest possible limit, however, since W_R mass increases there is some reduction in signal efficiency (seen in Figs. 5.4, 5.5) due to various factors. This causes the limit to worsen at the upper end of the W_R mass.

All regions in the boosted and resolved analysis are considered in the fit and limit-setting

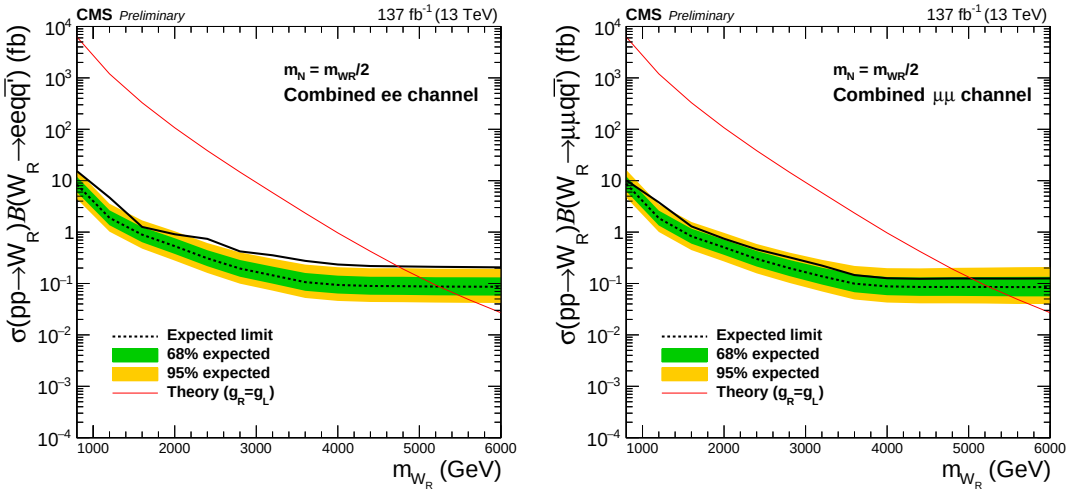


Figure 8.6: The upper limit on $\sigma(pp \to W_R) \times BR(W_R \to ee(\mu\mu)jj)$ cross section limit are shown for both of lepton flavors with the combined data of all three years. The N_R mass at each point is 0.5 times the W_R. This limit is dominated by the resolved signal region.

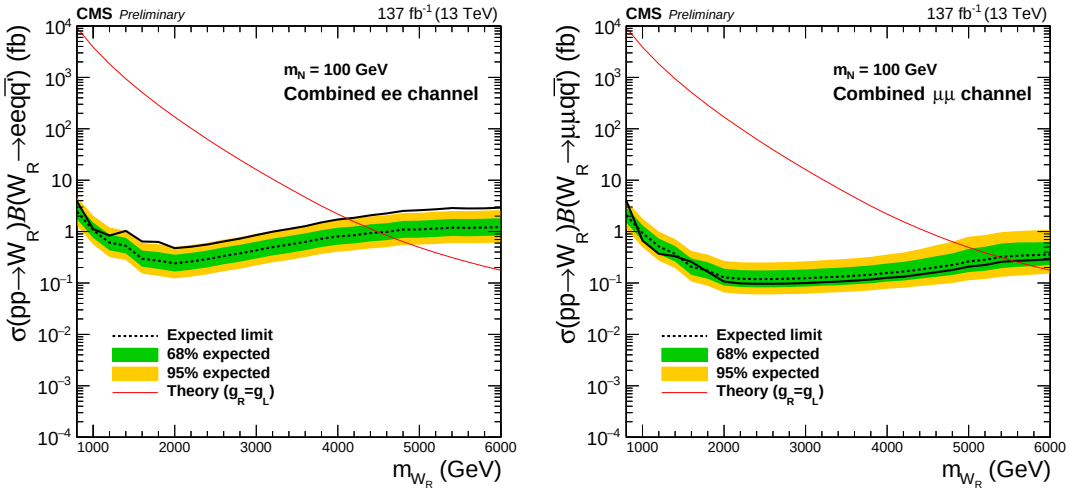


Figure 8.7: The upper limit on $\sigma(pp \to W_R) \times BR(W_R \to ee(\mu\mu)jj)$ cross section limit are shown for both of lepton flavors with the combined data of all three years. The N_R mass at each point is 100 GeV. This limit is dominated by the boosted signal region.

to produce the combined limit shown in Fig. 8.6 for each lepton flavor. For this graph, the N_R/W_R mass ratio is fixed at 0.5. In Fig. 8.7 the same calculation is shown where the N_R mass is fixed at 100 GeV.

8.3.3 Two Dimensional Limits

Two dimensional (2D) limits are produced using a grid of W_R–N_R mass points with the hypothesis that $g_{\mathrm{L}} = g_{\mathrm{R}}$. The colors represent the cross-section limits and the contours can be thought of as representing the point at which the theoretical cross-section shown in the 1D limit crosses over the measured limit, or crosses over one of the expected limit contours. The two-dimensional limits are shown in Fig. 8.8 and Fig. 8.9. The two dimensional limits are shown for both searched lepton flavors. The colors show the ratio between the observed cross-section limit at the given point and the $g_{\mathrm{L}} = g_{\mathrm{R}}$ prediction. Several contours are shown. The resolved and boosted contours are shown separately and combined. It can be seen that the electron channel observed excess results in a weaker than expected limit. While the muon channel limit is on the low side of the expected limit, it is important to remember that the observed limit is based on just a few events.

8.3.4 A Discussion of the Excess

The electron-flavor analysis shows a roughly two sigma excess in the resolved channel. The expected background distribution for the resolved electron channel for all three years is shown with the data overlayed. A total of thirteen events were observed in the last bin, exceeding the expectation by ~ 5 events.

While this excess could be indicative of some type of signal, it is important to additionally study any possible reconstruction or detector effects. It is possible that the excess comes from an electron reconstruction or related detector issue or it could be the result of an underestimated or unknown systematic discrepancy in the background simulations. The specific detector conditions at the time the events were recorded, and the details of the ECAL response to the recorded electrons were studied by the ECAL detector performance group and no significant suspects have been determined. It is impossible to determine whether any specific event is a background or signal event, however, it is worthwhile to note that the excess events are within a mass region previously searched in and excluded.

A three-dimensional display of one of the recorded events in the highest mass bin of the electron-resolved signal region is shown in Fig. 8.11. The objects reconstructed in this event are detailed in Table 8.5.

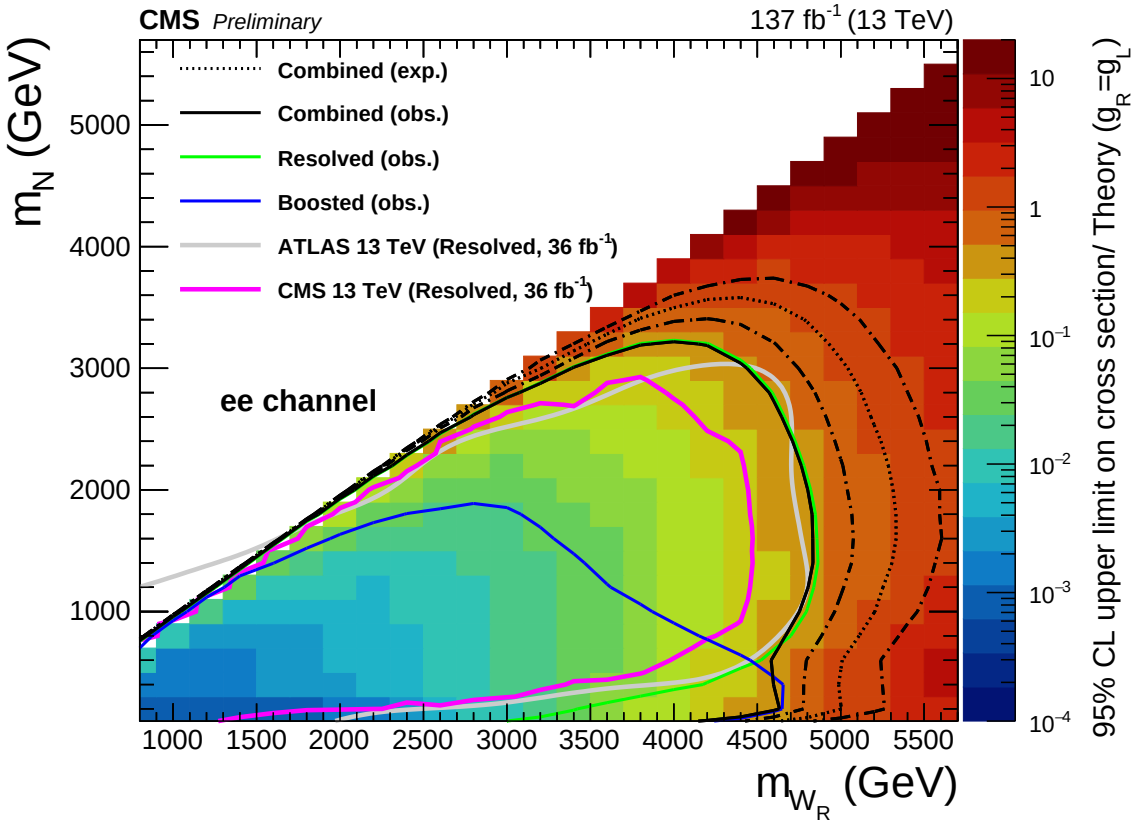


Figure 8.8: The upper limit on $\sigma(pp \to W_R) \times BR(W_R \to eejj)$ cross section limit are shown for the entire Run 2 dataset. The expected and observed exclusions are shown for the resolved, boosted, and combined analysis. The CMS and ATLAS' experiment's observed results with the 2016 dataset and a resolved signal selection are also shown.

Table 8.5: The object information of the resolved di-electron event, (RunNumber, LumiSection, EventNumber) = (278406:329:470624728). The event display is shown in Fig. 8.11.

Object ID	p_{T} (GeV)	η	ϕ	mass (GeV)
Positron	921.68	0.18	1.06	–
Electron	73.89	1.82	-1.96	–
Jet	805.76	-1.88	-1.57	47.71
Jet	528.99	1.18	3.08	33.32
$m(\ell\ell jj)$	4605.65 GeV			
$m(\ell\ell)$	705.90 GeV			

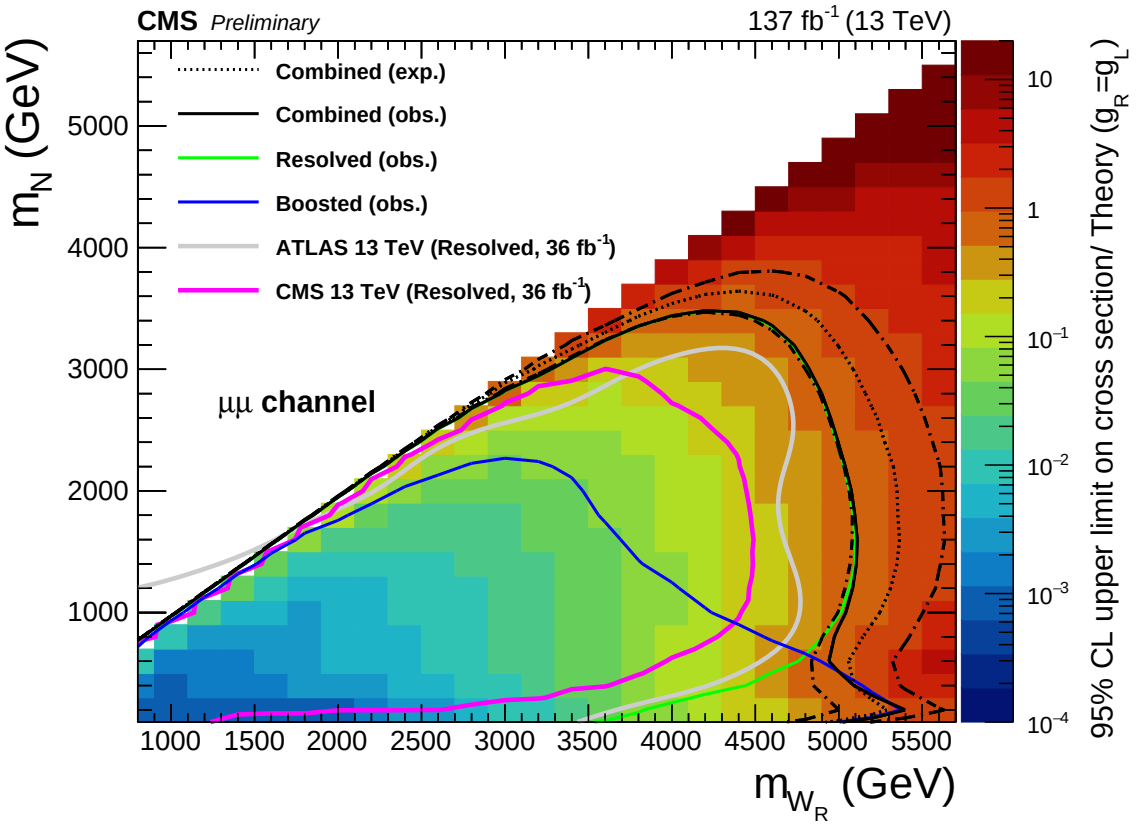


Figure 8.9: The upper limit on $\sigma(pp \to W_R) \times BR(W_R \to \mu\mu jj)$ cross section limit are shown for the entire Run 2 dataset. The expected and observed exclusions are shown for the resolved, boosted, and combined analysis. The CMS and ATLAS' experiment's observed results with the 2016 dataset and a resolved signal selection are also shown.

