UNIVERSITY OF CALIFORNIA, SAN DIEGO

A Search for Dark Particles Produced in Association with the Z Boson and Jets in 13 TeV Proton-Proton Collisions

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

in

Physics

by

Bobak Hashemi

Committee in charge:

Professor Frank Wüerthwein, Chair
Professor Avraham Yagil, Co-Chair
Professor Daniel Arovas
Professor John McGreevy
Professor David Meyer

2018

The dissertation of Bobak Hashemi is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____
Co-Chair

_____
Chair

University of California, San Diego

2018

DEDICATION

To my Mother and Father, who taught me to reach and keep reaching.

EPIGRAPH

The most incomprehensible thing about the universe is that it is comprehensible

—Albert Einstein

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Where to start with a project as large as a CMS analysis? The CMS collaboration is a group of thousands of scientists, engineers, technicians, students, and staff from over 200 institutes in over 40 countries. The collaboration built the CMS detector, and now operates and processes the data it delivers, allowing scientists like me to analyze the output of the proton-proton collisions. Without the thousands of people who have spent decades designing, troubleshooting, organizing, fund-raising, ect... the work in this dissertation would not be remotely possible. I thank my co-authors in the collaboration for their essential contributions to this work.

The CMS detector itself is located at CERN near Geneva Switzerland. Again here, we have thousands of people who have worked for decades ensuring that CMS has proton beams colliding in its belly at extremely high energy today. Again here it is impossible to overstate the absolute necessity of the LHC team and CERN staff to this work.

Within the CMS collaboration, there are several individuals whose contributions to this work are much more direct. Pablo Martinez Ruiz del Arbol, Christian Schomakers, Leonora Vesterbacka, and Sergio Sánchez Cruz all contributed to this analysis, providing numerical factors, cross validations, discussions about methodology, and so on. Here I must also mention the tremendous leadership and mentorship provided to me by Dominick Olivito and Vince Welke, both of whom not only contributed code and man-hours, but also gave me direction, engaged me intellectually, and taught me just about everything I know about this search. These individuals, along with myself, constitute the Edge-Z team within the CMS SUSY group.

Next, I would like to thank the rest of the members of the CMS SUSY group for their important comments, feedback, and peer-review of the search, specifically Cristina Botta and Lesya Shchutska who were the conveners while this search was being executed. In addition to the SUSY group, we received important feedback from many members of CMS at large, and a review committee consisting of Marcello Maggi, Niki Saoulidou, Hongxuan Liu, and Christopher West scrutinized our search more thoroughly and provided crucial comments about the presentation of

our methods.

I must also acknowledge SNT, an unofficial group of collaborators working within CMS including: P. Chang, M. Derdzinski, D. Gilbert, D. Klein, V. Krutelyov M. Masciovecchio, S. May, D. Olivito, F. Würthwein, A. Yagil, G. Zevi Della Porta, N. Amin, C. Campagnari, B. Marsh, S. Wang, L. Bauerdick, K. Burkett, O. Gutsche, S. Jindariani, M. Liu, H. Weber, F. Golf, I. Suarez. Almost all of the members of SNT have, to some degree, been important to the success of this project. Whether through discussion, contribution of code, or data processing, being part of the SNT team ensured I never could be stuck for too long on any issue.

One more time I would like to Dominick Olivito who has been both a cherished friend and irreplaceable mentor to me throughout my time at CMS. In addition to Dominick, my co-advisors Frank Wüerthwein and Avraham Yagil were incredibly patient and good-natured PIs who trained me to be a particle physicist and never let me forget the importance of clear presentation skills, practice, and clear-thinking.

I would also like to thank Dylan Gilbert, Philip Chang, Mark Derdzinski, and Nick Amin for important discussions about physics, software, and work in HEP. Mark and Nick also contributed to this work by processing data from CMS into more accessible data formats for the SNT group at large.

Finally, I would like to specifically thank the following senior members of SNT who have also been tremendous mentors to me: Giovani Zevi Della Porta, Philip Chang, Mia Liu, Hannsjorg Weber, Claudio Campagnari, Frank Golf, and Indara Suarez.

In chapter 2, I must also specifically thank my co-authors on CMS for contributing figure 2.26 showing the B-tag CSV discriminator, figure 2.24 showing the MET resolution after the type-1 correction, and figure 2.25 showing the efficacy of the MET filters.

In chapter 3, I must also specifically thank my co-authors on CMS for contributing figure 3.17 (a), showing the previous best limits on the strong GMSB SUSY model.

VITA

2012	B. S. in Physics, Pennsylvania State University
2014	M. S. in Physics, University of California, San Diego
2012-2015	Graduate Teaching Assistant, University of California, San Diego
2018	Ph. D. in Physics, University of California, San Diego

PUBLICATIONS

M.W.E. Smith, et. al., *The Astrophysical Multimessenger Observatory Network (AMON)*, Astroparticle Physics (2013) 2013: 45. https://doi.org/10.1016/j.astropartphys.2013.03.003.

The CMS collaboration, Sirunyan, A.M., Tumasyan, A. et al., *Search for new phenomena in final states with two opposite-charge, same-flavor leptons, jets, and missing transverse momentum in pp collisions at* $\sqrt{s} = 13$ *TeV*, J. High Energ. Phys. (2018) 2018: 76. https://doi.org/10.1007/s13130-018-7845-2

The CMS collaboration, Sirunyan, A.M., Tumasyan, A. et al., *Combined search for electroweak production of charginos and neutralinos in proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, J. High Energ. Phys. (2018) 2018: 160. https://doi.org/10.1007/JHEP03(2018)160

ABSTRACT OF THE DISSERTATION

A Search for Dark Particles Produced in Association with the Z Boson and Jets in 13 TeV Proton-Proton Collisions

by

Bobak Hashemi

Doctor of Philosophy in Physics

University of California, San Diego, 2018

Professor Frank Wüerthwein, Chair
Professor Avraham Yagil, Co-Chair

This thesis presents the results of a search for new physics in proton-proton collisions at the Large Hadron Collider, running with 13 TeV center of mass energy, using data gathered by the Compact Muon Solenoid (CMS). The search targets TeV mass-scale dark matter candidates and uses final states with two opposite-charge and same-flavor light leptons (electrons or muons) having dilepton mass consistent with the Z boson, at least 2 hadronic jets, and at least 100 GeV of transverse momentum imbalance. The thesis contains a historical introduction to particle physics, brief reviews of the standard model and supersymmetry, an in-depth discussion of the acquisition

of data by the Large Hadron Collider and CMS detector, and a pedagogical overview of the analysis methods used. No statistically significant deviation is found from the expected standard model background. The search results are presented and interpreted in the context of several simplified models of supersymmetry, including a model of Gauge Mediated Supersymmetry-Breaking (GSMB) with gluino production and models with Electroweakino production. The excluded mass ranges for these models are advanced by approximately 50-100% with respect to the best previous searches. This work represents the current state of the art for their exclusion.

Chapter 1

Introduction

Particle physics aims to characterize the fundamental constituents of matter and the mechanisms by which they interact. The current state of the art theory is called *The Standard Model of Particle Physics* (henceforth, just the standard model), which is written in the framework of *Quantum Field Theory* (QFT).

Though the standard model has been shown to make accurate predictions in a wide range of particle physics experiments, there are no shortage of open questions as to the origin of its structure. Furthermore, the standard model does not attempt to describe gravitation, and the observation of massive neutrinos prove definitively that the standard model must be incomplete. Another observation which strongly implies a deficiency is the evidence of at least one so-called *dark matter* particle, an abundant particle which does not couple to the photon with any non-trivial strength. None-the-less, it is clear that the theory which supersedes the standard model must reduce to it in the appropriate limit, and its successes should not be downplayed.

This thesis presents the search for a dark matter candidate which couples either directly or indirectly to the Z boson. The final states studied include multiple hadronic jets, transverse momentum imbalance, and a pair of opposite-charge same-flavor leptons having dilepton mass consistent with the pole mass of the Z boson. These final states are motivated by the strengths of

the CMS detector as well as simplified models of supersymmetry, a systematic framework for extending the standard model.

1.1 Historical Context

Nearing the turn of the 20th century, the debate on the existence of elementary particles had still not been completely settled. But in the following 80 years, a wildly rich and successful theory of matter would be developed, culminating in what is called the standard model, which has been called the best tested theory in science.

1.1.1 The evolution of particle physics

J.J. Thomson is credited with finding the very first elementary particle, [54] the electron. Thomson observed that cathode rays, now known to be streams of electrons, would bend under the influence of a magnetic field. This implied that rays were a beam of particles, as opposed to some sort of aether phenomena, the competing theory at the time, because they responded in accordance with the known force law for charged particles. Thomson was able to deduce the charge to mass ratio for the electrons and compared it to the known values for ionic gases. Thomson found that the electron had a $\frac{q}{m}$ which was strikingly smaller than any ionic gas, 3 orders of magnitude smaller than hydrogen ions.

Thomson's student, Ernest Rutherford pushed the field of particle physics further when he discovered the nucleus, and later the proton. Rutherford famously shot alpha particles (hydrogen nuclei) at a thin sheet of gold. He found that some particles were deflected backwards, though the vast majority passed through the sheet. This implied the gold was actually a lumpy collection of heavy (when compared to the alpha particles) particles, rather than a continuous material. Albert Einstein postulated the photon soon after, due to the photoelectric effect and Max Planck's famous solution to the ultraviolet catastrophe. The Neutron was discovered by James Chadwick in the

1930s after hearing about a series of experiments performed by German and French physicists which found that alpha particles striking beryllium would create a penetrating radiation that was not influenced by electric fields.

Also in the 1930s, Hideki Yukawa used the size of the atomic nucleus to predict the existence of pions, now known to be the carriers of the strong force whose existence counteracts the electromagnetic repulsion of protons in the nucleus. When looking for pions, cosmic ray experiments found muons, and mistakingly believed they had confirmed Yukawa's theory. Muons were eventually shown to not interact much with nuclei, and so they were ruled out as nuclear force carriers. Pions are produced in large numbers in the upper atmosphere, but tend to disintegrate on their way to the ground. They were eventually discovered by comic ray experiments performed in the Andes mountains.

During the same time period, Dirac predicted the existence of antiparticles using his mathematically correct, but philosophically misguided, theory of holes. The positron was discovered by Anderson when he exposed a cloud chamber to a magnetic field and discovered electron/positron pair production. Another bit of elementary particle physics uncovered during this fruitful 1930s era was that of the existence of the neutrino, which was famously theorized by Wolfgang Pauli in order to salvage energy conservation in beta decay. Enrico Fermi unified Pauli's idea with the discovery of the neutron by hypothesizing that beta decay was the decay of a neutron into a proton, an electron, and a neutrino. The electron anti-neutrino was finally discovered by observing reverse beta capture. Further observations showed muons could decay to electrons, but only in association with two neutrinos. This lead to the notion that neutrinos must carry "lepton flavor," and that there must be a neutrino for each lepton.

By the 1950s, particle accelerators began to appear, and so too did the observation of a "zoo" of new particles in cloud chamber experiments. These particles included some mesons which had a *strangely* long lifetime, like the K^0, and a slew of other heavy particles. Observations suggested strange particle were always produced with another strange partner, but there was no

such restriction on their decays. That fact, in addition to the long lifetime of strange particles, hinted that their decays were mediated by a different force than their creation. This lead to the notion of conservation of strangeness whereby each particle was assigned a strangeness of 1, 0, or -1. Along with conservation of baryon number (theorized by Stückelberg to stabilize the proton), conservation of lepton number, and conservation of charge, a series of discrete symmetries began to appear for elementary particle interactions.

Using charge and strangeness values for particles, Murray Gell-Mann developed the famous "eightfold way," a scheme for laying out particles in geometrical patterns based on their charge and strangeness values. Arranging the particles in this way inferred the existence of a heavy particle with negative charge and -3 strangeness, now called the Ω^-, which had not been observed. The observation of the Ω^- ushered in an era of particle physics which was based on symmetry principles, a trend which still dominates the field to this day.

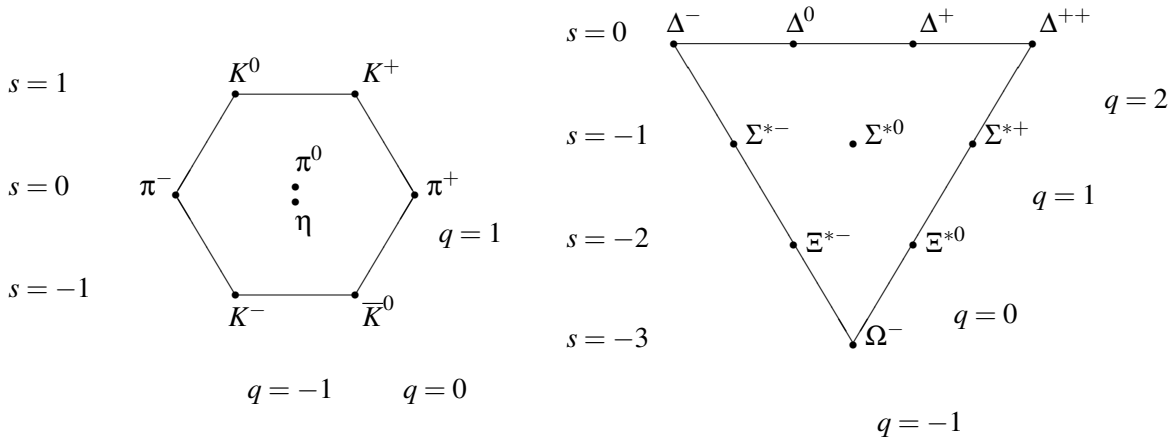


Figure 1.1: The meson octet (left) and the baryon decuplet (right). Murray Gell-Mann arrayed elementary particles in these patterns and used them to predict the existence of the Ω^-. This sparked a revolution in elementary particle physics based on discrete symmetries. Modified from [68]

Gell-Mann and Zweig independently proposed that the origin of these patterns were due to the fact that hadrons are made up of *quarks*. The quark model was successful at reproducing the predictions of the eightfold way, and gave a mechanism for why particles with certain electric charges, masses, and decay pathways exist while others don't. Additionally, deep inelastic scattering experiments, much the same as Rutherford's, showed that the charge inside protons also seem to be collected in lumps, and in a way that is consistent with the three fractionally charged partons described in the quark model. [50] The quark model eventually became encoded in Quantum Chromodynamics, which took its modern form in the early 1970s with the discovery of asymptotic freedom. The gauge bosons in QCD are called gluons. In 1979, gluon emission, and therefore gluons themselves, was experimentally observed in three jet final states using electron-positron collisions at DESY. [51]

In the 1930s, the theoretical problem of infinities in quantum field theory was systematically studied, specifically in the context of the self-energy of the electron. These problems were addressed by charge and mass renormalization, ideas born at the Shelter Island Conference in 1947, addressing the Lamb shift and anomalous magnetic moment of the electron. By the 1950s, much of the theoretical foundation for Quantum Field Theory had been laid in pursuit of a theory of Quantum Electrodynamics (QED). Julian Schwinger and Sin-itiro Tomonaga independently developed a formalism based on operator mathematics. This approach was shown to be equivalent to a formalism developed by Richard Feynman using path integrals by Freeman Dyson in 1949.[1] The theoretical progress during this time explained the observation of the Lamb shift in the spectrum of hydrogen, and the electron's anomalous magnetic moment, a computation which is experimentally verified to better than 1 part per billion and has elevated the standard model to be called the best tested theory in science. In formulating the equivalence of theories, Dyson laid out criteria to decide whether a theory was renormalizable, an important feature expected from any physical QFT.

[1] The prominent physicist Oppenheimer was so sure that Feynman's ideas were wrong that Dyson's proof earned him a lifetime appointment at the Institute for Advanced Study in Princeton without ever needing to earn a PhD.

In the 1950s an enormous theoretical breakthrough was made by Yang and Mills, who developed the concept of a Gauge theory. The gauge group SU(2)xU(1) was found to be at the heart of the unified theory of the electromagnetic and weak interactions, now called electroweak theory. Steven Weinberg invoked the Higgs mechanism in 1967 to bring the theory into modern form, showing that a massless theory in the high energy limit could "spontaneous break" into a theory with a massless photon but massive Ws and Zs at laboratory energy scales. [2] Weinberg's theory also required the existence of a scalar boson called the Higgs. [89] His theory was partially confirmed with the discovery of the W and Z bosons at the UA1 and UA2 experiments at CERN in 1983. The final prediction of the electroweak theory was verified in 2012 with the discovery Higgs boson decaying to 2 photons in 2012 by the CMS and ATLAS collaborations at the Large Hadron Collider. [40, 37] The search presented in this thesis was conducted within the CMS collaboration.

With the combination of QCD and electroweak theory, a total Lagrangian could be written with the symmetry group SU(3)xSU(2)xU(1), this is now what we called the standard model. But the particle content of the standard model still was being discovered into the next decade. In the 1970s, several breakthroughs were made. The discovery of the J/Psi meson in 1974 confirmed the existence of a fourth heavier quark, now called the charm. This brought the total number of known quarks to 4. However in the previous year experiments showed violation of CP symmetry in Kaon decays, and it was worked out by Kobayashi and Maskawa that a 4 quark model could not accommodate CP violation, at least 6 quarks were needed.

In 1975, the Tau lepton and its neutrino were discovered. This provided another hint that there were 6 quarks, as it seemed natural to have the same number of quarks and leptons. In 1977, the bottom quark was discovered at Fermilab, bringing the total number of known quarks 5. The top quark was finally detected 1995 in proton-antiproton collisions at Fermilab. This completed

[2] Weinberg's work was largely ignored for several years because the it was an example of a spontaneously broken non-Abelian gauge theory. The renormalizability of that type of theory was not shown until 1971 by 't Hooft and Veltman. [62] [51].

the currently understood quarks sector.

As described above, the standard model has successfully predicted the existence of the W, Z, and Higgs bosons, including their decays. It predicted the electron's anomalous magnetic moment to extremely high precisions. It predicted the existence of a third generation of quarks and leptons based on CP violation. Indeed, at the current state of affairs, the standard model it incorporates essentially all known particle phenomenology, with the notable exceptions of neutrino masses and dark matter.

1.2 The Standard Model

The standard model of particle physics is the combination of electroweak theory with quantum chromodynamics. In section 1.1.1 we covered some of the highlights in the experimental and theoretical progress made while constructing this theory. In this section, we will take a look at the particle content and interactions of the standard model.

Figure 1.2 is a table that summarizes all the particles in the standard model. In purple, the quarks are shown. Each quark comes in one of 3 colors and has an associated anti-quark which comes in one of 3 anti-colors. In other words, the 6 slots in the figure really correspond to 36 quantum fields and equally 36 different particles. Quarks carry baryon number $\frac{1}{3}$, and anti-quarks carry baryon number $\frac{-1}{3}$. A quark is called up-type if it is positively charged (like the up quark), and down-type if it is negatively charged. The gluon comes in 8 varieties, very roughly, each gluon carries a color and an anti-color. The charged leptons: electron, muon, and tau, each come in negatively charged and positively charged varieties, positively charged leptons are called the anti-leptons. In the standard model, the neutrinos each have a partner anti-particle, but it is not known whether anti-neutrinos exist or if neutrinos are their own anti-particles. Leptons carry lepton number $\frac{1}{3}$ and anti-leptons carry lepton number $\frac{-1}{3}$. The only other hidden detail is the W boson, which comes in both positively and negatively charged variates.

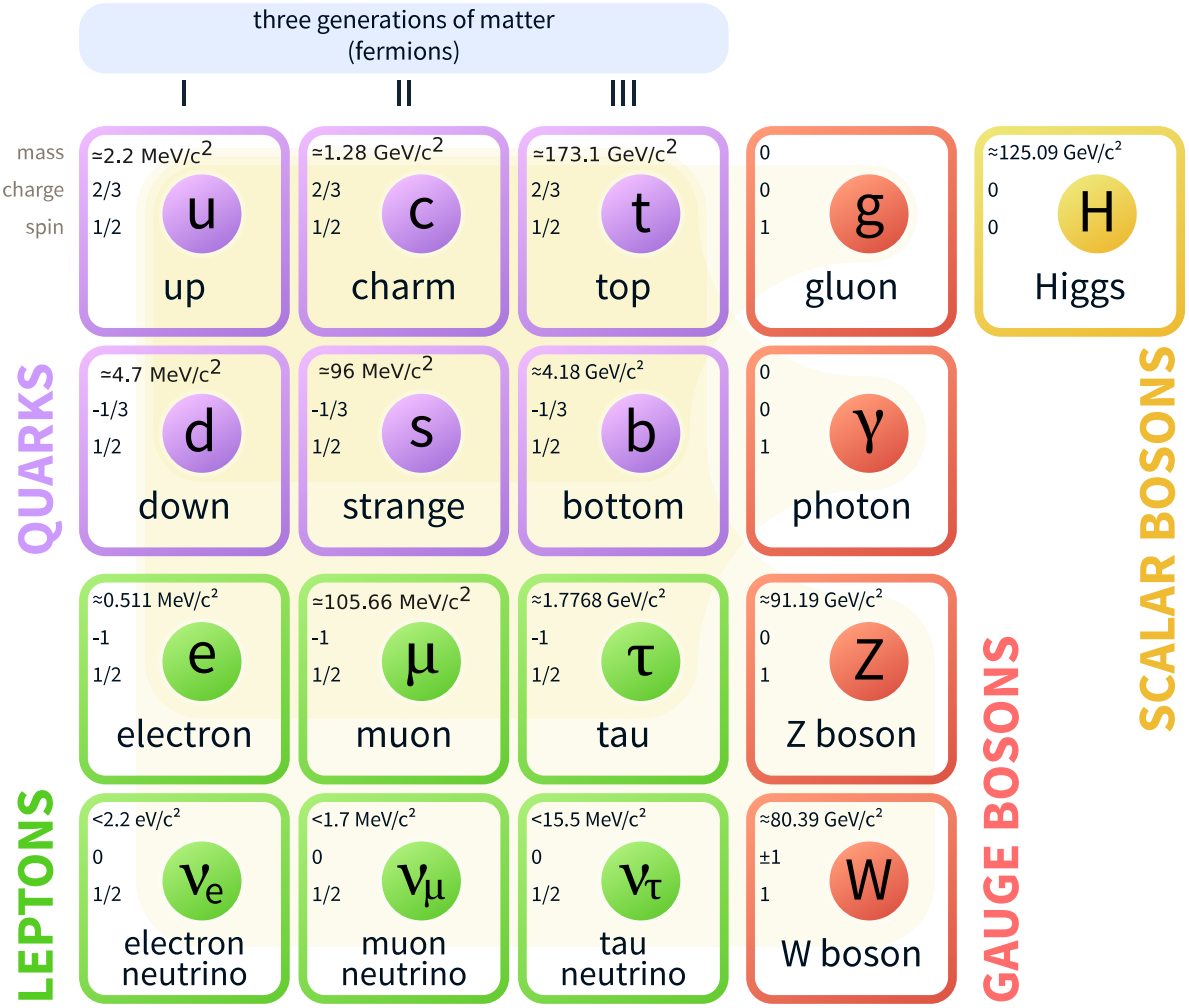


Figure 1.2: A table showing the particle content of the standard model of particle physics. Taken from [75]

Each particle in the standard model is associated with a tensor field [3] that permeates all of spacetime, excitations of the fields manifest in reality as particles. The motion of particles is characterized by the propagation of excitations in spacetime. Analogously, a rock thrown into a pond will cause the water level to shift slightly downward before it breaks the surface tension.

[3] A tensor a sort of generalization of a matrix, it includes real and complex numbers, vectors, matrices, and higher dimensional analogs of matrices. A tensor field associates one of these objects with each point in spacetime. For instance, a "scalar field" associates a complex number with each point in space and time. Electromagnetism associates the "Maxwell-stress tensor", a 4x4 matrix, with each point in spacetime.

The deviation of the water height from equilibrium is a type of excitation, and it travels outward from the location the rock landed.

The particles in the standard model are broken up into bosons and fermions. The fermions are the quarks and the leptons, the bosons are photon, gluon, W, Z, and Higgs. The difference between bosons and fermions is characterized by their spin, which is intrinsic angular momentum a particle has even when it is at rest. Fermions have spin whose magnitude is a half integer multiple of $\hbar$, e.g. $\frac{\hbar}{2}$, $\frac{3\hbar}{2}$, etc... All standard model fermions have spin $\frac{\hbar}{2}$, typically we use units where $\hbar = 1$ and say standard model fermions have *spin half*. Bosons have spin whose magnitude is an integer multiple of $\hbar$, e.g. $0, \hbar, 2\hbar$, etc... All the gauge bosons in the standard model have spin 1, the only boson with non unity spin is the Higgs, which has spin 0.

Every excitation of a quantum field takes some amount of energy. For a massive particle, the minimum amount of energy needed to make an excitation is called the mass of the particle. For instance, the minimum amount of energy needed in the muon field to make an excitation (a muon) is approximately 105 MeV. While the particle exists, this energy is trapped at the location of the particle in space. Quantum fields can exchange energy, such an event is called an *interaction*. In an interaction, the energy stored in one or more quantum fields is funneled into one or more other quantum fields. For instance, a muon can decay to an electron, a muon neutrino, and an electron anti-neutrino, meaning that the energy stored in the muon was redistributed into the electron field, the muon neutrino field, and the electron anti-neutrino field.

All interactions in QFT are local, meaning that energy can only be exchanged at roughly the same point in space and time. As an example, a muon on mars in 1970 can not create an electron and two neutrinos today in Geneva, Switzerland. Further, all interactions in the standard model are mediated by the bosons. This means a fermion can not exchange energy directly with another fermion without exchanging some energy with an appropriate boson field. In the example of the muon decaying into an electron, a muon neutrino, and an electron anti-neutrino, first the muon decays into a W^- boson and a muon neutrino, then the W^- boson decays into an electron

and an electron anti-neutrino.

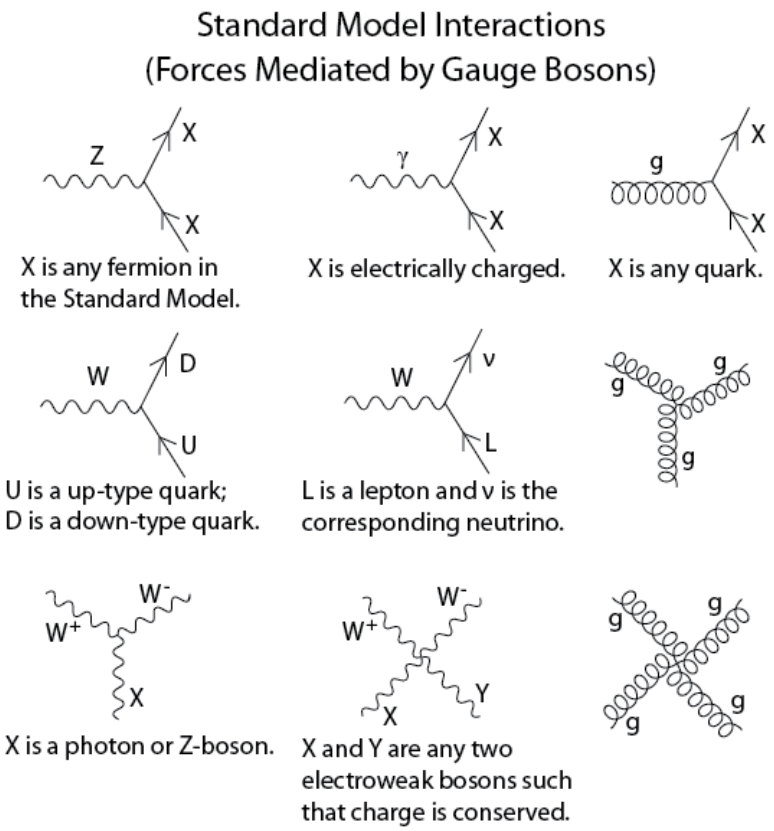


Figure 1.3: A list of the interactions in allowed in the standard model. Taken from [57]

Figure 1.3 shows the interactions allowed in the standard model. Roughly, the gluons mediate the strong interaction and can only interact with colored objects, meaning quarks and other gluons. The W and Z bosons mediate the weak interaction and can interact with any fermion as well as among themselves whilst conserving charge, with the exception of triple Z and quadruple Z interactions. The photon can interact with any charged particle.

The rules in 1.3 are called Feynman diagrams. These diagrams encode all of the discrete conservation laws in the standard model. The existence of any particle decay or production can be predicted with these rules by first checking whether these vertices can be connected together such that it turns the initial state into the desired final state, and then ensuring energy and momentum

conservation will not be violated in the process.

Protons

The LHC is a proton-proton collider. In these collisions, the initial state particles are any particles found in protons. Figure 1.4 shows the chance of an interaction with the different constituents of the proton in a collision, expressed as a function of the fraction of the proton's energy carried by the interacting fundamental particle.

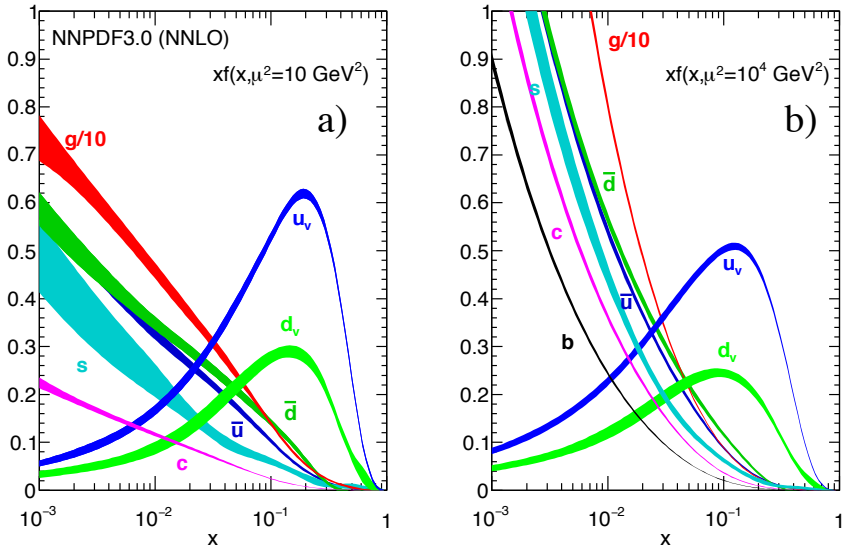


Figure 1.4: The probability of colliding with a quark, by flavor, or gluon with fraction x of the proton's total energy in a proton collision. The y-axis is the probability scaled by the energy fraction x. The gluon line has been scaled down by a factor of 10 to fit nicely with the rest of the curves, in other words, gluons are the most likely fundamental particle to interact in proton collisions. On the left (a), the curves are drawn for protons with 10 GeV of energy, on the right (b), with 100 GeV. Proton collisions at the LHC are done at approximately 7,000 GeV. The peaks for the u and d lines show the valence quark content. Taken from [76, sec. "Structure Functions"]

Of the quarks, notice the most likely objects are the u and d quarks, in fact, there is roughly twice the chance for a u quark than a d quark. This encodes the fact that the proton is a bound QCD state of two up quarks and one down quark. However, there is still some chance to get any quark when colliding protons, this is due to the famous "quark sea;" the energy bound up in the proton is in a superposition of particle states. One consequence of the sea for particle colliders

is that the valence quark content of the particles in the collision is almost inconsequential for producing particles much lighter than the center of mass energy of the collision.

As a specific example, the top two entries of the left row in figure 1.3 show that generating a W or Z boson from quarks requires at least one anti-quark. The naive picture of the proton as containing three quarks, uud, would predict then predict it is impossible to collide protons and produce a W or Z. However, because the W and Z are 100 times lighter than the center of mass energy of collisions, the chance of pulling an anti-up or anti-down quark from the sea with sufficient energy is large enough to compensate for the fact that there are no valiance anti-quarks.

In fact, the W and Z are only likely to be produced if the energy of the quarks sums to approximately 100 GeV. At large enough energies, using valence quark becomes undesirable for the production of W and Z bosons, and the cross section is dominated by sea quark seq antiquark collision. A small deviation in the W and Z cross sections between proton-proton and proton-antiproton collisions can be seen at around 2 TeV in figure 2.3. The small dip in the cross section there reflects the change of source for the data from $p - \bar{p}$ collisions at Tevatron to $p - p$ collisions at the LHC.

1.3 Problems with the Standard Model

Though the standard model has had incredible successes, there are still many open questions. The following list some observational and theoretical motivations for physics beyond the standard model. No attempt at completeness is made.

Observational Issues

1. **Dark Matter –** Observations in astronomy show that the ratio of luminous matter to non-luminous matter in large-scale astrophysical bodies does not agree with standard model predictions. A few key observations are: The rotation speed of objects in spiral galaxies

is higher than expected at the galactic edges, the commonly accepted conclusion drawn from this observation is the existence of a dark matter halo that encloses the galaxies. In larger scales, galaxy clusters also provide three independent mass tests: the rotational speed of galaxies in the cluster, the x-ray energy spectrum and flux, and the strength of gravitational lensing. All three of these measurements agree with the dark matter hypothesis with an approximate dark-matter to luminous matter ratio of 5 to 1. [14] Next, the effects of dark matter in the early universe create a signature in the cosmic microwave background polarization, the WMAP and Planck satellite experiments observed this signature and found it to be in excellent agreement with the ΛCDM dark matter model. [84] [12] Finally, the bullet cluster is a the result of a collision of two galaxies and gravitational lensing observations show that the center of mass of this cluster is well offset from the center of luminous mass, a prediction of dark matter without strong self-interactions. [22] [76, ch. 26]

2. **Gravitation** – There is no mechanism for gravitation in the standard model. Attempts to build a quantum theory of gravity using standard methods yields a non-renormalizable theory, inferring that the high energy physics is not understood. The quest to quantize gravitation has been ongoing for almost 100 years.

3. **Neutrino Masses** – Neutrinos are modeled as massless in the standard model, however neutrino oscillations have shown that they must have some mass. Massive neutrinos can be either Dirac or Majorana particles, current experiments are attempting the decide which is the appropriate model and to find the precise neutrino masses.

4. **Spacetime Expansion** – Observations show that spacetime is expanding, an effect first seen by Edwin Hubble, and is doing so at an accelerating rate. The most common and straightforward explanation for this fact is the inclusion of a cosmological constant term in general relativity, which, by dimensional analysis, is amenable to being interpreted as a

constant energy density that permeates spacetime. Therefore, it is sensible to identify the cosmological constant term with the energy density of the vacuum in QFT. [80] Although no direct calculation of this quantity exists for the standard model, back of the envelope calculations show that the vacuum energy of the standard model is expected to be between 50 and 120 orders of magnitude off from the measured value for the cosmological constant from astrophysical observations. This is known as the cosmological constant problem. [24] [90]

5. **Anomalous μ g-factor** – Although this result is still being tested, there are some indications that the muon magnetic moment differs from the standard model prediction by 3 standard deviations.[58] This result becomes somewhat more convincing when considering the electron magnetic moment is considered one of the most accurate predictions in science, as was previously discussed.

Theoretical Issues

1. **GUT Scale Renormalization** – There is precedent to assume that at high enough energies, all three forces in the standard model combine into a single force, a Lagrangian with this feature is called a grand unified theory (GUT). The energy at which this occurs is typically called the GUT scale. Using the methods of renormalization, one can find the GUT scale to be between 10^{15} and 10^{16} GeV. Using renormalization theory, the expected strength of the gauge interactions can be extrapolated to that energy, as can be seen in figure 1.6. However, the standard model does not seem to show that the interaction strengths will converge at this energy. Theorists take this as a hint that the model is incomplete.

2. **Naturalness and the Hierarchy problem** – There are several energy scales in the standard model. The two heaviest are the electroweak scale at roughly 100 GeV and the Planck scale at roughly 10^{18} GeV. The electroweak scale is the mass scale of the massive electroweak

bosons and the Higgs, this is where electrodynamics and the weak force combine into a single force. The Planck scale is where quantum gravitational effects are expected to become important and is the typical cutoff scale for standard model calculations[4]. The fact that these two scales are wildly different is called the Hierarchy problem, which is also sometimes presented as the difference between the strength of gravity and the weak interaction.

A complete exposition of theoretical aspects of the scale difference is beyond the scope of this thesis, however, the flavor of the issue is that the wildly different scales create odd interplays and "fine-tuning" in standard model calculations. For instance, when computing the mass of the Higgs boson, the one-loop correction to the Higgs mass from a fermion f is

$$\Delta m_H = \frac{-|\lambda_f|^2}{8\pi^2}[\Lambda_{UV}^2 + ...]$$

where Λ_{UV} is the scale at which the theory is expected to break down, taken to be the Planck scale if no new physics exists between the standard model and quantum gravity. Therefore, the Higgs mass, measured to be 125 GeV, gets quantum corrections of order the Planck mass squared, some 30-odd orders of magnitude larger. In other words, the canonical theory requires that the sum of infinitely many terms, all of order 10^{36} should sum to some number which is order 10^2 GeV.

It is important to note that theorists disagree about the validity of this argument, as it may be reading too deeply into a perturbative calculation. However, the oddity of this calculation was one of the main driving influences behind the development of supersymmetry. The addition of supersymmetric particles cancels the quadratic dependence on the cutoff scale in this calculation.

3. **Why *these* parameters** – The standard model has 19 free parameters. The current theory

[4] if using a cutoff for regularization

provides no clues as to why those 19 numbers have the values they have.

4. **3 Generations** – Figure 1.2 shows that the fermions in the SM can be grouped into three very distinct generations. The particles in these different generations are identical in all respects except for their masses. This striking feature begs for an explanation, but like the choice of parameters, the standard model has nothing to say about why three generations of fermions exist.

5. **Strong CP Problem** – The strong interaction obeys a symmetry called CP: interaction rates are identical for a process if all particles are exchanged for their anti-particles, and the particle's spins are flipped. However, the weak force does not obey this symmetry, implying it is not a generic symmetry of nature. Further, the QCD Lagrangian is written in such a way that CP violation is readily accommodated, ensuring CP symmetry in QCD requires that parameters in the theory be fine tuned such that their CP violation cancels in a seemingly contrived manner. This, like the hierarchy problem above, is seen as an oddity which calls for a dynamical mechanism to ensure these parameters have a relationship that protects CP symmetry. The currently favored theoretical fix is the addition of the "axion", a particle whose equations of motion naturally enforce CP symmetry in QCD. [18]

6. **Inflation** – In addition to the standard expansion, Inflation is often invoked as a solution to several oddities in cosmological observations, e.g. the apparent thermal equilibrium between distant regions of the universe. Inflation requires the addition of a scalar field to the standard model called the inflaton.

7. **Unstable Universe** – Using standard model parameters, e.g. the Higgs and top quark masses, it is possible to compute the Higgs potential in the standard model. If the Higgs is at a local minimum of its potential, then it is possible that the universe is only meta-stable, implying the Higgs field could decay to its true vacuum state at any time, releasing a large amount of energy in the process and therefore destroying the universe. Recent analyses

show that the best fit to experimental observations imply a metastable universe, although stability is still possible within uncertainties. [20] The absurdity of this result can be considered motivation for a theory which gives less ridiculous predictions.

