A Search for Lepton Flavor Violating Decays of the Higgs Boson and a Measurement of W Boson Production using the CMS Detector at the LHC

by

Aaron Levine

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Physics)

at the

University of Wisconsin – Madison

2016

Defended on

Dissertation approved by the following members of the Final Oral Committee:

Sridhara Dasu · Professor of Physics

Wesley Smith · Professor of Physics

Other Member · Professor of Physics

Other Member · Professor of Physics

Other Member · Professor of Other Department

Abstract

Abstract Goes Here

Acknowledgements

This is where any acknowledgements would go.

Contents

List of Figures

List of Tables

Chapter 1

Theoretical Motivation

In high energy particle physics, the goal is to use high energy densities to probe small distance scales and investigate properties of fundamental particles and the forces that govern them. This thesis describes the search for an exotic, lepton flavor violating, interaction between a Higgs boson, a muon, and a tau. These concepts are defined in subsequent sections of this chapter. This interaction has never been observed in nature and is prohibited by the Standard Model, the most comprehensive theory of particle interactions to date. If observed, these lepton flavor violating couplings would represent signs of a deeper theory that would significantly enhance our perception of forces and particles. This thesis also contains the measurement of W boson production and associated jets. This process involves key forces and particles of the Standard Model, and precise measurements of the process will improve understanding of the existing theory. These measurements take place at the Compact Muon Solenoid (CMS) detector using the Large Hadron Collider (LHC).

1.1 The Standard Model

Our knowledge of particle interactions is summarized in a theory called the Standard Model. The theoretical framework of the Standard Model was developed over the course of the 20th century as more and more fundamental particles and forces were discovered and studied. The Standard Model is not without its flaws. It is a phenomenological theory that contains many free parameters, such as the masses of the particles, that can only be determined from experiment. It also fails to include the gravitational interactions, and thus only describes three of the four fundamental forces. The ad hoc design of the Standard Model and the lack of unification provides a strong motivation

for particle physicists to search for a more profound, comprehensive theory. The current state of particles physics takes a two pronged approach, with physicists making precision measurements of the Standard Model and searching for physics beyond the Standard Model. The goal is to measure the parameters of the Standard Model while searching for the theory's extension. This thesis contains both aspects. The W+jets production cross section measurement directly tests the Standard Model while the search for lepton flavor violating Higgs couplings tests hypotheses that extend the theory.

1.1.1 Elementary Particles

The fundamental particles in the Standard Model are summarized in Figure 1.1. The fundamental particles consist of fermions and bosons. The fermions are particles that have half integer spin, and the bosons have integer spin. Spin is the intrinsic quantized angular momentum of a particle. The bosons mediate the fundamental forces and provide mass. They will be discussed further in Section 1.1.2 and Section 1.1.3. The fermions consist of the quarks, leptons, and neutrinos.

There are three generations of quarks, with two flavors in each generation. Each quark generation consists of a quark with +2/3 of an electron charge (+2/3e) and a quark with -1/3 of an electron charge (-1/3e). Quarks also have a property known as "color." There are three possible colors for each quark: red, green, or blue. The concept of color is important for understanding the strong nuclear force, as discussed in Section 1.1.2. As shown in figure 1.1 the masses of the quarks vary significantly from one generation to the next.

Free quarks have never been observed in nature. Only color neutral combinations of quarks have been observed. For example, a red, green, or blue quark can form a stable state with a corresponding antiquark, with an anticolor of antired, antigreen, or antiblue. This state is known as a meson. The baryons consists of three quarks, one of each color. The proton (uud), and the neutron (udd), are both baryons. Matter than is composed of quarks is called hadronic matter. At CMS, hadrons are often found in collimated streams of particles known as jets.

There are three generations of leptons, each with a neutrino pair. The electron is extremely light and does not decay, but the muon and the tau are heavier and have lifetimes on the order of $10^{-6}s$ and $10^{-15}s$ respectively[58]. The mechanisms for lepton decay will be discussed in Section 1.1.2 . In the Standard Model, there is a conserved quantity known as the lepton number. Each lepton/neutrino pair, such as an e and an ν_e each will have an electron lepton number of 1, and the corresponding antiparticle pair will have a lepton number of -1. Figure 1.2 gives an example of
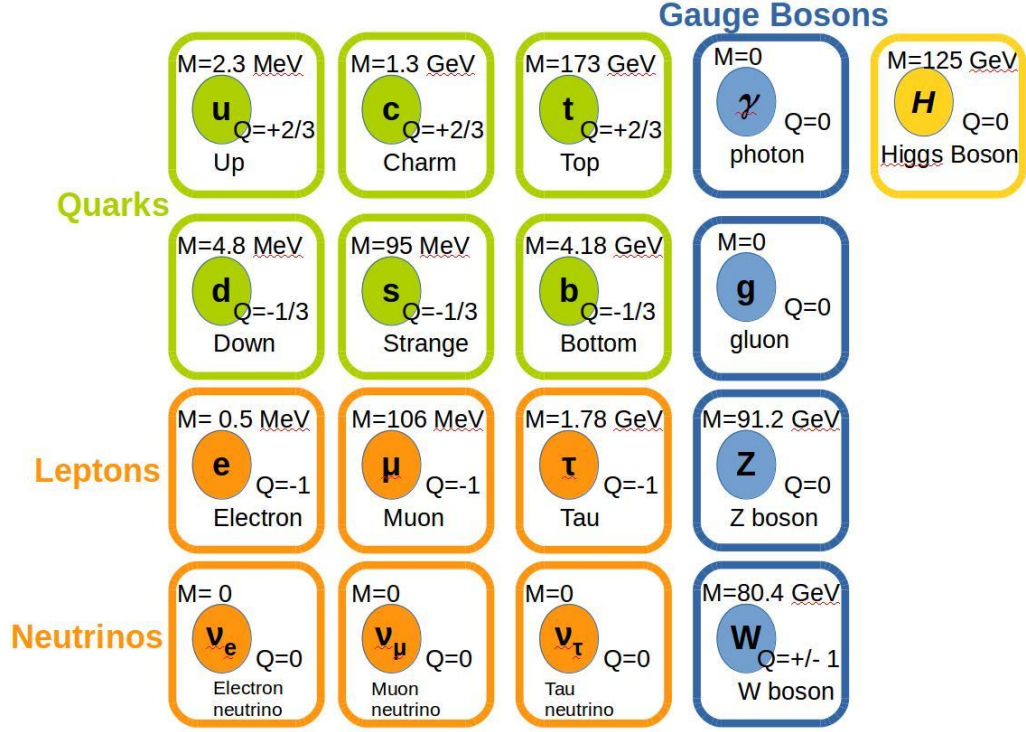
Figure 1.1: The fundamental particles of the Standard Model. Charges are expressed in units of electron charge.[58]

a decay that conserves lepton number. In the initial state only the tau exists and the tau lepton number of the system is one while the muon and electron lepton numbers of the system are 0. The decay produces a tau neutrino which also has a tau lepton number of one along with an electron and an antielectron neutrino, which have electron lepton numbers of one and negative one, respectively. So the final state also has a tau lepton number of one. Also note that lepton number and flavor are conserved at each vertex. In the lepton flavor violating couplings studied in this thesis, a muon and a tau share a vertex, which is not allowed in the standard model.

1.1.2 Elementary Forces

Force	Relative Strength	Mediator
Strong	1	Gluon
Electromagnetic	10^{-3}	Photon
Weak	10^{-11}	W,Z Boson

Table 1.1: Fundamental forces of the Standard Model[45].
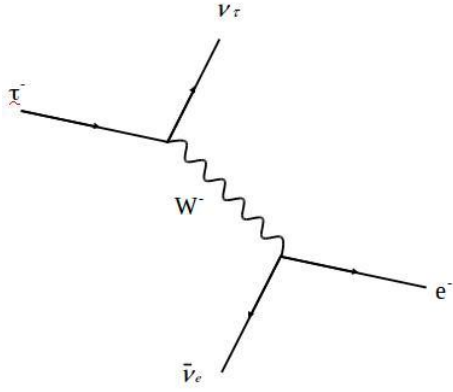
Figure 1.2: Decay of the tau lepton, as mediated by the W boson. The W boson is discussed further in Section 1.1.2.

The fundamental forces of the Standard Model are shown in Table 1.1. Note that gravity is not included in the table. Gravity is not contained in the standard model and has a strength on the order of 10^{-30} [18] relative to the weak nuclear force. It plays a negligible role in high energy physics.

The theory of the electromagnetic is know as quantum electrodynamics (QED). The Electromagnetic force is mediated by the photon. The photon is a spin 0 massless particle. Two examples of electromagnetic interactions are show in figure 1.3. This figure displays an example of a Feynman diagram, which are pictorial representations of particle interactions that are also used to calculate the matrix elements and amplitudes of the process. Matrix elements are used to compute the cross sections, which, as described in chapter2, are a measurement of the likelihood of the process occurring. The photon only interacts with particles that carry charge.
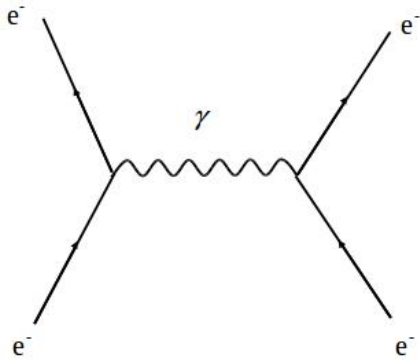
Figure 1.3: A photon mediating scattering between two electrons

The weak nuclear force is mediated by the W and Z bosons. The Z boson is a neutral spin 1 particle. It decays into pairs of quarks, leptons, or neutrinos. The W boson is a charged spin 1 particle. It has the special property of changing flavors of quarks and leptons. For example, the W boson can mediate the decay of a tau to an electron. This is illustrated in figure 1.2. Note that lepton number is still conserved in these types of interactions.

