SEARCH FOR HEAVY GAUGE BOSONS DECAYING TO TAU LEPTONS USING VECTOR BOSON FUSION PROCESSES IN PROTON-PROTON COLLISIONS AT CMS

By

Susan Catherine Olmsted

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Physics

May 10, 2024

Nashville, Tennessee

Approved:

Alfredo Gurrola, Ph.D.

Julia A. Velkovska, Ph.D.

Will E. Johns, Ph.D.

Paul D. Sheldon, Ph.D.

Norman H. Tolk, Ph.D.

Copyright © 2024 Susan Catherine Olmsted
All Rights Reserved

ACKNOWLEDGMENTS

I would like to acknowledge the support of my husband, Dakota Foster, and that of my cats, Chonk and Pancake.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

One primary goal of particle physics is to explain the fundamental physics of the universe: that is, to analyze the properties of particles and the forces by which they interact. Previously, physicists have used a theoretical framework called the Standard Model (SM) to describe these particles and force interactions. In many ways, the SM has been extremely successful, as it has proven to be consistent with many experimental observations. However, the SM fails to explain several recently discovered phenomena. Examples of these shortcomings include the absence of a candidate for dark matter (DM), the failure to account for matter-antimatter asymmetry, and the lack of explanation for force hierarchy asymmetry. Therefore, theoretical extensions to the SM, called beyond standard model (BSM) frameworks, have been introduced to offer solutions to these problems. Although the motivations and implications of these models can vary significantly, a common characteristic is the manifestation of new particles with TeV-scale masses. Such high-mass particles can be probed in proton-proton (pp) collisions at high-energy particle accelerators.

The Large Hadron Collider (LHC) is the world's largest and most powerful particle accelerator. Located near Geneva, Switzerland, the LHC aggregates the research of thousands of scientists, making it the largest experiment in the world. The collider consists of a 27-kilometer ring made of semiconductor magnets capable of producing the high-energy collisions necessary to produce the particles theorised to solve current physics problems. My research uses data collected from the Compact Muon Solenoid detector, which is located at one of several collision points along the LHC.

The following thesis will describe my search for a BSM TeV-scale theoretical particle called the Z' (*Z-prime*) boson via a rare particle interaction called Vector Boson Fusion. This search used data samples with an integrated luminosity of 137.1 fb^{-1} with pp collisions at $\sqrt{s} = 13$ TeV. This methodology had thus far not been attempted at a particle collider. Like the Z boson of the SM, a Z' is produced via interactions involving the weak force.

In this thesis, I will discuss the following topics in their respective chapters: background elementary particle physics information (Chapter 2); a description of physics beyond the standard model and an examination of previous research on Z' searches (Chapter 3); an in-depth description of the LHC and the CMS (Chapter 4); an explanation of the apparatus's triggers and our method of physics object reconstruction (Chapter 5); the analysis strategy for my channel of Z' decay (Chapter 6); the other considerations involved in my analysis (Chapter 7); and the results from my analysis (Chapter 8). Then, I will conclude by discussing the outcomes of my research (Chapter 9), including its impact on particle physics.

CHAPTER 2

Background Information: Particles and Particle Interactions

2.1 The Standard Model

Elementary particle physics studies the fundamental constituent particles of the universe and the fundamental forces that govern them. Our current understanding of particle physics is described by the Standard Model (SM). Although the SM does not account for every observed phenomenon in the universe, it does predict every confirmed elementary particle discovered so far. The SM is based on Quantum Field Theory, which is the framework that combines classical field theory, quantum mechanics, and special relativity.

The SM describes two types of elementary particles: bosons and fermions. Further, the interactions between these particles are governed by four fundamental forces: the gravitational force, the electromagnetic force, the weak force, and the strong force. The SM describes many physical phenomena and makes predictions consistent with most experimental results.

Fermions are distinct from bosons in that they have half-integer spins and therefore must obey the Pauli Exclusion Principle, which states that no two fermions can occupy the same quantum state. Fermions can be broken into quarks and leptons. Quarks are unique in that they have color charge and thus cannot exist on their own; they must exist in a bound state with other quarks or antiquarks. Leptons consist of the electrically charged leptons– electrons, muons, and taus– and their corresponding neutrinos (electron neutrinos, muon neutrinos, and tau neutrinos, respectively), which are neutral, extremely light particles.

Bosons have integer spins and are characterized mainly by the forces that they mediate. Bosons include the photon, which mediates the electromagnetic force; the W bosons and the Z boson, which mediate the weak force; and gluons, which mediate the strong force. The final boson, called the Higgs boson, is unique in that it does not govern a force, but exists as an excitation of the Higgs field, which allows matter to have mass.

The following sections discuss these types of particles in more detail, with a dedicated section for the Higgs boson, as its nature as a scalar boson leads to distinct consequences, and a section dedicated to the tau lepton and its decay modes, as these topics are particularly important for this study.

2.2 Fermions: Matter Particles

All particles that make up matter are considered fermions. Fermions come in two types: (1) quarks, which make up composite particles such as protons and neutrons, and (2) leptons, which consist of, for instance, electrons. Both types of fermions are further broken into “generations” ranging from one to three, with each

Table 2.1: List of forces, their relative strengths (normalized to that of the gravitational force), and the particles that carry them. Note that the graviton (g) is *not* an SM particle.

Fermions				
Generation	**Particle**	**Charge (in terms of e)**	**Spin**	**Mass (MeV/c^2)**
1	up; u	$+2/3$	$1/2$	~ 2.2
1	down; d	$-1/3$	$1/2$	~ 4.6
2	charm; c	$+2/3$	$1/2$	~ 1280
2	strange; s	$-1/3$	$1/2$	~ 96
3	top; t	$+2/3$	$1/2$	~ 173100
3	bottom; b	$-1/3$	$1/2$	~ 4180
1	electron; e	-1	$1/2$	0.511
1	electron neutrino; ν_e	0	$1/2$	$< 2.2 \times 10^{-6}$
2	muon; μ	-1	$1/2$	105.7
2	muon neutrino; ν_μ	0	$1/2$	< 0.170
3	tau; τ	-1	$1/2$	~ 1776.9
3	tau neutrino; ν_τ	0	$1/2$	< 15.5
Bosons				
SM?	**Particle**	**Charge (in terms of e)**	**Spin**	**Mass (MeV/c^2)**
Yes	photon; γ	0	1	0
Yes	W boson; $W^\pm$	$\pm$ 1	1	~ 80.4
Yes	Z boson; Z	0	1	~ 91.2
No	gluon; g	0	1	Unknown [55]
Yes	Higgs boson; H_0	0	0	~ 125.1

subsequent generation in general having more mass (though it should it be noted that the neutrinos' masses have not yet been measured, so this tendency is unconfirmed for neutrinos). Additionally, higher generations tend to have shorter lifetimes, although neutrinos can freely oscillate between one another (and in fact exist as a superpostion of all three types). For example, the muon, which is a second-generation lepton, is heavier and has a shorter lifetime than the first-generation lepton, the electron, does. According to the SM, quarks and leptons each come in six types, with two particles represented by each generation. There could, theoretically, be more generations of fermions.

The six types of quarks described by the SM are as follows: up (u), down (d), charm (c), strange (s), top (t), and bottom (b). The first generation of quarks are u and d, the second are c and s, and the third are t and b. To make up a composite particle, quarks must have a sufficiently neutral "color charge." Just as atoms must have an equal number of negatively-charge electrons and positively-charged protons to achieve a neutral electric charge for the atom overall, composite particles must consist of quarks that, together, neutralize the color charge. Each type of quark comes in three colors: red, green, and blue. It is important to note that these particles do not physically have color in the colloquial sense, but that color is used as an analogy to represent the phenomenon of several particles balancing one another out, similarly to how multiple colors of light can

balance out to white.

The six leptons, grouped with their subsequent generations, are electrons e and electron neutrinos (ν_e; muons μ and muon neutrinos ν_μ; and taus τ and tau neutrinos ν_τ). Taus in particular are important for this thesis, as the analysis strategy in this thesis focuses on Z' decaying to a pair of taus. Leptons are involved in weak interactions, and e, μ, and τ are involved in electromagnetic interactions. Neutrinos are neutrally charged, while e, μ, and τ particles are negatively charged. All electric charge is normalized to that of the electron, and μ and τ have electric charges equal to e. Further, e, μ, and τ particles have increasing masses with each generation, while neutrino, according to the SM, are massless. Neutrino masses have, however, been measured to be greater than zero (but still extremely light), representing a flaw in the SM.

Additionally, matter particles have oppositely-charged antimatter particle equivalents. For instance, the antiparticle equivalent of an electron is the positron, which has the same mass and spin of an electron but has a charge of about 1.6×10^{-19} Coulombs (as opposed to about -1.6×10^{-19} for electrons). Some particles without charge– for instance, the photon– are considered to be their own antiparticle. Antineutrinos have been found to have exclusively right-handed spins (the opposite neutrinos, which exclusively have left-handed spins) [54]. Additionally, a particle and its antiparticle may collide into one another to produce another particle, such as when an electron and a position annihilate to produce two photons. Antiparticles are generally represented by writing a bar over the symbol for its respective standard matter particle (i.e. an antiup quark is written $\bar{u}$).

Since quarks cannot exist on their own, they bind together via the strong force in composite particles called hadrons. As mentioned, this resulting hadron must be colorless. The neutral color charge is accomplished either by three quarks binding together (one "red," one "blue," and one "green") *or* by a particle and an *antiparticle* binding together (for instance, one "green" and one "antigreen" quark). A particle made of three quarks is called a baryon and has a half-integer spin; a particle made of a quark and an antiquark is called a meson and has an integer spin. An example of a common baryon that is important for this analysis is the proton, which is made of two up quarks and a down quark and is used in LHC collisions. An important meson for this analysis is the pion (π), which comes in three types: a positively charged pion π^+ made up of an up quark and an antidown quark ($u\bar{d}$)); a neutral pion π^0 made up of a mixture of up, down, antiup, and antidown quarks ($\frac{u\bar{u}-d\bar{d}}{\sqrt{2}}$); and a negatively-charged pion made of a down quark and an antidown quark ($d\bar{u}$). Pions are important in the hadronic decay of taus, which will be discussed in a later section. A visualization of the types of fermions and how they relate to bosons and fermions is shown in Figure hadrons.

Table 2.3 lists the most common possible decay modes of the τ lepton, along with their respective branching fractions (percent chance that the τ will decay to that mode).

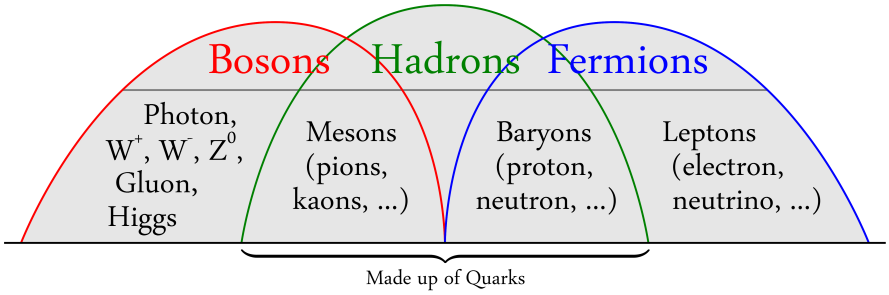


Figure 2.1: This figure shows the relationship between the types of hadrons, bosons, and fermions.

2.3 Gauge Bosons: Force Carrying Particles

As mentioned, bosons are force-carrying particles that have integer spins. The four gauge bosons are photons (γ) W bosons (which are represented by W^+ or W^- depending on the charge), Z bosons, and gluons (g). Photons propagate the electric force, giving particles charge. The $W^{+/-}$ and Z bosons propagate the weak force, which holds atoms together and is responsible for nuclear decays. Gluons propagate the strong force, which is responsible for keeping hadrons together and is responsible for color charge. Photons, Z bosons, and gluons have neutral electric charge, while W bosons have electric charge equal to that of an electron. Photons are massless, and gluons have an unknown mass with upper bounds of 1 MeV, 20 MeV, or 10^-10 MeV, depending on the experiment [55].

The four fundamental forces– listed in order of increasing strength– are the gravitational force, the electromagnetic force, the weak force, and the strong force. These forces, along with their corresponding strengths, range, and associated particles, are listed in Table 2.2 below. It should be noted that the SM accounts for neither the gravitational force nor the hierarchy of forces (i.e. their increasing strength), representing a shortcoming of the SM.

Table 2.2: List of forces, their relative strengths (normalized to that of the gravitational force), and the particles that carry them. Note that the graviton (g) is *not* an SM particle.

Force	**Relative Strength**	**Range**	**Particle**
Gravitational	10^0	∞ m	Graviton (G); Non-SM
Weak Force	10^{32}	10^{-18} m	W^+, W^-, and Z^0
Electromagnetic	10^{36}	∞ m	Photon (γ)
Strong Force	10^{38}	10^{-15} m	Gluons (g)

The comparatively small range of weak force is explained by the Yukawa interaction. The $W^\pm$ and Z bosons have very large masses. The Electromagnetic and gravitational forces' infinite ranges are thus due to the photon's masslessness and the gluon's extremely small mass. Conversely, the gluon's short range despite

its masslessness is explained by the Yukawa interaction and is due to the energy that would have to be used to overcome the strong force and separate quarks. The energy used to do this is so strong that it produces more quarks and antiquarks in the process, which bind to other quarks and produce more hadrons. The hadrons that are made up of these particles have a size of about 10^{-15} meters, explaining the force's range.

The electromagnetic force and the weak force first merged in 1961 as part of the Glashow-Weinberg-Salam (GWS) theory of electroweak processes [33]. The GWS theory was the first to predict the three vector bosons that carry the weak force: two charged bosons called the W^+ and W^- bosons, and one neutral boson called the Z boson. The theory also allowed the masses of these new particles to be calculated, with $M_W \approx$ 80 GeV and $M_Z \approx 90$ GeV. These masses are extremely heavy; for comparison, the rest mass of the proton is only about 0.938 GeV. The confirmation of the weak bosons and their masses demonstrates the validity of the GWS theory.

Through the GWS theory, the W and Z bosons are understood to mediate the weak force. Similarly, photons mediate the electromagnetic force. Comparatively, the weak force is only 10^{-4} as strong as the electromagnetic force, which is why it is considered "weak." However, this relative weakness can be attributed to the large mass of the intermediate vector bosons– *not* the intrinsic strength of the force. After all, the weak fine structure constant is about 1/29, which is larger than the electromagnetic fine structure constant (1/137). The weak force, therefore, has a *stronger* intrinsic coupling than the electromagnetic force. This discrepancy arises because the denominator of the propagator ($q^2 - M_W^2$) is very large at energies much lower than the W boson mass, resulting in a relatively weak force strength. However, since new colliders can reach higher energies, it is now possible to probe regimes in which the weak force dominates, including the BSM physics space with TeV-scale Z' bosons.

2.4 Higgs Boson: The Scalar Boson

The Higgs boson is a unique boson in that, although its spin is still an integer, the spin is 0 instead of 1, giving it distinct properties. The Higgs boson, referred to as a "scalar boson," was discovered in 2012, [2, 9] providing undeniable evidence of the Higgs mechanism. The Higgs mechanism was originally introduced to the SM to account for symmetry breaking.

Symmetry is an essential component of both the SM and in particle physics in general: under mathematical transformations, the feature represented by the symmetry must be preserved. For instance, charge symmetry refers to the consequence of the SM that, when the charge operator is exacted on a set of particles, the result is a set of particles identical to the input set in every way, except that they have opposite charge. These oppositely-charged particles exist for every charge-carrying particle in the SM and are called antiparticles.

Different symmetries are represented by different unitary groups ($U(n)$, where n is the number of degrees in the vector). The simplest unitary group, $U(1)$, refers to the rotational symmetry of a circle: that is, the field has angular symmetry and remains the same ("invariant") when the field undergoes a phase transformation.

Special unitary groups ($SU(n)$) are subgroups of $U(n)$ that are invariant. $S(n)$ groups can be used to represent the symmetry of the fundamental forces: $SU(2)$ represents symmetry in the electroweak interaction and $SU(3)$ represents color symmetry. Thus, the entire standard model can be summed up as the cross product of these groups: $SU(3)_C \times SU(2)_L \times U(1)_\Gamma$. This calculation, however, results in symmetry breaking: that is, the resulting cross product is not identical to the input. Therefore, the Higgs mechanism was introduced to account for this symmetry break.

In the SM, the group $SU(3)$ is generally written as $SU(3)_C$ to represent the "color gauge group." This group contains quarks are those are the particles that can have color charge. Quarks are represented in matrix form as follows:

$$q = \begin{pmatrix} q_r \\ q_g \\ q_b \end{pmatrix} \tag{2.1}$$

where q_r, q_g, and q_b represent quarks with red, green, and blue color charge respectively.

The $SU(2)$ group represents the weak isospin group and contains both leptons and quarks. However, the particles are represented differently whether they are left- or right-handed: left-handed versions of the particles are represented as doublets separated by generation, and right-handed particles are represented as singlets. Thus, quarks and leptons are represented as follows under the $SU(2)_L$ group:

$$Quarks \rightarrow \begin{pmatrix} u \\ d \end{pmatrix}_L \begin{pmatrix} c \\ s \end{pmatrix}_L \begin{pmatrix} t \\ b \end{pmatrix}_L u_R d_R c_R s_R t_R b_R \tag{2.2}$$

$$Leptons \rightarrow \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L e_R \mu_R \tau_R \tag{2.3}$$

where L represents the left-handed particles and R represents the right-handed particles. Note as well that this definition means that there are no right-handed neutrinos.

Because the SM states that physical states must be invariant under space-time phase transformations, the mediators of particle interactions must be massless. However, we know that some bosons (Z, $W^\pm$ do have mass, leading to a break in $SU(2)_{L)\times U(1)_\Gamma}$ symmetry. To solve this issue, physicists introduced the concept of the "Higgs mechanism." This concept allows particle interactions to couple to mass, resolving the broken symmetry. The Higgs mechanism operates in a vacuum state called the Higgs field and is mediated by the

Higgs boson.

2.5 Tau Lepton

2.5.1 Tau Decay Modes

Due to its high mass, τ is the only lepton that can decay both leptonically and hadronically, meaning that it can decay to either leptons or hadrons. About 35% of tau decays are lepton decays, while the other 65% of taus decay hardonically. Figure 2.2 demonstrates possible decay modes for τ.

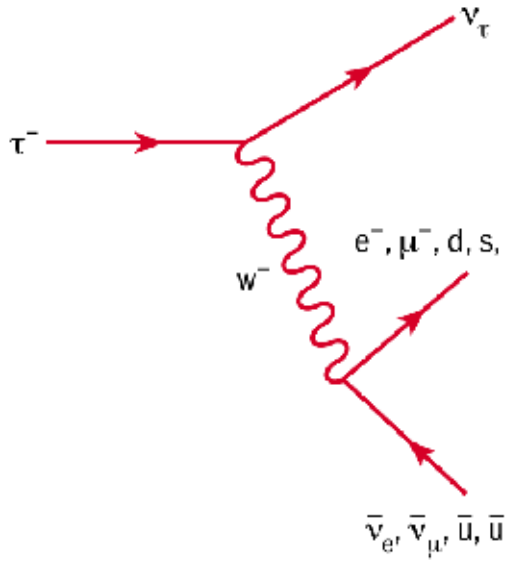


Figure 2.2: This figure shows possible decay modes for the tau lepton. Note that the end states include both hadronic and leptonic decays.

Note that the τ emits a W^- boson due to its relationship to the electroweak force, and that this boson must be the same charge as the τ to conserve charge. The W^- then decays to oppositely-charged fermions. When the τ emits the W^- boson, it must also decay to a ν_τ to conserve the concepts of "tau lepton number." Conservation of lepton number means that the decay products of an interaction must have the same lepton number value as the interaction began with. Here, τ^- and ν_τ have a tau lepton number of +1; μ^- and ν_μ have a muon lepton number of +1; and e^- and ν_e have an electron lepton number of +1. On the other hand, each particle's antiparticle has the respective lepton number of -1. An example of this phenomenon is shown in the decay of the a π^- described by the following equation:

$$\pi^- \rightarrow \mu^- + \bar{\nu}_\mu \tag{2.4}$$

where the antineutrino $\bar{\nu}_\mu$ balances out the muon electron number of the μ.

When a τ decays to quarks, those quarks bind to become hadrons (hence "hadronic" decay). This hadronic

decay is useful for identifying tau because, due to the tau's extremely short lifetime (about 2.9×10^{-13} seconds) and the detector's distance from the particle collision's primary vertex (about one meter), the tau can *only* be detected via its decay products, and the leptonic tau decay is impossible to distinguish from the direct production of electrons and muons without the tau. Hadronically decaying taus are written as τ_h. Table 2.3 lists the most common possible decay modes of τ_h, along with their respective "branching fractions," which is the percent chance that the τ_h will decay to that mode.

Table 2.3: This table lists decay modes of the τ lepton and their branching fractions

Decay Modes	**Branching Fraction**
$\pi^+\pi^0\nu_\tau$	25.4%
$e^-\nu_e\nu_\tau$	17.8%
$\mu^-\nu_\mu\nu_\tau$	17.4%
$\pi^-\nu_\tau$	11.1%
$\pi^-\pi^-\pi^+\nu_\tau$	9.5%
$\pi^-\pi^-\pi^+\nu_\tau$	9.2%
$\pi^-\pi^-\pi^+\pi^0\nu_\tau$	4.4%
$\pi^-\pi^0\pi^0\pi^0\nu_\tau$	1.1%

2.5.2 Hadronic Jets

To detect τ_h, then, one may look for the decay products typical of τ_h decay. However, a challenge arises due to the similarity between τ_h decay signatures and the decay signatures of quarks and gluons. This is because the partons (parts of protons) in particle collisions produce hardons via a process called "hadonization." This section describes the hadronization process.

As mentioned, protons are comprised of two u quarks and a d quark. However, high-energy collisions show that the makeup of protons is more complex than this would suggest. The bare quarks in a proton are held together by virtual gluons and virtual quarks. Virtual particles have similar characteristics to those of their non-virtual counterparts, but they exist for an extremely small amount of time as determined by the uncertainty principle. The virtual gluons and quarks (which can exist in the form virtual mesons) propagate the strong nuclear force, holding protons together. These virtual particles, along with the bare quarks, all make up the proton's partons. Figure 2.3 shows a representation of the partons in a proton, demonstrating just how many partons a proton can contain. Due to the virtual particle's extremely short existence, they do not contribute to the mass of the proton; however, the partons' spins do contribute to the total spin of the proton.

When protons collide at high energies, the partons within those protons interact and generate the hard scattering events that physicists aim to study. These events have the energy of the partons involved in the

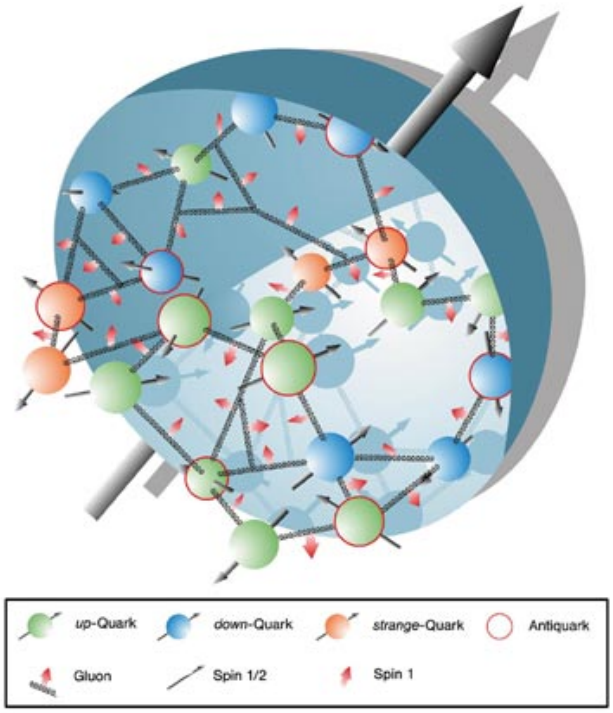


Figure 2.3: A representation of the partons in a proton.

event rather than the energy of the colliding protons, making the spectrum of energies we can examine much broader than we otherwise could. The disadvantage to this phenomenon is that the large number of hadrons within a proton then produce "jets."

Figure 2.4 shows how parton decay can produce a shower of hadrons. Here, one can see how a quark can emit a gluon, which can then produce hadrons. The resulting hadrons are called hadronic jets and are very similar to the result of hadronically decaying taus.

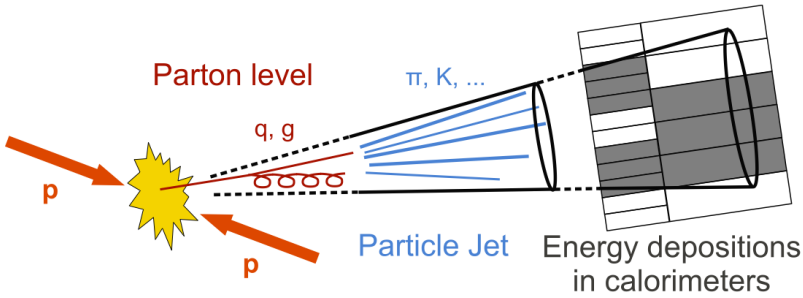


Figure 2.4: This figure shows how partons can produce a shower of hadrons.

Because the decay products of hadronically decaying taus resemble the products of hadronic jets, it is

necessary to examine the ways in which the processes are different, such as the transverse width of the structures. Since the width of a hadronic jet is related to the energy of the fragmenting hadrons, the width of the jet is relatively broad. Conversely, τ_h decays have a narrower, "pencil-like" due to their comparatively fewer and more-energetic final state particles. Figure 2.5 demonstrates the difference between the structures of hadronic jets and hadronically decaying taus.

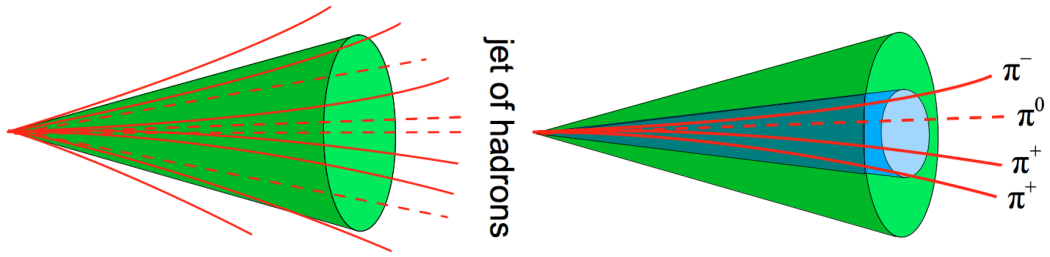


Figure 2.5: This figure shows the difference in the width of structures created by hadronic jets (Left) and hadronically decaying taus (Right).

Thus, isolating taus involves targeting this difference in decay structure. To identify τ_h, physicists define a narrow cone around the possible τ_h decay products and a larger cone with the width of hadronic jets. Decay structures that extend past the radius of the τ_h isolation radius are not considered to be τ_h candidates.

CHAPTER 3

Beyond Standard Model Physics

The following discusses theories that can account for the various shortcomings of the SM. These theories are considered "Beyond Standard Model (BSM)" theories in that they are meant to take the SM and expand it to correct for its observed shortcomings. Importantly, all of these models predict a Z' boson with high mass. This section also discusses the ways in which these theories have been tested.

3.1 Why Z'?

It is important to elaborate on the reasons why Z' is a good candidate for explaining SM shortcomings. Models that include Z' bosons can explain several inconsistencies between the predictions of the standard model and the results derived from observation. Below, I describe three motivations for Z'.

One motivation involves the calculation of the dark matter relic density, demonstrated by Figure 3.1. This figure shows that, in the high-energy environment of the early universe, dark matter was created and annihilated at approximately equal rates. As the universe expanded, the dark-matter-to-standard-matter processes began to freeze out. The theoretical prediction for the relic density depends on the properties of the dark matter particle but also on the mediators of dark matter interactions. Since dark matter is theorized to have strong coupling to the weak force, Z' could be a candidate for such a particle.

The second discrepancy is that of the R_D/R_{D^*} anomaly, which is summarized by Reference [7]. R_D and R_{D^*} are the ratio of branching factions, as defined below:

$$R_D = \frac{\Gamma(B \to D\tau\vec{\nu})}{\Gamma(B \to D(e/\mu)\vec{\nu})}, R_{D^*} = \frac{\Gamma(B \to D^*\tau\vec{\nu})}{\Gamma(B \to D^*(e/\mu)\vec{\nu})} \tag{3.1}$$

where B represents a meson with a b quark, D represents meson with a c quark, and D^* represents a meson with a D meson in an excited state. The SM predictions for R_D and R_{D^*} are 0.3 ± 0.011 and 0.254 ± 0.004 respectively, while the experimental results for R_D and R_{D^*} are about $0.397 \pm 0.040 \pm 0.028$ and 0.310 ± 0.017 respectively.

Figure 3.2 shows the latest results of the R_D/R_{D^*} anomaly, aggregating the results of several experiments in one plot. This plot shows the ratio of the branching fractions R_D and R_{D^*}, which contain an excess of semileptonic decays. Importantly, this ratio is sensitive to lepton flavor universality.

Finally, Z' can provide an answer for the notorious anomalous magnetic moment of the muon by provide additional contributions to the magnetic moment. The magnetic moment of a particle is associated with its

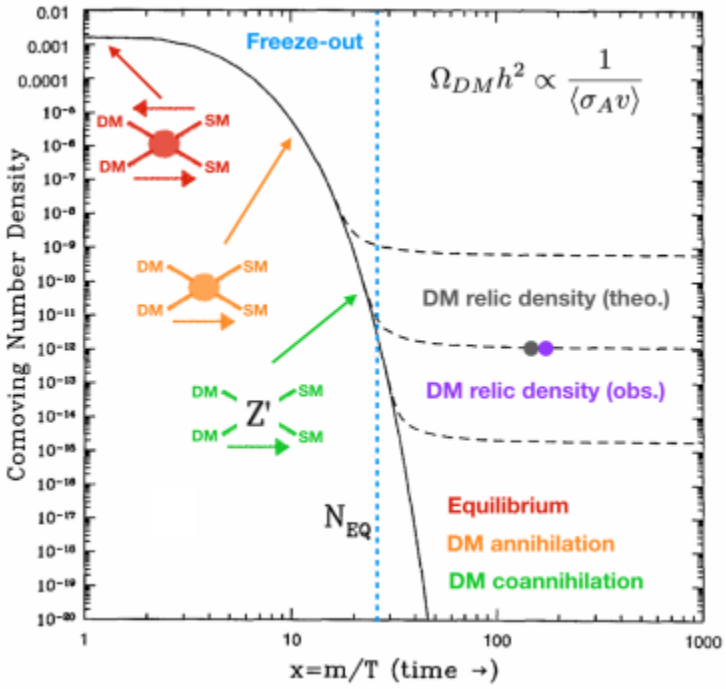


Figure 3.1: This plot shows the change of dark matter relic density over time.

g-factor, which indicated the strength the particle's own magnetic field and its rate of gyration in an external magnetic field. The SM predicts that the muon would have a g factor of 2, but the experimentally measured g is larger than 2, which is why it is called the muon g-2 anomaly to encapsulate this phenomenon. The muon magnetic anomaly $a_\mu = (g_\mu - 2)/2$ is measured to be about 116592061(41) $\times 10^{11}$, which is a 4.2 standard deviation difference from the SM prediction of 116591810(43) $\times 10^{-11}$ [52].

Magnetic moment is calculated by adding the contributions of all the possible processes that can occur. The Feynman diagram shown in Figure 3.3 shows one way by which a Z' particle could contribute to the muon's magnetic moment, providing an additional loop diagram and adding to the g factor. The electron magnetic moment, however, does *not* have a a discrepancy in experimental measurements. This suggests that the boson associated with extra Feynman diagram loops must either have an extremely small coupling to the electron, or none at all, meaning the boson must have non-universal couplings. Thus, a neutral electroweak gauge boson with non-universal couplings such as Z' could solve the phenomenon of the anomalous muon magnetic moment.

3.2 Z' Models

There are several models that predict heavy gauge bosons such as Z'. These heavy gauge bosons would be the gauge field of a new local broken symmetry. Examples of such models are ones with a new U(1) gauge

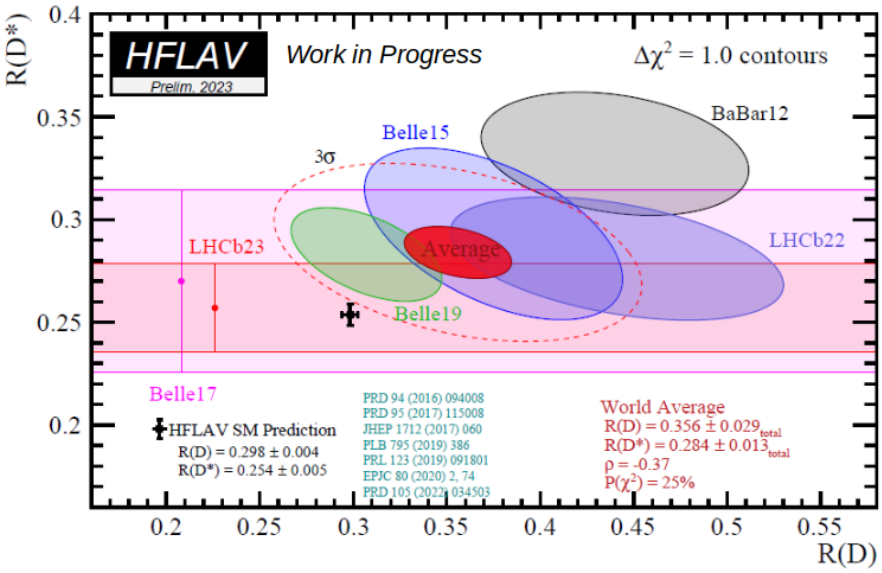


Figure 3.2: This plot shows the latest RD/RD* anomaly results, aggregating the work of several experiments. [4]

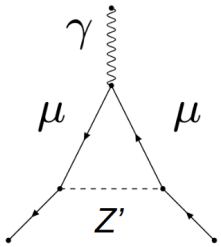


Figure 3.3: This is the Feynman diagram demonstrating the way by which Z' could contribute to the magnetic moment of the muon.

symmetry, Little Higgs models, and E6 Grand Unified Theories (GUTs).

In models with a new U(1) gauge symmetry, the Z' is the gauge boson of the broken symmetry. That is, the SM is extended by including an extra U(1) gauge group to the theory, i.e. the model would look something like the following: $S(U)(3) \times SU(2) \times U(1)_1 \times U(1)_2$. The symmetry-breaking components of this cross product are the mathematical representation of the model's bosons, as a boson is needed to restore symmetry. Any model that includes a U(1) gauge symmetry extension will have, at minimum, an octet of gluons, the $W^{\pm}$ bosons, the Z boson, the photon, and the Z' boson [5].

Little Higgs models also extend the SM with additional gauge groups. This family of models extend the SM in a way that the breaking of the global symmetry by gauge and Yukawa interactions generates Higgs mass and couplings at the TeV scale. Due to the symmetry inherent to the models, the additional Higgs mass

terms cancel; however, one or more Z' bosons are left behind. [42].

Kaluza-Klein (KK) models are characterized by an additional spacial dimension. The Z' bosons predicted by these models are excited states of a neutral, bulk gauge symmetry. [43]

Due to the breadth, scope, and implications of these models, it clear that the Z' particle is an important BSM particle candidate to investigate.

3.3 Established Research and Hypothesis

As mentioned, models attempting to address the incompleteness of the SM propose a heavy spin-1 neutral gauge boson (Z'). The typical discovery signature of a Z' is its decay to a pair of oppositely-charged leptons with no additional activity in the detector. In particular, if the Z' has enhanced couplings to third-generation fermions, then $Z' \to \tau\tau$ channel is an important mode for discovery.

Traditional searches for the Z' boson mainly target production via Drell-Yan (DY) processes of order α_{EW}^1 (i.e. $q\bar{q} \to Z' + 0j/1j$). An example of DY production of Z' is shown in Figure 3.4. The DY searches necessitate a sizable coupling for the Z' to light quarks ($g_{Z'qq}$). One hypothesis as to why the Z' boson has eluded discovery in the traditional DY searches conducted thus far is that the coupling for Z' to light quarks is actually smaller than that required for sensitivity in the DY searches. Therefore, it is important to develop new search methods that do not depend on $g_{Z'qq}$.

Figure 3.4: This is the Feynman diagram demonstrating Drell-Yan production of Z'.

The focus of this thesis is a new search for Z' bosons using events produced through vector boson fusion (VBF) processes (see Figure 3.5). Figure 3.5 shows the Feynman diagrams that represent VBF production of Z'. The tagging of events produced though VBF processes has been proposed as an effective experimental tool for dark matter (DM), heavy Majorana neutrinos, and electroweak supersymmetry (SUSY) searches at the LHC [28, 31, 29]. In particular, it has been shown that VBF processes become increasingly more important for probing large mass scales due to collinear logarithmic enhancements in the production cross-sections. Furthermore, the VBF Z' production rates do not depend on $g_{Z'qq}$ and thus may be an important mechanism for discovery to complement the DY Z' searches where current sensitivity is limited (more on this later). The

distinct nature of the VBF topology allows for easier distinguishing between desired signal and background processes (in particular QCD processes), which is also an important characteristic that can be utilized in Z' searches characterized by large QCD backgrounds, such as $Z' \to \tau\tau$.

The ATLAS and CMS Collaborations have reported results of unsuccessful searches for $Z' \to \tau\tau$ resonances in proton-proton (pp) collisions at $\sqrt{s} = 7$, 8, and 13 TeV and have excluded Z' masses up to 2.1 TeV [3, 21, 1, 47]. Those analyses investigated Z' production via DY processes of first order in the electroweak coupling (i.e. $q\bar{q} \to Z' + 0j$), with high signal acceptance to produce a "bump" in the reconstructed ditau invariant mass spectrum against a background distribution that falls steeply at higher masses. It is possible that these searches were unsuccessful in discovering the Z' boson because the DY process requires Z' to couple to the quarks relatively strongly; that is, the studies assumed that Z' coupling to quarks would be same value as the SM Z boson coupling to quarks. If the Z'-q coupling is 0.25 times the Z-q coupling, the upper bounds of Z' mass was found to be 0.5 TeV. Thus, this study examined a different discovery signatuve that assumes lower values for Z' coupling to quarks, which increased the discovery potential for Z'.