1.4 How Does Supersymmetry Help?

Supersymmetry (SUSY) is a framework which can be used to build infinitely many extensions of the SM. It can be traced back to the late 1960s, with the first physical models of SUSY in 4 dimensions discovered by Wess and Zumino in the 1974. [91, ch. 24] The key idea behind SUSY is the treatment of boson and fermion fields as pairs, called *superpartners*.[5] Superpartners are required to have exactly the same quantum numbers[6], except of course their spin. A Lagrangian is called supersymmetric if boson and fermion superpartners can be rotated into one another without changing its form.

It is useful now to introduce some terminology. Particles which appear in the standard model are called *baryonic matter*[7]. Superpartners to baryonic matter are called *sparticles*, a portmanteau of supersymmetric and particles. The superpartners of quarks and leptons are called squarks and sleptons respectively; here the "s" stands for scalar. For technical reasons based on the chiral nature of some standard model interaction, superpartners of fermions must have spin 0. Finally, the superpartners of the gauge bosons are given the suffix -ino, e.g. gluino and Zino, they are collectively called the gauginos.

[5] The most commonly cited review is [72]

[6] Mass and charge. As we will discuss in sec 1.4.3, the mass of the sparticles in physical models of SUSY are not actually identical to those in the standard model. This is possible in models where SUSY a broken symmetry.

[7] This is a bit of a misnomer given leptons and the gauge bosons carry no baryon number, but are still lumped into this classification

1.4.1 The Minimal Supersymmetric Standard Model Extension

Supersymmetry operators take bosons into fermions and visa versa. It is possible to construct SUSY theories where each standard model particle has multiple superpartners, typically the number of supersymmetries in a SUSY theory is denoted by $\mathcal{N}$. The minimal model of SUSY that incorporates the standard model has one supersymmetry and roughly one superpartner per standard model particle, $\mathcal{N} = 1$, and is called the Minimal Supersymmetric Standard Model (MSSM).

The particle content of the MSSM is not precisely double the particle content of the standard model. For technical reasons, the number of Higgs-like bosons in the MSSM is expected to be 5 in total. These include the known Higgs, being the lightest at 125 GeV, two charged Higgs called the H^+ and H^-, and two neutral Higgs called the A and H. In addition, gravitation can be incorporated into SUSY theories. The $\mathcal{N} = 1$ SUSY theory with gravitation is called the minimal supersymmetric theory of gravity (mSUGRA). These theories include massless gravitons and their superpartner, gravitinos.

1.4.2 R-parity

An attractive ad-hoc symmetry is added is to most supersymmetry models called R-parity. R-parity requires a multiplicatively conserved quantum number at each interaction vertex. Each particle is assigned a number

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

where B, L, and s, are the particle's baryon, lepton, and spin quantum numbers respectively. $P_R = +1$ for baryonic matter and $P_R = -1$ for sparticles. Without R-parity, SUSY models could allow for the violation of baryon and lepton numbers in laboratory decays, while no such decays

have ever been detected. [8] A particularly strong limit on R-parity violation comes from the proton lifetime, which is currently observed as $> 10^{32}$ years, 22 orders of magnitude larger than the age of the universe.

Another consequence of exact R-parity conservation is that there must be an even number of sparticles at each interaction vertex. With baryonic initial states, like at the LHC, SUSY particles would need to be pair produced. Additionally, heavy sparticle decay products must include at least one lighter sparticle. Taking this to its logical conclusion, the lightest supersymmetric particle (LSP) will be absolutely stable, i.e. will not be able to decay into standard model particles even if they are lighter, and must eventually appear as part of the decay chain when any SUSY particle is produced. This means that in models with R-parity conservation, the LSP could provide a ubiquitous dark matter candidate, provided it is electrically neutral. [72, sec. 6.2]

1.4.3 SUSY Breaking

If supersymmetry were an exact symmetry, then sparticles should have the same mass as their standard model superpartner. For instance, a selectron should exist with a mass of 0.511 MeV, such a particle should have been produced in laboratory experiments long ago. Because no such sparticles have been observed, there must be a breaking of SUSY in the vacuum state that we experience as our physical reality. Such a phenomenon is known as spontaneous symmetry breaking, a short review can be found in [87]. Theoretical investigations show that in most scenarios for SUSY breaking, the sparticles are amenable to having their own mass scale, meaning they do not vary in mass over more than an order of magnitude. This is an important feature because the non-observation of sparticles implies they are heavier than baryonic matter.

Another important feature of SUSY breaking is the possibility of mixing of electroweak gaugino and higgsino states, as well as inter-squark mixing and inter-slepton mixing. This means that the mass eigenstates seen in nature do not need to correspond to the SUSY eigenstates. As

[8] R-parity is sufficient, but not strictly necessary to accomplish this.

an example, if SUSY exists in nature, a collider like the LHC need not produce exactly selectrons or smuons, but rather particles which are some linear combination of those states.

The scale of SUSY breaking is related to its attractiveness as an extension to the standard model. The hierarchy problem can be boiled down to a "disturbing sensitivity of the Higgs potential" to any physics larger energy scale that shows up in extensions to the standard model. SUSY is a fix for this problem, given that the superpartner masses are not large compared to the Planck mass.[72, pg. 11] Using this criterion, SUSY breaking models estimate that the mass scale of SUSY particles should not be much larger than a few TeV. This is sometimes called "electroweak scale SUSY breaking," and it is main reason many scientists at CERN are looking at SUSY models to motivate searches for new particles.

Theoretical investigations of SUSY breaking show that it is not possible to accommodate SUSY breaking scenarios where all standard model particle masses are lighter than their superpartners without invoking a so-called "hidden sector", a group of particles which do not interact strongly with the standard model particles or their superpartners with large couplings. One class of these models is called gauge mediated SUSY breaking (GMSB), a schematic of which is shown in figure 1.5. In GMSB, SUSY is broken in the hidden sector, which couples to another auxiliary messenger sector. The messenger sector then can be coupled to the MSSM by normal $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge and gaugino interactions.[60]

Figure 1.5: A schematic diagram showing the structure of GMSB models. A hidden sector exists which is the source of SUSY breaking. The hidden sector interacts with a messenger sector, which in turn has interactions with the MSSM particles in a "flavor blind way". Taken from [60]

1.4.4 Why SUSY?

Below, we attempt to summarize the main motivations for distinguishing SUSY as an extension to the standard model:

- Supersymmetry offers a solution to the hierarchy problem described in section 1.3 by automatically including exactly canceling counter-terms for fermion loop corrections to the Higgs mass. In "natural" models of SUSY breaking, the expected mass scale of SUSY partners is right around the TeV range, and so these particles are likely accessible to the LHC if they exist.

- In models with R-parity conservation, supersymmetry provides a natural dark matter candidate. The lightest supersymmetric particle (LSP) should not be able to decay to regular matter because of the new conservation law, and so an electrically neutral LSP could be the dark matter we see in astrophysical observations.

- The minimal supersymmetric extension to the standard model seems to accommodate a high-energy GUT Lagrangian. As shown in figure 1.6, RG calculations show that the coupling constants related to the SM forces seem to unify near the GUT scale when analyzing the MSSM, while the coupling constants in the standard model do not intersect at a single point. This is seen as a sign that the MSSM could be the low energy theory associated with a GUT.

- Though not mentioned specifically in the previous discussion, some of the interest in SUSY comes from mathematical considerations. Mathematically, supersymmetry is the only known caveat to the Coleman-Mandula theorem, which states that the symmetry group of non-trivial QFTs must be written as

$$G_{\text{Poincaré}} \times G_{\text{internal}}.$$

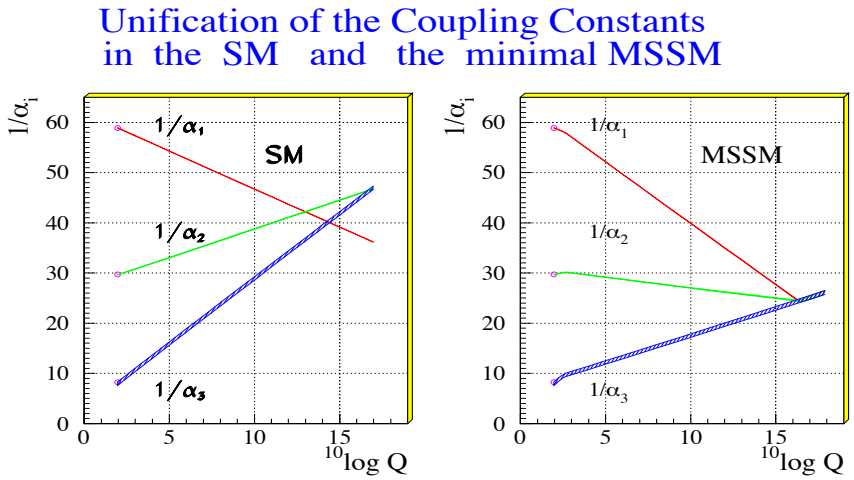


Figure 1.6: The strength of the strong[SU(3)] and electroweak [SU(2)xU(1)] coupling constants as a function of energy scale. On the left, the prediction derived from the standard model is shown, and on the right, the predictions from the MSSM. In a grand unified theory, all forces unify into a single force. Therefore, when extrapolating the strength of forces to higher energy scales using the RG equations, the strength of the couplings in a low energy effective theory underlying a high energy GUT are expected to meet at a single point, where all the forces unite have equal strength. In the standard model (i.e. assuming no new particles heavier than the top quark exist), there are 8 standard deviations separating a perfect fit, while the MSSM accommodates a unification point much more readily. This is interpreted as a sign that the MSSM might be the legitimate low energy approximation to the GUT of nature. Taken from [63]

This result is known as the Haag–Łopuszański–Sohnius theorem. Broadly speaking, supersymmetry is the only known caveat to a quite strong restriction on the structure of QFTs.

1.4.5 Simplified Models

The MSSM itself adds over 100 free parameters to the standard model, not including those that come from the SUSY breaking scheme or gravitation. The large parameter space dimensionality of realistic SUSY models make them extremely difficult to compare with observation. In order to address this issue, a series of simplified models has been developed which attempt to target specific phenomenology that commonly appears in both SUSY and other BSM models. In these models, only a small number particles are added with a single interaction. In typical cases,

only one or two of the sparticle masses and/or branching ratios are varied across the relevant parameter space. [43] [15] [67]

The simplified models used in this thesis are shown in sec 3.1.1.

1.5 Why focus on the Z with $E_{\mathrm{T}}^{\mathbf{miss}}$ final state?

As mentioned in the introduction, the ZMET final state is motivated partially by considerations about the detector, and partially by simplified supersymmetric models. Leptons at the LHC are rarely produced compared to hadronic jets,2.1.1 and are typically measured with high energy resolution (2.3.4, 2.3.5). This makes a leptonically decaying Z boson a great object to tag: the energy resolution is good and the standard model backgrounds for opposite-charge same-flavor leptons with dilepton mass near the Z pole mass are relatively small compared to the main production mode.

The main production mode of Z bosons in proton-proton collisions is the Drell-Yan process. The requirement of at least 2 jets and momentum imbalance in the final state suppresses the DY contamination of the background strongly. In fact, with those two cuts, the cross section is close to that of $t\bar{t}$ production. The background contribution to arbitrary kinematic distributions due to $t\bar{t}$ production can be predicted using the flavor symmetry of the decay. These background will then have very tight systematic uncertainties. In short, this final state provides a fairly clean signal experimentally.

On the theoretical side, this is a very generic search; an excellent feature for a physics analysis. In addition to the SUSY motivation, [73, 79, 74] encoded in the simplified models in sec 3.1.1, all standard model fermions have some coupling to the Z boson. In broad strokes, we can expect the production of any heavy resonance to have some chance to emit a Z boson, and likely have a long decay chain with many jets. If the Z decays leptonically, our final states should be sensitive to any such model.

1.5.1 Past results

An analysis in this final state has been published several times from the CMS collaboration, with the latest iteration in 2016.[81, 65, 66] The differences between the previous version and the analysis presented in this thesis are summarized below:

- The integrated luminosity analyzed increased by a factor of 15.
- Search regions were added to target SUSY production leading to final states contain an additional W or Z boson (VZ), and final states containing a Higgs boson (HZ). Interpretations in simplified models that produce these final states were also added. 3.4.3
- A correction is now applied to the photon sample used in the Z+Hadronic background prediction to subtract away events with real $E_{\mathrm{T}}^{\mathrm{miss}}$. 3.5.2
- A new method for the flavor symmetric background prediction was developed which uses same-sign events outside the Z mass window to predict the MET spectrum inside. 3.5.3

Chapter 2

The acquisition of data

2.1 The LHC

The Large Hadron Collider (LHC) is a particle accelerator and collider which runs underground near Geneva, Switzerland. Figure 2.1 shows the outline of the beam pipe under the greater Geneva area at the Swiss-French border. The beam pipe is 26.7 km long and is housed in a tunnel between 45 and 170 meters underground. This thesis is concerned with the proton-proton collisions at the LHC, which constitute the majority of the run-time.

The protons in these collisions are sourced from hydrogen gas, which has its electrons stripped at the CERN Meyrin facility and are then sent through several smaller accelerators before being injected into the LHC at 450 GeV. The LHC then accelerates the protons such that they achieve a kinetic energy of 6.5 TeV in the lab frame, the collisions between the beams then have a center of mass energy, $\sqrt{s}$, of 13 TeV [52], [21].

Protons are injected into the LHC in bunches, each containing approximately 10^{12} protons [7], and are accelerated in both clockwise and counterclockwise directions in two separate high vacuum beam pipes. About these beam pipes, superconducting dipole magnets guide the beams around the circular path, and quadrupole and octopole magnets ensure the beams stay focused

and collimated.

The proton bunches extend approximately 55 mm in length and are spaced such that the time between bunch crossings at any particular point in the beam pipe is approximately 25 ns at full energy and luminosity. At the four points, 1, 2, 5, and 8, distinguished in figure 2.1, the beams are crossed and the protons are given a chance to collide. In an average head-on bunch crossing, approximately 20 proton-proton pairs will collide and create deposits of energy in the detectors wrapped around the interaction points. For the analysis presented in this thesis, data was collected during the 2016 LHC run, corresponding to a usable integrated luminosity of 35.9 fb^{-1} collected by the CMS detector. Typical instantaneous luminosities during this time period were on the order of 10^{34} cm^{-2} s^{-1} $= 10^{-5}$ Hz fb^{-1}, as can be seen in figure 2.2 [36].

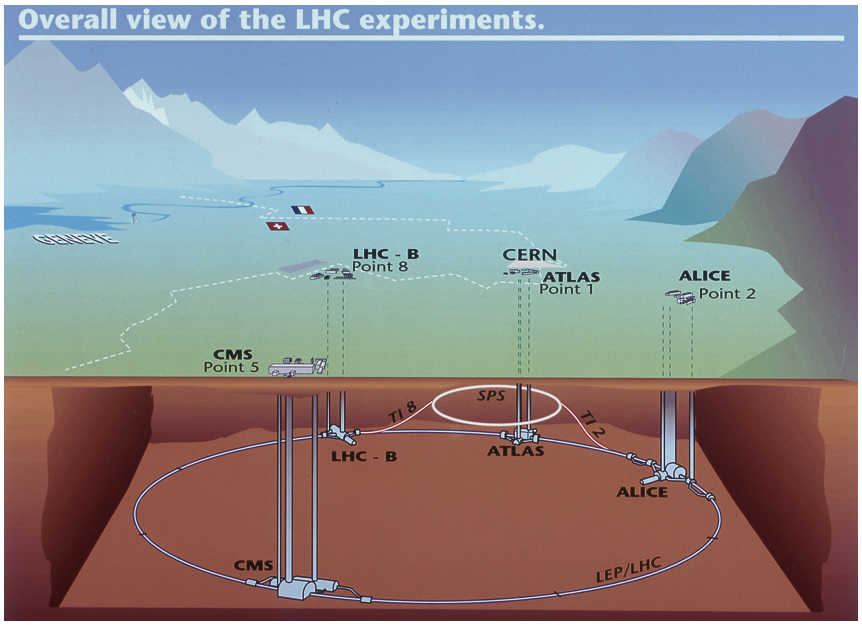


Figure 2.1: Outline of the LHC tunnel around the greater Geneva area. The accelerator is 26.7 km long and is housed in a tunnel between 45 and 170 meters underground. Taken from [88]

2.1.1 What gets made?

As mentioned in the previous section, the typical number of collisions leading to measurable energy deposits in the detector is 20 per bunch crossing. Figure 2.3 shows cross section

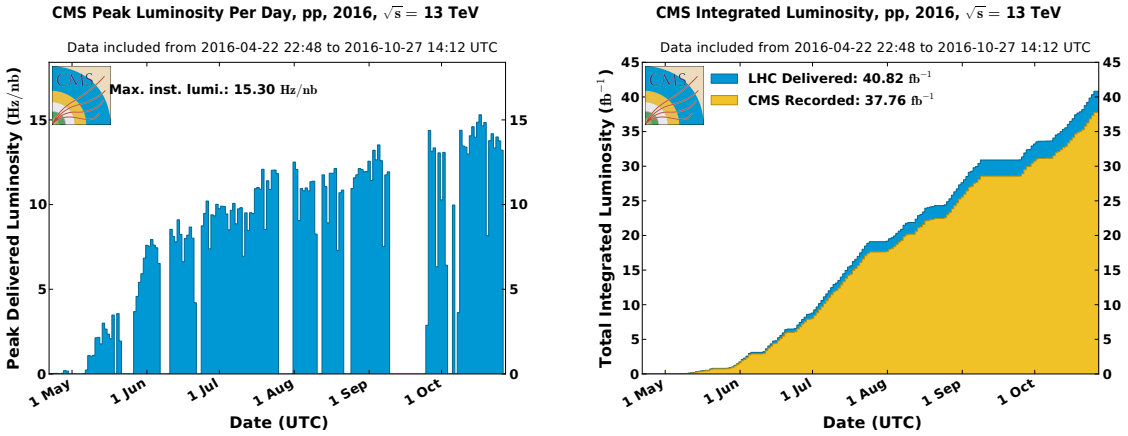


Figure 2.2: Peak instantaneous and integrated luminosity over time during the 2016 LHC proton-proton run delivered to the CMS detector. The difference between the 35.9 fb^{-1} used in this analysis and the shown 37.76 fb^{-1} comes from the omission of certain run periods where the detector was not operating optimally for the detection of leptons. Taken from [36].

for various proton-proton final states as a function of center of mass energy. Notice that the vast majority of the collisions result in low energy jet production or elastic scattering. The analysis presented in this thesis is concerned with the production of Z bosons, whose production cross section is denoted as σ_Z in the figure.

Figure 2.4 shows the measured value for the cross sections of a wider variety of processes by the CMS collaboration. At 13 TeV, the inclusive Z boson cross section is approximately 5×10^7 fb. Given the instantaneous luminosity of 10^{-5} Hz fb^{-1}, the production cross section corresponds to a rate of approximately 500 Z bosons produced per second at peak luminosity. Given the total integrated luminosity of 35.9 fb^{-1}, the entire CMS dataset during this time period contained approximately 2 billion Z bosons.

Notice that the vast majority of collisions create only colored particles in the prompt process. Therefore, the most likely effect of the extra 20 collisions in a bunch crossing is to produce soft hadronic jets. For instance, the chance to produce another Z boson in an event that already has a Z boson is roughly 20 in a million, or 1 in 50,000. However, it is still important that particles from other collisions do not contaminate an event. As will be discussed more thoroughly in section 2.3.2, charged particle tracks can be traced back to the beamline and clustered into

points of origin called a verticies. The vertex that is assigned the largest squared sum of transverse energy in an event is called the primary vertex. The chance that two high energy verticies will exist in a single event is low because the number of collisions per crossing is small compared the ratio of the cross section of interesting physics to the total proton-proton cross section.

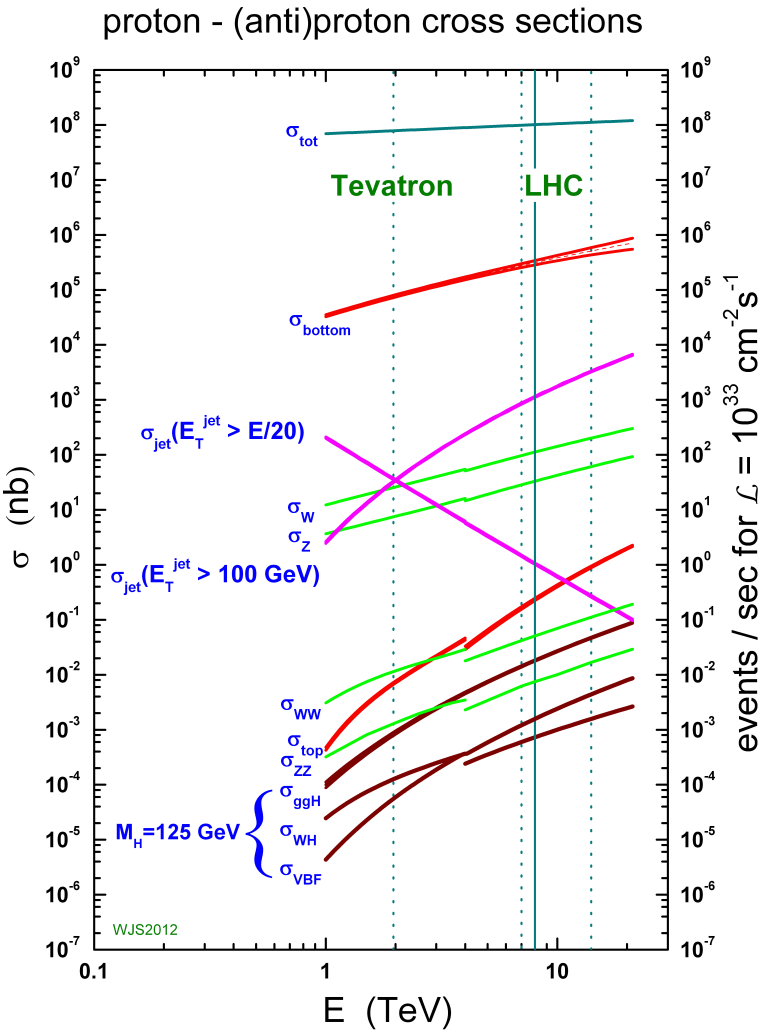


Figure 2.3: Production Cross Sections for proton-proton collisions. Cross sections at center of mass energy less than 4 TeV are taken from proton-antiproton collision data at the Tevitron, which leads to some discontinuity for some types of electroweak boson production.

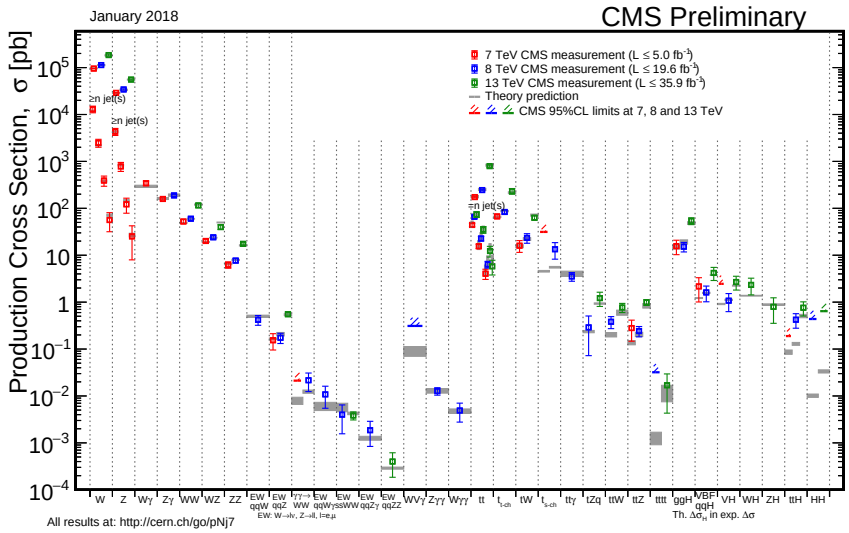


Figure 2.4: Cross sections measured by the CMS collaboration as of January 2018. Note that the cross section for Z+2 jets is close to the cross section for $t\bar{t}$ production. Additionally, note that for W and Z bosons, the cross section reduction in adding an additional jet is a factor between 5 and 10, consistent with the value of α_s at these energies. Taken from [35]

2.2 The CMS Detector

The Compact Muon Solenoid (CMS) is a general purpose detector at the LHC. The detector is shown in figure 2.5. It is the second largest detector at the LHC, weighing just under 15 kilotons. The envelope of the detector is a cylinder of radius 7.3 meters and length of 21.6 meters. The detector subsystems are embedded like an onion, sorted by radial distance from the beam pipe they are:

1. Silicon Pixel Tracker
2. Silicon Strip Tracker
3. Electromegnetic Calorimeter
4. Hadronic Calorimeter
5. Superconducting Solenoid
6. Muon System

The subsystems are broken into at least two regions. The *barrel* region is the central part of the detector, and it is built of mostly of detection modules that are oriented parallel to the beam pipe, since particles traveling through this part of the detector have more transverse than longitudinal momentum. The *end cap* contains modules oriented perpendicular to the beam pipe, since particles traveling through this part of the detector have more longitudinal momentum.

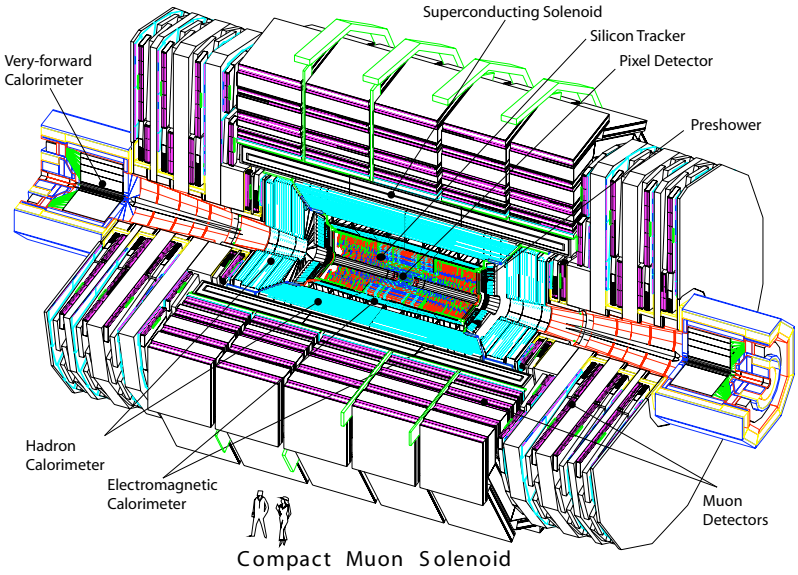


Figure 2.5: A cutaway drawing of the Compact Muon Solenoid. Taken from the CMS TDR [19].

As was shown in sec 2.1.1, most collisions at the LHC produce sprays of hadrons called jets. This reflects the fact that most interactions are between gluons. However, more rare electroweak processes, such as the production of the higgs boson, can lead to the creation of leptons. Therefore, measuring the energy spectra of leptons is of central importance for CMS analyses searching for new physics coupled to the electroweak sector; this thesis contains one such analysis.

In order to measure the momentum of charged particles, a large magnetic field is applied by a superconducting solenoid that is placed between the hadronic calorimeter and the muon system. The solenoid creates a roughly constant 3.8 Tesla magnetic field parallel to the beam pipe in the region of the detector filled by the tracking system and calorimeters. This magnetic field

will bend charged particles in accordance with the Lorentz force law and allow for a measurement of the particles momentum.

CMS was designed with several physics goals in mind, from finding the Higgs boson, to searches for dark matter and supersymmetry. The technical design report (TDR) [19] summarizes the physics and design goals for the detector. A short list follows:

1. The search for the Higgs boson, specifically in photon and muon final states[1]. This created a need for excellent muon and photon energy resolution and isolation. These requirements justify the advanced muon system and ECAL. Additionally searching for the Higgs in the $b\bar{b}$ and $\tau\bar{\tau}$ channels created the requirement for good offline b-tagging and τ-tagging capabilities, largely regulated by tracker resolution.

2. The search for supersymmetric (SUSY) particles. The main motivation for these searches is often in the context of R-parity conserving SUSY due to the natural dark matter candidate they provide as described in section 1.4.2. Dark particles are expected to only leave momentum imbalance in the detector, so there is need for good energy resolution for all physics objects.

3. The search for new massive vector bosons, typically dubbed a Z' search. Here dilepton (electron and muon) energy resolution are again of paramount importance.

4. The search for extra dimensions. The phenomenology of these models is very broad, but signatures can include all massive standard model particles and gravitons which leave a $E_{\mathrm{T}}^{\mathrm{miss}}$ signature.

5. Measurements furthering the precision Standard Model parameters. The production of top quarks is enhanced at the LHC compared to any previous colliders due to their large mass, which is non-trivial even at the TeV scale. Top quarks almost always decay to b-quarks which means the LHC can be used to study b physics.

[1] the diphoton channel was where it ultimately was found

6. In addition to proton-proton collisions, the LHC also collides lead ions which probe the thermodynamic properties of quantum chromodynamics (QCD), the theory of the strong nuclear force. These collisions typically produce hadronic jets and their p_{T} spectrum is of interest due to observations at RHIC [49].

2.2.1 Coordinate System

Throughout this document, a standard coordinate system is used, this system is a cylindrical coordinate system with the z axis oriented along the beam pipe. $z = 0$ is situated at the mid-point of the detector, 10.8 m from either edge. The $\theta = 0$ direction points toward the Jura mountains, with $\theta = 90^\circ$ pointing straight upwards, away from the center of the earth.

Rather than θ, we use the pseudorapidity, $\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right)$. $\eta = 0$ corresponds to $\theta = 90^\circ$ and η grows to infinity as θ goes to 0. The benefit of using η is that differences in η, $\Delta\eta$, between two particles are approximately invariant under Lorentz boosts along the beam axis, whereas differences in θ are not. The extent to which $\Delta\eta$ is equal in different reference frames is regulated by the mass of the particles, with equality achieved in the case where the mass to momentum ratio of the particles goes to 0, i.e. the high energy limit.[2] The full detector's fiducial area corresponds to about $|\eta| < 2.4$, which is about 10° off of the beampipe[3]. The positive x axis is defined as pointing to the center of the circle outlined by the LHC, and the positive y axis points towards the sky. The ϕ direction is the angle from the positive x axis to the positive y axis. Figure 2.6 shows this information visually.

2.2.2 The Inner Tracker

The inner tracker is the closest part of the CMS detector to LHC beamline and interaction point where protons collide. [25] It surrounds the interaction point with a length of 5.8m and

[2]Pages 6-8 in reference [48] shows that differences in rapidity are Lorentz invariant and that psuedorapidity and rapidity are equal in the high energy limit

[3]Note: this is limited by the tracking system, the calorimetry system extends further to about $|\eta| < 5.2$

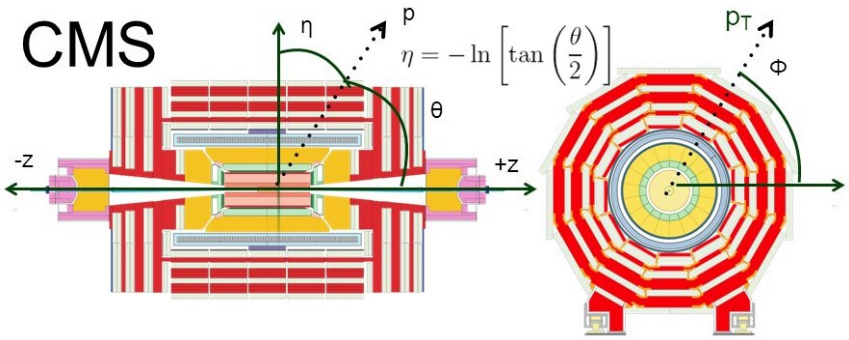


Figure 2.6: Cross sectional and transverse views of the CMS detector with η, ϕ, and θ coordinates shown. Taken from [1].

radius of 1.25m. The purpose of the system is to track charged particles to their vertex and measure the momentum of charged particles via the saggita in the particle arc due to the magnetic field, as outlined in B. The entire active detection area of the inner-tracker system is made of silicon, but can be broken into 2 main subsystems:

1. **pixel detectors** – The innermost part of the tracker system is a pixel-based detector composed of 1,440 thin modules, of cross section $100 \times 150 \mu\text{m}^2$, capable of measuring hits with fine granularity in 3 dimensions. This subsystem's main purpose is to aid in vertex reconstruction, which leverages its superior 3 dimensional spacial reconstruction capabilities.

2. **strip detectors** – The outer part of the inner tracker is composed of 15,148 strip modules. The main purpose of this subsystem is to track charged particles from the vertex into the ECAL and also to give one measure of their momentum.

Silicon detectors work on the principle of semi-conduction. When a charged particle passes through the material, electrons are kicked into the conduction band and drift, due to a bias voltage applied across the sample, towards electronics attached to the material that record the current. Because silicon has a relatively small band gap, the entire tracking system needs to be kept at low temperature, approximately -10° C, in order to keep the electrons stationary in the valence band in the presence of the bias voltage.

The tracker geometry is shown in figure 2.7. The pixel detectors are the closest to the beampipe and consist of 3 layers in the barrel region and 2 annuli in the endcap region. The three barrel layers are positioned in concentric cylinders about the beampipe at radii of 4.4 cm, 7.3 cm, and 10.2 cm respectively. The endcap annuli are placed at $|z| = 34.5$cm and 46.5 cm respectively and have an inner radius of 6 cm and an outer radius of 15 cm. The strip tracker consists of an inner barrel region (TIB), an outer barrel region (TOB), and an inner disk region (TID) in the barrel, and two endcap regions (TEC +/-).

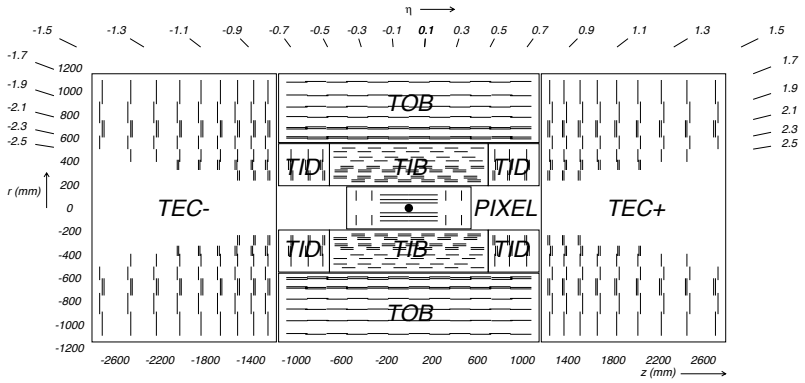


Figure 2.7: Schematic view of the CMS inner tracking system. Notice the detectors only cover $|\eta|$ ranges less than 2.5 and that particles traveling near $|\eta| = 1.5$ pass through the most material. Taken from [25].

Because one of the goals of the tracking system is to obtain a measure of momentum by tracking the natural motion of particles through space in a magnetic field, it is important that the interactions between charged particles and the tracking system do not change the motion, and likewise the energy, of the particles dramatically. For particles other than the electron and photon, the silicon tracker has mostly negligible effects on the energy as the amount of bremsstrahlung is inversely proportional to the mass of the particle to approximately the 6th power;[4] their main energy loss mechanism is through ionization.[76, sec. 33.2] However, for the electron and the photon, interaction with the tracker can cause bremsstrahlung radiation and pair production respectively.

[4]This can be seen in the classical theory of bremsstrahlung radiation as described in [61, pg. 464, eq. 11.75] by replacing the Lorentz factor γ with $\frac{E}{mc^2}$. In the case of an acceleration in an orthogonal direction, $\gamma^6 \rightarrow \gamma^4$.

To understand the magnitude of these effects, it is typical to look at the number of radiation lengths[5] of material in the tracker. The "material budget" of the tracker is shown in figure 2.8. Due to the large amount of non-sensitive material in the range $|\eta| \in [1.4, 1.6]$, leptons for this analysis are not considered in that range, as is explained in section 3.4.1. As can be seen in the figure, many η values correspond to high probability of radiation and pair production for electrons and photons respectively. We will explain in section 2.3.3 how the momentum of these types of particles is reconstructed given these issues.

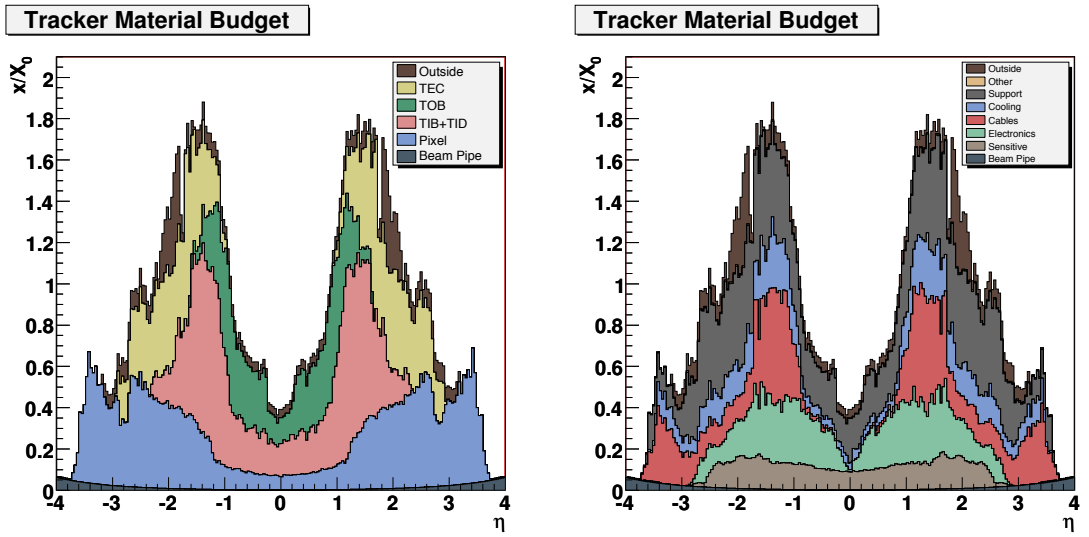


Figure 2.8: The material budget of the tracking system in radiation lengths. On the left, the material is broken down by tracker subsystem, on the right it is broken down by the type of component. As can be seen on the right plot, a large amount in non-sensitive material, like cables and support structure, is found near $|\eta| = 1.5$. Due to the increased material and the transition in the ECAL from barrel to endcap, electrons (and muons to preserve symmetry) in this analysis are not used if they are found in the window $|\eta| \in [1.4, 1.6]$. Taken from [25].