The theory of the strong nuclear force is known as quantum chromodynamics (QCD). The strong nuclear force is mediated by the gluon. Each gluon is a color doublet, with a corresponding color and anticolor. For example, a gluon could be red and antigreen, or blue and antired, et cetera. Gluons only interact with particles that share one of their color charges. Therefore, the quarks are the only particles involved in strong interactions. Virtual gluons can self interact which increases the strength of the color charge surrounding a quark. At small distance scales the virtual gluon cloud is penetrated and the effective color charge decreases. This effect is known as asymptotic freedom. [45] At larger distances, as a quark antiquark pair move apart from each other, the increase in potential energy allows new quark antiquark pairs to be created from the vacuum. The new quarks will be bound to the initial quarks by the strong nuclear force, which is why free quarks have never been observed. The coupling constant for QCD, α_s, is energy dependent and scales as $\alpha_s(Q^2) \propto 1/ln(\frac{Q^2}{\Lambda_{QCD}^2})$. Here Q^2 represents the momentum transfer of the process and Λ_{QCD} is a constant, referred to as the QCD scale, with a measured value of 214 MeV[58]. At the TeV scale collision energies of the Large Hadron Collider, α_s is small and interactions are modelled using perturbative QCD (pQCD). Leading order terms that depend on α_s and next to leading order terms that depend on α_s^2 dominant the computation and higher order terms can be neglected.

As mentioned in chapter 2.2 both the strong force and the weak force play important roles in the production of W bosons and associated jets. The strong force governs quark-gluon interactions and jet formation, while the weak force is responsible for the W boson couplings to initial start quarks and final state muons. Thus, a precision measurement of W+jets production is an powerful channel to evaluate the accuracy of the Standard Model.

1.1.3 The Higgs Boson

In classical physics, the equations of motion of a system are governed by the Lagrangian. The Lagrangian is defined as $L = T - V$ where T is the kinetic energy and V is the potential energy. In the Standard Model, we can also define a Lagrangian that governs all particle interactions[45]

However, when mass terms for W and Z bosons are added to the Lagrangian, the Standard Model breaks down. Loops formed by W and Z bosons contribute infinities to Standard Model calculations that cannot be removed. This means that the Standard Model becomes unrenormalizable.
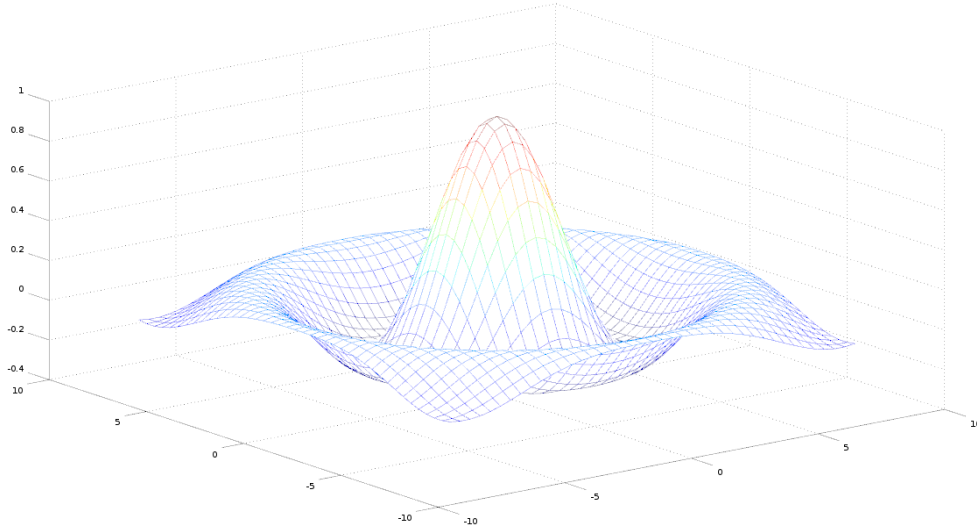
Figure 1.4: The potential of the Higgs Field. Note the nonzero minima.

To resolve this, we introduce a field with a potential shown in figure 1.4. This potential has a minimum that is not at the origin of the coordinate system. We can break the symmetry by shifting our reference point to the minimum of the potential and expanding for small deviations about the minimum. When this shift is introduced to the Lagrangian, it results in mass terms appearing for the W and Z bosons, the quarks and leptons, and a mass term for the particle associated with the field itself. That particle is known as the Higgs Boson. In addition to providing the theoretical framework to add mass terms to the Lagrangian, the Higgs Boson couples to every massive particle in the Standard Model. The couplings of the Higgs to the W and Z bosons cancel out the infinities mentioned earlier.[45] Thus, the field that naturally gives rise to masses in the Standard Model also cancels out the divergences that prevented us from simply placing mass terms in the Lagrangian to begin with. A Higgs Boson with a mass of 125 GeV was discovered at CERN on July 4 2012. [58]

1.2   Beyond the Standard Model

As mentioned in Chapter 1 Standard Model requires many free parameters determined by experiment and fails to unify gravity with the other four fundamental forces. Additionally, it fails

to account for neutrino oscillations[42][5][6] and it doesn't address the hierarchy problem. [14] A more fundamental theory must exist, and the recently discovered Higgs boson provides a rich environment for probing the existence of such a theory. Many proposed theories such as Randall-Sundrum models [60] or Two Higgs Doublet models [21] predict a Higgs boson with couplings that violate lepton flavor conservation. The strength of Higgs couplings are proportional to the masses of the particles involved, so the most logical place to start a search for lepton flavor violation would be a Higgs boson coupling directly to the two heaviest leptons: a muon and a tau.

There have been no direct searches for $H \to \mu\tau$ prior to the result contained in this dissertation. Null searches for $\tau \to \mu\gamma$ constrain the branching fraction for $H \to \mu\tau$ to $\mathcal{O}(10\%)$.[46]

Lepton flavor violating couplings may explain a discrepancy in the anomalous magnetic moment of the muon. The magnetic moment is a quantity defined for a current loop that describes the torque the current loop experiences in a magnetic field. Muons are point particles, so according to classical electrodynamics they lack a magnetic moment. However, in QED, loop interactions will effectively generate a magnetic moment, known as the anomalous magnetic moment of the muon. The current measured value differs from the expected standard model value by 2.87×10^{-9}, a significance of 3.6σ [19][58]. Loop corrections, as shown in Figure1.5 from a Higgs coupling to a muon and a tau could resolve this discrepancy [46].
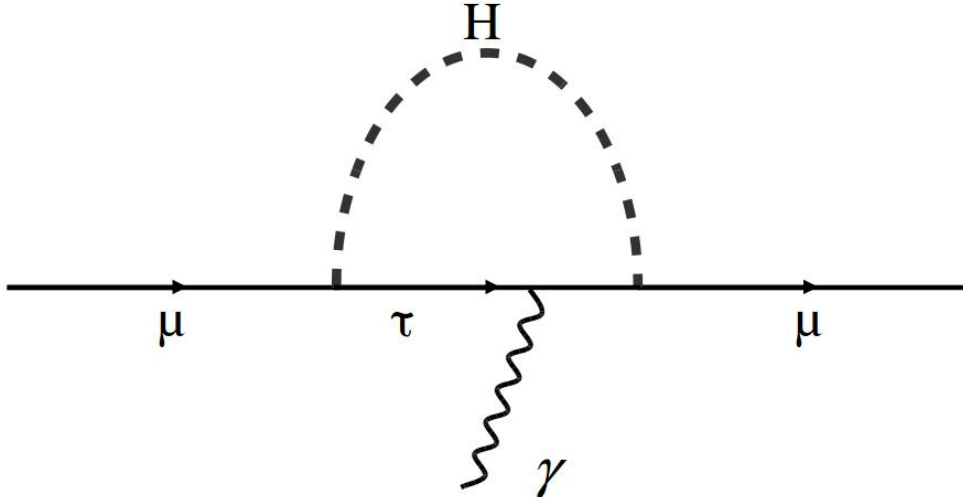
Figure 1.5: Loop contribution of lepton flavor violating couplings between muons and taus to the anomalous magnetic moment of the muon.

Chapter 2

Collider Phenomenology

Now that the underlying physics has been described in Chapter 1, the question remains: how exactly are W+Jets and lepton flavor violating Higgs events produced and studied? Before proceeding further, it is necessary to define some basic terms of collider physics. The cross section of a particle interaction is an effective area that gives a sense of the likelihood of the interaction. In collider physics, cross sections are measured in barns (b), where $1b = 10^{-24} cm^2$. The luminosity is defined as the rate of observed events divided by the cross section. The integrated luminosity is computed by integrating the luminosity with respect to time. Integrated luminosity is used to define how much total data a collider has supplied. These terms will be used for describing proton proton collisions, W boson production, and Higgs boson production in the coming sections of this chapter.

2.1 Proton-Proton Collisions

The Large Hadron Collider (LHC), which will be discussed further in Chapter 3, is a proton-proton collider. The proton, as mentioned in Chapter 1, is considered to be the bound state of three quarks: two up quarks and a down quark. However, proton-proton collisions cannot simply be thought of as six up and down quarks interacting. Inside the proton, massive amounts of quark and antiquark pairs pop in and out of existence. At the high energies and short distance scales probed by the LHC, these "sea quarks" play a dominant role in proton proton collisions. The gluons that bind elements of the quark/antiquark sea also play a large role in collisions. In fact, these gluons carry about 50% of the proton's momentum[45]. So when two protons collide at the LHC, it is effectively a collision between two clouds of gluons and quarks.These collisions can be studied via

the parton model. This model defines the protons as collections of point particles (partons) carrying a fraction x of the hadron's momentum. The distribution of x depends on the momentum scale Q^2 that is probed.

When these clouds of gluons and quarks collide, most of the collisions involve a small amount of momentum transfer between the particles involved. These are known as "soft" scattering. The important processes, such as W boson or Higgs boson production, require a "hard" scattering event. These types of events seldom happen more than once each time the clouds of gluons and quarks intersect. Therefore, for each interesting hard scattering event there will be many soft scattering events occurring at the same time. The soft scattering events are referred to as the "underlying event."