Additionally, the analysis assumes that Z' couples to the SM weak bosons. This makes VBF Z' production an important production mechanism with a unique topology: VBF production is identified by requiring events to contain a pair of jets that (1) have TeV scale dijet mass, (2) are widely separated (i.e. they have relatively high pseudorapidity), and (3) exist in opposite hemispheres of the CMS detector. The dijet pair boost the momentum of the Z' due to momentum conservation, which also improves sensitivity in the search for heavy Z' by helping the signal to stand out from the background. The analysis is also motivated by models with enhanced Z' couplings to third-generation fermions.

Further extensions to the SM that were proposed as an explanation for the high mass of the top quark also predict Z' bosons that couple more strongly to third-generation fermions than to the first or second generation fermions. Thus, this study focused on the $Z' \to \tau\tau$ search channels (Figure 3.5, Left). Our analysis team examined four $\tau\tau$ final states: $\tau_e\tau_\mu$, $\tau_e\tau_h$, $\tau_\mu\tau_h$, and $\tau_h\tau_h$, where τ_ℓ ($\ell = e, \mu$) and τ_h refers to the leptonic and hadronic decay modes of the τ lepton, respectively. Since the leptonic τ decays are indistinguishable from prompt production of electrons and muons, these four channels were studied using the $\ell\tau_h$, $e\mu$, and $\tau_h\tau_h$ decay modes. Additionally, since we assume the Z' couples to the SM vector bosons, the $Z' \to WW$ decay width (and therefore branching ratio) can be large, which further motivated the use of these channel to interpret the results for the $Z' \to WW$ scenario in which both W bosons decay to a lepton and neutrino (Figure 3.5, Right). Note that my portion of the analysis was the $\tau_h\tau_h$ channel, while the results of the other channels are published in Reference [35]. All four channels are used in the results section of this thesis for comparison.

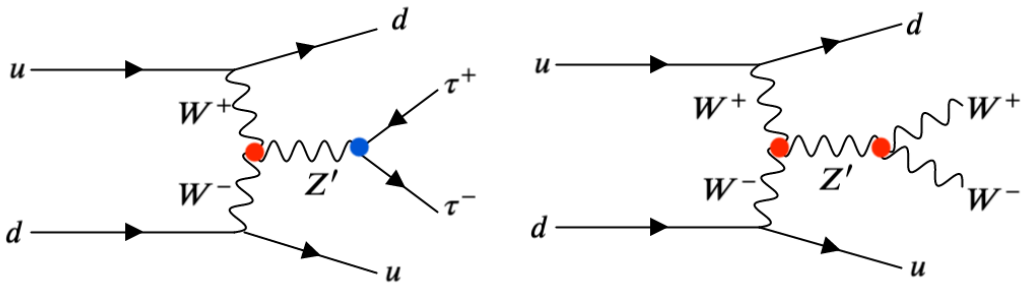


Figure 3.5: Representative Feynman diagrams for the Z' production via VBF processes, followed by decays to $\tau^+\tau^-$ (Left) and W^+W^- (Right). Each red solid circle represents the vertex for Z' coupling to bosons while the blue solid circle the vertex for Z' coupling to fermions.

3.4 VBF Z' model

The Vanderbilt Experimental High Energy Physics (VEHEP) group published a theory paper discussing other possible modes for producing a Z' through vector boson fusion (VBF) processes [32], which would be the best discovery mode for Z' if the Z'-q couplings are small (see Figure 3.6). VBF is a purely electroweak process characterized by two SM vector bosons (radiated from quarks involved in pp collisions) fusing into a new heavy Z' particle before decaying into two leptons. Now that the LHC is probing higher pp collision energies and acquiring data at higher integrated luminosities necessary for rare processes such as VBF, we have the opportunity to search for Z' production via VBF using pp data acquired by the CMS experiment. For example, the VBF Z' production cross-section is about 0.1 pb for m(Z') = 1 TeV. This would result in a VBF Z' production rate of about N = 1.363×10^4 events with the 137.1 fb^{-1} of pp data collected in 2016-2018. Previous feasibility studies suggest that it is possible to achieve a 5σ discovery reach for Z' masses below 1.5 TeV with an integrated luminosity of 137.1 fb^{-1}.

To allow for a broad and generic discussion about the importance of VBF Z' under various model assumptions, we consider a "simplified phenomenological approach" where the Z' mass and the Z' couplings are free parameters. Some differences with respect to the commonly used SSM Z' model is that the coupling $g_{Z'VV}$ to the heavy vector bosons ($V = W$ or Z) is allowed. Since the coupling is very model-dependent and we do not want to constrain this search, $g_{Z'VV}$ will be considered to be free parameter. The range will be constrained naturally by the limits of no $Z'VV$ coupling and a very broad Z' decay width. We define the coupling as a function of the maximal coupling, $g_{Z'VV} = \kappa_V g^{max}_{Z'VV}$, where $\kappa_V \leq 1$, to avoid unitarity violation. The dependence of the branching fraction with respect to the individual particles can be seen in Figure 3.7 (right) for one very specific subset of Z' models where the Z' couplings to SM fermions are SSM-like. We note that the branching fractions are model dependent parameters, which depend on every single coupling between Z' and all the SM particles. On one extreme, the $Z' \to \tau\tau$ branching fraction can be close to 100% if the Z' coupling

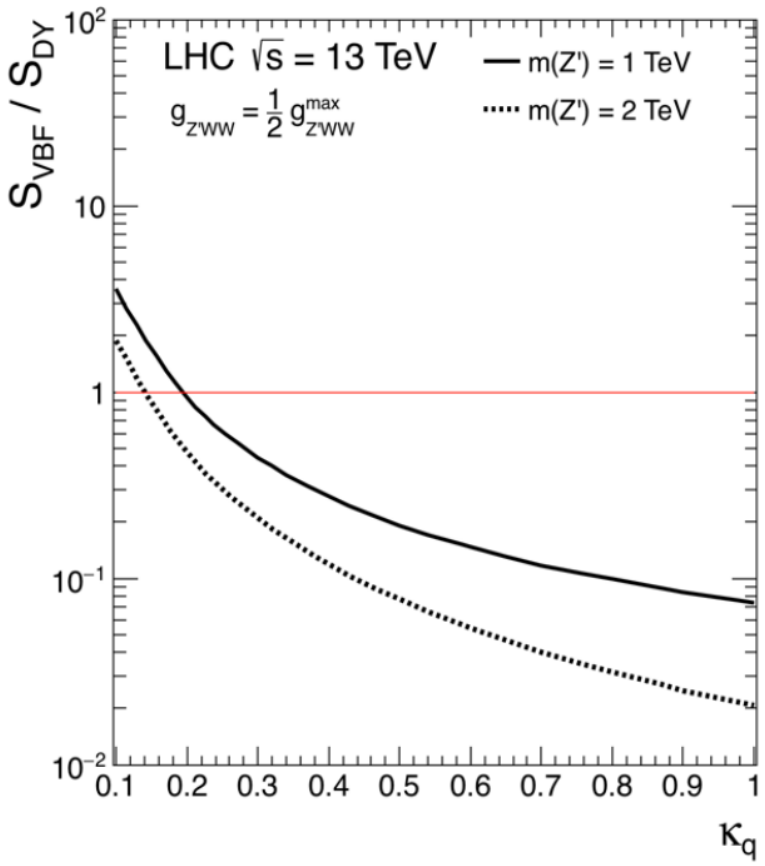


Figure 3.6: This figure the the ratio of the signal significance of VBF Z' to DY Z' as a function of κ_q [32]. Here, one can see that, for small κ_q, the ratio of the two signal significance values is higher.

to the SM weak bosons (κ_V) is relatively small with respect to the tau coupling, and the $Z' \to \tau\tau$ coupling is the most relevant fermion coupling. On the other hand, if the Z' is allowed to couple strongly to many other SM particles, such as the top quark, b-quark, etc., then the branching fractions can vary significantly. All of these scenarios are possible, and one can change the set of "good" models depending on the constraints one wants to impose (Yukawa hierarchy, explanation of the top mass, Higgs hierarchy problem, matter-antimatter asymmetry, etc.). These assumptions of the model only affect the $\sigma \times Br$ theory curve. Therefore, we present our results versus Z' mass, coupling, and branching fraction.

There is strong motivation to consider non-universal gauge couplings. For example, one simple model is the G(221) or topcolor model. The model splits fermions gauge groups into two SU_2: one for the light fermions of the first and second generation and one for the heavy third generation. The two groups mix and result in relatively unchanged SM gauge bosons and heavy gauge bosons. The small effects on the known measurements were extensively studied [10, 37, 30]. While the "light" bosons have a universal coupling to all fermions in this model, the new heavy gauge bosons would have a non-universal coupling. To parameterize

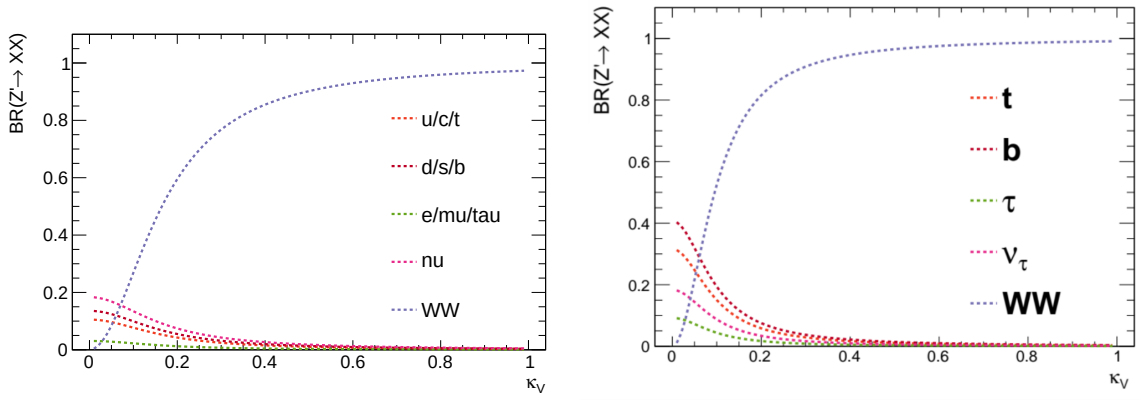


Figure 3.7: Branching fraction for a Z' as a function of κ_V. The fermion masses are neglected for simplicity. While the left side shows the branching fraction for universal couplings of the Z', the right side shows the branching fraction if no coupling to the light fermions is allowed.

this possibility, two couplings can be introduced: g_l and g_h, for the light fermions and the heavy fermions respectively. We use this simple parametrization to study a few key parameters and will define g_l and g_h as the deviation from the weak coupling. For smaller values of g_l, the branching fractions of the Z' change drastically, since the decay into light fermions is suppressed (see Figure 1 for one example scenario).

For the production and signal generation of a Z', we focus on the final state of one Z' with two jets. Figure 3.8 gives the decay width for a 1 TeV Z' for fully expressed couplings to the first and second generation fermions, while Figure 3.9 gives the decay width for a 1 TeV Z' for fully expressed couplings to the third generation fermions.

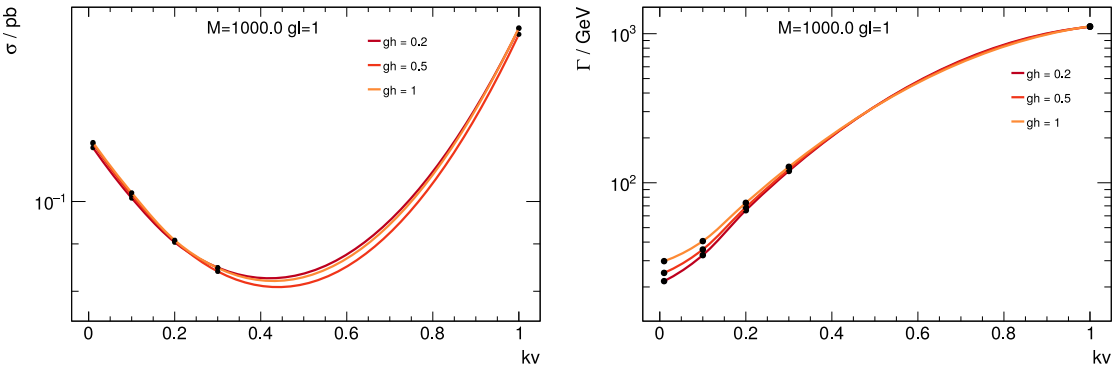


Figure 3.8: Production cross-section (Left) and decay width (Right) for 1-TeV Z' boson as a function of κ_V varying the coupling to the third generation fermions (g_h) at a fixed coupling to the first and second generation fermions ($g_l = 1$).

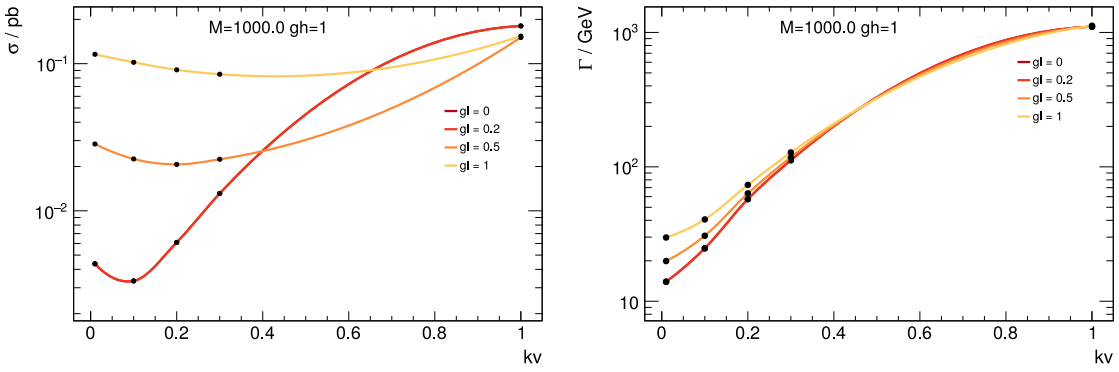


Figure 3.9: Production cross-section (Left) and decay width (Right) for 1-TeV Z' boson as a function of κ_V varying the coupling to the coupling to the first and second generation fermions (g_l) at at a fixed coupling to third generation fermions (g_h).

CHAPTER 4

Experimental Apparatus

4.1 Large Hadron Collider

The Large Hadron Collider is two 27-kilometer ring collider located near Geneva, Switzerland as part of the European Center for (CERN). The ring has cavities with alternating positive and negative electric fields and accelerate protons to speeds close to the speed of light. An overview of the detector (located 100 meters underground) is shown in Figure 4.1 (left), which demonstrates just how large the collider is. Additionally, a diagram (not to scale) shows the layout of the LHC in Figure 4.1 (right).

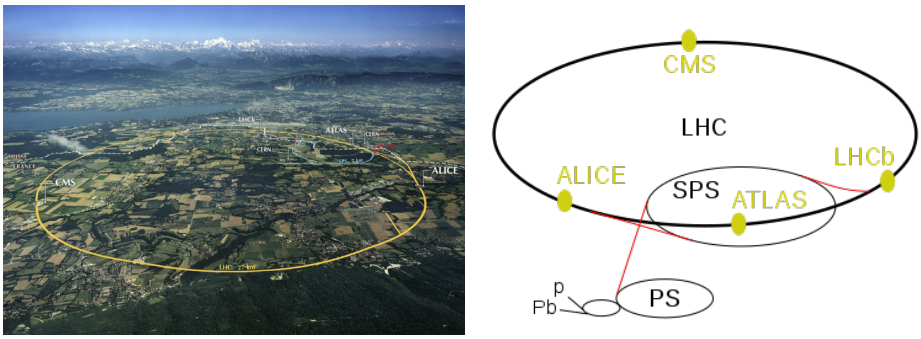


Figure 4.1: This figure shows two representations of the LHC. (Left) An aerial view of the LHC; (Right) A diagram of the LHC structures (not to-scale).

The LHC works by accelerating protons using radio waves and by steering the protons using a magnetic field. To acquire protons, hydrogen atoms are stripped of their electrons via an electric field. CERN engineers then switch the electric fields from positive to negative to pull the protons forward, switching the field at such a frequency that protons collect in bunches instead of a continuous stream. The protons then enter the radiofrequency cavities, which are metallic chambers in which radio waves interact with the passing particle bunches to give them energy, accelerating them along the beam pipe. Radio waves are required to attain the specific energy of a proton.

The collider's magnetic field must be extremely powerful– 8.3 Tesla, which is more than 100,000 times that of the Earth– because, as the particles become faster, they require more force to steer them along the ring. The magnets used in the machine are electromagnets with a superconducting coil that produces a current of 11,080 amperes. Dipole magnets are used to bend the particles along the ring, which quadrupole magnets focus the beams.

At four locations along the ring, the two sub-rings cross, allowing the opposite-direction protons to collide

into one another. These high-energy collisions can probe similar high-energy conditions of the early universe (according to Big Bang Cosmology). Additionally, the resulting collision from such high-energy particles can create relatively heavy particles, including previously unidentified ones. One of the four collision locations is the CMS, which is the detector that we will use for this project. The CMS detector measures positions, momenta, and energies of the resultant particles from the collision, allowing us to analyze the collision and search for new particles.

In particle physics, luminosity is the expected yield of events delivered per area per time. The luminosity can be measured in inverse barns per second ($b^{-1}s^{-1}$). The integrated luminosity (L) is the integral of the luminosity over the data-taking time frame. L is proportional to the probability that a particle interaction will occur.

The luminosity of a particular experiment is calculated using the following: $L = \frac{N}{\sigma}$, where N is the number of events and σ is the cross-section of the particle interactions. The cross-section of an interaction is a metric for the likelihood that the particle interaction will occur. In particle physics, the cross-section is usually measured in barns, where one barn equals 10^{-28} m^2. The cross-section of particle collision at the LHC is about 80 mb. Luminosity quantity is usually expressed in inverse femtobarns (fb^1) per second, where one barn is 10^{-28} m^2 per second. The intergrated luminosities for 2016, 2017, and 2018 are 35.9 fb^{-1}, 41.5 fb^{-1} , and 59.7 fb^{-1} respectively, making the full Run II luminosity 137.1 fb^{-1}.

The center-of-mass energy refers to the total energy of a collision in the center-of-mass reference frame. For Run II data, the center-of-mass energy was 13 TeV. This means that one can produce particles with a maximum total energy of 13 TeV.

Since each bunch crossing produces many particle interactions, multiple interactions are recorded at one time, called pile-up (PU) interactions. Resolving individual interactions from one another is a difficult but important aspect of analyzing LHC data. The number of pileup interactions has increased each year, so the effects of pile-up must be mitigated for each year individually.

The distributions of the average number of interactions per bunch crossing for each year are depicted in Figure 4.2. The minimum bias cross-sections σ_{in}^{pp} are also shown for the various luminosities. With so many interactions occurring per bunch crossing, we must keep these averages in mind when calculating the effects of pileup.

4.2 The Compact Muon Solenoid

The Compact Muon Solenoid (CMS) detector has a purposely descriptive name. The detector is 15 meters long and 21 meters high–which, for the amount of technology that it contains, is considered "compact." Further, the detector was created to specialize in the accurate detection of muons, which is evident in that

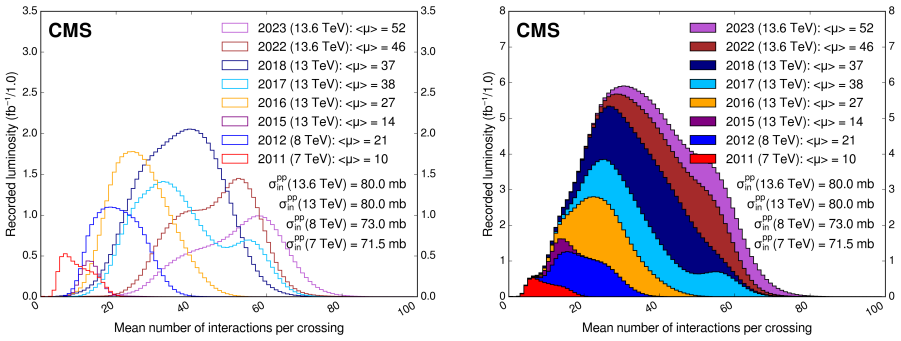


Figure 4.2: This figure depicts the mean number of parton interactions per bunch crossing, plotted per year individually (Left) and stacked atop one another (Right).

a large potion of the detector is dedicated to the muon chambers. Finally, the detector contains a powerful solenoid that produces a 3.8 Tesla magnetic field. This solenoid is useful for particle identification because it bends electrically-charged particles depending on their charges (positive or negative) and and their momenta. For instance, lower momenta particles that have charge will have higher-angle trajectories. Thus, this solenoid is necessary to determine information for particle identification. A representation of the CMS detector is shown in Figure 4.3.

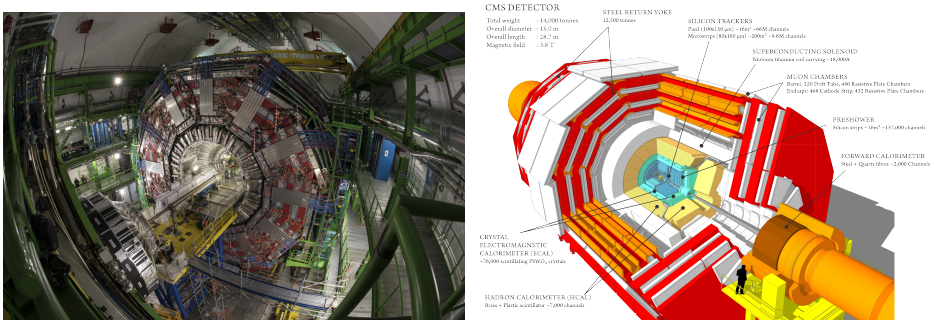


Figure 4.3: This figure shows two representations of the CMS detector. (Left) A real-life picture of the detector; (Right) A diagram showing a cross-section of the detector.

The CMS detector has four concentric sub-detectors: the tracker, the electromagnetic calorimeter (ECAL), the hadronic calorimeter (HCAL), and the muon chambers [26]. A diagram of these trackers is shown in Figure 4.4. The tracker is the inner-most sub-detector and collects data regarding a particle's path through the magnetic field, allowing us to calculate the particle's momentum. The particles that can be tracked in this sub-detector are muons, electrons, hadrons, and other fermion particles. As a charged particle passes through the silicon pixels, an electrical signal is produced, allowing the particle's position and angle from the z-axis (i.e. beam line) to be measured. Currently, the tracker only extends to pseudorapidity of 2.5, where pseudorapidity is defined as $\eta = -ln(tan(\frac{\theta}{2}))$. Here, θ represents the polar angle with respect to the z-axis. Thus,

the tracker is not present in the forward regions of the detector.

The goal of the ECAL is to measure the energies of resultant electrons and photons. The ECAL is made of heavy lead tungstate crystals, which are very dense (8.28 $\frac{g}{cm^3}$). The crystals are transparent and scintillate when electrons or photons pass through them, where to "scintillate" means to produce a photon burst proportional to the particle's energy. These crystals have high density and produce fast, short, well-defined bursts, allowing precise particle identification. Electrons and photons release all of their energy in this subdetector, for their masses are too low to pass all the way through.

Similar to the ECAL, the HCAL also measures the energy of incoming particles, only they measure the energy of hadrons instead of electromagnetic particles. The endcap and barrel components of the HCAL, which cover the pseudorapidity range $|\eta| < 3$, are comprised of alternating layers of absorber and scintillator materials, allowing the detector to measure the positions and energies of resultant hadrons. Extensive forward calorimetry (Hadron Forward Calorimeter, or HF) complements the barrel and endcap detectors by covering the pseudorapidity range $3.0 < |\eta| < 5.0$, which is a region that is not covered by the tracker. The HF subdetector is essential for measuring forward jets.

The muon chambers are the outermost and largest section of the detector, totaling 1400 individual muon chambers. Because these chambers are located beyond ECAL and HCAL, only heavier particles that were not stopped by the HCAL are able to penetrate through to the muon chambers. These chambers measure a particle's angle and position, allowing the particle's momentum to be calculated.

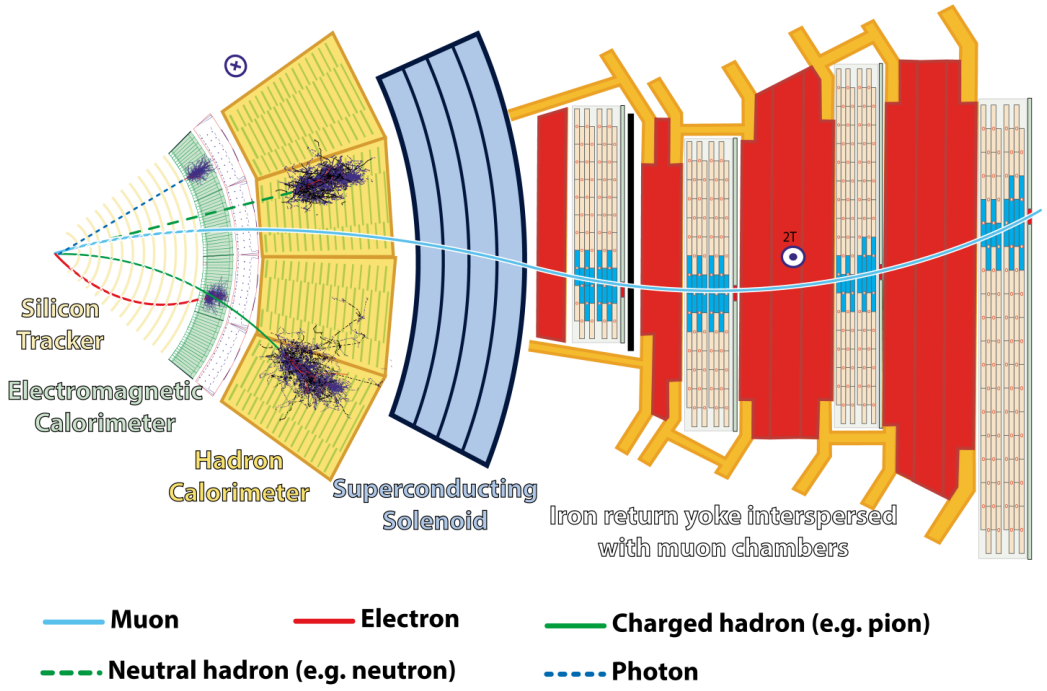


Figure 4.4: This figure shows how various particles move through the four CMS subdetectors.

4.3 Defining the CMS Coordinate System

Next, I will define important collider quantities that I will use in my data analysis.

A diagram of the CMS coordinate system is shown in Figure 4.5. Here, $\vec{p}$ refers to the momentum of a particular particle, while $\vec{p_T}$ refers to its transverse momentum. Transverse momentum (p_T) is the component of a particle's momentum that is in the the transverse direction, such that $p_T = \sqrt{p_x^2 + p_y^2}$. Since the pp system has total initial momentum of zero (i.e. the two proton beams are in opposite directions), the sum of the momentum vectors of particles produced in the collision must also equal zero.

The sum of these momentum vectors, however, does not equal zero. This is because some particle momenta cannot be detected by CMS, such as that from neutrinos. Thus, the missing transverse momentum vector ($\vec{p}^{miss}$) is the momentum in each event that is not accounted for by the detected particles. The magnitude of the transverse component of $\vec{p^{miss}}$ is denoted p_T^{miss}. The missing transverse energy of an interaction is called MET (sometimes also written as E_T^{miss}).

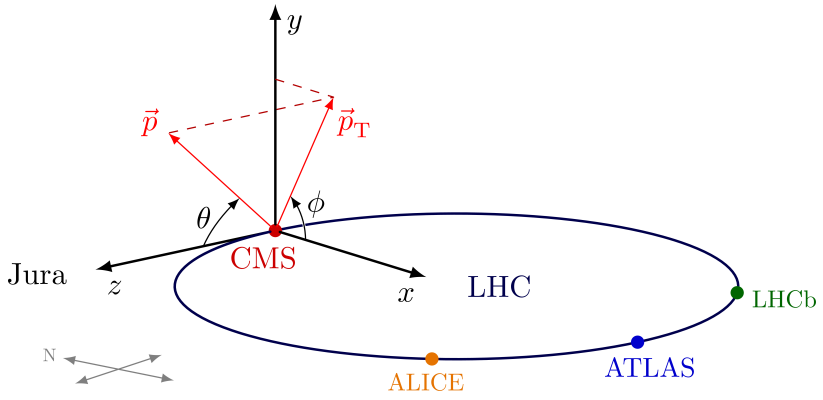


Figure 4.5: This figure shows the coordinate system used in the CMS detector. [40]

Pseudorapidity (η) is the coordinate that describes the angle between a particle and the particle beam, as shown in Figure 4.6 (Top). As you can see, η approaches ∞ as the angle approaches the beam line. η is calculated using the following equation:

$$\eta = -ln[tan(\frac{\theta}{2})]. \tag{4.1}$$

Values of eta for different angles of θ are shown in Figure 4.6 (Bottom).

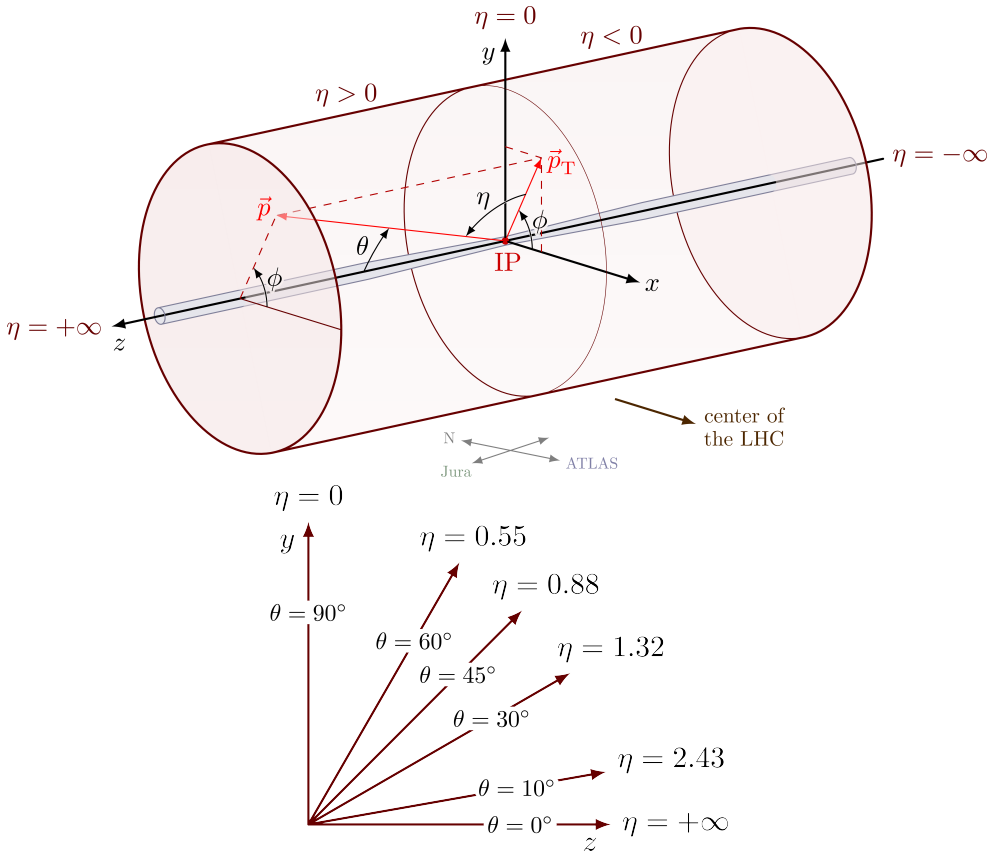


Figure 4.6: This figure shows a representation of pseudorapidity (η). (Top) η with respect to the CMS detector. (Bottom) Values of eta for various angles of θ. [40]

CHAPTER 5

Triggers and Physics Object Reconstruction

Now that I have discussed the physics underlying this analysis and the apparatus used to gather data, it is necessary to go into more detail regarding the specifics behind how the data is analyzed. First, I will explain how CMS event triggers work and discuss the specific triggers used in this analysis. Then, I will explain how those events are reconstructed to various physics objects.

5.1 Triggers

This thesis focuses on a search for VBF $Z' \to \tau\tau$, considering the channel in which both τ leptons decay hadronically ($\tau\tau \to \tau_h\tau_h$). Therefore, selecting important events for this analysis required the use of $\tau_h\tau_h$ (double hadronic tau) triggers.

Since the LHC creates billions of collisions per second, it is impossible to store the information produced from all the collisions. Therefore, it is important to focus on collecting information from the collisions that have the highest likelihood of contributing to the discovery of new physics. Thus, the LHC uses a system of triggers that make decisions regarding which particle collision information to record.

A proton bunch collides with another bunch every 25 nanoseconds, meaning the bunches collide at a rate of 40 megahertz. This means that, once one collision happens, another batch of particles will collide before the first batch passes all the way through the detector. Although the trigger system reduces the number of events recorded to about 100 events per second, it is still necessary to have a method of storing the data that accounts for the overlapping collision information. The CMS detector solves this problem via the use of a pipeline of detectors that synchronize to ensure that particles from different collisions can be associated to the correct event.

Information from particle collisions first encounter Level 1 (L1) triggers, which are completely hardware-based. L1 triggers decide whether to record the particle event information based on simple considerations. For instance, if a collision is measured to have an amount of energy very similar to that of an electron, and the analysis expects to see zero electrons, then the trigger will likely not record the information. The goal of this process is to reduce the amount of events recorded to about 10^5 Hz by ignoring events with characteristics that are not likely to produce new physics.

After the L1 triggers, the amount of event information is reduced further via High Level Triggers (HLT). HLT triggers are software-based and use more detector information to inform its decisions. For instance, HLT triggers can use both energy spread and particle path information to select interesting events. Thus, the

HLT triggers reduce the rate of recorded events to a manageable frequency.

The collider luminosity can change depending on the year that the data is collected, such that different HLTs are needed to collect data rate to something manageable. Therefore, trigger selection are important to consider: which triggers should we use to maintain consistency across years? Which triggers should we use if a certain trigger is used one year, but not another? Trigger selection for this analysis is described further in Section 5.1.6.

5.1.1 Trigger Types

Since this analysis focused on a search for VBF $Z' \to \tau_h\tau_h$, selecting important events required the use of $\tau_h\tau_h$ (double hadronic tau) triggers. Table 5.1 lists the specific double hadronic tau triggers used in this analysis. Additionally, the trigger efficiency is shown in Figure 5.1.

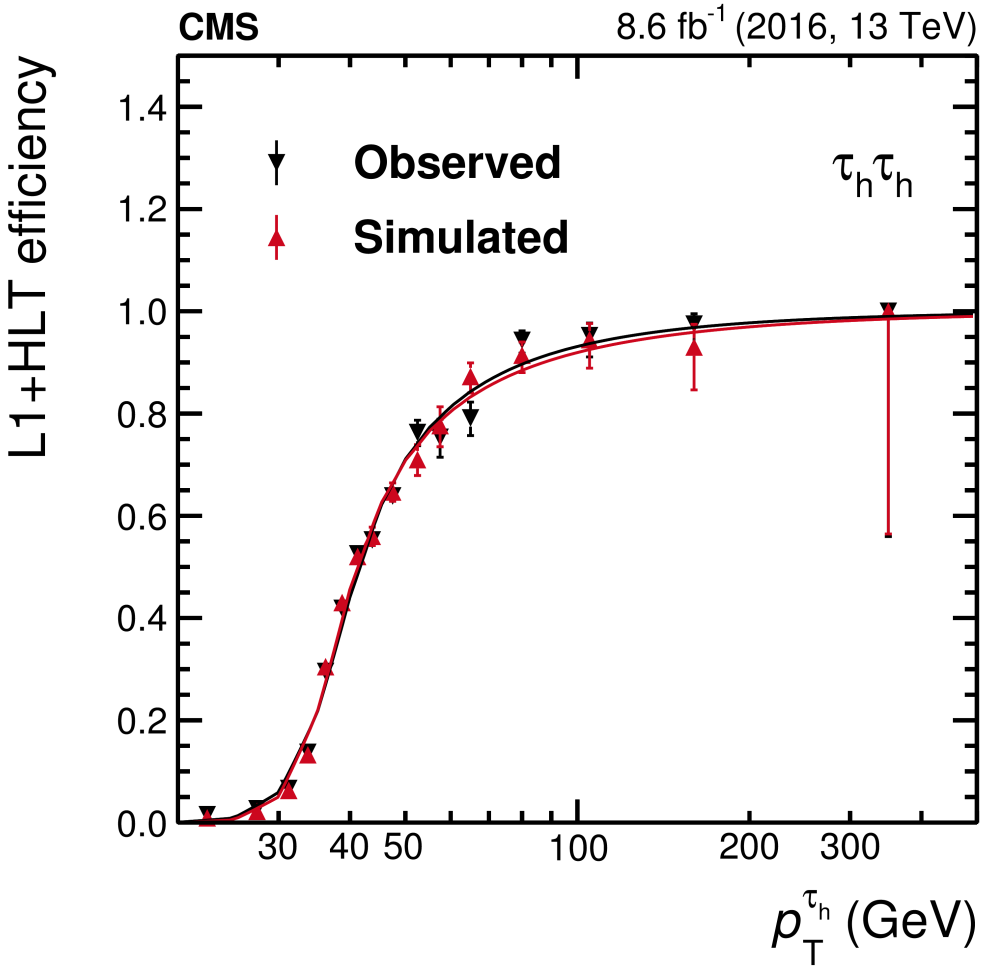


Figure 5.1: This figure shows the efficiency of the $\tau_h\tau_h$ trigger [45].