2.2.3 The Electromagnetic Calorimeter

The CMS Electromagentic Calorimeter (ECAL) is a nearly hermetic and homogeneous cylinder made of lead tungstate ($PbWO_4$) crystals with attached light measuring devices. The

[5] taken to be the distance at which a high energy electron is expected to lose $\frac{1}{e}$ of its energy [76, sec 34.4.2], or $\frac{7}{9}$ the mean free path for a high energy photon.

crystals are “truncated pyramids”, roughly rectangles of approximately 23 centimeters in length that taper slightly, from 26x26 mm^2 to 22x22 mm^2 in the barrel[26, pg. 4], to accommodate the curved shape of the ECAL in the ϕ direction and the angle at which the crystals are oriented to face the interaction point.[6] Schematic views of the ECAL can be seen in figures 2.9 and 2.10. As can be seen in the figures, the calorimeter is broken into two physical sections, the barrel region (EB) and the endcap region (EE). The preshower disk in front of the endcap region is immaterial for this search, but is there to help distinguish between neutral pions converting to a di-photon pair with small angle separation from a single high energy photon.

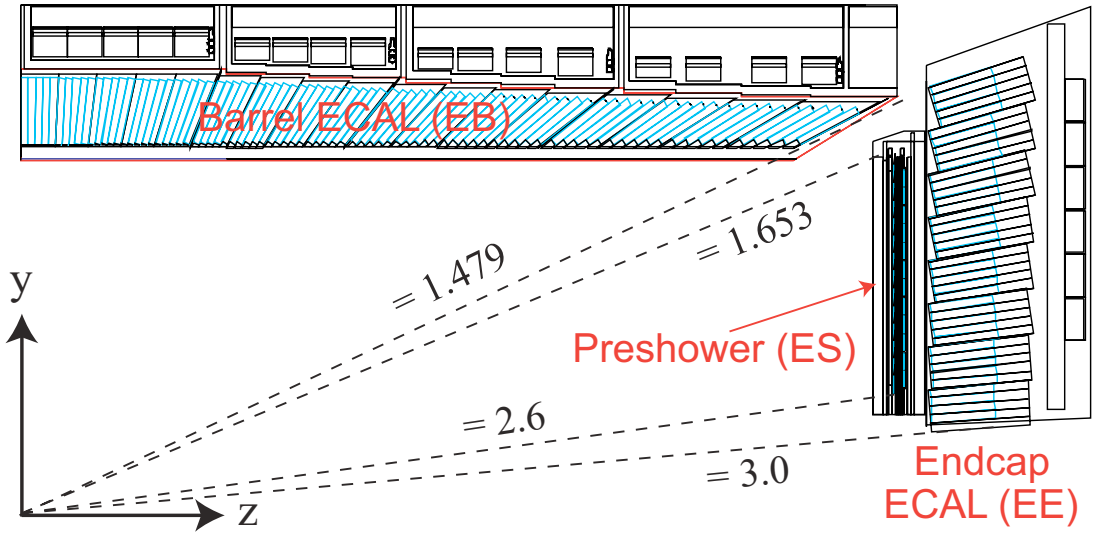


Figure 2.9: A cross sectional view of the electromagnetic calorimeter on the CMS detector. Notice the geometry of the crystals the transition region near $|\eta| = 1.5$. Taken from [19].

The purpose of the ECAL is essentially to give the most important measure of energy for electrons and photons. As explained in the previous section, the dynamics of electrons in material are quite distinct from heavier charged particles in material, the next lightest being a muon. Typical muon deposits in the ECAL are roughly 300 MeV[28], whereas electrons under 500 GeV tend to have almost all of their energy absorbed by the ECAL[25, sec. 4.10]. This is to say that charged particles heavier than an electron tend to pass right through the ECAL with little

[6] the crystals have a 3° angle with respect to the line that connects their incident face to interaction point in both the η and ϕ directions to allow for better coverage of the fiducial volume.

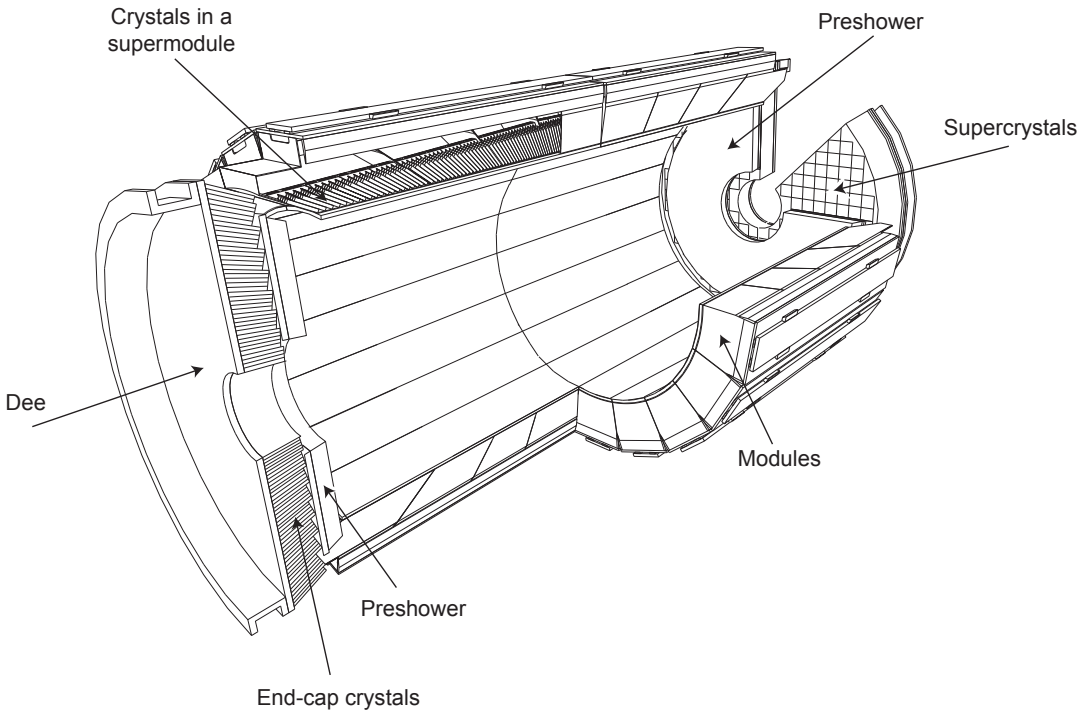


Figure 2.10: A cutaway of the electromagnetic calorimeter on the CMS detector. Taken from [25].

disturbance.

The ECAL operates on the principle of scintillation[69, ch. 7], and makes use of the fast scintillation time (80% of light emitted within 25 ns), high stopping power, and consequently small Molière radius of lead tungstate. The lead tungstate crystals have a length which corresponds to approximately 25 radiation lengths, this ensures that almost all of the energy carried by a high energy photon or electron will be radiated. The small Molière radius allows for good spacial resolution of energy deposits. [25, pg. 90] Two things happen when an electron or photon is incident on one of the crystals:

1. A cascade of electrons, positrons, and photons is created. This is due to the effects discussed in the previous section, photons will pair produce electron-positron pairs. Electrons and positrons undergo bremsstrahlung radiation in material creating high energy photons. The result is that these processes feedback on each other and the multiplicity of particles

explodes in the material creating an “electromagnetic shower”. A hypothetical shower imposed on an ECAL crystal is shown in figure 2.11.

2. Scintillation light is emitted by the lead tungstate due to interactions with charged particles and the light captured by the photo detectors attached to the crystal.

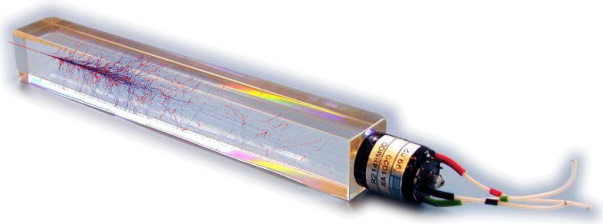


Figure 2.11: A lead tungstate crystal from the CMS ECAL endcap region with a vacuum phototriode attached. A hypothetical shower from an electron is superimposed on the image. Taken from [2].

From the amount of scintillation light, the energy of the incident particle can be reconstructed with high resolution. Figure 2.12 shows the energy measured using a 5x5 grid of ECAL crystals for 120 GeV electrons. From the figure, we can see that the ECAL energy resolution is excellent for electrons in this range, the standard deviation of the energy measurement being 1-2% of the energy.

2.2.4 The Hadronic Calorimeter

The CMS Hadronic Calorimeter (HCAL) is a sampling calorimeter made of interleaved layers of dense absorber metals with plastic scintillator material. The purpose of this subsystem is to measure the energy of the hadrons and massive charged particles that make up jets. The measurement of these objects are of particular importance for the reconstruction of $E_{\mathrm{T}}^{\mathrm{miss}}$, which is a central event-level object in this analysis.

Figure 2.13 shows the general design of an HCAL wedge in the barrel region. Layers of brass are interspersed with scintillator strips that read out the energy of secondary particles

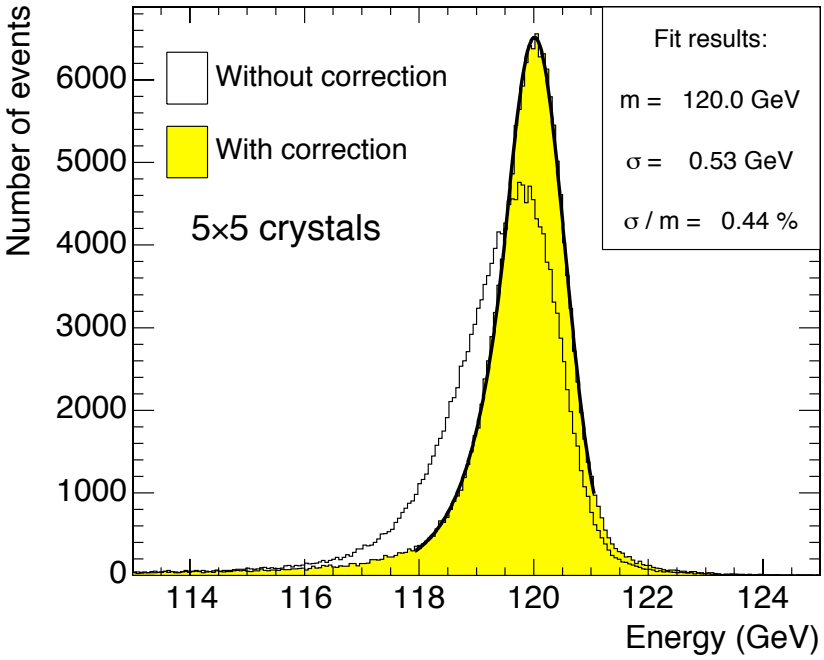


Figure 2.12: The energy measured from 120 GeV electrons taken from a 5x5 grid of ECAL crystals. Taken from [25].

created in the hadronic interaction. In terms of nuclear interaction lengths, the HCAL goes from about 5 nuclear interaction lengths at $\eta = 0$ up to about 10 at $\eta \geq 1.3$. Due to the shallow depth in the barrel, an outer layer of scintillator is added outside the magnet, essentially using the magnet as an absorber layer. This outer layer is called the HO and runs in the range $\eta \in [0, 1.3]$. Figure 2.14 shows a cross-section of the HCAL including the HB, HE, and HO. With the HO in place, the minimum number of interaction lengths across the HCAL is increased to almost 12 except in the transition region between the endcap and barrel.[25, pg. 138]

Sampling calorimeters again work on the principle of scintillation, however, HCALs are meant to measure energy from neutral particles as well as charged hadrons. The strategy used in sampling calorimeters is to place a dense material with many nuclei in front of the incident particles. Hadronic interactions are much more complicated than the photon and electron interactions in the ECAL because the initial state particles can vary.

The products of inelastic collisions in the HCAL include charged hadrons like the pion

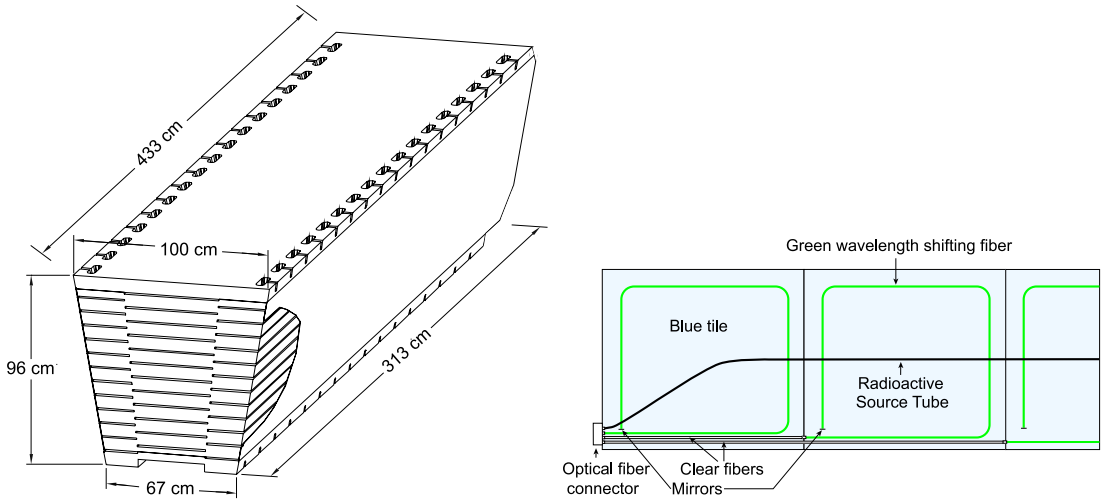


Figure 2.13: Shown on the left, a wedge of the HCAL in the barrel. The interior of the wedge is brass, and the shown slots have the scintillator modules (shown to the right) inserted so that secondary particles made in interactions with the brass will pass through them and have their energy measured. To the right, a scintillator module is shown, when a charged particle passes through the scintillator, light is generated and collected by the wavelength shifting fiber shown in green. When scintillation light is collected by the fiber, it releases green light that travels down to the clear fiber via total internal reflection and sent to be collected by photo sensors. Taken from [10].

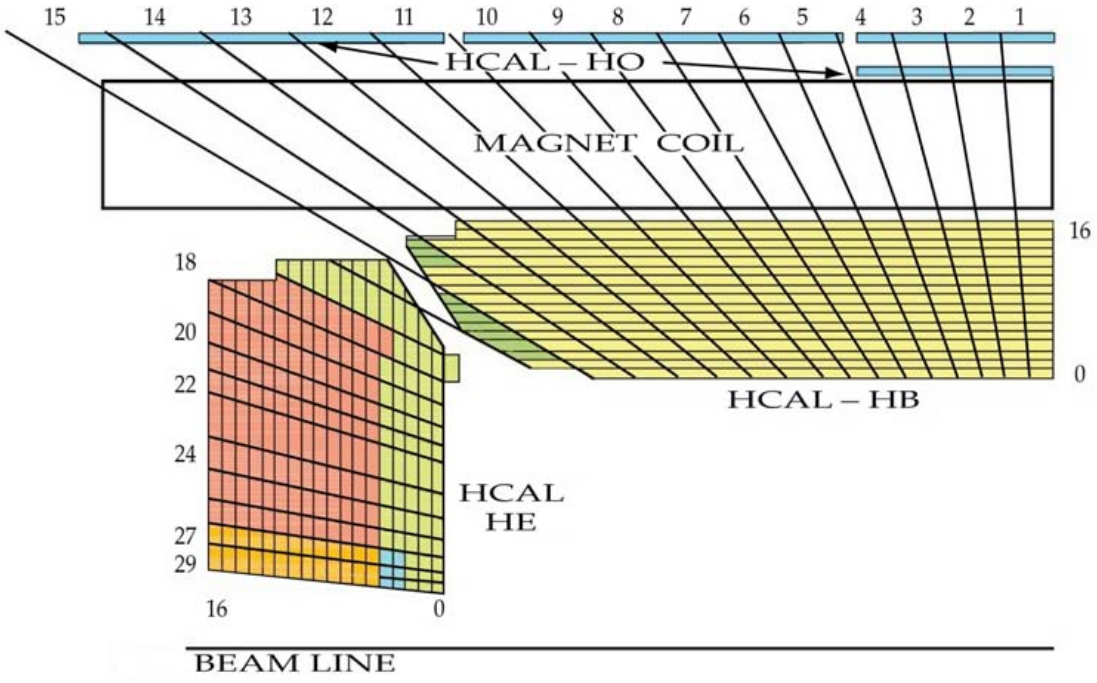


Figure 2.14: The a cross-sectional view of the CMS hadronic calorimeter. Taken from [25].

or protons, which deposit energy via ionization and excitation, and neutral particles like the η and π^0, which decay to photons and create electromagnetic cascades [76, sec 34.9.2] [59]. The cascade of particles then leads into the scintillating material where scintillation light is generated and collected. Again, the amount of light collected is proportional to the energy of the charged particle.

In the CMS HCAL, the absorber layers are made of brass, and the scintillator material is a proprietary plastic material, Kuraray SCSN81[7]. The HCAL is broken into three physical sections: the barrel (HB) extending to $|\eta| \leq 1.4$, the slightly overlapping endcap (HE) extending between $|\eta| \in (1.3, 3)$, and the slightly overlapping forward detector (HF) covering the range $|\eta| \in (2.9, 5)$. In this analysis, we do not use jets measured in the forward detector, that is we do not consider jets with $|\eta| > 3$.

2.2.5 The Muon System

The outer-most detection subsystem on CMS are the muon chamber. Muons literally put the "M" in CMS, they are an extremely important particle for LHC physics for several reasons. [3, sec 1.2]

1. Their production is associated with the electroweak sector, which includes the Higgs boson.

2. They have a long enough lifetime that they can make it through the entire detector, making them easily identifiable and allowing for good momentum resolution due to their long track.

3. Their radiative losses in the tracking system and ECAL are much smaller than electrons due to their heavier mass. They also don't deposit much energy in the HCAL because they don't interact via the strong force.

[7] excluding the very first layer where the absorber metal is stainless steel and the plastic scintillator is Bicron BC408. The very last layer also has stainless steel as the absorber metal.

The muon system on CMS is made up of three different types of detectors, all based on gas ionization.

1. **Drift Tubes (DTs)** – Layered thin tubes with conducting cathode walls and a anode wire which runs centrally through the length of the tube. The tube is filled with Argon and CO_2 gas. A drift tube is shown to the left in figure 2.15, the length of the tube is not shown. Due to the extent in one direction, drift tubes can not give spacial resolution in all three dimensions. A typical module of these tubes has 3 layers, oriented such that the middle layer runs perpendicular to the outer layers in order to reconstruct all the position in all three dimensions.

2. **Cathode Strip Chambers (CSCs)** – Six layers of anode wire planes interleaved with seven layers of cathode strips, arrayed in a cross-hatch pattern filled with gas. This configuration allows for all three dimensions of the muon's position to be reconstructed from a single module by cross referencing the voltage spikes in the strips against the voltage spikes in the wires. A schematic view of this type of detector is shown to the right in figure 2.15.

3. **Resistive Plate Chambers (RPCs)** – Parallel plate capacitors made of insulating material and filled with gas. When a muon passes through the RPC, the gas is ionized and an electron shower is generated which passes through the capacitor plates and is collected by a set of strips outside the capacitor. These detectors are interspersed throughout the entire system and mostly used to help with triggering as the response time and refractory period of these detectors are very fast compared to the 25 ns bunch crossing time.

Figure 2.16 shows the general layout of the muon system. In the barrel region, approximately $|\eta| < 1.2$, the muon system is made up of drift tubes interleaved with RPCs. In the endcap region, approximately $|\eta| \in (1.2, 2.4)$, the drift tubes are replaces by CSC modules due to the higher magnetic field and higher radiation dose in this region.

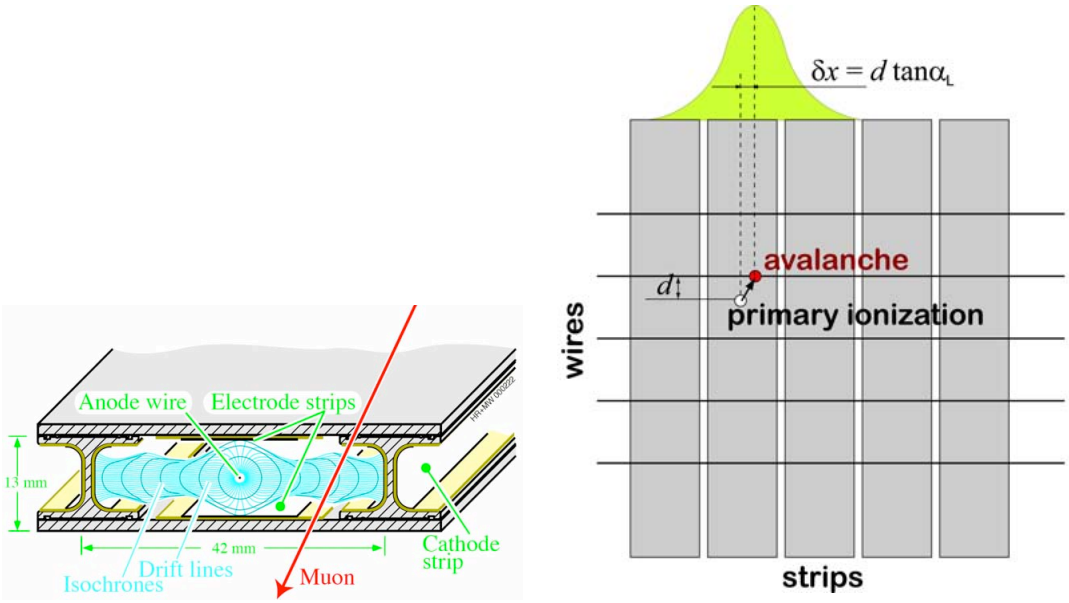


Figure 2.15: **Left:** A cross sectional head on view of a CMS drift tube, the tube is much longer in the omitted direction than is shown. A hypothetical path for a muon is drawn, the particle would ionize the gas in the chamber and create a spike in the potential difference between the anode and cathode as free electrons and charged ions are collected. **Right:** A top down view of a CSC module. A hypothetical muon hit is shown at the spot labeled "primary ionization". A CSC module has the ability to reconstruct all three dimensions of the muon's position by cross referencing which strips collected ions and which wires collected free electrons.Taken from [25].

2.2.6 Event Triggering

A raw CMS event takes roughly 1 MB of disk space, given the rate at which bunches are crossed, storing the output from every bunch crossing would require reading out, analyzing, and storing upwards of 40 TB of data per second.[27] The final event readout rate of the CMS detector is designed to be closer to 400 Hz, which would correspond to roughly 400 MB of data per second, with the main bottleneck being offline data storage capacity and processing speed. In 2016, the actual readout rate was approximately 1 GB/s.

In order to reduce the event rate from 40 MHz down to 400 Hz, the CMS detector employs a two-stage trigger system. The first stage, called the level 1 or L1 trigger, takes the rate down to approximately 100 kHz, a process done in hardware using application specific integrated circuits

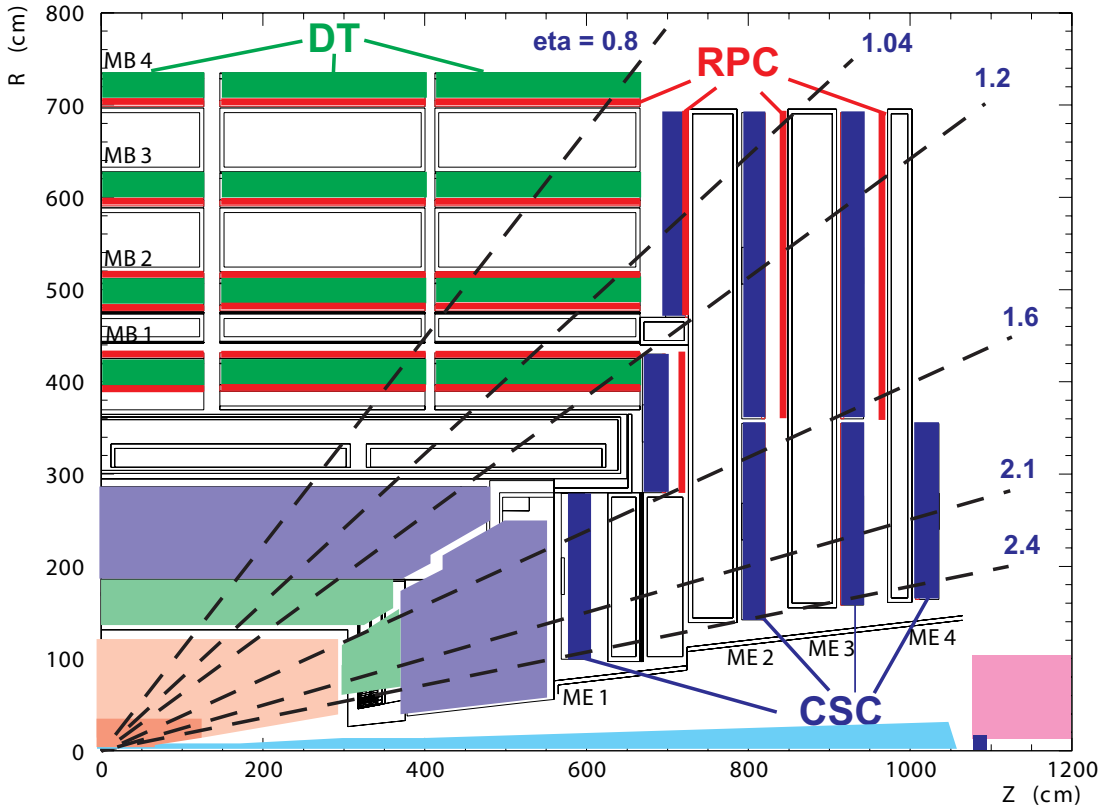


Figure 2.16: A cross-sectional view of the CMS muon system. In the barrel, with $|\eta| < 1.2$, drift tubes (DTs) and resistive plate chambers (RPCs) are used. In the end cap, $|\eta| \in (1.2, 2.4)$, RPCs and cathode strip chambers (CSCs) are used. Taken from [19].

(ASICs) and field programmable gate arrays (FPGAs) for maximum speed. Events selected by the L1 system are then passed to the high level trigger system (HLT) which reconstructs physics objects in software and makes a better determination as to whether the event should be written to disk. The HLT system brings the event rate down to 100s of Hz.

From the point of view of the L1 system, electrons and photons are identical. Decisions at level 1 for electrons and photons are made by the e/γ L1 trigger, which mainly looks for large isolated ECAL energy deposits without a corresponding HCAL deposit. For the muon triggers, information from CSCs, DTs, and RPCs are used. Built in electronics on the detector perform track finding algorithms on hits in the DTs and CSCs independently. "Track primitives" are created by special track finding hardware newly added for 2016. The track primitives are matched

with timing and hit data from the RPCs to make decisions about muon quality and energy at L1. The processed information is sent to the global trigger system to make the final decision as to whether the event will be sent to the HLT farm.

The HLT system in 2016 consisted of a processing farm with approximately 22,000 CPUs. The turn around rate for event classification was approximately 160 ms per event. The machines in the HLT perform object identification and energy reconstruction in a way that matches the offline software suite as much as possible, though there are some optimizations made that introduce small differences to accommodate event processing at the 100 kHz rate. The HLT system is organized around the concept of "trigger paths", which are algorithms that run object reconstruction and makes selections on these objects. A trigger path will evaluate to either true or false on an event by event basis.

The full list of CMS HLT trigger paths used in this analysis is presented in table 2.1.

The high level triggers are generally based on isolation and transverse momentum criteria. For instance, the trigger HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v* roughly requires that the event have at least one muon above 17 GeV of p_{T} and another above 8 GeV. The tags TrkIsoVVL refer to isolation requirements on the muon and DZ refers to a requirement that when their tracks terminate close to one another in the z (beamline) direction traced back to the beamline. The v* at the end of these triggers denotes that we use the latest version of the software that implements the trigger.

Trigger efficiencies

Due to the low precision in the energy measurements at level 1, and slight differences in reconstruction, ID, and isolation requirements between the HLT and offline level, objects near the p_{T} thresholds of a high level trigger do not always cause the event to be stored, we say that the triggers are not 100% efficient. In other words, sometimes an event with an electron that is reconstructed to have 17 GeV of p_{T} offline will not pass the Ele17 trigger because it was

either not selected at level 1, or it was reconstructed to have less p_T than 17 GeV by the HLT reconstruction software.

The rate at which this occurs can be treated statistically. It is mainly a function of the offline reconstructed p_T because all of the triggers used in this analysis are based on the p_T of some object. As an example, figure 2.17 shows the L1 trigger efficiency for the single muon L1 25 GeV trigger in the 2017 dataset[8] at the p_T threshold of 25 GeV (the target threshold for the single muon trigger). As can be seen in the figure, muons with $p_T \leq 8$ GeV (as reconstructed offline) are almost never selected by the 25 GeV L1 trigger. The trigger never selects 100% of the muons, even at very high energies. However, after about 30 GeV, the percentage of muons selected plateaus. We say that the trigger is *fully efficient* at 30 GeV to denote this feature, which is common to all trigger turn-on curves.

Trigger efficiencies enter this analysis in two places. First, MC used in this analysis has the trigger efficiencies built in, that is MC events are tagged as passing for failing triggers at rates equal to the trigger efficiencies. Second, the difference in the efficiencies for the electron and muon trigger paths in this analysis are part of what is corrected by the $R_{SF/DF}$ factor, which will be described in section 3.5.3. In short, this analysis uses the symmetry between production rates of electrons and muons in certain physics processes to make certain background predictions. That symmetry is broken by the fact that, along with different offline ID requirements, the trigger efficiencies for electrons and muons at the same p_T will, in all likelihood, not be identical.

[8] this analysis uses the 2016 dataset, the figure is only for illustration

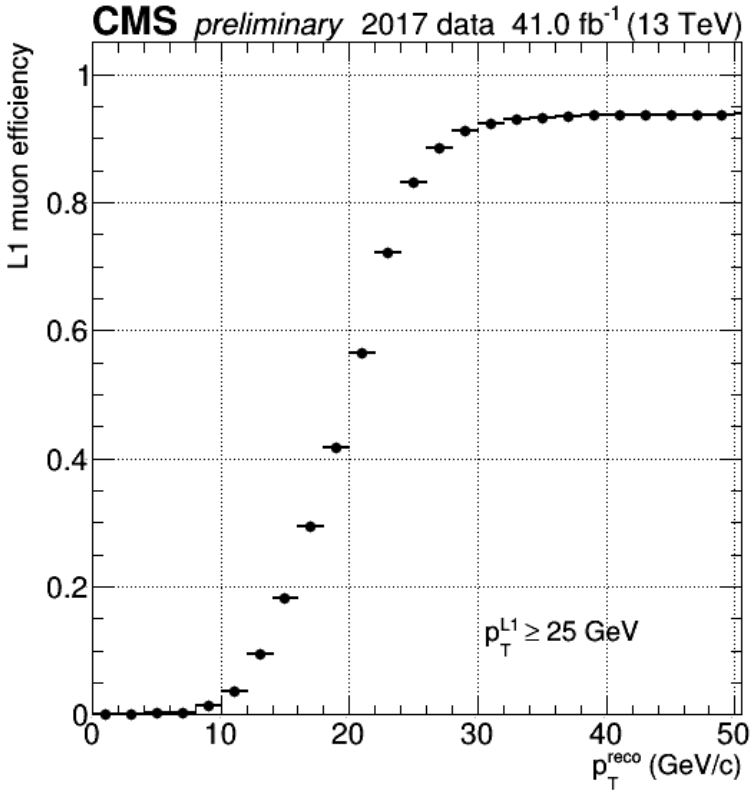


Figure 2.17: The trigger efficiency for the L1 muon 25 GeV trigger. The system attempts to pass all muons with 25 GeV of p_{T} to the HLT system, though the fast calculation with on-board hardware at L1 does not capture all muons at this threshold. The plot is made by selecting events with two muons reconstructed offline where the leading p_{T} muon has p_{T} above 27 GeV and has passed the L1 25 GeV trigger. The second muon is then checked to see whether it also passed the L1 criteria. This method is called tag-and-probe. As can be seen, the trigger is never at 100% efficiency, but a plateau exists around 30 GeV where we call the trigger *fully efficient*. Taken from [34].

Table 2.1: List of all triggers used in this analysis.

trigger name
di-muon triggers
–HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_v*
–HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v*
–HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_v*
–HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v*
–HLT_Mu27_TkMu8_v*
–HLT_Mu30_TkMu11_v*
di-electron triggers
–HLT_Ele17_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*
–HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*
–HLT_DoubleEle33_CaloIdL_GsfTrkIdVL_v*
–HLT_DoubleEle33_CaloIdL_GsfTrkIdVL_MW_v*
eμ triggers
–HLT_Mu17_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v*
–HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL_v*
–HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v*
–HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL_DZ_v*
–HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*
–HLT_Mu8_TrkIsoVVL_Ele17_CaloIdL_TrackIdL_IsoVL_v*
–HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_v*
–HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v*
–HLT_Mu30_Ele30_CaloIdL_GsfTrkIdVL_v*
–HLT_Mu33_Ele33_CaloIdL_GsfTrkIdVL_v*
single-γ triggers
–HLT_Photon22_R9Id90_HE10_IsoM_v*
–HLT_Photon30_R9Id90_HE10_IsoM_v*
–HLT_Photon36_R9Id90_HE10_IsoM_v*
–HLT_Photon50_R9Id90_HE10_IsoM_v*
–HLT_Photon75_R9Id90_HE10_IsoM_v*
–HLT_Photon90_R9Id90_HE10_IsoM_v*
–HLT_Photon120_R9Id90_HE10_IsoM_v*
–HLT_Photon165_R9Id90_HE10_IsoM_v*
–HLT_Photon165_HE10_v*
single-μ triggers
–HLT_IsoMu24_v*
–HLT_IsoTkMu24_v*

2.3 Physics Objects

After energy deposits are read out from the detector they are used to reconstruct *physics objects* that can then be further analyzed to discover new particles and interactions. The layout of the CMS detector was chosen carefully to allow for the identification of several groups of particles:

1. **Jets** – sprays of particles containing mostly hadrons and photons. An ideal jet will leave multiple tracks in the tracker that point back to the same spot on the beamline and deposit a large amount of energy in the HCAL and a smaller amount of energy in the ECAL as well.

2. **Isolated Photons** – single photons that were likely produced in the hard collision, rather than being radiated by a charged particle in the detector. An ideal single photon will leave no tracks and deposit energy in only the ECAL

3. **Isolated Electrons** – single electrons that were likely produced in the hard collision, rather than being produced as a secondary particle in the detector (for instance in pair production). An ideal isolated electron will leave a hit in every layer of the tracker and deposit energy only in the ECAL along the line traced out by the tracker hits. Bremsstrahlung photons are typically radiated from electrons in the tracker, therefore we should also expect kinks in electron tracks accompanied by energy deposits in the ECAL associated with the kinks.

4. **Muons** – expected to leave hits in every layer of the inner tracker. When extrapolated through the detector, the inner tracker hits should match up with standalone muon tracks in muon system. Nearly no energy should be deposited in the ECAL or HCAL

5. **Tagged Jets** – jets that have been tagged as originating from a specific particle, for instance a b-quark or a τ-lepton. These objects typically use information from the jet tracks to give a measure of how likely a specific jet was to represent the decay dynamics of the specific particle in question.

In order to reconstruct these objects from energy deposits in the detector, CMS employs an algorithm called particle flow[83] which will be outlined in section 2.3.3. However, the first step in object generation is the vertex and track reconstruction.

2.3.1 Track Reconstruction

Track reconstruction is the process of turning hits in the tracker into proposed trajectories for charged particles. The track finding software used by CMS is called the Combinatorial Track Finder [44, pg. 12], which is based off of the Kalman filter [56] track fitting method. The basic idea is to work iteratively, first building the more "obvious" tracks for high p_{T} particles, and then removing their hits from the collection thus making the next track easier to find.

Each iteration begins with a "seed track." In the first iteration, the seed tracks are built from collections of 3 hits, one in each layer of the pixel tracker, with curvature tolerance such that the particle p_{T} is larger than 0.8 GeV. In later iterations, seed tracks are built from hit groups that are missing a layer in the pixel tracker, or even from hit groups solely in the strip tracker to accommodate the possibility of secondary decays.

The seed tracks are extrapolated into the next layer of the tracker, taking into account the uncertainty in the track's position and momentum, by assuming a constant magnetic field and disregarding the possibility of multiple scattering and energy loss when the particle is not within a tracker layer[9]. After the expected trajectory is built, hits within some number of standard deviations of the expected trajectory in the nearest tracker layer are collected into groups of possible energy deposits for the particle[85]. This process can include the possibility of adding a "ghost hit" if no hit was found in a module along a particular possible trajectory.

A χ^2 test is then performed between the expected trajectory and the hit groups, taking into account the hit and trajectory uncertainties. The number of hits in a typical event is enough

[9]Energy loss is modeled statistically when the particle is in the tracker material, which has the effect of broadening the track uncertainty.

that most seed trajectories can be associated with multiple paths at each layer. After possible trajectories are built out of the hit groups, only a small number of possible trajectories are extrapolated into the next layer based on their χ^2 values and number of ghost hits. This process is repeated until track quality parameters, for instance too many ghost hits, can no longer be satisfied at the next layer. After the hits in the detector are organized into tracks, another fit is applied to smooth the track trajectory using a Kalman filter fit. During this stage, hits which are outliers can be disassociated with the trajectory and added back the general collection of hits. The smoothing step occurs again until all outliers have been removed[44, sec 4.3].

2.3.2 Vertex Selection

Each event recorded at CMS has a *primary vertex*, which is roughly defined as the point along the beamline at which the highest energy tracks emerged. This process involves first selecting the highest quality tracks from track reconstruction, then clustering them into sets that seem to originate from roughly the same point along the beamline, and finally fitting the location of the vertex based on the cluster end-points.

Tracks are selected based on the significance of the transverse impact parameter to the beamline ($\frac{|d_0|}{\delta d_0}$ <5), number of hits ($\geq$ 2 pixel, $\geq$ 5 pixel+strip), and the χ^2 value from the track fit ($<$ 20). Again the vertex selection is an iterative process, at first all tracks are assumed to come from one vertex, then at each step vertices can be split. In practice, this is achieved with a process called *deterministic annealing* (DA), briefly outlined in appendix A, which is based on ideas from thermodynamics of minimizing free energy at some temperature.

Once vertices are identified and tracks are assigned to a vertex, vertices with at least two tracks are fitted using an adaptive vertex fitter [55] to compute position and quality quantities such as the position of the vertex in space and weights for the likelihood that the tracks associated with the vertex genuinely began there. The weight, w_i, of a track i is close to 1 if its position is near the vertex and close to 0 if its position is several standard deviations away. The quantity n_{dof}

is used a quality parameter defined as

$$n_{\text{dof}} = -3 + 2 \sum_{i=1}^{\#\text{ tracks}} w_i.$$

When more than one primary vertex can be found in an event, the vertex with the highest scalar sum of p_{T}^2 for all the tracks associated with the vertex is chosen. In the analysis presented in this thesis, we only select events where the primary vertex has $n_{\text{dof}} > 4$, and the vertex is located within 24 cm from the center of the detector along the beamline direction and within 2 cm from the beamline in the transverse plane.