In addition to the hard scattering process of interest, there will be extra proton-proton interactions in each event which produce additional hadrons. These secondary interactions are referred to as "pileup". The additional hadrons produced from the secondary interactions threaten to cloud the hard scattering process at the primary vertex. Methods of vetoing and subtracting pileup contributions to the event are discussed in section 5.

As quarks and gluons move apart after the collision, the increase in potential energy from the strong interaction produces a stream of collimated hadrons, known as a jet. Jet formation is shown in Figure 2.1. First, energetic quarks radiate gluons, which split into quarks. These quarks may still be energetic enough to radiate gluons, and the process continues. This is known as parton showering. Detailed simulations are described in chaper 4. After the conclusion of parton showering, hadronization occurs. QCD confinement requires that the quarks and gluons are naturally grouped into colorless clusters. These clusters then separate into groups of hadrons, which is known as hadronization. The identification of jets is discussed in Chapter 5.

2.2   W Boson Production: Mention prior results

W bosons are produced by deep inelastic scattering between partons, as show in Figure 2.2. The cross section of this process depends on the parton distribution functions. Examples of CTEQ6 parton distribution functions are show in Figure ??[59]. These parton distribution functions are used in PYTHIA (chapter 4).

W bosons decay to leptons and quarks at equal rates. The W may decay to an up/down quark pair, a charm/strange quark pair, or a lepton/lepton neutrino pair. The presence of the muon
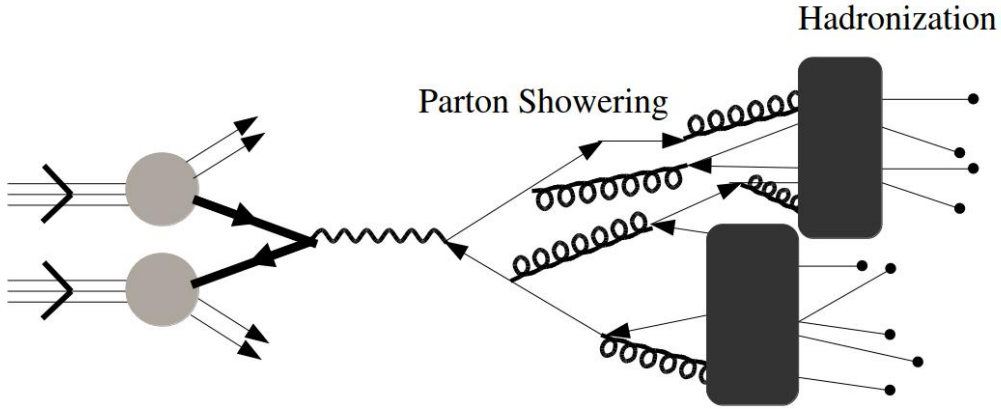
Figure 2.1: Diagram of jet formation occurring after a hard scatter event. First quarks and gluons are radiated in parton showering. Then they combine into colorless cluster before condensing into the hadrons that make up jets.
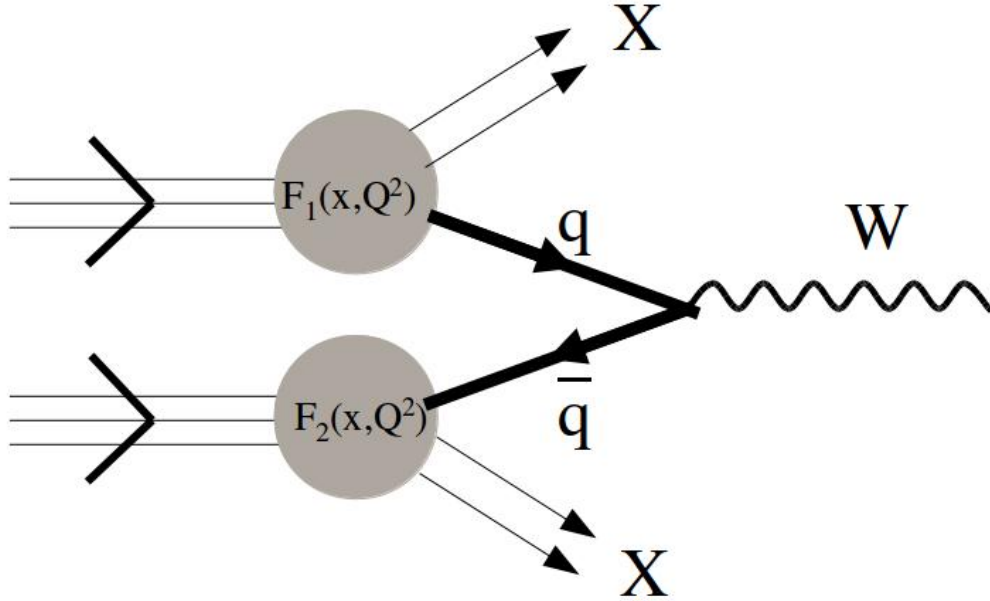
Figure 2.2: Diagram of W Boson production. The function F_1 and F_2 are the parton distribution functions of the incoming protons.

neutrino in the final state is necessary to conserve lepton flavor in the standard model, as discussed in Chapter 1. The decay to a bottom/top quark pair is suppressed because of the high mass of the top quark. Each quark has three flavors, and there are three total leptons, so ultimatly W decay has six hadronic final states and three leptonic final states. This means that W bosons decay to hadrons approximately 67% of the time and decay to leptons about 33% of the time. As discussed
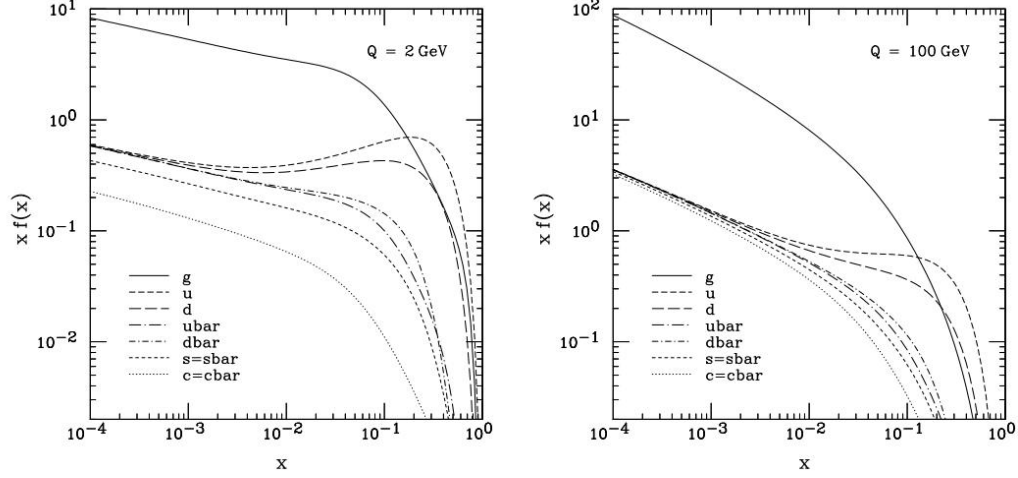
Figure 2.3: Parton distribution functions in the CTEQ6 scheme[59]. As the energy Q of the collision increases, the quarks involved carry less and less of the proton's momentum and the applicability of pQCD increases.

in Chapter 3 and Chapter 5, due to CMS's ability to reconstruct muons, we look for W bosons decaying to a muon and a muon neutrino. Because of lepton universality, the branching fraction of $W \to \mu\nu_\mu$ is about 11%.

W boson production is studied based on the number of jets produced in association with the W boson. An example of a single radiated jet is show in Figure 2.4. The cross section for W boson from two partons scattering is given by equation 2.1 [18]. Here, V_{ud} is a matrix element that defines the probability for up-down quark transition, G_F is a fundamental constant, M_W is the mass of the W boson, and $\hat{s}$ is the momentum of the diquark system. The hats on $\hat{s}$ and $\hat{\sigma}$ indicate that these values apply to one parton scatting off of another parton. To calculate the full cross section of W production we must integrate over the parton distribution functions of the quarks involved.

$$\hat{\sigma}(ud \to W) = 2\pi |V_{ud}|^2 \frac{G_F}{\sqrt{2}} M_W^2 \delta(\hat{s} - M_W^2) \qquad (2.1)$$

Previous results from CMS measured at 7 TeV center of mass collisions[48] are show in Figure 2.5.

2.3 Higgs Boson Production

As discussed in Chapter 1, Higgs boson couplings are mass dependent. The more massive the particle, the more likely it will couple to the Higgs. Top quarks, with a mass of 172 GeV, are the
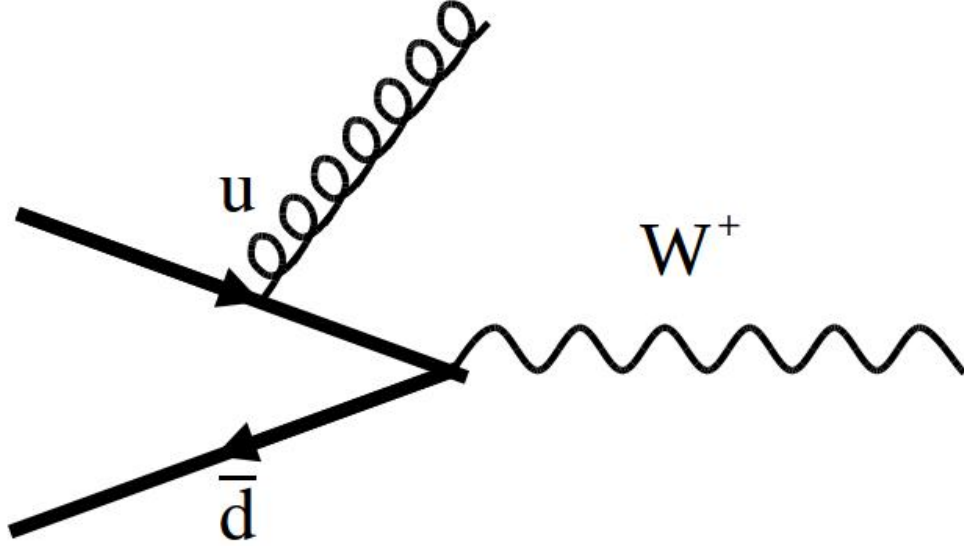
Figure 2.4: Example of the production of a W plus one jet event.

heaviest standard model particle. As a result, the dominant method of Higgs production is when two energetic gluons fuse together via create a top quark loop, as shown in Figure 2.6. This is called gluon-gluon fusion.