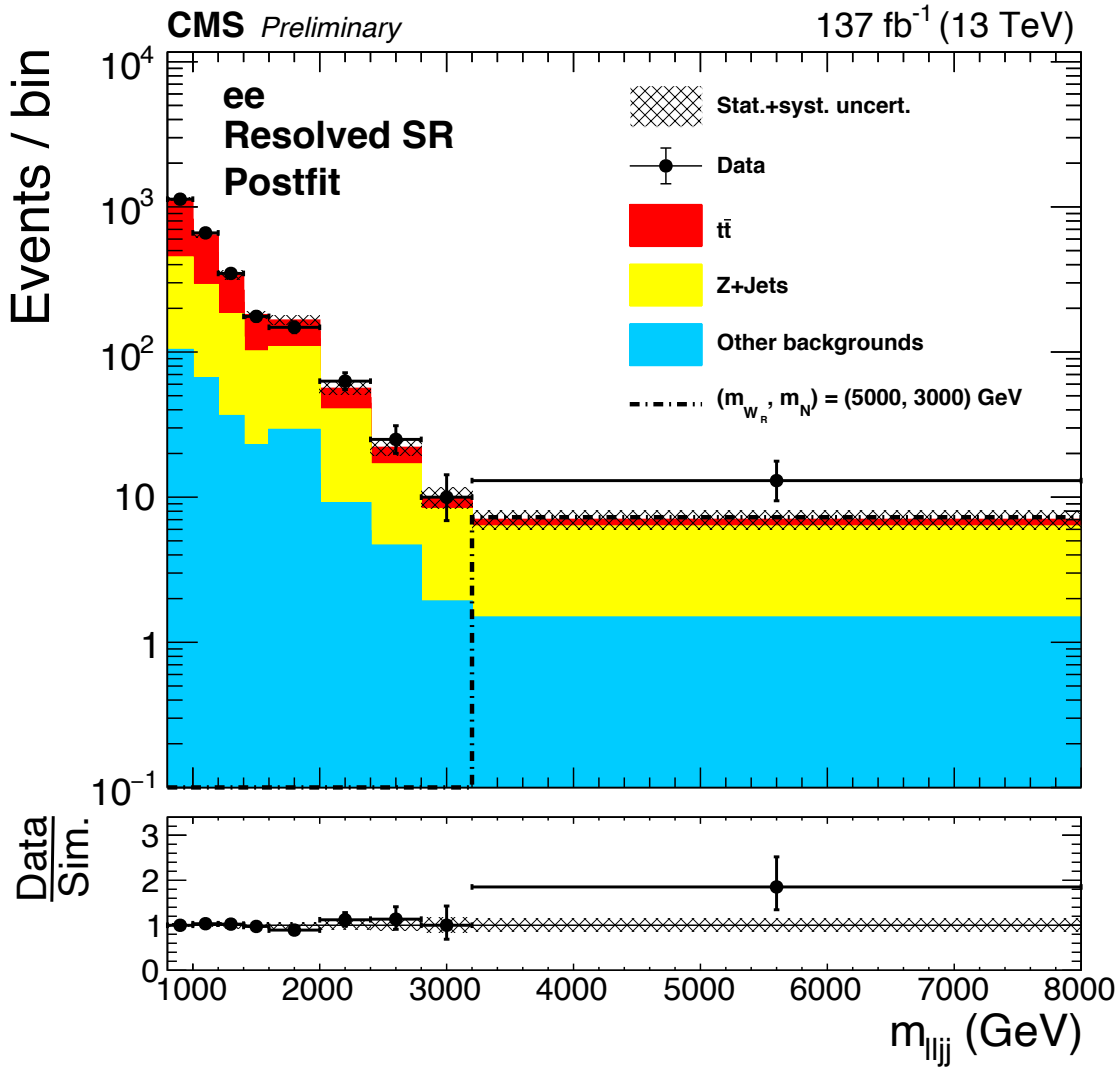


Figure 8.10: The expected and measured mass spectrum is shown in the resolved electron channel. An excess can be seen in the last bin with 13 events observed and ~ 8 expected.

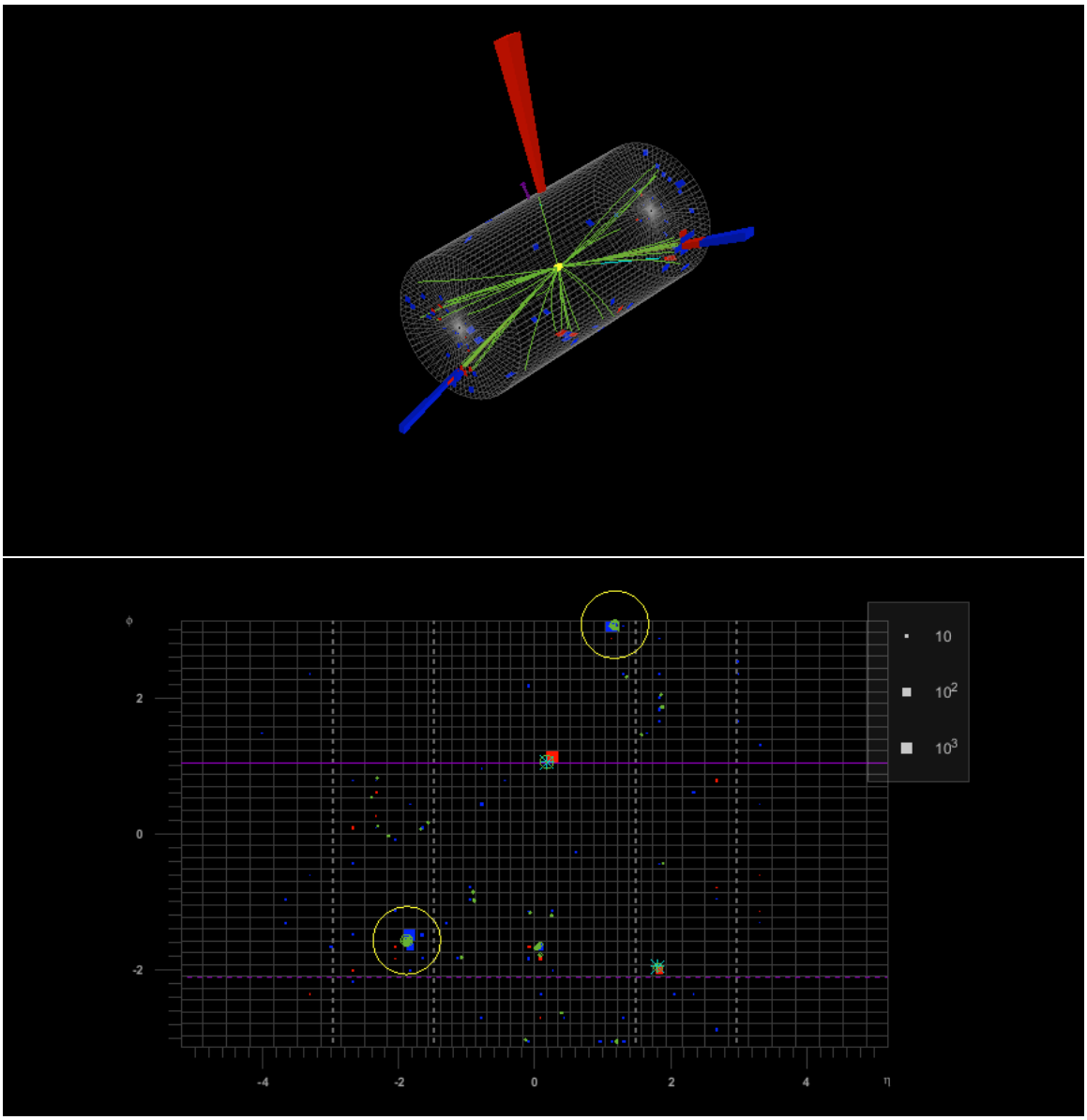
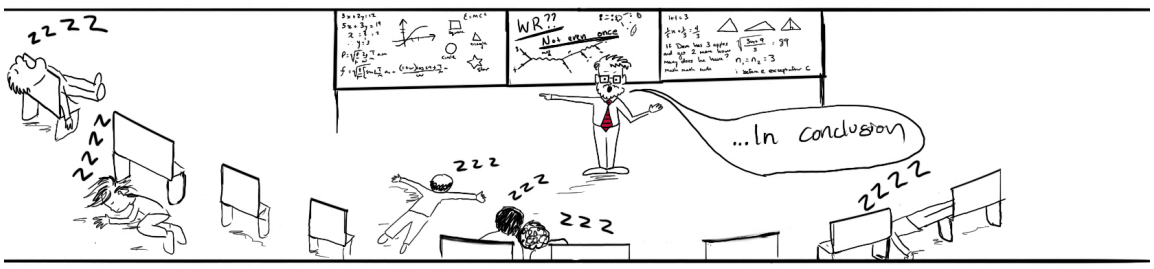


Figure 8.11: Event display of the resolved di-electron event, (RunNumber, LumiSection, EventNumber) = (278406:329:470624728). Two perspectives are shown, the top shows a 3D representation of the detector and the bottom shows an unwrapped version with $\eta - -\phi$ coordinates. Tracks corresponding with the primary collision vertex (yellow) are shown in green, energy deposited in the ECAL is shown with red bars, and energy deposited in the HCAL is shown with blue bars. The reconstructed electrons are circled in yellow.

Chapter 9

Conclusion



The most complete limits on the W_R decaying to two muons or two electrons were detailed in this analysis. No significant deviations from the Standard Model were observed in the combined $\sqrt{s} = 13\,\text{TeV}$ dataset collected over 2016, 2017, and 2018 for a total of $137\,\text{fb}^{-1}$. The limit on the W_R mass is set at $4-5$ TeV and $4-5.5$ TeV for the electron and muon flavor search respectively for N_R masses ranging between $100-3000$ GeV.

Theoretical motivations for left-right symmetric additions to the standard model have been presented and their historical origins [1] have been discussed in Chapter 2. After covering the design and performance of the CMS detector in Chapter 3, the new breakthroughs in boosted-object reconstruction were covered in Chapter 4. Leveraging these exciting techniques, the W_R signal acceptance was greatly increased and a complete analysis strategy for boosted and resolved W_R signals, decaying to two leptons and two jets, was shown in Chapter 5. The reconstruction and identification of all of the necessary particles in the CMS detector was presented in Chapter 6. All of the standard model backgrounds of this analysis were discussed in Chapter 7 as well as the detailed study of two dominant backgrounds, $\text{t}\bar{\text{t}}$ and Z/γ^* . The computation of the upper-limit on the W_R cross-section was detailed

and the characteristics of a small excess in the resolved-electron channel were discussed in Chapter 8.