These convoluted names refer to specific trigger information. HLT refers to the trigger level (high-level

trigger), and DoubleMediumIsoPFTau3* refers to the specific trigger that recorded the event. These triggers used Double Medium Isolation, which means that the trigger used the "medium" level tau isolation criteria. The τ_h reconstruction algorithm is described in a later section. The criteria used to identify genuine τ_h candidates can be modified to be stringent depending on the sensitivity of the study with respect to the probability of a jet being misidentified as a τ_h. More-stringent ("tighter") criteria have lower τ_h identification efficiency, and less-stringent ("looser") criteria have higher τ_h identification efficiency.

The label Trk1 refers to the number of charged particles reconstructed for the HLT-level τ_h object. The eta2p1 label means that the trigger requires the HLT-level τ_h candidate to have $|\eta| < 2.1$.

Table 5.1: Ditau triggers for the analysis

Year	Double Hadronic Tau Trigger
2016	HLT_DoubleMediumIsoPFTau32(35)(40)_Trk1_eta2p1
	HLT_DoubleMediumCombinedIsoPFTau35(40)_Trk1_eta2p1
2017	HLT_DoubleMediumChargedIsoPFTau35(40)_Trk1_eta2p1
	HLT_DoubleMediumChargedIsoPFTau35(40)_Trk1_TightID_eta2p1
2018	HLT_DoubleMediumChargedIsoPFTau35(40)_Trk1_eta2p1
	HLT_DoubleMediumChargedIsoPFTauHPS35(40)_Trk1_eta2p1

5.1.2 Trigger Efficiency in MET Primary Data

Although the ditau triggers are used in the signal region, other triggers are used to select data for control samples used for background estimation studies. The following two subsections discuss those additional triggers. Table 5.2 lists the specific single-muon triggers used in this analysis. The trigger efficiency for single-muon triggers can be found in Reference [38]. IsoMu2* refers to the specific HLT used.

Table 5.2: Single muon triggers for the analysis

Year	Single Muon Trigger
2016	HLT_IsoMu24 || HLT_IsoTkMu24
2017	HLT_IsoMu27
2018	HLT_IsoMu24

The single-muon trigger efficiencies for events with genuine prompt muons are measured by the CMS Muon Physics Object Group (POG). Those measurements show that the simulated samples model the efficiencies observed in data well, showing data-to-MC trigger efficiency ratios consistent with 1 (within 1-2% accuracy) [48]. Therefore, we studied the performance of these single-muon triggers for real data events with fake muons, which are harder to correctly model in simulation. To perform these studies, we used data sets that were obtained without using these single-muon triggers to provide unbiased studies of the performance

of the trigger. Therefore, we use the E_T^{miss} primary dataset (PD), which is obtained using E_T^{miss} triggers.

The plot in Figure 5.2 shows the result for the turn-on curve in data. From Figure 5.2, one can see a $p_T(\mu)$ selection for the signal region of $p_T(\mu) > 30$ GeV lies firmly on the plateau for the trigger efficiency curve in data, although the trigger efficiency at the plateau is approximately 80% for events with fake muons.

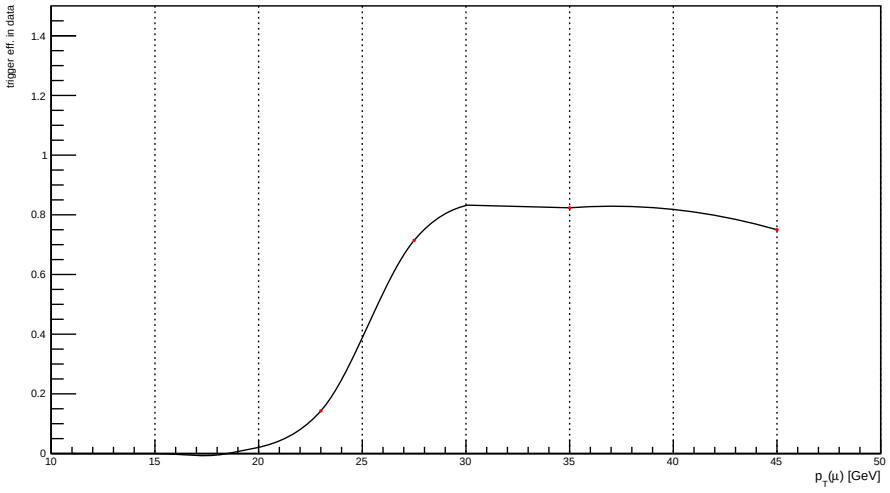


Figure 5.2: Trigger efficiency for muons in $t\bar{t}$-enriched sample in E_T^{miss} PD.

5.1.3 Effect of VBF selections on trigger efficiency

To see the effect of VBF selections on the single-muon trigger efficiency for genuine prompt muons, a comparison is made using trigger efficiency in data with VBF selections and without VBF selections using the single-muon trigger with Z$\rightarrow \mu\mu$ events. Trigger efficiency (ε is calculated using a reference trigger: $\varepsilon = N(ref and probe)/N(ref)$, where $N(ref)$ is the number of events measured in the reference trigger and $N(ref and probe)$ is the number of events that fire both the reference trigger and the trigger whose efficiency we are measuring. From Figure 5.3, one can see that the VBF selections do not change the trigger efficiency. Additionally, the single-muon trigger efficiency for genuine prompt muons is greater than 90% for muons with $p_T > 30$ GeV, thus motivating the use of this criteria for the background estimations of this analysis, which will be described in greater detail later.

5.2 Physics Object Reconstruction

5.2.1 Jet Reconstruction

In this analysis, we used the particle-flow (PF) technique [44] to improve the jet p_T and angular resolution. The PF technique combines information from different subdetectors to produce a mutually exclusive col-

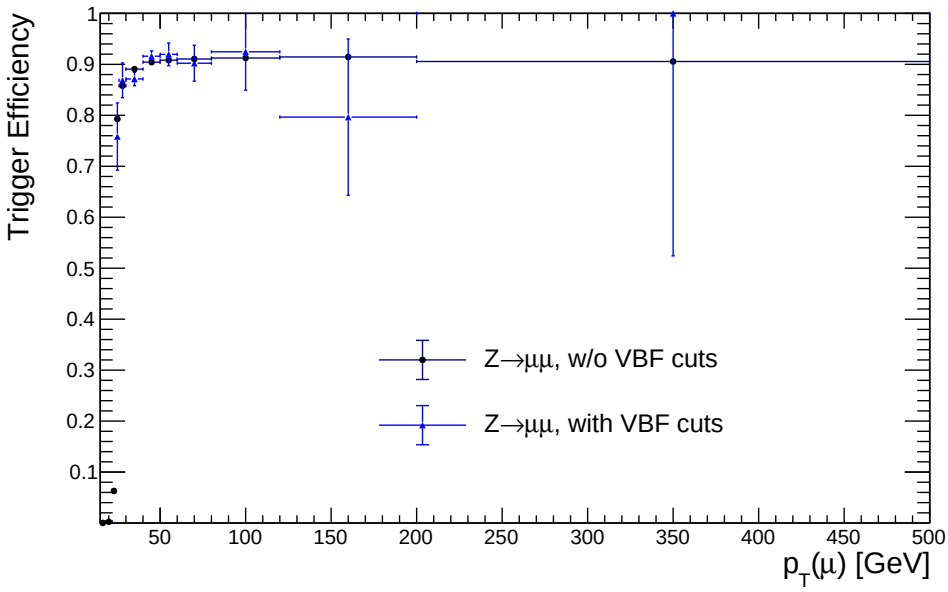


Figure 5.3: Trigger efficiency for muons in Z events (Z$\rightarrow \mu\mu$) with and without our VBF requirements, showing no visible change.

lection of particles (namely muons, electrons, photons, charged hadrons and neutral hadrons) that are used as input for the jet clustering algorithms. Jets are clustered using the anti-k_T algorithm [8], with a distance parameter of $R = 0.4$. We use these AK4 jets from the collection "slimmedJets," which are made from ak4PFJetsCHS.

The jets require energy corrections obtained using simulated events that are generated with PYTHIA, processed through a detector simulation based on GEANT, and confirmed with in situ measurements of the p_T balance. The overall jet-energy corrections depend on the η and p_T values of jets. The jet-energy corrections are applied by L1 FastJet, L2 Relative, and L3 Absolute corrections. To remove the extra energy in jets from underlying event (UE) and pileup (PU), the L1 FastJet corrections use the event-by-event UE/PU densities. The L2 and L3 corrections use jet balancing and photon+jet events to improve and provide a better energy response as a function of jet p_T and η. For data, additional residual corrections are applied.

Jets are required to have $p_T > 30$ GeV and $|\eta| < 5.0$. For the identification of jets, the "Loose" ("Tight") PF ID is used in 2016 (2017,2018) in this analysis. The jet reconstruction and ID efficiency in simulation is $>$99% for the entire η and p_T range.

5.2.2 b-Jet Tagging

A b-jet is a jet with a decay that includes a b quark. In this analysis, b-tagged jets are used to obtain $t\bar{t}$ enriched control samples to estimate the $t\bar{t}$ rate in the signal region.

For the 2016 analysis, the combined secondary vertex v2 algorithm [13] is used to identify jet as originating from hadronization of a b-quark (discriminator name "pfCombinedInclusiveSecondaryVertexV2BJetTags"). This algorithm combines reconstructed secondary vertex and track-based lifetime information to build a MVA-based discriminator to distinguish between jets from b-quarks and those from charm or light quarks and gluons. For 2017 & 2018, the Deep CSV algorithm is used to identify a jet as originating from hadronization of a b-quark [13].

The minimum thresholds on these discriminators define loose, medium, and tight operating points with a misidentification probability of about 10%, 1%, and 0.1%, respectively, for an average jet p_T of about 80 GeV. We would like to make an important note, as is highlighted by the b-tagging POG, that the b-tagging efficiencies and scale factors can vary depending on the year, run, and topology. For example, in this analysis we find that the b-tagging efficiency in $t\bar{t}$ MC is about 40% in 2016, 60% in 2017 and 50% in 2018 [14] [12]. The extent of this effect depends on several factors, including the amount of hadronic energy flow spread out over the detector, which is dictated by whether or not events pass or fail the VBF requirements. The end result is that the yields don't always scale from year to year with the difference in luminosity. The prescription is used for the b-tagging and mis-tagging scale factors and efficiencies. They are applied using the method called "Event reweighting using scale factors only" [18].

5.2.3 Muon Reconstruction and Identification

Muon reconstruction is a multistep process that begins with the information gathered from the muon subdetector. First, standalone muons are reconstructed from hits in the drift tube (DT) and cathode strip (CSC) chambers. Hits from the innermost muon stations are combined with hits in the other muon segments using the Kalman fitting technique [23]. The standalone muon trajectory is reconstructed by extrapolating from the innermost muon station to the outer tracker surface. This standalone muon trajectory is then used to find a matching track reconstructed in the inner silicon tracker. Finally, standalone muons and matching silicon tracks are used to perform a global fit resulting in a "global" muon.

Global muons are reconstructed by combining tracker muons from the inner silicon tracker and standalone muons from the muon chambers. Once a muon is required to have matching tracks in the inner and outer detectors, the main source of background consists of charged hadrons that leave a signature in the inner silicon tracker while also penetrating through the hadronic calorimeter and creating hits in the muon chambers. Charged hadrons that penetrate the hadronic calorimeter and leave hits in the muon system will deposit significant energy in the calorimeters. Therefore, a calorimeter compatibility algorithm can be used to significantly reduce the number of charged pion fakes. However, calorimeter compatibility is not used in this analysis due to our uncertainty of the performance of such algorithms in the presence of high PU.

The presence of punch-throughs often occur due to pions from the fragmentation of quarks and gluons. These punch-throughs can often be discriminated against by requiring isolation. Similarly, non-prompt muons from heavy flavor decays and decays in flights are expected to be within jets and can be discriminated against by imposing an isolation requirement. Muon identification is described in more detail in [15].

Isolated muons are required to have minimal energy from PF neutral and charged candidates in a cone of $\Delta R < 0.4$ around the lepton trajectory. This method of muon isolation is described in Reference [45], which states that muon isolation (I^{μ}_{rel}) is calculated as follows:

$$I^{\mu}_{rel} = \frac{\sum p_T^{charged} + max(0, \sum p_T^{neutral} + \sum p_T^{\gamma} - p_T^{PU})}{p_T^{\mu}} \tag{5.1}$$

where $p_T^{charged}$, $p_T^{neutral}$, p_T^{γ}, p_T^{PU} and p_T^{μ} are the scalar transverse momenta within the ΔR of charged hadrons, neutral hadrons, photons, pileup, and muons respectively. We used the PF-based combined relative isolation with $\Delta\beta$ corrections for pileup, where $\Delta\beta$ is the factor that accounts ratio of charged hadron energy to photon energy withing inelastic pp collisions [45].

Table 5.3 shows the complete list of μ identification criteria. We use the "Tight" identification working point and relative isolation < 0.15 in this analysis. In all channels, the identification and isolation used follows the POG recommended criteria. The muon trigger/identification efficiencies and scale factors used to correct the MC expectations in these analyses have been taken from the Muon POG [17].

Table 5.3: μ ID Selections

Cut
recoMu.isGlobalMuon()
muon::isPFMuon()
recoMu.globalTrack()-> normalizedChi2()< 10
recoMu.globalTrack()-> hitPattern().numberOfValidMuonHits()> 0
recoMu.numberOfMatchedStations()> 1
fabs(recoMu.muonBestTrack()-> dxy(vertex-> position()))< 0.2
fabs(recoMu.muonBestTrack()-> dz(vertex-> position()))< 0.5
recoMu.innerTrack()-> hitPattern().numberOfValidPixelHits()> 0
recoMu.innerTrack()-> hitPattern().trackerLayersWithMeasurement()> 5

5.2.4 Tau Reconstruction and Identification

The challenge in identifying hadronically decaying taus is discriminating against generic quark and gluon QCD jets which are produced with a cross–section several orders of magnitude larger. CMS has developed several algorithms to reconstruct and identify hadronically decaying taus based on Particle Flow (PF) objects. Information from multiple subdetectors within the experiment is combined to efficiently discriminate hadronic taus against the main background sources. For this analysis, the tau POG-recommended DeepTau

algorithm is used. DeepTau combines information from the high-level reconstructed tau features together with the low-level information from the inner tracker, calorimeters and muon subdetectors using particle flow candidates & electrons and muons reconstructed within the tau isolation cone [50]. DeepTau also takes advantage of more inclusive tau reconstruction decay mode definitions that were recently developed in CMS: three charged prongs + 0 or 1 π^0 with relaxed matching conditions added to previously reconstructed 1 charged prong + 0, 1, 2 π^0 and three charged prongs + 0 π^0 with tight matching conditions.

The DeepTau algorithm at CMS takes a series of input variables. The first is one global event variable: the average energy deposition density (ρ). Next, there exist 42 high-level event variables used during tau reconstruction or that have been proven to provide strong discriminating power in previous tau POG studies. Finally, there exists information about 4-momentum, track quality, relation with the primary vertex, calorimeter clusters, and muon stations that can be used for each candidate reconstructed within the tau signal or isolation cone, where available. Depending on the candidate type, there are 7 to 27 variables to be used for each PF candidate. There are 37 variables corresponding to each electron or muon candidate.

Candidates belonging to the inner and outer cones are separated and split into two grids with $\eta \cdot \phi$ cell size of 0.02·0.02(0.05·0.05) for the inner (outer) cone. The following considerations apply to the network architecture. High-level variables and each input cell are pre-processed by a few fully connected dense layers, where a "dense layer" refers to a layer of the neural network that calculates a value using information from every neuron in the previous layer. For the inner (outer) grid, the pre-processed cell data are fed into 5 (10) 2D convolutional layers with 3·3 window size, which result in 64 features that are passed to the next step. All features from previous steps are combined and passed through five dense layers. Probabilities of the reconstructed tau candidate being electron, muon, quark or gluon jet, or hadronic tau are estimated by the four neural network outputs.

To discriminate against muons, DeepTau taus are required to pass the lepton rejection discriminator which requires the lead track of the tau not be associated with a global muon signature. In order to discriminate against electrons, DeepTau taus are required to pass an MVA discriminator which uses the amount of HCAL energy associated to the tau with respect to the measured momentum of the track (H/p). Additionally, the MVA discriminator considers the amount of electromagnetic energy in a narrow strip around the leading track with respect to the total electromagnetic energy of the tau. Finally, DeepTau taus must not reside in the ECAL cracks. The identification and isolation used follows the Tau POG recommended criteria [41]. The exact discriminator names and working points are listed and described in their respective sections.

For consistency, information about the HPS algorithm for taus is included to follow. HPS makes use of PF jets as inputs to an algorithm that uses strips of clustered electromagnetic particles to reconstruct neutral pions. The electromagnetic strips ("neutral pions") are combined with the charged hadrons within the PFJets

to attempt to reconstruct the main tau decay modes outlined in Table 5.4.

Table 5.4: Reconstructed Tau Decay Modes

HPS Tau Decay Modes
Single Charged Hadron + Zero Strip
Single Charged Hadron + One Strip
Single Charged Hadron + Two Strips
Two Charged Hadrons
Three Hadrons

The single hadron plus zero strips decay mode attempts to reconstruct $\tau \to \nu\pi^{\pm}$ decays or $\tau \to \nu\pi^{\pm}\pi^{0}$ decays where the neutral pion has very low energy. The single hadron plus one or two electromagnetic strips attempts to reconstruct tau decays that produce neutral pions where the resulting neutral pion decays produce collinear photons. Similarly, the single hadron plus two strips mode attempts to reconstruct taus that decay via e.g. $\tau \to \nu\pi^{\pm}\pi^{0}$ where the neutral pion decays to well separated photons resulting in two electromagnetic strips. The three hadrons decay mode attempts to reconstruct tau decays that occur via $\rho(770)$ resonance. In the case of high-p_T τ leptons, their visible decay products have sufficient boost, producing very collimated tracks and thus overlapping tracker hits (merged tracks), which makes it difficult to efficiently reconstruct the three hadrons decay mode (i.e. sometimes PF reconstructs two tracks instead of three tracks). Therefore, the two charged hadrons decay mode attempts to recover some of the inefficiency for identifying 3-prong high-p_T τ leptons.

We note, however, that the two-prong decay mode is not utilized in this analysis due to the large jet background (jets can primarily be reconstructed as two-prong τ_h candidates). In all cases, electromagnetic strips are required to have $E_T > 1$ GeV/c. Additionally, the particle flow charged hadrons are required to be compatible with a common vertex and have a net charge of $|q| = 1$. Figure 5.4 shows the performance of different hadronic tau identification working points.

The "Tight" working point is a balance of real τ_h efficiency ($\sim 50\%$) and hadron jet rejection rate ($>99\%$ for jets with $p_T > 20$ GeV). We note that the "Tight" WP is used for the τ_hs against electrons and muons as well.

5.2.5 Tau Energy Scale and Resolution

Since the resolution and scale of our mass reconstruction depends on the effectiveness of the τ_h reconstruction, in this section we summarize our studies on τ_h response and resolution. We define the response as the relative difference between the transverse momentum of a reconstructed τ_h (that has passed all τ_h ID discriminators) and the transverse momentum of a generated τ_h that has been matched ($\Delta R < 0.2$) to the reconstructed tau. We can see from Figure 5.5 that the response distribution contains a narrow Gaussian like component in the

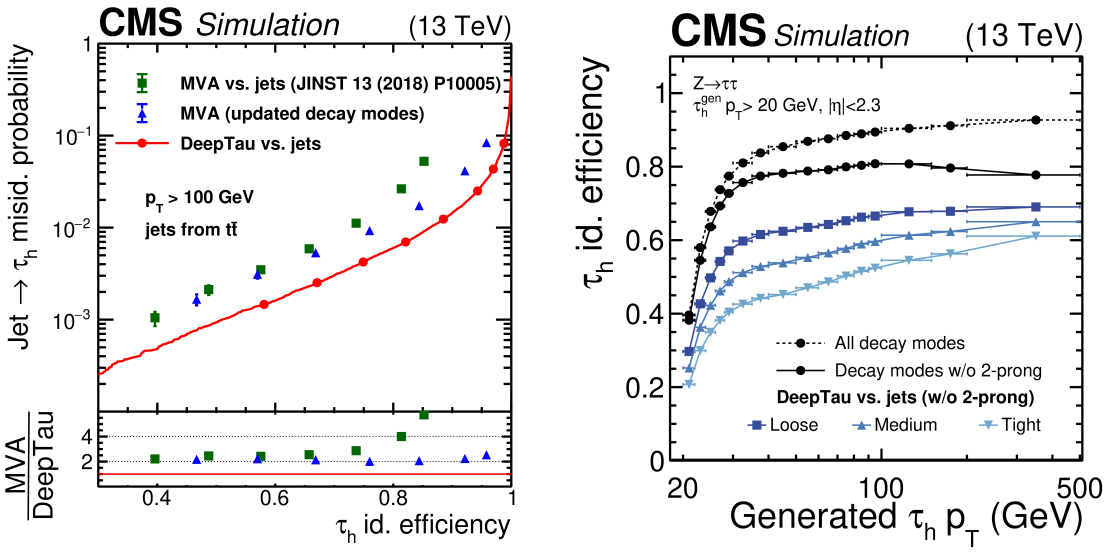


Figure 5.4: This figure demonstrates the performance of difference hadronic tau identification working points [53].

bulk, in addition to a relatively long tail (compared to electrons and muons).

While the tails become less substantial at high p_T, the Gaussian-like bulk of the response distributions broaden at high p_T. The systematics section describes how the uncertainty on τ_h energy scale and resolution impacts the predictions for BG and signal in our main search regions.

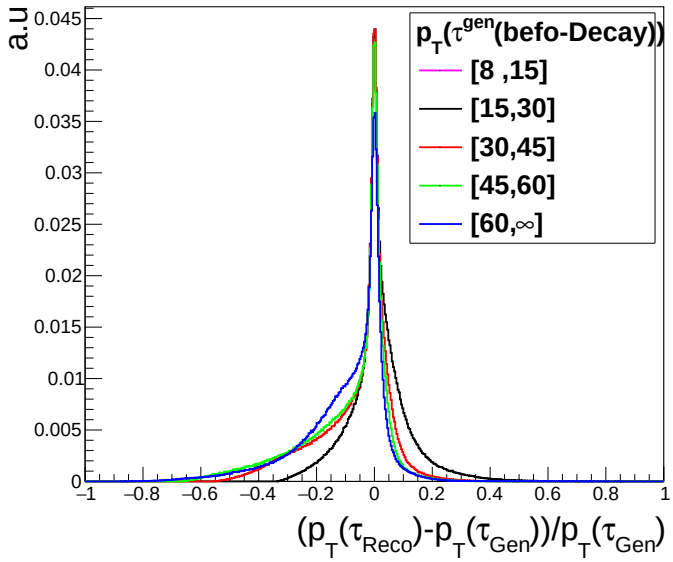


Figure 5.5: Relative p_T resolution of reconstructed τ_h candidates.

CHAPTER 6

Strategy for Hadronic Ditau Final State

The following section details our search for Z' via VBF with the $\tau_h\tau_h$ (also called ditau) final state. This includes a summary of the analysis strategy, the signal optimization studies, the background estimation strategy with all the control region studies, and the summary of the final predictions in the signal region.

6.1 Analysis Strategy

Pairs of tau leptons decay to the following final states: $\mu\mu$, ee, $e\mu$, $e\tau_h$, $\mu\tau_h$, and $\tau_h\tau_h$. Their respective branching ratios are listed in Table 6.1. My analysis studies the $\tau_h\tau_h$ final state, which has the largest branching fraction. As mentioned, however, hadronically decaying taus run the risk of being confused with hadronic jets. Thus, a major focus of this analysis will be accurate background estimation.

Table 6.1: Ditau Decay Branching Ratios

Decay Modes	Branching Fraction
$\mu\mu$	3.1%
ee	3.1%
$e\mu$	6.2%
$e\tau_h$	23.1%
$\mu\tau_h$	22.5%
$\tau_h\tau_h$	42.0%

To find ditau particle decays, I searched for events that have two τ_h candidates that are oppositely charged and have nearly-opposite p values (that is, they travel in opposite directions). Since this interaction decays to up to four neutrinos, high E_T^{miss} was expected and accounted for. Additionally, the high presence of neutrinos means that the reconstructed Z' mass, defined as $\tau_h\tau_h + E_T^{miss}$, has a broad peak, as opposed to the narrow peaks seen in $Z' \to \mu\mu$ and $Z' \to ee$ searches. Figure 6.1 below shows a comparison of the $Z' \to ee$ and $Z' \to \tau\tau \to \tau_h\tau_h$ search signatures, demonstrating the relative broadness of the reconstructed Z' mass peak.

My strategy to resolve such a broad signal was (1) to optimize parameter selections for the strongest Z' signal, and (2) to model the background processes for VBF $Z' \to \tau\tau$ as accurately as possible so that, even with uncertainties, I would be able to resolve a Z'. Parameter optimization consisted of varying individual parameters (for instance, dijet mass) while keeping the other parameters fixed, and with VBF cuts implemented. Signal significance was then compared to find the parameters that give the highest chance of resolving a Z' signal.

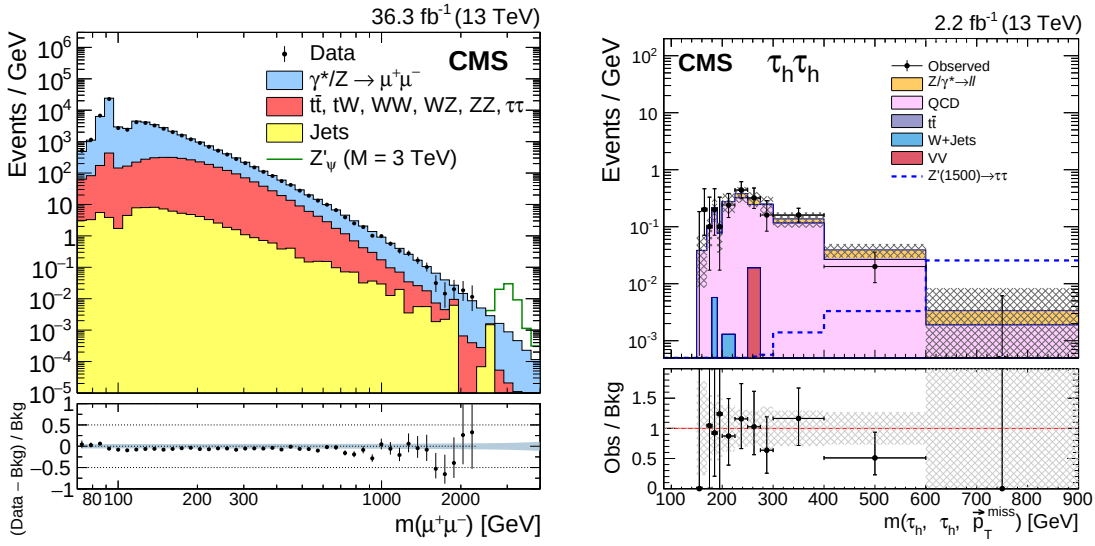


Figure 6.1: The figure demonstrates the relative broadness of the invariant reconstructed mass plot of ditau decay verses dimuon decay. (Left) Dimuon final state mass plot. (Right) Ditau final state mass plot.

The $Z \to \tau\tau$ interaction is both a background interaction for the ditau channel and an interaction that we can use to verify the efficiency of the reconstruction methods and background estimation with genuine τ_h candidates. CMS data is used to verify background estimations via a process called a data-driven approach. Data-driven estimates use real data in control regions orthogonal to the signal region to estimate the events in the signal region via a transfer factor, which is described in greater detail in a later section.

Vector boson fusion (VBF) has distinct kinematics: large difference in pseudorapidity between jets, and high invariant dijet mass (TeV-scale). The difference in pseudorapidity ($\Delta\eta$) is calculated using the following:

$$\Delta\eta_{jj} \to |\eta_{j1} - \eta_{j2}| \tag{6.1}$$

where η_{j1} and η_{j2} refer to the η of each jet respectively. The invariant dijet mass can be calculated as a function of $\Delta\eta_{jj}$:

$$m_{jj}^2 = 2p_{T,j1}p_{T,j2}(cosh(\Delta\eta_{jj}) \tag{6.2}$$

Thus, we can isolate potential VBF events by requiring high $\Delta\eta$ and m_{jj}. This is particularly useful in reducing the number of QCD events, as those events typically have low $\Delta\eta_{jj}$ and low m_{jj}. This phenomenon is demonstrated in Figure 6.2.

The VBF restrictions used for this analysis were to require at least two jets with $p_T > 30$ GeV, and having $m_{jj} > 500$ GeV and $\Delta\eta_{jj} > 4.2$. Additionally, a cut of at least 30 GeV of E_T^{miss} was required to match a ditau

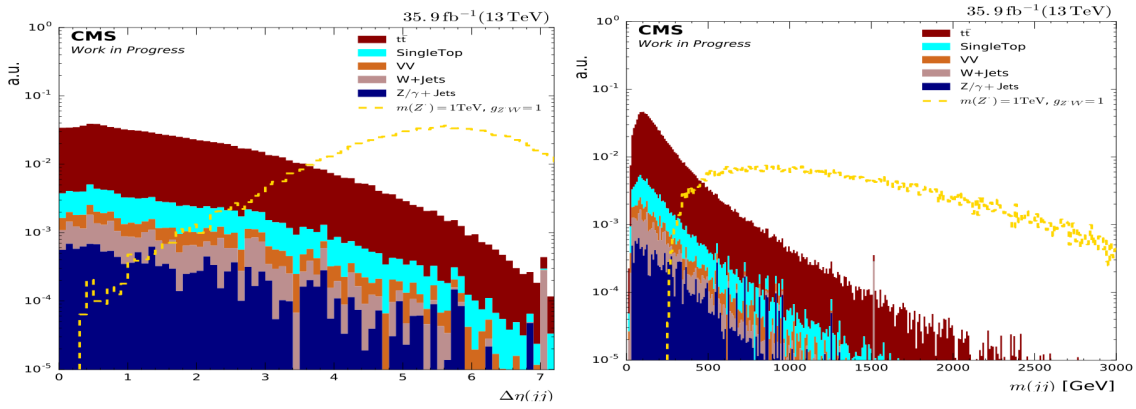


Figure 6.2: This figure demonstrates the strength of the Z' signal with respect to (Left) $\Delta\eta_{jj}$ and (Right) m_{jj}.

final state. Objects in the central part of the detector ($|\eta| < 2.1$) are reconstructed using kinematic selections specific to each object; such selections are referred to as "central selections."

In this analysis, I defined background-enriched control regions (CR) that are orthogonal (not-overlapping) to the signal region (SR), such that the CR would have a negligible presence of signal events. An example of such an orthogonal CR is that of the $Z \to \tau\tau$ control region, which has a $m(\tau_h, \tau_h, E_T^{miss}) < 150$ GeV. Once I create a CR targeting a particular background, I use the CR to determine a scale-factor (SF), which is the ratio between the observed data and the background estimated via simulation, which can be calculated as follows:

$$SF_{Process} = \frac{N^{Data}_{Process}}{N^{MC}_{Process}} = \frac{N^{Data}_{CR} - N^{MC}_{Not process}}{N^{MC}_{Process}} \tag{6.3}$$

Where $SF_{Process}$ is the scale factor for a particular process; $N^{Data}_{Process}$ is the number of events for the process present in data for that control region; $N^{MC}_{Process}$ is the number of events for the process present in the MC simulation for that control region; N^{Data}_{CR} is the number of events in data for that control region; and $N^{MC}_{Not process}$ is the number of events from MC simulation that are not from that particular process.

The SFs can then be applied to the predictions from simulation to account for the difference between simulation and observation. This method of estimating the background event contribution of a particular process ($N^{SR,estimate}_{process}$) is summed up in the following equation:

$$N^{SR,\text{estimate}}_{process} = N^{SR,\text{ MC}}_{process} \times SF_{non-VBF} \times SF_{VBF} \tag{6.4}$$

where $N^{SR,\text{MC}}_{process}$ represents the predicted signal region yield for a particular process (for instance, $Z \to \tau\tau$) in simulation, while $SF_{non-VBF}$ and SF_{VBF} represent corrections to the central selection and VBF selection efficiencies respectively.

6.2 Event Simulation

6.2.1 Monte Carlo Event Generation

Next, it is important to understand how we simulate events to estimate the expected background so that we can compare that background to data. To maintain the integrity of our results, the signal region data is kept blinded until we have estimated the expected background as accurately as possible. Therefore, we must use tools to simulate events in a way that accounts for as many factors as possible.

The statistical technique we used to generate decay events based on their statistical likelihood to occur is the "Monte Carlo" (MC) method. MC methods rely on random sampling to obtain numerical results. In mathematics, for example, MC can be used to solve integration problem numerically, which is useful in cases in which solving the equation mathematically is difficult, too slow, or impossible. For statistical analysis, MC techniques may be used when the inputs for the model have a high amount of uncertainty.

MC event generation at CMS has two steps. First, a purely mathematical model based on SM or BSM theories generates events most likely to be produced for a given number of pp collisions. Second, a program is used to simulate how interactions with the CMS detector itself would effect the detected particles and therefore the resulting measurements. Thus, the event generator creates MC samples with simulated detector interactions that can be used to compare with the data.

To make the MC samples, we must consider the initial energy the interactions. As mentioned previously, particle interactions at the LHC consist mainly of partons interacting with one another. Since partons can include numerous virtual particles (particularly gluons), it is impossible to know the initial energy of any particular interaction. Thus, the parton energies must be estimated as a portion of the total energy of the proton (E_p). This portion of the proton energy can be described as xE_p, where x is a number between 0 and 1. The value x can be described by parton distribution functions (PDF). The PDF provides the probability that a given parton has a fraction x of the total proton energy. PDFs are measured by a variety of previous experimental results. [51]

When considering expected particle interactions at the level of pure theory, calculating input parameters can be straight-forward with varying accuracy. However, to achieve greater accuracy, it may be necessary to consider next-to-leading order (NLO) matrix elements that account for additional particle interactions. For instance, NLO calculations can account for initial state quarks exchanging gluons. In this way, NLO methods incorporate additional Feynman diagrams (i.e. $qg \to Zq$ and $qg \to Zq$), increasing the sample's statistical accuracy.

The event generator PYTHIA was used to create LO events, meaning that it used only the "tree level" Feynman diagrams (diagrams that contain no closed loops). The program MADGRAPH was also used

in conjunction with PYTHIA to account for hard parton radiation that results from parton showering and hadronization.

The program GEANT4 was used to simulate particle interactions with the CMS detector. The program accounts for the effect of detector geometry and materials on event measurement. This step allows for more-accurate simulation of particles within the detector itself.

Finally, MC sample creation must account for pile-up interactions (interactions from other pp collisions not associated with the collision of interest) within the detector. One method of estimating the effects of PU is to overlay real data on top of the MC sample, which is called "data mixing." The data events used for data mixing must pass certain triggers and are timed in a way such that the event has most likely resulted from pileup.

6.2.2 Samples Used

The 13 TeV data collected by the CMS detector in year 2016, 2017, and 2018 is used in this analysis. The first nine tables to follow show the collision datasets used for tau and single muon preferred datasets (PDs) in 2016, 2017, and 2018. The total integrated luminosity of the collision data samples is 137.1 fb^{-1}.

Official CMS approved MC samples, known as nanoAODv6, for years 2016, 2017, and 2018 were used for all Standard Model processes. The leading order generators, Pythia8 and Madgraph, were used for signal and background MC production. The predicted background yields in simulation were determined using NLO or NNLO cross-sections, while the signal yields and distributions in all plots shown were normalized using the leading order cross-sections. For processes in which an inclusive sample is utilized in addition to HT-binned samples, cross-cleaning was performed with a filter that selects for the appropriate HT-range at generator-level to avoid incorrectly doubling the yields.