Vector boson fusion is the dominant mechanism of two jet Higgs boson production. It is shown in Figure 2.7 Because of conservation of color, the two final state quarks must be the same color as their corresponding initial state quarks. This will force the jets produced by the final state quarks apart at a wide angle. A graphic of all the Higgs production channels is shown in Figure 2.8[37].

There is no known Higgs branching ratio to guide this search. This is the first direct measurement of $H \to \mu\tau$. The kinematics of this search are similar to the kinematics $H \to \tau\tau$ search carried out at CMS [27]. As discussed in chapter 5.4.2, tau leptons are detected at CMS via their hadronic decay mode (τ_h) or their leptonic decay modes (τ_e, τ_μ). Therefore, the final state objects, a muon and a tau, are exactly the same between $H \to \mu\tau$ and the $H \to \tau_{mu}\tau$ channel of $H \to \tau\tau$, making $H \to \tau\tau$ a major background of the search. A major difference between $H \to \tau_{mu}\tau$ and $H \to \mu\tau$ is the distribution of missing transverse energy (MET) in the event. MET, which is discussed further in chaper 5.5, arises from neutrinos that CMS cannot detect. As shown in figure 1.2, neutrinos are always associated with tau decays. The $H \to \tau_{mu}\tau$ channel will contain a muon
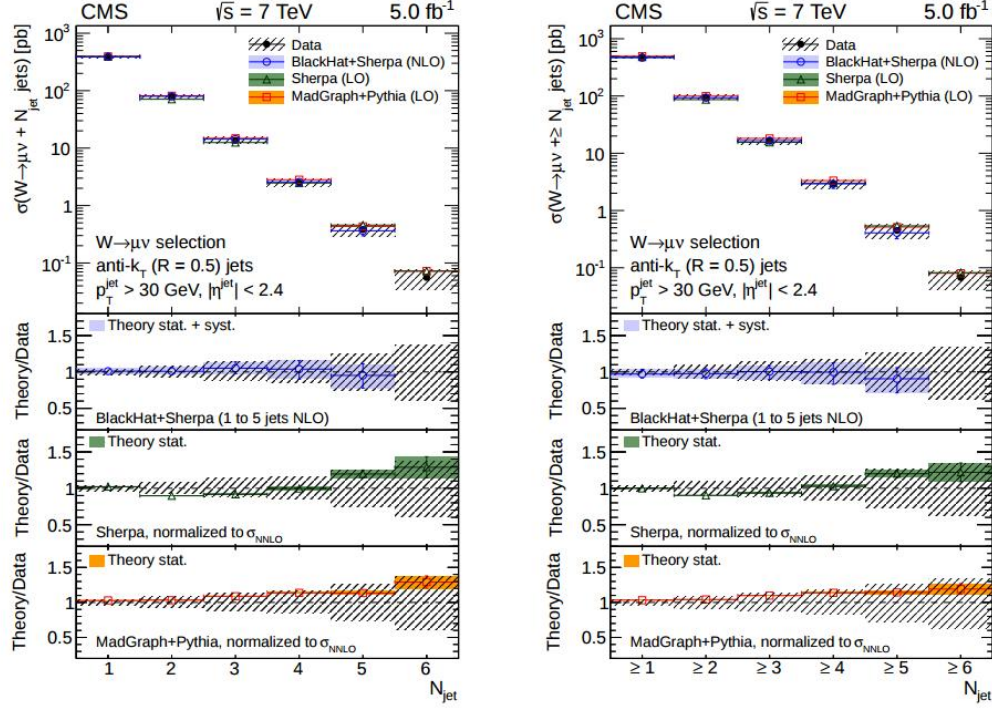
Figure 2.5: W+jets production at CMS in 2011. Cross sections are show in units of picobarns. The left plot shows exclusive jet binned (= NJets) results and the right plot shows inclusive jet binned (≥ NJets) results. Comparisons between theoretical predictions and data are shown in the ratio plots. Agreement is within statistical uncertainty.

neutrino associated with the muonic decay of the tau, but $H \rightarrow \mu\tau$ contains no such neutrino. The only neutrinos will be associated with the tau decay. The MET distribution will be used to define the signal region in chapter 6. A graphic of Higgs boson decay probabilities is shown in Figure 2.9[13]

2.4 Kinematics Definitions

Before proceeding further it is necessary to define relevant kinematic variables. The protons are accelerated up to almost the speed of light, so a relativisitic treatment must be used to study the collisions. In special relativity, the energy of the system, E, is defined as $E^2 = m^2 + p^2$. Here the speed of light, c, is set to 1, in accordance with natural units. The energy momentum 4-vector is defined as $\hat{P} = (E, \vec{p})$, and the dot product is defined such that $\vec{P}P = E^2 - p^2$, which is the lorentz invariant mass of the particle. The square of the sum of four momentum vectors gives the invariant mass of a system of particles. Lorentz invariant quantities are useful because they are
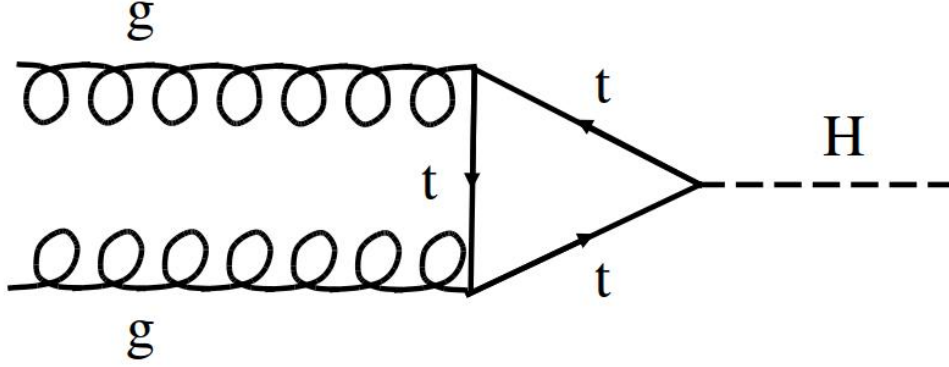
Figure 2.6: Higgs production via gluon gluon fusion.
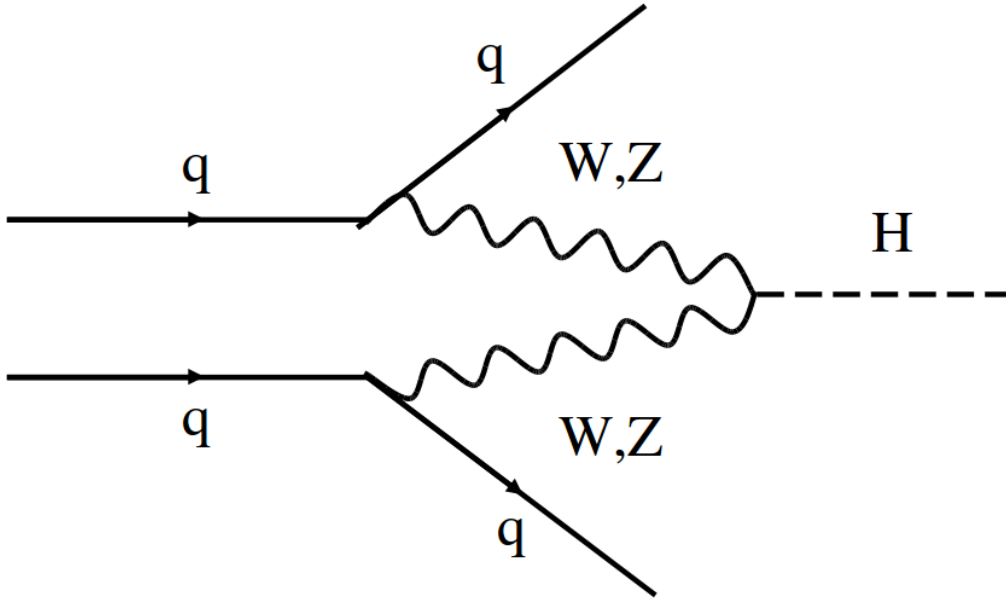
Figure 2.7: Higgs production via vector boson fusion.

frame of reference independent.

For two colliding particles, p_1 and p_2 we can define a quantity: $s = (p_1^2 + p_2^2)$, which is the Lorentz invariant square of the total center of mass energy of the event.