Ultimately, the cross-section limits set in this analysis are broadest, in terms of W_R–N_R mass range, to date. Some work could be done to better understand the non-prompt background contribution to the boosted-electron analysis channel. However, the overall significance of this analysis for leveraging new boosted-object techniques, in combination with sifting through the largest amount of data for a W_R search at CMS yet, remains.

References

[1] C.S. Wu, E Ambler, R.W. Hayward, D.D. Hoppes, and R.P. Hudson. Experimental test of parity conservation in beta decay. *Physical Review Letters*, 105:1413–1415, 02 1957.

[2] S Chatrchyan, V Khachatryan, AM Sirunyan, et al. Search for heavy neutrinos and w(r) bosons with right-handed couplings in a left-right symmetric model in pp collisions at sqrt[s]=7tev. *Physical review letters*, 109(26):261802, 2012.

[3] Vardan Khachatryan. Search for heavy neutrinos and w bosons with right-handed couplings in protonproton collisions at s = 8 tev. *European Physical Journal. C, Particles and Fields*, 74(11), 11 2014.

[4] A. M. Sirunyan et al. Search for a heavy right-handed W boson and a heavy neutrino in events with two same-flavor leptons and two jets at $\sqrt{s} = 13$ TeV. *JHEP*, 05:148, 2018.

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

[6] A. Salam. *Elementary particle theory. Relativistic groups and analyticity. Proccedings of the Eighth Nobel Symposium held May 19-25, 1968 at Aspenasgarden, Lerum, in the county of Alvsborg, Sweden. Edited by Nils Svartholm.* Almqvist & Wiksell; John Wiley Stockholm, New York, London [etc.], 1968.

[7] H. A. Lorentz. Simplified Theory of Electrical and Optical Phenomena in Moving Systems. *Koninklijke Nederlandse Akademie van Wetenschappen Proceedings Series B Physical Sciences*, 1:427–442, 1898.

[8] A. Einstein. Zur elektrodynamik bewegter krper. *Annalen der Physik*, 322(10):891–921, 1905, https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.19053221004.

[9] Paul Adrien Maurice Dirac and Ralph Howard Fowler. The quantum theory of the electron. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 117(778):610–624, 1928, https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.1928.0023.

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

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

[12] E. Fermi. Versuch einer theorie der β-strahlen. i. *Zeitschrift für Physik*, 88(3):161–177, Mar 1934.

[13] J. Orear, G. Harris, and S. Taylor. Spin and parity analysis of bevatron /tau/ mesons. *Physical Review (U.S.) Superseded in part by Phys. Rev. A, Phys. Rev. B: Solid State, Phys. Rev. C, and Phys. Rev. D*, 102, 6 1956.

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

[15] 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. *Physical Review - PHYS REV X*, 105:1415–1417, 02 1957.

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

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

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

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

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

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

[22] G. Aad, T. Abajyan, B. Abbott, 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.

[23] S. Chatrchyan, V. Khachatryan, A.M. Sirunyan, 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.

[24] G. D. Rochester and C. C. Butler. Evidence for the Existence of New Unstable Elementary Particles. *Nature*, 160:855–857, December 1947.

[25] M. Gell-Mann. A schematic model of baryons and mesons. *Phys. Letters*, 8, 2 1964.

[26] G Zweig. An SU_3 model for strong interaction symmetry and its breaking; Version 1. Technical Report CERN-TH-401, CERN, Geneva, Jan 1964.

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

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

[29] J. E. Augustin et al. Discovery of a Narrow Resonance in e^+e^- Annihilation. *Phys. Rev. Lett.*, 33:1406–1408, 1974. [Adv. Exp. Phys.5,141(1976)].

[30] 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, 1977.

[31] G. Arnison et al. Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c**2 at the CERN SPS Collider. *Phys. Lett.*, B126:398–410, 1983. [,7.55(1983)].

[32] P. Bagnaia et al. Evidence for Z0 —¿ e+ e- at the CERN anti-p p Collider. *Phys. Lett.*, B129:130–140, 1983. [,7.69(1983)].

[33] G. Arnison et al. Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at s**(1/2) = 540-GeV. *Phys. Lett.*, B122:103–116, 1983. [,611(1983)].

[34] M. Banner et al. Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider. *Phys. Lett.*, B122:476–485, 1983. [,7.45(1983)].

[35] F. Abe, M. G. Albrow, S. R. Amendolia, et al. Evidence for top quark production in pp collisions at s = 1.8 tev. *Physical review D: Particles and fields*, 50(5):2966–3026, 1 1994.

[36] S. Abachi et al. Search for high mass top quark production in $p\bar{p}$ collisions at $\sqrt{s} =$ 1.8 TeV. *Phys. Rev. Lett.*, 74:2422–2426, 1995, hep-ex/9411001.

[37] M. Tanabashi, K. Hagiwara, K. Hikasa, et al. Review of particle physics. *Phys. Rev. D*, 98:030001, Aug 2018.

[38] J. C. Pati and A. Salam. Lepton number as the fourth "color". *Phys. Rev. D*, 10:275, 1974.

[39] R. N. Mohapatra and J. C. Pati. A natural left-right symmetry. *Phys. Rev. D*, 11:2558, 1975.

[40] R. Jr. Davis, D.S. Harmer, and K.C. Hoffman. Search for neutrinos from the sun. *Phys. Rev. Lett.*, 1 1968.

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

[42] R.N. Mohapatra and G. Senjanovic. Neutrino mass and spontaneous parity nonconservation. *Phys. Rev. Lett.*, 44(14), 4 1980.

[43] Ettore Majorana. Teoria simmetrica dellelettrone e del positrone. *Nuovo Cim.*, 14:171–184, 1937.

[44] Paul Langacker and S. Uma Sankar. Bounds on the mass of W_R and the $W_L - W_R$ mixing angle ζ in general $\mathrm{SU}(2)_L \times \mathrm{SU}(2)_R \times \mathrm{U}(1)$ models. *Phys. Rev. D*, 40:1569–1585, Sep 1989.

[45] Albert M Sirunyan et al. Search for high-mass resonances in dilepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV. *JHEP*, 06:120, 2018, 1803.06292.

[46] Albert M Sirunyan et al. Search for high mass dijet resonances with a new background prediction method in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Journal of High Energy Physics*, 2019, 1911.03947.

[47] M. Aaboud et al. (ATLAS Collaboration). Search for heavy majorana or dirac neutrinos and right-handed w gauge bosons in final states with two charged leptons and two jets at s=13 tev with the atlas detector. *J. High Energy Phys.*, 016, 2019.

[48] Wikipedia. Cp violation. In *CP violation*. Wikimedia Foundation, Inc., 2019. https://en.wikipedia.org/wiki/CP_violation.

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

[50] Nick Ellis. The atlas years and the future. In *Symposium on High Energy Physics*. University of Birmingham School of Physics and Astronomy, July 2002. http://www.ep.ph.bham.ac.uk/general/outreach/dowellfest/.

[51] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Parton distributions for the LHC. *Eur. Phys. J.*, C63:189–285, 2009, 0901.0002.

[52] Tai Sakuma and Thomas McCauley. Detector and event visualization with SketchUp at the CMS experiment. *Journal of Physics: Conference Series*, 513(2):022032, jun 2014.

[53] V Veszpremi. Operation and performance of the cms tracker. *Journal of Instrumentation*, 9:C03005, 03 2014.

[54] 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.

[55] S. Abdullin, ITEP /Moscow, V. Abramov, et al. The cms barrel calorimeter response to particle beams from 2-gev/c to 350-gev/c. *Eur.Phys.J.C60:359-373,2009*, 60(3), 1 2009.

[56] A.M. Sirunyan, Armen Tumasyan, Wolfgang Adam, et al. Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s}$= 13 TeV. *JINST*, 13(CMS-MUO-16-001. CMS-MUO-16-001-003. 06):P06015. 53 p, Apr 2018.

Replaced with the published version. Added the journal reference and the DOI. All the figures and tables can be found at http://cms-results.web.cern.ch/cms-results/public-results/publications/MUO-16-001 (CMS Public Pages).

[57] S. Chatrchyan, V. Khachatryan, A.M. Sirunyan, et al. Precise mapping of the magnetic field in the cms barrel yoke using cosmic rays. *Journal of Instrumentation*, 5(03), 3 2010.

[58] Manimala Mitra, Richard Ruiz, Darren J. Scott, and Michael Spannowsky. Neutrino Jets from High-Mass W_R Gauge Bosons in TeV-Scale Left-Right Symmetric Models. *Phys. Rev.*, D94(9):095016, 2016, 1607.03504.

[59] A. M. Sirunyan, A. Tumasyan, W. Adam, et al. Inclusive search for a highly boosted higgs boson decaying to a bottom quark-antiquark pair. *Phys. Rev. Lett.*, 120:071802, Feb 2018.

[60] M. Tanabashi et al. Review of Particle Physics. *Phys. Rev.*, D98(3):030001, 2018.

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

[62] Jesse Thaler and Ken Van Tilburg. Identifying boosted objects with n-subjettiness. *Journal of High Energy Physics*, 2011(3):15, Mar 2011.

[63] Christopher Brust, Petar Maksimovic, Alice Sady, et al. Identifying boosted new physics with non-isolated leptons. *Journal of High Energy Physics*, 2015(4):79, Apr 2015.

[64] Simone Marzani, Lais Schunk, and Gregory Soyez. The jet mass distribution after soft drop. *The European Physical Journal C*, 78(2):96, Feb 2018.

[65] V. Khachatryan, A.M. Sirunyan, A. Tumasyan, et al. Search for w decaying to tau lepton and neutrino in protonproton collisions at s=8tev. *Physics Letters B*, 755:196 – 216, 2016.

[66] Daniele Bertolini, Philip Harris, Matthew Low, and Nhan Tran. Pileup Per Particle Identification. *JHEP*, 10:059, 2014, 1407.6013.

[67] A M Sirunyan et al. Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$TeV. *JINST*, 13:P06015, 2018, 1804.04528.

[68] Bodek, A., Dyne, A., Han, J.Y., Sakumoto, W., and Strelnikov, A. Extracting muon momentum scale corrections for hadron collider experiments. *Eur. Phys. J. C*, 72(10):2194, 2012.

[69] A.M. Sirunyan, A. Tumasyan, W. Adam, et al. Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s}$=13 TeV. *Journal of Instrumentation*, 13(06):P06015–P06015, jun 2018.

[70] A. M. Sirunyan et al. Particle-flow reconstruction and global event description with the CMS detector. *JINST*, 12(10):P10003, 2017, 1706.04965.

[71] V. Khachatryan. Performance of electron reconstruction and selection with the cms detector in proton-proton collisions at s = 8 tev. *Journal of Instrumentation*, 10(06), 6 2015.

[72] EGamma POG. Electron energy smearing and scale corrections. `https://twiki.cern.ch/twiki/bin/viewauth/CMS/EGMSmearer`.

[73] EGamma POG. Egamma pog electron efficiencies and scale factors. `https://twiki.cern.ch/twiki/bin/view/CMS/EgammaIDRecipesRun2#Electron_efficiencies_and_scale`.

[74] EGamma POG. Heep electron id and isolation. `https://twiki.cern.ch/twiki/bin/viewauth/CMS/HEEPElectronIdentificationRun2#Selection_Cuts_HEEP_V6_0_Recomme`.

[75] Muon POG. `https://twiki.cern.ch/twiki/bin/view/CMS/MuonHLT2016`. Muon HLT in 2016.

[76] Muon POG. `https://twiki.cern.ch/twiki/bin/view/CMS/MuonHLT2017`. Muon HLT in 2017.

[77] Muon POG. `https://twiki.cern.ch/twiki/bin/view/CMS/MuonHLT2018`. Muon HLT in 2018.

[78] CMS collaboration. Measurements of inclusive and differential Z boson production cross sections in pp collisions at $\sqrt{s}$ = 13 TeV. CMS Physics Analysis Summary CMS-PAS-SMP-15-011, CERN, 2016.