Table 6.2: Collision Data Samples: Tau (2016)

Physics Sample	Official CMS Datasets
Run 2016Bv1 Tau Run2016B-25Oct2019	*/Tau/Run2016B_ver1-Nano25Oct2019_ver1-v1/NANOAOD*
Run 2016Bv2 Tau Run2016B-25Oct2019	*/Tau/Run2016B_ver2-Nano25Oct2019_ver2-v1/NANOAOD*
Run 2016C Tau Run2016C-25Oct2019	*/Tau/Run2016C-Nano25Oct2019-v1/NANOAOD*
Run 2016D Tau Run2016D-25Oct2019	*/Tau/Run2016D-Nano25Oct2019-v1/NANOAOD*
Run 2016E Tau Run2016E-25Oct2019	*/Tau/Run2016E-Nano25Oct2019-v1/NANOAOD*
Run 2016F Tau Run2016F-25Oct2019	*/Tau/Run2016F-Nano25Oct2019-v1/NANOAOD*
Run 2016G Tau Run2016G-25Oct2019	*/Tau/Run2016G-Nano25Oct2019-v1/NANOAOD*
Run 2016H Tau Run2016H-25Oct2019	*/Tau/Run2016H-Nano25Oct2019-v1/NANOAOD*

Leading order $t\bar{t}$, W + jets, and DY + jets Madgraph samples were used and the predicted yields in simulation were determined using NLO or NNLO cross-sections. Because the MC simulated samples contain a pileup (PU) distribution that does not match that of data, the MC needs to be properly weighted to fit the PU

Table 6.3: Collision Data Samples: Tau (2017)

Physics Sample	Official CMS Datasets
Run 2017B Tau Run2017B-25Oct2019	*/Tau/Run2017B-Nano25Oct2019-v1/NANOAOD*
Run 2017C Tau Run2017C-25Oct2019	*/Tau/Run2017C-Nano25Oct2019-v1/NANOAOD*
Run 2017D Tau Run2017D-25Oct2019	*/Tau/Run2017D-Nano25Oct2019-v1/NANOAOD*
Run 2017E Tau Run2017E-25Oct2019	*/Tau/Run2017E-Nano25Oct2019-v1/NANOAOD*
Run 2017F Tau Run2017F-25Oct2019	*/Tau/Run2017F-Nano25Oct2019-v1/NANOAOD*

Table 6.4: Collision Data Samples: Tau (2018)

Physics Sample	Official CMS Datasets
Run 2018A Tau Run2018A-25Oct2019	*/Tau/Run2018A-Nano25Oct2019-v1/NANOAOD*
Run 2018B Tau Run2018B-25Oct2019	*/Tau/Run2018B-Nano25Oct2019-v1/NANOAOD*
Run 2018C Tau Run2018C-25Oct2019	*/Tau/Run2018C-Nano25Oct2019-v1/NANOAOD*
Run 2018Dv2 Tau Run2018D-25Oct2019	*/Tau/Run2018D-Nano25Oct2019_ver2-v1/NANOAOD*

Table 6.5: Collision Data Samples: Muon (2016)

Physics Sample	Official CMS Datasets
Run 2016Bv1 SingleMu Run2016B-25Oct2019	*/SingleMuon/Run2016B_ver1-Nano25Oct2019_ver1-v1/NANOAOD*
Run 2016Bv2 SingleMu Run2016B-25Oct2019	*/SingleMuon/Run2016B_ver2-Nano25Oct2019_ver2-v1/NANOAOD*
Run 2016C SingleMu Run2016C-25Oct2019	*/SingleMuon/Run2016C-Nano25Oct2019-v1/NANOAOD*
Run 2016D SingleMu Run2016D-25Oct2019	*/SingleMuon/Run2016D-Nano25Oct2019-v1/NANOAOD*
Run 2016E SingleMu Run2016E-25Oct2019	*/SingleMuon/Run2016E-Nano25Oct2019-v1/NANOAOD*
Run 2016F SingleMu Run2016F-25Oct2019	*/SingleMuon/Run2016F-Nano25Oct2019-v1/NANOAOD*
Run 2016G SingleMu Run2016G-25Oct2019	*/SingleMuon/Run2016G-Nano25Oct2019-v1/NANOAOD*
Run 2016H SingleMu Run2016H-25Oct2019	*/SingleMuon/Run2016H-Nano25Oct2019-v1/NANOAOD*

Table 6.6: Collision Data Samples: Muon (2017)

Physics Sample	Official CMS Datasets
Run 2017B SingleMu Run2017B-25Oct2019	*/SingleMuon/Run2017B-Nano25Oct2019-v1/NANOAOD*
Run 2017C SingleMu Run2017C-25Oct2019	*/SingleMuon/Run2017C-Nano25Oct2019-v1/NANOAOD*
Run 2017D SingleMu Run2017D-25Oct2019	*/SingleMuon/Run2017D-Nano25Oct2019-v1/NANOAOD*
Run 2017E SingleMu Run2017E-25Oct2019	*/SingleMuon/Run2017E-Nano25Oct2019-v1/NANOAOD*
Run 2017F SingleMu Run2017F-25Oct2019	*/SingleMuon/Run2017F-Nano25Oct2019-v1/NANOAOD*

Table 6.7: Collision Data Samples: Muon (2018)

Physics Sample	Official CMS Datasets
Run 2018A SingleMu Run2018A-25Oct2019	*/SingleMuon/Run2018A-Nano25Oct2019-v1/NANOAOD*
Run 2018B SingleMu Run2018B-25Oct2019	*/SingleMuon/Run2018B-Nano25Oct2019-v1/NANOAOD*
Run 2018C SingleMu Run2018C-25Oct2019	*/SingleMuon/Run2018C-Nano25Oct2019-v1/NANOAOD*
Run 2018D SingleMu Run2018D-25Oct2019	*/SingleMuon/Run2018D-Nano25Oct2019-v1/NANOAOD*

distribution observed in data. The reweighting of MC events is performed by determining the probabilities to obtain n interactions in data ($P_{data}(n)$) and MC ($P_{MC}(n)$) and using the event weights

$$w_{PU}(n) = \frac{P_{data}(n)}{P_{MC}(n)} \tag{6.5}$$

to reweigh MC events based on the number of interactions. The recommended minbias cross-section of 69.2 *mb* is used to determine these weights.

The following corrections have been applied to the MC predictions: (*i*) trigger efficiency weights to match the efficiency in data as recommended by the POGs; (*ii*) b-tagging SFs as recommended by the POG; (*iii*) pileup weights to match the number of interactions per bunch crossing in data (using recommended minibus cross-section); (*iv*) DY Z/W corrections (see subsequent sections) to correct the modeling of the Z/W boost (and, therefore, the correlated initial state radiation kinematics); (*v*) POG-recommended lepton identification scale factors.

6.3 Signal Optimization Studies

To have the best chance of finding the Z' particle, it is necessary to use the most optimal search parameters, i.e. the search parameters that allow the Z' to stand out from its background processes. The best signal possible is found by calculating each parameter's signal significance, which is defined by $\frac{S}{\sqrt{S+B}}$, where S stands for the event contribution in the signal and B stands for the event contribution of the background. The optimal selection criteria were found by varying each selection/parameter individually and comparing the resulting signal significance.

The optimization plots for the $\tau_h\tau_h$ channel are shown in Figure 6.3. These plots demonstrate how the signal significance varies with the each parameter cut value. Thus, to achieve highest signal significance, the optimal selection chosen was chosen. For instance, the dijet mass cut of $m_{jj} > 500$ GeV was chosen. For some parameters, other factors had to be considered in addition to the signal significance. For instance, the optimization plot for the τ_h ID working point suggests a loose working point, but I used a tighter working point to avoid fake taus. The final optimal selections and parameters used in the VBF Z' study are listed in Table 6.8.

The optimized selections in the $\tau_h\tau_h$ channel are similar that of other VBF Z' channels in our analysis, with the main difference being the higher p_T thresholds on τ_h candidates, driven by the per leg trigger turn-on curve for the di-τ_h trigger. Additionally, I veto events containing additional electrons and muons to select for only hadronic τ decay. The $m(\tau_h, \tau_h, E_T^{miss})$ mass plot shown in Figure 6.4 (right) is the expected background and signal prediction (for a benchmark signal sample) with the optimized signal region selections. As mentioned previously, the presence of a Z' signal is extracted with a shape based fit of the expected vs. observed $m(\tau_h, \tau_h, E_T^{miss})$ distributions, with the signal showing up as a broad excess in the tails.

Table 6.8: This table lists the selections optimized for best signal significance.

$\tau_h\tau_h$ Channel	
Parameter	**Requirement/Selection**
$p_T(\tau_h)$	> 70 (motivated by trigger selection)
Tau η	< 2.1
E_T^{miss}	> 30
τ_h ID	"Tight" DeepTau isolation with decay mode finding (new DMF))
$p_T(j)$	> 30
Jet ID	"Loose" (2016); "Tight" (2017-2018)
Trigger	HLT_DoubleMediumIsoPFTau*; HLT_DoubleMediumCombinedIsoPFTau*
Event Cleaning	$E_T^{miss}\,filters; instrumental background$
$N(\tau_h)$	$= 2$
$q(\tau_{h1}) \cdot q(\tau_{h2})$	< 0
$\lvert\Delta\eta_{jj}\rvert$	> 4.2
m_{jj}	> 500 GeV
$\eta_{j_1} \times \eta_{j_2}$	< 0
$N(\mathrm{b-jet})$	$== 0$
$N(\mathrm{e},\mu)$	$== 0$
$\Delta R(\tau_h, j)$	> 0.4
$\Delta R(\tau_h, e)$	> 0.3
$\Delta R(\tau_h, \mu)$	> 0.4

6.4 Signal efficiencies for Ditau Channel

Signal efficiencies are shown in Figure 6.5 for the four search channels of the $Z' \to \tau\tau$ analysis. These efficiencies are a function of the Z' mass. The signal efficiencies are calculated for both universal coupling to light and heavy fermions (blue line) and enhanced coupling to heavy fermions (red line) scenarios and are compared in each figure for each final state. To explain the dependence of signal efficiency on coupling of Z' to vector bosons (κ_V), Figure 6.5 shows the signal efficiency as a function of κ_V : 0.1 (Top), 0.5 (Middle) and 1.0 (Bottom). Both the $Z' \to \tau\tau$ and $Z' \to WW$ decay modes are considered in Figure 6.5.

6.5 Relevant SM Background Processes

Next, it is necessary to to model the background processes. These are processes that are able to pass the requirements that we make on the events. In other words, these are events that can appear to fake a Z'. Figure 6.6 shows the Feynman diagrams of the three major background processes in the search: Drell-Yan production of $Z \to \tau\tau$, W+Jets, and quantum chromodynamics (QCD).

Drell-Yan makes up about 61% of our total background due to its production of two genuine τ_h candidates and two jets (from initial state radiation) that can satisfy VBF criteria (see Figure 6.6 [Left]). W+Jets background, on the other hand, results from a single genuine tau and a "fake" tau due to an extra jet (with two other jets that can satisfy VBF criteria).

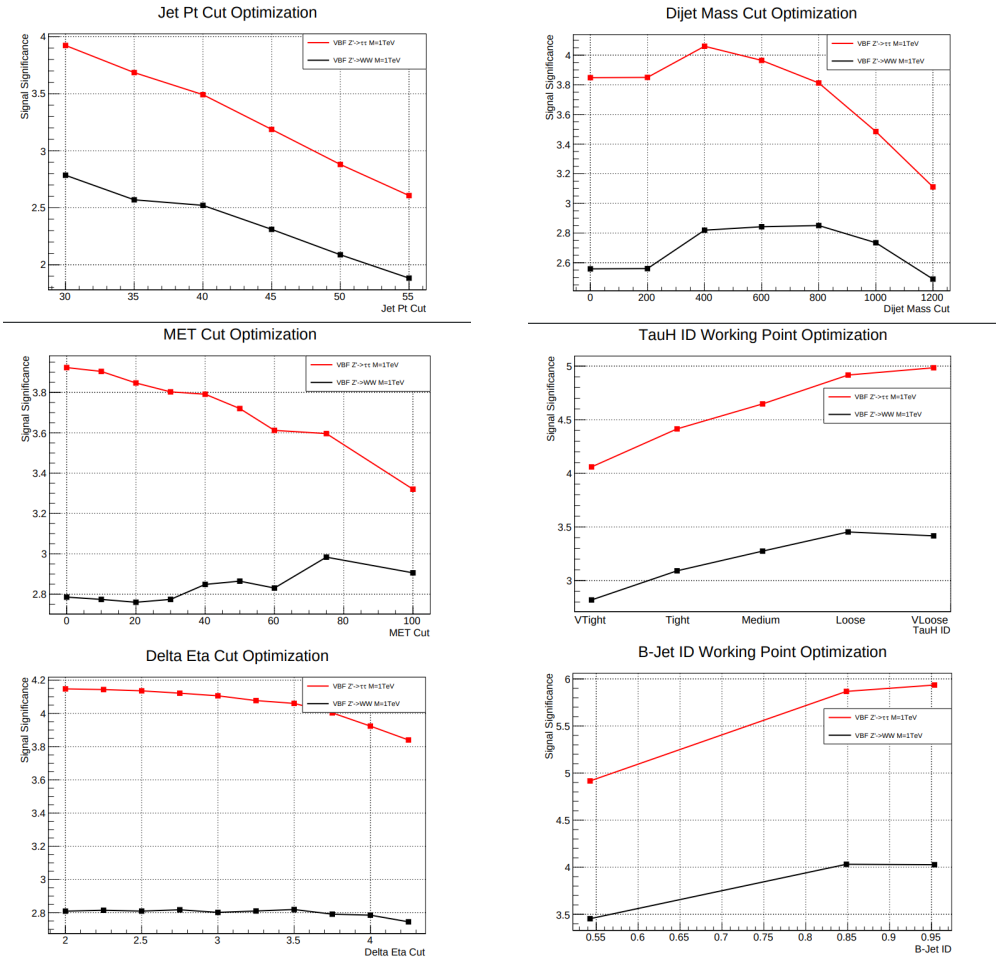


Figure 6.3: These plots compare the signal significance of a variety of parameter values.

The final major background contributor results from QCD processes in which jets fake both of the taus. Although QCD background contribution is minimized due to VBF criteria on the jets, enough QCD events meet the VBF criteria that this background remains the third-largest contributor to the background (see Figure 6.6 [Right]).

6.6 Background Estimation Strategies

Our background estimation strategy, then, is as follows: (1) Define control regions that are dominated by each of the background processes and have little contribution from the signal, (2) Use data in the control region to measure the dijet invariant mass shapes and the probabilities that these background events will meet our VBF selection criteria, and (3) For small background contributions, dijet and ditau mass shapes may be taken from simulation. With enough precision, it should be possible to resolve a Z' apart from the background, even when accounting for uncertainties.

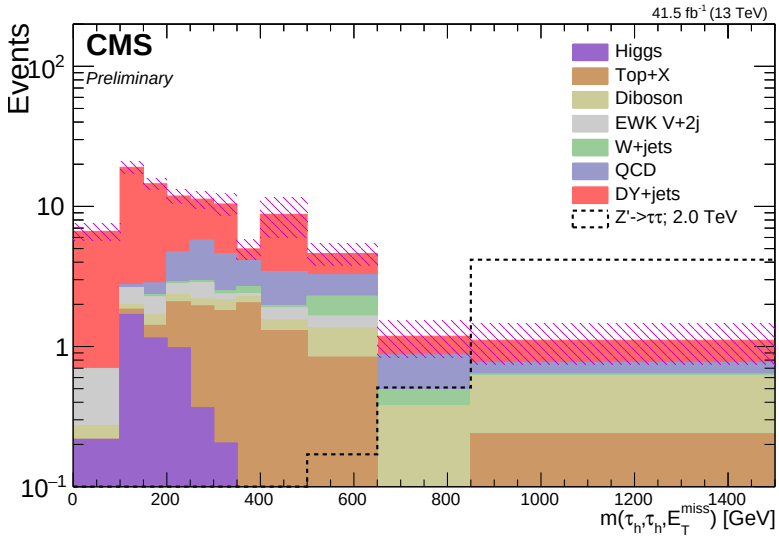


Figure 6.4: (Mass plot for 2017 MC with optimal selections applied.

To estimate the DY+Jets background, we used CRs (orthogonal to the SR), obtained using events at low $m(\tau_h, \tau_h, E_T^{miss}$, with and without VBF criteria. The equation used in this estimation was as follows:

$$N_{DY}^{SR,\text{ estimate}} = N_{DY}^{SR,\text{ MC}} \times SF_{non-VBF} \times SF_{VBF} \tag{6.6}$$

$N_{DY}^{SR,\text{MC}}$ represents the predicted signal region yield in simulation, while $SF_{non-VBF}$ and SF_{VBF} represent corrections to the central selection and VBF selection efficiencies respectively. With this approach, we factorize the modeling and mismodeling of the backgrounds arising from the central and VBF selections. $SF_{non-VBF}$ is typically measured in control regions without (or with inverted) VBF criteria, which means the level of agreement between data and MC in those CRs is due to the modeling of the central selection efficiencies. Further, SF_{VBF} is measured in control regions with the VBF criteria applied, and these control regions already have corrections for the mismodeling of the non-VBF efficiencies (using $SF_{non-VBF}$). We note that the contribution from pure electroweak Z+jets processes (e.g. VBF) is small with respect to the mixed QCD-QED DY+jets process and the total background ($\sim 4\%$ of the total background), and thus this background is taken directly from simulation.

W+jets becomes a background when the W decays to a τ lepton and neutrino (resulting in genuine τ_h and E_T^{miss}), one jet from initial state radiation fakes a τ_h, and two additional initial state radiation (ISR) jets pass the VBF criteria. Since this represents a high-order process (at least 3 ISR jets), and since the combined probability for a jet to fake a τ_h and two ISR jets pass the VBF criteria is small, the W+jets background represents a very small contribution to the signal region ($< 5\%$). The W+Jets background was estimated

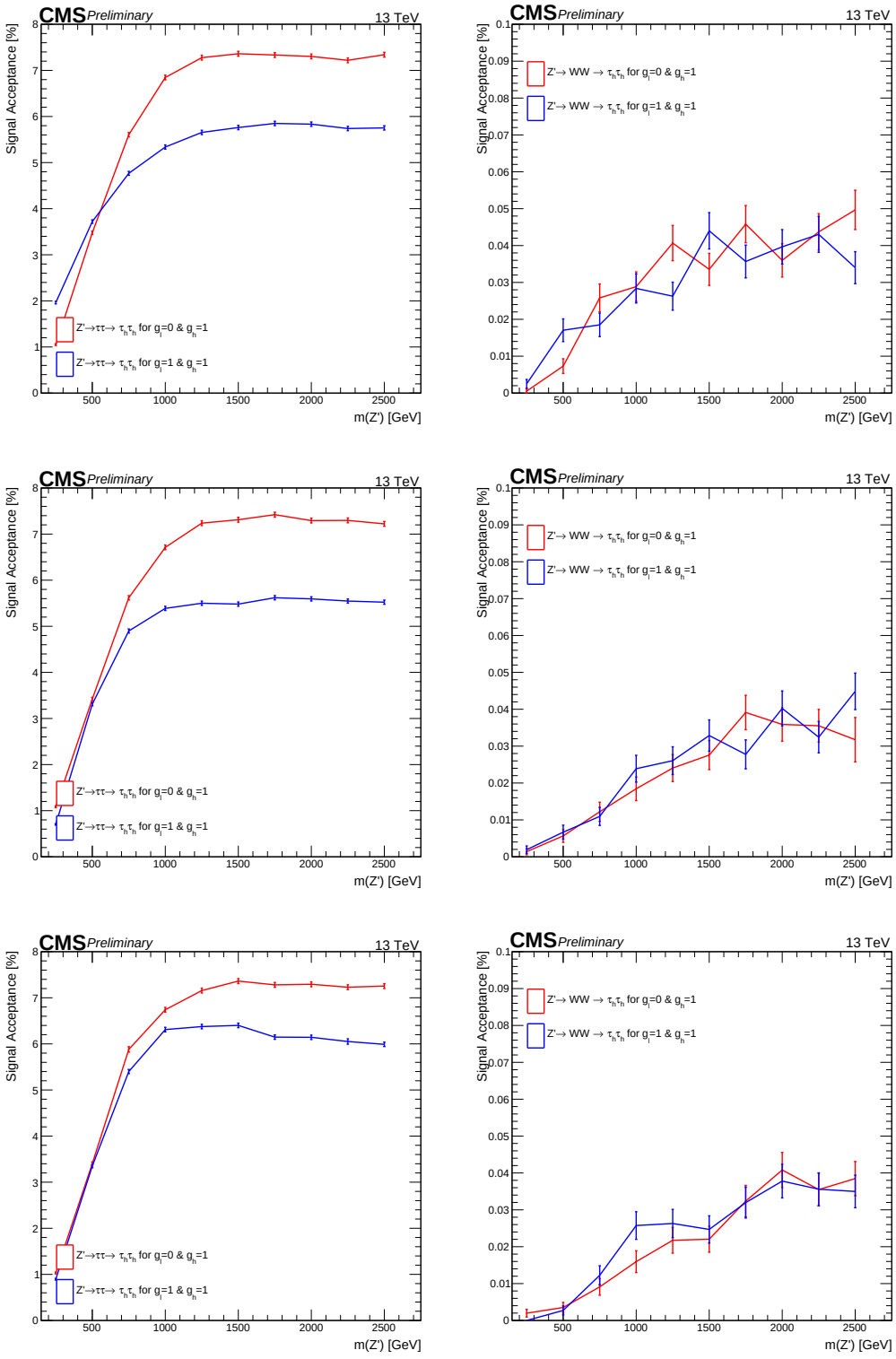


Figure 6.5: Signal efficiencies for $\tau_h\tau_h$ final state for VBF Z' decaying to $\tau\tau$ (left) and WW (right) with κ_V : 0.1 (top), 0.5 (middle) and 1.0 (bottom)

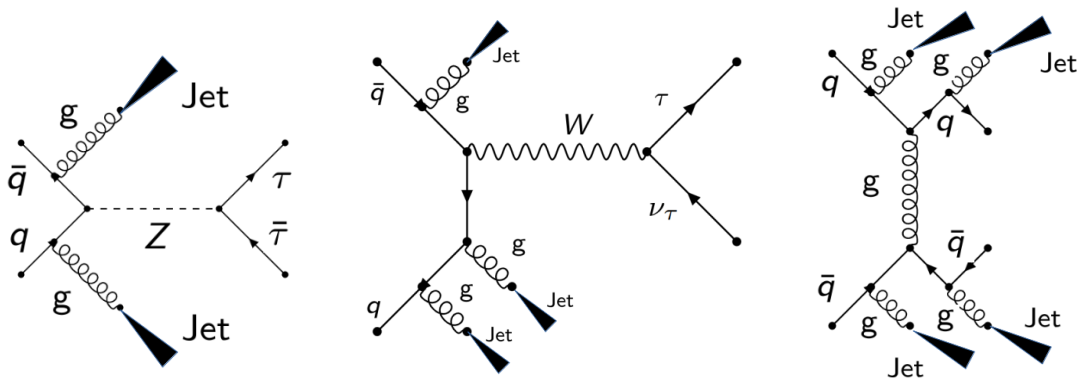


Figure 6.6: This plot shows Feynman diagrams for this study's major background processes: DY+Jets (Left), W+Jets (Middle), and QCD (Right).

using a single muon control region to measure SF_{VBF}, since the $W \to \tau\nu$ vs. $W \to \mu\nu$ decay mode does not change the VBF efficiency.

Before measuring this SF_{VBF} for W+jets, we first confirm good modeling of the central selections by the W+jets MC. With this strategy, the estimate of the W+jets background yield in the SR can be calculated using the following equation:

$$N_{wjets}^{SR,MC} = N_{wjets}^{SR,\mathrm{MC}} \times SF_{VBF} \tag{6.7}$$

The number of MC events in the W+jets control region can be written as follows:

$$N_{wjets,MC} = \sigma_{MC} \times Lumi_{MC} \times eff_{DiTauCuts,MC} \times eff_{VBFCuts,MC} \tag{6.8}$$

The cross-section and luminosity are already measured by other CMS groups, so $\sigma_{Data} = \sigma_{MC}$ and $Lumi_{Data} = Lumi_{MC}$. Thus,

$$N_{wjets,Data} = \sigma_{Data} \times Lumi_{Data} \times eff_{DiTauCuts,Data} eff_{VBFCuts,Data} \tag{6.9}$$

Dividing these equations, we get the following:

$$\frac{N_{wjets(SR),Data}}{N_{wjets(SR),MC}} = \frac{eff_{DiTauCuts(Data)}}{eff_{DiTauCuts(MC)}} \times \frac{eff_{VBFCuts(Data)}}{eff_{VBFCuts(MC)}} = SF_{DitauCuts} \times SF_{VBFCuts} \tag{6.10}$$

This equation allows us to find the $N_{wjets}^{SR,Data}$ given $N_{wjets}^{SR,MC}$, $SF_{DitauCuts}$, and $SF_{VBFCuts}$.

Additionally, we can use the control region containing a single muon to find the VBF scale factor:

$$N_{wjets(CR),MC} = \sigma_{MC} \times Lumi_{MC} \times eff_{SingleMuonCuts,MC} \times eff_{VBFCuts,MC} \tag{6.11}$$

and

$$N_{wjets(CR),Data} = \sigma_{Data} * Lumi_{Data} \times eff_{SingleMuonCuts,Data} \times eff_{VBFCuts,Data} \tag{6.12}$$

Similarly to the above procedure, we divide the two equations to find

$$\frac{N_{wjets(CR),Data}}{N_{wjets(CR),MC}} = SF_{SingleMuonCuts} \times SF_{VBFCuts} \tag{6.13}$$

We may therefore use the single-muon CR to calculate a VBF scale factor. To do this, however, we must first factor out the contribution of $SF_{SingleeMuonCuts}$. Thus, I ran the single muon control region with failing and passing VBF cuts to find $SF_{SingleMuonCuts}$.

Small backgrounds such as $t\bar{t}$ were estimated using simulation. Where possible, we used validation samples for the small backgrounds to assign systematic uncertainty and/or give additional confidence that these small backgrounds can be taken from simulation. QCD multijet background will be estimated in a fully data-driven way using like-sign ditau events, re-weighted using an OS-to-LS transfer factor:

$$N_{QCD}^{SR,\text{ estimate}} = N_{QCD}^{LS,\text{ CR}} \times R_{OS/LS} \tag{6.14}$$

In our study, $R_{OS/LS}$ is measured using low-E_T^{miss} events with inverted τ_h isolation.

6.7 Control Region Results

The following section contains the results of the background estimation studies.

6.7.1 Estimated DY+Jets Background using a Z to Ditau Control Region

The first $Z \to \tau\tau$ control region is created using analogous selections to those of the signal region, except we require a mass selection m$(\tau, \tau, E_T^{miss}) <$ 150 GeV and inverted VBF selections. Events passing the "inverted VBF criteria" are the set of events that have 0 dijet candidates passing the VBF criteria. This can be any event which does not have at least two jets passing the p_T, η, and ID cuts, or an event with two good jets, but all of the dijet pairs fail either the opposite hemisphere requirement, the dijet mass requirement, or the $\Delta\eta_{jj}$ cut. The 2016 results of this study are shown in Figure 6.7 and Table 6.9. Given these results, the 2016 scale factor was calculated to be $SF_{non-VBF} = 1.02 \pm 0.06$.

The second $Z \to \tau\tau$ control region is the same region as the first (including the ditau mass selection), except with passing VBF selections. The 2016 results are shown in Figure 6.8. There is good level of agreement between data and simulation, within the statistical uncertainty (about 20%), with a corresponding scale factor of $SF_{VBF} = 0.96 \pm 0.2$.

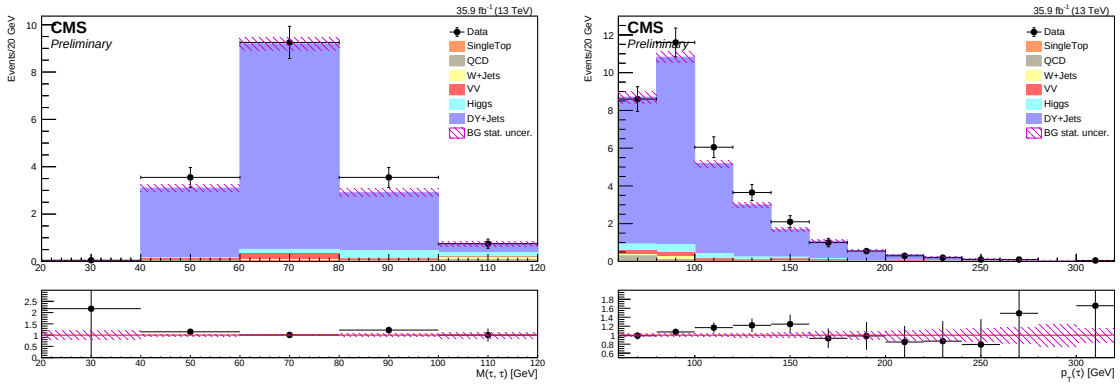


Figure 6.7: (Left) 2016 $m(\tau_h, \tau_h)$ plot for the $Z \to \tau\tau$ inverse VBF control region. (Right) $p_T(\tau_h)$ distribution of $Z \to \tau\tau$ inverse VBF control region.

Table 6.9: Event output table for the $Z \to \tau\tau$ control regions for 2016, given $m(\tau_h, \tau_h, E_T^{miss}) < 150$ GeV.

Process	Pass VBF, OS	Inverse VBF, LS	Inverse VBF, OS
Data	31.0	9.0	343.0
DY + jets	30.7 ± 2.4	5.3 ± 0.9	289.7 ± 7.4
SingleTop	0.0 ± 0.0	0.3 ± 0.2	0.2 ± 0.2
VV	1.0 ± 0.2	0.3 ± 0.0	8.7 ± 0.8
W + jets	0.5 ± 0.4	0.6 ± 0.3	2.9 ± 1.0
Higgs	-	-	15.0 ± 0.4
QCD	-	-	2.7 ± 0.0
SF	0.96 ± 0.2	-	1.02 ± 0.06

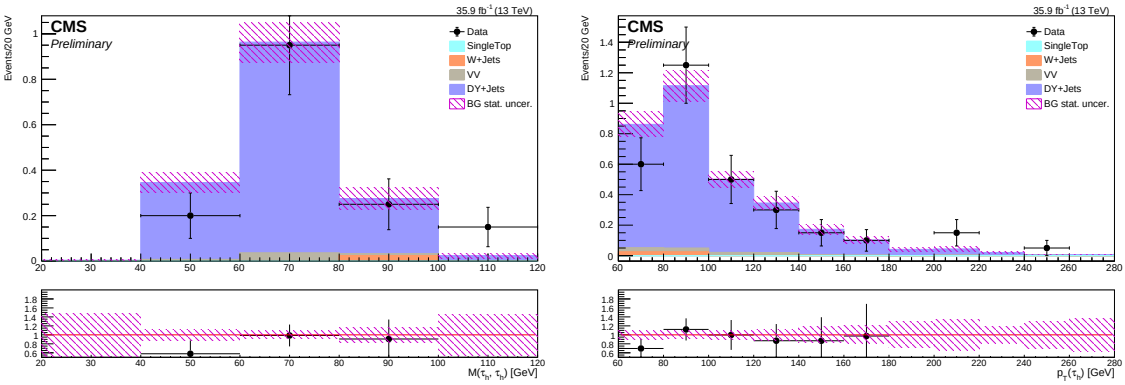


Figure 6.8: The $m(\tau_h, \tau_h)$ (left) and $\eta(\tau_h)$ (right) plots for the $Z \to \tau\tau$ control region passing VBF selections control region.

The results for the DY+jets studies using 2017 MC and data samples are shown in Figures 6.9-6.15 and Table 6.10. Note that the bottom plots represent the ratio pads, which plot the data over the predicted background. The results using 2018 MC and data samples are shown in Figures 6.16-6.18 and Table 6.11. The 2017 and 2018 studies show similar conclusions as the 2016 studies, with scale factors that are consistent with unity within the statistical uncertainties. We remind the reader that the *SF*'s from Tables 29-31 are used

to estimate the DY+jets background yield in the signal region.

We also note that the QCD contribution to the "Pass VBF" CR is negligible. We confirm this by observing a data yield of 0 in the like-sign Pass VBF samples. The QCD yield in the "Inverse VBF" DY+jets control is still small ($\approx 3\%$), but non-zero. Although small, this contribution is estimated in a similar fashion as described previously, using the QCD yields in the like-sign control regions scaled by the OS-to-LS transfer factor (we'll discuss this procedure in more detail in sections to follow).

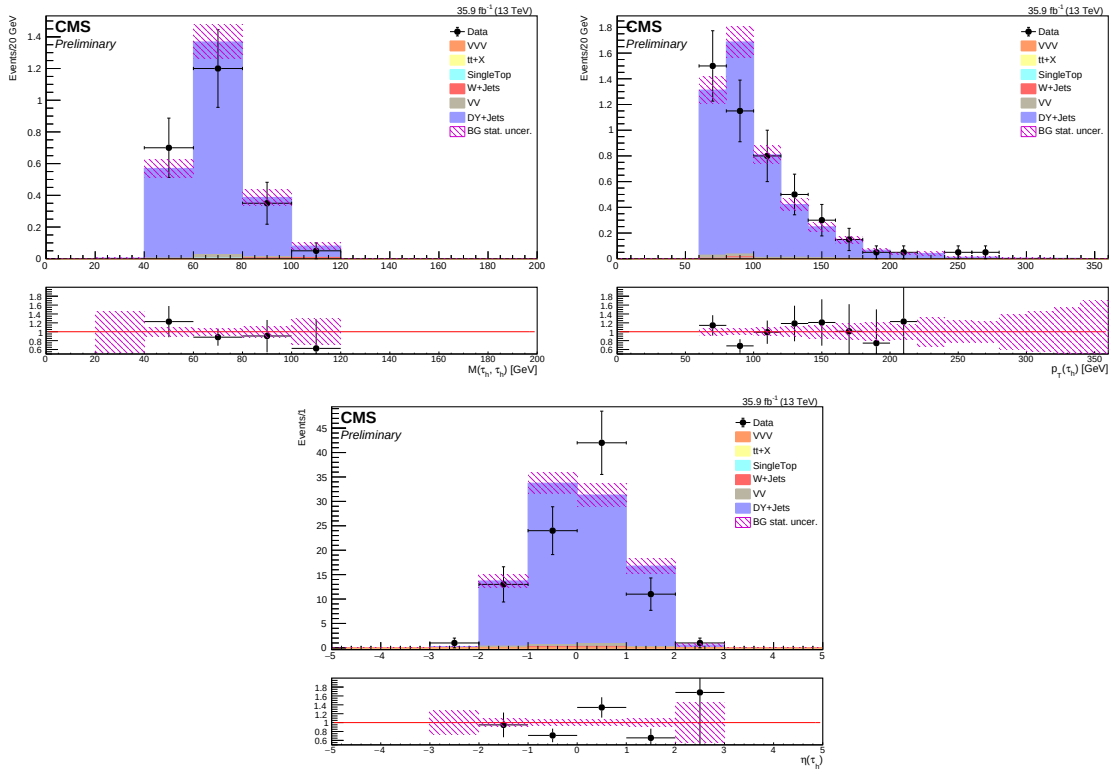


Figure 6.9: These plots represent the 2017 $Z \to \tau\tau$ control region for passing VBF with opposite sign $\tau_h\tau_h$ in the final state. (Top Left) $m(\tau_h, \tau_h)$ plot. (Top Right) $p_T(\tau_h)$ plot. (Bottom) $\eta(\tau_h)$ plot.

6.7.2 Estimating W+Jets Background using a Single Muon Control Region

To estimate the W+Jets background in the $\tau_h\tau_h$ signal region, we used a single muon control region. This control region was chosen to minimize the effects of QCD (QCD can't be entirely decoupled from W+jets in the $\tau_h\tau_h$ channel) and to extract VBF efficiency scale factors for Z/W + jets. This control region was created by using the SingleMuon PD and selecting events passing a single-muon trigger with p_T threshold of 24 or 27 (depending on the year), and requiring one muon with $|\eta| < 2.1$, $p_T > 35$ GeV, passing tight ID, and $60 < m_T(\mu, E_T^{miss}) < 100$. The passing VBF selections requirement is determined the same as for the signal region selections. The results of this background study are shown in Figures 6.15-6.16 and Table 6.12 for

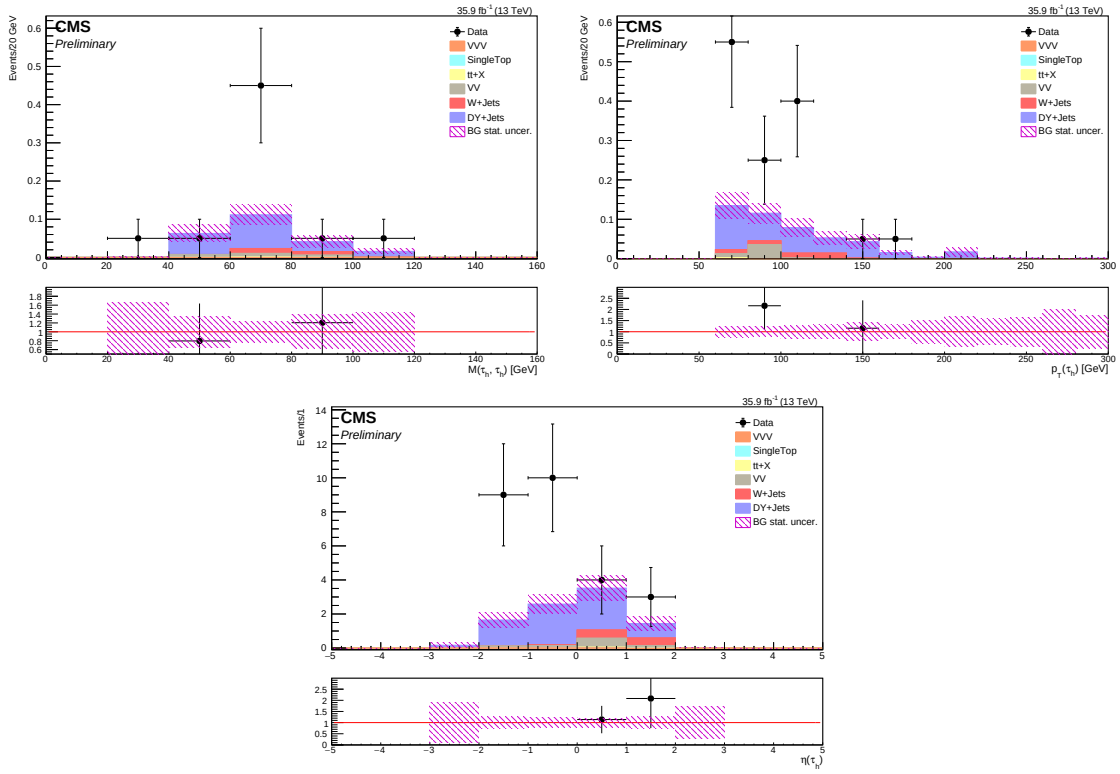


Figure 6.10: These plots represent the 2017 $Z \to \tau\tau$ control region for inverse VBF with like-sign $\tau_h\tau_h$ in the final state. (Top Left) $m(\tau_h, \tau_h)$ plot. (Top Right) Tau $p_T(\tau_h)$ plot. (Bottom) $\eta(\tau_h)$ plot.

Table 6.10: Event output table for the $Z \to \tau\tau$ inverse VBF control region for 2017, given $m(\tau_h, \tau_h, E_T^{miss}) <$ 150 GeV.

Process	Pass VBF, OS	Inverse VBF, LS	Inverse VBF, OS
Data	46.0	13.0	455.0
DY + jets	47.2 ± 3.0	3.7 ± 0.8	402.7 ± 9.2
SingleTop	0.0 ± 0.0	0.0 ± 0.0	0.2 ± 0.2
VV	0.7 ± 0.1	0.5 ± 0.2	8.8 ± 0.6
VVV	0.0 ± 0.0	0.0 ± 0.0	0.3 ± 0.1
W + jets	0.2 ± 0.1	0.5 ± 0.3	5.6 ± 2.0
tt+X	0.0 ± 0.0	0.0 ± 0.0	0.2 ± 0.0
Higgs	-	-	17.3 ± 0.7
QCD	-	-	10.2 ± 0.6
SF	0.95 ± 0.16	-	1.08 ± 0.07

2016 MC samples and data, 6.17-6.18 and Table 6.13 for 2017 MC samples and data, and 6.19-6.20 and Table 6.14 for 2018 MC sample and data.