2.3.3 Particle flow

The particle flow algorithm is in charge of turning tracks and energy clusters from the calorimeters into physics objects with energy, momentum, and location information. It achieves this goal through an iterative process of linking energy deposits, much like the tracking algorithm outlined above in sec 2.3.1. Energy deposits in the Calorimeters, tracks in the inner tracker, and tracks in the muon chambers are all linked together pairwise based on various criteria [83, sec. 4]. The exact criteria for particle flow identification can be found in detail in the previous citation. In broad strokes, the energy deposits in the detector are turned into electrons, photons, muons, charged hadrons, and neutral hadrons based on the following criteria:

Electrons are identified when tracks in the inner tracker are linked with energy deposits in the ECAL. When such a link is found, a Gaussian Sum Filter (GSF) [11] algorithm reconstructs the track because kinks in electron tracks due to bremsstrahlung radiation are not handled well by the standard Kalman Filter tracking as it does not account for momentum changes due to radiation. An energy deposit in the ECAL without a corresponding track is reconstructed as a photon. Muons are reconstructed by comparing tracker muons, tracks in the inner tracker with at least one associated hit in the muon system, with standalone muons, reconstructed tracks in the

muon system. Reconstruction of a muon requires a tracker muon and a standalone muon with compatible trajectories and the absence of a large energy deposit in the calorimeters. Finally, energy deposits in the HCAL are considered neutral hadrons, and tracks which leave energy deposits in the HCAL, but not the ECAL, are considered charged hadrons.

After the energy deposits are clustered into particles, we use the anti-KT algorithm with a radius of $R = 0.4$ in (η, ϕ) space to cluster particles into jets.[82]

2.3.4 Electron Measurement Pipeline

High energy electrons have energy measurements based mainly on their ECAL deposit. When a charged particle moves through the tracker, a momentum measurement can be found as a function of its curvature as described in the appendix B. However, due to the likelihood of bremsstrahlung radiation, the sagitta of an electron track has a high chance of being malformed, containing kinks caused when a high energy photon is emitted. Since the sagitta momentum measurement assumes the free motion of a charged particle in a magnetic field with constant energy, the track from an electron that emits a high energy photon can not always be used to determine the electron's original momentum.

When the electron reaches the ECAL, almost all of its energy is deposited within a small region. For instance, a 120 GeV electron typically deposits 95% of its energy within a 5x5 crystal array. However, when passing through the tracker, an electron will radiate 33% of its energy on average before it reaches the ECAL, this figure inflates to about 86% of the energy for electrons that pass through the densest parts of the detector near $|\eta| = 1.4$ (seen in fig. 2.8). Due to the bending of the electron in the magnetic field, bremsstrahlung photons will tend to spread out in the ϕ direction, with the spread in the η direction being mostly negligible [5, sec. 4.1].

In addition to the momentum measurement, the Kalman Filter track building method is poorly suited for electron tracks as it also assumes the motion of constant energy charged particle in a magnetic field. In order to construct electron tracks, CMS also utilizes another track finding

algorithm, developed specifically for the reconstruction of electron tracks called the Gaussian Sum Filter (GSF)[11]. The GSF algorithm can be intuited roughly as several Kalman Filters running in parallel, each assuming different rates of radiation.

A high energy photon radiated by an electron will leave an energy deposit in the ECAL and change the curvature of the electron's trajectory. These deposits can typically be found by tracing tangents from the electron's trajectory to the ECAL, due to the high likelihood that bremsstrahlung photons will be emitted along the electrons momentum vector.

To measure the total energy of the electron, the energy of the bremsstrahlung photons should be included as well, and so reconstruction software clusters ECAL energy deposits into *superclusters*. Track reconstruction with the GSF algorithm is similar to that of the Kalman filter described in sec. 2.3.1, but, in addition to seeding tracks from pixel hits, ECAL superclusters are also used as seeds and tracks leading into the end of a supercluster are extrapolated backwards towards the beamline assuming both charge hypotheses. Ultimately, the electron's momentum is reconstructed using a weighted combination of the track momentum and supercluster energy. Electrons with energy less than 15 GeV use almost exclusively the pixel track momentum, while electrons with energy greater than 250 GeV use the supercluster measurement exclusively.

A series of corrections are applied to the supercluster energy. These corrections are derived in simulation where true electron energies and momenta are known to perfect precision. They include the effect of deposited energy not being properly associated the supercluster, loss of energy due to gaps in and between detector modules, pileup, and so on.[5, sec 4.8]

Differences in data and simulation are due mainly to imperfect modeling of the tracker material. Figure 2.18 shows the dilepton mass of Z→ee events in both data and simulation. The events in data are estimated to have 2% background contamination[26, sec. 4.5.1] and are selected by requiring two opposite charge electrons with dilepton mass in the range 60-120 GeV, and where each lepton has at least 25 GeV of p_{T}. These comparisons are made for different ranges of η and p_{T} as well as instantaneous luminosity and "quality class" based on the scale and number

of bremsstrahlung photon deposits found in the electron's supercluster. Although the agreement between data and simulation is excellent, a correction is applied to data so that the overall energy scale matches that seen in simulation.

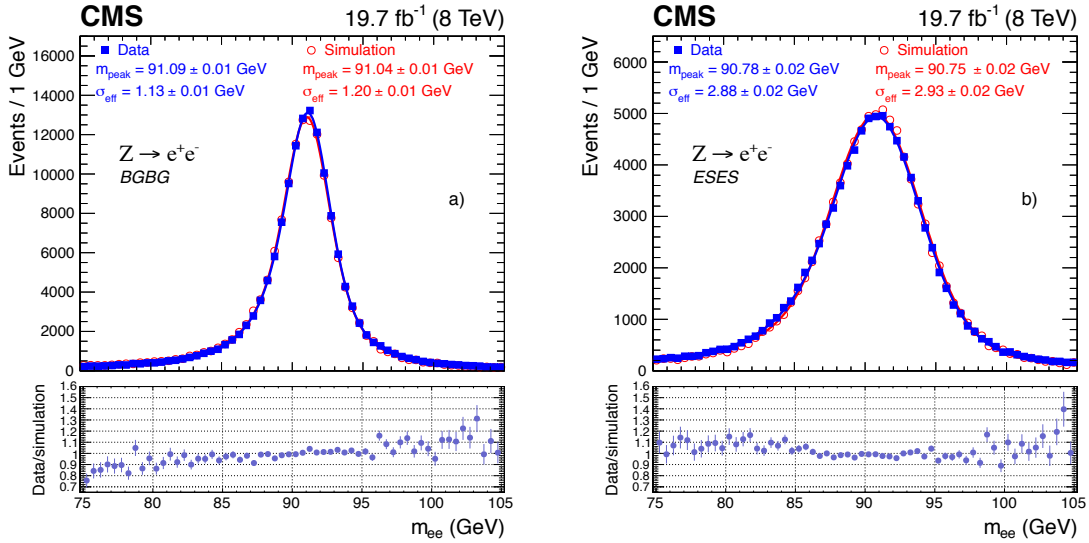


Figure 2.18: Data and simulation of Z to di-electron events at 8 TeV. On the left, only electrons in the barrel which have not undergone any bremsstrahlung radiation are plotted, these should be the easiest class of electrons to measure. To the right, only electrons in the endcap with multiple subclusters in their ECAL supercluster are plotted, these should be the hardest class to measure. In both cases there is excellent agreement between data and simulation meaning that the corrections applied to data derived from simulation are reasonable. Additionally, the peak Z masses match the established value to approximately the percent level in both cases. Taken from [5].

The agreement between data and simulation is important as it shows excellent end-to-end agreement between data and simulation, which includes the calculation of matrix elements, simulation of hadronization, and detector modeling.[10] Figure 2.19 shows the final measured energy resolution for electrons studied at $\sqrt{s} = 7$ TeV. To find the resolution, $Z \to ee$ candidates are chosen and the dilepton mass is compared against the Breit-Wigner distribution expected for the Z mass from theory. The resolution is broken up as a function of *R9*, which is roughly a measure of how concentrated the electron's energy was in the supercluster, with 1 meaning all

[10] further described in sec 2.4

the energy was in a 3x3 crystal grid. From the figure, we can see that the typical uncertainty for electrons in this analysis is at the few percent level.

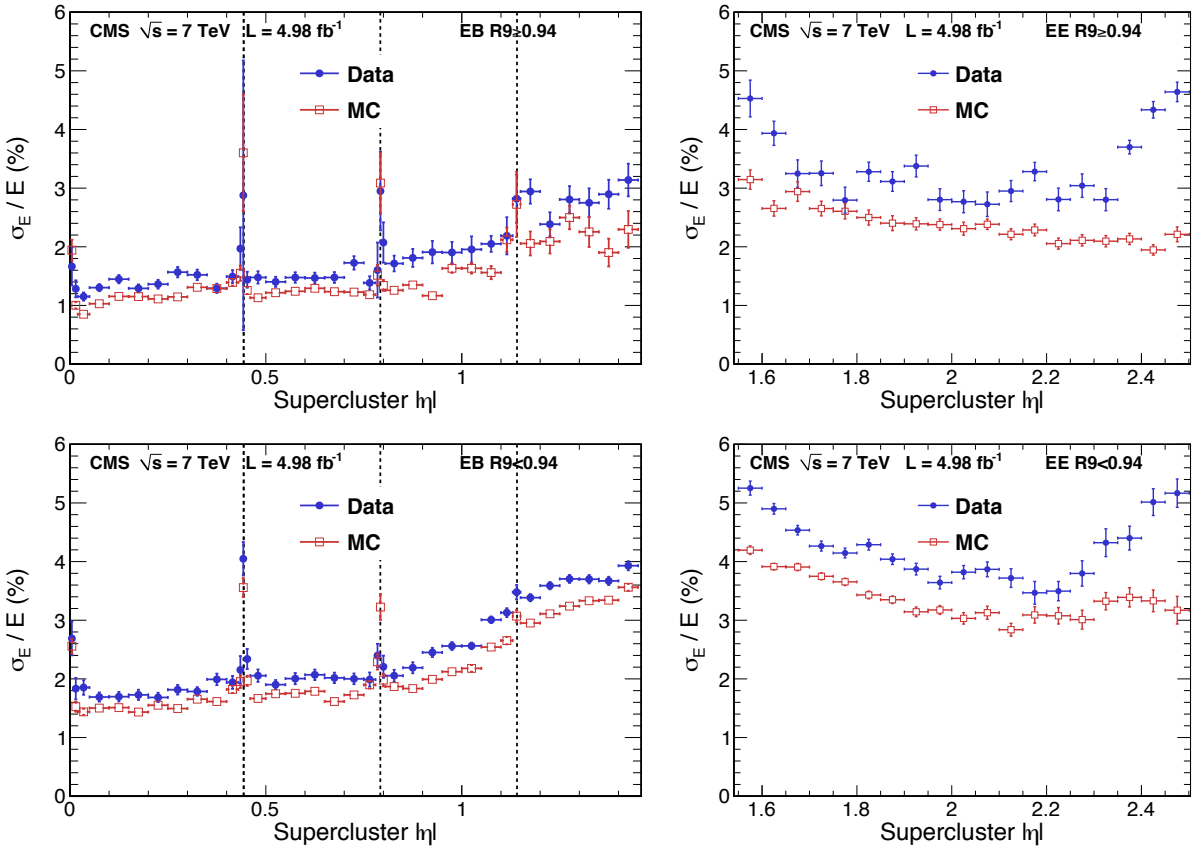


Figure 2.19: The energy resolution on electrons as a function of pseudorapidity. To find the resolution, $Z \rightarrow ee$ events are chosen in data and Drell-Yan simulation. The dilepton mass of those events are compared to the theoretical prediction of a Breit-Wigner line-shape for the Z mass peak, convolved with a Gaussian function to absorb the non-Z background tail distribution. Taken from [26].

2.3.5 Muon Measurement Pipeline

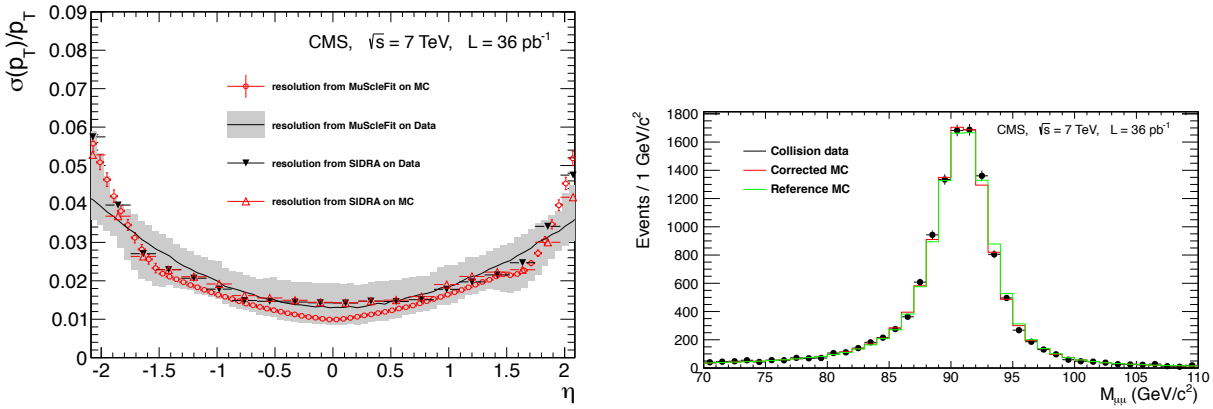
Muon reconstruction starts with tracks from the inner and tracker and muon system. Muon identification in particle flow can come from several sources, but in this analysis we only use "global muons".

Tracks found in the muon system are called "standalone muons." For each standalone

muon track, a collection of possible tracks in the inner silicon tracker that could correspond to the standalone muon are identified. These tracks are then propagated forward to a common plane with the standalone muon track which is propagated backwards to the same plane. Position information on that plane, along with momentum measurements and uncertainties are used to select the best match. [41, sec 3] [38, sec 5.1] After a match is found, the Kalman Filter track fitting method is applied from scratch to the hits in the inner tracker and muon system to smooth the overall combined track, this constitutes the global muon track used in this analysis.

Muon momenta are computed from the track sagitta, as explained in appendix B. At low energies, the segment of the track in the tracker provides the best momentum resolution due to the tighter density of the inner tracker modules. At energies above approximately 200 GeV, the tracks in the muon system provide a better measurement due to their longer lever arm.

The largest sources of energy mismeasurement comes from misalignment of the tracker and muon modules, incorrect modeling of the magnetic field, and mis-modeling of the tracker material.[41, sec. 6] The muon energy resolution measured during the 7 TeV run can be see in figure 2.20a. The solid black line shows the agreement between data the theoretical Breit-Wigner line shape for the muon momentum resolution. Figure 2.20b shows the agreement between data and Drell-Yan simulation. For our analysis, the expected momentum resolution for muons is at the percent level. More detail about the precise methodology can be found in the previous reference.



(a) Momentum resolution

(b) Di-muon mass spectrum

Figure 2.20: **(a):** The momentum resolution muons in our expected p_T range measured in the CMS detector at $\sqrt{s} = 7$ TeV. The solid black line, labeled MuScleFit, shows the resolution obtained by fitting a smearing function[41, sec. 6.1] to the Breit-Wigner line shape. **(b):** Data and simulation of Z to di-muon events at 7 TeV. The black dots represent data from the detector, while the green line-shape represents the expected line shape for $Z \rightarrow \mu\mu$ events using the baseline detector simulation. Taken from [41].

2.3.6 Photon Measurement Pipeline

Photons are identified as ECAL deposits that do not have an associated track in the tracker. The main source of contamination in the photon collection comes from the decay of π^0 mesons to two photons.[46] Many of the requirements in the sections 3.3.5 are made to mitigate the likelihood this will occur, for instance the tight photon isolation and the requirement that there not be a relatively non-trivial energy deposit in the HCAL near a photon energy deposit.

Photons have a sizable chance of converting into an electron-positron pair in the detector as was shown in figure 2.8. Therefore, electron pairs passing certain criteria, e.g. small angular separation, are also considered photons.

Figure 2.21 shows the expected energy resolution for photons as predicted by simulation. The resolution for photons is again around the percent level, very similar to the lepton resolutions shown in the previous sections. Without referencing the figure above, the energy resolution of photons could have been anticipated from the figures in section 2.3.4. This is because most electrons will have some fraction of their energy measured as bremsstrahlung photons. In other

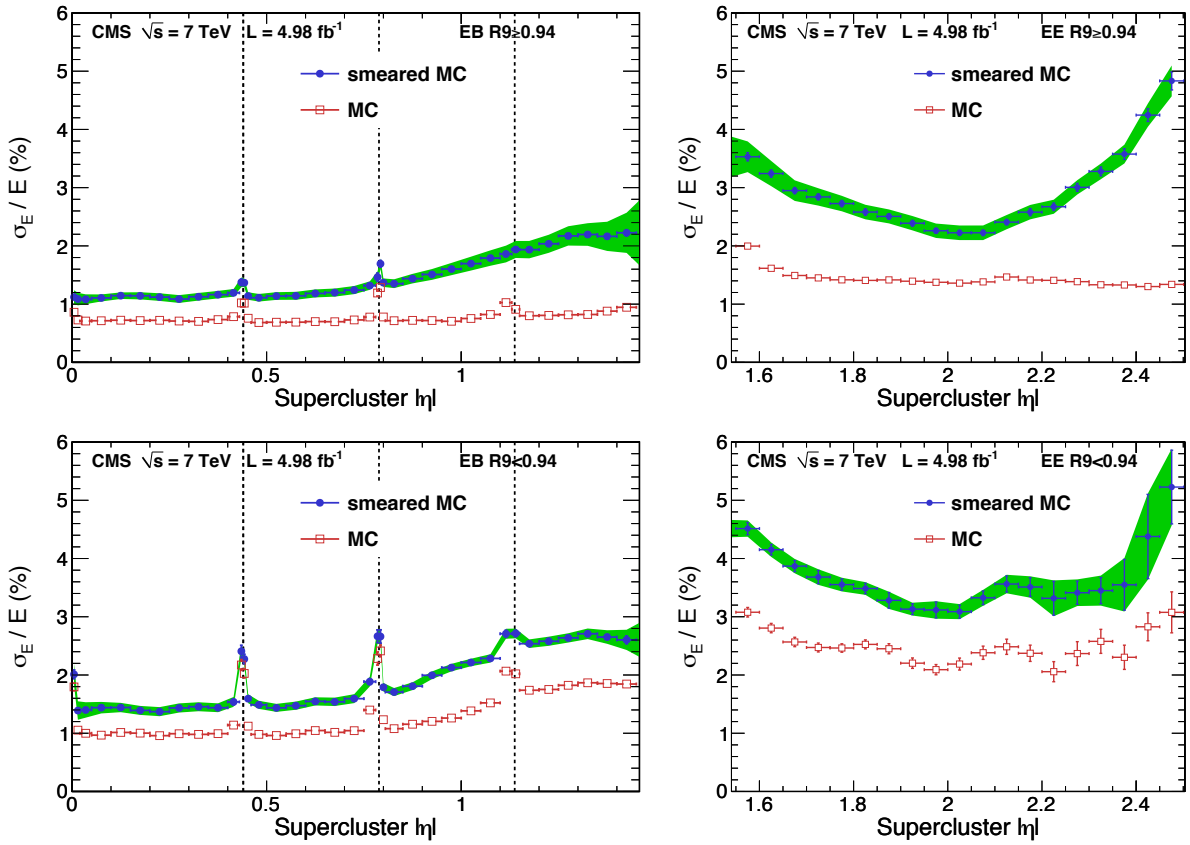


Figure 2.21: The energy resolution on photons based on predictions from simulations. Taken from [26].

words, the energy resolution of electrons is determined, to a moderate degree, by the energy resolution of photons.

Figure 2.22 also shows another test of the photon energy resolution. In the figure, Drell-Yan to dimuon events are selected with photon final state radiation. That the mass of the 3-object system matches the Z peak in simulation is another strong indicator that the energy resolution for photons in simulation is well-modeled.

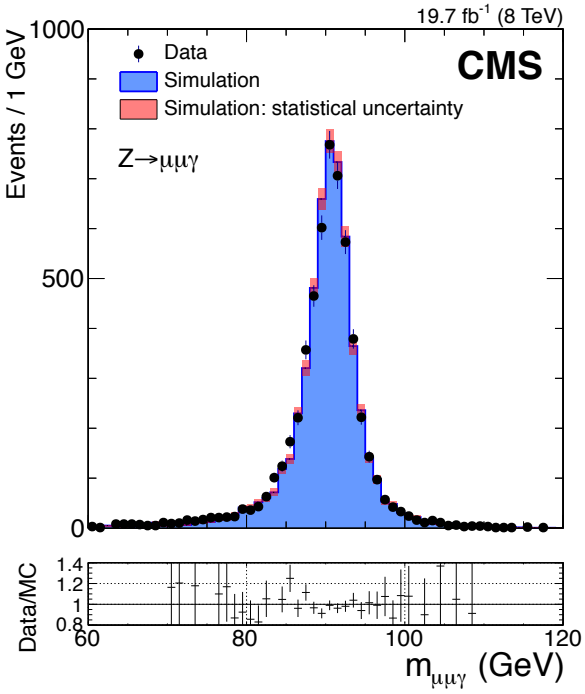


Figure 2.22: The mass of the $\mu\mu\gamma$ system in $Z \to \mu\mu$ events with γ final state radiation. The events are selected by requiring an opposite charge muon pair, each with $p_{\mathrm{T}} > 10$ GeV, $m_{ll} > 35$ GeV, and a photon passing loose selection criteria with $p_{\mathrm{T}} > 10$ GeV. The agreement between data and simulation shows further evidence that the photon energy resolution is in line with that of muons and electrons. Taken from [46, sec. 5].

2.3.7 Jets

The energy measurement pipeline of photons, muons, and electrons were described above. The final particle flow objects to consider are charged hadrons, which are reconstructed from tracks that lead to energy deposits in the HCAL, and neutral hadrons, which are energy deposits in the HCAL that do not have any associated tracks.

After the particle flow algorithm identifies all particle candidates in an event, jets are reconstructed by an algorithms that clusters particle candidates. This analysis uses the anti-k_T algorithm with a radius parameter of $R = 0.4$[23] [16].

Roughly, the algorithm starts with the largest p_{T} particle in the event and then attempts to add more particles to make a jet. The decision whether to add two particles is based on comparing the number of distance parameters between the particles in η-ϕ space, $\frac{(\delta\eta)^2+(\delta\phi)^2}{R=0.4}$, to

the momentum of the higher p_{T} particle in the pair. If energy deposits are added together, they are then subsequently treated as a single particle. If no other deposits can be found that are close enough relative to the momentum of a cluster, the cluster is finalized as a jet.

Jet energies are the hardest to get correct, this is mainly due to the difficulty in correcting for pileup, specifically inability to do targeted pileup rejection against neutral hadrons that come from auxiliary collisions. To mitigate the effects of pileup energy, charged hadrons not associated with the primary vertex in an event are not used for the anti-k_T clustering. Several other corrections are applied to jet energies, the suite of these corrections are called the JECs (Jet Energy Corrections).

Detail about the JECs can be found in references [64], [39], [30], [4], and [31]. The corrections are organized into several steps defined below:

1. **L1 Pileup** – Correct for energy added to jets due to pileup. The rough idea is to find the average amount of energy deposited in the detector by pileup collisions as a function of the instantaneous beam luminosity and region of the detector. An offset correction is found based on the jet p_{T}, area, and location in the detector. The measurement is actually performed in simulated QCD multijet events with and without pileup overlay from the minimum bias dataset.[11] The detector is broken into a grid in the $\eta - \phi$ plane, and the difference between the energy in the grid with and without the pileup overlay, Δp_{T}, is computed. The median of this difference across all grid squares in an event is called ρ, and it is combined with the jet η, p_{T}, and area (A), to find a formula for the offset correction by fitting a function described in the reference provided at the end of this bullet. The best parameterization is found for different ranges of the instantaneous luminosity of the zero bias events, making the correction a function of the jet p_{T}, η, and area, as well as the instantaneous luminosity for the event and the median grid energy ρ.

[11] The minimum bias dataset is a bank of simulated events meant to represent pileup. The energy deposits in these events are added to the QCD MC in order to simulate real pileup conditions on top of the simulated event.

Data events get an additional correction for the difference between the detector and its simulation by comparing the energy deposited in random cones around the detector in zero bias trigger events against QCD multijet simulation.[12] [64, sec. 4.3].

2. **L2L3 Response** – Correct for the response of the detector using MC jets. QCD dijet events are simulated and corrections are derived as a function of the jet p_T and η so that the reconstructed jet p_T is equal to the generated p_T and the response of the detector is uniform with respect to p_T and η.[39, sec. 6.3] [64, sec. 5]

3. **L2L3 Residual** – Correct for the difference between the MC and data. All above corrections are applied to events with a photon and jets or a Z boson which decays into a dielectron or dimuon pair. Because the energy resolutions for the electromagnetic objects are much better than the resolution for jets, the jet energies can be corrected such that the total transverse momentum in the event is zero.[39, sec. 6.3] [64, sec. 6]

After all the corrections are applied, the final jet p_T resolution comes out to be approximately 10% [39, sec 7.2], as can be seen in figure 2.23. In the figure, the dotted red line represents the p_T resolution measured in simulated dijet events under the hypothesis of transverse momentum balance, and the black dots represent the same for dijet events in data. There is excellent agreement between the resolutions in simulation and data in the barrel, and reasonable agreement in the endcap.

2.3.8 MET Reconstruction

A central object in this analysis is the transverse momentum imbalance, also called E_T^{miss} for missing transverse energy.[13] The E_T^{miss} is defined as the negative of the sum of all the

[12] The zero bias trigger collects collision data at complete random, meaning it should be populated largely by low energy QCD events.

[13] In particle colliders, the mass energy for particle's of interest is typically much smaller than the kinetic energy. In this limit, the energy and momentum are the same.

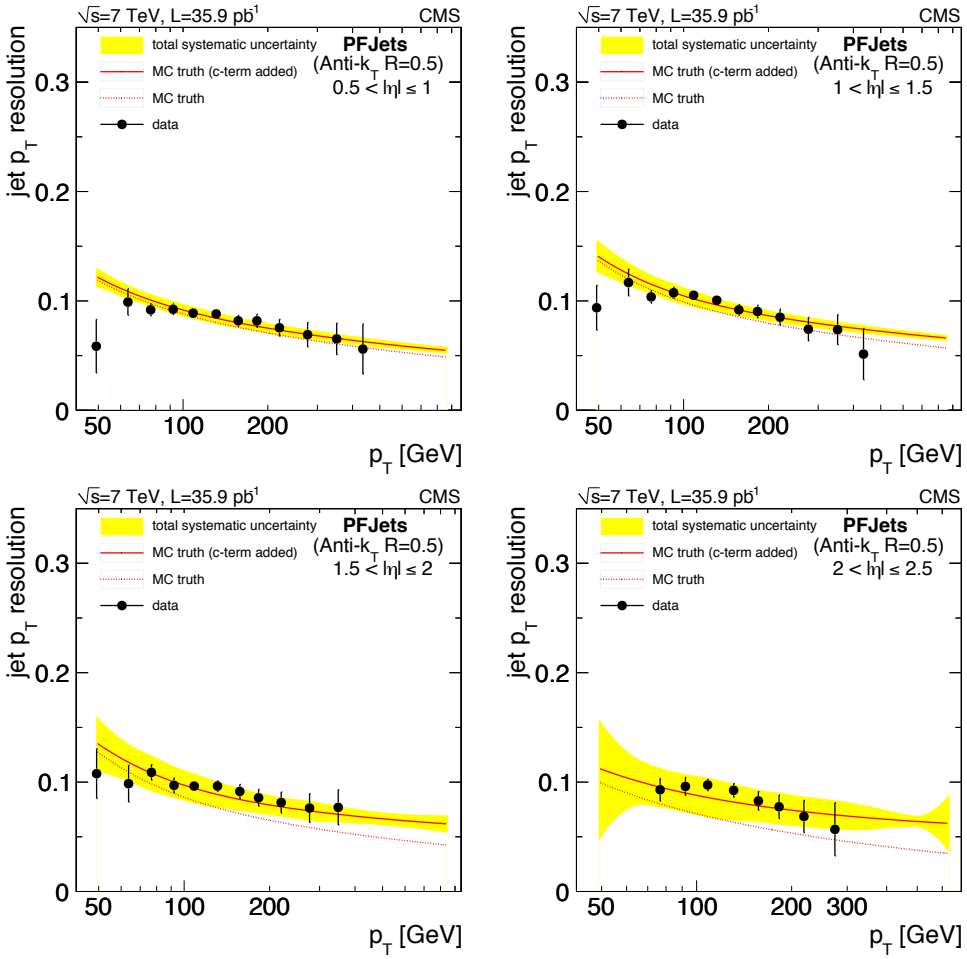


Figure 2.23: The p_{T} resolution for Anti-k_T jets in different η regions of the detector. The black dots represent the typical size of the correction needed to bring dijet events to transverse momentum balance. The red dotted line represents the difference in simulation between the generated jet p_{T} and the p_{T} reconstructed after detector simulation. The solid red line represents the same as the dotted red line, except with a small residual correction for the different resolutions found data and MC. Taken from [39].

transverse momentum of particles associated with the primary vertex:

$$\vec{E}_{\mathrm{T}}^{\mathrm{miss}} = -\sum_{\text{all objects}} \vec{p}_{\mathrm{T}}$$

In other words, the addition of an object with transverse momentum equal to the $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ vector would bring any event into transverse momentum balance.

$E_{\mathrm{T}}^{\mathrm{miss}}$ is generated from three main sources:

1. **Real $E_{\mathrm{T}}^{\mathbf{miss}}$** – Also called genuine $E_{\mathrm{T}}^{\mathrm{miss}}$, is $E_{\mathrm{T}}^{\mathrm{miss}}$ that comes from particles which are not detected by the CMS detector. In the case of the standard model, the only particle which is expected to leave no signature in the detector are neutrinos. For the sake of this thesis, real $E_{\mathrm{T}}^{\mathrm{miss}}$ is identical to the energy of neutrinos in the event.

2. **Fake $E_{\mathrm{T}}^{\mathbf{miss}}$** – Also called artificial $E_{\mathrm{T}}^{\mathrm{miss}}$, is $E_{\mathrm{T}}^{\mathrm{miss}}$ that comes from mismeasurements of physics objects. Given the percent level energy resolutions on photons and the light leptons, fake $E_{\mathrm{T}}^{\mathrm{miss}}$ in this thesis is assumed to come only from jet mismeasurements.

3. **Anomalous $E_{\mathrm{T}}^{\mathbf{miss}}$** – This is $E_{\mathrm{T}}^{\mathrm{miss}}$ which comes from various sources of noise in the detector, e.g. a dead cell in the ECAL or when a particle directly interacts with light guides or photomultipliers [6, sec 5]. A set of *MET filters*, which will be outlined in section 2.3.9, are applied to remove these spurious sources. Figure 2.25 shows the effect of using these filters with the first 12.9 fb^{-1} of the data collected for this analysis.

The Type 1 Correction

As has been demonstrated in sections 2.3.4, 2.3.5, and 2.3.6, the momentum resolution of electrons, muons, and photons is at the single digit percent level. In section 2.3.7, we showed the energy resolution for jets is several times larger, typically in the 10% range. Therefore, in events with multiple jets and no real or anomalous $E_{\mathrm{T}}^{\mathrm{miss}}$, it is expected that the bulk of the transverse momentum imbalance will be due to mismeasurements of the jet energies.

As was explained in section 2.3.7, a series of corrections are applied to the jets in order to mitigate the uncertainty associated with jet energies. Roughly, when the corrected jet energies are used to compute the $E_{\mathrm{T}}^{\mathrm{miss}}$, we call this quantity the *Type 1 Corrected MET*.

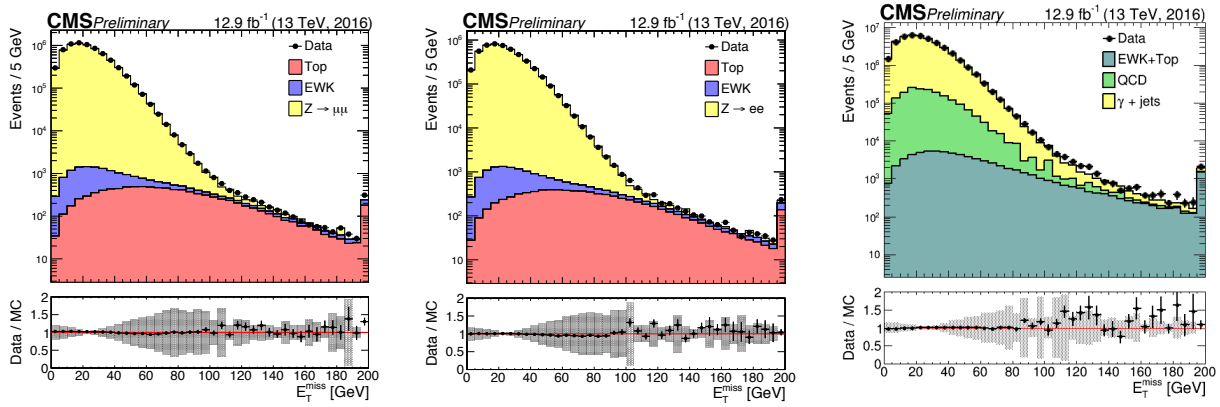
In practice, the corrections are only applied to jets which deposited less than 90% of the jet's energy into the ECAL, and muons in jets are subtracted from the jet energy before it is corrected, then added back so that only the energy in the jet not associated with the muon is

corrected. The purpose of these two caveats is that the JECs are somewhat dependent on flavor composition. For instance, a spray of pions will have a different detector response than a spray of pions, photons, and muons. The JECs are derived via dijet events, which means they are likely seeded by gluons whose decay chains tend to not include muons.

Agreement Between Data and Simulation

The modeling of both real and fake $E_{\mathrm{T}}^{\mathrm{miss}}$ in this analysis is done largely via data driven techniques which are not sensitive to $E_{\mathrm{T}}^{\mathrm{miss}}$ modeling in simulation. However, in some cases, namely the electroweak subtraction in sec 3.5.2, the Z+ν background in sec 3.5.4, and the Z+Hadronic closure test in sec 3.5.2, we do rely on simulation for $E_{\mathrm{T}}^{\mathrm{miss}}$ modeling.

Figure 2.24 shows the agreement in the type-1 corrected $E_{\mathrm{T}}^{\mathrm{miss}}$ profile between data and simulation in events with a Z boson that decays to an electron-positron pair (2.24a) and a $\mu^+\mu^-$ pair (2.24b). Additionally, figure 2.24c shows the agreement between data and simulation for events with an isolated photon. In general, there is excellent agreement the bulk, while most bins are within one standard deviation.



(a) Z→ e^+e^- (b) Z→ $\mu^+\mu^-$ (c) Isolated photons

Figure 2.24: Agreement between data and simulation for (a) Z→ *ee*, (b) Z→ *μμ*, and (c) isolated photon events. The gray bands on the ratio plot show the uncertainty due to the JECs. There is excellent agreement in the bulk, and most bins are within the expected one-σ fluctuation in the tails. Taken from [6].

2.3.9 MET Filters

As explained in sec 2.3.8, various low probability phenomena can cause large energy mismeasurements leading to large $E_{\mathrm{T}}^{\mathrm{miss}}$ values in events. The following are a list of *MET filters* applied as event level vetos in this analysis, they attempt to tag events with such spurious signals.

- CSC tight beam halo 2016 filter
- HBHE noise filter
- HBHE iso noise filter
- ee badSC noise filter
- ECAL dead cell trigger primitive filter
- JetMET muon fraction filter

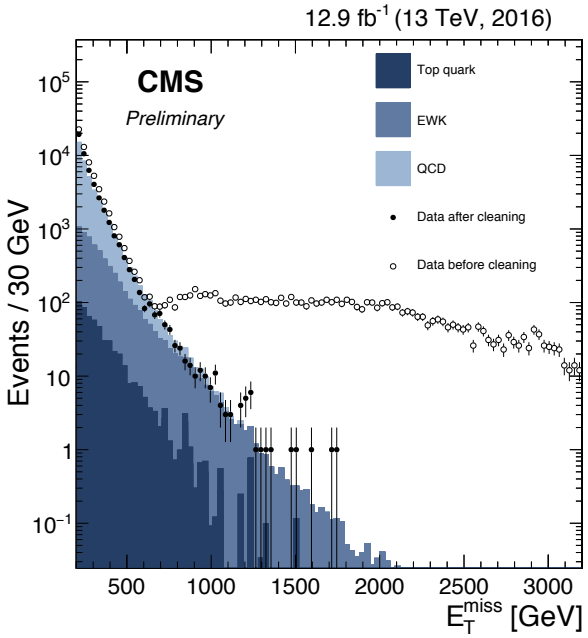


Figure 2.25: The effect of $E_{\mathrm{T}}^{\mathrm{miss}}$ filter application on dijet events in the first good 12.9 fb^{-1} of 2016 data taken by CMS. The filters used in this figure are slightly different that those used in this analysis. Taken from [6].

More detail on these filters can be found in reference [6, sec 5]. A brief summary follows:

The HBHE noise filter looks at timing and pulse shape information as well as geometrical information about the readout boxes and photodiodes to reject noise signals from the HCAL. The HBHE iso noise filter rejects events based on topological information from the tracker, ECAL, and HCAL. The ECAL dead cell trigger primitive and ee badSC noise filters reject events that have extremely large energy contributions from single crystals or small bunches of crystals in the ECAL. The CSC tight beam halo filter removes events with beam halo, a phenomenon where a proton collision occurs before the interaction point, causing particles to travel through the detector collinear to the beam at constant ϕ, in "the beams halo." The JetMET muon fraction filter removes events where a jet is found anti-aligned to the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector containing a muon that carries at least 50% of the energy in the jet. This filter suppresses events with muons that are assigned too large a p_{T}.

Figure 2.25 shows the effects of the applying these met filters to a sample of dijet events.[14]

2.3.10 B-Tagging

This analysis attempts to tag jets which contain a b-quark emitted from the primary vertex. The combined secondary vertex, version 2 (CSVv2) algorithm is used to accomplish this task. It is based off of the CSV algorithm described in reference [42].

Bottom quarks will generally hadronize into B mesons or B baryons which have a long enough lifetime that they travel a few millimeters from the beamline before decaying, creating a secondary offset vertex. Therefore, b-tagging methods tend to focus on the identification of secondary vertices, with specific track multiplicities and mass, as well as several other kinematical conditions on the shape and composition of jets[86]. The full implementation of CSVv2 includes machine learning algorithms that take in a suite of variables characterizing a jet's tracks, momen-

[14] A slightly different filter was used to reject events with muons assigned an incorrectly large p_{T}. But no muons that pass which the filter used in fig 2.25 were found after applying our baseline selections in data

tum, impact parameters, etc... and outputs a number, called the discriminator, that characterizes the probability the jet contained a b-quark. These full description is available in reference [33, section 5.1.2].