[79] CMS collaboration. CMS luminosity measurements for the 2016 data taking period. CMS Physics Analysis Summary CMS-PAS-LUM-17-001, CERN, 2017.

[80] CMS collaboration. CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV. CMS Physics Analysis Summary CMS-PAS-LUM-17-004, CERN, 2018.

[81] CMS collaboration. CMS luminosity measurement for the 2018 data taking period at $\sqrt{s} = 13$ TeV. CMS Physics Analysis Summary CMS-PAS-LUM-18-002, CERN, 2017.

[82] Jon Butterworth, Stefano Carrazza, Amanda Cooper-Sarkar, et al. PDF4lhc recommendations for LHC run II. *Journal of Physics G: Nuclear and Particle Physics*, 43(2):023001, jan 2016.

[83] A L Read. Presentation of search results: theCLstechnique. *Journal of Physics G: Nuclear and Particle Physics*, 28(10):2693–2704, sep 2002.

[84] 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(2):435 – 443, 1999.

Appendix A

Individual Year Distributions

This appendix provides the signal and control region $m(\ell\ell jj)$ and $m(\ell J)$ distributions for each year. The pre-fit and post-fit distributions are shown for the Drell-Yan control regions, the flavor-sideband control regions and the signal regions for boosted and resolved N_R scenarios.

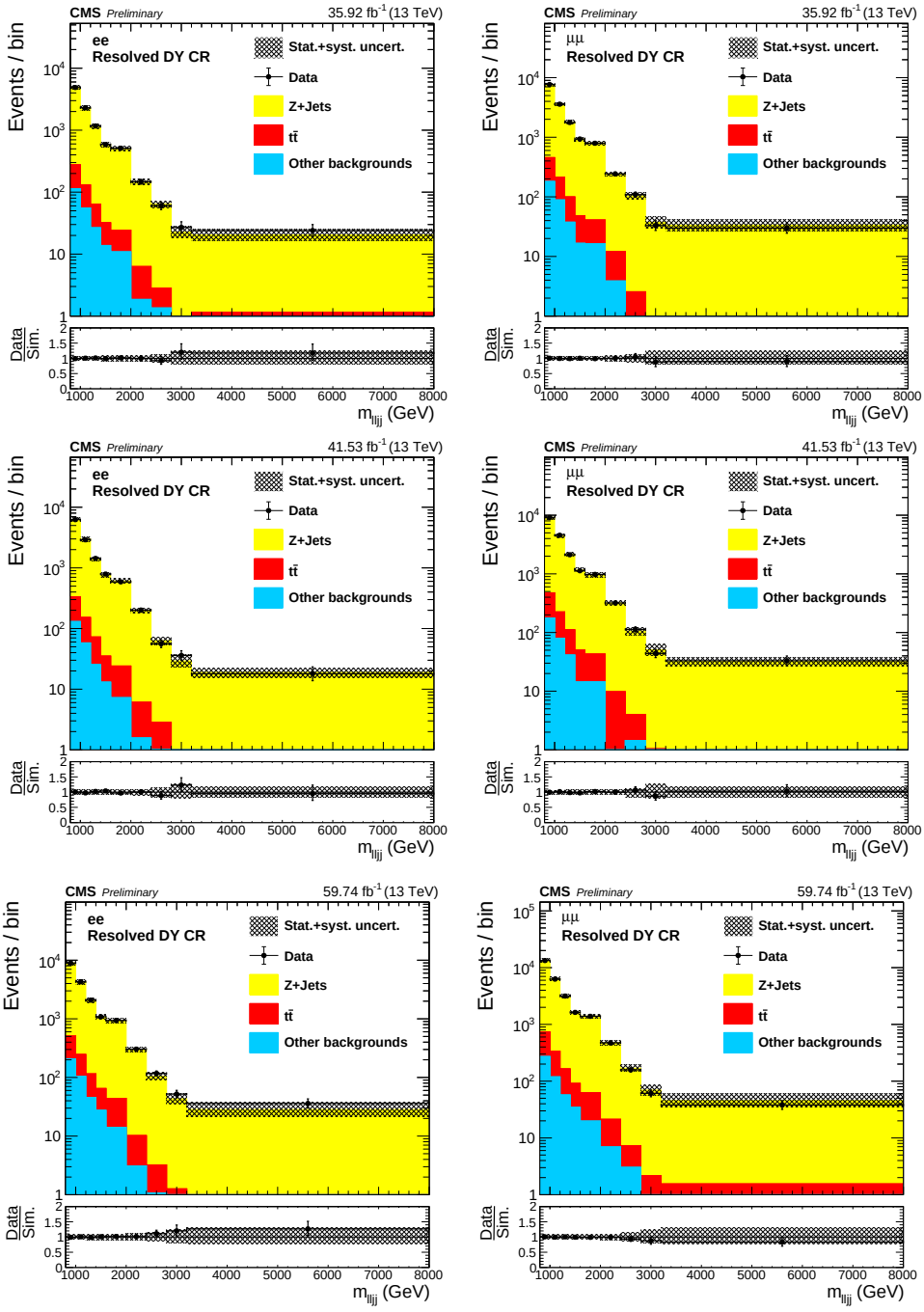


Figure A.1: The pre-fit $m(\ell\ell jj)$ and $m(\ell J)$ distribution in the low $m_{\ell\ell}$ control regions with three years individually. Results for the di-electron (di-muon) channel is shown on the left (right), for the resolved selection.

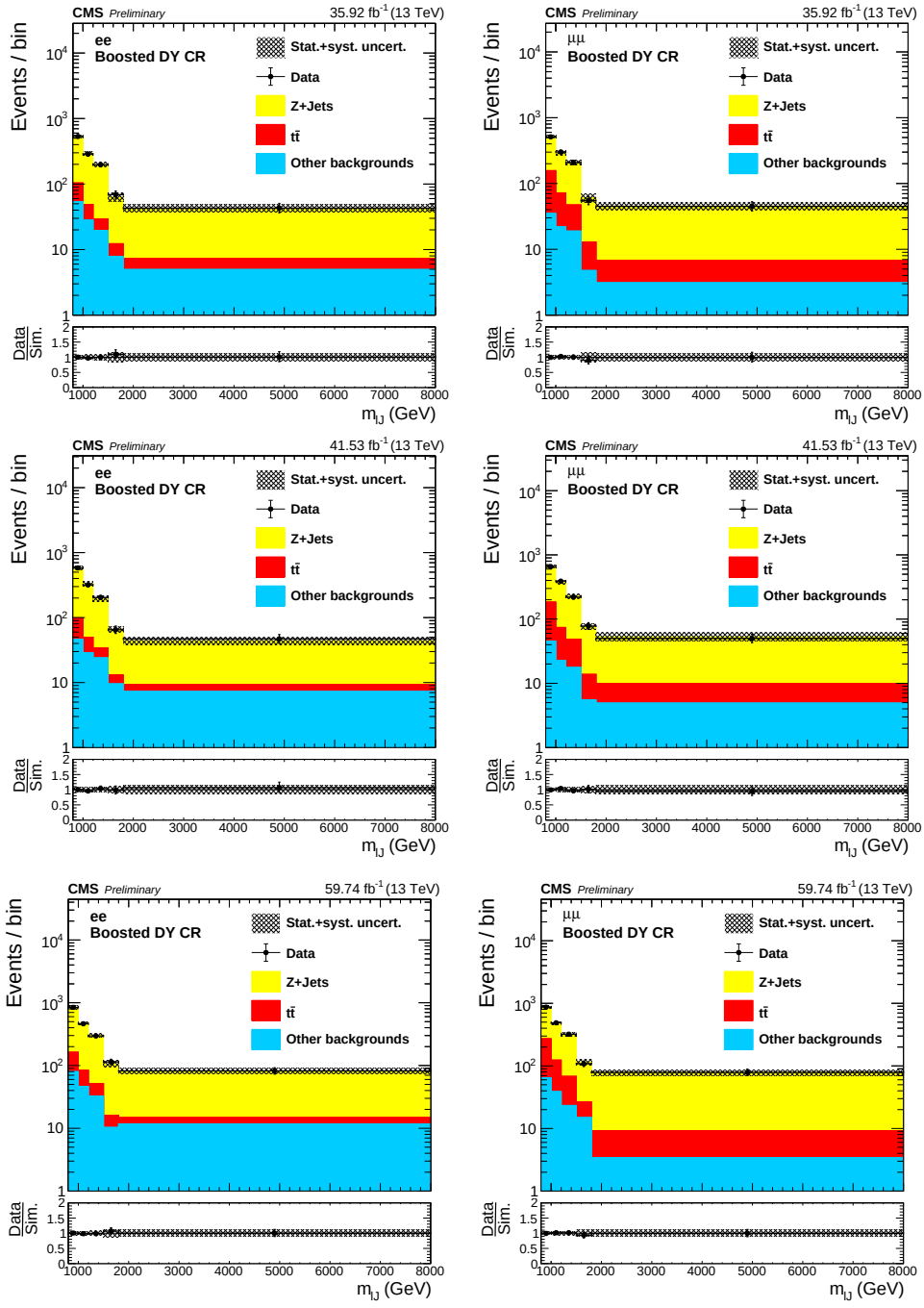


Figure A.2: The pre-fit $m(\ell\ell jj)$ and $m(\ell J)$ distribution in the low $m_{\ell\ell}$ control regions with three years individually. Results for the di-electron (di-muon) channel is shown on the left (right), for the boosted selection.

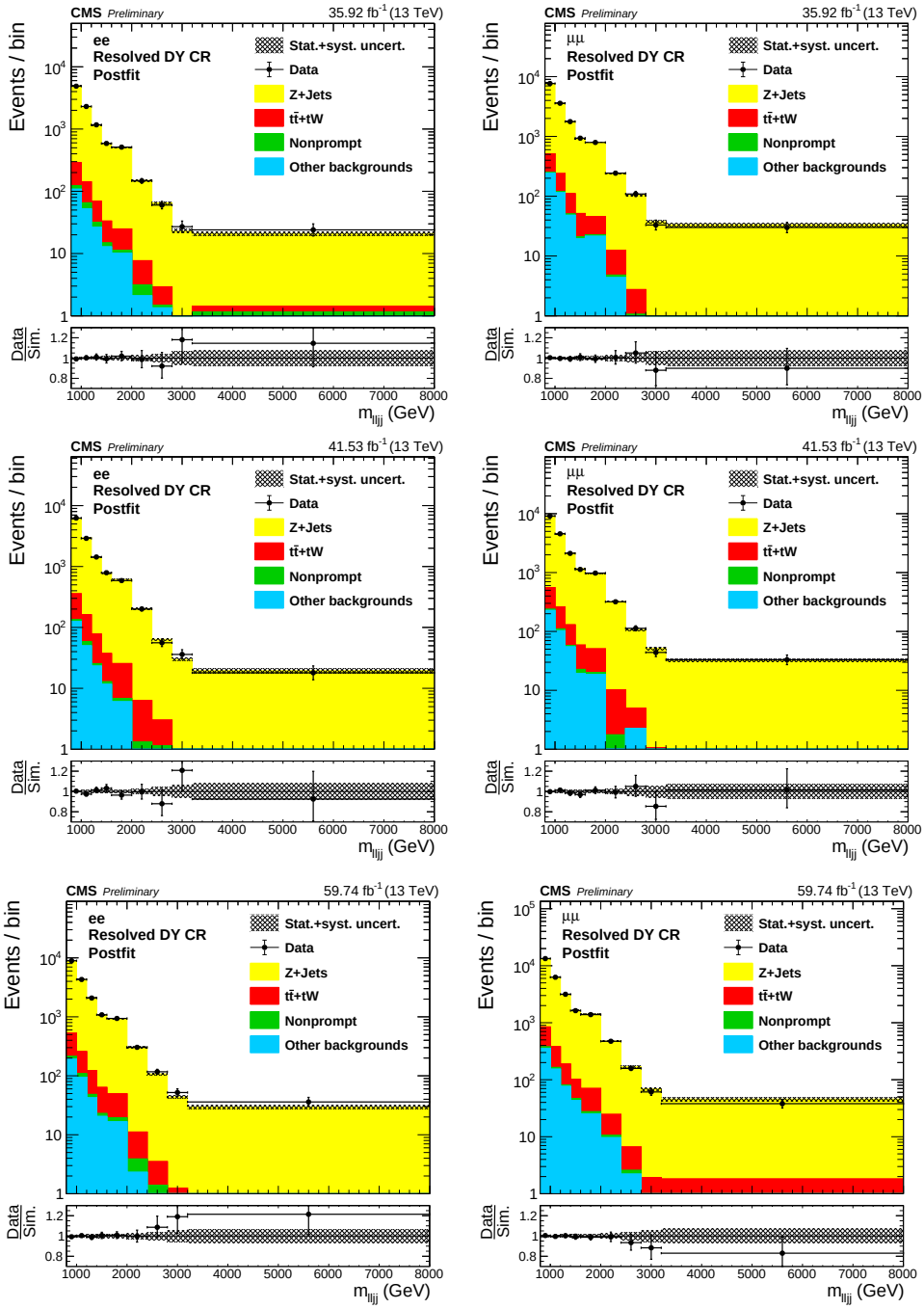


Figure A.3: The post-fit $m(\ell\ell jj)$ and $m(\ell J)$ distribution in the low $m_{\ell\ell}$ control regions with three years individually. Results for the di-electron (di-muon) channel is shown on the left (right), for the resolved selection.