We note that the W + jets background yields shown in Tables 6.12- 6.14 are derived directly from the MC samples.

The QCD multijet background is not shown in Tables 6.12- 6.14 because it is negligible after VBF selec-

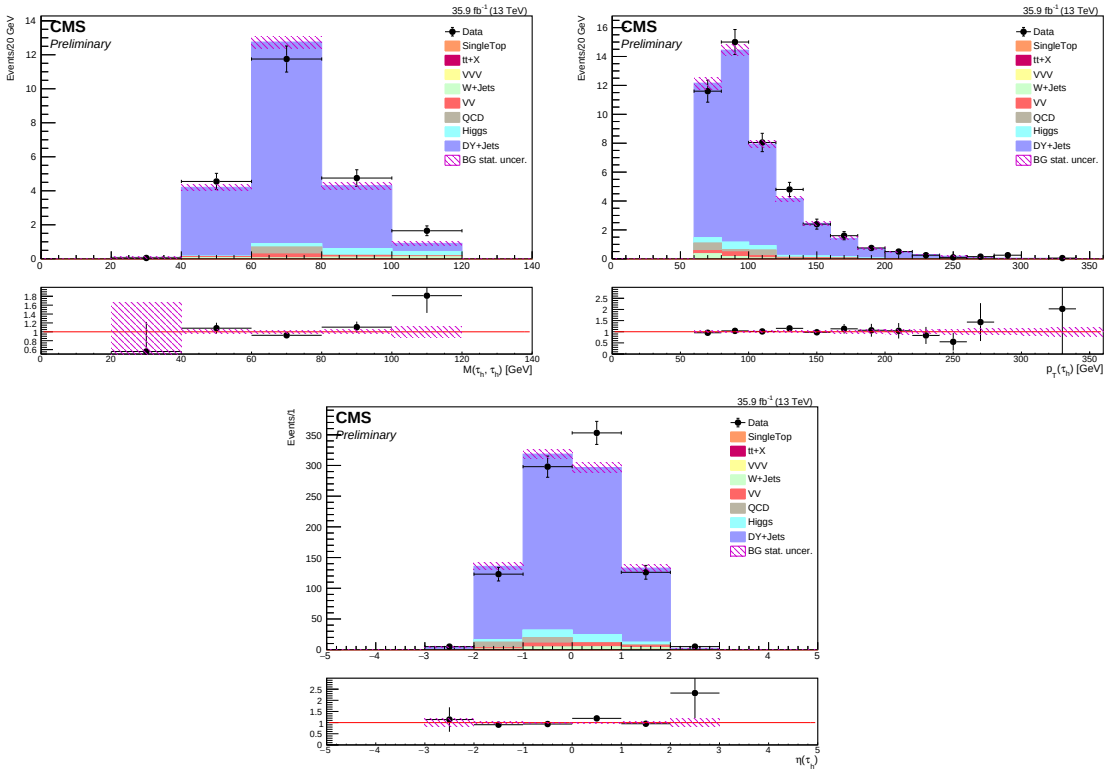


Figure 6.11: These plots represent the 2017 $Z \to \tau\tau$ control region for inverse VBF with opposite sign $\tau_h\tau_h$ in the final state. (Top Left) $m(\tau_h, \tau_h)$ plot. (Top Right) $p_T(\tau_h)$ plot. (Bottom) $\eta(\tau_h)$ plot.

Table 6.11: Event output table for the $Z \to \tau\tau$ inverse VBF control region for 2018, given $m(\tau_h, \tau_h, E_T^{miss}) <$ 150 GeV.

Process	Pass VBF, OS	Inverse VBF, LS	Inverse VBF, OS
Data	66	17	643
DY + jets	55.2 ± 3.6	7.1 ± 1.4	636.9 ± 13.3
Single Top	0.0 ± 0.0	0.1 ± 0.1	1.0 ± 0.5
VV	1.5 ± 0.3	0.4 ± 0.2	16.0 ± 0.8
VVV	0.0 ± 0.0	0.0 ± 0.0	0.3 ± 0.1
W+Jets	0.0 ± 0.0	1.4 ± 0.9	5.3 ± 2.2
tt+X	0.1 ± 0.0	0.0 ± 0.0	0.3 ± 0.0
QCD	-	-	9.3 ± 0.3
SF	1.17 ± 0.16	-	0.96 ± 0.04

tions, because of the low jet$\to \mu$ fake rate, m_T mass window cut around m_W, and the power of the VBF cuts at suppressing QCD. There are 0 events passing cuts with the QCD MC samples, with a very small error on 0. We have also validated that QCD is indeed small in this CR using non-isolated muon samples scaled by isolation transfer factors.

From Figures 6.15- 6.20 and Tables 6.12-6.14, one can see that the Inverse VBF CRs show good agreement between MC and data, and thus the central selection efficiencies are well modeled, with scale factors

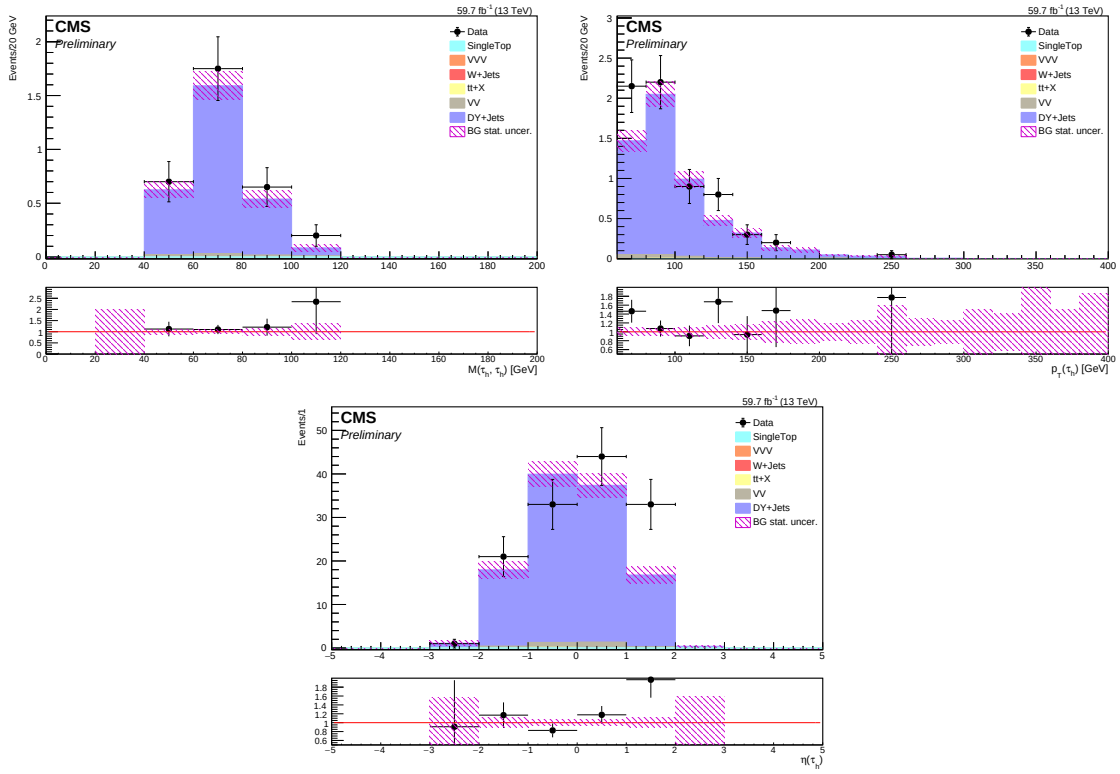


Figure 6.12: These plots represent the 2018 $Z \rightarrow \tau\tau$ control region for passing VBF with opposite-sign taus in the final state. (Top Left) $m(\tau_h, \tau_h)$ plot. (Top Right) $p_T(\tau_h)$ plot. (Bottom) $\eta(\tau_h)$ plot.

Table 6.12: Event yield and scale factor obtained from W($\rightarrow \mu\nu$)+Jets control region in 2016.

Process	2016 Inverse VBF	2016 Pass VBF
Data	5364081	433635
DY + jets	1039045.8 $\pm$ 1136.4	94218.6 $\pm$ 339.9
Single Top	69673.6 $\pm$ 96.1	10702.8 $\pm$ 33.4
VV	50472.4 $\pm$ 87.2	3609.9 $\pm$ 22.9
W+Jets	4374239.3 $\pm$ 7393.4	430922.3 $\pm$ 2566.6
SF	0.96 $\pm$ 0.002	0.75 $\pm$ 0.005

Table 6.13: Event yield and scale factor obtained from W($\rightarrow \mu\nu$)+Jets control region in 2017.

Process	2017 Inverse VBF	2017 Pass VBF
Data	6471919	555089
DY + jets	1610979.0 $\pm$ 1861.6	147009.3 $\pm$ 552.3
Single Top	64651.2 $\pm$ 89.2	8460.3 $\pm$ 27.0
VV	55029.5 $\pm$ 96.7	3878.7 $\pm$ 25.7
VVV	319.1 $\pm$ 3.1	30.4 $\pm$ 1.0
W+Jets	4883104.2 $\pm$ 9256.7	534785.2 $\pm$ 3385.4
tt+X	674.1 $\pm$ 3.5	81.8 $\pm$ 1.3
SF	0.971 $\pm$ 0.002	0.740 $\pm$ 0.005

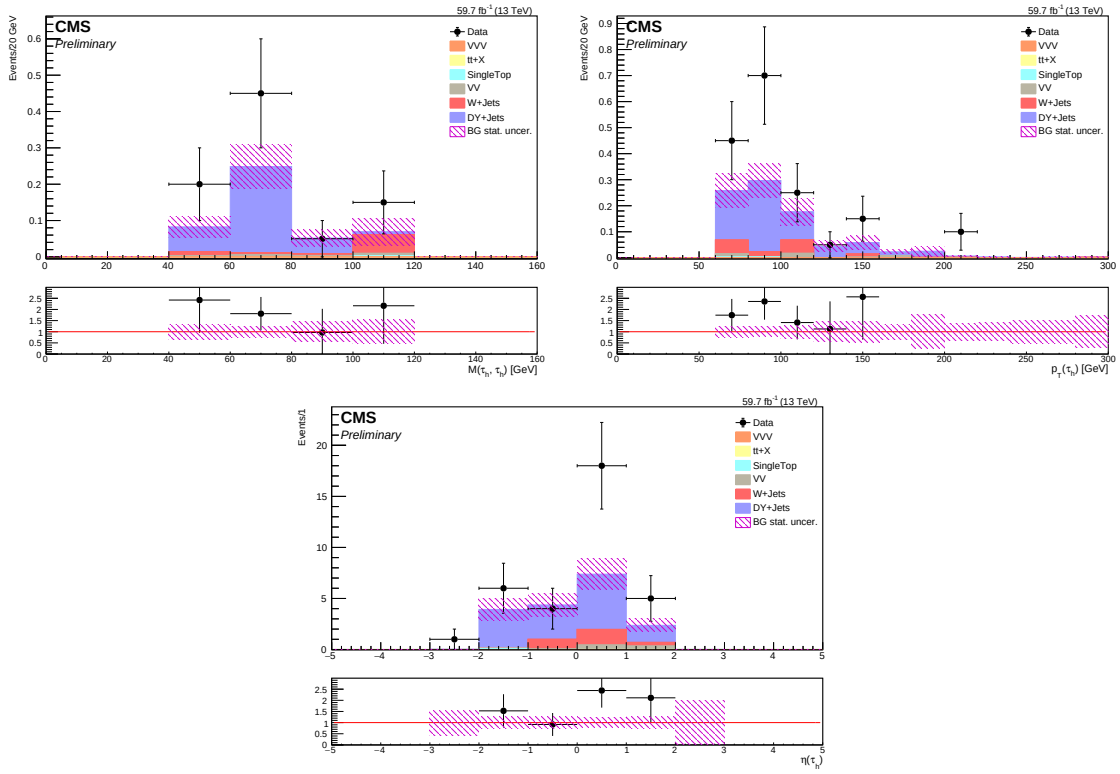


Figure 6.13: These plots represent the 2018 $Z \to \tau\tau$ control region for inverse VBF with like-sign $\tau_h\tau_h$ in the final state. (Top Left) $m(\tau_h, \tau_h)$ plot. (Top Right) $p_T(\tau_h)$ plot. (Bottom) $\eta(\tau_h)$ plot.

Table 6.14: Event yield and scale factor obtained from W($\to \mu\nu$)+Jets control region in 2018.

Process	2018 Pass VBF
Data	705474
DY + jets	125606.8 $\pm$ 580.4
Single Top	10917.0 $\pm$ 34.5
VV	3919.2 $\pm$ 20.2
VVV	37.0 $\pm$ 1.3
W+Jets	520215.9 $\pm$ 3714.2
tt+X	115.7 $\pm$ 2.1
SF	1.085 $\pm$ 0.008

near unity for all years. On the other hand, and as expected, the Pass VBF CRs show some disagreement between MC and data. The level of agreement varies depending on the year. The VBF efficiency scale factors are about 0.74-0.75 for 2016 and 2017, but $SF_{VBF} = 1.09 \pm 0.01$ in 2018. The improved modeling in 2018 is due to improved jet energy resolution scale factors for low-p_T forward jets in 2018. In 2016-2017, forward jet energy resolution factors are not p_T-dependent, whereas the 2018 jet energy resolution factors for high-η jets were indeed measured as a function of p_T. This is clearly observed the low-p_T jet disagreement in Figures 6.16 and 6.18 for 2016 and 2017 data. Figure 6.20 for 2018 shows better agreement at low $p_T(j)$.

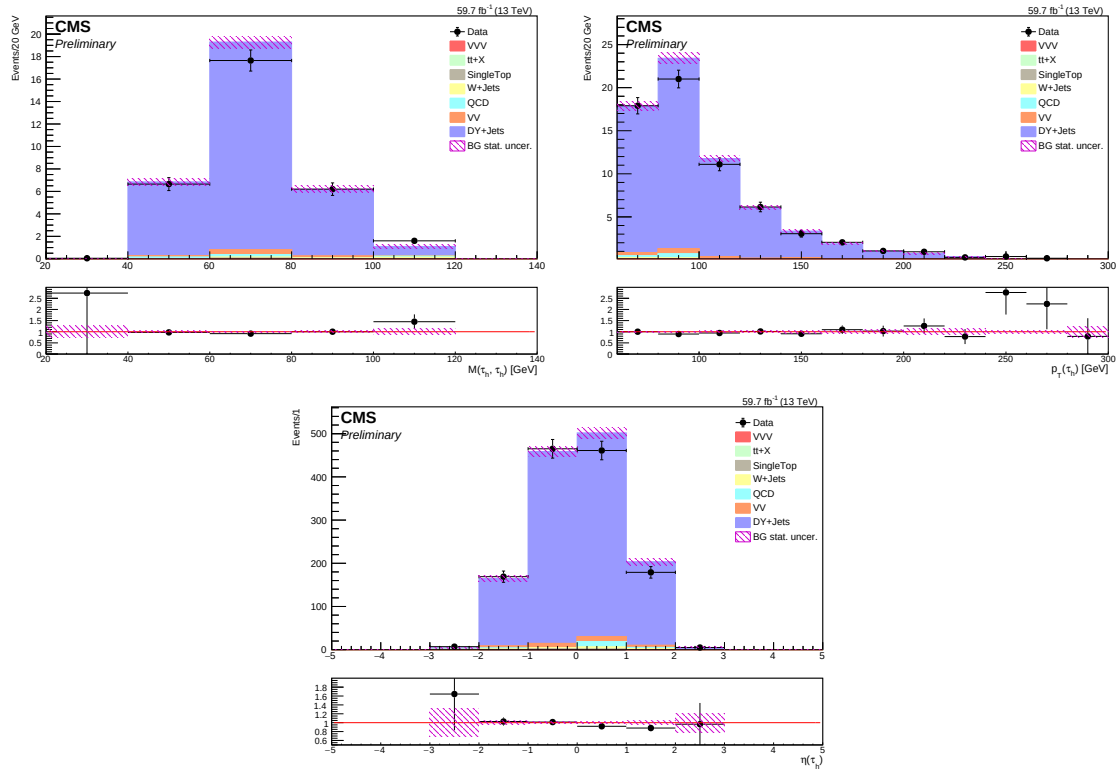


Figure 6.14: These plots represent the 2018 $Z \rightarrow \tau\tau$ control region for inverse VBF with opposite-sign taus in the final state. (Top Left) $m(\tau_h, \tau_h)$ plot. (Top Right) $p_T(\tau_h)$ plot. (Bottom) $\eta(\tau_h)$ plot.

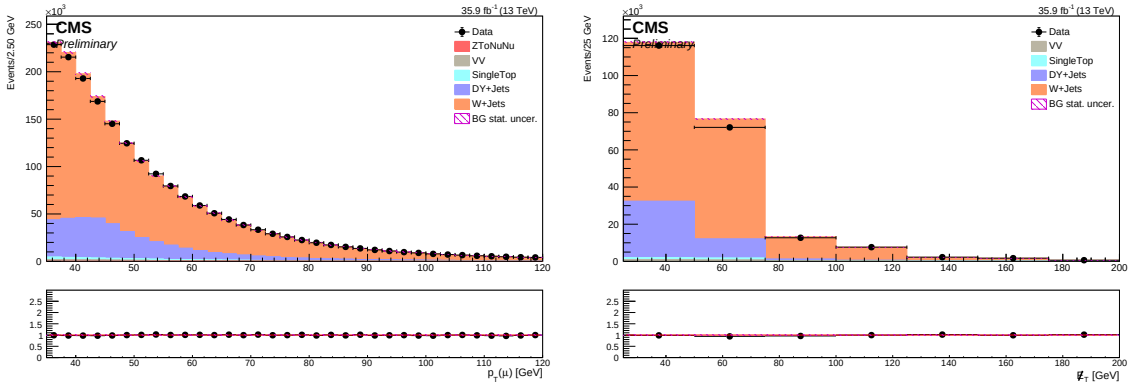


Figure 6.15: These plots represent the 2016 inverse VBF single μ CR. (Left) Muon p_t plot. (Right) E_T^{miss} *plot*.

6.7.3 Estimating QCD using ABCD Control Region Method

In the non-VBF $Z' \rightarrow \tau\tau$ searches, QCD is a major background process. Although the VBF selections help significantly reduce the QCD contribution, it is still necessary to estimate this background using a data-driven approach. We use an ABCD method to estimate this background: Region A is the signal region (SR); Region B is defined with the same selections as the SR, except that it requires a like-sign (LS) ditau pair rather than an opposite-sign (OS) pair; Region C requires an OS ditau pair and uses events that fail the VBF criteria;

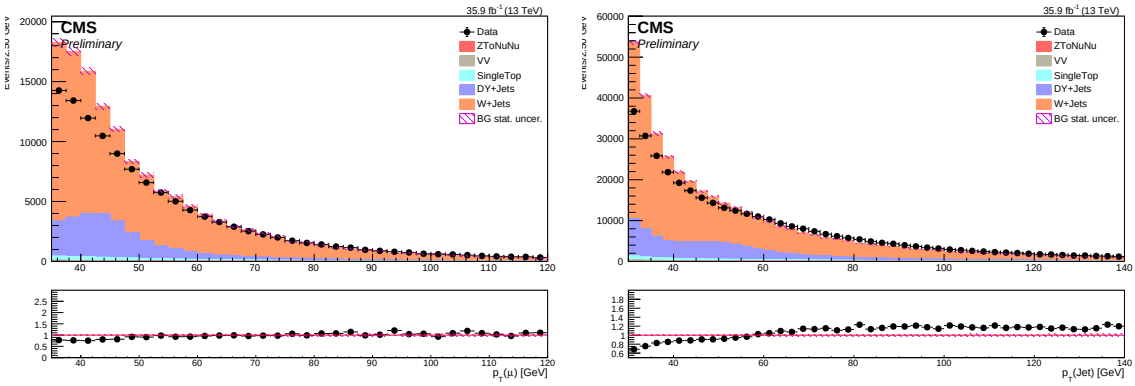


Figure 6.16: These plots represent the 2016 passing VBF single muon CR. (Left) Muon p_t. (Right) E_T^{miss} plot.

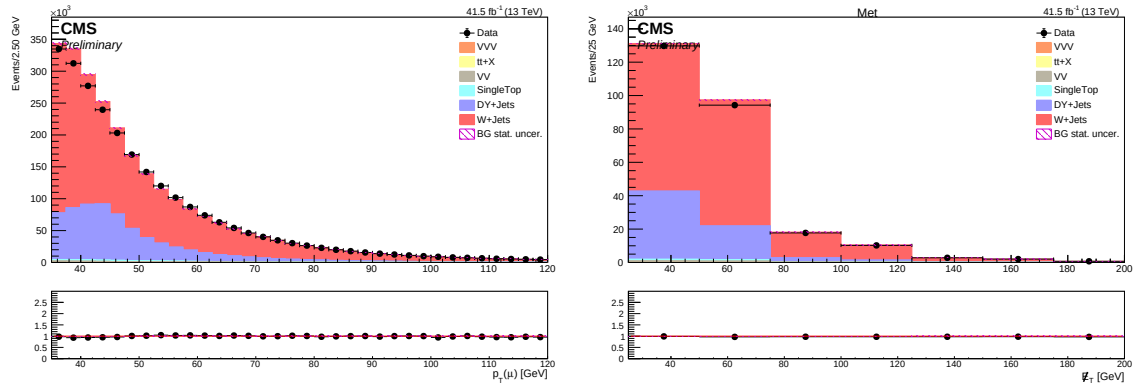


Figure 6.17: These plots represent the 2017 inverse VBF single muon CR. (Left) Muon p_t. (Right) E_T^{miss} plot.

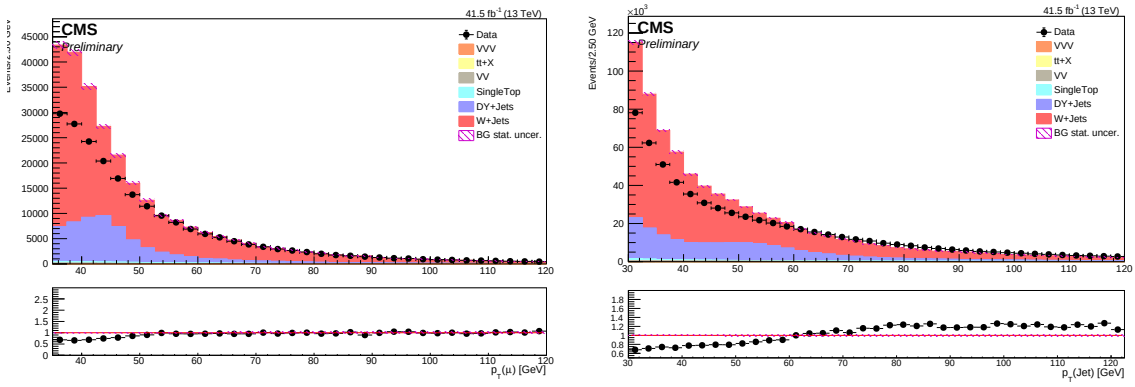


Figure 6.18: These plots represent the 2017 passing VBF single muon CR. (Left) Muon p_t. (Right) Jet p_T plot.

and Region D requires LS ditau pairs that fail VBF criteria. This array of regions can be visualized using Figure 6.21. To estimate the QCD background, then, we calculate an OS-to-LS transfer factor ($R_{OS/LS}$) using Regions D and C, and then we use that factor to scale down the QCD yield determined from Region B. The

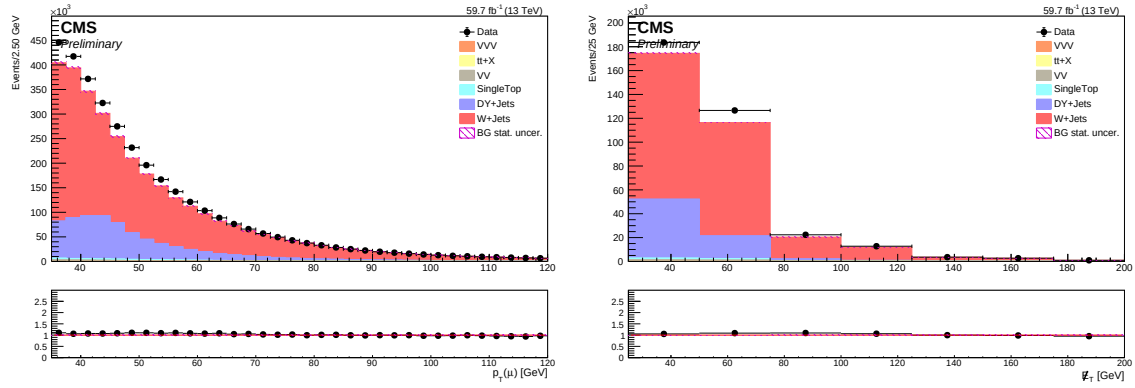


Figure 6.19: These plots represent the 2018 inverse VBF single muon CR. (Left) Muon p_t. (Right) E_T^{miss} plot.

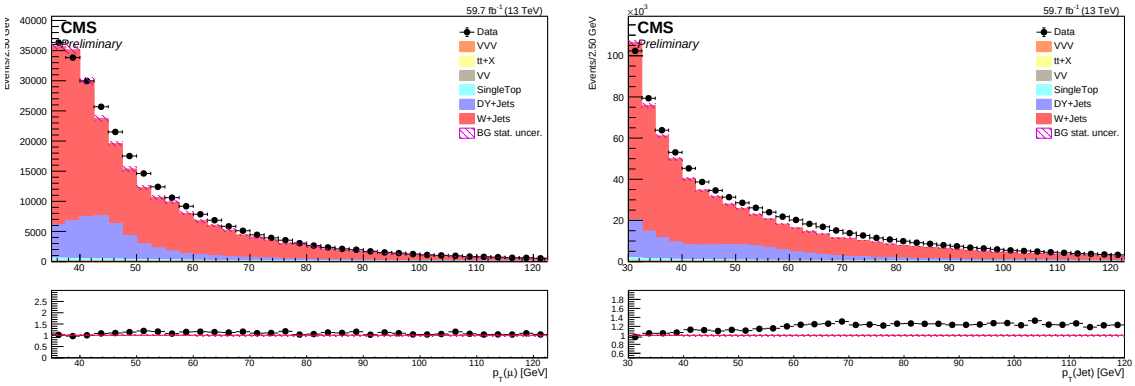


Figure 6.20: These plots represent the 2018 passing VBF single muon CR. (Left) Muon p_t. (Right) Jet p_T plot.

result estimates the QCD background for Region A, the signal region.

Region B results using 2016 MC and data samples are shown in Figure 6.22 and Table 6.15. From these results, it is clear that Region B has high purity of QCD. The QCD yield and kinematics shapes in Region B are determined by subtracting non-QCD contributions (using MC) from the data.

Similarly, 2016 Region D results are shown in Figure 6.23 and Table 6.15, and 2016 Region C results are shown in Figure 6.24 and Table 6.15. To ensure high QCD purity, an inverted tau isolation was required for these control regions. As with Region B, the QCD yield and kinematics shapes in these regions were determined by subtracting non-QCD contributions from the data.

Figures 6.25-6.26 and Table 6.16 show the QCD Regions B, C, and D results for 2017 MC and data samples. These results show similar results between run periods.

Figures 6.27-6.28 and Table 6.17 show the QCD Region B, D, and C results for 2018 MC and data samples. These results also show similar results between run periods.

Given the results from Region C and Region D, a transfer factor $R_{OS/LS}$ was calculated using the following

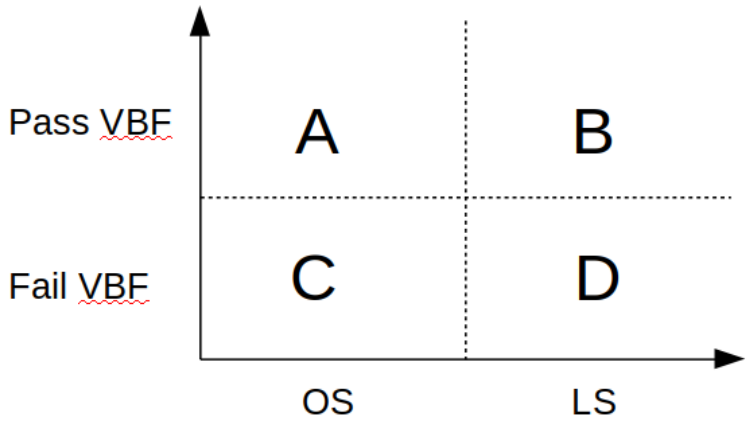


Figure 6.21: Visualization of the ABCD control region method used for the QCD background estimation. Note that Region A is the signal region. Regions C and D require the fail VBF and inverted tau isolation criteria to ensure high QCD purity and are used to develop a transfer factor that, when applied to Region B, estimates the QCD background contribution in Region A.

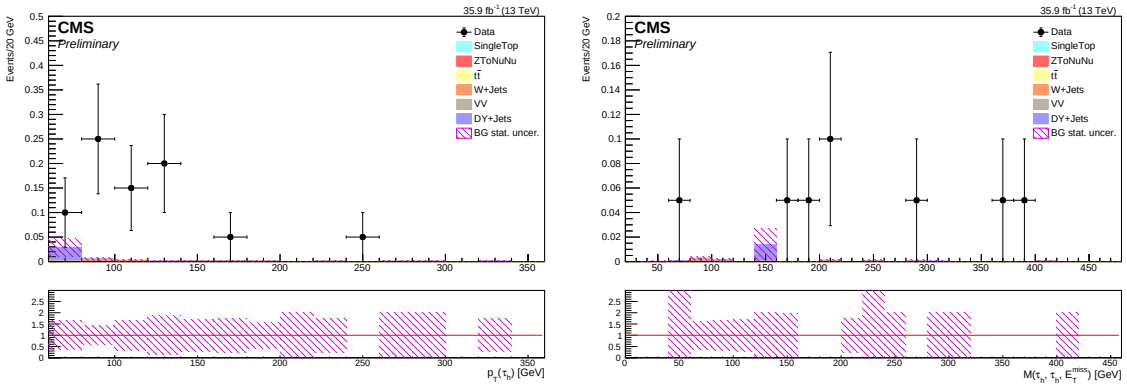


Figure 6.22: Tau p_T (left) and ditau mass (right) plots for QCD Region B using 2016 MC and data samples.

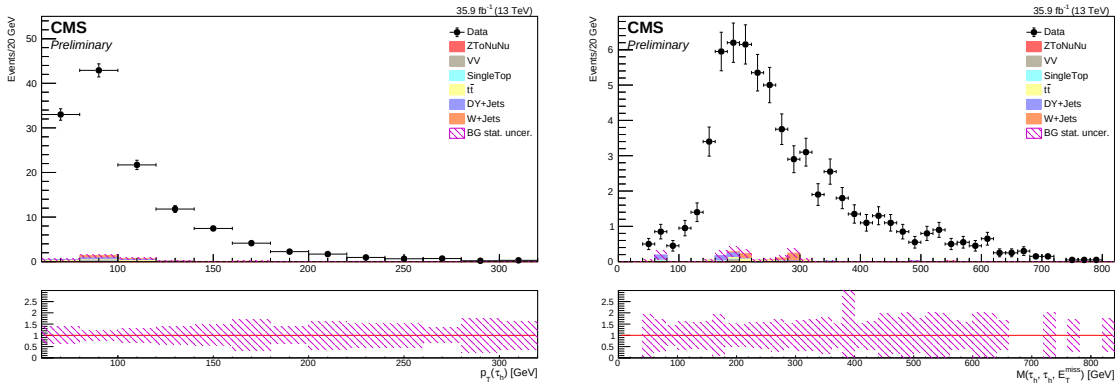


Figure 6.23: Tau p_T (left) and ditau mass (right) plots for QCD Region D using 2016 MC and data samples.

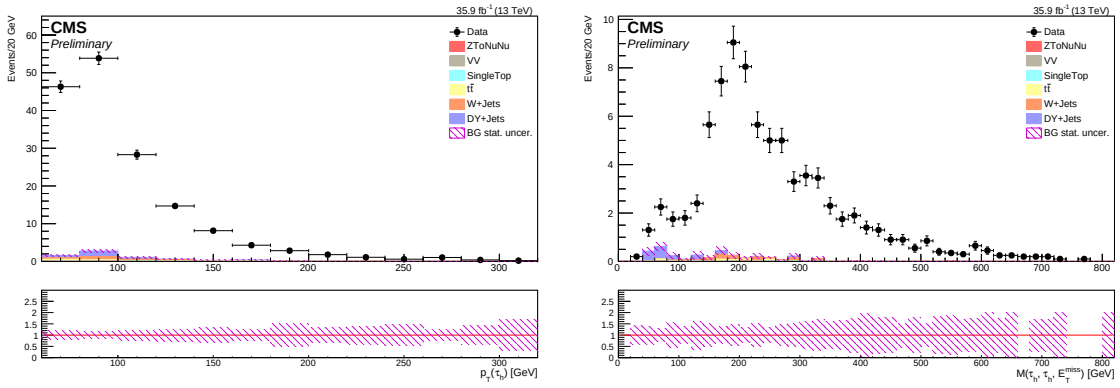


Figure 6.24: Tau p_T (left) and ditau mass (right) plots for QCD Region C using 2016 MC and data samples.

Table 6.15: Event output table for 2016 QCD Regions B, D, and C.

Process	QCD CR B	QCD CR C	QCD CR D
Data	14.0	2565.0	2006.0
DY + jets	1.1 ± 0.4	49.7 ± 3.9	13.6 ± 2.0
Single Top	0.1 ± 0.1	4.9 ± 0.9	1.2 ± 0.4
VV	0.2 ± 0.0	5.5 ± 1.3	1.2 ± 0.4
W+Jets	1.6 ± 0.8	42.7 ± 4.3	9.6 ± 1.9

Table 6.16: Event output table for 2017 QCD Regions B, D, and C.

Process	QCD CR B	QCD CR C	QCD CR D
Data	13.0	3084.0	2238.0
DY + jets	1.4 ± 0.5	47.2 ± 4.0	13.9 ± 2.0
Single Top	0.0 ± 0.0	3.2 ± 0.7	1.7 ± 0.5
VV	0.1 ± 0.0	4.0 ± 0.8	2.2 ± 1.1
VVV	0.0 ± 0.0	0.1 ± 0.1	0.0 ± 0.0
W+Jets	1.2 ± 0.8	47.1 ± 6.7	12.9 ± 2.9
tt+X	0.0 ± 0.0	0.0 ± 0.1	0.0 ± 0.0

Table 6.17: Event output table for 2018 QCD Regions B, D, and C.

Process	QCD CR B	QCD CR C	QCD CR D
Data	21	3004	2288
DY + jets	3.3 ± 1.2	74.2 ± 6.2	14.9 ± 2.5
Single Top	0.0 ± 0.0	4.0 ± 0.9	1.4 ± 0.5
VV	0.2 ± 0.1	3.5 ± 0.5	1.3 ± 0.3
VVV	0.0 ± 0.0	0.1 ± 0.1	0.0 ± 0.0
W+Jets	0.5 ± 0.5	49.2 ± 7.6	11.2 ± 3.1
tt+X	0.0 ± 0.0	0.0 ± 0.1	0.2 ± 0.1

equation:

$$R_{OS/LS} = \frac{N^C_{\text{Data}} - N^C_{\text{non-QCD, MC}}}{N^D_{\text{Data}} - N^D_{\text{non-QCD, MC}}} \tag{6.15}$$

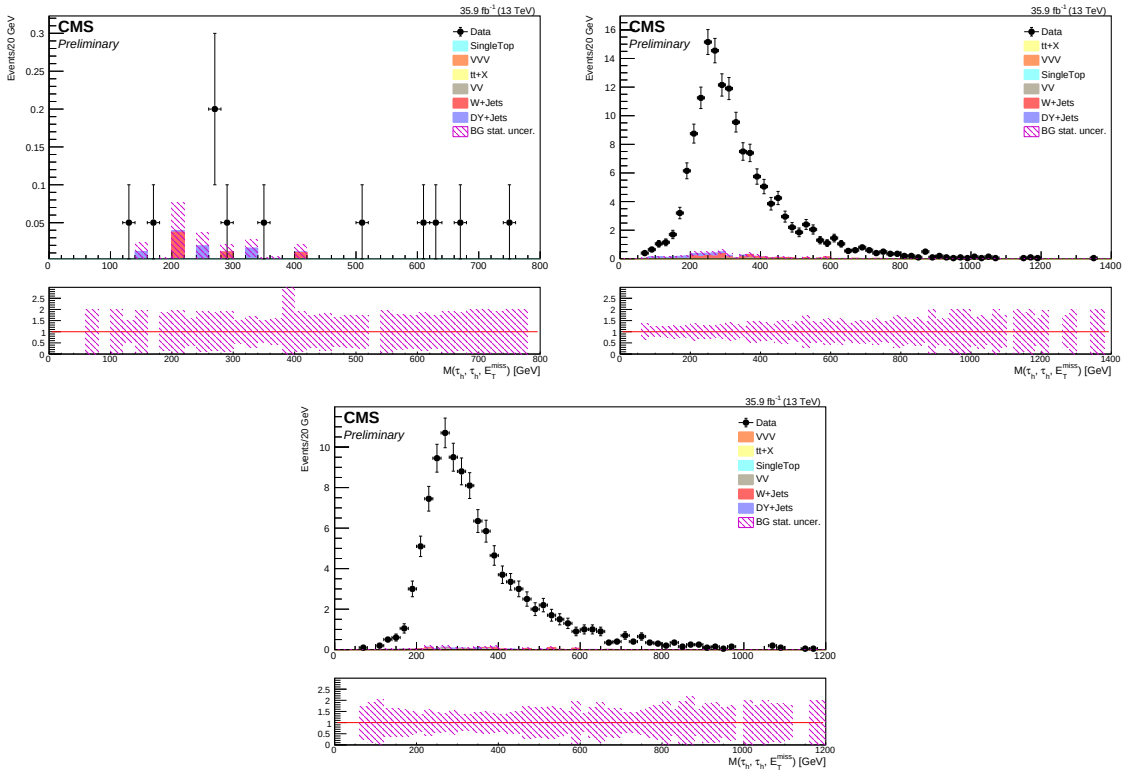


Figure 6.25: These plots represent the ditau mass for 2017 results for control regions B, C, and D. (Top Left) Region B. (Top Right) Region C. (Bottom) Region D.