Figure 2.26 shows the CSVv2 discriminator scores for jets separated by flavor in an inclusive multijet sample. In this analysis, a jet is called a b-jet if it has a discriminator value larger than 0.8484. This corresponds to a false positive rate of approximately 1% for light quarks and gluons.

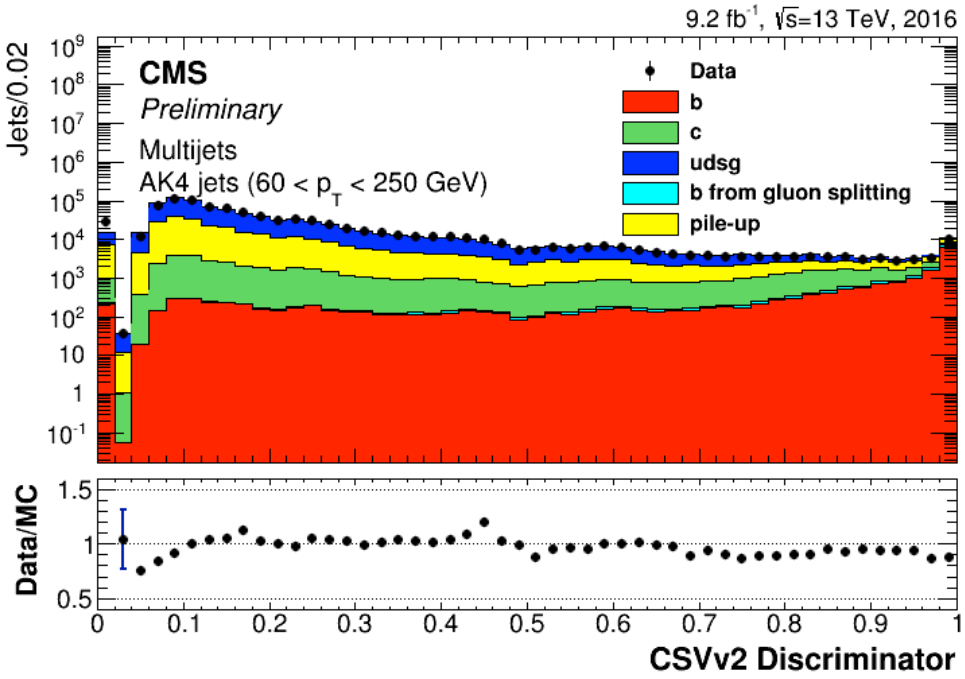


Figure 2.26: The count of jets assigned different CSVv2 discriminator values from an inclusive multijet sample. The value of the discriminator used in this analysis to tag a b-jet is 0.8484, which corresponds to a false positive rate of 1%. Taken from [32].

2.4 Physics Simulation

The physics simulation used in this thesis are listed in table 2.2. All simulation is created in three steps. First an event generator computes matrix element for various final states. The

event generator used for each sample is listed in the table. The events are then interfaced with Pythia 8.2 for hadronization and parton showering.

Table 2.2: Generators used to create physics simulation.

Process	Generator
Drell Yan + jets	MadGraph 5
γ + jets	MadGraph 5
W + jets	MadGraph 5
$T\bar{T}$ + jets	Powheg v2 (dilepton), MadGraph 5 (single letpon)
Single Top	MadGraph_aMC@NLO 2.2, Powheg v2 (s-channel)
WZ	Powheg v2
WW	Powheg v2
ZZ	Powheg v2
VVV	MadGraph_aMC@NLO 2.2
Wγ + jets	MadGraph 5
γ Z + jets	MadGraph 5
$T\bar{T}Z$	MadGraph_aMC@NLO 2.2
$T\bar{T}W$	MadGraph 5
TWZ	MadGraph 5
SUSY Signal Models	MadGraph 5

The detector simulation for standard model processes is based on a full GEANT4 [13] simulation of the CMS detector. SUSY signal simulation is interfaced with the fastsim toolkit. [9] The cross sections for some standard model processes are corrected for beyond leading order effects by matching to computations from the FEWZ package [71]. All simulated physics samples are processed through the same reconstruction software used for real physics events in data, including a simulation of the hardware triggers. In addition, pileup is simulated in events by overlaying the energy deposits found in events from the minimum bias MC sample, a sample of simulated low energy QCD events made to represent pileup collisions. This allows for high quality simulated events that can be directly compared with real data.

2.5 Datasets

The data used in this thesis was collected by the CMS detector at the Large Hadron Collider in 2016, it corresponds to 35.9 fb^{-1} of integrated luminosity. Events are seeded by triggers selecting for double electron/photon, double muon, single muon, and single photon final states. The precise selection for these events are described in sections 3.4 and 3.3.

2.6 Acknowledgments

In the previous chapter, I outlined the LHC and CMS detector, including an outline of the signal processing and discussions about performance of the detector. I would like to take time to acknowledge the many members of the CMS collaboration, the LHC team, the CERN staff, and countless others who have done such an amazing job building, documenting, and operating these unbelievable machines and tools. In addition to building and operating the hardware, the CMS collaboration has provided a tremendous software suite which abstracts the inner working of the detectors and allows scientists like me to focus on physics. Without the tremendous effort of the CMS collaboration, the work in this thesis would be impossible.

I must also specifically thank my co-authors on CMS for contributing figure 2.26 showing the B-tag CSV discriminator, figure 2.24 showing the MET resolution after the type-1 correction, and figure 2.25 showing the efficacy of the MET filters.

Chapter 3

A search for new physics in events with a Z boson, missing transverse energy, and jets

3.1 Motivations

This search for new physics is motivated by observations of dark matter in astrophysical and cosmological data as well as the hierarchy problem. Supersymmetry provides a solution to these two problems at once in theories with 'R-parity' conservation due to the existence of a lightest supersymmetric particle (LSP) which can not decay to standard model particles. If the LSP is not charged, then it is a natural candidate for dark matter.

The hard part of searching for dark matter is that it is known to not interact via the strong or electromagnetic forces with any non-trivial cross section. Therefore, the only signature dark particles will leave in a detector like CMS is transverse momentum imbalance, which is the central variable we use in this search.

3.1.1 Motivating Models

We use several simplified SUSY models to motivate our search regions and interpret our results. Feynman diagrams of the additional operators in these models are shown in figures 3.1 and 3.2. In each model, one or more of the new particle masses are free parameters. Setting maximal masses for these particles that are compatible with our observations is the intent of this style of SUSY search. These models are simulated with different values for the free mass parameters using the madgraph package as discussed in sec. 2.4. A scan is then done over the simulated mass points to check whether the expected signature for any mass point describes the data better than the background only hypothesis, this is the content of section 3.7.3.

In fig. 3.1 we have a model of gauge mediated SUSY breaking (GMSB) where the LSP is a gravitino and is assumed to have the arbitrary low mass of 1 GeV. The gluino and neutralino masses are free parameters. We refer to this model as "strong SUSY" because two gluinos are produced which cause the initial cascade of particles. The production of gluinos in this model tends to lead to many hadronic jets in the final state. The "strong signal regions", defined in sec 3.4.3, are built to target this model.

In fig. 3.2, we have three models which we refer to as "electroweak SUSY". In these models, electroweak superpartners are created in the collision without associated jets. The free parameters in these models are the masses of the electroweak superpartners. Due to the different branching ratios for Higgs and W/Z decays, two different signal regions are defined to target the left two diagrams (the VZ region) and right diagram (the HZ region). In models with a gravitino, it is again assumed to have a mass of 1 GeV.

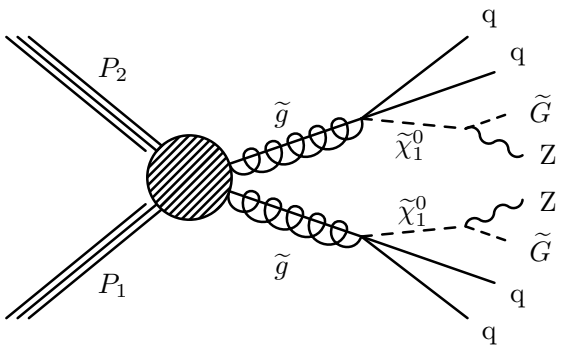


Figure 3.1: The Feynman diagram for the additional term added to The Standard Model lagrangian to produce the simplified supersymmetric model used to interpret this in analysis in the context of "strong SUSY". In this model of GSMB, the gravitino is the LSP with an assumed mass of 1 GeV. Two gluinos are produced in the collision which decay to neutral electroweak superpartners that couple to the Z boson. Because the model assumes the production of gluinos, large amounts of hadronic activity is expected.

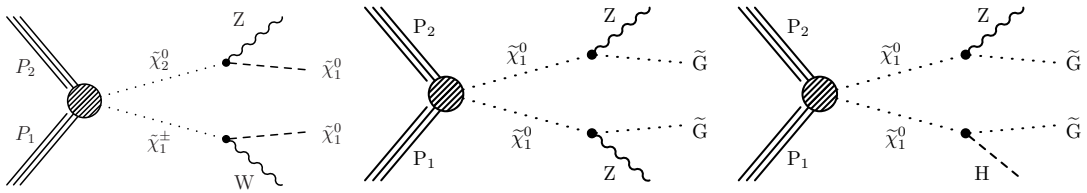


Figure 3.2: The Feynman diagrams for the electroweak SUSY models. In each diagram, electroweak superpartners are created in the collision and decay to either electroweak or Higgs bosons. The right two models again represent GMSB with a gravitino that is assigned a mass of 1 GeV.

3.2 Analysis Strategy

3.2.1 Background Considerations

Leptonic Final States

The Z boson can decay to any fermion. In theory, one could perform this analysis in an all hadronic final state, which might seem advantageous as the Z will decay to hadrons approximately 10 times as often as it will decay to light leptons.[76] However, there are several reasons leptonic final states are highly advantageous.

In hadron colliders, the most common types of final states are those with only hadronic activity, as can be seen in figure 2.3; leptonic final states provide a much cleaner population in which to search for Z bosons. Additionally, as referenced in 2.3.4, 2.3.5, and 2.3.8, the fidelity of energy measurements for the light leptons is much better than for jets at CMS. This provides great advantage for background discrimination; leptons from Z boson decays will tend to have a very specific dilepton mass, a quantity reconstructed from the momentum measurements.

The decays of the Z produce two opposite sign, same flavor fermions. A further benefit of the leptonic channel is that flavor and charge identification is fairly easy for the light leptons but nearly impossible to identify for jets at the current state of the art (for instance, one can not say with high confidence that a jet was produced by a positively charged charm quark).

Therefore, even though the branching ratio to light leptons from Z bosons is lower than the branching ratio to hadronic final states, the better energy resolution, lower background rates, and flavor identification make the leptonic final states far more powerful.

Background Sources

When using the leptonic final states, there are essentially two other background sources of leptons we must consider, these are γ and W decays. Due to the high mass of the Z boson, any $\gamma \to \ell^{\pm}\ell^{\mp}$ events can easily be vetoed since these events should have dilepton mass peaking at 0

GeV.

The W boson will decay into leptons only with their complimentary neutrino, this means that in order to select a pair of opposite charge and same flavor charged light leptons, there must be at least two W bosons in the event. Because the decays of the W will be independent, there is only a 50% chance that the two leptons will have the same flavor. As will be discussed later, this makes the background prediction for these types of events straightforward as events with two different flavor leptons can be used to model essentially any kinematical distribution.

It turns out the most common source of W bosons is through the production of two top quarks which decay to a bottom quark and a W boson. To reject these events, we will use the MT2 variable, described in sec 3.2.1. Further, the dilepton mass in $t\bar{t}$ events will be distributed essentially flatly across the Z mass window. The overall dilepton mass distribution will be a falling distribution with a small bump for the Z Boson.

Finally, due to the instability of the τ lepton, we neglect this channel from the search. τ leptons are much harder to identify than the light leptons, and further they can decay to light leptons in a flavor symmetric manner, so are predicted by that background channel's prediction method.

Hadronic Activity Requirements

As mentioned above, the major production mode of opposite sign dilepton pairs with dilepton mass near 91 GeV is from Drell-Yan production of Z bosons. The diagram for this type of process is shown in fig. 3.3. As can be seen in the figure, the leading order diagram for this process has no free quarks or gluons. That there are no free colored particles means that the vast majority of these events will likewise come without any jets.

Higher order diagrams with ISR or FSR take a cross-section production hit of roughly 1/5 due to the additional factor of the QCD coupling, α_s, in the phase space integral. This means that we can suppress the DY background by a factor of almost 100 by requiring at least 2 jets in an

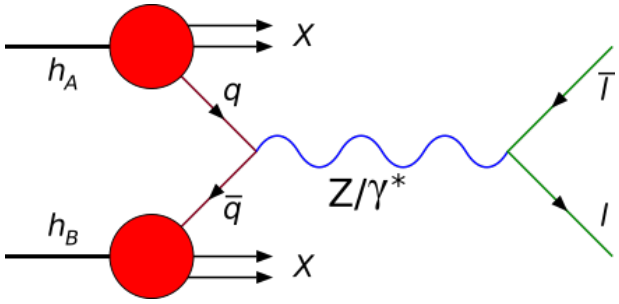


Figure 3.3: A diagram showing the Drell-Yan process. A quark and an anti-quark annihilate into an off-shell γ or Z boson, which in turn decays to a pair of opposite-charge same-flavor leptons.

event. This requirement is adopted in all search regions to keep our Drell-Yan background count low. Further, the background prediction for 0 or 1 jet events would require a completely different methodology due to different sources of $E_{\mathrm{T}}^{\mathrm{miss}}$ in high jet multiplicity final states. Therefore, this analysis limits itself to final states with at least 2 jets.

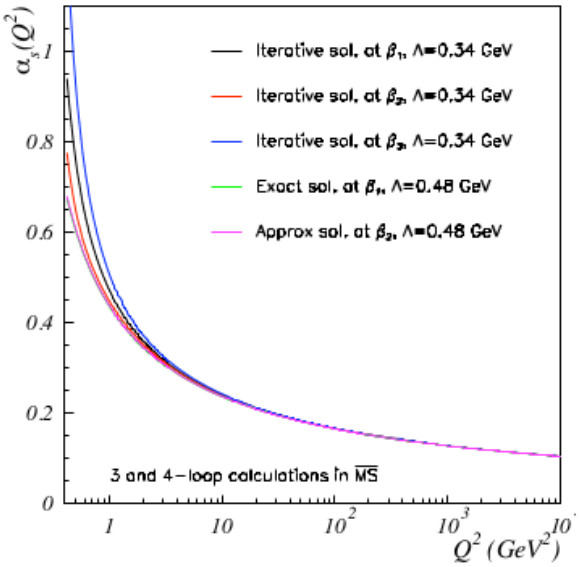


Figure 3.4: The QCD coupling constant computed to different orders and different cutoff scales. The production of the Z boson is much more likely at center of mass energy near the mass of the Z. This means that α_s is between $\frac{1}{5}$ and $\frac{1}{10}$, which can be viewed as the zeroth order multiplicative correction to the DY with a single ISR or FSR jet cross section

In addition, the strong SUSY model shown in fig. 3.1 anticipates lots of hadronic activity.

Therefore, in the strong search regions, we also bin in the number of jets[1] and require considerable hadronic activity by selecting events with H_T[2] above 200 or 500 GeV for the regions with and without b-tags respectively. This further rejects Drell-Yan events.

$E_{\mathrm{T}}^{\mathrm{miss}}$ Binning

Dark particles are expected to leave no signature in the detector except transverse momentum imbalance. In order to get maximal sensitivity to the different mass points for the SUSY models, we bin the search regions in $E_{\mathrm{T}}^{\mathrm{miss}}$.

B-Tagging

We bin search regions in number of b-tags, which are outlined in detail in sec 2.3.10. For the region targeting the HZ model, this is because the Higgs primarily decays to a pair of b-quarks. For the VZ model, we require there be no b-tags as it suppresses the $t\bar{t}$ background. For the strong search regions, we separate regions based on whether they have any b-tags, this because it causes the background composition to change dramatically; in regions with a b-tag, the dominant background will be $t\bar{t}$, whereas in regions without any b-tags the dominant background will be Z+jets. In turn, this allows us to use a slightly more aggressive MT2 cut in regions with a b-tag.

MT2

The MT2 variable is defined as

$$\mathrm{MT2} = \min_{E_{\mathrm{T}}^{\mathrm{miss}}\ \mathrm{splittings}} \left[\max \left\{ \mathrm{M_T}(p_1, \not{p}_1), \mathrm{M_T}(p_2, \not{p}_2) \right\} \right]$$

[1] This has the added benefit of creating search regions which are dominated by only one of the two dominant backgrounds, $t\bar{t}$ and Drell-Yan.

[2] H_T is the scalar sum of transverse energy for all jets in an event, above certain thresholds.

where $\not{p}_1$ and $\not{p}_2$ are decompositions of the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector, and $\mathrm{M_T}$ is the transverse mass.

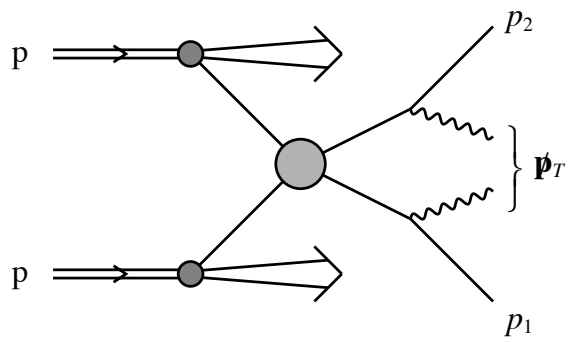


Figure 3.5: The canonical picture of a decay with an MT2 endpoint, taken from the original paper[70].

When a pair of W bosons each decay into a lepton and a neutrino, this value should not be able to be larger than the mass of the W boson. We can see this is true because for the correct splitting of the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector into the momenta of the two neutrinos, the $\mathrm{M_T}$ for each lepton is bounded by the W mass. Therefore by summing over all possibilities, we will ensure the outer minimum will select at most the mass of the W (there is no lower limit besides 0).

However, in the case of an arbitrary decay, for instance in DY or in signal models with dark matter, there is no need for this value to be smaller than the mass of the W and generally higher values will be found. Therefore, we can use this quantity as a handle for rejecting events where the leptons come from two W bosons. We require in our strong search regions that MT2 be at least 80 GeV, just about the mass of the W boson. This mainly rejects the $t\bar{t}$ background.

MT2b

Another variable used in this analysis is MT2b, or sometimes MT2($\ell b \ell b$), which can only be computed in events with 2 jets that are b-tagged. MT2 can be computed for any pair of 4-vectors along with a $E_{\mathrm{T}}^{\mathrm{miss}}$ vector. In the case of $t\bar{t}$ decays, shown in figure 3.6, one can see that by summing the 4-vectors for the lepton and b-quark, an MT2 end point should be found at the top mass, so long as the correct b-quark can be associated with the correct lepton.

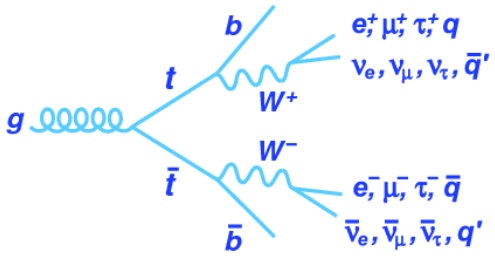


Figure 3.6: Diagram of $t\bar{t}$ production. An MT2-like endpoint can be found at the top mass if the lepton and b-quark momentum vectors are added so that choosing the $E_{\mathrm{T}}^{\mathrm{miss}}$ splitting that corresponds to the actual neutrino momenta will recover the transverse 4-vector of the top at the root of the decay. Figure taken from [8].

The same power of MT2 to reject semi-invisibly decaying pair produced particles can be used to reject $t\bar{t}$ events in a population with 2 b-jets by trying the two different combinations of b-jet/lepton pairing and selecting the minimum value out of those two. Again, in the case of $t\bar{t}$ decay, MT2b defined in this way will have the mass of the top quark as one of the possible combinations, whilst there is no such requirement for a general process. MT2b is used only in the electroweak HZ region.

3.3 Object Selection

After the particle flow algorithm identifies lepton, jet, and photon candidates, and before events can be classified as belonging to the various search, control, and closure regions, an additional set of selections are applied in order to further purify the analysis object collection. The purpose of this section is to define these selections.

3.3.1 Lepton ID

There are essentially two ways to get "fake" leptons in the particle-flow lepton population:

1. A set of calorimeter and tracker hits from hadronic activity mimics the geometry expected

by an lepton closely enough that a lepton is reconstructed from what should be called, for instance, a charged hadron and a photon.

2. A colored particle, normally a charm or bottom quark, emits an electroweak boson that decays into a real lepton. In physics parlance, this is called *non-prompt* lepton, as it is a secondary decay (*prompt* meaning created at the primary vertex).

Although (2) is a "real" lepton in the colloquial sense, they are background objects that contaminate the lepton population in the context of most analyses, which aim to study particles produced directly in proton-proton interactions. To guard against these fake leptons, an additional set of cuts, called an ID, are required to be passed by each candidate before it is considered a "good" lepton.

In this analysis, we identify two "working points" for leptons, tight and loose. These categories are defined based the lepton flavor and a set of cuts described below. Any lepton which passes the criteria to be classified as tight would necessarily pass the criteria to be classified as loose.

3.3.2 Lepton Isolation

In addition to the ID, isolation requirements are also necessary for leptons to be added to the analysis selection. In this analysis we use the variable miniRelIso, which is defined as the energy inside a variable size cone, determined by the lepton candidates p_{T}, about the lepton divided by the leptons p_{T}. The cone size is set by p_{T} in the following manner (units in GeV):

$$\Delta R = \begin{cases} 0.2 & p_{\mathrm{T}} < 50 \\ \frac{10}{p_{\mathrm{T}}} & 50 \leq p_{\mathrm{T}} < 200 \\ 0.05 & p_{\mathrm{T}} \geq 200 \end{cases} \cdot$$

The miniRelIso variable is then defined as the energy of all particle flow candidates inside the cone of size ΔR with effective area[3] corrections applied divided by the p_{T} of the lepton.

$$\text{miniRelIso} = \frac{\mathrm{E}_{\text{cone}}}{p_{\mathrm{T}}}.$$

3.3.3 Electron ID and Isolation

The electron candidates are required to pass the following cuts in addition to the particle flow identification:

Table 3.1: Requirements for electron identification in addition to particle flow. d_0 and d_z represent the closest distance of the lepton track to the primary vertex in the x-y plane and along the z axis respectively. The SIP3D variable is the impact parameter significance, $\frac{\sigma_{\text{ip}}}{\text{ip}}$, where the impact parameter is the closest distance of the lepton track to the primary vertex in 3 dimensions. The conversion veto attempts to reject photon conversions into electrons.[45, sec. 5.3]

Cut variable	Requirement
p_{T}	> 10 GeV
d_0 (w.r.t. 1st good PV)	< 0.05 cm
d_z (w.r.t. 1st good PV)	< 0.1 cm
miniRelIso	< 0.10
abs(SIP3D)	< 8
maxLostHits	$== 0$
Conversion Veto	must pass
Spring 2016 POG MVA	see below
miniRelIso	< 0.10

The tight and loose criteria for electrons is based entirely on the MVA output. The POG MVA is a boosted decision tree (BDT)[4] prepared by the CMS EGamma Physics Object Group (POG). The BDT is trained on simulated Z$\rightarrow e^+e^-$ events where electrons are considered to be real if the candidate can be matched to an electron emitted from the Z, and fake otherwise. An

[3] See sec 2.3.7

[4] A boosted decision tree is a sort of classifier that assigns a real number to a well-formed tuple of data. The algorithm tries to construct a map such that the number output for tuples in the same category are close to each other, then that number can be used discriminate between categories. Technical details are beyond the scope of this thesis, but a good review can be found [53]. In the case of the electron ID MVA, the categories are real or fake and the numbers lie between -1 and 1.

additional validation sample of mostly real electrons in data is constructed from e^+e^- events where the dilepton mass is within 7.5 GeV of the Z pole mass, $|M_{ll} - M_Z| < 7.5$ GeV, and each lepton is required to be isolated such that the energy in a cone around the electron must be less than 10% of the total energy in the cone (including the electron). A sample of mostly fake leptons in data is constructed by requiring a third lepton candidate in these events which has inverted isolation criteria and the additional requirement that the $E_{\mathrm{T}}^{\mathrm{miss}}$ in the event is less than 25 GeV to suppress WZ events.

Distributions of various kinematical quantities are constructed from simulation and data that have reason to discriminate between real and fake electrons, for instance the spread of the calorimeter hits in η. Some of these distributions are shown in figure 3.7. The MVA is trained on the simulated events then checked against the data samples described above for consistency. Finally electrons in the barrel region and endcap regions are partitioned and each set is used to train an MVA specifically targeting the detector subsystems there [5].

The working points for electrons based on the MVA output are shown below:

Table 3.2: Electron identification working points used in this analysis.

pseudorapidity region	momentum [GeV]	loose WP	tight WP
$0 < \|\eta\| < 0.8$	$10 < p_{\mathrm{T}} < 15$	-0.86	0.77
$0 < \|\eta\| < 0.8$	$15 < p_{\mathrm{T}} < 25$	$-0.96+0.10 * \frac{p_{\mathrm{T}}-15}{10}$	$0.52+0.25 * \frac{p_{\mathrm{T}}-15}{10}$
$0 < \|\eta\| < 0.8$	$p_{\mathrm{T}} > 25$	-0.96	0.52
$0.8 < \|\eta\| < 1.479$	$10 < p_{\mathrm{T}} < 15$	-0.85	0.56
$0.8 < \|\eta\| < 1.479$	$15 < p_{\mathrm{T}} < 25$	$-0.96+0.11 * \frac{p_{\mathrm{T}}-15}{10}$	$0.11+0.45 * \frac{p_{\mathrm{T}}-15}{10}$
$0.8 < \|\eta\| < 1.479$	$p_{\mathrm{T}} > 25$	-0.96	0.11
$1.479 < \|\eta\| < 2.5$	$10 < p_{\mathrm{T}} < 15$	-0.81	0.48
$1.479 < \|\eta\| < 2.5$	$15 < p_{\mathrm{T}} < 25$	$-0.95+0.14 * \frac{p_{\mathrm{T}}-15}{10}$	$-0.01+0.49 * \frac{p_{\mathrm{T}}-15}{10}$
$1.479 < \|\eta\| < 2.5$	$p_{\mathrm{T}} > 25$	-0.95	-0.01

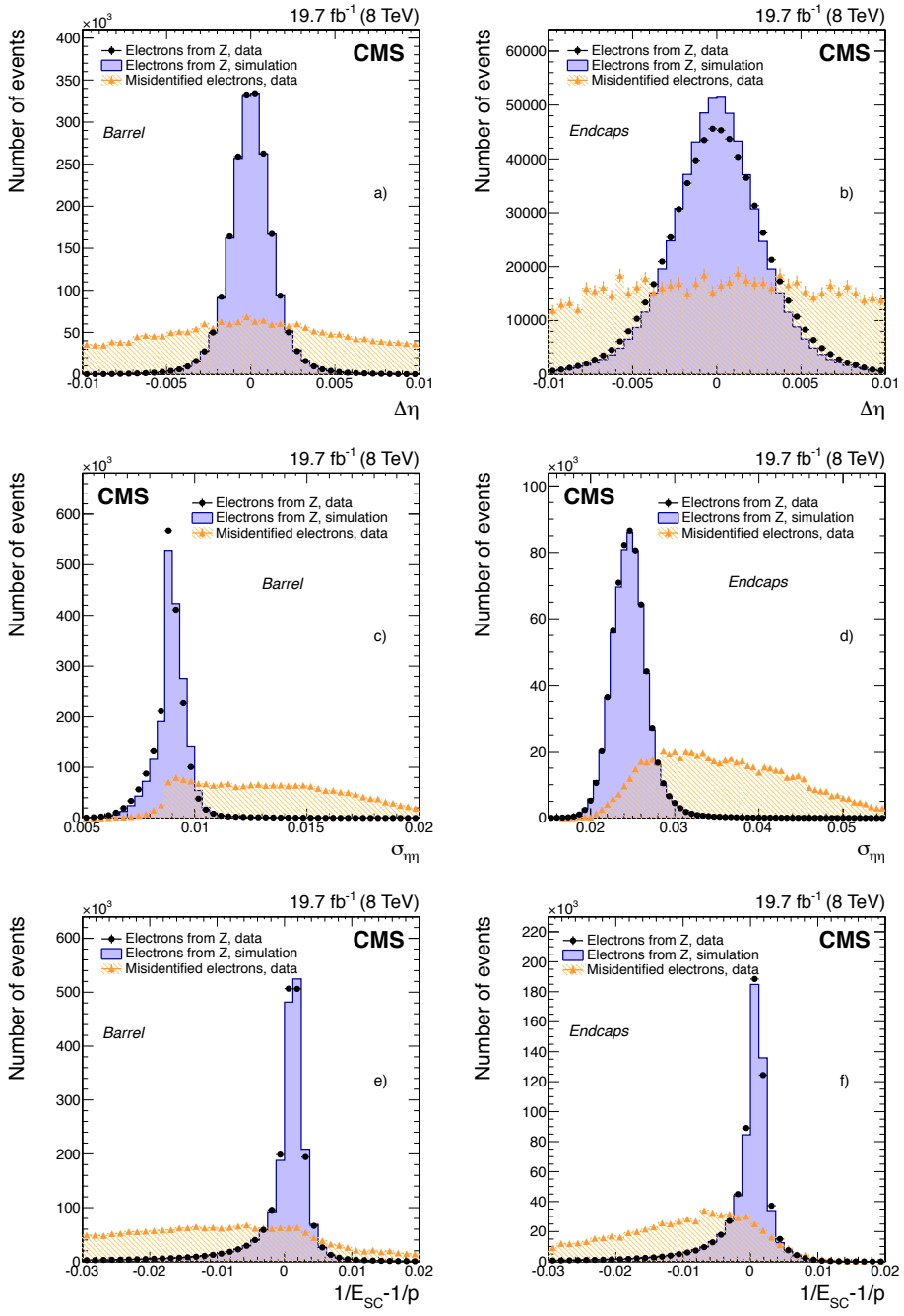


Figure 3.7: Some distributions used in the electron MVA that help discriminate between real (prompt) and fake (non-prompt) electrons. Simulation provides a guaranteed way to tag electrons as prompt or not-prompt, and so is used to train the MVA. However, it is important that data and simulation agree in these variables. $\Delta\eta$ and $\sigma_{\eta\eta}$ are variables that characterize the spread of detector hits associated with the electron in the η direction. E_{SC} is all the energy deposited into the ECAL which is associated to the electron, and p is the electron's momentum.

3.3.4 Muon ID and Isolation

The muon ID is based purely on a cut-based approach, no MVA is used. Again, the selection criteria is based upon the muon POG, which is documented in more detail at [29]. The selection used for the muon ID and isolation are defined below:

Table 3.3: Summary of the muons selection requirements. d_0 and d_z represent the closest distance of the lepton track to the primary vertex in the x-y plane and along the z axis respectively. The SIP3D variable is the impact parameter significance, $\frac{\sigma_{ip}}{ip}$, where the impact parameter is the closest distance of the lepton track to the primary vertex in 3 dimensions.

Good Muon Requirements		
Quantity	Tight Requirement	Loose Requirement
Quality Muon	Must pass	Must pass
Fraction of valid tracker hits	> 0.8	> 0.8
d_0 (w.r.t. 1st good PV)	< 0.05 cm	< 0.05 cm
d_z (w.r.t. 1st good PV)	< 0.1 cm	< 0.1 cm
SIP3D	< 8	< 8
miniRelIso	< 0.20	< 0.40

The criteria to be a "Quality Muon" is given below:

Table 3.4: Summary of Quality Muon requirements. The segment compatibility is an internal CMS variable which characterizes how well the global muon track matches track segments in the muon system. The global track χ^2 tests the fit of the muon track fits the hits in the tracker and muon system. The χ^2_{kink} represents the probability that the muon was the result of a decay in flight.

Quality Muon Requirements	
Quantity	Requirement
Segment compatibility	> 0.451
or	
Normalized global-track χ^2	< 3
Tracker-Standalone position match	< 12
Kink finder	< 20
Segment compatibility	> 0.303

3.3.5 Photon Selection

Photons are used in this analysis as part of the Z+Hadronic background prediction. The details of this prediction are in 3.5.2. The bulk of these selections are in order to ensure there is no inefficiency for the trigger for γ+jets events.

- $p_{\mathrm{T}} > 25$ GeV
- $|\eta| < 2.4$
- No matching pixel track (pixel veto)
- There must be a jet candidate of $p_{\mathrm{T}} > 10$ GeV matched to the photon within $\Delta R < 0.3$. The matched jet is required to have a neutral electromagnetic energy fraction of at least 70%.
- We reject photons which have an electron of at least $p_{\mathrm{T}} > 10$ GeV within $\Delta R < 0.2$ in order to reject conversions from electrons from W decays which are accompanied by real $E_{\mathrm{T}}^{\mathrm{miss}}$.
- We reject photons which are aligned with the $E_{\mathrm{T}}^{\mathrm{miss}}$ to within 0.4 radians in phi.
- To ensure full efficiency with respect to the isolated photon triggers used, we apply the following additional cuts:
  - ratio of the energy of the energy of the photon's 3X3 supercluster to the photons 5X5 supercluster (R9) > 0.92
  - $\frac{H}{E} < 0.2$
  - hollow track isolation < 3 GeV
  - photons with $|\eta| < 1.4$ have ECAL pfcluster iso < 3 GeV + p_{T} (γ) * 0.0053
  - photons with $|\eta| < 1.4$ have HCAL pfcluster iso < 7 GeV + p_{T} (γ) * 0.014
  - photons with $|\eta| > 1.6$ have ECAL pfcluster iso < 3 GeV + p_{T} (γ) * 0.0034
  - photons with $|\eta| > 1.6$ have HCAL pfcluster iso < 7 GeV + p_{T} (γ) * 0.0139

3.3.6 Jet Selection

Jets are selected from the particle flow selection and are refined via "charged hadron subtraction" and the Summer16_23Sep2016V3 "jet energy corrections" described in 2.3.7. In this analysis, we distinguish between jets and "b-tagged jets." B-tagging is described in 2.3.10, but it comes down to a numerical quantity assigned to each jet called its *b-tag csv value*. If the csv value is large enough, a jet is considered likely to have been seeded by a bottom quark.

Table 3.5: Summary Of Good Jet Requirements. The photon veto and cut on the fraction of energy found in the ECAL without an associated track (neutral EM fraction) attempts to veto jets that are truly photons. The lepton veto and cut on the fraction of energy from the ECAL with an associated track (charged hadron and charged EM fraction) attempt to veto jets that are truly leptons. The cuts on neutral hadron fraction and number of constituents ensure the jets are really sprays of hadronic particles. The cut on p_{T} attempts to remove jets from pileup collisions from consideration.

Jet Selections		
Quantity	Cut Value	
$\|\eta\|$	< 2.4	
Lepton Veto	Not within $\|\Delta R\| < 0.4$ of a tight lepton	
Photon Veto	Not within $\|\Delta R\| < 0.4$ of a good photon	
Neutral Hadron Fraction	< 0.99	
Neutral EM fraction	< 0.99	
Charged hadron fraction	> 0	
Charged multiplicity	> 0	
Charged EM fraction	$< .99$	
Number of constituents	> 1	
Quantity	Jet Requirement	B-tagged Jet Requirement
p_{T}	> 35 GeV	> 25 GeV
B-tag CSV	None	> 0.8484

3.3.7 Isolated Tracks

In addition to vetoing events which have extra loose leptons, a typically more inclusive veto is applied to kill events which might have an extra prompt lepton based on the identification of an isolated charged object.

To distinguish these types of objects, we introduce the notion of "track isolation," defined as the total energy of all the particle flow charged candidates that trace back to within $|dz| < 0.1$ cm of the primary vertex within a cone of $|\Delta R| < 0.3$ about the lepton.[5]

There are two types of isolated tracks, those that identified by particle flow as charged hadron candidates, and those that are identified by particle flow as electron or muon candidates. The criteria on top of the particle flow identification for these two sources are listed in the tables below:

Table 3.6: Summary Of Isolated Track Requirements.

Isolated Track Selections		
Quantity	Charged Hadronic Requirement	Light Lepton Requirement
p_{T}	> 10 GeV	> 5 GeV
Vertex Association[6]	PVTight or PVUsedInFit	PVTight or PVUsedInFit
Track Isolation	< 8 GeV	< 8 GeV
Track Isolation / p_{T}	< 0.2	< 0.1
$\lvert dz \rvert$	< 0.1 cm	< 0.1 cm

3.4 Event Selection

The selection of events can be broken into several stages. Though each of these regions will be expanded upon in the following sections, in broad strokes there 4 separate final states used to accomplish this analysis. These final states define the signal regions and 3 control regions used to conduct this search:

Search Regions (SR) – Events with two opposite charge and same flavor light leptons build our search regions, we have either a e^+e^- or a $\mu^+\mu^-$ pair in each event.

[5]Notice that the miniRelIso we defined above was defined as a relative isolation, the energy in a cone divided by the p_{T} of the object in question, track isolation as defined here has units of energy.

[6]PVTight means the particle was closer to the primary vertex than any other reconstructed vertex. PVUsedInFit means the particle was used to define the primary vertex.

Flavor Symmetric Control Regions – Events with two opposite charge and different flavor light leptons build our flavor symmetric control region, we have either a $e^+\mu^-$ or a μ^+e^- pair in each event. This region is used to predict the flavor symmetric background.

γ+Jets Control Region – Events with a single photon are used to construct the γ+jets control region used for the Z+jets background prediction.

EWK Contamination Closure Region – Events with a photon and muon are used to check the modeling of the "EWK contamination" in the $E_{\mathrm{T}}^{\mathrm{miss}}$ Templates prediction, described in section 3.5.2.

The CMS datasets which seed these events are described in 2.5.

3.4.1 Dilepton Selection

The following are the requirements for events with two light leptons, i.e. for the search regions and for the flavor symmetric control regions.

Light lepton candidates are tagged by the particle flow algorithm described in section 2.3.3. In addition to the list below, leptons must also pass preselection ID and isolation requirements which differ depending on their flavor, but are described in 3.3.3 and 3.3.4 above.

- The (sub)leading lepton in each event must have at least (20) 25 GeV of transverse momentum. These points were selected so that the event would be in the "trigger turn-on" described in 2.2.6.
- The pseudorapidity for each lepton must be within the inner tracker's fiducial area, namely $|\eta| < 2.4$
- No lepton should be in the "dead zone", described in section 2.2.2, this is the range $1.4 < |\eta| < 1.6$

- For the search region, events must have exactly one pair of opposite charge same flavor (OCSF) light leptons, for the flavor symmetric control region events must have exactly one pair of opposite charge and different flavor (OCDF) light leptons.
- The mass constructed out of the sum of lepton vectors (dilepton mass) must be between 86 and 96 GeV.
- The transverse momentum of the dilepton system must be at least 25 GeV. This is in order to ensure parity with the γ selections used in the Z+hadronic background prediction.