The rapidity y of a particle is defined as: $y = \frac{1}{2}ln(\frac{E+p_z}{E-p_z})$, where p_z is the component of the momentum parallel to the beam axis. The rapidity gap between two particles, Δy_{12}, is lorentz invariant. The pseudorapidity η is defined as : $\eta = \frac{1}{2}ln(\frac{p+p_z}{p-p_z})$. This can be expressed in terms of polar angle (Section 3.2) as $\eta = ln(cot(\theta/2))$. The pseudorapidity is equal to the rapidity if the rest
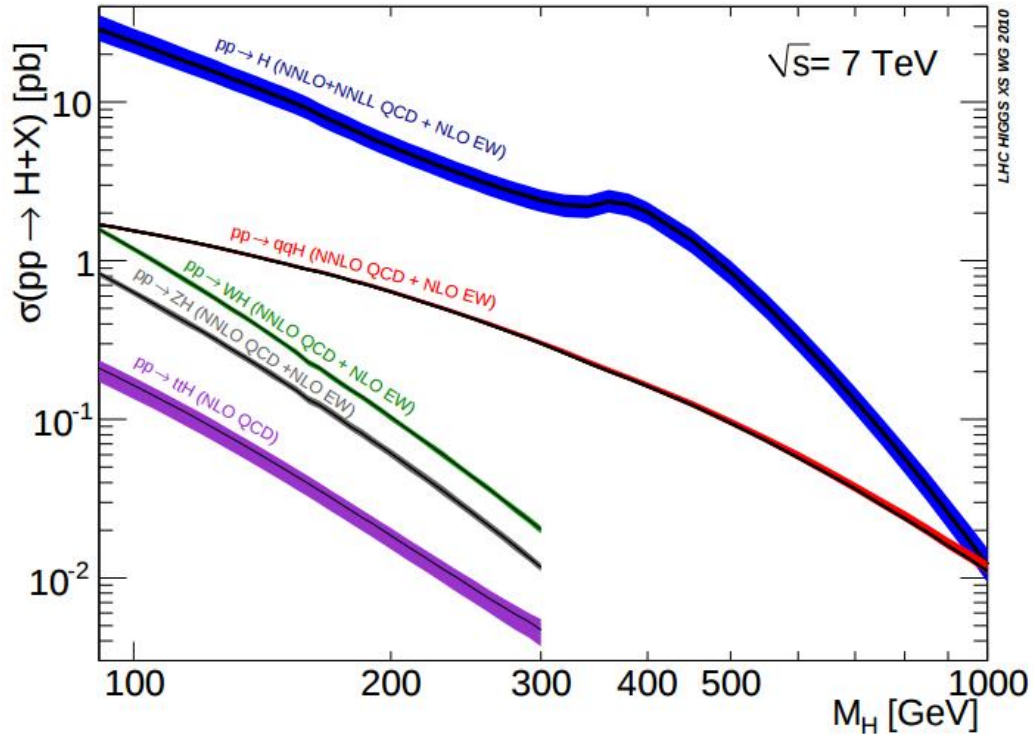
Figure 2.8: Higgs production cross sections[37]. At a mass of 125 GeV, the two largest production modes are gluon gluon fusion and vector boson fusion.

mass is 0. At high energy collisions at the LHC, $E \gg m$ and $y \approx \eta$.
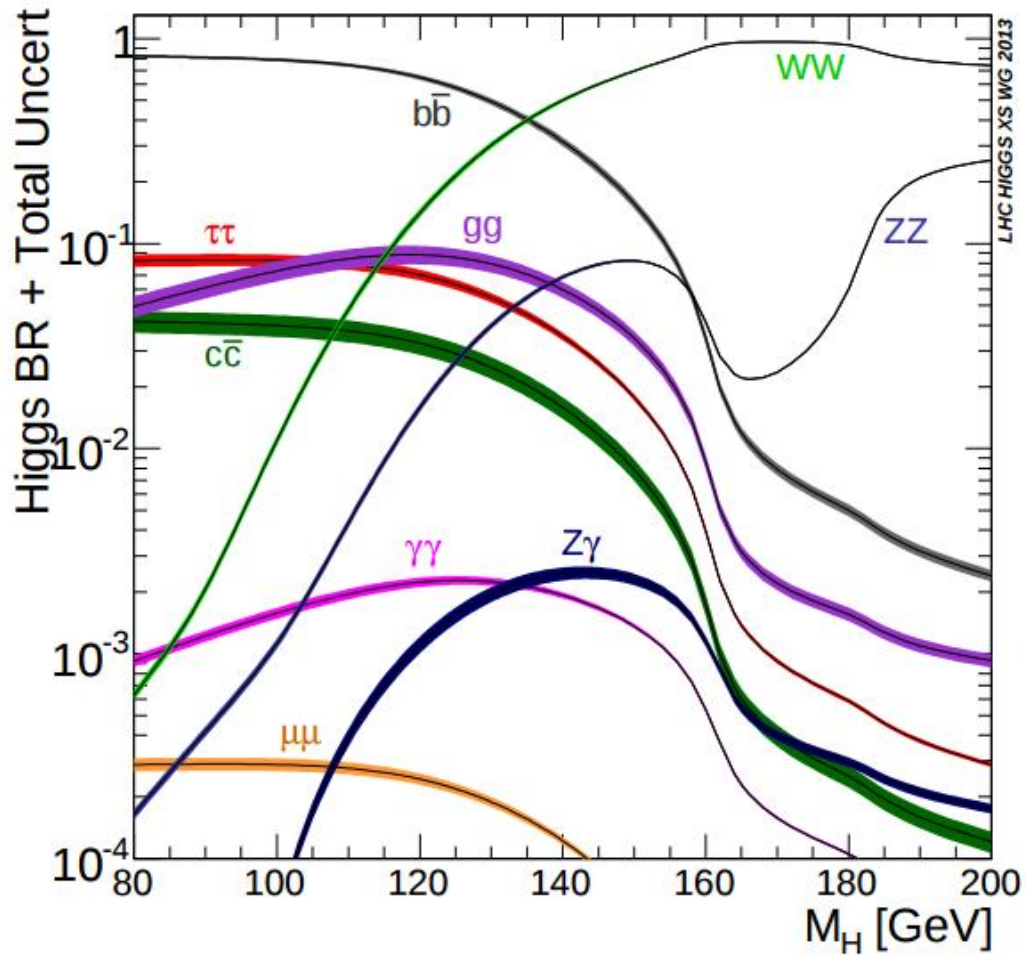
Figure 2.9: Higgs decay branching ratios[13]. The mass of the Standard Model Higgs is 125 GeV.

Chapter 3

Experimental Design: The Headings below are self explanatory

3.1   LHC

3.2   CMS

3.2.1   Overview

3.2.2   Tracker

3.2.3   ECAL

3.2.4   HCAL

3.2.5   Muon System

3.2.6   Trigger

Chapter 4

Event Simulation

Particle interactions at CMS are computationally intensive to model. The strong nuclear force plays a dominant role in proton-antiproton collisions, but calculations involving QCD depend on large numbers of terms that define interactions between many quarks and gluons. At short distances, on the order of a femtometer, we can define the momentum scale Q to be much greater than Λ_{QCD}, as defined in chapter ?? This means that the effects of QCD can be calculated perturbatively (pQCD), which means that high order terms can be neglected and the calculations can be simplified. However, there are large amounts of soft radiation at Λ_{QCD}, as explained in chapter 2.1, which necessitates the use of computational software.   The interactions of the collision products with the the detector also need to be modelled via simulation. Accurate models of physical processes at CMS are vital for testing existing theories and searching for new ones, so accurate modelling of these processes is very important for CMS and for particle physics in general.

4.1   Monte Carlo Event Generation

Physical processes at CMS are simulated using a class of software called Monte Carlo generators. These programs are named after the location of the famous casino because Monte Carlo software leans heavily on random number generation to simulate the kinematic distribution and decay chains of the event products. When using Monte Carlo software to simulate collisions, the user must specify the center of mass energy, the initial colliding particles, and the desired final products. Additional parameters can be defined by the user, such as the hadronization scale discussed in chapter 4.1.3. The three main components of Monte Carlo simulation are matrix element computation, parton

showering, and hadronization.

4.1.1   Matrix Elements

Once the initial and final state particles are specified, a series of Feynman diagrams are created. From the matrix elements discussed in chapter 1, production amplitudes are computed for the process. However, this calculation provides only a very basic picture of the event and neglects soft radiation at the pQCD scale.

4.1.2   Parton Showering

As discussed in chapter 2, in high energy collisions protons can be modeled as collections of partons where the partons are pointlike particles carrying a particular fraction of the proton's momentum. Parton distribution functions, as discussed in chapter 2, provide a model of how the protons will interact in a collision.

After the proton-proton collisions, Sudakov Form Factors[58] are computed, which represent the probablity of a parton splitting into multiple partons. A low momentum bound for splitting is defined, and all partons above this threshold are randomly split in accordance with the probability of splitting. Color is properly accounted for at each vertex. Parton showers simulate QCD radiation emitted by quarks in the form of gluons, or a gluon splitting into two quarks.

At this point it is necessary to reconcile the matrix element computation, which represents high energy hard scattering, with parton showers, which model soft radiation. Two methods are available. The matrix element and parton shower method (ME+PS) and the next to leading order and parton shower method (NLO+PS)[58]. In the ME+PS method, matrix elements are computed for the fundamental process with the addition of n partons. The additional partons are required to be separated by a specified transverse momentum threshold. The momentum threshold is chosen to be at the upper limit of pQCD. In this way, the event can be computed accurately at large angle via matrix element methods, and then parton showering algorithms can be applied to the additional partons in the event. The ME+PS method is good for simulating events with many hard jets that are well separated. These kind of jets are simulated much better with tree level computations rather than lower energy pQCD parton showering. The next to leading order and parton shower method (NLO+PS) extends to parton shower method to next to leading order to QCD.