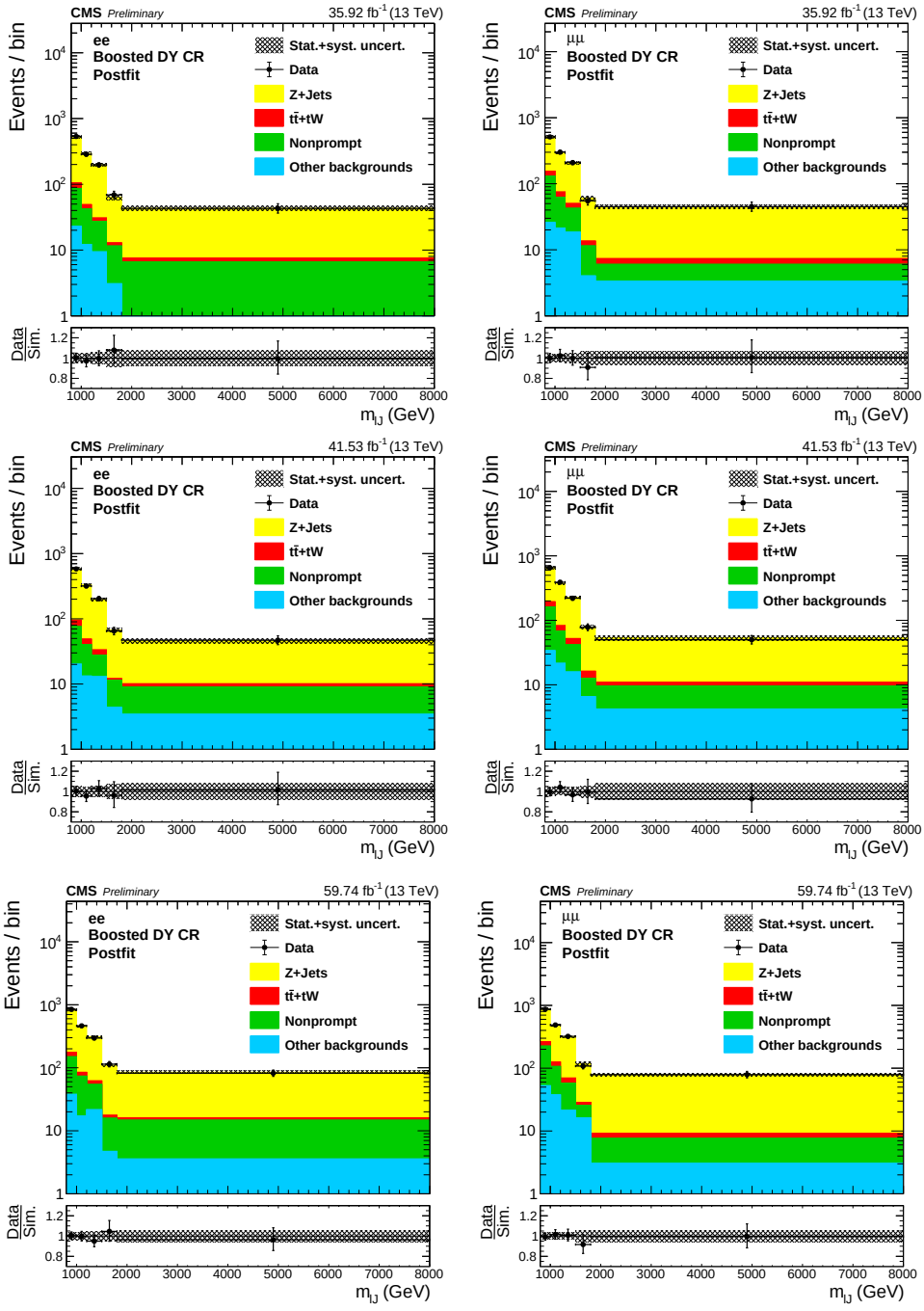


Figure A.4: The post-fit $m(\ell\ell jj)$ and $m(\ell J)$ distribution in the low $m_{\ell\ell}$ control regions with three years individually. Results for the di-electron (di-muon) channel is shown on the left (right), for the boosted selection.

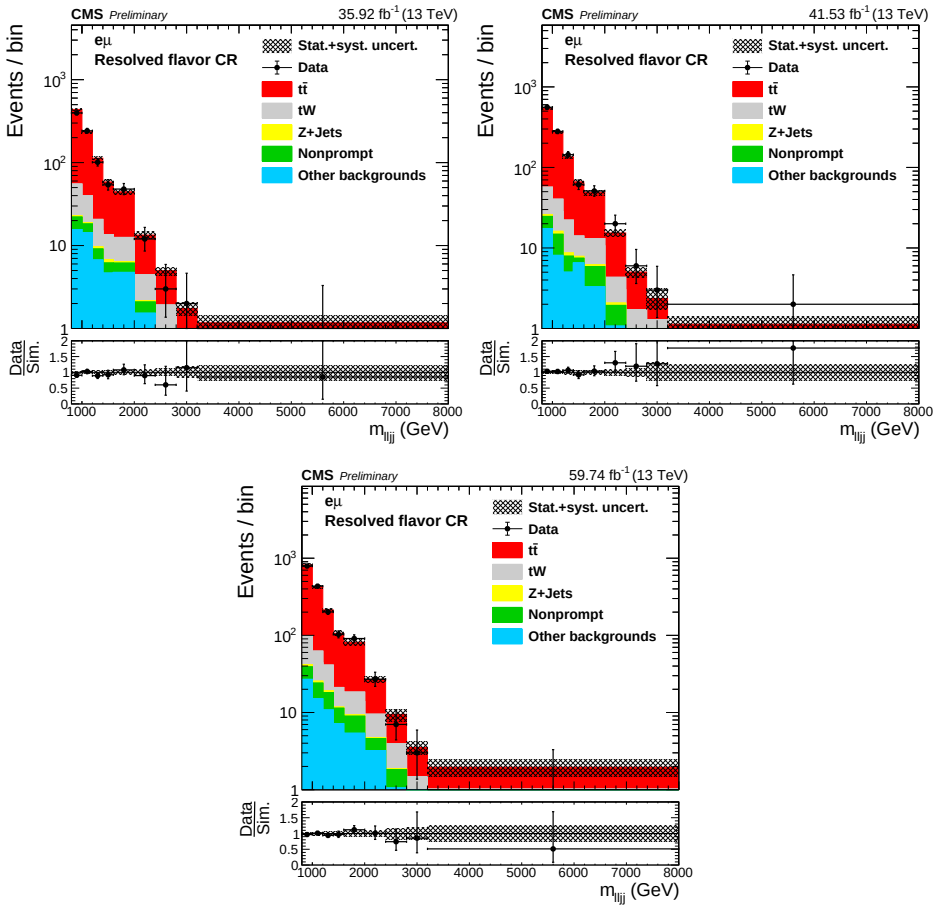


Figure A.5: The pre-fit $m(\ell\ell jj)$ and $m(\ell J)$ distribution in the flavor-sideband control regions with three years individually for the resolved selection.

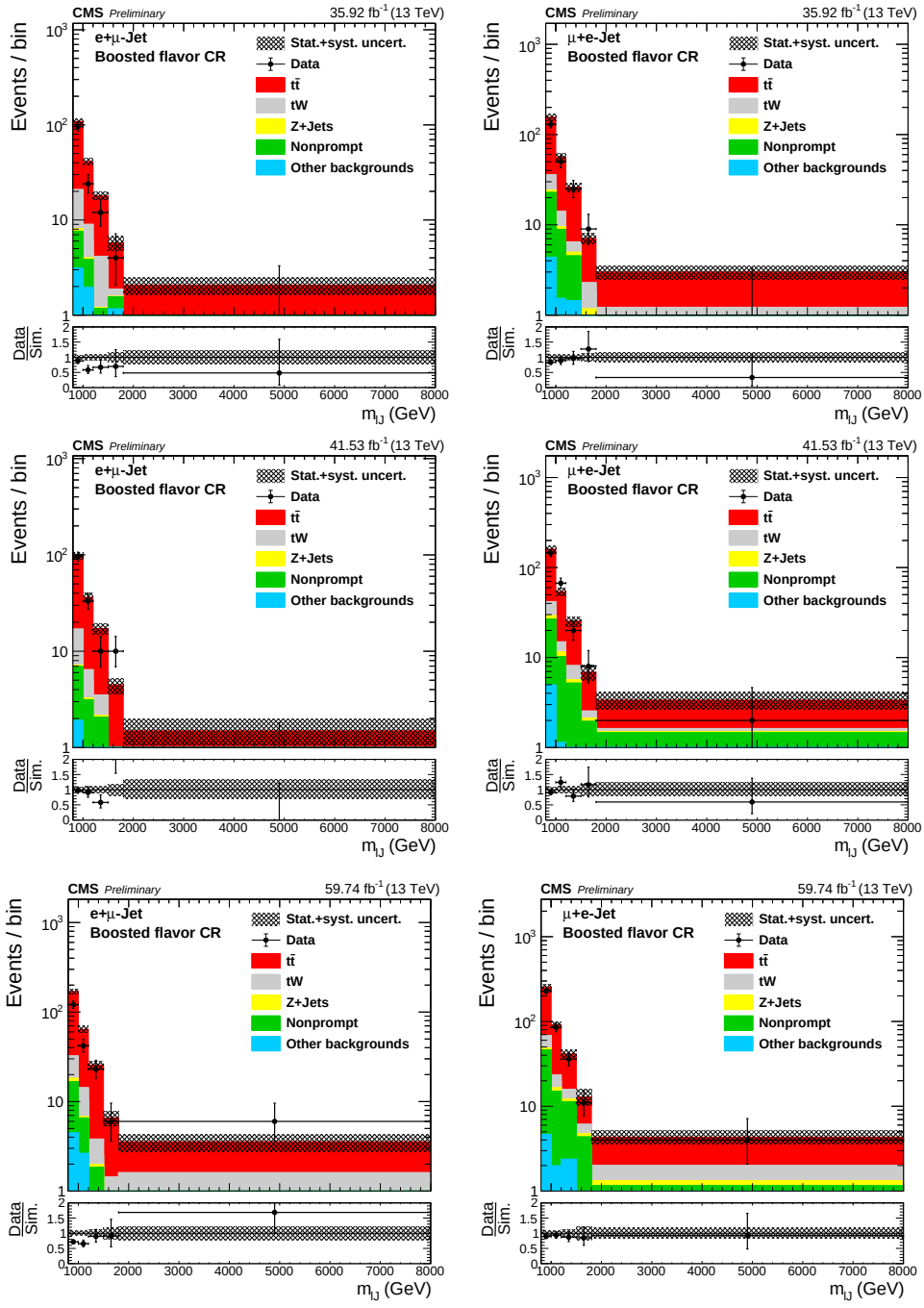


Figure A.6: The pre-fit $m(\ell\ell jj)$ and $m(\ell J)$ distribution in the flavor-sideband control regions with three years individually for the boosted selection. muon(electron)-in-jet selection is shown on the right(left)