3.4.2 Event Vetos

Events are typically vetoed[7] across all search, control, and closure regions if any of the following are true:

Isolated Track Veto – Events with an isolated track, defined in 3.3.7.

$E_{\mathrm{T}}^{\mathbf{miss}}$ Filters – Events where any of the "$E_{\mathrm{T}}^{\mathrm{miss}}$ Filters", described in section 2.3.9, are true.

Jet/$E_{\mathrm{T}}^{\mathbf{miss}}$ alignment – Events where the $E_{\mathrm{T}}^{\mathrm{miss}}$ is aligned with the leading or subleading jet in the ϕ direction up to 0.4. Events with $|\Delta\phi(\mathrm{jet}_{1,2}, E_{\mathrm{T}}^{\mathrm{miss}})| < 0.4$ are not accepted.

Dilepton events specifically, those within the SRs and flavor symmetric control regions, are vetoed under the following conditions:

Lepton Cone Isolation – The two tight leptons are within a cone of $\Delta R < 0.1$ of each other.

Extra Lepton Veto – There is an additional loose or tighter lepton in the event. In other words, event with three or more loose IDed leptons are vetoed. IDs are defined in 3.3.3 and 3.3.4.

[7] precise definitions of the regions will be given below

3.4.3 Search Regions

The search is broken into two distinct branches based on the motivating models listed in sec 3.1.1. The strong and electroweak search regions are summarized in the table below:

Table 3.7: Summary of signal region selections.

Search selections					
Strong Search Regions					
Region	N_{jets}	$N_{\text{b-tags}}$	H_T	$M_{T2}(\ell\ell)$	E_T^{miss} binning
SRA b-veto	2–3	= 0	$>$ 500 GeV	$>$ 80 GeV	[100,150,250,∞]
SRB b-veto	4–5	= 0	$>$ 500 GeV	$>$ 80 GeV	[100,150,250,∞]
SRC b-veto	≥ 6	= 0	-	$>$ 80 GeV	[100,150,∞]
SRA b-tag (SRAb)	2–3	≥ 1	$>$ 200 GeV	$>$ 100 GeV	[100,150,250,∞]
SRB b-tag (SRBb)	4–5	≥ 1	$>$ 200 GeV	$>$ 100 GeV	[100,150,250,∞]
SRC b-tag (SRCb)	≥ 6	≥ 1	-	$>$ 100 GeV	[100,150,∞]
Baseline	≥ 2	-	-	$>$ 80 GeV	$>$ 100 GeV
Electroweak Search Regions					
Region	N_{jets}	$N_{\text{b-tags}}$	dijet mass	M_{T2}	E_T^{miss} binning
VZ	≥ 2	= 0	$M_{jj} < 110$ GeV	$M_{T2}(\ell\ell) > 80$ GeV	[100,150,250,350,∞]
HZ	≥ 2	= 2	$M_{b\bar{b}} < 150$ GeV	$M_{T2}(\ell b \ell b) > 200$ GeV	[100,150,250,∞]

Table 3.8: Additional selections for the search regions.

Cut	Value
Leptons	$e^{\pm}e^{\mp}$ or $\mu^{\pm}\mu^{\mp}$ with $p_T > 25(20)$ GeV for the (sub)leading lepton
Dilepton Selections	All outlined in sec 3.4.1
Vetoes	Extra lepton
	Isolated Track
	Lepton Cone
	E_T^{miss} Filters
	Jet/E_T^{miss} Alignment

In addition to these selections, all dilepton selections, outlined above in sec 3.4.1, and extra lepton vetoes, outlined above in sec 3.4.2, are applied.

3.4.4 Flavor Symmetric Control Region

A flavor symmetric control region in constructed for every search region by flipping the requirement for an OCSF dilepton pair to a OCDF dilepton pair and widening the dilepton mass window from ± 5 GeV about the Z pole mass to > 20 GeV. Table 3.10 summarizes below:

Table 3.9: Summary of selections for the flavor symmetric control regions.

Cut	Value
Baseline	Start with search regions from tables 3.7 and 3.8
Flavor Selection	Flip same flavor requirement, i.e. select $e^{\pm}\mu^{\mp}$ pair
Dilepton Mass	Extend range to > 20 GeV

3.4.5 γ+Jets Control Region

In order to do the Z+jets background prediction, γ+jets events must be chosen in kinematic regions that mimic the search regions. The photon events must pass precisely the cuts for the signal region, listed in sec 3.4.3, to be used in constructing the prediction for that region, the details of which are described in sec 3.5.2. In addition to these event level selections, the photons themselves must pass the criteria in "Photon Selection", sec 3.3.5.

One caveat is the MT2 variable. In order to construct these in events without leptons, we simulate the decay of a Z boson, with momentum equal to that of the leading photon in the event, to two leptons. Those leptons are then used to construct the MT2 variable for each photon event. In the case of MT2b, all photon events are simply allowed to pass the cut. In future instances of this analysis, MT2b could also use the simulated leptons.

3.4.6 EWK Subtraction Closure Region

The EWK Subtraction Closure Region is constructed to check the performance of physics simulation at predicting the $E_{\mathrm{T}}^{\mathrm{miss}}$ profile in Wγ events. The region is defined in table 3.11.

Table 3.10: Summary of selections for the flavor symmetric control regions.

Cut	Value
Baseline	Start with search regions from tables 3.7 and 3.8
Flavor Selection	Flip same flavor requirement, i.e. select $e^{\pm}\mu^{\mp}$ pair
Dilepton Mass	Extend range to > 20 GeV

Table 3.11: Definition of the Electroweak Subtraction Closure Region

Cut	Value
μ	Exactly one with $p_{\mathrm{T}} > 25$ GeV
γ	At least one with $p_{\mathrm{T}} > 25$ GeV
$E_{\mathrm{T}}^{\mathrm{miss}}$	> 50 GeV
N_{jets}	≥ 2
$M_T(\mu, E_{\mathrm{T}}^{\mathrm{miss}})$	> 30 GeV
$\Delta\phi(\gamma, E_{\mathrm{T}}^{\mathrm{miss}})$	> 0.4
Vetoes	$E_{\mathrm{T}}^{\mathrm{miss}}$ filters

Here $M_T(\mu, E_{\mathrm{T}}^{\mathrm{miss}})$ is the transverse mass[8] of the μ+$E_{\mathrm{T}}^{\mathrm{miss}}$ system. $\Delta\phi(\gamma, E_{\mathrm{T}}^{\mathrm{miss}})$ is the angle between the photon and the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector in the x-y plane.

3.4.7 $R_{SF/DF}$ Measurement Region

The $R_{SF/DF}$ measurement region is defined in table 3.12

Table 3.12: Definition of the $R_{SF/DF}$ measurement region.

Cut	Value
Leptons	Exactly 2 passing selections in sec. 3.4.1 on dilepton selection
$E_{\mathrm{T}}^{\mathrm{miss}}$	$\in [100, 150]$ GeV
N_{jets}	≥ 2
$M_{\ell\ell}$	$\notin [70, 110]$ GeV
Vetoes	$E_{\mathrm{T}}^{\mathrm{miss}}$ filters
	Jet/$E_{\mathrm{T}}^{\mathrm{miss}}$ alignment

[8] The mass of a 4-vector with the z-component set to 0.

3.4.8 κ Measurement Regions

The measurement of the κ transfer factor is done in several regions. The baseline region, strong regions, HZ, and VZ, regions are defined in table 3.7, for the measurement of κ the same-flavor requirement is flipped to require $e\mu$ pairs in each event. The "strong region with bs" is the sum of all strong search regions that have a b-tag, i.e. SRAb, SRBb, and SRCb, and "strong region, b-veto" is the sum of strong search regions without a b-tag.

3.4.9 Z+ν Control Regions

These regions are used to normalize the MC samples used in the Z+ν background prediction. Three control regions are constructed for the three different processes. For all these regions, leptons must pass the normal analysis selections listed in sec 3.4.1 on dilepton selections and in secs 3.3.3 and 3.3.4 on electron and muon ID and isolation respectively.

Table 3.13: Definition of the WZ, ZZ, and TTZ control regions. These regions are used to normalize the MC in the Z+ν prediction.

Cut	WZ Value	ZZ Value	TTZ Value
Leptons	Exactly 3 with $p_{\mathrm{T}} > 25, 20, 20$ GeV	Exactly 4 with $p_{\mathrm{T}} > 25, 20, 20, 20$ GeV, second pair must have $\mathrm{M}_{\ell\ell} > 20$ GeV	Exactly 3 with $p_{\mathrm{T}} > 25, 20, 20$ GeV
$N_{\mathrm{b-tags}}$	0	-	Exactly two with $p_{\mathrm{T}} > 25$ GeV
$E_{\mathrm{T}}^{\mathrm{miss}}$	< 60 GeV	-	> 30 GeV
N_{jets}	≥ 2	≥ 2	≥ 2
Vetoes	$E_{\mathrm{T}}^{\mathrm{miss}}$ filters	$E_{\mathrm{T}}^{\mathrm{miss}}$ filters	$E_{\mathrm{T}}^{\mathrm{miss}}$ filters
	Jet/$E_{\mathrm{T}}^{\mathrm{miss}}$ alignment	Jet/$E_{\mathrm{T}}^{\mathrm{miss}}$ alignment	Jet/$E_{\mathrm{T}}^{\mathrm{miss}}$ alignment

3.5 Background Estimation Methods

As discussed in 3.2.1, after the $\mathrm{M}_{\ell\ell}$ and jet selections, the main backgrounds for this search can be broken into three categories, each with their own prediction methods. For any lepton which comes from a flavor symmetric process (i.e. the path from colored particles to an analysis lepton went through a W boson), there should be a kinematically equivalent event in the

same region of phase space but with the same flavor requirement flipped to requiring different flavor leptons. For any event where the dilepton pair comes from a Z boson, e.g. Drell-Yan or WZ (with W$\rightarrow q\bar{q}$) production, all of the $E_{\mathrm{T}}^{\mathrm{miss}}$ in the event will be due to mismeasurements as there are no neutrinos in this population. As will be discussed, the $E_{\mathrm{T}}^{\mathrm{miss}}$ profile in events with a Z and jets should be well modeled by the $E_{\mathrm{T}}^{\mathrm{miss}}$ profile in events with a single photon and an equal number of jets. To predict the count of these types of events we extrapolate from γ+jets events with a few corrections examined below in 3.5.2. Finally, the last possible source of dilepton pairs are events where the pair comes from a Z decay, but a neutrino is produced so that there is genuine $E_{\mathrm{T}}^{\mathrm{miss}}$, e.g. ZZ (with one Z$\rightarrow \nu\nu$) or WZ (with W$\rightarrow l\nu$ and a lost lepton). This final source typically has much lower cross sections due to the fact that there must be at least two electroweak bosons produced to have a dilepton pair and a neutrino and often times a lepton must be lost due to the strong extra lepton vetos we impose.

3.5.1 Data-Driven Predictions

Every background prediction we make is based on countless assumptions.[9] Our assumptions come in two forms:

1. Assumptions grounded in established physics, e.g. the kinematics of same flavor dilepton pairs should be drawn from the same distributions as the kinematics of opposite flavor pairs if the leptons came from W bosons.

2. Assumptions about our tools, e.g. our simulated data is actually representative of the physics it's trying to simulate.

The point of our endeavor in science is to leverage, test, and propose new assumptions of the first kind, while mitigating the effects of assumptions of the second kind. To that end, I'd like

[9] I apologize in advance for getting philosophical in this section.

to draw some focus to the notion of a "data-driven" prediction, which are used for the majority of the background methods in this analysis.

We bin in $E_{\mathrm{T}}^{\mathrm{miss}}$, the only signature we expect from a theoretical dark particle. As described in sec. 2.3.8, $E_{\mathrm{T}}^{\mathrm{miss}}$ comes in two forms, genuine and artificial. Every event has contributions from both of these sources. However, Z+jets events are dominated by artificial $E_{\mathrm{T}}^{\mathrm{miss}}$, Z+$\nu$ should be dominated by genuine $E_{\mathrm{T}}^{\mathrm{miss}}$, and the flavor symmetric background is somewhere in between since the primary contributor to that background has both many jets and multiple νs.

Artificial $E_{\mathrm{T}}^{\mathrm{miss}}$ has been a historically difficult thing to simulate correctly. Although all simulated events are run through a full GEANT simulation of the CMS detector, the response of the detector changes slightly over time and less systematic issues often arise which can not be modeled, e.g. issues with the beam can cause higher rates of pile-up than expected. Due to these issues, it is more optimal to have a data-driven prediction method for sources that might be dominated by artificial $E_{\mathrm{T}}^{\mathrm{miss}}$. In other words, we'd like to have prediction methods that don't rely on detector simulation, even better if they don't rely on physics simulation.

3.5.2 Z + Hadronic

The Z+Hadronic prediction is based on the observation that the $E_{\mathrm{T}}^{\mathrm{miss}}$ profile in events with only hadronic activity (jets) accompanied by a dilepton pair coming from a Z boson will have $E_{\mathrm{T}}^{\mathrm{miss}}$ almost entirely due to mismeasurements of the jets. Figure 3.8 represents the central idea behind the method.

In an event with only a leptonically decaying Z and jets, the $E_{\mathrm{T}}^{\mathrm{miss}}$ should come almost entirely from jet energy mismeasurements. If each jet in an event has a true amount of energy $\hat{E}_{x,y,z}$ in the x, y, or z direction, and a measured amount of energy $E_{x,y,z}$, the $E_{\mathrm{T}}^{\mathrm{miss}}$ in the event will be:

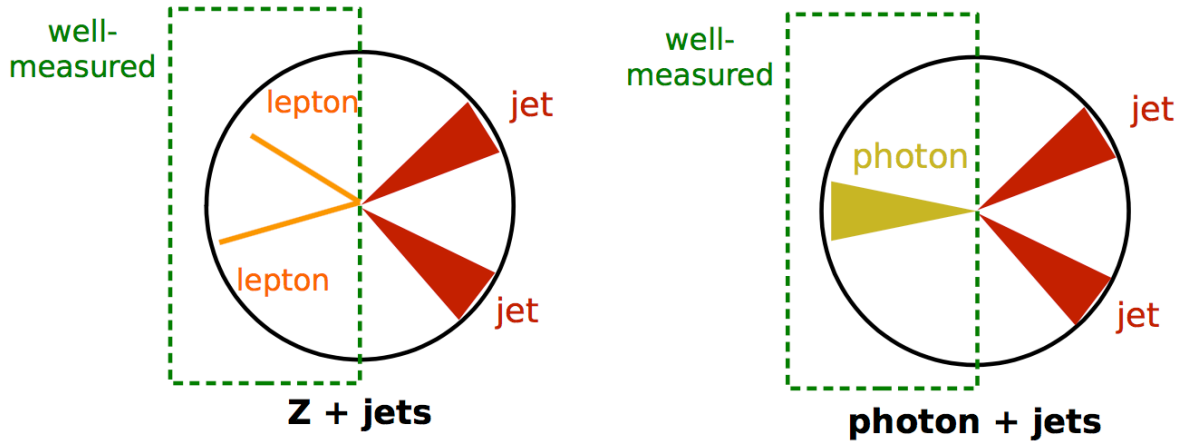


Figure 3.8: The central idea behind the Z+Hadronic background prediction. A well measured photon acts as a proxy for a well-measured Z boson (reconstructed by its leptonic decay products), then the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution for the events is a function of configuration and energy of the jets in the event.

$$E_{\mathrm{T}}^{\mathrm{miss}} = \sqrt{\left(\sum_{\mathrm{jets}} \hat{E}_x - E_x\right)^2 + \left(\sum_{\mathrm{jets}} \hat{E}_y - E_y\right)^2} \tag{3.1}$$

Given the probabilistic nature of mismeasurements, each terms $\hat{E}_{x,y} - E_{x,y}$ in the sums of eq 3.1 can be modeled as a random variable sampled from a Gaussian. The central value of that Gaussian, i.e. how much each jet is mismeasured, is to first order well correlated with the true jet energy which is in turn well correlated with the measured jet energy as explained in sec 2.3.7. Further, if jets tend to clump up in a single hemisphere of the detector, the error in over(under)measurements add, while if they are spread about the detector geometry, those errors will tend to cancel.

Therefore, the number of jets, their relative orientations, and their absolute energy scale will change the $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum. The upshot of these points is that we should not expect that the $E_{\mathrm{T}}^{\mathrm{miss}}$ profile in γ+jets events to be precisely the same as in Z+jets events unless these parameters,

at least, are nearly identical between the populations.

There is no ideal way to correct for all differences. When this method was incepted[77], the original author constructed fine-grained bins based on the number of jets within particular p_{T} windows. In our updated method, we found that simply correcting the p_{T} distribution of the γ+jets events in each signal region yields a $E_{\mathrm{T}}^{\mathrm{miss}}$ profile with good enough agreement to the Z+jets sample. The ultimate test of this method comes from the "closure test" outlined in section 3.5.2. The correction of the p_{T} distribution goes as follows:

1. Generate a sample of 'clean'[10] γ+jets events which pass roughly the same selections as the dilepton events, but with the requirement for two leptons replaced by the requirement for one photon.[11]

2. Take the clean photon sample and bin events by p_{T} in the intervals between (GeV): [25, 33, 40, 55, 85, 105, 135, 180, 250, ∞]. These are chosen to be just wide enough that there a high number of photon events in each bin, but fine enough to capture the p_{T} dependence of the jet configurations.

3. Generate a sample of 'clean' Z+jets events

4. Normalize the area of the clean Z+jets p_{T} distribution and clean γ+jets to unity

5. Assign a weight to each photon event such that the bin has the same area as the corresponding Z+jets p_{T} bin, use these weights when constructing any other distribution in the γ+jets sample, for instance the $E_{\mathrm{T}}^{\mathrm{miss}}$.

When we have a γ+jets sample that has been p_{T} reweighted to match the Z+jets sample, we can then assume the $E_{\mathrm{T}}^{\mathrm{miss}}$ shape in the γ+jets events models the $E_{\mathrm{T}}^{\mathrm{miss}}$ shape in the Z+jets sample to some degree of accuracy we will test in the next section. However, the overall number

[10] Generating a clean sample of Z+jets events and γ+jets events requires some work for the data samples, this will be discussed in the following sections.

[11] precise definition in sec 3.4.5

γ+jets events is not related to the number of Z+jets events in a simple way. To actually extract a prediction from the γ+jets $E_{\mathrm{T}}^{\mathrm{miss}}$ shape, we normalize the γ+jets $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution to the Z+jets distribution in the $E_{\mathrm{T}}^{\mathrm{miss}}$ 50-100 bin. This bin is chosen because we expect any new physics to be at higher $E_{\mathrm{T}}^{\mathrm{miss}}$, and the cut on MT2 in the strong signal regions ends up depleting the population of Z+jets events for $E_{\mathrm{T}}^{\mathrm{miss}}$ below 50 GeV. The upshot is that normalizing to a low $E_{\mathrm{T}}^{\mathrm{miss}}$ bin provides us with a way to essentially correct the γ+jets cross section to the Drell-Yan cross section. When the cross section and p_{T} shape have been corrected, the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution of the γ+jets events constitutes our Z+jets background prediction.

p_{T} Reweighting Closure Test

In this section we show the results of the closure test meant to check how well the p_{T} reweighting procedure mitigates the differences in jet kinematics and configuration for the Z+hadronic background prediction. For this test, Z+jets MC is compared to γ+jets MC in each signal region. The procedure enumerated in the previous section is applied to both samples and then the $E_{\mathrm{T}}^{\mathrm{miss}}$ shape is compared.

We use this test to set the expected fluctuation, or systematic uncertainty, for this method. A percent uncertainty is extracted for the 100-150 $E_{\mathrm{T}}^{\mathrm{miss}}$ bin and bins above 150 GeV by choosing the larger of the following:

1. The deviation of ratio between the Z+jets prediction and γ+jets prediction from unity
2. The statistical uncertainty on the ratio

Figure 3.9 and Table 3.14 summarizes the results.

Table 3.14: Numerical representation of the data in figure 3.9.

SR	sample	100.00-150.00	150.00+
SRA	Z Jets	18.89±0.69	1.55±0.24
	Photon Jets	16.26±0.53	1.38±0.16
	Ratio	1.16±0.06	1.12±0.21
SRAb	Z Jets	10.70±1.07	0.52±0.08
	Photon Jets	8.94±0.60	0.52±0.11
	Ratio	1.20±0.14	1.01±0.26
SRB	Z Jets	15.93±0.69	1.51±0.18
	Photon Jets	14.19±0.46	1.48±0.13
	Ratio	1.12±0.06	1.02±0.15
SRBb	Z Jets	6.13±0.59	1.05±0.11
	Photon Jets	5.84±0.39	1.06±0.16
	Ratio	1.05±0.12	0.99±0.18
SRC	Z Jets	4.12±0.43	0.39±0.07
	Photon Jets	3.58±0.26	0.35±0.07
	Ratio	1.15±0.15	1.12±0.29
SRCb	Z Jets	1.41±0.17	0.28±0.06
	Photon Jets	1.19±0.14	0.28±0.06
	Ratio	1.19±0.20	1.01±0.31
VZ	Z Jets	39.20±3.06	1.63±0.40
	Photon Jets	36.59±2.28	1.86±0.25
	Ratio	1.07±0.11	0.88±0.24
HZ	Z Jets	3.85±1.18	0.19±0.05
	Photon Jets	2.13±0.24	0.19±0.05
	Ratio	1.80±0.59	0.98±0.34

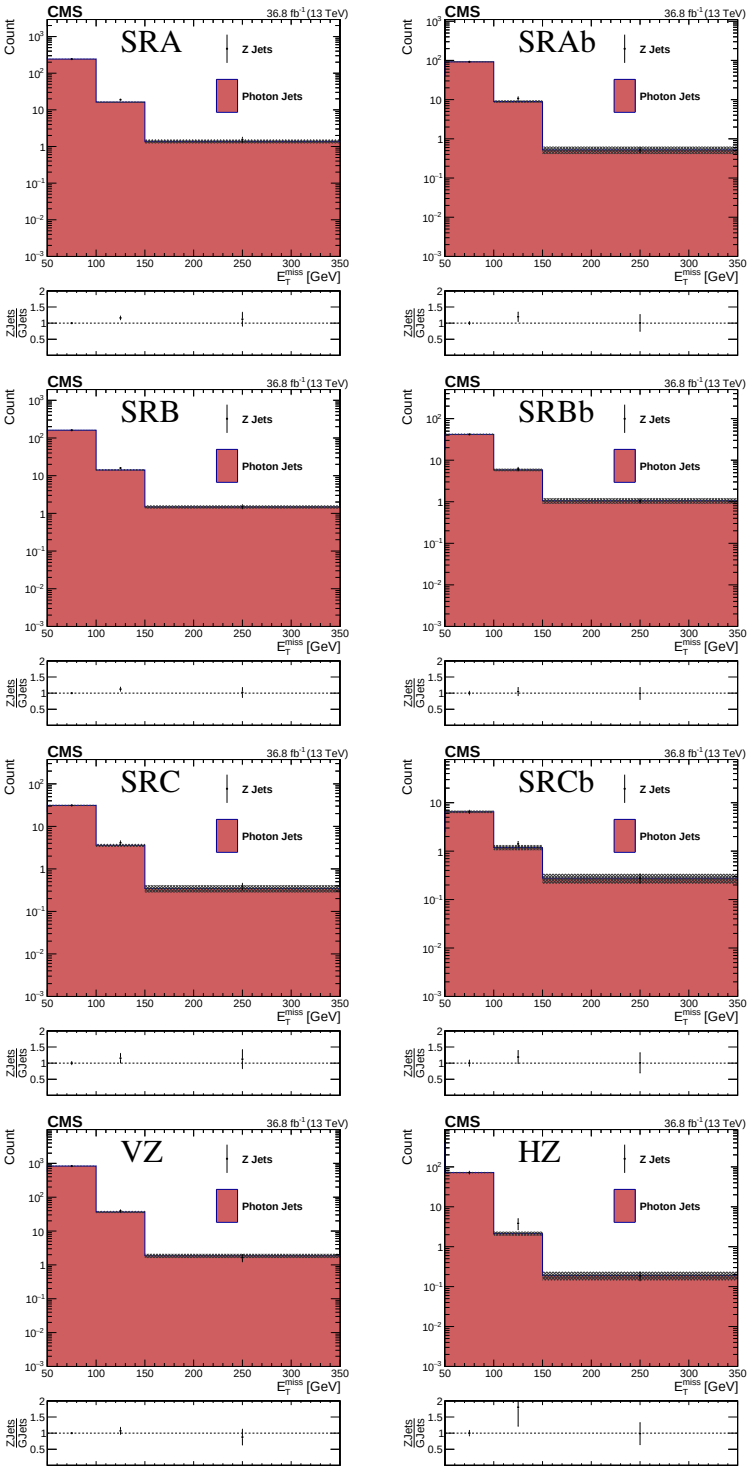


Figure 3.9: The results of the closure test to assess the efficacy of the p_{T} reweighting for the Z+Hadronic background prediction. In black, the Z+jets events and in red, the p_{T} reweighted γ+jets events. Notice that the counts in the 50-100 $E_{\mathrm{T}}^{\mathrm{miss}}$ bin are identical by construction as the γ+jets events are normalized in this bin. This data is summarized in table 3.14.

From the table, most of the uncertainties for the strong search regions are less than 30%. There is one bin with large uncertainty in the electroweak HZ region, but this is somewhat due to the rarity of Z+jets events in that region, as can be seen by the comparatively large uncertainty on the ratio in that bin as well.

Contamination From Events With Real $E_{\mathrm{T}}^{\mathrm{miss}}$ (Electroweak Subtraction)

In the previous sections, we discussed the principles behind the Z+hadronic background methods and quantified the uncertainty inherent in this prediction due to the fundamental issue of different jet configurations and energy profiles between the two populations. Now we turn to the issue of creating a clean sample of γ+jets events.

The baseline selection of γ+jets data is detailed in sec 3.4.5, our initial γ+jets sample is all events which pass those cuts. This method hinges on that fact that our sample of photon events have non-zero $E_{\mathrm{T}}^{\mathrm{miss}}$ only due to energy mismeasurements, i.e. there are no neutrinos in the events. However, the following processes are identified as the non-negligible physics that can lead to a final state which contain photons, jets, and neutrinos:

1. Single W production with γ radiation, where the W$\rightarrow \ell\nu$ and the lepton is lost.
2. Single Z production with γ radiation, where the Z$\rightarrow \nu\nu$.
3. $t\bar{t}$ production with a single top decaying leptonically and a lost lepton.

In order to create a pure sample of γ+jets events, we simulate these processes in MC and pass the generated events through the same selection criteria as the γ+jets events. Before the γ+jets sample is reweighted to match the p_{T} spectrum of the Z+jets sample, the p_{T} profile of these contaminating events are subtracted. In addition, when making the $E_{\mathrm{T}}^{\mathrm{miss}}$ prediction, the $E_{\mathrm{T}}^{\mathrm{miss}}$ of the contaminating events is subtracted from the photon sample.

To estimate the uncertainty associated with this process, we perform a check on the performance of the simulated events at predicting the $E_{\mathrm{T}}^{\mathrm{miss}}$ and p_{T} profiles in a $\gamma+\mu$ control

region fully defined in 3.4.6. This region is mainly focused on checking the MC performance for events with a single W boson that decays to a μ along with γ radiation, this includes item (3) above as the top decays through a W.[12]

The $E_{\mathrm{T}}^{\mathrm{miss}}$ and p_{T} profiles in these events are shown in figure 3.10. From this figure, we can see an envelope of 30% around unity in the ratio encapsulates all of the points in the $E_{\mathrm{T}}^{\mathrm{miss}}$ profile and almost all point in the p_{T} profile. We assign a one-sigma expected fluctuation to this method of 30%. To apply this uncertainty, we count the number of events subtracted from the γ+jets sample in each $E_{\mathrm{T}}^{\mathrm{miss}}$ bin individually and take 30% of that number as the contribution to the uncertainty band associated with this source.

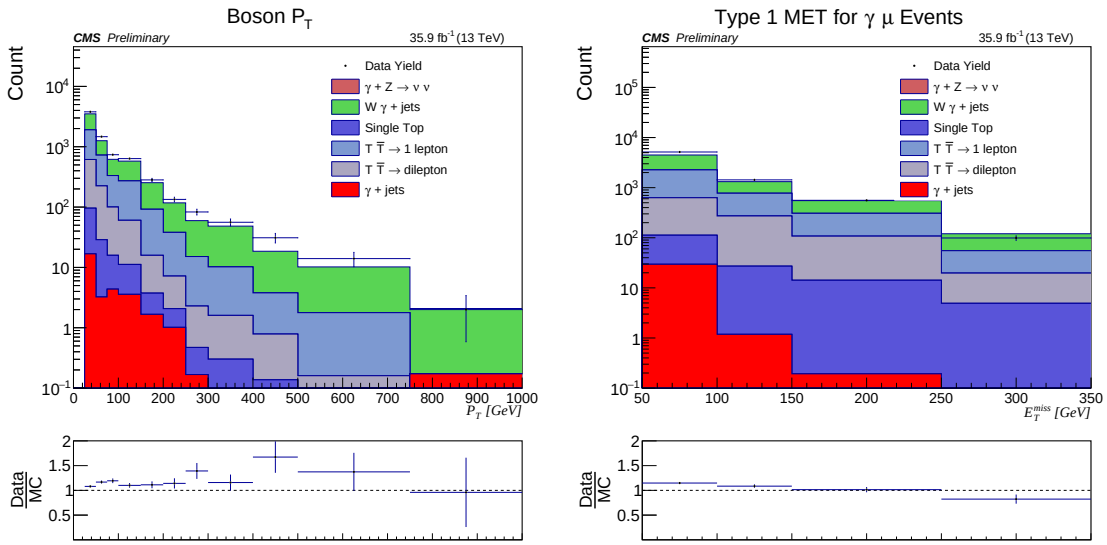


Figure 3.10: The p_{T} (left) and $E_{\mathrm{T}}^{\mathrm{miss}}$ (right) profiles in simulation and data for the $\gamma\mu$ control sample, defined in sec 3.4.6. The agreement between data and simulation is used to extract the expected fluctuation for the process of removing events with neutrinos from the γ+jets population. Given that the ratio is never deviates from unity by more than 30% in the $E_{\mathrm{T}}^{\mathrm{miss}}$ profile, and mostly so in the p_{T}, we use 30% as the systematic uncertainty for this process.

[12]The Z$\rightarrow \nu\nu$ simulation modeling is assumed to have similar performance as the relevant physics, the photon radiation probability and p_{T} profile, should be the same across both samples because the same MC generator, hadronizer, and reconstruction software are used in both cases.

Algorithm For The Prediction

Now that we have addressed the fundamental premise of this method and all the complications, we can give a full algorithm for the prediction:

1. Construct the population of single photon + jets events in the augmented search regions with dilepton quantities replaced by photon selection, full selection in sec 3.4.5.

2. Simulate $\gamma+\nu$ events in MC, create a 'clean' population of only γ+jets events by subtracting the kinematic distributions of the $\gamma+\nu$ sample from any distribution of interest (in this case p_{T}).

3. Take the 'clean' photon sample and bin events by p_{T} in the intervals between (GeV): [25, 33, 40, 55, 85, 105, 135, 180, 250, ∞]. These are chosen to be just wide enough that there a high number of photon events in each bin, but fine enough to capture the p_{T} dependence of the jet configurations.

4. Create a 'clean' Z+jets sample by taking the dilepton data that pass search region criteria, described in sec 3.4.3, and subtracting the other background predictions from any distribution of interest (in this case p_{T}). Note that the dilepton events are not binned in $E_{\mathrm{T}}^{\mathrm{miss}}$ at this point, all other cuts are applied.

5. Normalize the area of the clean Z+jets p_{T} distribution and clean γ+jets to unity.

6. Assign a weight to each photon event such that the bin has the same area as the corresponding Z+jets p_{T} bin.

7. Construct the $E_{\mathrm{T}}^{\mathrm{miss}}$ profile of the cleaned γ+jets sample where each event is given the weight derived in the previous step.

8. Find the ratio of the $E_{\mathrm{T}}^{\mathrm{miss}}$ 50-100 bin for the clean Z+jets sample vs. the clean γ+jets sample, apply this ratio across the entire cleaned γ+jets $E_{\mathrm{T}}^{\mathrm{miss}}$ profile

Systematics

The total uncertainty for this prediction is broken into 4 uncorrelated sources.

$$\sigma^2_{\text{net}} = \sigma^2_{\text{Normalization}} + \sigma^2_{\text{EWK Sub}} + \sigma^2_{\text{Statistical}} + \sigma^2_{p_{\text{T}}\text{ Reweighting}}$$

- **Normalization** – The uncertainty due to the cross section normalization of the γ+jets events to the Z+jets events in the $E_{\text{T}}^{\text{miss}}$ 50-100 bin. This is a percent uncertainty found on a bin-by-bin basis. The percentage is found by dividing the statistical uncertainty by the total count of events in the Z+jets $E_{\text{T}}^{\text{miss}}$ 50-100 bin. If the 50-100 $E_{\text{T}}^{\text{miss}}$ bin has a count of N events, a background prediction of b in some higher $E_{\text{T}}^{\text{miss}}$ bin will get an uncertainty contribution of $\sigma_{\text{Normalization}} = b\frac{\sigma(N)}{N}$ due to this source, where $\sigma(N)$ is the Poisson uncertainty for N events.

- **EWK Sub** – The uncertainty due to the imperfect modeling of the $\gamma+\nu$ events when subtracting them from the γ+jets sample. A 30% uncertainty is associated with the modeling of p_{T} and $E_{\text{T}}^{\text{miss}}$ as was shown in figure 3.10 and the section therein. This 30% is applied to the total count of events subtracted from the $E_{\text{T}}^{\text{miss}}$ bin and scaled by the normalization from the $E_{\text{T}}^{\text{miss}}$ 50-100 bin. If m_s events were subtracted from the γ+jets events in some $E_{\text{T}}^{\text{miss}}$ bin (after p_{T} reweighting), and n is the normalization applied to the γ+jets sample to correct the cross section, the total uncertainty contribution to the $E_{\text{T}}^{\text{miss}}$ bin due to this source is $\sigma_{\text{EWK Sub}} = 0.3 \cdot n \cdot m_s$.

- **Statistical** – The uncertainty due to the limited statistics of the γ+jets events. This is derived as a percent uncertainty on the p_{T} reweighted count of γ+jets events in a $E_{\text{T}}^{\text{miss}}$ bin. If a $E_{\text{T}}^{\text{miss}}$ bin has a background prediction of b which was calculated from a γ+jets $E_{\text{T}}^{\text{miss}}$ bin with statistical uncertainty σ_γ, the total uncertainty contribution to the $E_{\text{T}}^{\text{miss}}$ bin due to this source is $\sigma_{\text{Statistical}} = \sigma_\gamma \cdot b$

- p_{T} **Reweighting** – The uncertainty due to differences in the kinematics of jets in the photon and Z samples. The percent uncertainty for each $E_{\mathrm{T}}^{\mathrm{miss}}$ bin due to this source was summarized in table 3.14. If a $E_{\mathrm{T}}^{\mathrm{miss}}$ bin has a background prediction of b and the percent uncertainty due to the p_{T} reweighting is p, then the total uncertainty contribution to the $E_{\mathrm{T}}^{\mathrm{miss}}$ bin due to this source is $\sigma_{p_{\mathrm{T}}\ \mathrm{Reweighting}} = p \cdot b$.

A summary of all the Z+jets background predictions across all search bins as well as their systematics is shown below in table 3.15.

Table 3.15: A summary of all prediction counts and uncertainties for the Z+jets predictions. The number "ratio" is the particular uncertainty associated for the process divided by the total uncertainty for the prediction in the $E_{\mathrm{T}}^{\mathrm{miss}}$ bin.

SR	MET Bin	Prediction	Closure (ratio)	Normalization (ratio)	Statistical (ratio)	EWK Sub (ratio)
SRA	100-150	13.61 ± 3.14	2.72 (0.87)	1.00 (0.32)	1.14 (0.36)	0.34 (0.11)
	150-250	2.45 ± 0.87	0.64 (0.74)	0.18 (0.21)	0.36 (0.42)	0.42 (0.49)
	250+	3.26 ± 2.36	0.85 (0.36)	0.24 (0.10)	2.16 (0.91)	0.39 (0.16)
SRAb	100-150	8.21 ± 2.10	1.64 (0.78)	0.89 (0.42)	0.92 (0.44)	0.27 (0.13)
	150-250	1.19 ± 0.54	0.31 (0.57)	0.13 (0.24)	0.24 (0.45)	0.35 (0.65)
	250+	0.51 ± 0.27	0.13 (0.48)	0.05 (0.20)	0.15 (0.56)	0.18 (0.64)
SRB	100-150	12.81 ± 2.35	1.54 (0.65)	1.19 (0.51)	1.30 (0.56)	0.16 (0.07)
	150-250	0.89 ± 0.34	0.13 (0.39)	0.08 (0.24)	0.22 (0.66)	0.20 (0.59)
	250+	0.38 ± 0.20	0.06 (0.29)	0.04 (0.18)	0.12 (0.63)	0.14 (0.69)
SRBb	100-150	7.74 ± 3.11	0.93 (0.30)	1.32 (0.42)	2.65 (0.85)	0.23 (0.07)
	150-250	4.04 ± 3.33	0.73 (0.22)	0.69 (0.21)	3.16 (0.95)	0.25 (0.07)
	250+	0.10 ± 0.14	0.02 (0.13)	0.02 (0.12)	0.04 (0.28)	0.14 (0.94)
SRC	100-150	1.24 ± 0.43	0.19 (0.43)	0.27 (0.64)	0.27 (0.63)	0.04 (0.09)
	150+	0.13 ± 0.11	0.04 (0.36)	0.03 (0.27)	0.05 (0.51)	0.08 (0.74)
SRCb	100-150	0.14 ± 0.47	0.03 (0.06)	0.05 (0.10)	0.04 (0.08)	0.47 (0.99)
	150+	0.00 ± 0.33	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.33 (1.00)
VZ	100-150	29.27 ± 4.42	3.22 (0.73)	1.13 (0.26)	2.15 (0.49)	1.81 (0.41)
	150-250	2.87 ± 2.09	0.69 (0.33)	0.11 (0.05)	0.40 (0.19)	1.93 (0.92)
	250-350	1.00 ± 0.73	0.24 (0.33)	0.04 (0.05)	0.24 (0.33)	0.64 (0.88)
	350+	0.29 ± 0.30	0.07 (0.23)	0.01 (0.04)	0.08 (0.26)	0.28 (0.94)
HZ	100-150	2.90 ± 2.39	2.32 (0.97)	0.34 (0.14)	0.41 (0.17)	0.24 (0.10)
	150-250	0.26 ± 0.19	0.09 (0.47)	0.03 (0.16)	0.08 (0.44)	0.14 (0.75)
	250+	0.09 ± 0.07	0.03 (0.45)	0.01 (0.16)	0.05 (0.77)	0.03 (0.42)

3.5.3 Flavor Symmetric Background

The flavor symmetric background accounts for all cases where even a single lepton in the final dilepton pair was produced through the decay of a W boson. The basic idea is that the W has almost exactly the same decay rate to electrons and muons, therefore any event where a lepton was produced in a W decay should be mirrored by another event where the lepton was produced with the other flavor. In other words, if a physics process where a dilepton pair is produced and *at least* one lepton came from a W boson occurs at some frequency f, a process where the dilepton pair has different flavor should occur with the same rate, f. The same argument holds for any flavor-symmetric process. For instance, ZZ production with two lost leptons, one from each Z, is also a flavor symmetric background because the chance to lose a muon is no different from the chance to lose an electron.[13]

Furthermore, given that the mass of the electron and the mass of the muon are many orders of magnitude less than the dilepton mass requirement of roughly 90 GeV, the typical energy scales and orientation of light leptons in flavor symmetric processes should be the same regardless of whether the leptons are paired with a same flavor or different flavor partner. This means that we can use the kinematical distributions in events with different flavor lepton pairs to predict the kinematical distributions for events with a same flavor lepton pair.