4.1.3   Hadronization

After parton showering, the event consists of the hard final products and many soft partons, as defined in chapter 2.1. At this point, the partons must transform into color singlet final state hadrons. One way to the this is the Lund string model[**?**]. In this model, quark and anti-quark pairs are connected by color "strings" with a potential $V(r) = \kappa r$, where r refers to the distance between the quarks and κ is a dimensionless constant that defines the strength of the potential. Since the potential is directly proportional to the distance between the quark pairs, the energy of the system increases as the quarks move further apart. Eventually, the energy of the system is enough to generate an additional quark/anti-quark pair, which effectively breaks the string into two separate string, as shown in figure 4.1.The p_T of the quark or anti-quark is $< p_T^2 > = \kappa/\pi$ and the Lund fragmentation function[58] defines the fraction of the longitudinal momentum of the endpoint particle that is imparted to its recently produced neighbor. In this fashion, the kinematic variables of the produced hadrons are known, and the shower continues until an energy scale cutoff is reached.
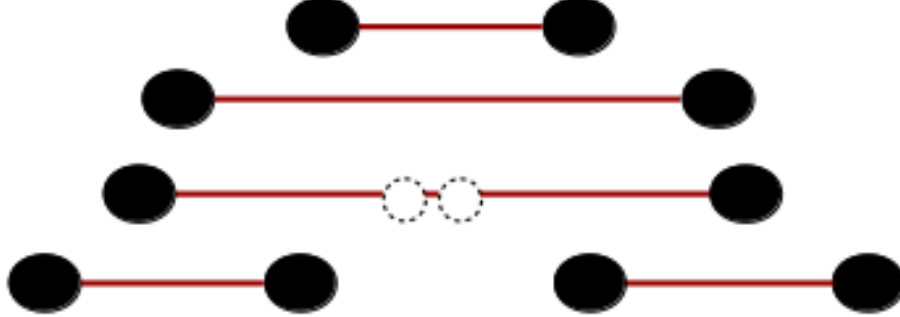
Figure 4.1: An illustration of the Lund string model. As the quark/anti-quark pair move further apart, the increase in potential energy creates an additional quark/anti-quark pair.

4.1.4   Monte Carlo Generator Software

A variety of different Monte Carlo generators are used at CMS. MADGRAPH [10] is used to compute matrix elements. Next to leading order (NLO) matrix elements can be computed with the aMC@NLO[10] version of MADGRAPH or with POWHEG[8, 40, 57]. NLO calculations are vital for accurately depicting physical processes such as W+Jets or Z+Jets. These calculations include quark and gluon loops that complement the fundamental hard scattering process. However, MAD-

GRAPH and POWHEG do not include parton showering and hadronization, which are necessary for accurate modelling of proton-proton collisions and jet formation, as discussed in chapter 2.1. The output of the matrix element generators are piped into PYTHIA, which models parton showers and hadronization using the Lund String Model. While PYTHIA is a powerful tool for calculating parton showering and hadronization, it is only a leading order (LO) generator, so it must receive matrix element results MADGRAPH or POWHEG to give results at the desired degree of precision. After hadronization, some heavy states may still need to decay. The τ lepton is too short lived to be directly observed in the detector, so any Monte Carlo simulation must decay the τ further. TAUOLA [65] is interfaced with PYTHIA to provide an accurate view of τ decay by taking into account τ helicity and polarization. The Monte Carlo for FeMtobarn (MCFM) program [24][26][25] is used to calculate the expected cross sections of the Monte Carlo simulation.

4.2   Detector Simulation

After the simulation of the physical process, complete with parton showering and hadronization, it is necessary to model the interactions of the finale state particles with the CMS detector. This is done using GEANT4[9]. First, an accurate model of the CMS detector must be built in GEANT4, defining both the geometry and material components of the detector. The simulation is then carried out in two steps: tracking and detector response. The tracking step simulates the passage of particles through matter, modelling the energy lost based on the particles and the detector material. The next step is to model the detector response. This will simulate the signal that each event will create. After the the GEANT4 simulation has completed, the output is converted to the same software format as the actual data and then reconstructed with the same algorithms discussed in chapter 5. The fully simulated events will retain all of the original Monte Carlo scattering information, so we can now understand what a particular physical process will look like from the point of view of the CMS detector. This allows us to study and improve the reconstruction of data, as discussed in chapter 5.

Chapter 5

Event Reconstruction

Studying physical processes requires reconstructing the particles in each event from the detector response. This is accomplished via the particle flow algorithm, which is discussed in Section 5.1. Specific applications to physics objects are discussed in subsequent sections of this chapter. In particular, the important objects for $H \rightarrow \mu\tau$ are muons, electrons, hadronic tau deposits, jets, and missing energy. The important objects for W+Jets are muons, jets, and missing energy.

5.1 Particle Flow Overview

As discussed in Section 3.2, the CMS detector is divided into many layers, each of which is responsible for measuring a unique part of an event. The detector response from each layer can be thought of as a building block for an event. Particle flow uses a series of algorithms to link these blocks to reconstruct the particles in the event. More specifically, particle flow algorithms link tracks in the tracker to energy deposits in the calorimeters or tracks in the muon system. Energy deposits that have no tracks associated with them are due to neutral particles. Ultimately, effective tracking algorithms are necessary before most particles can be identified.

Tracking algorithms build up the full tracks iteratively. Track construction begins by identifying hits in the pixel layer. There must be hits in at least three pixel layers with an energy of at least 200 MeV [30] These collections of hits are used to seed a track. The track is then propagated to subsequent layers of the tracker by a Kalman Filter technique [41]. A χ^2 fit is then used to identify the silicon strip hits that best correspond to the pixel triplets. All hits in the track are then removed from the list of hits to consider, and the algorithm is repeated for subsequent collections

of pixel triplets and silicon strip deposits. This is shown in figure 5.1.
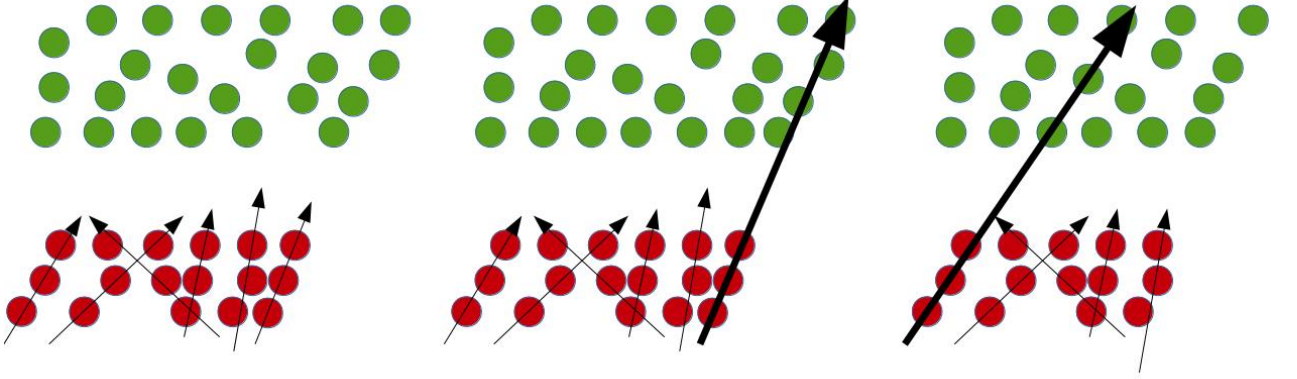
Figure 5.1: A visual depiction of the iterative tracking algorithm. First, pixel triplets (in red) are identified. Secondly, the best match of pixel hits and track hits is identified. Finally, the identified track is removed and the algorithm finds the second best track candidate.

Nuclear interactions in the tracker material can cause tracks to appear displaced. These interactions are identified by linking three tracks to a common secondary vertex. A secondary vertex is a vertex that is displaced from the primary vertex of the collision. The primary vertex is identified by selecting the vertex with the highest p_T^2 sum of its associated tracks. A primary track is required to link the secondary vertex to the primary vertex. The tracks linked to the secondary vertex are required to have an invariant mass greater than 200 MeV.

After identifying the tracks, their positions are extrapolated to associate them with energy deposits in other parts of the detector. Due to the magnetic field, the radius of the track will depend on the particle's momentum. The momentum of the track is compared to the energy measured in the calorimeters or the muon system to determine if the track was associated with the deposit. In this way, tracks and calorimeter deposits are linked together to reconstruct each particle of the event. Each particle has its own unique reconstruction conditions, which are described below.

5.2   Electrons

Electrons are identified by associating tracks to objects known as "superclusters." As electrons interact with the detector, they will lose kinetic energy in the form of radiated photons. This is known as bremsstrahlung radiation [44]. These radiated photons are not affected by the magnetic

field of CMS. As a result, the electron energy will be spread out in (η, ϕ) in the ECAL because the radiated photons are unaffected by the magnetic field.

Tracks from the standard iterative tracking procedure are used as electron candidates. Their p_T is required to be greater than 2 GeV. The electrons may not interact substantially with the detector, so if the track is propagated to the ECAL (Section 3.2.3) and if its p_T matches the ECAL cluster deposit, then the track is pre-identified as an electron. Otherwise the track is fit with a Gaussian-sum filter (GSF) to account for the non-Gaussian energy loss [4] A boosted decision tree [47] is then used to select the final GSF electron track candidates.

Photons are neutral particles and will therefore not leave any tracks in the tracker. Bremsstrahlung photons are reconstructed by examining tangents to the tracks at each tracker layer, where the bremsstrahlung radiation may have occurred as the electrons passed through the layer. These tangent tracks are extrapolated to the ECAL. If any of the extrapolated bremsstrahlung photon tracks are associated with clusters in the ECAL then those ECAL clusters are considered part of the electron supercluster. A similar procedure is used to add clusters associated with bremsstrahlung photon e^+e^- pair production. Tracks associated with the supercluster have a high probability of being associated with jets or other particles. A more sophisticated method is needed to properly associate electron tracks with the supercluster. Tracks linked to cluster if extrapolated position overlaps with η/ϕ position of cluster.

The GSF electron tracks are then matched to an ECAL supercluster of at least 10 GeV. No more than 10% of the ECAL should be found in a topologically linked section of the HCAL. A boosted decision tree [47] takes as inputs the ECAL energy, HCAL energy, GSF track quality, among other parameters, and is trained to identify electrons. If there is no associated GSF track the particle flow candidate is defined as a photon. The reconstruction efficiency for electrons is about 90% for $p_T > 20$ GeV. [49]

In the 2012 analyses at $\sqrt{s} = 8$ TeV, a cut based ID is used [49]. The cuts used differ between the calorimeter barrel ($|\eta| < 1.479$) and the endcap ($1.479 < |\eta| < 2.5$). In the barrel, the extension of the shower in the η direction ($\sigma_{\eta\eta}$) is required to be less than 0.01. The ratio of the energy deposits (H/E) between the HCAL (section 3.2.4) and the ECAL (section 3.2.3) must be less than 0.12. The distances from the vetex, d_0 and d_Z (section 3.2) must be less than 0.02 cm and 0.2 cm respectively. The supercluster energy E_{SC} and the track momentum at the closest point to the vertex (p) are also used as discriminators, with the requirement that $1/E_{SC} - 1/p < 0.05$. In

the endcap the $\sigma_{\eta\eta}$ cut is loosened to 0.03 and the H/E cut is tightened to 0.10.