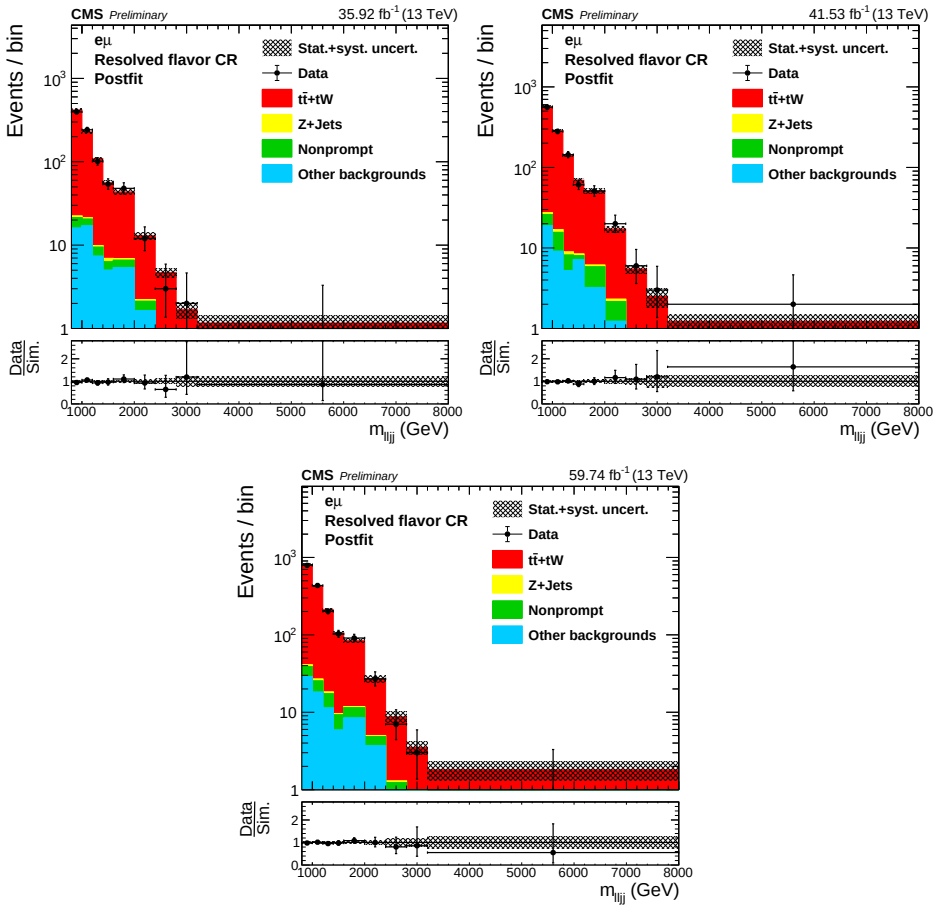


Figure A.7: The post-fit $m(\ell\ell jj)$ and $m(\ell J)$ distribution in the flavor-sideband control regions with three years individually for the resolved selection.

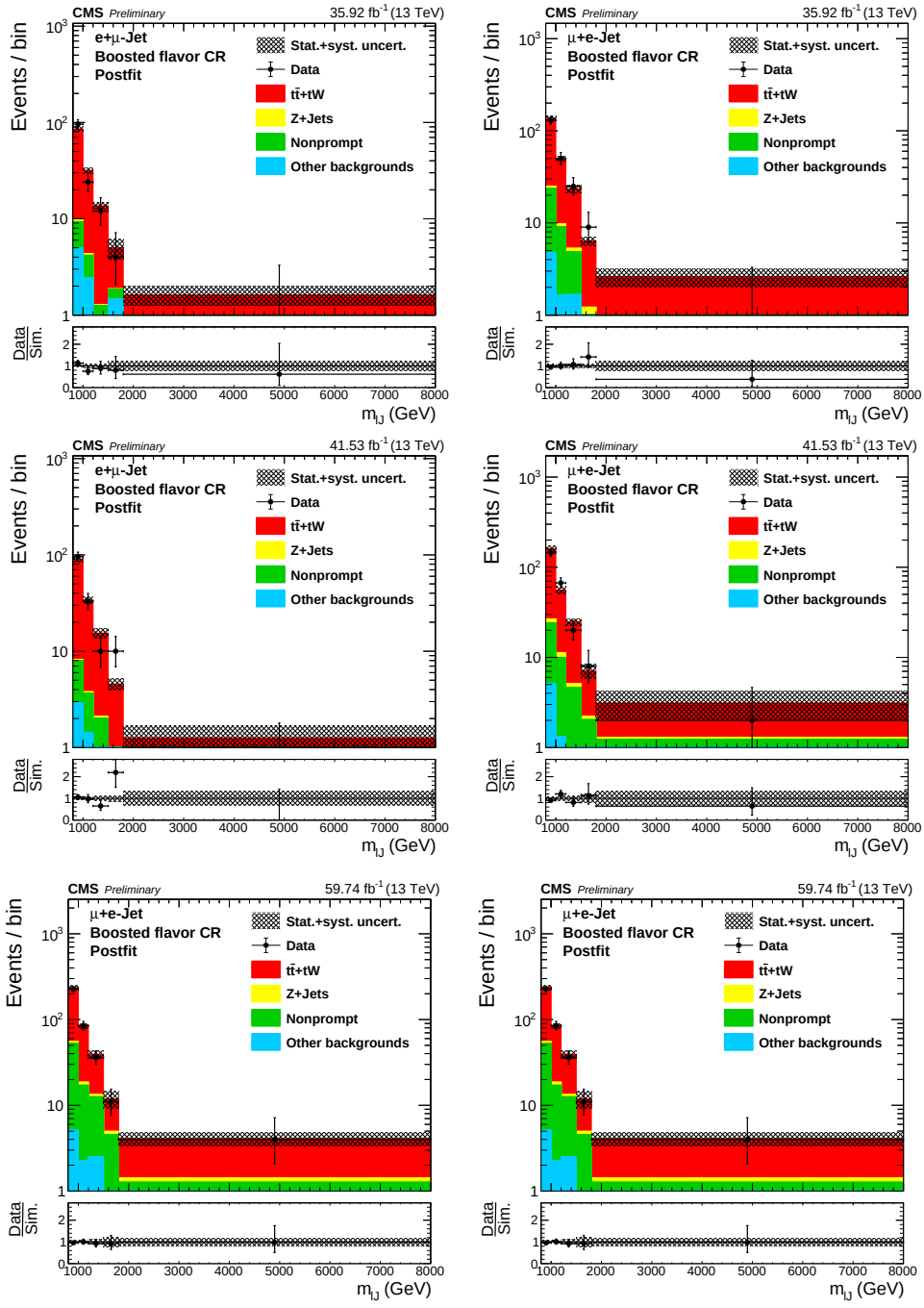


Figure A.8: The post-fit $m(\ell\ell jj)$ and $m(\ell J)$ distribution in the flavor-sideband control regions with three years individually for the boosted selection. muon(electron)-in-jet selection is shown on the right(left)

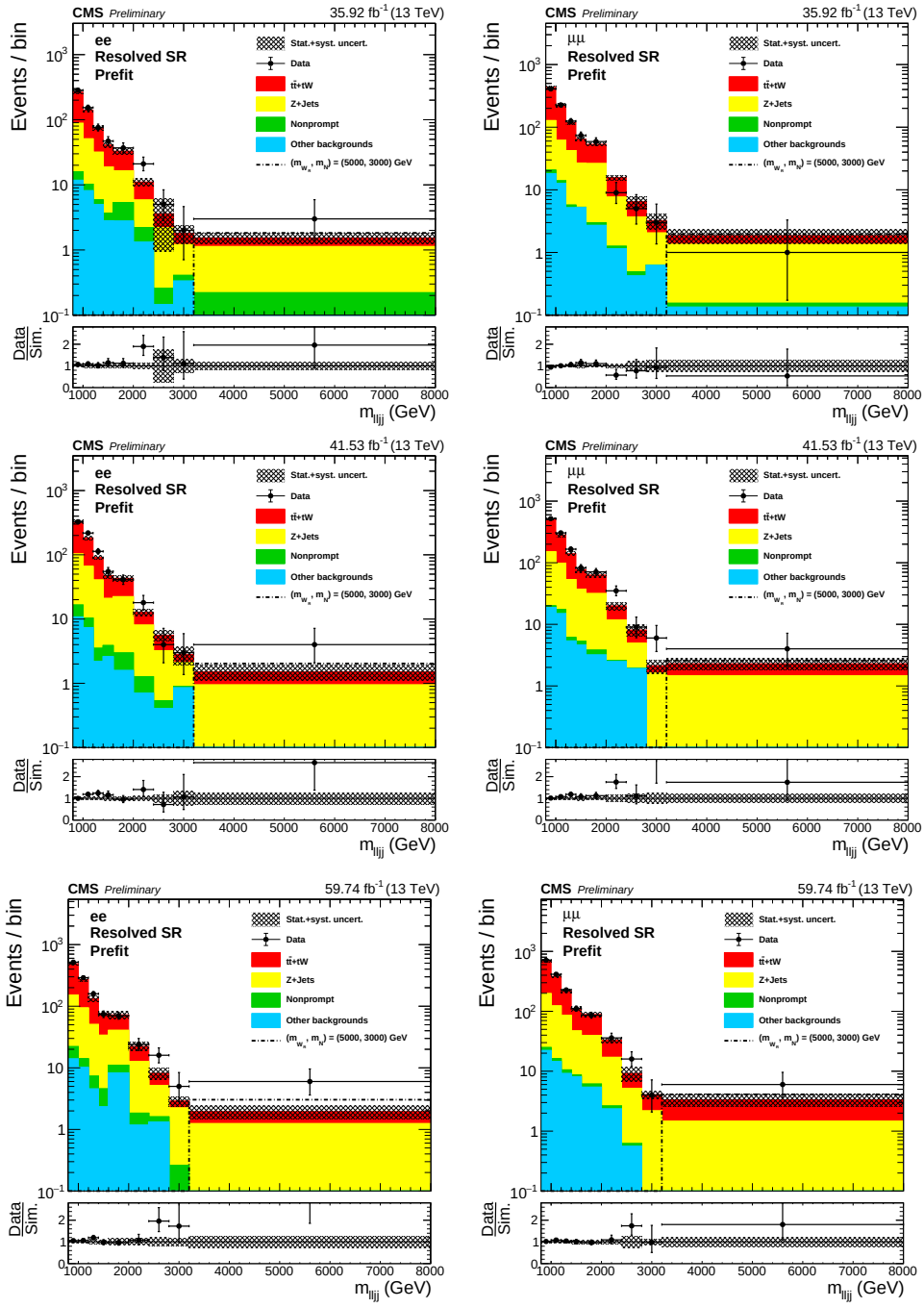


Figure A.9: The pre-fit $m(\ell\ell jj)$ and $m(\ell J)$ distribution in the signal regions with three years individually. Results for the di-electron (di-muon) channel is shown on the left (right), for the resolved selection.