In a totally idealized scenario, the flavor symmetric background prediction would be extremely simple. First, we would invert the selection for same flavor lepton pairs to require different flavor pairs in each $E_{\mathrm{T}}^{\mathrm{miss}}$ bin of our search regions. Then the count in each $E_{\mathrm{T}}^{\mathrm{miss}}$ for the different flavor pairs could be taken as the expected count for the search bin. However, as with any background prediction method, there are complications to our idealized scenario.

The first is that the *reconstruction efficiency* for electrons and muons is not identical,

[13]This is a rare process and those lost-lepton probabilities are not precisely identical, but the reader should understand that the existence of a W boson is not necessary to make a background process flavor symmetric. It is simply the most likely scenario as $t\bar{t}$ production is the only large cross section flavor-symmetric process in our search space.

meaning that there is a slight asymmetry between how likely the CMS detector and software suite are to find a muon vs. how likely they are to find an electron at some particular energy. This effect is caused by different detection methods used for electrons and muons. It manifests as different trigger efficiencies for electrons and muons at the same p_{T}, as briefly explained in sec 2.2.6, as well the differences in the electron and muon ID requirements outlined in sections 3.3.3 and 3.3.4. We correct for this effect with a factor called $\mathrm{R}_{\mathrm{SF/DF}}$, which is the subject of the next section.

The second is that with few enough events, the symmetry between same-flavor and different-flavor pairs can be thrown off by statistical fluctuations.[14] For this analysis, the expected number of flavor symmetric events in most search bins is less than 2. In order to ensure we have enough statistics for each prediction, the mass window for the acceptance of different flavor events is enlarged.

With all corrections applied, the final number of events predicted in a $E_{\mathrm{T}}^{\mathrm{miss}}$ bin is

$$N_{\mathrm{SF}} = \kappa \cdot \mathrm{R}_{\mathrm{SF/DF}} \cdot N_{\mathrm{DF}}, \tag{3.2}$$

where κ is a corrective factor associated with the extrapolation from the enlarged mass window into the Z mass window, and $\mathrm{R}_{\mathrm{SF/DF}}$ is a corrective factor associated with the asymmetry in electron and muon efficiencies.

The $\mathrm{R}_{\mathrm{SF/DF}}$ Factor

The factor $\mathrm{R}_{\mathrm{SF/DF}}$ in eq. 3.2 is a transfer factor which is meant to scale the number of different-flavor events in the flavor symmetric control regions, essentially the search regions with their same flavor requirement inverted, to the number of same-flavor events in associated search bin. This factor is derived through a direct measurement in the $\mathrm{R}_{\mathrm{SF/DF}}$ control region, defined in sec. 3.4.7 as $E_{\mathrm{T}}^{\mathrm{miss}} \in [100, 150]$, $N_{\mathrm{jets}} = 2$, and $\mathrm{M}_{\ell\ell} \notin [70, 110]$.

[14] Tossing a coin should yield as many heads as tails, however it is just as likely to get HH *or* TT in tossing a coin twice as it is to get one of each. To chance to get an uneven percentage of heads or tails in a series of throws decreases with number of throws. The symmetry has more power with more statistics.

The measurement of this factor is done in both data and $t\bar{t}$ MC, but the value MC is used only as a sanity check on the method. In order to set an uncertainty associated with this method, the $R_{SF/DF}$ control region is binned further by various parameters that are used to construct the search regions. The variation in the measured value across these bins is used as the expected fluctuation of the method. These checks are shown in figure 3.11.

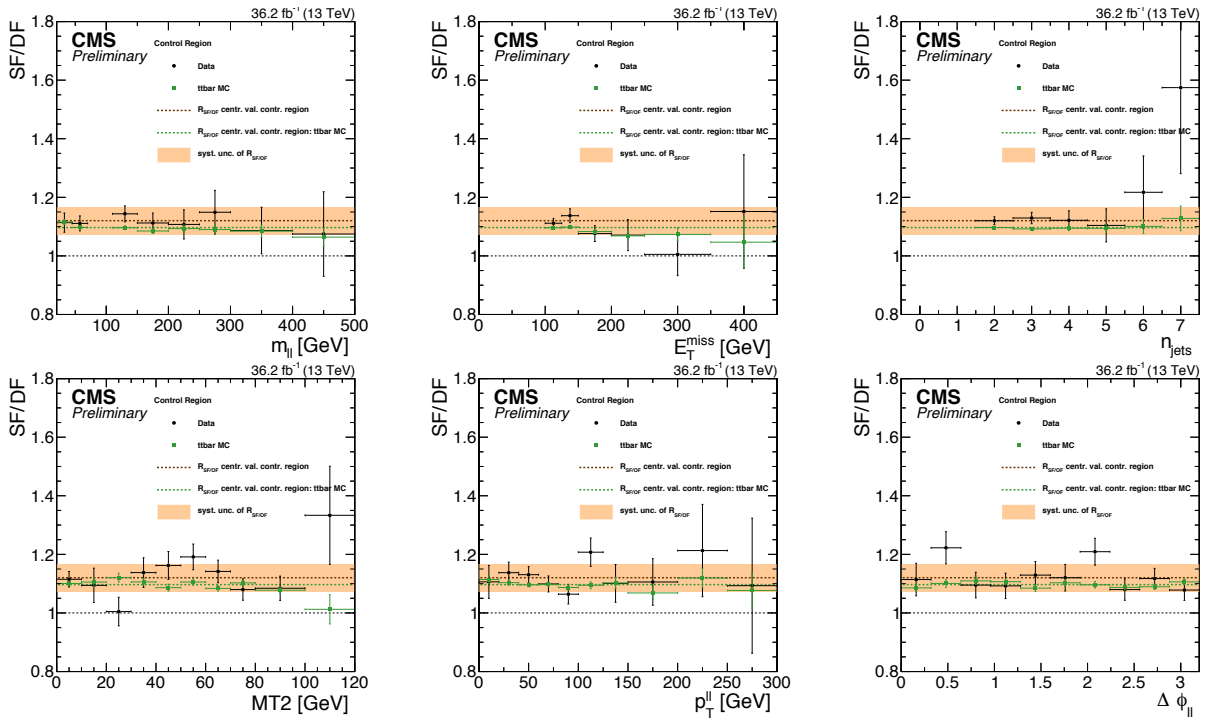


Figure 3.11: The measurement of $R_{SF/DF}$ and its uncertainty. The $R_{SF/DF}$ control region is further binned along the variables shown, and $R_{SF/DF}$ is measured as they vary. The orange band shows the one-σ expected fluctuation set by this method. As can be seen, the majority of data points fall within this orange band. $R_{SF/DF}$ is relatively well behaved as a function of these kinematic and event-level variables.

The final value of $R_{SF/DF}$ measured in data is 1.107 ± 0.046 (which is the value used in the final prediction), and the cross check in MC gives 1.090 ± 0.005.

Enlarging The Mass Window

In order to keep the statistical uncertainty reasonable, we enlarge the flavor symmetric control region's $M_{\ell\ell}$ window from ± 5 GeV about the Z mass to > 20 GeV. Then we must apply a transfer factor, denoted by κ, which gives the ratio of number of events expected in this enlarged window, to the number of events expected in the Z mass window.

$$\kappa = \frac{\text{Num Different Flavor Events in Z Mass Window}}{\text{Num Different Flavor Events with } M_{\ell\ell} > 20 \text{ GeV}}$$

Furthermore, although $M_{\ell\ell}$ was chosen precisely because it is a relatively flat variable for $t\bar{t}$ events near the search regions, there are going to be some kinematic differences introduced by this asymmetry. In order to set an uncertainty associated with this method, we measure κ in several control, search, and amalgamated search regions, defined in sec. 3.4.8. Simulated $t\bar{t}$ events are also used to measure κ in these regions as a sanity check on the method to ensure there is little to no background contamination in the flavor symmetric control region.

The measured value of κ is 0.065 ± 0.02 The κ central value is then chosen to fit the various measured values. The one-σ expected fluctuation is chosen such most values measured in data and all values measured in MC are within the band. Figure 3.12 summarized the measurements of κ across all search and control regions in both data and MC.

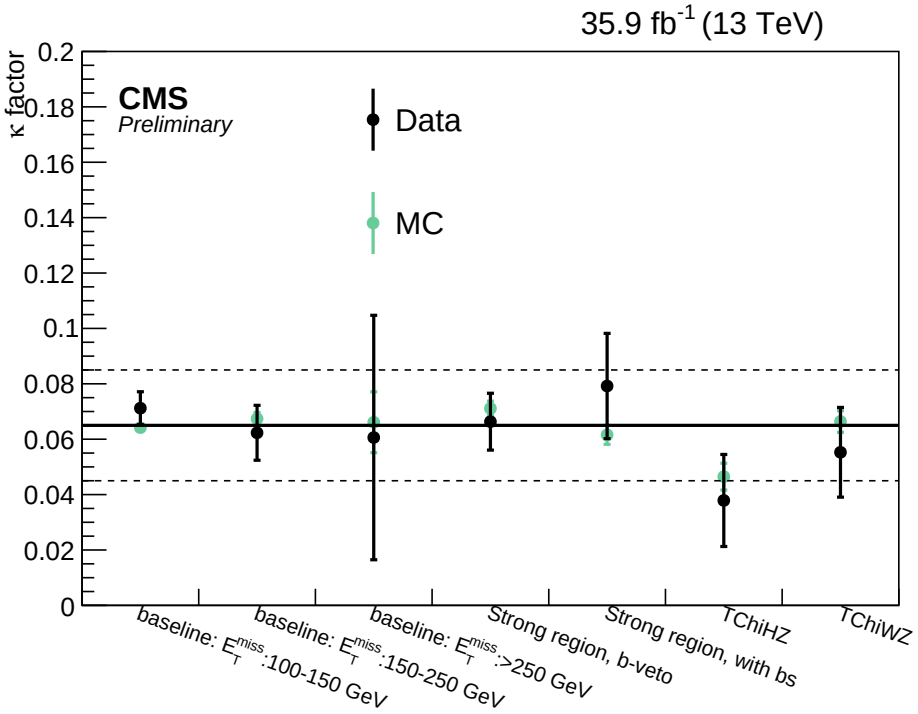


Figure 3.12: The measurement of κ across various search and control regions. The value measured in data and MC agree within uncertainty for all regions. The solid line is the central value chosen for κ and the dashed lines represent the one-σ expected fluctuations.

Systematics

The total uncertainty for this prediction is broken into 3 uncorrelated sources.

$$\sigma^2_{\text{net}} = \sigma^2_{R_{SF/DF}} + \sigma^2_{\kappa} + \sigma^2_{\text{Statistical}}$$

The quantities $\sigma_{R_{SF/DF}}$ and σ_{κ} are derived by taking the difference in the prediction when varying $R_{SF/DF}$ and κ between the two ends of their measured windows reported above. $\sigma_{\text{Statistical}}$ is the Poisson uncertainty on the number of events in the flavor symmetric control region associated the search bin scaled by the product of $R_{SF/DF}$ and κ.

3.5.4 Z + ν

Table 3.16 summarizes the non-negligible background processes considered that can produce the Z+ν signature. The simulated events are only accepted when the dilepton pair can be matched back to a pair that came from a Z boson and there is a neutrino in the final state. This selection ensures there is no double counting background events that should be predicted by the flavor symmetric or Z+hadronic predictions.

Table 3.16: A summary of the Z+ν background sources. These processes are simulated as described in sec 2.4. Only events which have a prompt neutrino and where the pair of selected leptons can be matched to a Z boson are considered to ensure orthogonality with the Z+jets and flavor symmetric background predictions.

Category	Processes
WZ	WZ$\rightarrow 3\ell + \nu$
ZZ	$2\ell + 2\nu$
TTZ	Z$\rightarrow 2\ell$ and Z$\rightarrow 2\nu$
VVV	ZZZ (inclusive decays)
	WZZ (inclusive decays)
	WWZ (inclusive decays)

Normalization Factors

Monte Carlo cross section are notoriously unreliable for low cross-section physics, especially in extreme regions of phase space used for new physics searches. In order to check the cross sections in the simulated events, the ZZ, WZ, and TTZ simulation are compared against data in specific control regions constructed to be sensitive to these processes. The full definitions of these control regions are given in sec 3.4.9. Figure 3.13 summarizes the results of these comparisons.

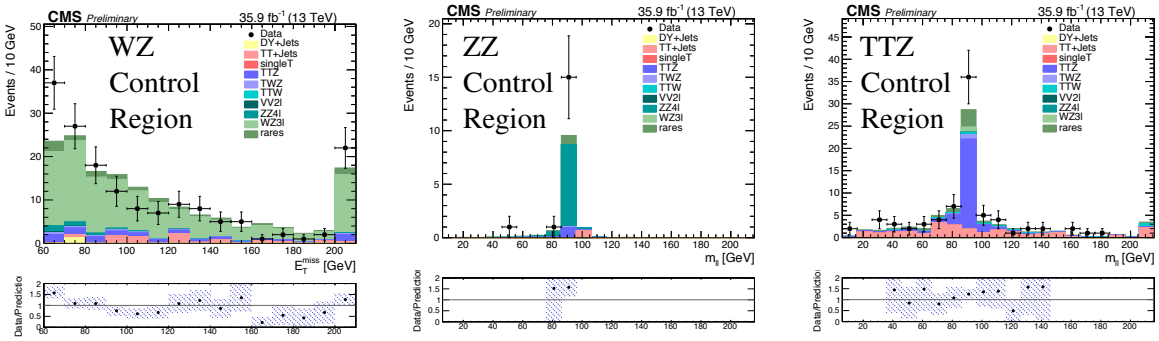


Figure 3.13: Data and MC are compared in WZ, ZZ, and TTZ control regions in order to check whether the cross section in MC is compatible with what we see in data. A normalized is derived from simulation such that the number of events in the control region in data and MC agree. The ratio is extracted in the 86 GeV $> m_{ll} >$ 96 GeV in the ZZ and TTZ regions, and in 60 GeV $> E_{\mathrm{T}}^{\mathrm{miss}} >$ 200 GeV in the WZ region.

Due to the low cross section, difficulty of isolation, and low impact on the results, a rigorous analysis on the performance of simulation for these processes is not undertaken. Rather, a large conservative uncertainty is chosen based on the historical performance of simulating these models. A summary of derived normalization factors and the associated uncertainty in MC modeling is given in table 3.17.

Systematics

The total uncertainty for this prediction is broken into 2 uncorrelated sources.

Table 3.17: A summary of the Z+ν background sources' normalization factors and uncertainties. For the WZ, ZZ, and TTZ processes, normalization factors were found to correct the MC cross section, the VVV processes have low enough cross section that this is neglected. Due to the difficulty of isolating these processes and their low impact on search results, flat uncertainties are taken based on our level of confidence in MC performance.

Sample	Normalization Factor	Percent Uncertainty
WZ	1.06	30%
ZZ	1.71	50%
TTZ	1.36	30%
VVV	1	50%

$$\sigma^2_{\text{net}} = \sigma^2_{\text{MC Performance}} + \sigma^2_{\text{Statistical}}$$

The statistical uncertainty is taken from the MC, the "MC Performance" uncertainty is given in table 3.17 above.

3.6 Results

3.6.1 Strong Search Regions

In this section, we show the search results for the strong signal regions. Figure 3.14 shows the results visually, each search region is shown binned in $E_{\text{T}}^{\text{miss}}$, the background predictions are broken down as described in 3.5 and uncertainty bands are shown as a hashing that captures all sources described above. The number of events in data seen are shown as black dots in the figures. The same data is shown numerically in table 3.18.

In the strong search regions, the data and background predictions match extremely well. Almost all bins are within the expected 1-σ fluctuation, with the exception of the MET tail bins which have some small downward fluctuations, the largest being in SRB.

Table 3.18: Numerical results for the strong search regions are shown. The precise definitions of these regions are shown in table 3.7. See figure 3.14 for the corresponding plots. The agreement in the MET 50-100 bin is by construction as that bin is used to normalize the Z+Hadronic prediction. No statistically significant deviation from the standard model prediction is found.

SRA	$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	50-100	100-150	150-250	250+
	Z+Hadronic	208.5 ± 16.1	13.6 ± 3.1	2.5 ± 0.9	3.3 ± 2.4
	FS	$0.4^{+0.3}_{-0.2}$	$0.4^{+0.3}_{-0.2}$	$0.2^{+0.2}_{-0.1}$	$0.2^{+0.2}_{-0.1}$
	Z+ν	1.1 ± 0.4	0.8 ± 0.3	1.4 ± 0.4	2.4 ± 0.8
	Sum	$210.0^{+16.1}_{-16.1}$	$14.8^{+3.2}_{-3.2}$	$4.0^{+1.0}_{-1.0}$	$5.9^{+2.5}_{-2.5}$
	Data	210	23	5	4
SRAb	$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	50-100	100-150	150-250	250+
	Z+Hadronic	92.2 ± 10.4	8.2 ± 2.1	1.2 ± 0.5	0.5 ± 0.3
	FS	$1.9^{+0.7}_{-0.7}$	$2.3^{+0.8}_{-0.8}$	$1.7^{+0.7}_{-0.6}$	$0.1^{+0.2}_{-0.1}$
	Z+ν	1.9 ± 0.4	1.9 ± 0.4	2.0 ± 0.5	1.8 ± 0.6
	Sum	$96.0^{+10.4}_{-10.4}$	$12.4^{+2.3}_{-2.3}$	$4.9^{+1.0}_{-1.0}$	$2.5^{+0.7}_{-0.7}$
	Data	96	14	7	1
SRB	$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	50-100	100-150	150-250	250+
	Z+Hadronic	130.1 ± 12.8	12.8 ± 2.3	0.9 ± 0.3	0.4 ± 0.2
	FS	$0.3^{+0.2}_{-0.2}$	$0.4^{+0.3}_{-0.2}$	$0.4^{+0.3}_{-0.2}$	$0.1^{+0.2}_{-0.1}$
	Z+ν	0.6 ± 0.2	0.3 ± 0.1	0.7 ± 0.2	1.2 ± 0.4
	Sum	$131.0^{+12.8}_{-12.8}$	$13.6^{+2.4}_{-2.4}$	$2.0^{+0.5}_{-0.5}$	$1.6^{+0.5}_{-0.4}$
	Data	131	10	4	0
SRBb	$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	50-100	100-150	150-250	250+
	Z+Hadronic	37.9 ± 6.7	7.7 ± 3.1	4.0 ± 3.3	0.1 ± 0.1
	FS	$0.7^{+0.4}_{-0.3}$	$1.4^{+0.6}_{-0.5}$	$1.1^{+0.5}_{-0.4}$	$0.2^{+0.2}_{-0.1}$
	Z+ν	1.3 ± 0.4	2.0 ± 0.5	2.3 ± 0.6	1.0 ± 0.3
	Sum	$40.0^{+6.8}_{-6.8}$	$11.1^{+3.2}_{-3.2}$	$7.4^{+3.4}_{-3.4}$	$1.3^{+0.4}_{-0.3}$
	Data	40	10	5	0
SRC	$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	50-100	100-150	150+	
	Z+Hadronic	23.8 ± 5.5	1.2 ± 0.4	0.1 ± 0.1	
	FS	$0.1^{+0.2}_{-0.1}$	$0.4^{+0.3}_{-0.2}$	$0.1^{+0.2}_{-0.1}$	
	Z+ν	0.2 ± 0.1	0.1 ± 0.1	0.5 ± 0.2	
	Sum	$24.0^{+5.5}_{-5.5}$	$1.7^{+0.5}_{-0.5}$	$0.7^{+0.3}_{-0.2}$	
	Data	24	4	0	
SRCb	$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	50-100	100-150	150+	
	Z+Hadronic	9.9 ± 3.7	0.1 ± 0.5	0.0 ± 0.3	
	FS	$0.1^{+0.2}_{-0.1}$	$0.0^{+0.1}_{-0.0}$	$0.3^{+0.2}_{-0.2}$	
	Z+ν	0.0 ± 0.1	0.6 ± 0.2	0.6 ± 0.2	
	Sum	$10.0^{+3.7}_{-3.7}$	$0.8^{+0.5}_{-0.5}$	$0.9^{+0.5}_{-0.4}$	
	Data	10	2	2	

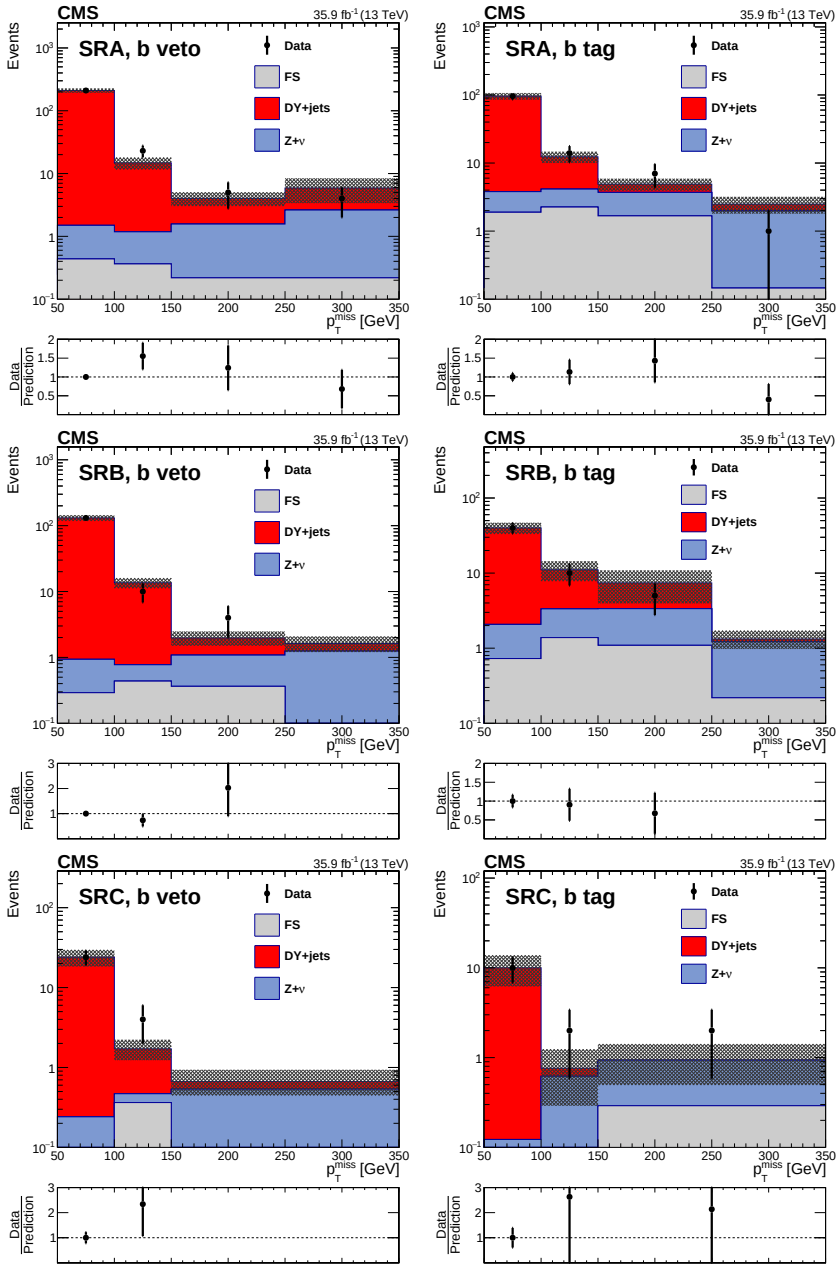


Figure 3.14: Results for the strong search regions are shown. The precise definitions of these regions are shown in table 3.7. The black dots represent the observed counts in the signal bins, with statistical uncertainty bands shown. The colored sections represent the predictions for the labeled background, and the hatch bands represent the total systematic and statistical uncertainty for the background prediction. The uncertainty band on the ratio reflects the straightforward propagation of the uncertainties for a ratio, incorporating both the observation and background uncertainties. Note the yields in the MET 50-100 bin agree perfectly by construction as that bin is used to normalize the Z+Hadronic prediction. The numerical yields corresponding to these plots are shown in table 3.18.

3.6.2 Electroweak Search Regions

The following shows the results for the electroweak search regions.

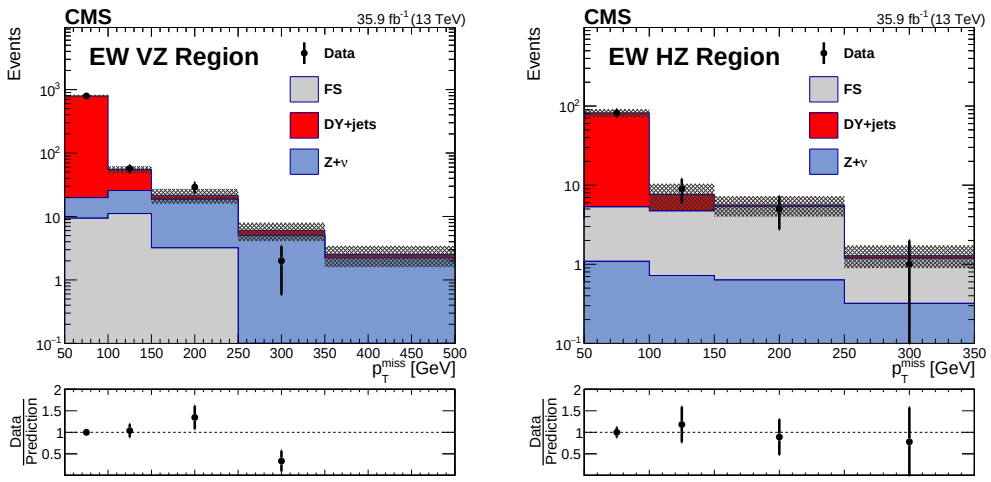


Figure 3.15: Results for the electroweak search regions are shown. The precise definitions of these regions are shown in table 3.7. The black dots represent the observed counts in the signal bins, with statistical uncertainty bands shown. The colored sections represent the predictions for the labeled background, and the hatch bands represent the total systematic and statistical uncertainty for the background prediction. The uncertainty band on the ratio reflects the straightforward propagation of the uncertainties for a ratio, incorporating both the observation and background uncertainties. Note the yields in the MET 50-100 bin agree perfectly by construction as that bin is used to normalize the Z+Hadronic prediction. The numerical yields corresponding to these plots are shown in table 3.19.

The results show good agreement between the standard model background prediction and the results in data. In the HZ region, all bins are within the 1-σ expected fluctuation. In the VZ region, a very small upward fluctuation is found in the MET 150-250 bin followed by downward fluctuations in the MET tails.

Table 3.19: Numerical results for the electroweak search regions are shown. The precise definitions of these regions are shown in table 3.7. See figure 3.15 for the corresponding plots. The agreement in the MET 50-100 bin is by construction as that bin is used to normalize the Z+Hadronic prediction. No statistically significant deviation from the standard model prediction is found.

VZ	$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	50-100	100-150	150-250	250-350	350+
	Z+Hadronic	773.2±31.9	29.3±4.4	2.9±2.1	1.0±0.7	0.3±0.3
	FS	$9.4^{+3.0}_{-3.0}$	$11.1^{+3.6}_{-3.6}$	$3.2^{+1.1}_{-1.1}$	$0.1^{+0.2}_{-0.1}$	$0.1^{+0.2}_{-0.1}$
	Z+ν	10.4±2.6	14.5±4.0	15.5±5.1	5.0±1.8	2.2±0.9
	Sum	$793.0^{+32.2}_{-32.2}$	$54.9^{+7.0}_{-7.0}$	$21.6^{+5.6}_{-5.6}$	$6.0^{+1.9}_{-1.9}$	$2.5^{+0.9}_{-0.9}$
	Data	793	57	29	2	0
HZ	$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	50-100	100-150	150-250	250+	
	Z+Hadronic	76.7±9.4	2.9±2.4	0.3±0.2	0.1±0.1	
	FS	$4.2^{+1.4}_{-1.4}$	$4.0^{+1.4}_{-1.4}$	$4.7^{+1.6}_{-1.6}$	$0.9^{+0.4}_{-0.4}$	
	Z+ν	1.1±0.3	0.7±0.2	0.6±0.2	0.3±0.1	
	Sum	$82.0^{+9.5}_{-9.5}$	$7.6^{+2.8}_{-2.8}$	$5.6^{+1.6}_{-1.6}$	$1.3^{+0.4}_{-0.4}$	
	Data	82	9	5	1	

3.7 Signal Interpretations

As mentioned previously in 3.1.1, there are several simplified models we use to interpret the search results in the context of electroweak scale supersymmetry. Each signal model has a cross section for production that is set by the masses of the new particles in that model, and we assume that the addition of a process to the standard model only increases the final state counts in sensitive regions (i.e. that there is no negative interference of new physics on the relevant known physics).

Events are generated using physics simulation, described in 2.4, for an interval of *mass points* in each model. Then the compatibility of the observed data are considered under the null-hypothesis and under the hypothesis that the signal model is present in nature. A confidence level is set us using the CL_s method described in [78], the details of which are outlined in the next section.

3.7.1 Statistical Treatment

This analysis uses the CMS Higgs Combine Tool [17] to produce our confidence intervals. A brief description of the statistical methods used are presented in this section, more detail can be found in [17], [47], and [78].

In any given search bin, a prediction is made for the expected standard model background as described in 3.5. Each prediction method is characterized by several largely uncorrelated sources of uncertainty, called *nuisance parameters*, whose one σ uncertainty bands are combined in quadrature to give a one σ band for the background prediction. The set of all such parameters in this analysis is denoted by $\vec{\theta}$. Each nuisance is modeled as a random variable that multiplies the relevant prediction. The distribution from which this random variable is assumed to be sampled is shown in table 3.20 for every nuisance in this analysis.[15]

[15] Further details about the uncertainty sources considered for the background predictions can be found in 3.5, and for the signal model prediction in 3.7.2.

The Likelihood of an Observation

The data in any signal bin can be seen as being sampled from a Poisson distribution. Since each search region is independent, the likelihood of observing the results we've seen in data is the product of the likelihoods to see the outcome in each search bin:

$$\mathcal{L}(\text{observation}) = \prod_{i \in \text{Search Bin}} \frac{\lambda_i^{n_i} e^{-\lambda_i}}{n_i!}. \tag{3.3}$$

where each λ_i and n_i are the expected and observed number of events in search bin i respectively. If the signal does not exists in nature (the background only hypothesis), then $\lambda_i = b$, the background prediction for the bin. If the signal does exist in nature then $\lambda_i = s + b$, the sum of the signal and background predictions[16]. However, due to the probabilistic nature of the background and signal predictions, λ itself must be modeled by a distribution of possible values.

$$\lambda_i(\vec{\theta}, \mu) = \mu s(\vec{\theta}) + b(\vec{\theta}).$$

We've also added the *signal strength* μ to the definition of λ for later mathematical convenience. μ is a parameter that allows us to interpolate any further mathematical constructions between the signal+background and null hypothesis by varying from 1 to 0.

The nuisance parameters are chosen such that they act as multiplicative factors on the prediction value, that is to say for K relevant nuisance parameters[17] and a central value denoted by $\widetilde{b}$,

$$b(\theta) = \widetilde{b} \times \theta^0 \times ... \times \theta^K,$$

[16] again, we assume no negative interference

[17] Not all nuisance parameters effect every prediction. For instance, if b were the prediction of the Z+Hadronic background in SRA, there would be 4 multiplicative factors listed in table 3.20

and the same relationship holds for $s(\theta)$. The probability that a nuisance parameter takes a specific value is drawn from a probability distribution that was fit to auxiliary measurements designed to capture the expected variation of the nuisance. The distribution chosen for each nuisance used in this analysis is given in table 3.20.

Taking into account the probabilistic nature of $\lambda(\vec{\theta},\mu)$, the likelihood in eq 3.3 should be updated with a multiplicative factor that accounts for the probability of certain set of values for $\vec{\theta}$. The likelihood that we observe our data given a specific nuisance configuration and signal strength μ

$$\mathcal{L}(\text{observation}|\vec{\theta},\mu) = P(\vec{\theta}) \left(\prod_{i \in \text{Search Bin}} \frac{\lambda_i(\vec{\theta},\mu)^{n_i} e^{-\lambda_i(\vec{\theta},\mu)}}{n_i!} \right), \tag{3.4}$$

where $P(\vec{a})$ is the probability that the nuisance vector has the value $\vec{a}$. Given that we assume the nuisances are completely uncorrelated, $P(\vec{\theta})$ will be a product of terms for each nuisance with distributions chosen from table 3.20. As an example, if the only nuisances considered were the luminosity and the closure for the low $E_{\text{T}}^{\text{miss}}$ bin of the SRA Z+Hadronic prediction, where the value of the luminosity multiplicative factor is denoted by $\theta[0]$ and the closure factor is denoted by $\theta[1]$, the full likelihood function for the GMSB signal model would read

$$\begin{aligned}
\mathcal{L}(\vec{\theta},\mu|\text{data}) = \\
&\underbrace{\frac{1}{\sqrt{2\pi}\ln(1.026)} e^{-\frac{(\ln\theta[0])^2}{2(\ln 1.026)^2}} \frac{1}{\theta[0]}}_{\text{Log Normal Distribution for Luminosity}} \times \underbrace{\frac{1}{\sqrt{2\pi}\ln(1.2)} e^{-\frac{(\ln\theta[1])^2}{2(\ln 1.2)^2}} \frac{1}{\theta[1]}}_{\text{Log Normal Distribution for SRA Low } E_{\text{T}}^{\text{miss}} \text{ Closure}} \times \\
&\underbrace{\frac{\lambda_i(\vec{\theta},\mu)^{n_{\text{SRA}}} e^{-\lambda_i(\vec{\theta},\mu)}}{n_{\text{SRA}}!}}_{\text{Poisson Distribution for SRA Bin Count}} \times \ldots \times \underbrace{\frac{\lambda_i(\vec{\theta},\mu)^{n_{\text{SRCb}}} e^{-\lambda_i(\vec{\theta},\mu)}}{n_{\text{SRCb}}!}}_{\text{Only the 6 Strong Search Regions Contribute}}
\end{aligned} \tag{3.5}$$

Table 3.20: Systematic Uncertainty Assumed Shapes.

Prediction	Nuisance	Distribution
Z + Hadronic	Closure	log Normal
	EWK Subtraction	log Normal
	γ Normalization	log Normal
	γ Statistics	log Normal
Flavor Symmetric	κ Value	log Normal
	$R_{SF/DF}$ Value	log Normal
	Same Sign Statistics	Γ
Z + ν	Closure	log Normal
	MC Statistics	log Normal
Signal	MC Statistics	log Normal
	B-tag eff, heavy flavor	log Normal
	B-tag eff, light flavor	log Normal
	Jet Energy Scale	log Normal
	Lepton Trigger Efficiency	log Normal
	Lepton ID/Iso Efficiency	log Normal
	ISR Modeling	log Normal
	Luminosity	log Normal
	Fastsim MET Modeling	log Normal
	Generator Scale variations	log Normal
	PU reweighting	log Normal

The nuisances for the background predictions were described in the previous chapter, a breif description of the signal nuisances is given in section 3.7.2.

Setting Exclusion Limits

In order to exclude a mass point for a signal model, we adopt the CL_s method. The basic premise is that in order for a mass point to be excluded at the 95% confidence level, we require that, under some metric of probability, the background only hypothesis is 20 times more likely than the signal+background hypothesis[18].

The metric of probability we use is based upon the following formalism:

First, we construct the so-called profile-likelihood ratio, a "test statistic" defined as

[18] $\frac{1}{20}$ being the left of 5% from 95%

$$q_\mu = -2\ln\left(\frac{\mathcal{L}(\text{observation}|\vec{\theta}_\mu,\mu)}{\mathcal{L}(\text{observation}|\vec{\theta}_{\hat{\mu}},\hat{\mu})}\right) \tag{3.6}$$

where

$$\vec{\theta}_a = \vec{\theta} \text{ such that } \mathcal{L}(\text{observation}|\vec{\theta},a) \text{ is maximal, and}$$
$$\hat{\mu} = \mu \text{ such that } \mathcal{L}(\text{observation}|\vec{\theta}_\mu,\mu) \text{ is maximal. } [19]$$

In the above, we only search over values of $\hat{\mu} \leq \mu$, if the the global maximum is found outside this region, q_μ is set to 0. The denominator in eq 3.6 is the global maximum of the likelihood function over all values of $\vec{\theta}$ and μ, where μ is less than the μ in the numerator. The numerator in eq 3.6 is the maximum likelihood of the observed data over all possible θ configurations, but for a fixed μ. If a value of μ is very likely, then we expect the θ parameter to ensure that correlations between bins are not double counted in their effect on the compatibility of the observation. For instance, if the luminosity measurement was 3% low for our dataset and no new physics exists, the hope is that θ_0 will have the luminosity factor of 1.03.

Notice that q_μ is bounded by 0 below and unbounded above. If an observation point is highly likely w.r.t the global maximum, then q_μ is 0. To define the 95% confidence interval, it is easiest to think in terms of simulated data. [20]

Let q_μ^{obs} be the value of q for some μ given the actual results of our experiment shown in 3.6. We construct a distribution of the possible values for q_μ under the null hypothesis by using the likelihood in eq. 3.4 to give simulated results for an ensemble experiments and compute q_μ

[19] This means that $\mathcal{L}(\text{observation}|\vec{\theta}_{\hat{\mu}},\hat{\mu})$ is the global maximum of the likelihood

[20] although in practice analytic approximations are made in order to do the following calculations

for each simulated result. We assume the distribution of q_μ to tend larger for larger values of μ since the probability of measuring a new physics process should decrease with the processes cross section in the null hypothesis, and high values of q are associated with unlikely data. The percentage of simulated experiments where q_μ is larger than q_0^{obs} is called CL_b, the "confidence level" of the background only hypothesis. In short, we compute the percentage of experiments that give a less likely outcome than what we found.