In the 2015 analyses at $\sqrt{s} = 13$ TeV, a boosted decision tree is used to optimize the electron selection. MVA discriminators are used to require an electron efficiency of 90%.

Electrons are required to be isolated from additional particles in the event. The p_T of charged hadrons, neutral hadrons, and photons is summed within a cone of radius $\Delta R < 0.3$ around the electron. Corrections are applied for additional particles not associated with the primary vertex (pileup). The particle flow isolation is defined as $I_{PF}^e = \left(\sum^{Ch.had} p_T + max(0, \sum^{N.had.} p_T + \sum^{\gamma} p_T - 0.5 \sum^{PU} p_T \right) / p_T^e$. The factor of 0.5 corresponds to an approximate average ratio of neutral to charged particles, as measured in jets [1]. Alternatively, the effective area method can be used to estimate the pileup. The contribution from the pileup, p_T^{PU}, is defined as $p_T^{PU} = \rho A_{eff}$. Here ρ is the average energy density in the event, defined as the median of the energy density distribution for particles within the area of any jet (section 5.4.1) in the event.[49] FASTJET [23] is used to determine A_{eff}, the effective area of of the clustered jet.

5.3   Muons

Muons are highly penetrating particles [58] that deposit a minimal amount of energy in the ECAL and HCAL before depositing the remainder of their energies in the muon system. The muon system is discussed further in Section 3.2.5.

Muons are identified in three different ways. Standalone muons are identified based on hits in the DT, CSC, and RPC elements of the muon system. Hits in the DT's and CSC's are used as a seeds for a KF fit of the DT, CSC, and RPC hits. The resulting track is defined as a standalone muon. [29] If a standalone muon track is matched to a tracker track, then the two tracks are combined using the KF technique to create a global muon track. At high $p_T > 200$ GeV global muons have improved momentum resolution compared to standalone muons.

Tracker muons are constructed by extrapolating all tracker tracks with $p_T > 0.5$ GeV and $p > 2.5$ GeV to the muon system. If the track matches a small segment of DT and CSC hits, the track is considered to be a tracker muon. Tracker muons improve efficiency for muons with low $p_T < 5$ GeV because low energy muons won't register enough hits in the muon system to be identified as standalone or global muons.

The muon algorithms are very efficient, with approximately 99% of muons successfully identified in the tracker or the muon system. [29] To reduce the fake rate due to hadrons punching

through the calorimeters, global muons are required to be isolated, with no more than 10% of their energy measured by the calorimeters or the tracker inside a cone of $\Delta R < 0.3$. Additionally, charged hadrons may be reclassified as muons if their is a large discrepancy between their track p_T and the sum of their particle flow linked calorimeter deposits.

Additional tight identification criteria are applied to muons used in the analyses in sections 6 and 7 Global muons are required to be within $d_{xy} < 0.2$ cm and $d_z < 0.5$ cm of the primary vertex. Muons are required to have recorded at least one hit in the pixel system (section 3.2.2) and hits in at least six tracker layers. A particle flow based isolation requirement is defined for muons, analogous to the definition for electrons in section 5.2. $I_{PF}^{\mu} = \left(\sum^{Ch.had} p_T + max(0, \sum^{N.had.} p_T + \sum^{e,\gamma} p_T - 0.5 \sum^{PU} p_T \right) / p_T^{\mu}$. The isolation is computed in a radius of $\Delta R < 0.4$.

5.4 Hadrons

Hadrons are identified after the unique signatures of electrons, photons, and muons are already identified and removed from the list of remaining particle flow candidates. Neutral hadrons will only deposit energy in the calorimeters, while charged hadrons have associated tracks. For charged hadrons, the uncertainty in p_T returned by the track fit must be smaller than the calorimeter energy resolution for charged hadrons, described in Sections 3.2.3 and 3.2.4. The hadronic deposits are used to reconstruct jets and hadronic tau decays.

5.4.1 Jets

After hadronic deposits are reconstructed via particle flow, the anti-k_t [22] algorithm is used to cluster the deposits and define jets. A distance metric between hadronic particles i and j is defined: $d_{ij} = min(\frac{1}{k_{ti}^2}, \frac{1}{k_{tj}^2})\frac{\Delta_{ij}^2}{R^2}$. Here k_t is the transverse momentum and $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)$, where y and ϕ are the rapidity and azimuth of the particle, as defined in Chapter 2. At CMS, a radius of $R = 0.5(0.4)$ is used at 8 TeV (13 TeV). The radius is chosen to mimize the effect of pileup on the energy of the clustered jet. The distance between particle i and the beam is defined as $d_{iB} = \frac{1}{k_{ti}^2}$.

Beginning from a particle j, neighboring particles i are combined while $d_{ij} < d_{ib}$. Otherwise, i is considered the seed for a different jet. The results of this algorithm are show in Figure 5.2. The anti-k_t algorithm is infrared safe and collinear safe [22] which means that the jet clustering process is not affected significantly by low energy soft radiation. The anti-k_t algorithm is applied to jets at

CMS via the FASTJET framework[23].
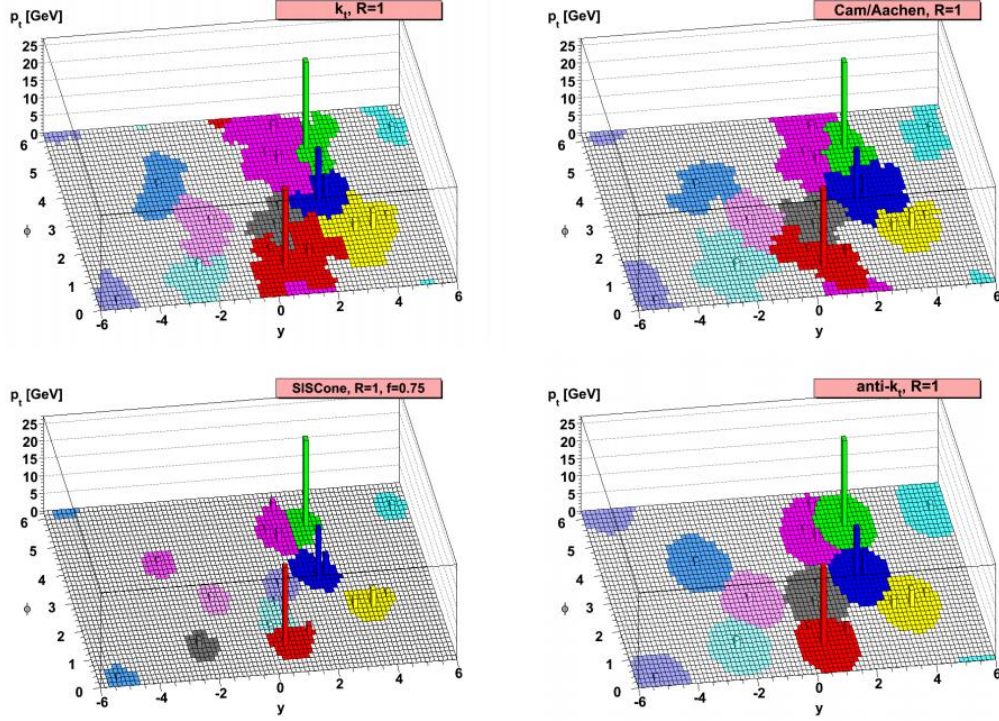
Figure 5.2: Performance of the anti-k_t algorithm. [22] Note that the anti-k_t algorithm is the only algorithm that generates a circular hard jet.

Charged hadron subtraction is used to reduce the dependence of the jets on pileup (Section 2.1) The majority of jet energy is carried by charged hadrons. For example, at $p_T = 100$ GeV charged hadrons are responsible for 65% of the jet energy. Photons carry 25% of the energy and neutral hadrons carry 10% of the energy. The charged hadron constituents of the jets still have track information stored via the particle flow algorithm. Any charged hadrons associated with pileup vertices instead of the primary vertex are retroactively removed from the jet. [53]

Additionally, three levels of corrections are applied to the jets. L1 corrections use the average p_T density per unit area to subtract energy associated with a combination of the underlying event, electronic noise, and pileup [53]. L2L3 corrections use Monte Carlo to calibrate the differences between reconstructed jets and generated jets in bins of p_T and η.

Jets which originate from pileup rather than the hard scattering process are identified and removed. A boosted decision tree is trained using a Z+jets Monte Carlo sample produced with MADGRAPH and PYTHIA. Jet shower shape and vertex related variables are used as inputs to

identify reconstructed jets which are not matched to a generator level hard scattering process. [2]

Jets which originate from b quark decays are referred to as b-jets. B quarks produced in the hard scattering event can travel a measurable distance before decaying. B-jets are identified by finding jets which originate from a secondary vertex, displaced in d_{xy} from the primary vertex. Secondary vertices are required to share less than 65% of their associated tracks with the primary vertex and have at least a 3σ radial distance significance from the primary vertex. The Combined Secondary Vertex (CSV) algorithm is used to identify B-jets. It uses vertex and tracker related variables to identify b-jets and is trained with a $t\bar{t}$ sample. B-tagging efficiencies are found to be 85%. [28]

5.4.2   Taus

As discussed in Chapters 1 and 2, taus can decay leptonically to electrons or muons, or hadronically to quarks. Taus are identified separately depending on their decay products, so the identification of leptonic tau decays has already been discussed in Sections 5.3 and 5.4, which describe the identification of electrons and muons respectively. This section will discuss the identification of hadronic taus. As discussed in Chapter 2, taus can decay hadronically in three dominant ways: To a charged pion, to a charged pion and a neutral pion, or to three pions. These decays have branching fractions of 11.6%, 26.0%, and 9.8%, respectively.