The result of this calculation is shown in Figure 6.29. The transfer factor has a $m(\tau_h, \tau_h, E_T^{miss})$ dependence. Thus, the transfer factor is derived as a function of $m(\tau_h, \tau_h, E_T^{miss})$ by fitting the $R_{OS/LS}$ vs. $m(\tau_h, \tau_h, E_T^{miss})$ points to a second-degree polynomial. This mass-dependent factor can be applied to QCD Region B to estimate the QCD background of the signal region. However, the shape of the QCD $m(\tau_h, \tau_h, E_T^{miss})$ distribution is derived from Region D, which contains more statistics than Region B. A dedicated closure test is performed for our QCD background estimation strategy, and is discussed in Appendix 6.8.1.

The ABCD technique for predicting QCD has additional challenges when calculating uncertainty. The total error was calculated using the following equation:

$$Err_{total} = \sqrt{Err^2_{Stat.err.,B} + Err^2_{Stat.err.,D} + Err^2_{t_f}}, \tag{6.16}$$

where $Err_{Stat.err.,B}$ is the statistical error in Region B, $Err_{Stat.err.,D}$ is the statistical error in Region D, and Err_{t_f} is the error in transfer factor.

Since the transfer factor is represented by a second-degree polynomial ($f(x) = a + bx + cx^2$), its uncer-

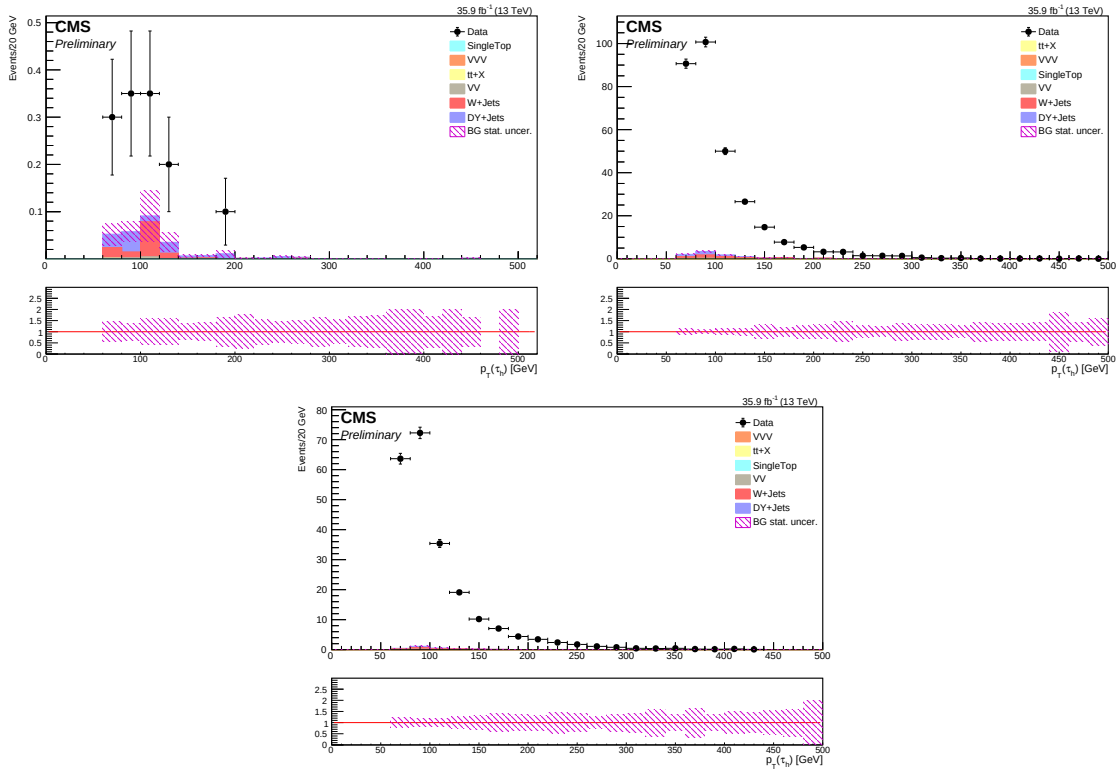


Figure 6.26: These plots represent the tau p_t for 2017 results for control regions B, C, and D. (Top Left) Region B. (Top Right) Region C. (Bottom) Region D.

tainty must be propagated as follows:

$$\partial f = \sqrt{(\frac{\partial f}{\partial a})^2\delta_a^2 + (\frac{\partial f}{\partial b})^2\delta_b^2 + (\frac{\partial f}{\partial c})^2\delta_c^2 + 2\frac{\partial f}{\partial a}\frac{\partial f}{\partial b}\delta_{ab} + 2\frac{\partial f}{\partial a}\frac{\partial f}{\partial c}\delta_{ac} + 2\frac{\partial f}{\partial b}\frac{\partial f}{\partial c}\delta_{bc}} \tag{6.17}$$

such that Err_{t_f} can be calculated using

$$Err_{t_f} = \sqrt{Err_a^2 + x^2 Err_b^2 + x^4 Err_c^2 + 2x\delta_{ab} + 2x^2\delta_{ac} + 2x^3\delta_{bc}}. \tag{6.18}$$

Consider, for instance, Bin #11 in Figure 6.33. Here, $Err_{Stat.err.,B}$ is 19.7%, $Err_{Stat.err.,D}$ is 10.85%, and Err_{t_f} is 6.49%. Therefore, we calculate the total error in QCD for Bin #11 to be $\sqrt{(0.197)^2 + (.1085)^2 + (0.0649)^2} = 23.4\%$. This procedure was used to calculate the error in QCD for all bins.

6.8 Closer Tests

Additionally, closure tests were performed to prove that our background estimation methods are accurate. These methods choose orthogonal control regions that should have no enrichment of the $Z' \to \tau_h\tau_h$ signal and compare the predicted background measurements to actual data. Since this data should contain only the

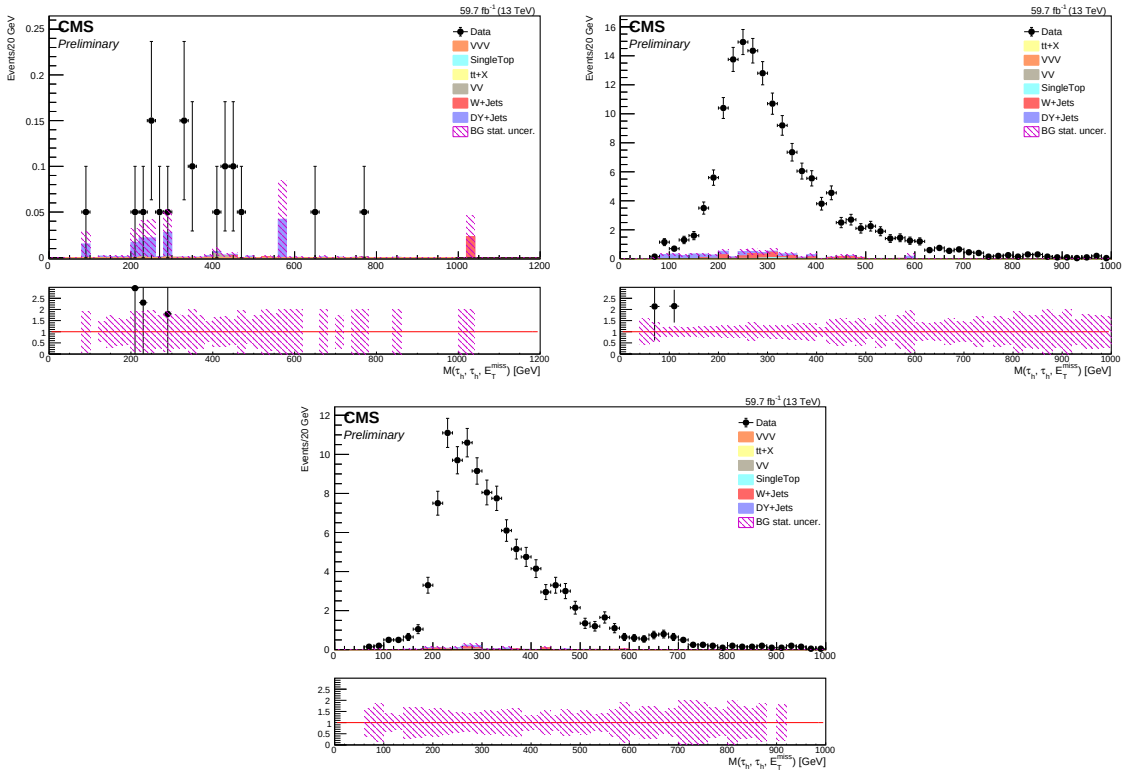


Figure 6.27: These plots represent the ditau mass for 2018 results for control regions B, C, and D. (Top Left) Region B. (Top Right) Region C. (Bottom) Region D.

background processes, the measurements should be consistent if the methods work.

6.8.1 QCD Estimation Closure Test

The following section contains closure tests confirming the accuracy of the QCD estimation strategy used in the $Z' \to \tau\tau$ signal region analysis. The closure tests use an orthogonal set of control regions to the signal region studies: instead of the $E_T^{miss} > 30$ GeV parameter used for the signal region and QCD regions B, C, and D above, these tests will use $E_T^{miss} < 30$ GeV in new control regions titled A', B', C', and D', where each prime region's only change from the original is this difference in the E_T^{miss} parameter, and the A' control region corresponds to the signal region, except with $E_T^{miss} < 30$ GeV.

The first closure test method uses all four QCD closure test control regions. As with the QCD estimation method above, this test uses Regions C', and D' to calculate the OS-to-LS transfer factor, which is then applied to Region B' to estimate the QCD contribution in Region A'. The resulting mass plots for Region A' in in each year are shown in Figure 6.30 below. The corresponding table is found in the table below. Since the data is consistent with the MC estimation, this background estimation method results in closure.

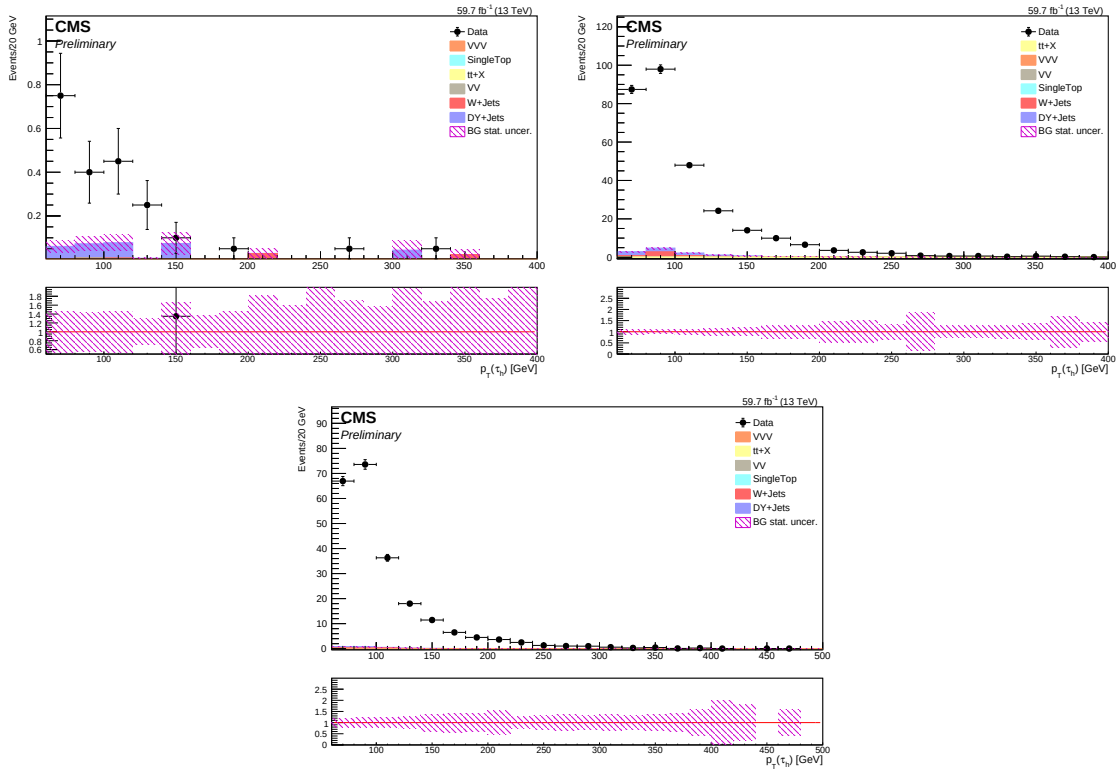


Figure 6.28: These plots represent the tau p_t for 2018 results for control regions B, C, and D. (Top Left) Region B. (Top Right) Region C. (Bottom) Region D.

Process	$E_T^{miss} < 30$ GeV; Pass VBF (2016-2018)
Data	96
DY + jets	58.00 ± 6.86
W + jets	3.82 ± 1.88
EWK Z+2j	4.35 ± 1.14
EWK W+2j	0.52 ± 0.37
Higgs	1.20 ± 0.09
Single Top	0.35 ± 0.2
$t\bar{t}$	2.39 ± 0.48
VV	1.22 ± 3.06
QCD	35.48 ± 7.86
Total BG	107.34 ± 7.86

Table 6.18

6.8.2 Drell-Yan Estimation Closure Test

A closure test for the Drell-Yan background prediction was performed using the same cuts as those of the signal region, except with an inverted E_T^{miss} cut ($E_T^{miss} < 30$ GeV) and with an additional cut on ditau mass ($m(\tau_h, \tau_h) < 150$ GeV) to maximize Drell-Yan purity. The QCD contribution in this region is extracted using a data-driven method similar to the QCD estimation method used in the SR. The DY+jets SFs are applied to

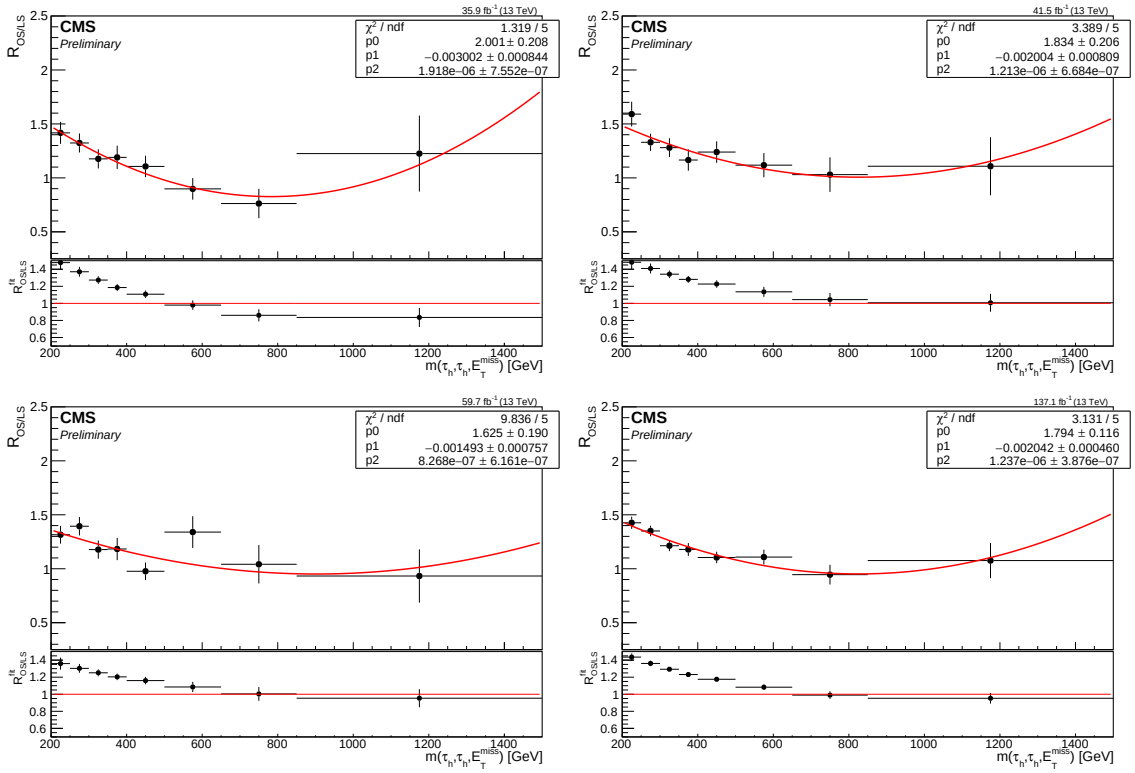


Figure 6.29: Transfer factor $R_{OS/LS}$ calculated using QCD control regions C and D for 2016 (Top Left), 2017 (Top Right), 2018 (Bottom Left), and all years combined (Bottom Right) MC and data samples. The ratio pad depicts the transfer factor, derived by fitting the $R_{OS/LS}$ vs. $m(\tau_h, \tau_h, E_T^{miss})$ points to the first degree polynomial and shows that the transfer factor is dependent on $m(\tau_h, \tau_h, E_T^{miss})$ across all three years.

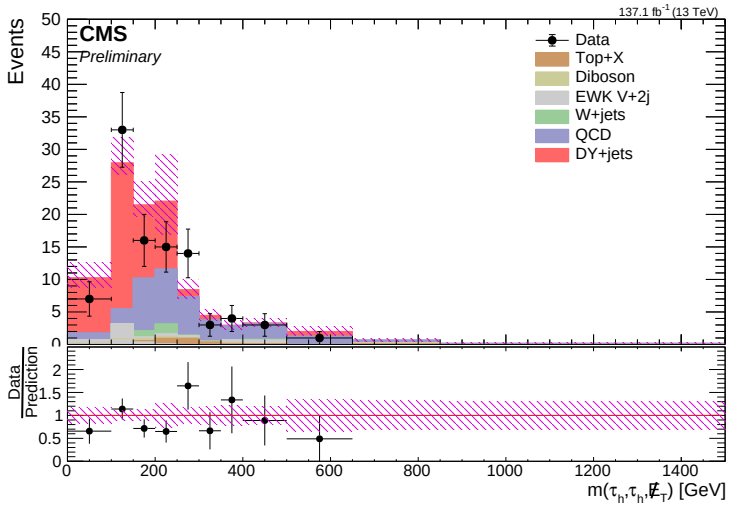


Figure 6.30: This plots represent the closure test for QCD estimation in the $Z' \to \tau\tau$ signal region in all three years (2016-2018).

correct the MC prediction for mismodelling of the selection efficiencies.

Figure 6.31 its corresponding table below show the closure test plot and corresponding output table including MC and data from all years (2016-2018). One can note that the resulting region has 82% purity of DY+Jets. The agreement between MC and data is observed to be good within statistical and systematics uncertainties.

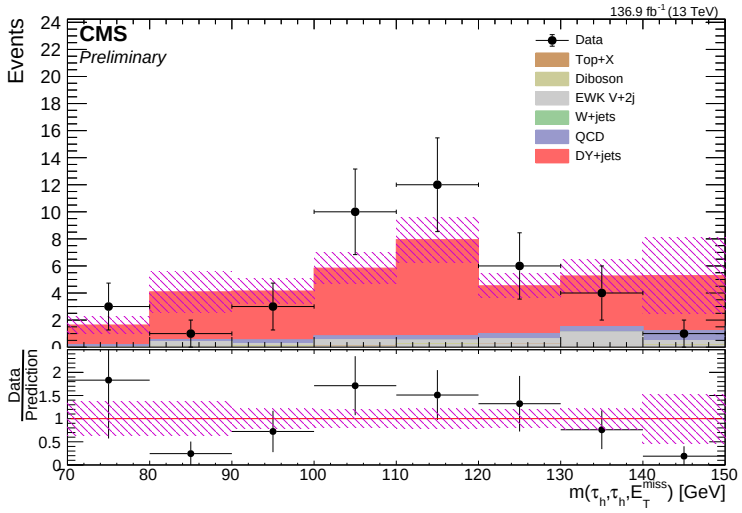


Figure 6.31: This plots the first closure tests for Drell-Yan estimation in the $Z' \to \tau\tau$ signal region using MC and data from all three years (2016-2018).

Process	$E_T^{miss} < 30$ GeV; $m_\tau m_\tau < 150$ GeV; Pass VBF (2016-2018)
Data	40
DY + jets	32.38 ± 3.29
W + jets	0.0 ± 0.0
EWK Z+2j	2.89 ± 0.97
EWK W+2j	0.0 ± 0.0
Higgs	0.77 ± 0.07
Single Top	0.0 ± 0.0
$t\bar{t}$	0.0 ± 0.0
VV	0.26 ± 0.07
QCD	3.2 ± 0.29
Total BG	39.46 ± 3.44

Table 6.19

6.9 Total Estimated Background in Signal Region vs. Data

Finally, we can examine the unblinded data and compare it to the estimated background. Figure 6.33 shows the total estimated background contribution for the $\tau_h\tau_h$ search region for 2016 (left), 2017 (middle) and 2018 (right). Table 6.20 shows the event yield of each predicted background in the signal region for three years.

The DY+jets and W+jets background predictions in the signal region are determined by applying the

derived SFs ($SF_{non-VBF}$ and SF_{VBF}) to the default Z/W + jets prediction from MC: $N^{SR,estimate} = N^{SR,MC} \times SF_{non-VBF} \times SF_{VBF}$. The QCD multijet background is derived in a data-driven way using the like-sign CRs scaled by the OS-to-LS transfer factor.

Since the data in the signal region is consistent with the SM background prediction, there is no significant sign of new physics.

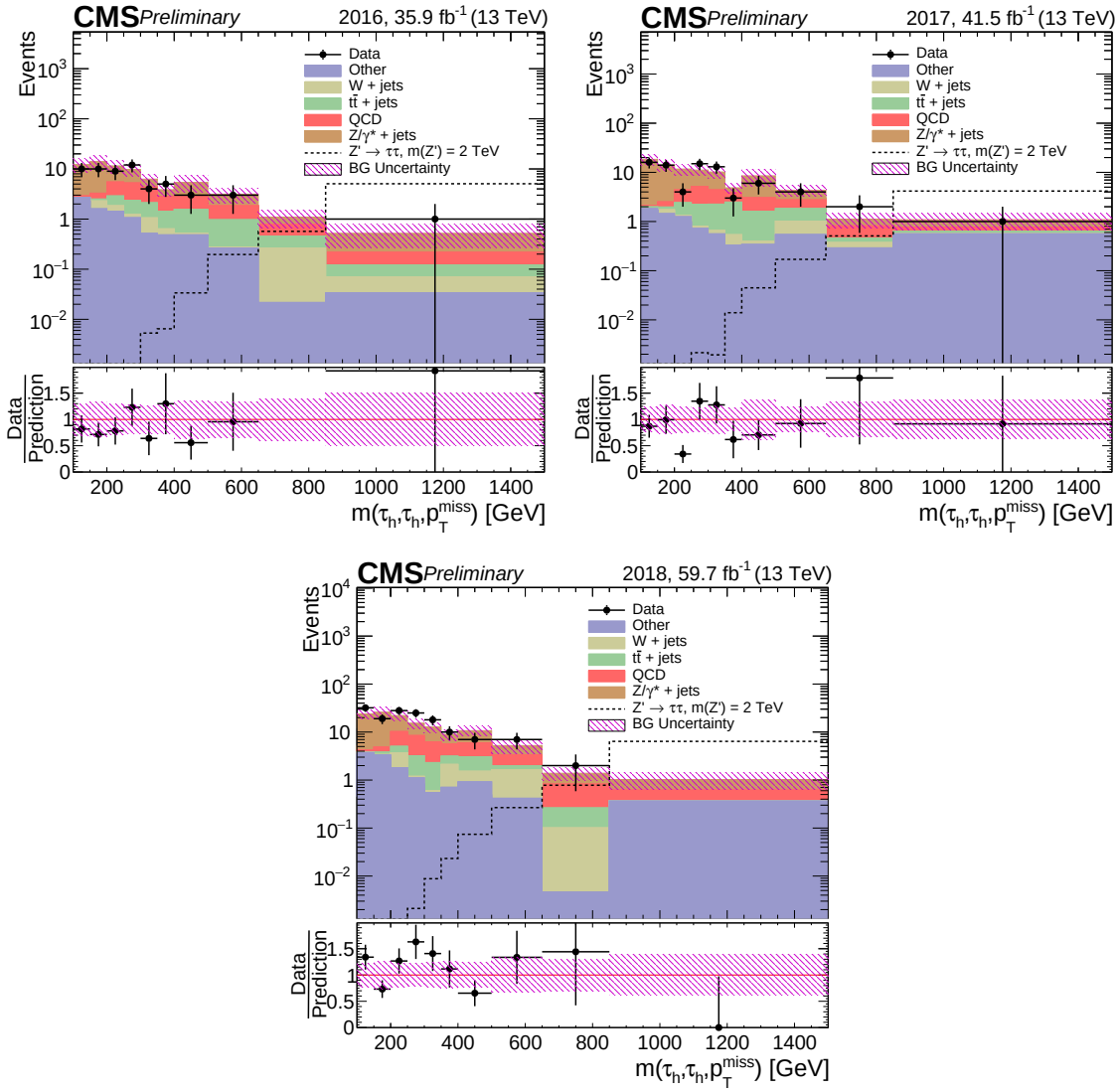


Figure 6.32: Total estimated and observed background contributing for $\tau_h\tau_h$ final state in 2016 (Top Left), 2017 (Top Right) and 2018 (Bottom).

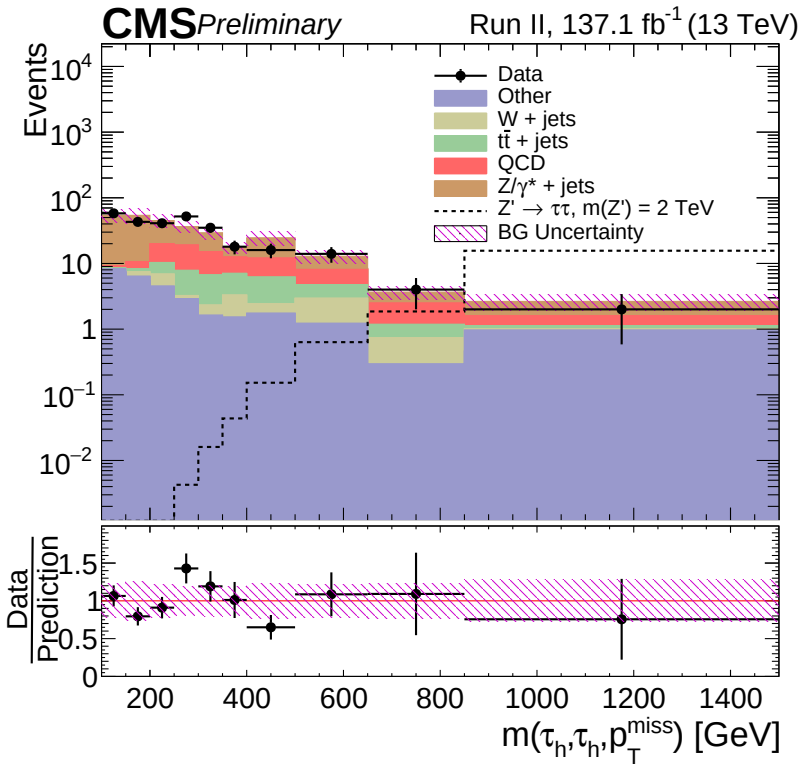


Figure 6.33: Total estimated and observed background contributing for $\tau_h\tau_h$ final state for all years (2016-2018).

Process	2016	2017	2018
Data	63	85	150
DY + jets	39.3 ± 4.6	57.9 ± 5.5	99.0 ± 8.1
W + jets	2.3 ± 0.7	1.2 ± 0.3	6.9 ± 2.8
EWK Z+2j	3.3 ± 1.7	3.4 ± 0.7	6.7 ± 1.3
EWK W+2j	0 ± 0	0.2 ± 0.2	0.9 ± 0.5
Higgs	5.3 ± 0.2	4.7 ± 0.3	9.5 ± 0.3
Single Top	0.6 ± 0.3	0.4 ± 0.2	1.4 ± 0.6
$t\bar{t}$	6.2 ± 0.6	8.6 ± 0.8	8.6 ± 0.9
VV	3.0 ± 0.5	3.0 ± 0.6	3.0 ± 0.4
QCD	12.9 ± 2.1	12.0 ± 2.0	23.5 ± 2.6
Total BG	72.8 ± 5.3	91.5 ± 6.0	159.5 ± 9.1

Table 6.20: Event yield of each predicted background in the signal region for $\tau_h\tau_h$ final state.

CHAPTER 7

Interpretation of Results

This chapter discusses the interpretations of the results of the analysis.

7.1 Upper Limit Calculations

To interpret the results of the analysis, it is useful to calculate an upper limit on the cross-section of the Z' signal. That is, in the case that a Z' signal is not found, we can still calculate the lowest possible VBF $Z' \to \tau\tau$ cross-section that can be excluded, given its absence in our results. We can calculate such upper limits via the "profile likelihood ratio test," which uses a likelihood function to represent the probability that a given set of fit parameters will match that of a given model [27].

The profile likelihood ratio test works by defining a null hypothesis (H_0) using only background processes. The null hypothesis is then compared to a model H_1 that includes both background and hypothesized signal processes. When setting limits, the roles are reversed, with the background plus signal set used for H_0 and the background-only set used for H_1. Ultimately, this test quantifies the agreement between the hypothesis H and the actual data, where the hypothesis is excluded if the calculated p-value is below a certain threshold.

To construct a likelihood function, we define a variable μ as the strength of a signal, where the estimated BG has a strength of 0 and the hypothesized signal has a strength of 1. Systematic uncertainties are accounted for via nuisance parameters (NP) θ.

We use $\vec{n} = (n_1, ..., n_N)$ to represent the histogram of the phase space in which we search for the signal, regarding some kinematic variable. We use the expression $u_i(\theta)$ to represent the expectation values in terms of the nuisance parameters. Further, the expectation values for n_i are represented as $E|n_i| = \mu s_i + b_i$, where s_i and b_i are the mean number of entries for a certain bin i for signal and background respectively.

Then, we can define a second set of parameters, $\vec{m} = (m_1, m_2, ..., m_M)$, which we can use to put constraints on the nuisance parameters. Here, the value of m_M is the number of entries in the M bin. The expectation values for m can then be expressed as $E|m_i| = u_i(\vec{\theta})$.

Thus, we can use the following expression to represent the likelihood function:

$$\mathscr{L}(\mu, \vec{\theta}) = \prod_{j=1}^{N} \frac{(\mu s_j + b_j)^{n_j}}{n_j!} e^{-(\mu s_j + b_j)} \prod_{k=1}^{M} \frac{u_k^{m_k}}{m_k!} e^{-u_k}, \tag{7.1}$$

which is the product of the Poisson probabilities for all bins.

The profile likelihood ratio is calculated using the following equation:

$$\gamma(\mu) = \frac{\mathscr{L}(\mu, \hat{\hat{\vec{\theta}}})}{\mathscr{L}(\hat{\mu}, \hat{\vec{\theta}})} \tag{7.2}$$

where $\hat{\hat{\vec{\theta}}}$ is the value of $\vec{\theta}$ for which $\mathscr{L}$ is maximized for a specific μ. The function $\mathscr{L}(\hat{\mu}, \hat{\vec{\theta}})$ in the denominator represents the maximized value of $\mathscr{L}$ where $\hat{\mu}$ and $\hat{\vec{\theta}}$ are used as ML estimators. Thus, $\mathscr{L}(\mu, \hat{\hat{\vec{\theta}}})$ is the "profile likelihood" as a function of μ for particular (defined) NPs, and $\mathscr{L}(\hat{\mu}, \hat{\vec{\theta}})$ is the likelihood function as a function of both μ and $\hat{\theta}$.

The null hypothesis H_0 was defined as having only BG processes with no signal, such that the strength parameter $\mu = 0$. We compared H_0 to the test statistic t_μ, which is defined as follows:

$$t_{mu} = -2\gamma(\mu). \tag{7.3}$$

An excess representing a 5σ p-value would qualify as significant enough to indicate a new particle discovery.

In the event that an excess is not found, we can still calculate confidence levels (CL_s) to determine upper limits on the signal's cross-sections:

$$CL_s = \frac{CL_{s+b}}{CL_b} = \frac{p_{s+b}}{1 - p_b} \tag{7.4}$$

where p_b is the p-value for the background-only data set and p_{b+s} is the p-value for the background plus signal data set. These p-values are calculated by taking the integral of the PDFs of the sets. Then, we can vary the strength of the signal μ until we get a confidence level of $CL_s \leq 0.5$, indicating that the associated cross-section has a confidence level of 95%.

As with the background estimation step, the signal region data remained blinded until the expected limits were determined. Thus, expected limits are set using MC samples created to predict the Z' signal. The data was then unblinded, and we compared the expected limits to the observed limits.

7.2 Systematics

The following systematics have been considered in the analysis:

Parton Distribution Functions (PDF): The NNPDF3.0 LO and NLO [6] parton distribution functions (PDFs) are used in the event generation for 2016. Version 3.1 is used for 2017 and 2018. The systematic effect due to imprecise knowledge of the parton distribution functions is determined by comparing these with the default PDF and variations within the family of parametrization [39]. We used the maximal deviation from the central value and obtained a value of 6.0%.

Initial State Radiation (ISR): We applied uncertainty to the ISR weights bin-by-bin for the boost of the Z/W boson. The relative uncertainties are summarized in Table 7.1.

Table 7.1: Event Weight Uncertainties by Z/W-Boost

Z/W-Boost Bin	Weight Uncertainty
Bin 1: 0-50 GeV	2.23%
Bin 2: 50-100 GeV	1.21%
Bin 3: 100-150 GeV	1.09%
Bin 4: 150-200 GeV	1.18%
Bin 5: 200-300 GeV	1.29%
Bin 6: 300-400 GeV	1.95%
Bin 7: 400-600 GeV	3.23%
Bin 8: 600-1000 GeV	14.45%

The first step in investigating the uncertainty associated with the ISR weights was to determine the statistical uncertainty bin-by-bin for the boost of the Z-boson. The uncertainty on the weight (combination of Poisson error in data and the statistical uncertainty on MC) varies with the p_T of the Z-boson and ranges from 1-3% at low Z-p_T to 11% at high Z-p_T. These values are only the uncertainty on the weights and do not necessarily reflect how varying the weights by these uncertainties affects the signal and BG yields and shapes in a given region (CR or SR). To determine the effect of these weight uncertainties on the BG/signal yields and distributions, we varied the weights up by 1σ and down by 1σ, where σ is the actual uncertainty on the weight. This method was used to examine the possible systematic effect on the mass distribution in the signal region for Z+Jets and W+Jets in the $e\mu$-channel. The plots in Figure 7.1 demonstrate this method, where red is the mass distribution obtained with the central value of the weights, green is the variation of the weights down by 1σ, and blue is for the variation of the weights up by 1σ. From these plots, we derive a 5% systematic uncertainty, independent of the reconstructed $m(\tau_h, \tau_h, E_T^{miss})$ mass, to be applied to Z+Jets and W+Jets to account for the uncertainty associated with the ISR reweighting.

Luminosity: We considered a 1.2 - 2.5% uncertainty on the measured luminosity [49, 22, 24].

Pileup: The uncertainty due to pileup is obtained by shifting the minimum bias cross-section up and down by 4.6% [19].

Trigger, Reconstruction, and Selection: An overall uncertainty is applied for the trigger uncertainties. We consider 3.0% uncertainty due to the trigger efficiency, which is a result of a fit to the trigger efficiency turn-on curve after VBF selections. The 3% is the uncertainty on the fit at the plateau, which is driven by the statistics in data as shown in figure 7.2. The uncertainty for reconstruction of muons and electrons is up to about 2% each depending on η or p_T.

***b*-Tagging Efficiency:** We consider the uncertainty on the mis-tag rate as measured by the b-tagging POG

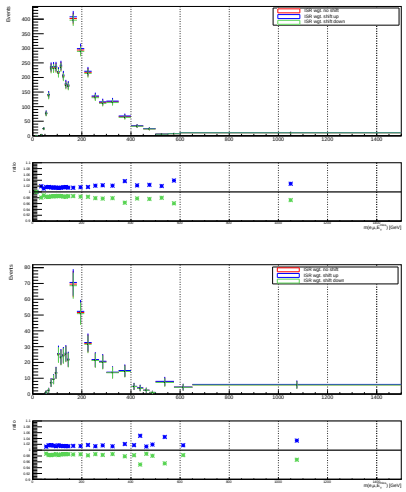


Figure 7.1: $m(e,\mu,MET)$ for Drell-Yan (top) and for W+Jets (bottom) in the Signal Region

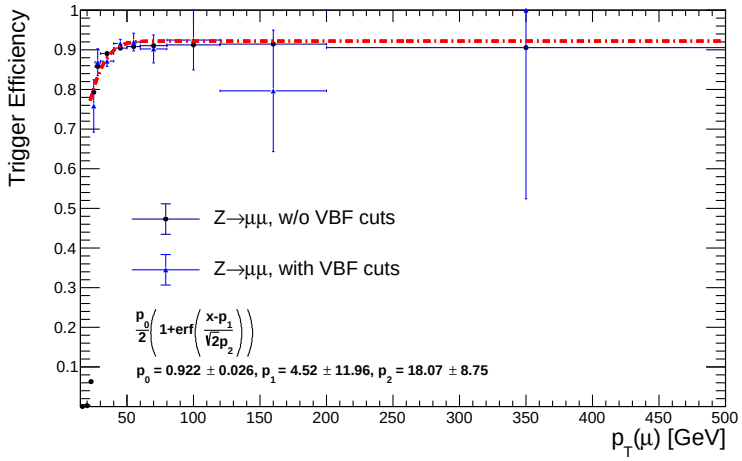


Figure 7.2: Fit to the trigger efficiency turn-on curve with $Z \rightarrow \mu\mu$ + VBF cuts in data. The fit shows about 3% uncertainty on the fit at the plateau

[20] (about 10%). For the case of our signal, the systematic uncertainty on the requirement of 0 jets mis-tagged as b-jets is determined by propagating the 10% uncertainty on the mis-tag rate through the following equation (which represents the signal efficiency for requiring 0 jets mis-tagged as b-Jets):

$$\varepsilon^{NBtag<1} = 1 - \sum_{n=1} P(n) \cdot \sum_{m=1}^{n} C(n,m) \cdot f^m \cdot (1-f)^{n-m} \tag{7.5}$$

where $P(n)$ is the probability to obtain n additional jets (non-tau and non-lepton) in the event, $C(n,m)$ the combinatorial of *n choose m*, and f the mis-tag rate. The probability to obtain at least one additional jet in the event is 10%. Therefore, based on the above equation, the mis-tag rate and uncertainty, and the probability to

obtain at least one additional jet we calculate a systematic effect of 1% on our signal due to the mis-tag rate. The b-tagging/mis-tagging systematics are considered 100% correlated across MC based backgrounds with similar composition (e.g. W + jets and DY + jets where there are typically no real b-jets), but completely uncorrelated to backgrounds that have different composition (e.g. $t\bar{t}$ vs. DY + jets). We note the b-tagging uncertainties are evaluated on a per-event basis, following the POG recommendation and as referenced in the object reconstruction section.