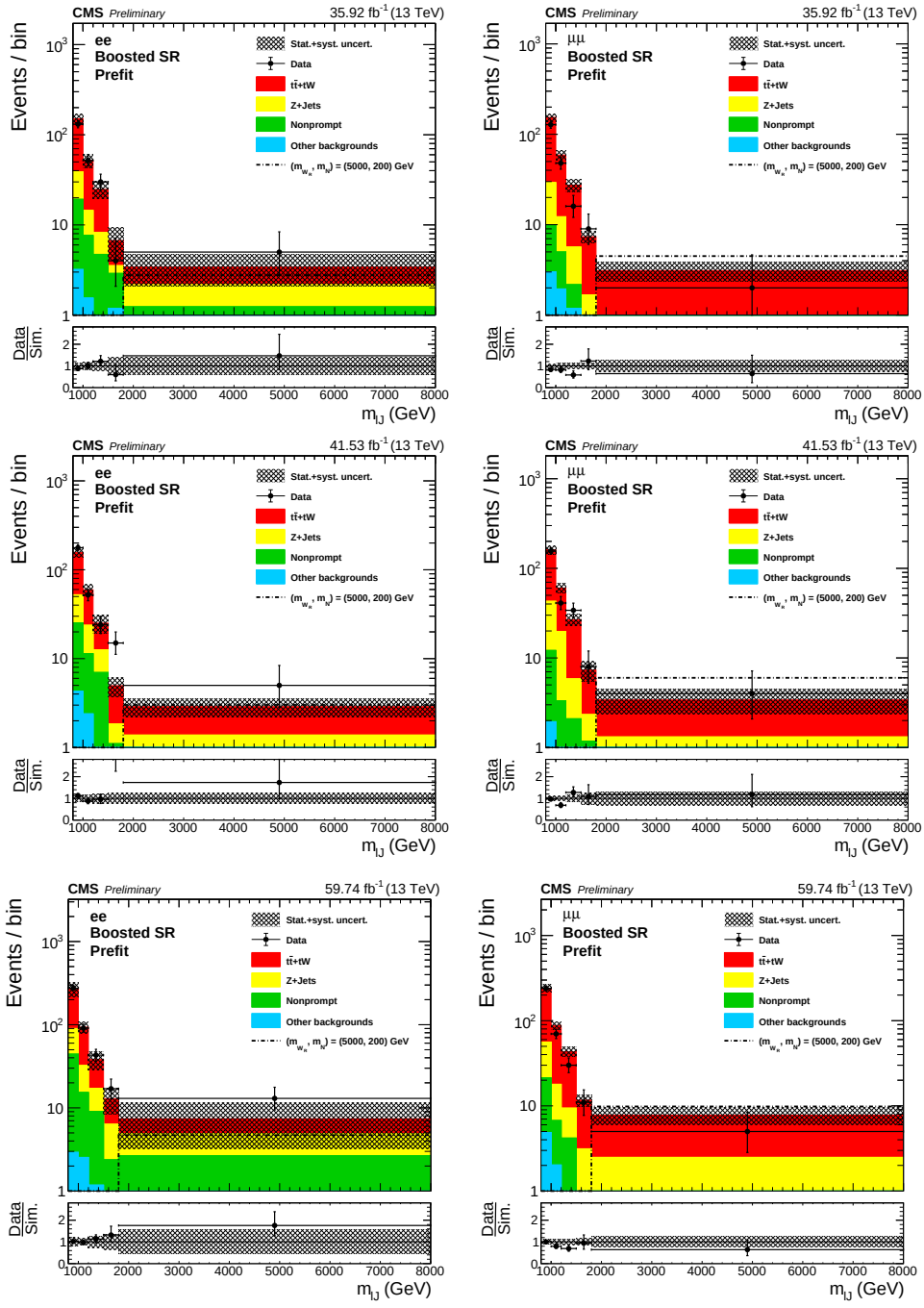


Figure A.10: The pre-fit $m(\ell\ell jj)$ and $m(\ell J)$ distribution in the signal regions with three years individually. Results for the di-electron (di-muon) channel is shown on the left (right), for the boosted selection.

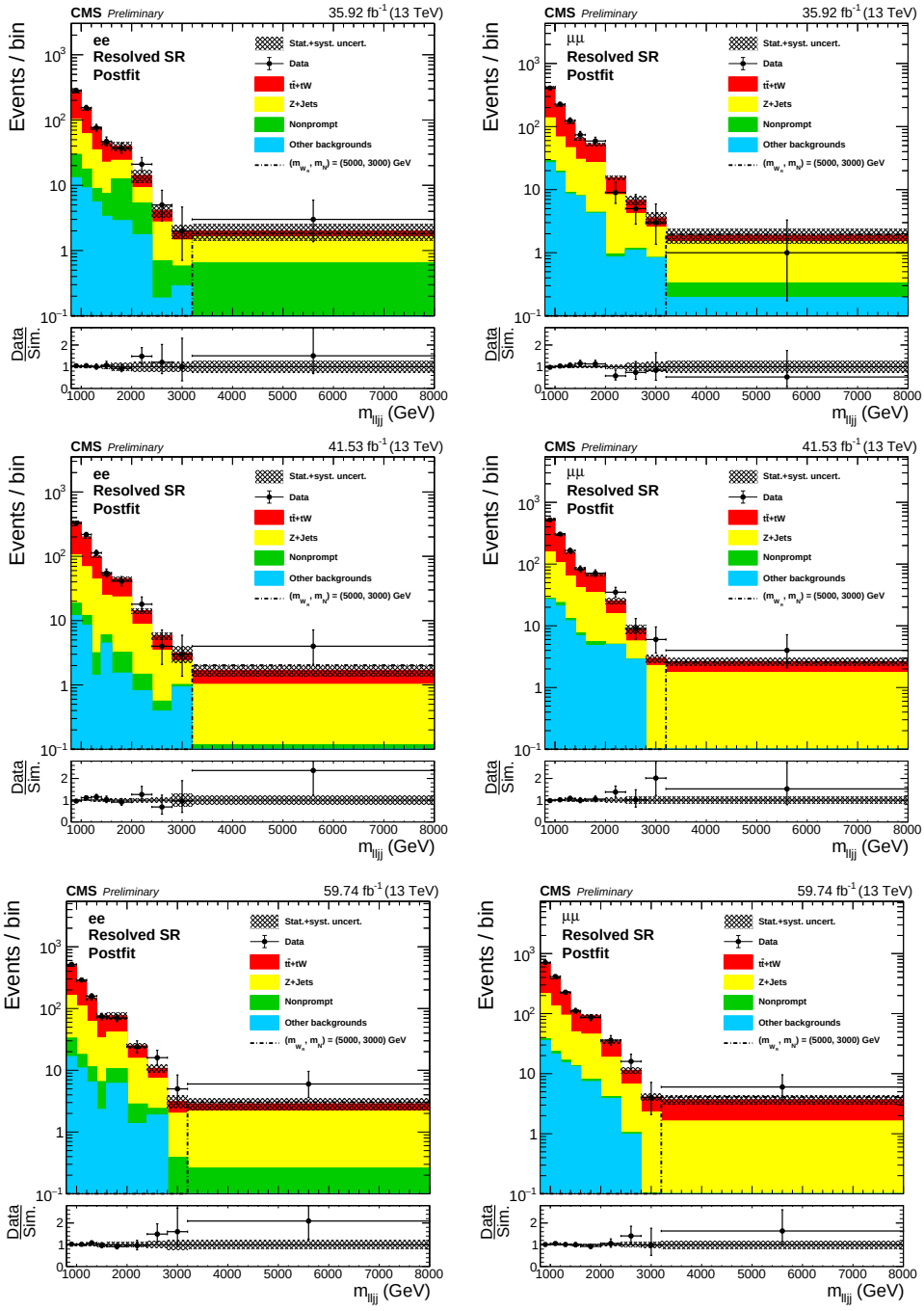


Figure A.11: The post-fit $m(\ell\ell jj)$ and $m(\ell J)$ distribution in the signal regions with three years individually. Results for the di-electron (di-muon) channel is shown on the left (right), for the resolved selection.

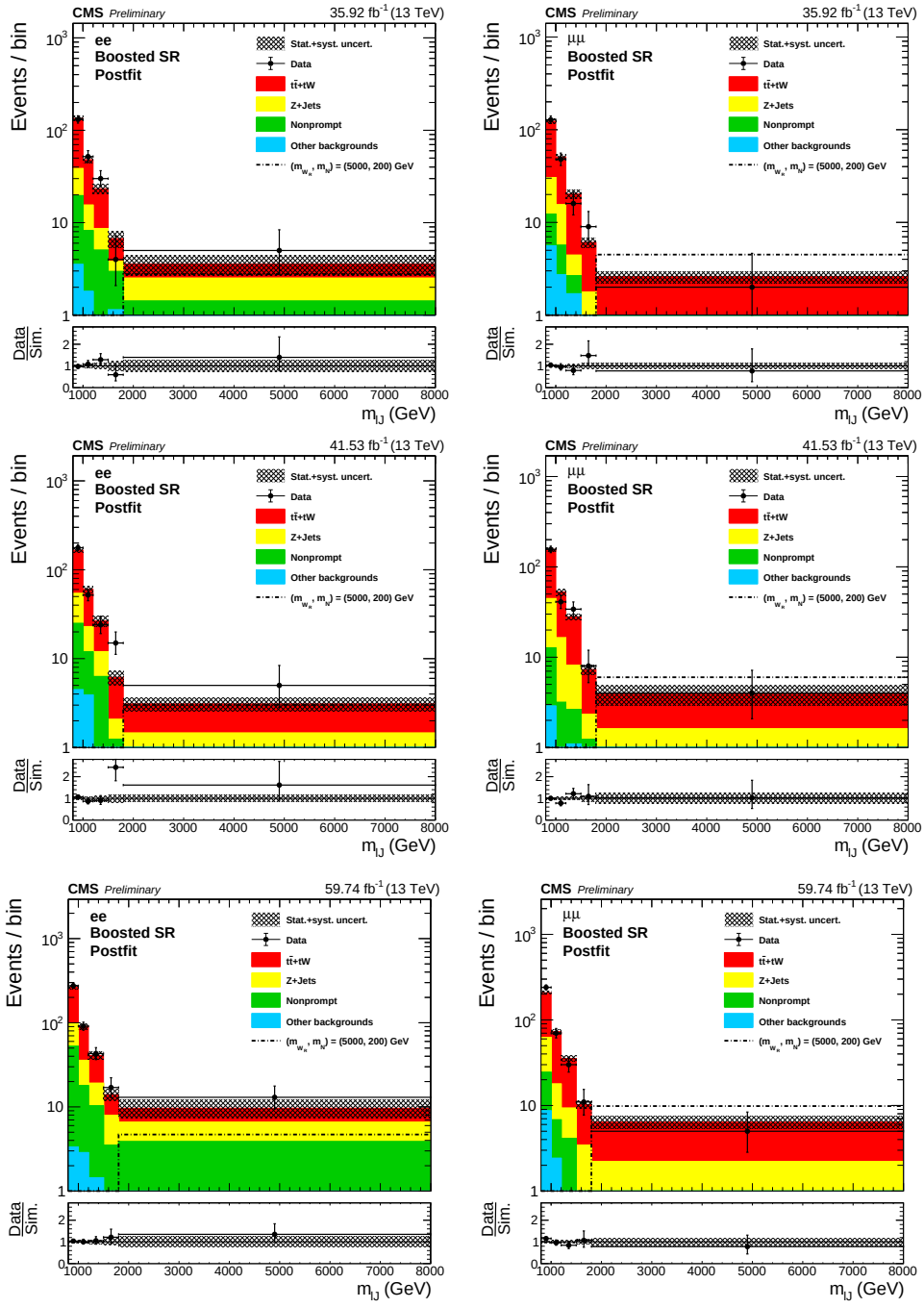


Figure A.12: The post-fit $m(\ell\ell jj)$ and $m(\ell J)$ distribution in the signal regions with three years individually. Results for the di-electron (di-muon) channel is shown on the left (right), for the boosted selection.

Appendix B

All Backgrounds

This appendix presents a list of all of the backgrounds simulations used by this analysis (Table B.1, Table B.2, and Table B.3). Each sample is generated for every year and for 2017 and 2018 only relevant detector conditions are changed. However, the Monte-Carlo simulation generator configurations differ slightly for some samples in 2016. The standard model process, its cross-section, and the effective number of years of luminosity generated is shown.

Table B.1: The list of simulation samples and corresponding cross-sections used in the 2016 analysis.