We repeat this procedure again, only in this case using the background+signal hypothesis to generate our simulated experiments, i.e. μ is not required to be 0 in λ. The percentage of simulated experiments where q_μ is larger than q_μ^{obs} is called CL_{s+b}, the "confidence level" of the signal+background hypothesis.

Finally, the value of μ for which

$$\text{CL}_s = \frac{\text{CL}_{s+b}}{\text{CL}_b} \leq 0.05$$

is called the exclusion limit for the mass point. If the value of μ for which this condition is met is 1 or more, the mass point can not be excluded at the 95% level. If the value of μ is less than 1, then we call the mass point excluded.

In the final exclusion plots, an "expected limit" is normally drawn as well. To construct this limit, we follow the same limit setting procedure described above, but replace q_μ^{obs} with the q value for a simulated result of our experiment assuming the background only hypothesis. This process is then repeated to produce a distribution of excluded μ values. The central value and one-σ expected limits are then taken from this distribution.

3.7.2 Systematics in Signal Yield

In order to interpret these results in the context of electroweak scale SUSY, the simplified SUSY models in figure 3.1 and 3.2 are simulated using the madgraph package as explained in sec

2.4. The simulated events are then run through the same software used to reconstruct and select real data events.

Table 3.21 is a list of nuisance parameters associated with the simulation of SUSY models. Although many of these could also be applied to the other simulated background predictions, to wit: the Z+ν background, other nuisances would dominate these mostly few-percentage level effects.[21]

Table 3.21: Systematic uncertainties of the expected signal yield.

Source	Value (%)
MC Statistics	±1–60
B-tag eff, heavy flavor	±3
B-tag eff, light flavor	±2
Jet Energy Scale	±1–15
Lepton Trigger Efficiency	±3
Lepton ID/Iso Efficiency	±7
ISR Modeling	±0–7
Luminosity [7]	±2.6
Fastsim MET Modeling	±1–60
Generator Scale variations	±3
PU reweighting	±3
Total uncertainty on signal	±10–60

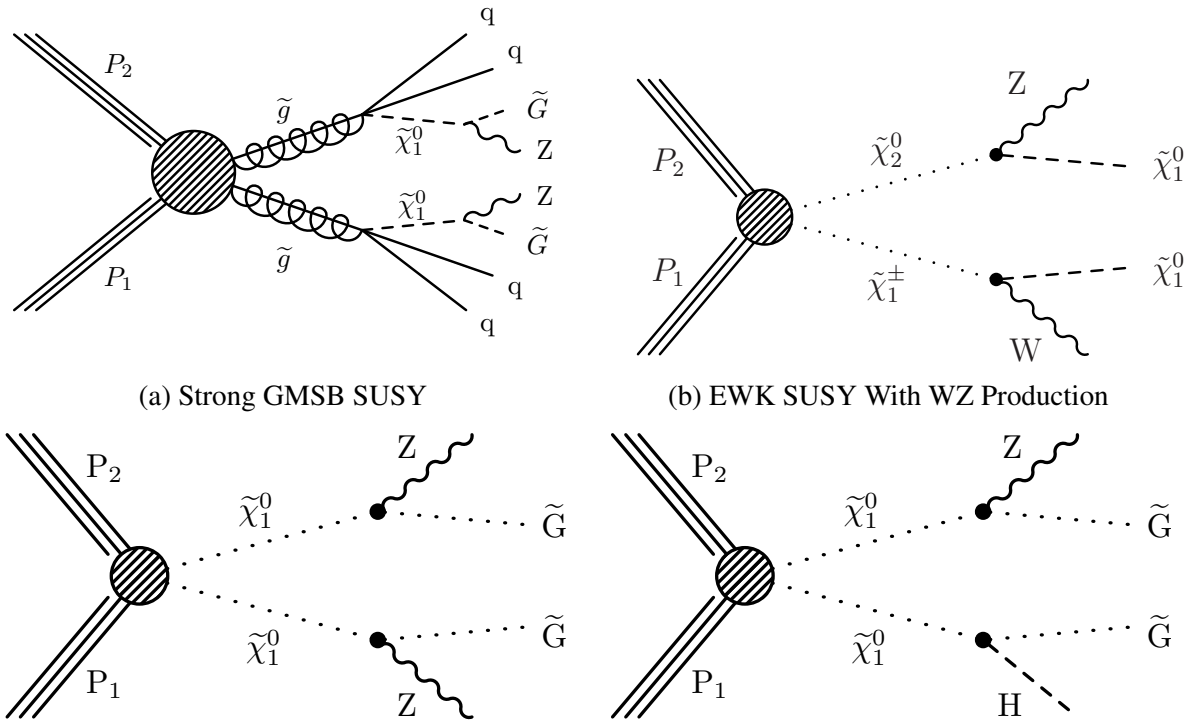
A brief description of the nuisances follows: The MC statistics uncertainty is the statistical uncertainty on number of signal events generated. The B-tag efficiency for heavy (light) flavor jets characterizes the difference between MC and data jet reconstruction for jets originating from bottom and charm (up, down, and strange) quarks. The jet energy scale nuisance is due to the uncertainty on the correction applied to jet energies. The lepton efficiencies are due to the simulated modeling of the acceptance of leptons. The ISR modeling is due to uncertainty in modeling the initial state radiation in simulation. The luminosity nuisance is due to uncertainty in the amount of luminosity provided to the CMS detector. The Fastsim MET modeling is due to the lower quality fast simulation of the CMS detector used for signal simulation, and its effect on the

[21] Fastsim $E_{\mathrm{T}}^{\mathrm{miss}}$ modeling is unique to the SUSY simulation and statistics vary by sample.

MET resolution. The generator scale variations characterize the uncertainty in QFT parameters, like the proton pdfs, when computing matrix elements. Finally, the pileup reweighting nuisance is due to the uncertainty in amount of pileup added to signal events.

3.7.3 Exclusion Limits

In this section, the results in 3.6 are interpreted in the context of the simplified supersymmetric models described in 3.1.1. For the sake of convenience, the Feynman diagrams for these models are reproduced in figure 3.16.



(a) Strong GMSB SUSY

(b) EWK SUSY With WZ Production

(c) EWK SUSY With ZZ Production

(d) EWK SUSY With HZ Production

Figure 3.16: Feynman diagrams for the SUSY models used in interpreting the results of this analysis. More exposition on the properties of these models is found in sec 3.1.1.

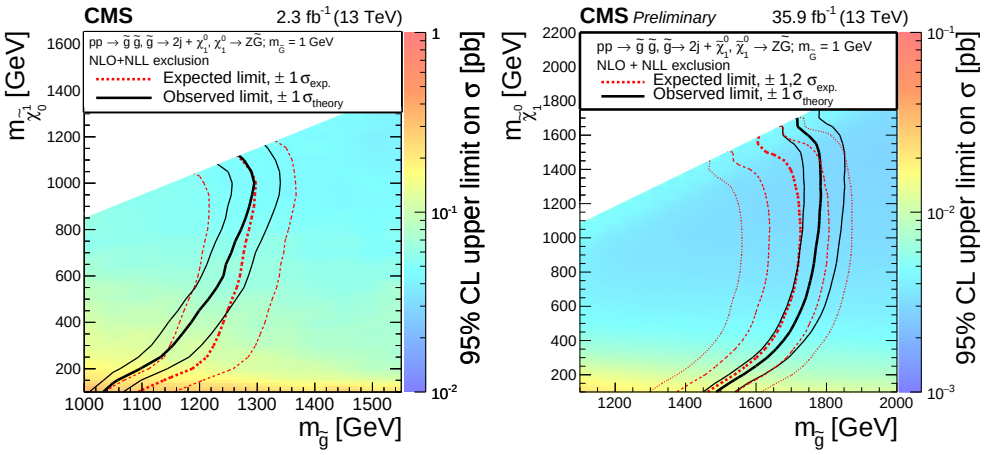
For the strong production model, shown in fig. 3.16a, the exclusion limits are shown in figure 3.17, which combines data from all the strong search regions in the likelihood function. The free parameters in this model are the gluino and $\widetilde{\chi}_1^0$ masses, the gravitino mass is set to 1 GeV. The

bulk of the sensitivity comes from the high jet multiplicity and high $E_{\mathrm{T}}^{\mathrm{miss}}$ search regions, SRB(b) and SRC(b). Due to the downward fluctuations in most of the high $E_{\mathrm{T}}^{\mathrm{miss}}$ bins, our observed limit is slightly better than the expected limit. In a previous CMS result probing similar final states,[66] this simplified model was excluded at the 95% CL for gluino masses roughly below 1.2-1.3 TeV for $\widetilde{\chi}_1^0$ masses below 1 TeV. This analysis advances the limits significantly, by roughly 400 GeV in gluino mass for similar $\widetilde{\chi}_1^0$ masses, as can be seen in figure 3.17.

In the compressed spectrum (where the $\widetilde{\chi}_1^0$ mass is close to the gluino mass), the neutralino is expected to carry away more energy from the gluino than the quarks. In the extremely compressed regions where the $\widetilde{\chi}_1^0$ and gluino masses are only separated by about 100 GeV, it's possible that one of the jets can be missed. This makes SRB/SRBb more important regions at the top of the plot. The data shows a downward fluctuation in the high $E_{\mathrm{T}}^{\mathrm{miss}}$ bins of those regions, and so the observed limits tend to move towards higher masses than the expected limits. At large mass splittings, i.e. points near the x axis, SRC and SRCb are the most important regions due to the high jet multiplicity. There is a small upwards fluctuation in SRCb which offsets the downward fluctuations in SRB, SRBb, and SRC; this causes the observed limit to be closer to the expected limit in the bottom of the plot.

Figure 3.18 shows the exclusion limits for the electroweak WZ model in fig. 3.16b. For this model, data from the electroweak regions are used in the likelihood, however almost all the discriminatory power comes from the VZ region. The free parameters in this model are the mass of the $\widetilde{\chi}_1^0$ and $\widetilde{\chi}_1^{\pm}$, the $\widetilde{\chi}_2^0$ mass is set equal to that of the $\widetilde{\chi}_1^{\pm}$. Downward fluctuations in the high $E_{\mathrm{T}}^{\mathrm{miss}}$ bins for that region again make the observed limits stronger than the expected limits. For light LSP mass points, the model is excluded for $\widetilde{\chi}_1^{\pm}$ and $\widetilde{\chi}_2^0$ masses up to 600 GeV, better than double the previous limits. In the high $\widetilde{\chi}_1^{\pm}$ and $\widetilde{\chi}_2^0$ mass range, the excluded LSP mass range extends to around 250 GeV.

Figure 3.19, shows the exclusion limits for the electroweak ZZ model shown in fig. 3.16c. In this model, the only free parameter is the mass of the $\widetilde{\chi}_1^0$, the gravitino mass is assumed to be 1



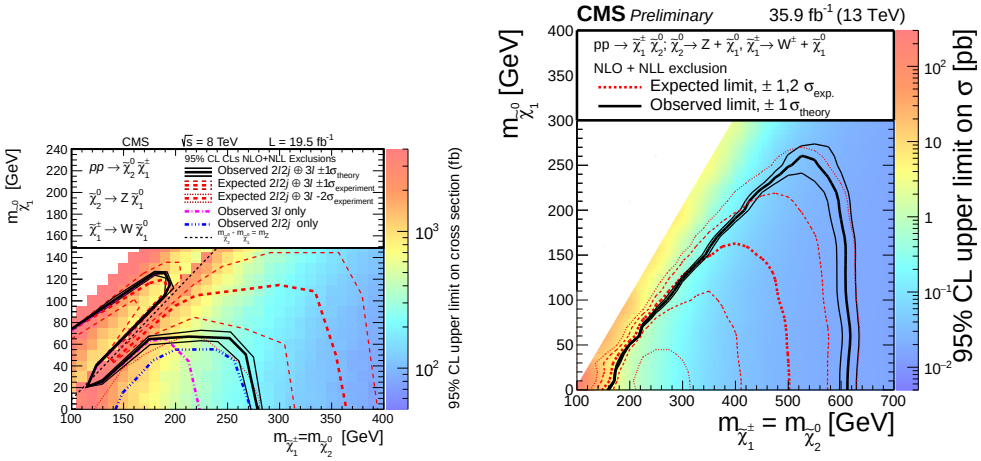
(a) Previous best limits on this model

(b) Limits set by this analysis

Figure 3.17: The limits set on the strong GMSB SUSY model from fig. 3.16a are shown to the right. To the left, we show the previous best limit on this model set by CMS in 2016.[66] The solid black and bold dotted red lines trace out the largest masses where the signal strength μ is less than 1 at 95% confidence for observed and expected data respectively as described in section 3.7.1, mass points to the left of these line are excluded. The data excludes this model at the 95% CL for gluino masses below 1450 GeV for low LSP masses, and below 1750 GeV for large LSP masses.

GeV. The thick pink line is the theoretical cross section for this model as a function of the free mass parameter, its thickness shows the uncertainty on the theory calculation. The solid black and dotted black lines show the observed and expected limits respectively. The point where the pink line and the solid black line cross is the observed exclusion limit for the $\widetilde{\chi}_1^0$ mass at 95% confidence. This result pushed the previous limits on the $\widetilde{\chi}_1^0$ mass from about 375 GeV to about 675 GeV.

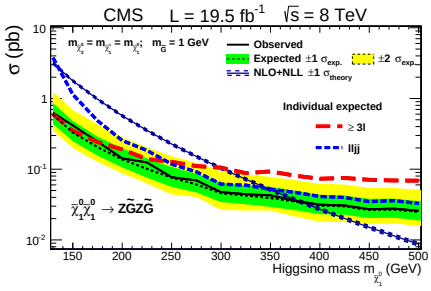
Finally, figure 3.20, shows the exclusion limits for the electroweak HZ model shown in fig. 3.16d. The only free parameter in this model is the $\widetilde{\chi}_1^0$ mass. This model assumes a 50% branching ratio of the $\widetilde{\chi}_1^0$ to higgs and Z bosons. The likelihood for this model is constructed using data from the VZ and HZ regions. The likelihood also takes into account that the signature of the ZZ model above should be produced with $\frac{1}{2}$ the frequency as that of the HZ model since the 50% branching ratio to Z bosons means that the ZZ final state should be produced as well. Points to the left of the crossing point between the pink and solid black line are excluded at the



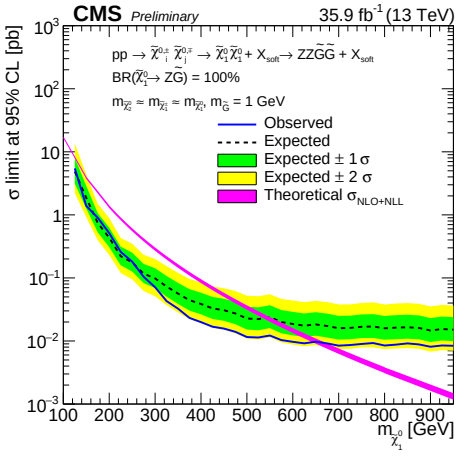
(a) Previous best limits on this model (b) Limits set by this analysis

Figure 3.18: The limits set on the electroweak WZ model in fig. 3.16b. In this limit, data from the electroweak VZ and HZ search regions are utilized, though the VZ region has by far the larger sensitivity to this model. The solid back and thick dashed red lines show the observed and expected limits respectively as described in section 3.7.1, mass points below these lines are excluded. These results push the excluded $\widetilde{\chi}_1^{\pm}$ and $\widetilde{\chi}_2^0$ masses to 600 GeV from 275 GeV for low mass LSPs, and excludes LSP mass points for up to about 250 GeV for high $\widetilde{\chi}_1^{\pm}$ and $\widetilde{\chi}_2^0$ masses, up from about 60.

95% level. This analysis pushed the limits for the $\widetilde{\chi}_1^0$ mass in this model from 250 GeV to about 500 GeV.

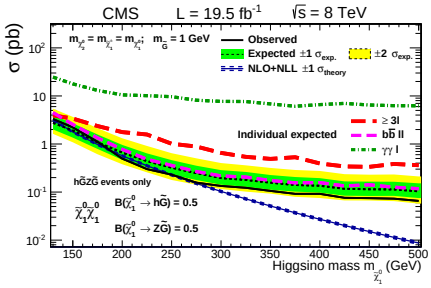


(a) Previous best limits on this model

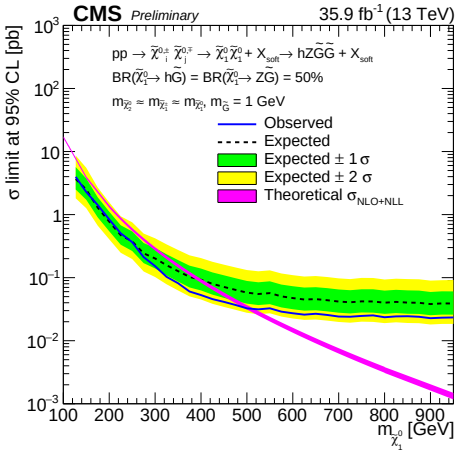


(b) Limits set by this analysis

Figure 3.19: The limits set on the Electroweak ZZ model shown in fig. 3.16c. In this limit, data from the electroweak VZ search region is utilized and the $\widetilde{\chi}_1^0$ is assumed to decay 100% to the Z. On the right, the point where the pink line crosses the solid black line marks the highest mass excluded. Where the pink line crosses the dotted line marks the expected limit. The downward fluctuation in the VZ region high $E_{\mathrm{T}}^{\mathrm{miss}}$ bin again caused the observed limit to be better than the expected limit. The limit set on the $\widetilde{\chi}_1^0$ mass increased by about 300 GeV from the previous result on the left.



(a) Previous best limits on this model



(b) Limits set by this analysis

Figure 3.20: The limits set on the Electroweak HZ model shown in fig. 3.16d. In this limit, data from the electroweak VZ and HZ search regions are combined and the branching ratios for $\widetilde{\chi}_1^0$ are assumed to be 50/50 between the Higgs and Z.

3.8 Acknowledgments

In the previous chapter, I presented a search for new physics which relied heavily on the work of several collaborators. This work would have been impossible without the leadership of Dominick Olivito, Pablo Martinez Ruiz del Arbol, and Vince Welke. In addition, the analysis relied on direct contributions from Christian Schomakers, Leonora Vesterbacka, and Sergio Sánchez Cruz. I would also like to thank CMS members (and my co-chairs) Frank Wüerthwein and Avraham Yagil for their mentorship and support, their group has developed this search over many years. In addition to my direct collaborators, the CMS collaboration itself consists of over 4000 members who built the infrastructure to acquire data, provided crucial feedback and internal review of analysis methodology, and supplied indispensable tools for this analysis, e.g. the b-tagging software suite and the MVA for electron identification. I would like to acknowledge all members of CMS, past and present, for their tremendous contributions to this work, enabling a few people with laptops to search for new fundamental particles.

I must also specifically thank my co-authors on CMS for contributing figure 3.17 (a), showing the previous best limits on the strong GMSB SUSY model.

Chapter 4

Conclusions

In this thesis, we presented a search for dark particles which couple to the Z boson by searching in proton-proton collision final states with two opposite-charge and same-flavor light leptons, hadronic jets, and transverse momentum imbalance. No statistically significant deviation from the standard model predictions were found in the search regions. Interpretations of the search results were formulated with respect to simplified models of supersymmetry. The results presented represent the state-of-the-art for the exclusion of those simplified models.

For the model of GMSB, in figure 3.16a, gluino mass limits were extended by approximately 50%, between 400-500 GeV, at similar neutralino masses. For the model with the WZ final state, in figure 3.16b, electroweakino masses were extended 325 GeV from 275 GeV out to 600 GeV at low LSP masses, and LSP mass ranges excluded were extended out to 250 GeV. For the model with the ZZ final state, in figure 3.16c, the exclusion on the neutralino masses was extended from 375 GeV to 600 GeV. Finally, for the model with the HZ final state, in figure 3.16d, the exclusion on the neutralino mass was extended from 275 GeV to 500 GeV.

Though the standard model has enjoyed enormous success explaining and predicting the structure we see in fundamental interactions, observations of dark matter and energy from astrophysics, gravitation, and neutrino masses show that the standard model must not be the

complete theory of nature. Supersymmetry has been distinguished as an extension to the standard model for several reasons, highly motivated by naturalness in models of TeV scale SUSY breaking, predicting new particles near the TeV scale.

With no new particles found thus far and exclusion limits for simplified SUSY models pushing into the TeV scale, the theoretical consideration of naturalness and TeV scale SUSY breaking are becoming less and less tenable. However, there are still other indications, e.g. unification of forces and Higgs metastability, that we expect to see some new particles in accessible mass ranges.

Other surprises can also be on the horizon. In the coming years we expect to see the birth of gravitational wave astronomy, multimessenger astronomy, the high luminosity LHC, and progress in the neutrino sector. All of these experimental avenues will test completely new regimes of physics and have promise to usher in a revolution of understanding about fundamental questions. The prevailing wisdom today among theoretical physicists is that the standard model is the low energy limit of a more complete theory. The community's ultimate goal is still to find a unified theory of everything that includes all known phenomenology including gravitation, the particles in the standard model, dark matter, and spacetime expansion.

Appendix A

Deterministic Annealing

The free energy analog is

$$F = -T \sum_{i}^{\#\text{ tracks}} p_i \log \left(\sum_{k}^{\#\text{ vertices}} \rho_k \exp \left[-\frac{\left(z_i^t - z_k^v\right)^2}{T(\sigma_i^z)^2} \right] \right)$$

where p_i is the weight of the track i, a measure of its quality from 0 to 1, z_i^t and z_k^v are the positions of the track i and vertex k along the beamline with σ_i^z being the uncertainty in the position of the track, and ρ_k is a measure of the vertex quality. The upshot of this method is that as the temperature T is reduced, the optimal number of vertices to accommodate the tracks begins to increase with each track getting its own vertex at $T = 0$. The final number of vertices is set at special temperature of $T = 4$, which was identified as a good compromise between incorrect splitting and vertex resolution. A track i is assigned a probability to come from a vertex k by a Maxwell-Boltzmann distribution

$$p_{ik} = \frac{\rho_k \exp \left[-\frac{\left(z_i^t - z_k^v\right)^2}{T(\sigma_i^z)^2} \right]}{\sum_{j}^{\#\text{ vertices}} \rho_j \exp \left[-\frac{\left(z_i^t - z_j^v\right)^2}{T(\sigma_i^z)^2} \right]}.$$

The final vertex assignment is done by taking the highest probability vertex for each track

at $T = 1$.

Appendix B

The motion of particles in a magnetic field

Magnetic fields can not change the magnitude of velocity for a charged particle due to the Lorentz force law containing a cross product that includes $\vec{v}$,

$$\vec{F} = q\vec{v} \times \vec{B}.$$

This means that a charged particle in a magnetic field feels an acceleration which attempts to turn it in a circle at constant velocity. It is well known that this type of centripetal acceleration must have magnitude $a = \frac{v^2}{r}$. Therefore,

$$\frac{mv^2}{r} = q\vec{v} \times \vec{B}$$

by collecting terms, we can find the momentum of the particle in terms of the radius of rotation and the magnetic field

$$p = qBr.$$

This relationship still holds in the relativistic case when $p = mv \rightarrow \gamma mv$.

The radius of motion caused by the magnetic field in a particle detector can not be

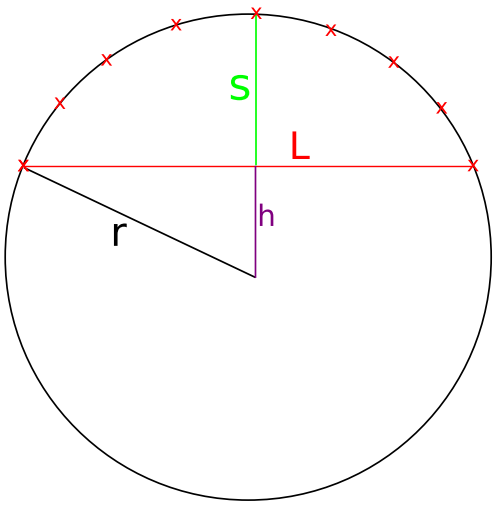


Figure B.1: The Sagitta of the circle. The red x symbols represent where a particle detector might sample the position of a particle moving along the segment whose endpoints coincide with the chord L. The sagitta, s, is what is measured directly by detectors and used to estimate the momentum of charged particles in a magnetic field.

measured directly. Particle detectors essentially sample the motion of the particle at various points along its trajectory. Figure B.1 shows an example of a circular arc which represents the trajectory of a charged particle, the red x symbols represent points where a hit was recorded by the detector. To get a measure of the radius of curvature, the sagitta of the track can be measured between the two terminal hits of the track, which is the maximal deviation of the particle's track from a straight line.

Using

$$h+s=r \text{ and}$$
$$\left(\frac{L}{2}\right)^2+h^2=r^2$$

yeilds

$$s=r\left(1-\sqrt{\left(1-\frac{L^2}{4r^2}\right)}\right).$$

Expanding to first order about $\frac{L^2}{4r^2} = 0$, which is regulated by the momentum of the tracks, we can solve for r to find

$$r = \frac{L^2}{8s}.$$

This can be replaced into the above expression for the momentum to yeild

$$p = \frac{qBL^2}{8s}.$$

The realtive uncertainty on the momentum is then

$$\delta p = \frac{qL^2}{8s}\delta B - \frac{qBL^2}{8s^2}\delta s + \frac{2qBL}{8s}\delta L.$$

Or, writing in terms of p:

$$\frac{\delta p}{p} = \frac{\delta B}{B} - \frac{\delta s}{s} + \frac{2\delta L}{L}$$

Therefore, a longer track and a higher magnetic field provide better momentum resolution.

Bibliography

[1] Cms y totem estudiarán la resonancia a 750 gev vía ct-pps. http://francis.naukas.com/2016/05/16/cms-y-totem-estudiaran-la-resonancia-a-750-gev-gracias-a-ct-pps/. Accessed: 2018-04-17.

[2] Electromagnetic calorimeter. https://cms-docdb.cern.ch/cgi-bin/PublicDocDB/ShowDocument?docid=12030. Accessed: 2018-04-23.

[3] *The CMS muon project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997.

[4] CMS JEC Run I legacy performance plots. Sep 2015.

[5] Performance of electron reconstruction and selection with the cms detector in proton-proton collisions at $\sqrt{s}$ = 8 tev. *Journal of Instrumentation*, 10(06):P06005, 2015.

[6] Performance of missing energy reconstruction in 13 TeV pp collision data using the CMS detector. Technical Report CMS-PAS-JME-16-004, CERN, Geneva, 2016.

[7] CMS Luminosity Measurements for the 2016 Data Taking Period. Technical Report CMS-PAS-LUM-17-001, CERN, Geneva, 2017.

[8] T. Aaltonen et al. Observation of Single Top Quark Production and Measurement of —Vtb— with CDF. *Phys. Rev.*, D82:112005, 2010.

[9] S Abdullin, P Azzi, F Beaudette, P Janot, A Perrotta, and the CMS Collaboration. The fast simulation of the cms detector at lhc. *Journal of Physics: Conference Series*, 331(3):032049, 2011.

[10] S. Abdullin and etal. Design, performance, and calibration of cms hadron-barrel calorimeter wedges. *The European Physical Journal C*, 55(1):159–171, May 2008.

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

[12] P. A. R. Ade et al. Planck 2015 results. XIII. Cosmological parameters. *Astron. Astrophys.*, 594:A13, 2016.

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

[14] Steven W. Allen, August E. Evrard, and Adam B. Mantz. Cosmological parameters from observations of galaxy clusters. *Annual Review of Astronomy and Astrophysics*, 49(1):409–470, 2011.

[15] Daniele Alves and etal. Simplified models for lhc new physics searches. *Journal of Physics G: Nuclear and Particle Physics*, 39(10):105005, 2012.

[16] Ryan Atkin. Review of jet reconstruction algorithms. *Journal of Physics: Conference Series*, 645(1):012008, 2015.

[17] The ATLAS, CMS Collaborations, and The LHC Higgs Combination Group. Procedure for the LHC Higgs boson search combination in Summer 2011. Technical Report CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11, CERN, Geneva, Aug 2011.

[18] Frank T. Avignone. Viewpoint: Homing in on axions? URL: https://physics.aps.org/articles/v11/34, 2018.

[19] G L Bayatian and etal. *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*. Technical Design Report CMS. CERN, Geneva, 2006. There is an error on cover due to a technical problem for some items.

[20] A. V. Bednyakov, B. A. Kniehl, A. F. Pikelner, and O. L. Veretin. Stability of the electroweak vacuum: Gauge independence and advanced precision. *Phys. Rev. Lett.*, 115:201802, Nov 2015.

[21] Michael Benedikt, Paul Collier, V Mertens, John Poole, and Karlheinz Schindl. *LHC Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2004.

[22] Gianfranco Bertone, Dan Hooper, and Joseph Silk. Particle dark matter: evidence, candidates and constraints. *Physics Reports*, 405(5):279 – 390, 2005.

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

[24] Sean M. Carroll. The cosmological constant. *Living Reviews in Relativity*, 4(1):1, Feb 2001.

[25] S Chatrchyan and etal. The CMS experiment at the CERN LHC. The Compact Muon Solenoid experiment. *JINST*, 3:S08004. 361 p, 2008. Also published by CERN Geneva in 2010.

[26] Serguei Chatrchyan et al. Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in pp Collisions at $\sqrt{s} = 7$ TeV. *JINST*, 8:P09009, 2013. [JINST8,9009(2013)].

[27] Sergio Cittolin, Attila Rácz, and Paris Sphicas. *CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and data-acquisition project*. Technical Design Report CMS. CERN, Geneva, 2002.

[28] CMS Collaboration. Measurement of the muon stopping power in lead tungstate. *Journal of Instrumentation*, 5(03):P03007, 2010.

[29] CMS Collaboration. Baseline muon selections. https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2, 2015.

[30] CMS Collaboration. Recommended jet energy corrections and uncertainties for data and mc. https://twiki.cern.ch/twiki/bin/view/CMS/JECDataMC, 2015.

[31] CMS Collaboration. Introduction to jet energy corrections at cms. Revision 6, 2016.

[32] CMS Collaboration. Performance of b-Tagging Algorithms in Proton Collisions at 13 TeV using the 2016 Data. Jul 2016.

[33] CMS Collaboration. Identification of heavy-flavour jets with the cms detector in pp collisions at 13 tev. *Journal of Instrumentation*, 13(05):P05011, 2018.

[34] CMS Collaboration. Level-1 muon trigger performance with the 2017 data. Private Twiki – Revision 3, 2018.

[35] CMS Collaboration. Physics results combined. Public Twiki – Revision 75, 2018.

[36] CMS Collaboration. Public cms luminosity information. Public Twiki – Revision 131, 2018.

[37] The ATLAS Collaboration. 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.

[38] The CMS Collaboration. Muon reconstruction in the cms detector. *CMS Internal Analysis Note (available to CMS members only)*, 2009. CMS AN-2008/97.

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

[40] The CMS Collaboration. 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.

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

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

[43] The CMS Collaboration. Interpretation of searches for supersymmetry with simplified models. *Phys. Rev. D*, 88:052017, Sep 2013.

[44] The CMS Collaboration. Description and performance of track and primary-vertex reconstruction with the cms tracker. *Journal of Instrumentation*, 9(10):P10009, 2014.

[45] The CMS Collaboration. Performance of electron reconstruction and selection with the cms detector in proton-proton collisions at $\sqrt{s} = 8$ tev. *Journal of Instrumentation*, 10(06):P06005, 2015.

[46] The CMS Collaboration. Performance of photon reconstruction and identification with the cms detector in proton-proton collisions at $\sqrt{s} = 8$ tev. *Journal of Instrumentation*, 10(08):P08010, 2015.

[47] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C*, 71(2):1554, Feb 2011.

[48] E. Daw. "lecture 7 - rapidity and pseudorapidity. `http://www.hep.shef.ac.uk/edaw/PHY206/Site/2012_course_files/phy206rlec7.pdf`, March 2012. Pages 6 through 8.

[49] David d'Enterria. Quark-gluon matter. *Journal of Physics G: Nuclear and Particle Physics*, 34(7):S53, 2007.

[50] Tommaso Dorigo. Electron scattering and the proton structure, 2010. http://www.science20.com/quantum_diaries_survivor/electron_scattering_and_proton_structure.

[51] John Ellis. The discovery of the gluon. *International Journal of Modern Physics A*, 29(31):1430072, 2014.

[52] Lyndon R Evans and Philip Bryant. LHC Machine. *JINST*, 3:S08001. 164 p, 2008. This report is an abridged version of the LHC Design Report (CERN-2004-003).

[53] Yoav Freund, Robert Schapire, and Naoki Abe. A short introduction to boosting. *Journal-Japanese Society For Artificial Intelligence*, 14(771-780):1612, 1999.

[54] J. J. Thomson M.A. F.R.S. Xl. cathode rays. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 44(269):293–316, 1897.

[55] R. Fruhwirth, W. Waltenberger, and P. Vanlaer. Adaptive vertex fitting. *J. Phys.*, G34:N343, 2007.

[56] R. FrÃŒhwirth. Application of kalman filtering to track and vertex fitting. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 262(2):444 – 450, 1987.

[57] Garyzx. File:standard model feynman diagram vertices.png, 2011. URL: https://commons.wikimedia.org/wiki/File:Standard_Model_Feynman_Diagram_Vertices.png.

[58] W. Gohn. The muon g-2 experiment at Fermilab. In *18th International Workshop on Neutrino Factories and Future Neutrino Facilities Search (NuFact16) Quy Nhon, Vietnam, August 21-27, 2016*, 2016.

[59] Dan Green. *The physics of particle detectors*. Cambridge University Press, 2010.

[60] Michael Grefe. Gauge mediation. `http://www.desy.de/~mgrefe/files/Grefe_GaugeMediation.pdf`, 2010.

[61] David J. Griffiths. *Introduction to Electrodynamics*. Prentice Hall, 3rd edition, 1999.

[62] G.'t Hooft. Renormalizable lagrangians for massive yang-mills fields. *Nuclear Physics B*, 35(1):167 – 188, 1971.

[63] D. I. Kazakov. Supersymmetry on the Run: LHC and Dark Matter. *Nucl. Phys. Proc. Suppl.*, 203-204:118–154, 2010.

[64] V. Khachatryan and etal. Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV. *Journal of Instrumentation*, 12(02):P02014–P02014, feb 2017.

[65] Vardan Khachatryan et al. Search for Physics Beyond the Standard Model in Events with Two Leptons, Jets, and Missing Transverse Momentum in pp Collisions at sqrt(s) = 8 TeV. *JHEP*, 04:124, 2015.

[66] Vardan Khachatryan et al. Search for new physics in final states with two opposite-sign, same-flavor leptons, jets, and missing transverse momentum in pp collisions at $\sqrt{s}$ = 13 TeV. *Submitted to: JHEP*, 2016.

[67] Sabine Kraml, Suchita Kulkarni, Ursula Laa, Andre Lessa, Wolfgang Magerl, Doris Proschofsky-Spindler, and Wolfgang Waltenberger. Smodels: a tool for interpreting simplified-model results from the lhc and its application to supersymmetry. *The European Physical Journal C*, 74(5):2868, May 2014.

[68] Laurascudder. Eightfold way (physics), 2007. URL: https://en.wikipedia.org/wiki/Eightfold_Way_(physics) Files: Meson_octet.png Baryon_decuplet.png.

[69] William R. Leo. *Techniques for Nuclear and Particle Physics Experiments*. Springer-Verlag, 1987.

[70] C.G Lester and D.J Summers. Measuring masses of semi-invisibly decaying particle pairs produced at hadron colliders. *Physics Letters B*, 463(1):99 – 103, 1999.

[71] Ye Li and Frank Petriello. Combining qcd and electroweak corrections to dilepton production in the framework of the fewz simulation code. *Phys. Rev. D*, 86:094034, Nov 2012.

[72] Stephen P. Martin. A Supersymmetry primer. 1997. [Adv. Ser. Direct. High Energy Phys.18,1(1998)].

[73] Konstantin T. Matchev and Scott Thomas. Higgs and z-boson signatures of supersymmetry. *Phys. Rev. D*, 62:077702, Sep 2000.

[74] Patrick Meade, Matthew Reece, and David Shih. Prompt decays of general neutralino nlsps at the tevatron. *Journal of High Energy Physics*, 2010(5):105, May 2010.

[75] MissMJ. File:standard model of elementary particles.svg, 2006. URL: https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg.

[76] C. Patrignani et al. Review of Particle Physics. *Chin. Phys.*, C40(10):100001, 2016.

[77] Victor Pavlunin. Modeling missing transverse energy in $v+$jets at cern lhc. *Phys. Rev. D*, 81:035005, Feb 2010.

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

[79] Joshua T. Ruderman and David Shih. General neutralino nlsps at the early lhc. *Journal of High Energy Physics*, 2012(8):159, Aug 2012.

[80] Svend E. Rugh and Henrik Zinkernagel. The quantum vacuum and the cosmological constant problem, 2001.

[81] Et. Al. S. Chatrchyan. Search for physics beyond the standard model in events with a z boson, jets, and missing transverse energy in pp collisions at sqrt(s) = 7 tev. *Physics Letters B*, 716(2):260 – 284, 2012.

[82] Matthias Schröder and the CMS collaboration. Performance of jets at cms. *Journal of Physics: Conference Series*, 587(1):012004, 2015.

[83] A.M. Sirunyan and etal. Particle-flow reconstruction and global event description with the cms detector. *Journal of Instrumentation*, 12(10):P10003, 2017.

[84] D. N. Spergel, R. Bean, O. DorÃ©, M. R. Nolta, C. L. Bennett, J. Dunkley, G. Hinshaw, N. Jarosik, E. Komatsu, L. Page, H. V. Peiris, L. Verde, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. Three-year wilkinson microwave anisotropy probe (wmap) observations: Implications for cosmology. *The Astrophysical Journal Supplement Series*, 170(2):377, 2007.

[85] A. Strandlie and W. Wittek. Propagation of covariance matrices of track parameters in homogeneous magnetic fields in CMS. 2006.

[86] M Strassler. B-tagging: Identifying jets from bottom quarks. `https://profmattstrassler.com/articles-and-posts/particle-physics-basics/the-known-apparently-elementary-particles/jets-the-manifestation-of-quarks-and-gluons/b-tagging-identifying-jets-from-bottom-quarks/`, March 2012.

[87] F. Strocchi. Spontaneous symmetry breaking in quantum systems. *Scholarpedia*, 7(1):11196, 2012. revision #138826.

[88] AC Team. The four main LHC experiments. Jun 1999.

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

[90] Steven Weinberg. The cosmological constant problem. *Rev. Mod. Phys.*, 61:1–23, Jan 1989.

[91] Steven Weinberg. *The Quantum Theory of Fields*, volume 3. Cambridge University Press, 2000.