To determine the effect of the b-jet SF on the BG/signal yields and distributions, we vary the SFs up by 1σ and down by 1σ, where σ is the actual uncertainty on the SF, to examine the possible systematic effect on the mass distribution in the 1b, passing VBF region for $t\bar{t}$ inclusive in the e/μ-channel. This is the plot in Figure 7.3, where red is the mass distribution obtained with the central value of the weights, green is the variation of the SFs down by 1σ, and blue is for the variation of the SFs up by 1σ. From this plot in Figure 7.3, we derive a 2% uncertainty to be applied conservatively to backgrounds with real b-jets to account for the uncertainty associated with using the b-jet SFs.

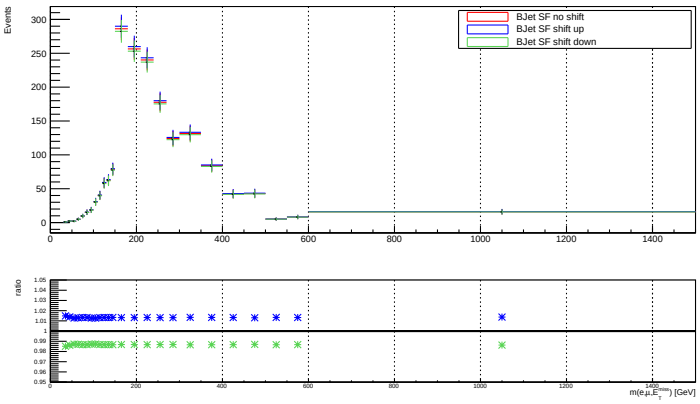


Figure 7.3: $m(e,\mu,MET)$ for $t\bar{t}$ Inclusive in the Signal Region

Electron Energy Scale: We consider the effect on the signal acceptance efficiency of 1% (2.5%) shift on the electron energy scale in the barrel (endcap) region. The systematic effect is $< 1\%$. These values follow from the CMS EGamma POG recommendations [25].

Muon Momentum Resolution: We consider 4% value (independent of p_T and η) as the momentum resolution uncertainty to cover the full momentum-η phase relevant for our analysis. According to the Muon POG recommendations [11], the muon momentum resolution uncertainty depends on p_T and η. For a muon with momentum of 100 (500) GeV and $|\eta| < 1.2$, the recommended resolution uncertainty is 1.5% (3.6%). For muons with $1.2 < |\eta| < 2.1$, the recommended resolution uncertainty was 2% and 3.9% for the same 100 GeV and 500 GeV momentum values. In our analysis, we have chosen to be conservative and use a

flat 4% value (independent of p_T and η). The difference between our conservative approach and the exact recommendation is insignificant in the limit calculations (about 0.1%) since the processes are dominated by other effects.

Muon Momentum Scale: We considered the effect on the signal acceptance efficiency of a 5% momentum scale uncertainty on the muon momentum. The systematic effect is $< 1\%$. These values follow from the Muon POG recommendations [11] [16] [15]. In recommendations from Muon POG, the muon scale uncertainties depend on whether we have a significant fraction of muons with $p_T > 200$ GeV. If so, the scale uncertainties are based on the GE method [46]. It is important to note that we are probing Z' masses between a few hundred GeV and a few TeV. For the lower mass Z' hypotheses, the p_T values of the muons are about 35 GeV on average (1/6 of the Z' mass of 200 GeV $\rightarrow$ each tau takes half the Z' mass, then it decays to a muon and two neutrinos, such that the muon takes 1/3 of the 1/2). For the higher mass Z' hypothesis of 2 TeV, the p_T of the muons are about 330 GeV on average. Table 7.2 below shows the fraction of muons with $p_T > 200$ GeV for a few example Z' mass scenarios.

Z' **Mass (GeV)**	$p_T > 200$	$p_T > 350$	$p_T > 500$	$p_T > 750$
500	8%	1%	0.2%	0.1%
750	22%	2%	0.1%	<0.05%
1500	56%	27%	9%	0.2%
2500	75%	54%	35%	14%

Table 7.2: Fraction of muons as a function of μ p_T.

One can see that some Z' mass scenarios would require the GE-based scale uncertainties, and others wouldn't. Therefore, we assigned a muon momentum scale uncertainty which would cover this large range of p_T, which accounts for that two types of uncertainty (Rochester and GE based uncertainties) in our p_T range. Based on the Muon POG recommendations [11], for $|\eta| < 1.2$, the GE based scale uncertainties range from $< 1\%$ at p_T = 500 GeV to 2% at TeV scale p_T values. For $1.2 < |\eta| < 1.6$, the GE based scale uncertainties range from 1.5% at p_T = 500 GeV to 3% at p_T = 1.5 TeV. For higher η muons, the scale uncertainty can be 3% for similar ranges of p_T. Our initial thought was to use a flat 3% scale uncertainty (independent of p_T and η), since it would conservatively cover the relevant phase space of the GE-based recommendations for high p_T. However, in the end we further increased our scale systematic uncertainty to 5% due to calibration studies performed in a W+jets control region. The standard candle W+jets control region is obtained using the muon primary datasets, requiring events to fire the isolated muon triggers used in this analysis, further selecting exactly one well-identified muon ($p_T > 35$ GeV, $|\eta| < 2.1$, passing Tight ID and nominal isolation), $E_T^{miss} > 30$ GeV, and with a muon-E_T^{miss} system having transverse mass compatible with the W-mass Jacobian peak (50-110 GeV). The W+jets control region stacked plot is shown below in

Figure 7.4.

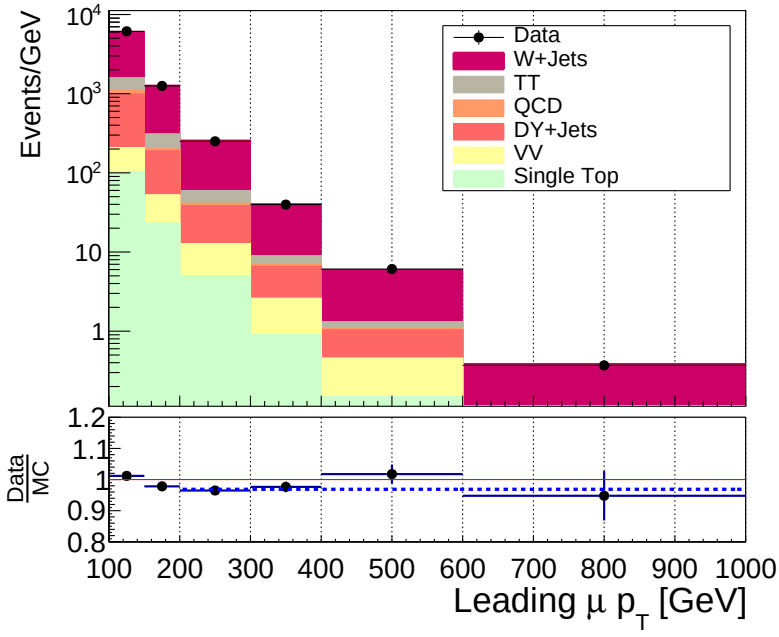


Figure 7.4: Leading μ p_T in a W+jets control region.

In this control region, we found that even after using best known cross-section (with k-factors), best measured lumi, scale-factors/corrections for mismodeling of MC efficiencies (including trigger, ID, etc.), we still observed a small disagreement between data and corrected backgrounds (about 4%). This small disagreement is highlighted in the ratio pad by the fit using the 0th degree polynomial 0.96 ± 0.006, which yielded a X^2/ndf value of 3.05/3). We studied the possibility that this residual discrepancy could be conservatively covered by a 5% muon momentum scale uncertainty and found this to be the case. Although it is possible that this small residual disagreement can be due to other effects, we point out that increasing our scale uncertainty to 5% results in a small change in the signal yields for all masses (as shown in Table 7.3) and an insignificant change to our limit calculations.

Z' **Mass (GeV)**	**S(MES up)/S(nominal)**	**S(MES down)/S(nominal)**
2500	1.01	0.991
2000	1.012	0.987
1500	1.021	0.975
750	1.017	0.978

Table 7.3: Ratio of event yields with 5% shift in muon momentum scale.

Jet Energy Scale: We considered the effect of a 2-5% jet energy scale uncertainty on the signal acceptance (depending on the η and p_T of the considered jet as prescribed by the *JetMET POG*) [36]. For example, for a 5% scale uncertainty, the jet 4-momentum is scaled by a factor of $k = 1.05$ or 0.95 ($p_{smeared} = k \cdot p_{default}$)

and variables are recalculated using $p_{smeared}$. In Figure 7.5, the blue ratio is for the up-shifted distribution to the central. The green ratio is for the down-shifted distribution to the central. Binning used here is optimized in the SR. Systematics for the varying of the JES (5%) amount to 10 – 15% on average.

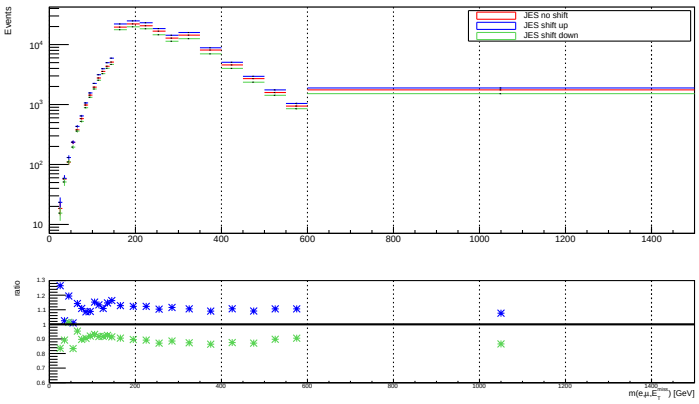


Figure 7.5: $m(e,\mu,MET)$ for $t\bar{t}$ Inclusive in the Signal Region

Jet Energy Resolution: We used the data-driven methods to estimate the backgrounds. There could be some effect on the background estimation due to the mismodeling of the non-targeted background, i.e., if the CR does not have 100% purity, then the calculation of data minus non-targeted MC background can be off a bit. This effect has been taken into account as part of our systematic uncertainties on the background predictions. In cases where MC based backgrounds must be subtracted off, the uncertainties in the MC backgrounds due to the listed systematic uncertainties (i.e. JER) are propagated throughout the subtraction and used to assign a systematic uncertainty on the background prediction. The effect of JER uncertainty in the MC backgrounds is estimated by scaling the resolution up-down with the uncertainty in the scale factor provided by the JetMET group [34]. In Figure 7.6, the blue ratio represents the up-shifted distribution to the central, and the green ratio represents the down-shifted distribution to the central. Systematics for the varying of the JER amount to 10-15% on average. We implement shape based uncertainties in the fit for $m(l_1,l_2,E_T^{miss})$ based on JES.

MET: The uncertainty on MET for our signal process is driven by the jet energy scale (non-tau jets) (JES), light lepton energy/momentum scale (LES), and unclustered energy (UCE). The systematic effect from MET due to JES and LES is included in the JES and LES systematic uncertainties described above. We find that a 10% uncertainty on the unclustered energy results in at most a 0.5% fluctuation on the signal acceptance.

Jet $\rightarrow \tau_h$ fake background in $l\tau_h$ final state: To estimate this background in $e\tau_h$ and $\mu\tau_h$ final states, fake ratio from DY+Jets control region are estimated. W+Jets control region is chosen, which has different quark-gluon jet concentration, to estimate the second set of fake factors. The effect of jet flavour on Jet $\rightarrow \tau_h$ fake background is estimated on both shape and normalization in both $e\tau_h$ and $\mu\tau_h$ final states for 2016, 2017 and

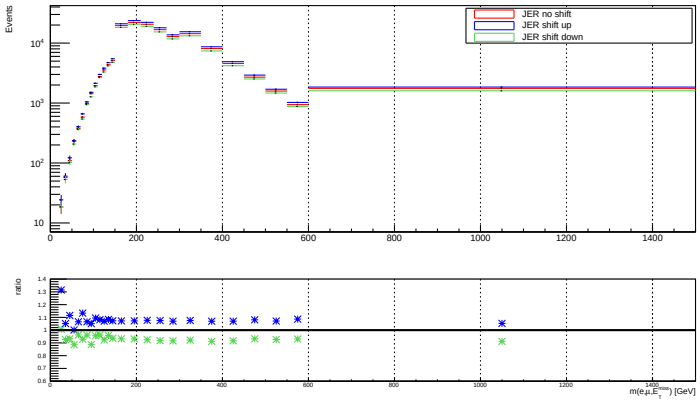


Figure 7.6: $m(e,\mu,MET)$ for $t\bar{t}$ Inclusive in the Signal Region

2018 and is considered as systematic uncertainty. The uncertainty is considered uncorrelated across years and channels.

For the $l\tau_h$ channels, the difference between the jet$\rightarrow\tau_h$ fake ratios in W + jets and Z + jets control regions depends on p_T (and thus reconstructed mass) and can be as large as about 20% in some mass bins as shown in Figure 7.7. These bin-by-bin uncertainties are also treated as shape systematics in the limit/significance calculations.

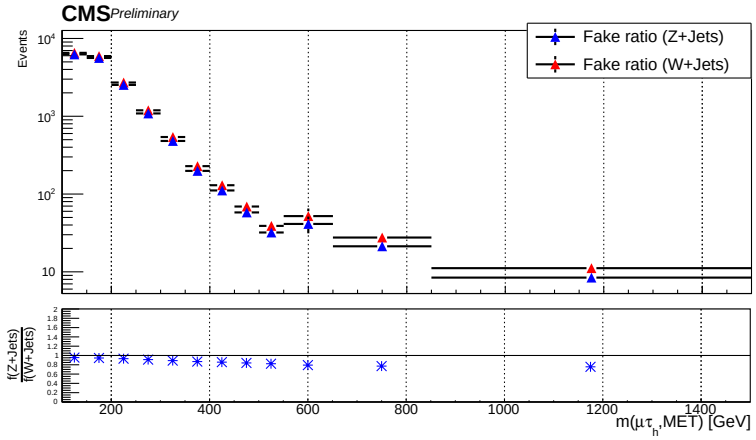


Figure 7.7: Jet $\rightarrow\tau_h$ fake background using fake ratio from W+Jets control region and DY+Jets control region.

7.3 Uncertainties in the $\tau_h\tau_h$ final state:

For the $\tau_h\tau_h$ channel, the uncertainties on the non-QCD background estimates come from the uncertainties on the SFs ($SF_{non-VBF}$ and SF_{VBF}). These can be as large as 17% and as small as $< 1\%$ depending on the background. The uncertainties on the data-driven QCD depend on the statistics in the like-sign sidebands.

These uncertainties can be large (close to 100%) in high mass bins where there are minimal data stats in like-sign.

Source channel:	QCD $\mu\tau_h$,eτ_h,$\tau_h\tau_h$,eμ	W $\mu\tau_h$,eτ_h,$\tau_h\tau_h$,eμ	DY $\mu\tau_h$,eτ_h,$\tau_h\tau_h$,eμ	$t\bar{t}$ $\mu\tau_h$,eτ_h,$\tau_h\tau_h$,eμ	VV $\mu\tau_h$,eτ_h,$\tau_h\tau_h$,eμ	Signal $\mu\tau_h$,eτ_h,$\tau_h\tau_h$,eμ
Lumi	L,L,L,L	L,L,L,L	L,L,L,L	L,L,L,L	L,L,L,L	L,L,L,L
μ ID	–,–,–,–	2,< 1,< 1,2	2,< 1,< 1,2	2,< 1,< 1,2	2,< 1,< 1,2	2,< 1,< 1,2
e ID	–,–,–,–	< 1,2,< 1,2	< 1,2,< 1,2	< 1,2,< 1,2	< 1,2,< 1,2	2,< 1,< 1,2
τ_h ID	–,–,–,–	< 1,< 1,5,< 1	5,5,9,< 1	2.5,2.5,7,< 1	4.5,4.5,9,< 1	5,5,9,< 1
Trigger	–,–,–,–	3,3,3,3	3,3,3,3	3,3,3,3	3,3,3,3	3,3,3,3
b ID	–,–,–,–	1,1,1,1	1,1,1,1	2,2,2,2	1,1,1,1	1,1,1,1
JES	–,–,–,–	s,s,s,s	s,s,s,s	s,s,s,s	s,s,s,s	s,s,s,s
TES	–,–,–,–	s,s,s,s	s,s,s,s	s,s,s,s	s,s,s,s	s,s,s,s
MMS	–,–,–,–	5,< 1,< 1,5	5,< 1,< 1,5	5,< 1,< 1,5	5,< 1,< 1,5	5,< 1,< 1,5
EES	–,–,–,–	< 1,1,< 1,1	< 1,1,< 1,1	< 1,1,< 1,1	< 1,1,< 1,1	< 1,1,< 1,1
MET uncl.	–,–,–,–	< 1,< 1,< 1,< 1	< 1,< 1,< 1,< 1	< 1,< 1,< 1,< 1	< 1,< 1,< 1,< 1	< 1,< 1,< 1,<
pdf	–,–,–,–	4.8,4.8,4.8,4.8	4.2,4.2,4.2,4.2	4.2,4.2,4.2,4.2	3.5,3.5,3.5,3.5	6,6,6,6
scale	–,–,–,–	1,1,1,1	1,1,1,1	3.5,3.5,3.5,3.5	1,1,1,1	2,2,2,2
bin-by-bin stat.	s,s,s,s	s,s,s,s	s,s,s,s	s,s,s,s	s,s,s,s	s,s,s,s
Closure+Norm.	< 20,< 20,< 17,–	< 20,< 20,1,1	3,6,< 21,1	< 20,< 20,–,< 20	–,–,–,–	–,–,–,–

7.4 Jet enegy resolution studies

As shown in Figures 7.8 and 7.9 for the leading jet p_T and η in the Z($\rightarrow \mu^+\mu^-$)+jets control samples for 2016 and 2017 (Table 7.4), a large discrepancy between data and simulation is observed at low jet p_T and the tracker-HF transition region (2.5< $|\eta|$ <3.0). There is also an important disagreement in the p_T^{miss} distribution in this control sample for 2016 and in a slightly less pronounced for 2017 (Figure 7.10). The p_T^{miss} in these distributions corresponds to the JER type-I p_T^{miss}. The method used to apply JER corrections for the distributions in these control samples consists of a combination of the scaling and stochastic methods, currently recommended by the CMS JME group: if there is a gen-level jet matching a reconstructed jet, then the scaling method is utilized. Otherwise, the stochastic method is used to derive the correction factor for JER (c_{JER}). We refer to this method as *nominal JER smearing*.

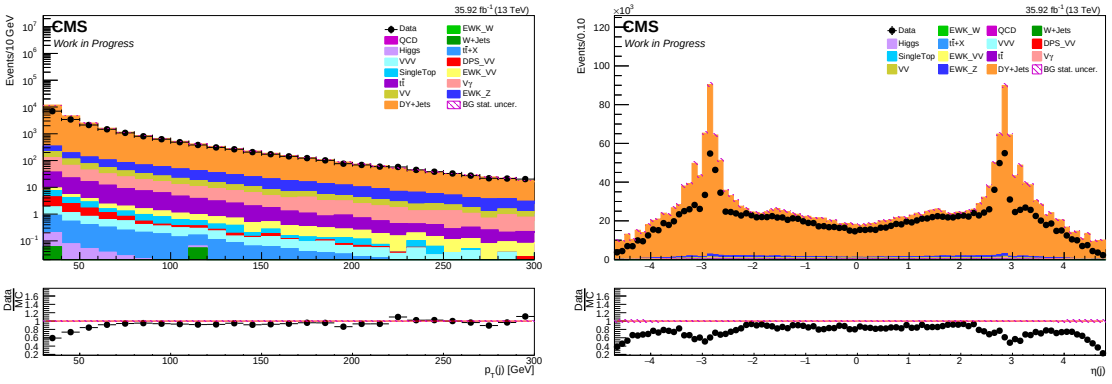


Figure 7.8: Distributions for the jet p_T and η in the Z+jets control region in 2016 after applying VBF selections.

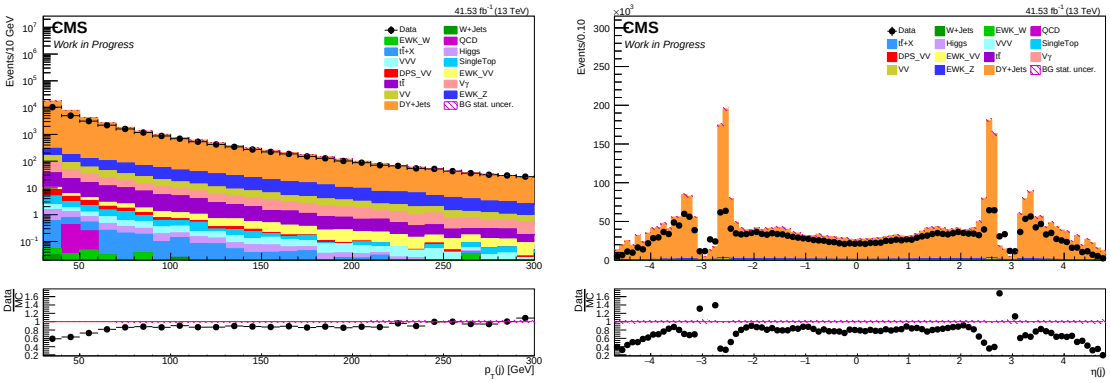


Figure 7.9: Distributions for the jet p_T and η in the Z+jets control region in 2017 after applying VBF selections.

	Object	Selection cuts
Central	Trigger	HLT_IsoMu24 or HLT_IsoMu27
	Muon selection	$N(\mu) = 2$, $p_T(\mu) > 30$ GeV, $\lvert\eta(\mu)\rvert < 2.1$, tight ID, rel. iso. < 0.15, $q(\mu_1) \times q(\mu_2) \times 0$
		$m(\mu,\mu) \in$ [80,100] GeV
	E_T^{miss} selection	> 0 GeV
	Electron veto	$N(e) = 0$, $p_T(e) > 10$ GeV, $\lvert\eta\rvert < 2.5$, loose ID (cut-based)
	Tau veto	$N(\tau) = 0$, $p_T(\tau) > 20$ GeV, $\lvert\eta\rvert < 2.3$, tight anti-e/μ discr., prong: 1or3hps
	b-jet veto	$N(b) = 0$, $p_T(b) > 20$ GeV, $\lvert\eta\rvert < 2.4$, Medium DeepCSV WP
VBF	Jet selection	$N(j) \geq 2$, $p_T(j) >$ **30 GeV**, $\lvert\eta(j)\rvert < 5.0$,
		loose (2016)/ tight (2017) PF ID, $\Delta R(e/\mu, j) > 0.4$
	Dijet selections	$\Delta\eta(j_1, j_2) >$ **3.8**, $\eta(j_1) \times \eta(j_2) < 0$, $m(jj) >$ **500 GeV**

Table 7.4: Event selections for Z($\rightarrow \mu^+\mu^-$)+jets control sample with VBF selections.

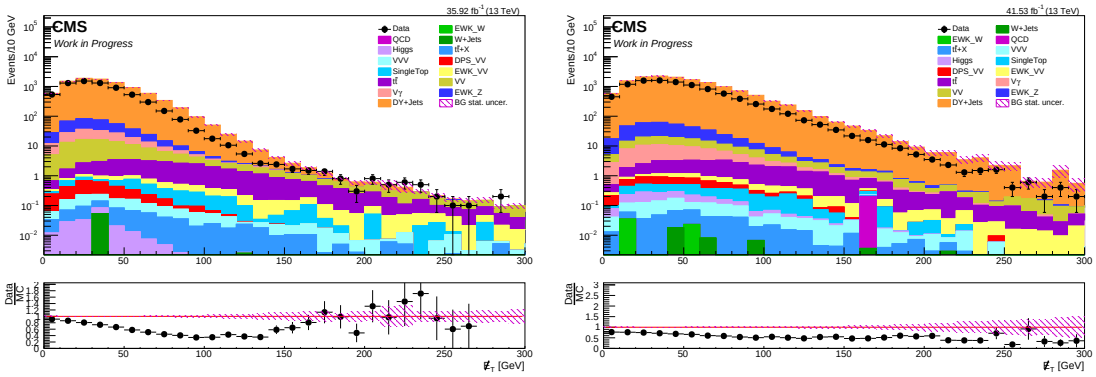


Figure 7.10: Distributions for missing transverse energy p_T^{miss} in the Z+jets control region after VBF selections for (left) 2016 and (right) 2017 data.

In this control region we do not expect to measure any real p_T^{miss}. Therefore, given the definition of $\vec{p_T^{miss}}$, any p_T^{miss} will come from the mis-measurement of the ISR jet activity, giving a starting point to investigate the origin of this discrepancy. For now we will focus in the 2016 Z($\rightarrow \mu^+\mu^-$)+jets control sample. The first step taken for this purpose was to not include the JER corrections and study the data-to-MC agreement in the

p_T^{miss} distribution for this case. As shown in Figure 7.11, the agreement is practically consistent with unity in the p_T^{miss} distribution and the discrepancy at low jet p_T and forward jets is reduced.

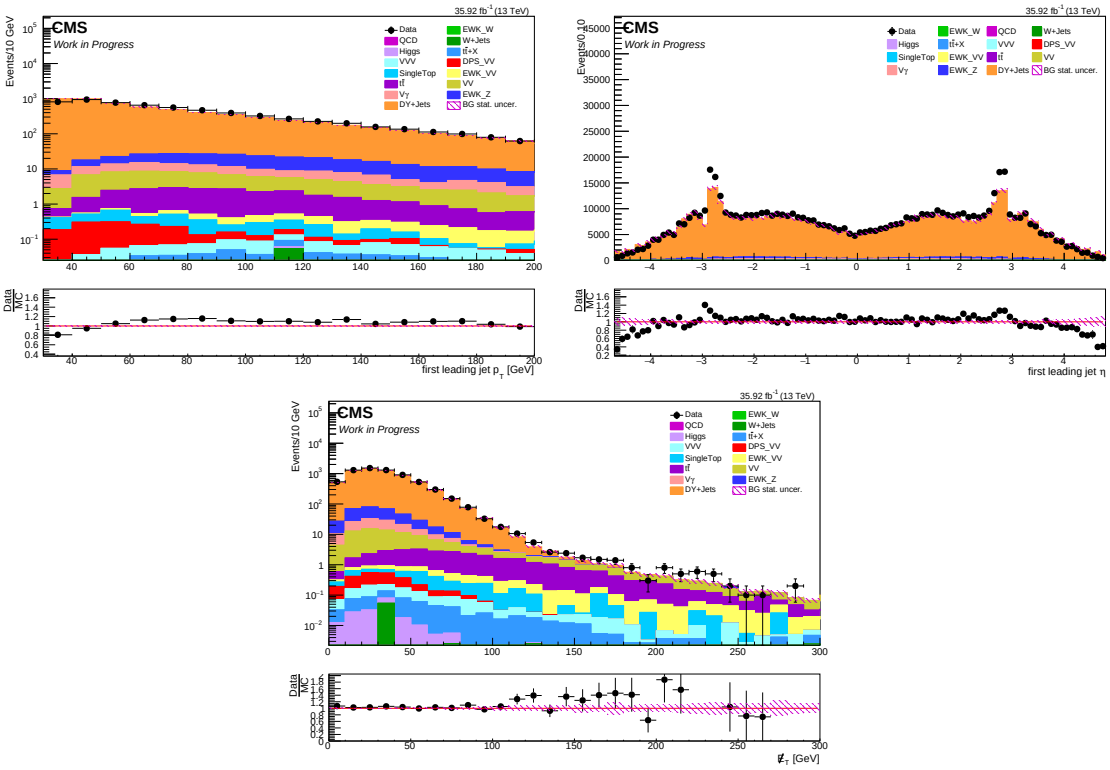


Figure 7.11: Distributions for the leading jet p_T (Top Left), η (Top Right), and p_T^{miss} (Bottom) in the 2016 Z($\rightarrow \mu^+\mu^-$)+jets control region without JER smearing corrections after applying VBF selections.

A quick study to confirm this hypothesis, we decided to obtain a similar Z($\rightarrow \mu^+\mu^-$)+jets control sample, but in this case only considering events with jets in the 2.5$< |\eta(j)| <$3.0 region, where the overcorrection from JER in simulation was the largest. Here, we compared the fraction of genuine and pileup jets in simulation, where a *genuine jet* is defined as a reconstructed jet **matched** to a gen-level jet and a *pileup jet* is a reconstructed jet **not matched** to a gen-level jet. In addition, properties such as the p_T and η of both kinds of jets are compared, as well as the c_{JER} (JER correction factor), s_{JER} (data-to-MC core resolution scale factor) and σ_{JER} (relative jet p_T resolution as measured in simulation) variables (Figures 7.12).

It is observed that, the number of unmatched jets is larger than that of matched jets in general, indicating that in this region we predominantly have pileup jets. The data-to-MC core resolution scale factor s_{JER}, which has only $\eta(j)$ dependence in 2016, shows that more unmatched jets will receive larger s_{JER} corrections, which explains the broad range of nominal c_{JER} observed for this type of jets. Finally, in Figure 7.13, one can see that most of the jets in the high η region correspond to low p_T jets.

Therefore, we confirmed that the disagreement observed in these three distributions (jet p_T, η and p_T^{miss})

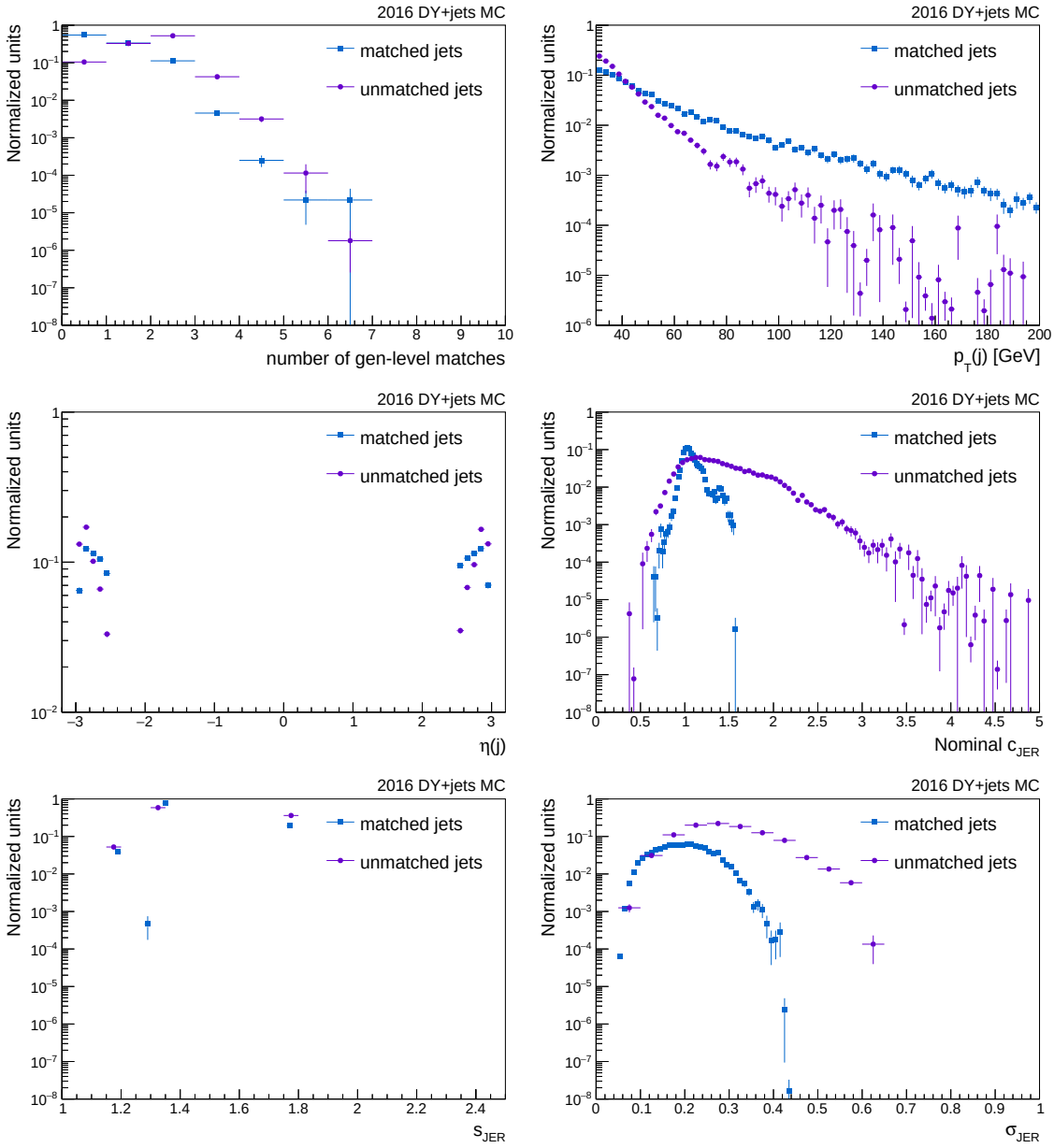


Figure 7.12: Distributions comparing jet multiplicity, jet p_T, η, nominal c_{JER}, s_{JER} and σ_{JER} for generator level matched (genuine) and unmatched (pileup) jets in the $Z(\rightarrow \mu^+\mu^-)$+jets control sample, only considering jets with $2.5 < |\eta| < 3.0$. A jet p_T vs. η comparison was also performed.

is correlated and is related to the JER corrections. Most jets in the $2.5 < |\eta| < 3.0$ are low p_T jets and the fraction of pileup jets is larger than that of genuine jets and pileup jets are receiving larger corrections for JER than genuine jets. The effect of this effect is very small in the analysis as the major backgrounds are estimated from data in each final state. The data-MC SF for small backgrounds estimation are taken into account and any effects due to these factors are incorporated in them. Dedicated closure tests are performed

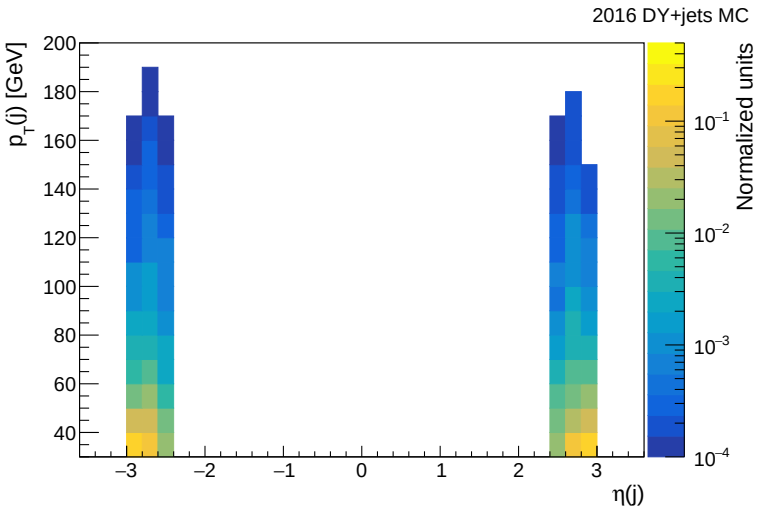


Figure 7.13: Comparison of jet p_T vs. η in the Z($\rightarrow \mu^+\mu^-$)+jets control sample, only considering jets with $2.5 < |\eta| < 3.0$.

for the validity of the SFs used and systematic uncertainties are considered accordingly.

CHAPTER 8

Results

We used the reconstructed Z' mass to set limits on the production cross-section of Z' decaying to $\tau\tau$ pairs and WW pairs. Separate limits are determined for each κ_V (coupling of Z' to vector bosons) for universal and non-universal coupling to light (g_l) and heavy (g_h) fermions.

We calculated expected limits at 95% CL on the signal cross-section using the method described in the previous chapter. The multibin method is used, using binned histograms as input. The nuisance parameters for the uncertainties are modelled with log normal distribution in the likelihood function.

In the results presented below, it is important to note that the branching fractions are model-dependent parameters and depend on every single coupling between Z' and all the SM particles. For example, on one extreme example, the $Z' \to \tau\tau$ branching fraction can be close to 100% if the Z' coupling to the SM weak bosons (κ_V) is relatively small with respect to the tau coupling, and the $Z' \to \tau\tau$ coupling is large and the only other relevant coupling. Conversely, if the Z' is allowed to couple strongly to many other SM particles, such as the top quark and the b-quark, then the branching fractions can vary significantly. All of these scenarios are probable and one can change the set of "good" models depending on the constraints one wants to impose (yukawa hierarchy, explanation of the top mass, Higgs hierarchy problem, matter-antimatter asymmetry, etc.). However, experimentally, the model's assumptions only affect the $\sigma \times Br$ theory curve in the expected limit plots. For this reason, we have plots of the expected exclusions versus Z' mass, coupling, and branching fraction. In this way, a theorist can choose their favorite model with a given branching fraction and extract the corresponding bound. Since the choice of Br is arbitrary, in some of the 1D limit plots, we use a benchmark scenario that visually gives the best representation of the results.