Process	Generator	Cross-section (pb)	Effective Years	Cross-section computed order
DY ($10 < m(\ell\ell) < 50$ GeV)	TuneCUETP8M1_13TeV-madgraphMLM-pythia8	18610	0.05	NLO
DY ($m(\ell\ell) > 50$ GeV)	TuneCUETP8M1_13TeV-madgraphMLM-pythia8	6077.22	0.67	NNLO
DY ($m(\ell\ell) > 50$ GeV, H_T 70 – 100 GeV)	TuneCUETP8M1_13TeV-madgraphMLM-pythia8	208.977	1.29	NNLO
DY ($m(\ell\ell) > 50$ GeV, H_T 100 – 200 GeV)	TuneCUETP8M1_13TeV-madgraphMLM-pythia8	181.30	0.52	NNLO
DY ($m(\ell\ell) > 50$ GeV, H_T 200 – 400 GeV)	TuneCUETP8M1_13TeV-madgraphMLM-pythia8	50.4177	0.65	NNLO
DY ($m(\ell\ell) > 50$ GeV, H_T 400 – 600 GeV)	TuneCUETP8M1_13TeV-madgraphMLM-pythia8	6.98394	5.25	NNLO
DY ($m(\ell\ell) > 50$ GeV, H_T 600 – 800 GeV)	TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1.68141	105.04	NNLO
DY ($m(\ell\ell) > 50$ GeV, H_T 800 – 1200 GeV)	TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.775392	118.05	NNLO
DY ($m(\ell\ell) > 50$ GeV, H_T 1200 – 2500 GeV)	TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.186222	109.61	NNLO
DY ($m(\ell\ell) > 50$ GeV, H_T 2500 – ∞ GeV)	TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.00438495	3119.7	NNLO
$t\bar{t}$ (semi-leptonic)	TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8	365.34	11.63	NNLO
$t\bar{t}$ (leptonic)	TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8	88.29	24.95	NNLO
W+Jets (leptonic)	TuneCUETP8M1_13TeV-madgraphMLM-pythia8	61526.7	0.01	NNLO
W+Jets (leptonic, H_T 70 – 100 GeV)	TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1637.1	0.17	NNLO
W+Jets (leptonic, H_T 100 – 200 GeV)	TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1627.45	0.21	NNLO
W+Jets (leptonic, H_T 200 – 400 GeV)	TuneCUETP8M1_13TeV-madgraphMLM-pythia8	435.237	0.38	NNLO
W+Jets (leptonic, H_T 400 – 600 GeV)	TuneCUETP8M1_13TeV-madgraphMLM-pythia8	59.1811	1.12	NNLO
W+Jets (leptonic, H_T 600 – 800 GeV)	TuneCUETP8M1_13TeV-madgraphMLM-pythia8	14.5805	8.72	NNLO
W+Jets (leptonic, H_T 800 – 2300 GeV)	TuneCUETP8M1_13TeV-madgraphMLM-pythia8	6.65621	7.82	NNLO
W+Jets (leptonic, H_T 1200 – 2500 GeV)	TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1.60809	5.12	NNLO
W+Jets (leptonic, H_T 2500 – ∞ GeV)	TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.0389136	219.98	NNLO
WW	TuneCUETP8M1_13TeV-pythia8	118.7	0.23	NLO
WZ	TuneCUETP8M1_13TeV-pythia8	47.13	0.59	NLO
ZZ	TuneCUETP8M1_13TeV-pythia8	16.523	1.67	NLO
WWW	TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.2086	32.03	NLO
WWZ	TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.1651	42.16	NLO
WZZ	TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.05565	123.46	NLO
ZZZ	TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.01398	496.33	NLO
ttW (leptonic)	TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	0.2043	294.36	NLO
ttW (hadronic)	TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	0.4062	57.11	NLO
ttZ	madgraphMLM-pythia8	0.5407	508.88	NLO
Single-top, s-channel	PSweights-amcatnlo-pythia8	3.36	18.61	NLO
Single-top, tW-channel – anti-top	powheg_TuneCUETP8M1	19.20	0.78	NNLO
Single-top, tW-channel – top	powheg_TuneCUETP8M1	19.20	0.77	NNLO
Single-top, t-channel – anti-top	PSweights-powhegV2-madspin	80.95	13.35	NLO
Single-top, t-channel – top	PSweights-powhegV2-madspin	136.02	13.73	NLO

Table B.2: The list of simulation samples and corresponding cross-sections used in the 2017 analysis.

Process	Generator	Cross-section (pb)	Effective Years	Cross-section computed order
DY ($10 < m(\ell\ell) < 50$ GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	18610	0.10	NLO
DY ($m(\ell\ell) > 50$ GeV)	TuneCP5_13TeV-amcatnloFXFX-pythia8	6077.22	2.40	NNLO
DY ($m(\ell\ell) > 50$ GeV, H_T 70 – 100 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	208.977	1.08	NNLO
DY ($m(\ell\ell) > 50$ GeV, H_T 100 – 200 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	181.30	1.83	NNLO
DY ($m(\ell\ell) > 50$ GeV, H_T 200 – 400 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	50.4177	6.05	NNLO
DY ($m(\ell\ell) > 50$ GeV, H_T 400 – 600 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	6.98394	43.34	NNLO
DY ($m(\ell\ell) > 50$ GeV, H_T 600 – 800 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	1.68141	154.01	NNLO
DY ($m(\ell\ell) > 50$ GeV, H_T 800 – 1200 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	0.775392	118.91	NNLO
DY ($m(\ell\ell) > 50$ GeV, H_T 1200 – 2500 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	0.186222	99.48	NNLO
DY ($m(\ell\ell) > 50$ GeV, H_T 2500 – ∞ GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	0.00438495	2832.12	NNLO
$t\bar{t}$ (semi-leptonic)	TuneCP5_13TeV-powheg-pythia8	365.34	7.25	NNLO
$t\bar{t}$ (leptonic)	TuneCP5_PSweights_13TeV-powheg-pythia8	88.29	0.26	NNLO
W+Jets (leptonic)	TuneCP5_13TeV-madgraphMLM-pythia8	61530	0.02	NNLO
W+Jets (leptonic, H_T 70 – 100 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	1637.1	0.33	NNLO
W+Jets (leptonic, H_T 100 – 200 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	1627.45	0.62	NNLO
W+Jets (leptonic, H_T 200 – 400 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	435.237	1.25	NNLO
W+Jets (leptonic, H_T 400 – 600 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	59.1811	6.0	NNLO
W+Jets (leptonic, H_T 600 – 800 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	14.5805	40.62	NNLO
W+Jets (leptonic, H_T 800 – 1200 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	6.65621	91.69	NNLO
W+Jets (leptonic, H_T 1200 – 2500 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	1.60809	454.2	NNLO
W+Jets (leptonic, H_T 2500 – ∞ GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	0.0389136	64690.39	NNLO
WW	TuneCP5_13TeV-pythia8	118.7	2.47	NLO
WZ	TuneCP5_13TeV-pythia8	47.13	3.43	NLO
ZZ	TuneCP5_13TeV-pythia8	16.523	3.82	NLO
WWW	TuneCP5_13TeV-amcatnlo-pythia8	0.2086	27.7	NLO
WWZ	TuneCP5_13TeV-amcatnlo-pythia8	0.1651	36.46	NLO
WZZ	TuneCP5_13TeV-amcatnlo-pythia8	0.05565	108.17	NLO
ZZZ	TuneCP5_13TeV-amcatnlo-pythia8	0.01398	430.6	NLO
ttW	TuneCP5_13TeV_madgraphMLM_pythia8	0.4611	335.04	NLO
ttZ	TuneCP5_13TeV_madgraphMLM_pythia8	0.5407	431.9	NLO
Single-top, s-channel	TuneCP5_PSweights_13TeV-amcatnlo-pythia8	3.36	63.83	NLO
Single-top, tW-channel – anti-top	TuneCP5_13TeV-powheg-pythia8	19.2	5.49	NNLO
Single-top, tW-channel – top	TuneCP5_PSweights_13TeV-powheg-pythia8	19.2	5.38	NNLO
Single-top, t-channel – anti-top	TuneCP5_PSweights_13TeV-powheg-pythia8	80.95	1.3	NLO
Single-top, t-channel – top	TuneCP5_13TeV-powheg-pythia8	136.02	1.27	NLO

Table B.3: The list of simulation samples and corresponding cross-sections used in the 2018 analysis.

Process	Generator	Cross-section (pb)	Effective Years	Cross-section computed order
DY ($10 < m(\ell\ell) < 50$ GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	18610	0.04	NLO
DY ($m(\ell\ell) > 50$ GeV)	TuneCP5_13TeV-amcatnloFXFX-pythia8	6077.22	0.28	NNLO
DY ($m(\ell\ell) > 50$ GeV, H_T 70 – 100 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	208.977	0.80	NNLO
DY ($m(\ell\ell) > 50$ GeV, H_T 100 – 200 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	181.30	0.35	NNLO
DY ($m(\ell\ell) > 50$ GeV, H_T 200 – 400 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	50.4177	1.05	NNLO
DY ($m(\ell\ell) > 50$ GeV, H_T 400 – 600 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	6.98394	1.73	NNLO
DY ($m(\ell\ell) > 50$ GeV, H_T 600 – 800 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	1.68141	25.72	NNLO
DY ($m(\ell\ell) > 50$ GeV, H_T 800 – 1200 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	0.775392	26.21	NNLO
DY ($m(\ell\ell) > 50$ GeV, H_T 1200 – 2500 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	0.186222	118.98	NNLO
DY ($m(\ell\ell) > 50$ GeV, H_T 2500 – ∞ GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	0.00438495	6849.62	NNLO
$t\bar{t}$ (semi-leptonic)	TuneCP5_13TeV-powheg-pythia8	365.34	4.65	NNLO
$t\bar{t}$ (leptonic)	TuneCP5_PSweights_13TeV-powheg-pythia8	88.29	12.19	NNLO
W+Jets (leptonic)	TuneCP5_13TeV-madgraphMLM-pythia8	61530	0.02	NNLO
W+Jets (leptonic, H_T 70 – 100 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	1637.1	0.29	NNLO
W+Jets (leptonic, H_T 100 – 200 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	1627.45	0.35	NNLO
W+Jets (leptonic, H_T 200 – 400 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	435.237	1.05	NNLO
W+Jets (leptonic, H_T 400 – 600 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	59.1811	1.73	NNLO
W+Jets (leptonic, H_T 600 – 800 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	14.5805	25.72	NNLO
W+Jets (leptonic, H_T 800 – 1200 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	6.65621	26.21	NNLO
W+Jets (leptonic, H_T 1200 – 2500 GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	1.60809	118.98	NNLO
W+Jets (leptonic, H_T 2500 – ∞ GeV)	TuneCP5_13TeV-madgraphMLM-pythia8	0.0389136	6849.62	NNLO
WW	TuneCP5_13TeV-pythia8	118.7	1.73	NLO
WZ	TuneCP5_13TeV-pythia8	47.13	2.36	NLO
ZZ	TuneCP5_13TeV-pythia8	16.523	2.73	NLO
WWW	TuneCP5_13TeV-amcatnlo-pythia8	0.2086	19.26	NLO
WWZ	TuneCP5_13TeV-amcatnlo-pythia8	0.1651	25.35	NLO
WZZ	TuneCP5_13TeV-amcatnlo-pythia8	0.05565	75.2	NLO
ZZZ	TuneCP5_13TeV-amcatnlo-pythia8	0.01398	299.34	NLO
ttW	TuneCP5_13TeV_madgraphMLM_pythia8	0.4611	465.28	NLO
ttZ	TuneCP5_13TeV_madgraphMLM_pythia8	0.5407	701.09	NLO
Single-top, s-channel	TuneCP5_PSweights_13TeV-amcatnlo-pythia8	3.36	89.36	NLO
Single-top, tW-channel – anti-top	TuneCP5_13TeV-powheg-pythia8	19.2	3.65	NNLO
Single-top, tW-channel – top	TuneCP5_PSweights_13TeV-powheg-pythia8	19.2	4.6	NNLO
Single-top, t-channel – anti-top	TuneCP5_PSweights_13TeV-powheg-pythia8	80.95	19.5	NLO
Single-top, t-channel – top	TuneCP5_13TeV-powheg-pythia8	136.02	22.8	NLO