Additionally, the full VBF Z' analysis involved several possible Z' decay modes, including Z' decaying to $\mu\tau_h$, $e\mu$, and $e\tau_h$. The results from those studies are shown here to compare to the $\tau_h\tau_h$ channel.

8.1 Limit Calculation

A shape-based analysis is performed using the $m_{l,l,E_T^{miss}}$ distributions as the fit discriminant to determine the likelihood of observing signal in the presence of the predicted background rate. The variable binning for the fit variable ($m_{l,l,E_T^{miss}}$) is chosen based on best signal significance and best limits. Therefore, the $\tau_h\tau_h$ channel has 10 bin, the $l\tau_h$ channels have 12 bins, and the $e\mu$ channel has 18 bins. Thus, there are 52 cards in total for the four channel per year, giving 156 total cards for three years for all the four channels.Correlations of systematic effects between signal, data and backgrounds are taken into consideration for each individual

channel, across channels.

Figure 8.1 shows the limits on the signal cross-section as a function of Z' mass for VBF Z' decaying to $\tau\tau$ and then further to $\tau_h\tau_h$ (Left) and $\mu\tau_h$ (Right) for κ_V (coupling of Z' to vector bosons) of 0.1 (Top), 0.5 (Middle) and 1.0 (Bottom), for enhanced coupling to heavy (g_h) fermions scenario. Similarly, Figure 8.2 shows the expected and observed limits for VBF Z' decaying to $\tau\tau$ and then further to $e\tau_h$ (left) and $e\mu$ (right) for κ_V (coupling of Z' to vector bosons) of 0.1 (top), 0.5 (middle) and 1.0 (bottom) for enhanced coupling to heavy (g_h) fermions scenerio. As can be seen from the plots, $\tau_h\tau_h$ is the most sensitive channel for VBF Z' decaying to $\tau\tau$.

8.2 Limits combining all channels

Figure 8.4 and 8.6 shows expected limits for the $\tau\tau$ channel and WW channel for different κ_V (coupling of Z' to vector bosons) for universal and non-universal coupling to light (g_l) and heavy (g_h) fermions. Figure 8.7 shows exclusion bounds at 95% CL as a function of branching fraction and $M_{Z'}$ for the $\tau\tau$ channel and WW channel for universal and non-universal coupling scenarios. As we can see from the limit plots in the $\tau\tau$ channel, we exclude Z' at 2.3 TeV with 20% branching fraction in case of maximum coupling to vector bosons (i.e. $\kappa_V = 1.0$).

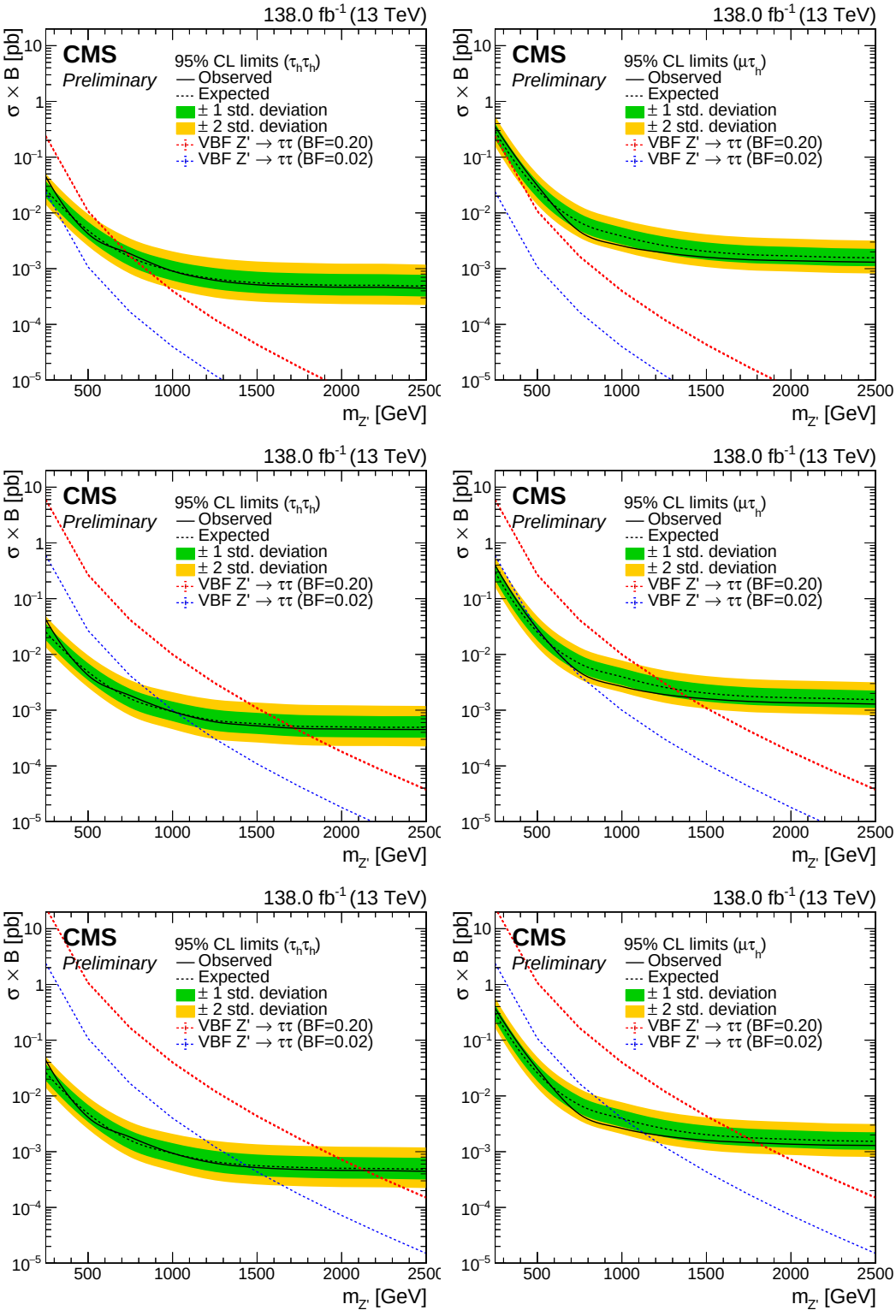


Figure 8.1: The 95% CL upper limits on the cross-section times Branching ratio for $\tau_h\tau_h$ (left) and $\mu\tau_h$ (right) for κ_V (coupling of Z' to vector bosons) : 0.1 (top), 0.5 (middle) and 1.0 (bottom) for enhanced coupling to heavy (g_h) fermions scenerio. The observed limits are also included.

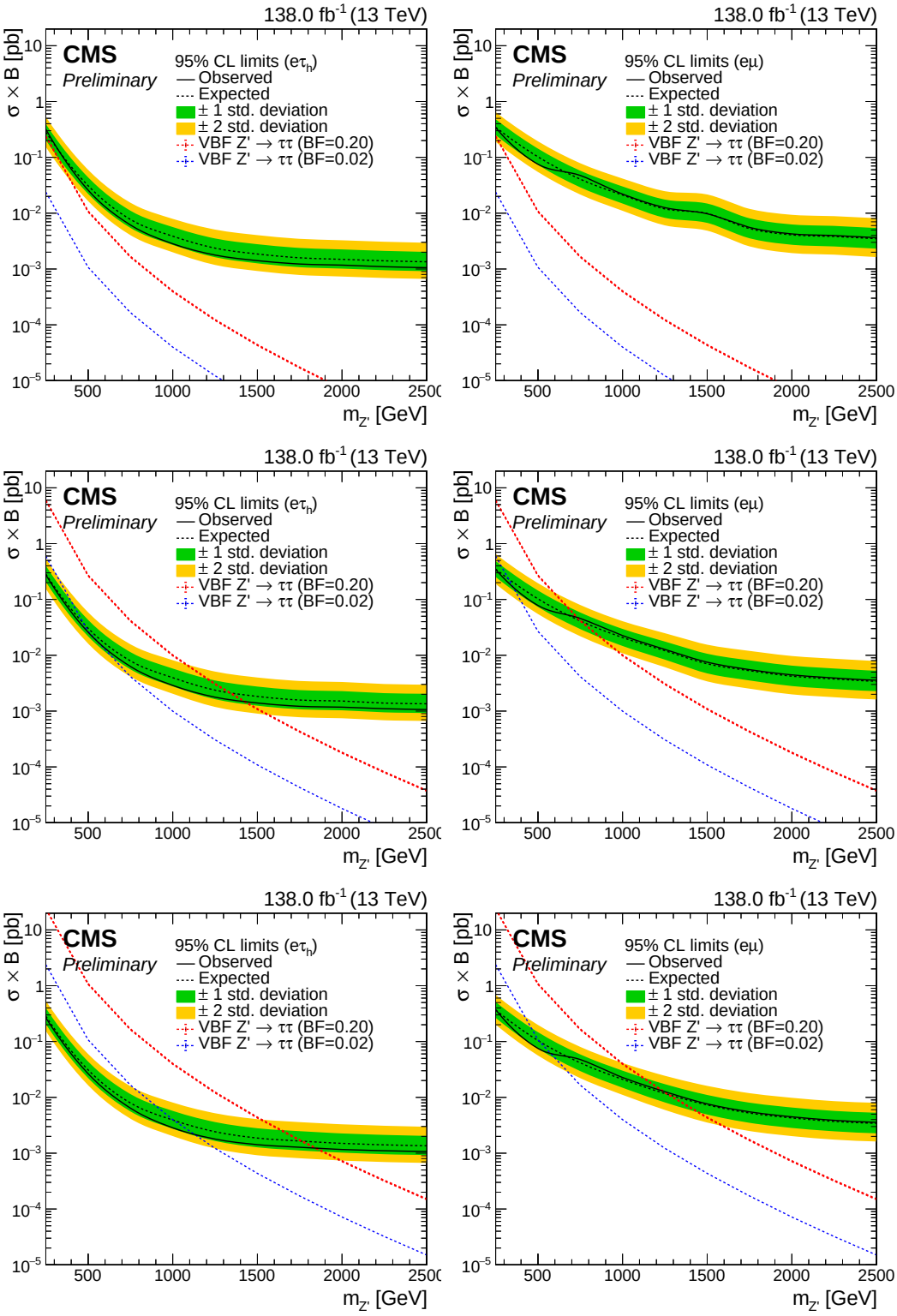
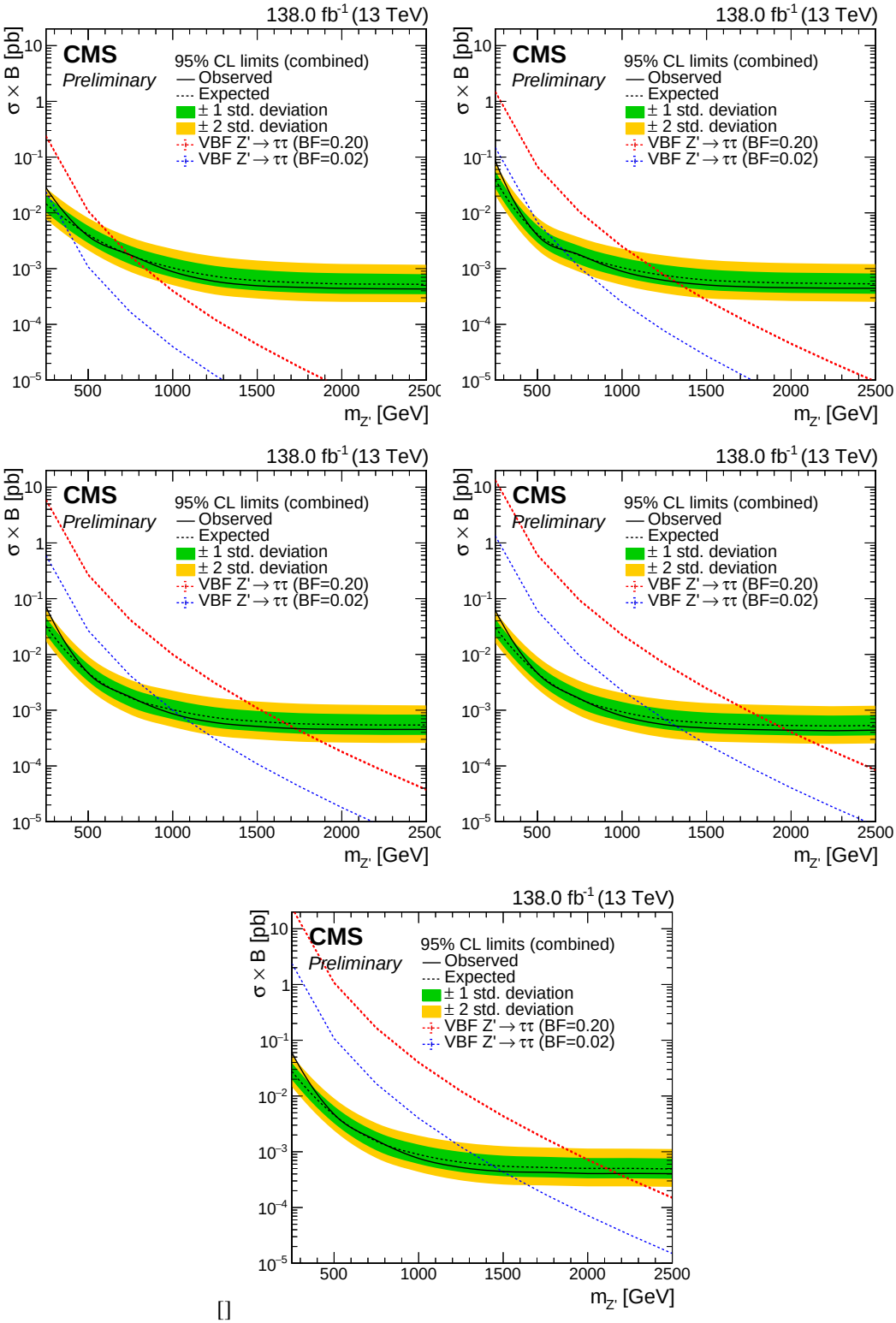


Figure 8.2: The 95% CL upper limits on the cross-section times Branching ratio for $e\tau_h$ (left) and $e\mu$ (right) for κ_V (coupling of Z' to vector bosons) : 0.1 (top), 0.5 (middle) and 1.0 (bottom) for enhanced coupling to heavy (g_h) fermions scenerio. The observed limits are also included.



[]

Figure 8.3: The 95% CL upper limits on the cross-section times Branching ratio for Z' decaying to $\tau\tau$ with universal coupling to light and heavy fermions for different κ_V : κ_V 0.1 (top-left), κ_V 0.25 (top-right), κ_V 0.50 (middle-left), κ_V 0.75 (middle-right) and κ_V 1.0 (bottom)

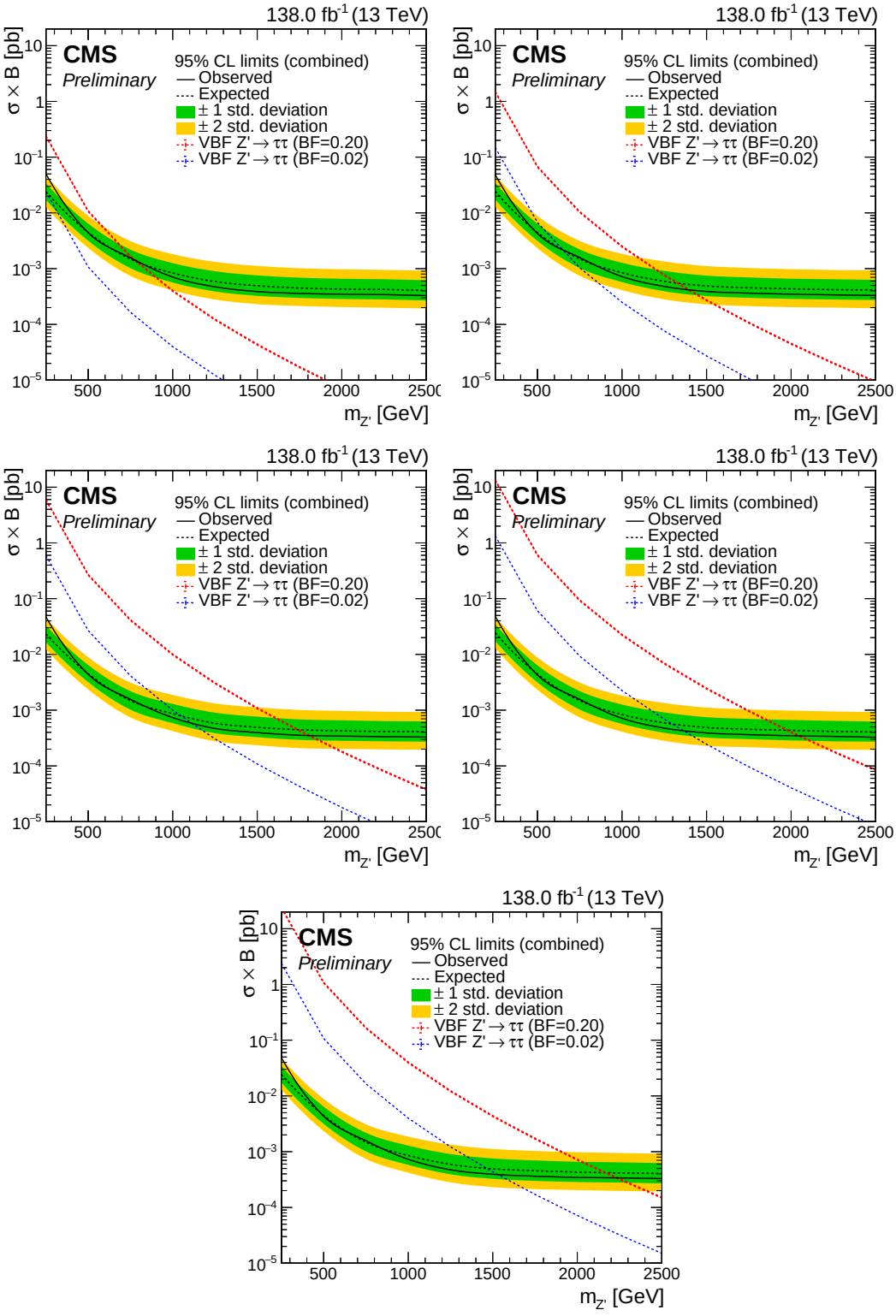


Figure 8.4: The 95% CL upper limits on the cross-section times Branching ratio for Z' decaying to $\tau\tau$ for suppressed coupling to light fermions for different κ_V : κ_V 0.1 (top-left), κ_V 0.25 (top-right), κ_V 0.50 (middle-left), κ_V 0.75 (middle-right) and κ_V 1.0 (bottom)

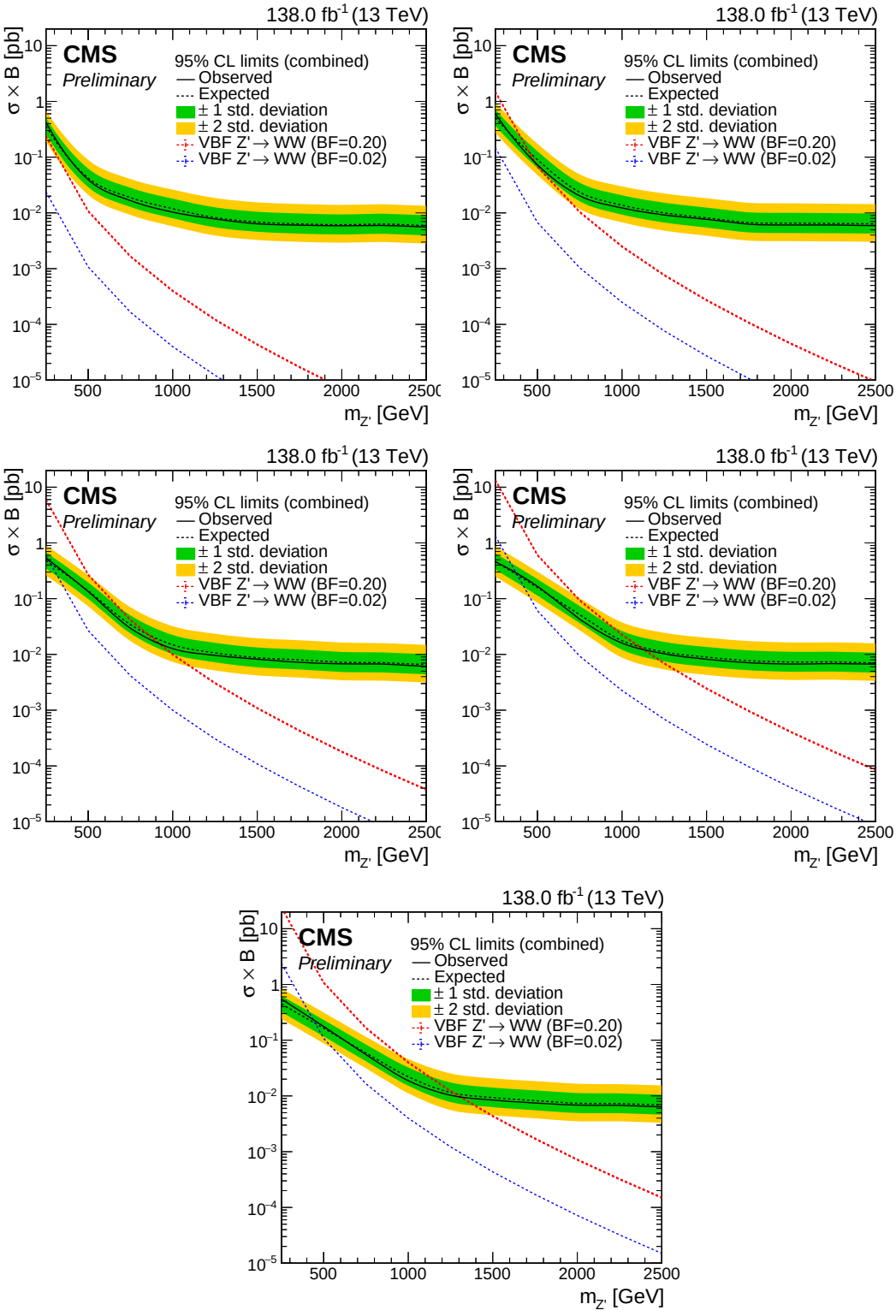


Figure 8.5: The 95% CL upper limits on the cross-section times Branching ratio for Z' decaying to WW with universal coupling to light and heavy fermions for different κ_V : κ_V 0.1 (top-left), κ_V 0.25 (top-right), κ_V 0.50 (middle-left), κ_V 0.75 (middle-right) and κ_V 1.0 (bottom)

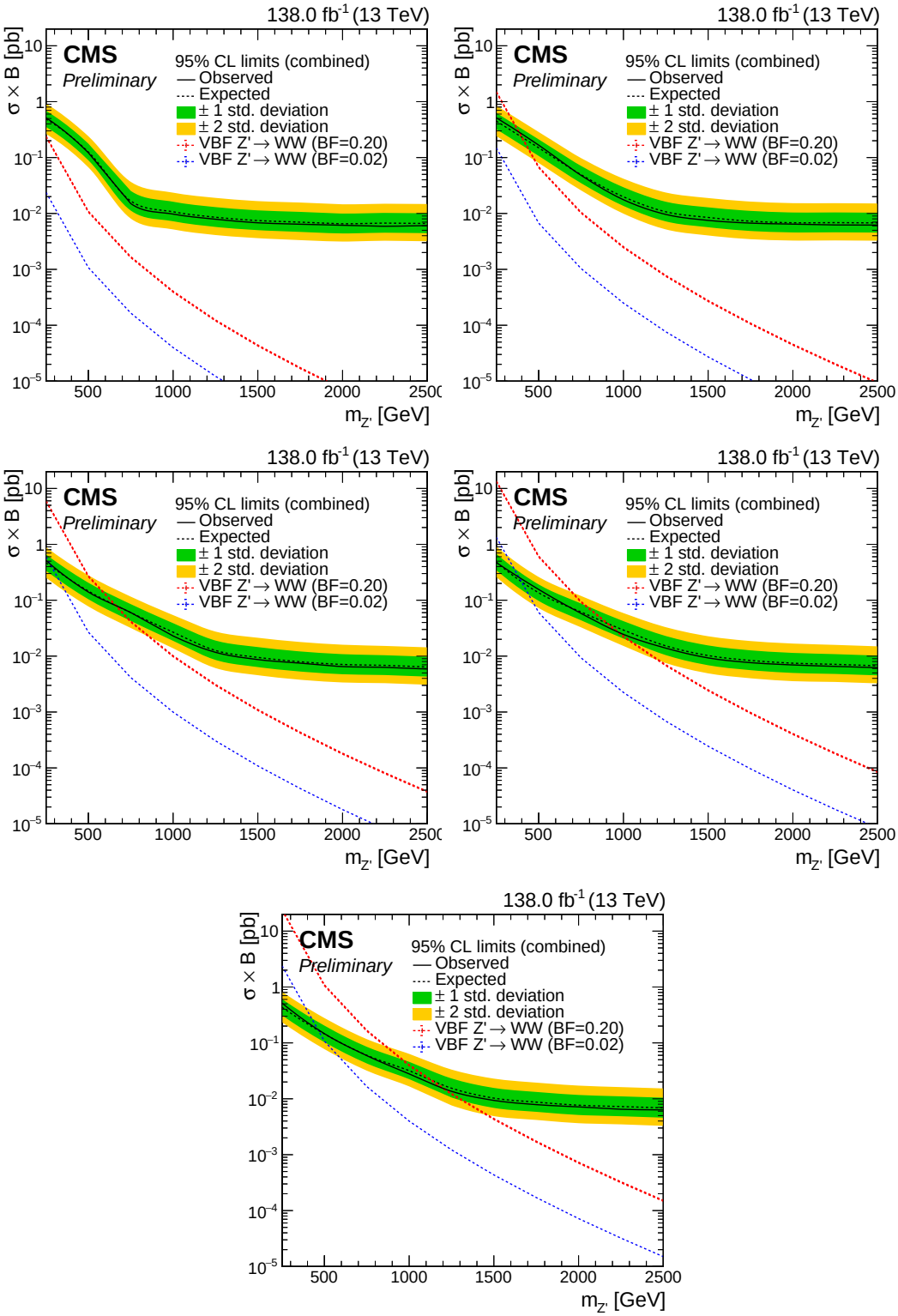


Figure 8.6: The 95% CL upper limits on the cross-section times Branching ratio for Z' decaying to WW for suppressed coupling to light fermions for different κ_V : κ_V 0.1 (top-left), κ_V 0.25 (top-right), κ_V 0.50 (middle-left), κ_V 0.75 (middle-right) and κ_V 1.0 (bottom)

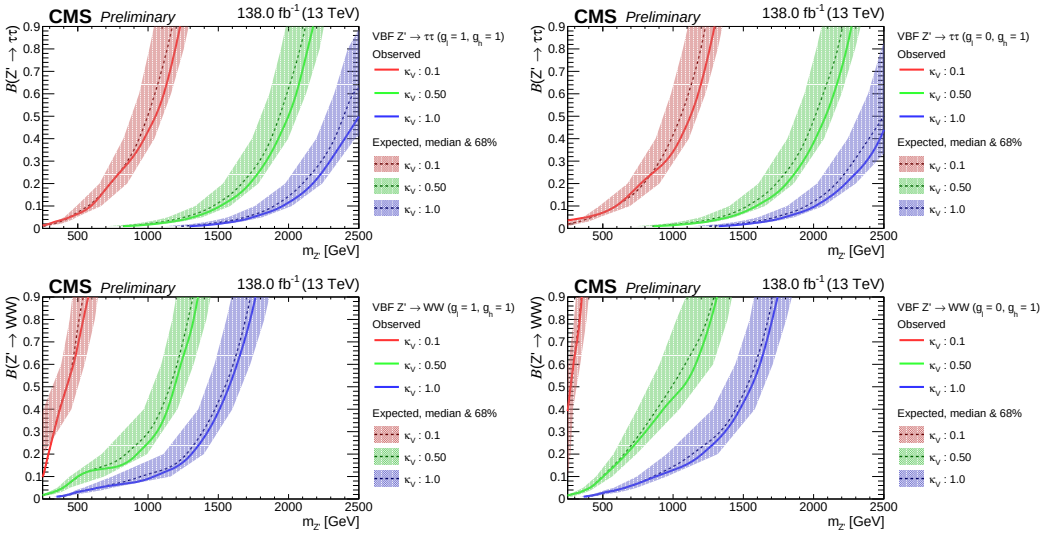


Figure 8.7: Expected and observed exclusion bounds at 95% CL as a function of branching fraction and $M_{Z'}$ for $\tau\tau$ channel (top) and WW channel (bottom) for universal (left) and non-universal (right) coupling scenarios. Regions to the left of these bounds are expected to be excluded.

CHAPTER 9

Conclusion

This thesis studied the possibility of a theoretical, TeV-scale, neutral spin-1 gauge boson called Z' produced via the process of vector boson fusion. The analysis used data from the CMS detector within the LHC over three years (2016-2018), corresponding to an integrated luminosity of 137 fb^{-1} and pp collisions with a center of mass energy of $\sqrt{s} = 13$ TeV. In particular, this thesis considered the $Z' \to \tau\tau$ channel. The invariant mass of the dilepton plus missing transverse momentum was used to search for the presence of signal as a broad enhancement above the background expectation. Background estimations were performed using Monte Carlo event generation methods that accounted for a variety of factors, including effects from the detector and pileup, to achieve the most accurate possible estimations for comparison to actual data. Signal optimization studies were performed on these MC samples to maximize the strength of the estimated signal against the background events. A dedicated background estimation was performed for each major background of the channel (DY+jets, W+jets, and QCD), and simulation alone was used to estimate the remaining minor background events.

Although data do not reveal evidence of a Z' boson, upper limits on the theoretical Z' cross-section were calculated using all VBF Z' search channels. In Z' models with non-universal fermion couplings, in particular models with Z' bosons that exhibit enhanced couplings to third-generation fermions, the presence of Z' bosons decaying to a tau lepton pair is excluded for Z' masses up to 2.45 TeV, depending on the Z' coupling to SM weak bosons, resulting in the most stringent limits on these models to date.

References

[1] M. Aaboud and et al. Search for new high-mass phenomena in the dilepton final state using 36 fb[1] of proton-proton collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Journal of High Energy Physics*, 2017(10), October 2017.

[2] G. Aad and et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):1–29, September 2012.

[3] G. Aad and et al. A search for high-mass resonances decaying to $\tau + \tau -$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *Journal of High Energy Physics*, 2015(7), July 2015.

[4] Y. Amhis and et al. Averages of b-hadron, c-hadron, and τ-lepton properties as of 2021. *Phys. Rev. D*, 107:052008, Mar 2023.

[5] Thomas Appelquist, Bogdan A. Dobrescu, and Adam R. Hopper. Nonexotic neutral gauge bosons. *Physical Review D*, 68(3), August 2003.

[6] Richard D. Ball et al. Parton distributions for the LHC Run II. *Journal of High Energy Physics*, 04:040, 2015.

[7] Debjyoti Bardhan, Pritibhajan Byakti, and Diptimoy Ghosh. A closer look at the RD and RD* anomalies. *Journal of High Energy Physics*, 2017(1), January 2017.

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

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

[10] Cheng-Wei Chiang, N. G. Deshpande, Xiao-Gang He, and J. Jiang. The Family $SU(2)_l$ x $SU(2)_h$ x $U(1)$ Model. *Phys. Rev. D*, 81:015006, 2010.

[11] CMS Collaboration. Muon physics object offline guide. https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMuons.

[12] CMS Collaboration. Physics Results BTV. https://cds.cern.ch/record/2627468.

[13] CMS Collaboration. Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV. *Journal of Instrumentation*, 13:P05011, 2018.

[14] CMS Collaboration. Performance of b tagging algorithms in proton-proton collisions at 13 TeV with Phase 1 CMS detector. https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsBTV#Physics_Analysis_Summary_and_Per, 2018.

[15] CMS Collaboration. Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV. *Journal of Instrumentation*, 13:P06015, 2018.

[16] CMS Collaboration. Performance of the reconstruction and identification of high-momentum muons in proton-proton collisions at $\sqrt{s}$ = 13 TeV. *Journal of Instrumentation*, 15(02):P02027–P02027, feb 2020.

[17] CMS Collaboration. Muon Physics Object Offline Guide. https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMuons.

[18] CMS Collaboration. Usage of b Tag Objects for 2010 Open Data Project. https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMuons. Accessed on December 1, 2023.

[19] CMS Collaboration. Utilities for Accessing Pileup Information for Data. https://twiki.cern.ch/twiki/bin/view/CMS/PileupJSONFileforData.

[20] CMS Collaboration. Performance of b-jet identification in CMS. *Journal of Instrumentation*, 8:P04013, 2013.

[21] CMS Collaboration. Search for new physics with high-mass tau lepton pairs in pp collisions at $\sqrt{s}$ = 13 TeV with the CMS detector. Technical report, CERN, Geneva, 2016.

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

[23] CMS Collaboration. Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s}$=13 TeV. *Journal of Instrumentation*, 13(06):P06015–P06015, jun 2018.

[24] CMS Collaboration. CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s}$ = 13 TeV. CMS Physics Analysis Summary CMS-PAS-LUM-18-002, 2019.

[25] CMS Collaboration. Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC. *Journal of Instrumentation*, 16:P05014, 2021.

[26] The CMS Collaboration. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08004, aug 2008.

[27] Glen Cowan. Use of the profile likelihood function in searches for new physics. pages 109–114, 2011.

[28] Andres G. Delannoy, Bhaskar Dutta, Alfredo Gurrola, Will Johns, Teruki Kamon, Eduardo Luiggi, Andrew Melo, Paul Sheldon, Kuver Sinha, Kechen Wang, and et al. Probing dark matter at the LHC using vector boson fusion processes. *Physical Review Letters*, 111:061801, 2013.

[29] Bhaskar Dutta, Alfredo Gurrola, Will Johns, Teruki Kamon, Paul Sheldon, and Kuver Sinha. Vector boson fusion processes as a probe of supersymmetric electroweak sectors at the LHC. *Physical Review D*, 87:035029, 2013.

[30] Lisa Edelhuser and Alexander Knochel. Observing nonstandard W' and Z' through the third generation and Higgs lens, 2014.

[31] Andrés Flórez, Kaiwen Gui, Alfredo Gurrola, Carlos Patiño, and Diego Restrepo. Expanding the reach of heavy neutrino searches at the LHC. *Physics Letters B*, 778:94–100, 2018.

[32] Andrés Flórez, Alfredo Gurrola, Will Johns, Young Do Oh, Paul Sheldon, Dylan Teague, and Thomas Weiler. Searching for new heavy neutral gauge bosons using vector boson fusion processes at the LHC. *Physics Letters B*, 767:126–132, April 2017.

[33] David J Griffiths. *Introduction to elementary particles; 2nd rev. version*. Physics textbook. Wiley, New York, NY, 2008.

[34] JetMET Group. Jet energy resolution: Official software tools for retrieving jer and scale factors. https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookJetEnergyResolution.

[35] Amandeep Kaur. *Search for heavy neutral gauge boson Z' using CMS detector at the LHC*. PhD thesis, Panjab University, 2022.

[36] V. Khachatryan, A. M. Sirunyan, A. Tumasyan, and W. Adam. Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV. *Journal of Instrumentation*, 12(02):P02014–P02014, February 2017.

[37] Yeong Gyun Kim and Kang Young Lee. Early LHC bound on the W' boson mass in the nonuniversal gauge interaction model. *Phys. Lett. B*, 706:367–370, 2012.

[38] Muon POG. L1 muon trigger public results. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/L1MuonTriggerPublicResults.

[39] P. Nadolsky et al. Implications of CTEQ global analysis for collider observables. *Physical Review D*, 78:013004, 2008.

[40] I. Neutelings. CMS coordinate system. https://tikz.net/axis3d_cms/.

[41] I. Ojalvo, M. Bluj, Y. Takahashi, and A. Nayak. Software guide for tau reconstruction. https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuidePFTauID.

[42] Maxim Perelstein. Little Higgs models and their phenomenology. *Progress in Particle and Nuclear Physics*, 58(1):247–291, January 2007.

[43] Thomas G Rizzo. Kaluza-Klein/Z' differentiation at the LHC and linear collider. *Journal of High Energy Physics*, 2003(06):021–021, June 2003.

[44] A. M. Sirunyan et al. Particle-flow reconstruction and global event description with the CMS detector. *Journal of Instrumentation*, 12:P10003, 2017.

[45] A. M. Sirunyan and et al. Performance of reconstruction and identification of leptons decaying to hadrons and v in pp collisions at s=13 TeV. *Journal of Instrumentation*, 13(10):P10005–P10005, October 2018.

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

[47] A. M. Sirunyan and et al. Search for high-mass resonances in dilepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Journal of High Energy Physics*, 2018(6), June 2018.

[48] A. M. Sirunyan and et al. Performance of the reconstruction and identification of high-momentum muons in proton-proton collisions at s = 13 TeV. *Journal of Instrumentation*, 15(02):P02027–P02027, February 2020.

[49] Albert M Sirunyan et al. Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV in 2015 and 2016 at CMS. *European Physical Journal C*, 81:800, 2021.

[50] Albert M Sirunyan et al. Identification of hadronic tau lepton decays using a deep neural network. 2022.

[51] Davison E. Soper. Parton distribution functions. *Nuclear Physics B - Proceedings Supplements*, 53(1–3):69–80, February 1997.

[52] Alexander Tewsley-Booth. The First Measurement of the Muon Anomalous Magnetic Moment from the Fermilab Muon $g-2$ Collaboration. 2022.

[53] A. Tumasyan and et al. Identification of hadronic tau lepton decays using a deep neural network. *Journal of Instrumentation*, 17(07):P07023, July 2022.

[54] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental Test of Parity Conservation in Beta Decay. *Phys. Rev.*, 105:1413–1415, Feb 1957.

[55] F.J. Ynduráin. Limits on the mass of the gluon. *Physics Letters B*, 345(4):524–526, 1995.