If the tau decays exclusively to charged hadrons, the hadronic tau (τ_h) is identified by comparing the invariant mass of the hadrons to the invariant mass of the tau. For example, if the tau decays to three charged pions, the invariant mass of the decay products must be between 0.8 GeV and 1.5 GeV, and the tracks must originate within $\Delta z < 0.4$ cm of the same event vertex [32], where Δz is the distance in the x-y plane of CMS, as discussed in Section 3.2. Mass windows of 50-200 MeV and 0.3-1.3 GeV are used for charged pion decays and charged pion plus neutral pion decays, respectively.

If a neutral pion is involved in the hadronic tau decay, then the hadron plus strips (HPS) method must be used to identify the taus. Neutral pions convert to energetic photons in the tracker which then produce electron positron pairs. These particles will be bent in opposite directions in the magnetic field of CMS and will deposit their energy in the ECAL in a strip in ϕ. The maximal distance between hadronic tau candidates and strips is $\Delta R = 3.0/p_T^\tau$. The strip sizes are 0.05 $\times$ 0.20 in the (η,ϕ) plane. The size is elongated in ϕ because the e^+e^- pair will be bent in ϕ due to

the magnetic field. In 13 TeV, the strip sizes are allowed to vary in a dynamic way that depends on the p_T weighted average of all e/γ objects in the strip. The strip size is allowed to vary from (0.05,0.05) to (0.15,0.30) in the (η,ϕ) plane.If a charged hadron from particle flow is linked to a strip in ECAL, and the invariant mass of the system is compatible with the hadronic tau mass, then the object is identified as a tau. The probability to reconstruct the tau decay mode is 80%, as measured in $Z \to \tau\tau$ events.

The particle flow tau isolation, I_{PF}^{τ}, is computed by summing momenta of particle flow candidates within a cone of $\Delta R = 0.5$ (0.4 at 13 TeV) around the tau. $I_{PF}^{\tau} = \sum^{Ch.had} p_T + \sum^{N.had.} p_T + \sum^{\gamma} p_T - 0.4576 \sum^{PU} p_T$ The 0.4576 factor is chosen to make the ID independent of pileup. The tight working point requires $I_{PF}^{\tau} < 0.8$ GeV and has an ID efficiency of 50 % in 8 TeV data. [3]

5.5   Missing Energy

There are some particles, for example neutrinos or dark matter candidates, that CMS cannot detect. They can be measured by applying the principle of conservation of energy. The transverse vector sum of all particle flow candidates in the event is computed. If CMS detected every particle, this sum would be zero. Therefore, the energy that CMS did not detect can be quantified by reversing the vector sum of the particles that were detected. If the jet energy corrections are included in the vector sum, the transverse missing energy (MET) is said to be Type 1 corrected [50]

Chapter 6

Lepton Flavor Violating decays of the Higgs Boson

A search for a lepton flavor violating (LFV) decay of a Higgs boson is performed in two channels: $H \to \mu\tau_e$ and $H \to \mu\tau_h$, where τ_h and τ_e are taus reconstructed in the hadronic and electronic decay channels respectively. As mentioned in section 2.3, these channels have a similar signature as the Standard Model $H \to \tau_\mu\tau_{had}$ and $H \to \tau_{mu}\tau_e$ searches, with a significant kinematic difference due to the muon arising directly from the Higgs decay. All the neutrinos in the event will arise from the tau decay. Because the tau is heavily boosted, the neutrinos will be collinear with the visible tau decay products. The term "visible decay products" is used to denote the tau decay products that CMS can detect directly, which are electrons, muons, and hadrons. Neutrinos register as missing energy in the event and are detected indirectly as discussed in section 5.5. The missing transverse energy in the event is the sum of the energy of all unidentified particles, so the individual neutrino momenta cannot be measured directly. In order to calculate the invariant mass (section 2.4) of the final state objects the collinear approximation must be used.

The collinear approximation[38] arises from the observation that since the mass of the Higgs is much greater than the mass of the tau, the tau decay products are highly boosted in the direction of the original tau. Therefore, the neutrino momentum can be approximated to be in the same direction as the other visible decay products of the tau. The component of the missing transverse energy in the direction of the visible tau decay products is used to estimate the transverse component of the neutrino momentum:

$$\vec{p}_T^{\,\nu} = \vec{E}_T^{miss} \cdot \hat{p}_T^{\tau_{vis}} \tag{6.1}$$

The fraction of the tau momentum carried by the visible tau decay products, $x_{\tau_{vis}}$, is given by:

$$x_{\tau_{vis}} = \frac{|\vec{p}_T^{\tau_{vis}}|}{|\vec{p}_T^{\tau_{vis}}| + |\vec{p}_T^{\nu}|} \tag{6.2}$$

The tau four momentum is then $\frac{1}{x_{\tau_{vis}}}(|\vec{p}^{\tau_{vis}}|, \vec{p}^{\tau_{vis}})$. The mass of the Higgs boson is much greater than the mass of the tau or muon or electron, so the approximation $M_H \gg m_\tau^2, m_l^2$ is valid. This yields the collinear mass equation:

$$M_H = M_{collinear} = \frac{M_{vis}}{\sqrt{x_{\tau_{vis}}}} \tag{6.3}$$

The Higgs mass has now been defined in terms of the mass of the visible decay products, which can be directly measured by CMS. As shown in figure 6.1 the collinear mass gives an improved mass resolution over the visible mass, in addition to peaking at the expected value of 125 GeV.
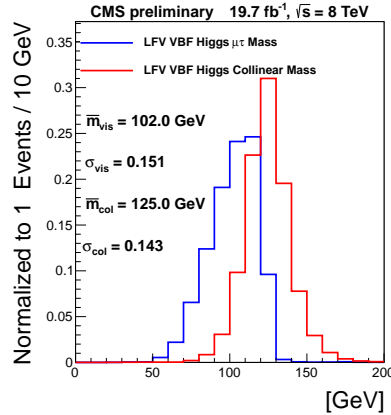
Figure 6.1: M_{vis} and $M_{collinear}$ shape comparisons after the pre-selection for the $H \to \mu\tau)h$ channel

6.1   8 TeV Backgrounds and Datasets

6.1.1   Data

The search is performed on the 2012, $\sqrt{s} = 8$ TeV dataset, comprising $19.7 fb^{-1}$ of data. The $H \to \mu\tau_{had}$ channel selection requires an isolated single muon trigger with a p_T^μ threshold of 24

GeV in the range $|\eta| < 2.1$, while the $H \rightarrow \mu\tau_e$ channel requires a muon-electron trigger with p_T threshold of 17 GeV ($|\eta| < 2.5$) and 8 GeV ($|\eta| < 2.4$), respectively.

6.1.2 LFV Higgs and SM Higgs

The signal decay $H \rightarrow \mu\tau$ is implemented via PYTHIA8. This version of PYTHIA8 contains decay mechanisms that account for tau polarization and spin, and the interface with Tauola is not needed[62]. The Standard Model $H \rightarrow \tau\tau$ background is modelled using PYTHIA 6 and Tauola. In addition to the generators, the full GEANT detector simulation is used for each Monte Carlo sample.

6.1.3 Z to tau tau

The $Z \rightarrow \tau\tau$ background is estimated using a particle flow embedding technique. A sample of $Z \rightarrow \mu\mu$ events is taken from data using a loose selection. The muons are then replaced with simulated tau decays reconstructed with the particle flow algorithm. The key features of the event topology, such as the jets, missing energy, and underlying event, are taken directly from data. Only the tau decays are simulated. The sample is normalized in accordance with the Monte Carlo expectation, using Drell Yan events generated with MADGRAPH and PYTHIA6.

6.1.4 Misidentified Leptons

Jets may be mistakenly identified as leptons. These misidentified leptons are referred to as "fakes." Misidentified leptons are produced by jets in W+Jets and QCD multi-jet events. The W+Jets background is particularly significant as the largest irreducible background in this analysis. If a jet is misidentified as a tau and the W boson decays muonically, then the kinematics of the event may be identical to the LFV signal: a muon, a tau, and missing energy opposite to the muon. This background is estimated using a data driven method. The same preselection cuts are used, but the isolation of one of the leptons is inverted, so that is it surrounded by other PF objects in the event. This enriches the contributions from W+jets and QCD multijet events and gives a larger sample of data statistics for estimating the background. The probability for particle flow objects to be misidentified as leptons is measured in an independant collision data set, and this probability is applied to the background enriched sample to compute the misidentified lepton background in the signal sample. This technique is shown schematically in Table 4 in which four regions are defined.

Regions I and III are the signal and background enriched regions, respectively, and Regions II and IV are the control regions, with the same selections as Regions I and III but with a same sign lepton requirement. The control regions are designed to maximize the number of fake lepton events while minimizing other backgrounds in order to directly test the accuracy of the fake rate method. This technique is illustrated in Table 6.1.

Table 6.1: Definition of the regions used to estimate the misidentified lepton background. The different regions have different requirements for the isolation and the relative charge of the two leptons $\ell_1^\pm$ and $\ell_2^\pm$, which can be e, μ or τ_h.

Opposite-sign leptons	Same-sign leptons
Region I	**Region II**
$\ell_1^\pm$(isolated)	$\ell_1^\pm$(isolated)
$\ell_2^\mp$(isolated)	$\ell_2^\pm$(isolated)
Region III	**Region IV**
$\ell_1^\pm$(isolated)	$\ell_1^\pm$(isolated)
$\ell_2^\mp$(non-isolated)	$\ell_2^\pm$(non-isolated)

In the $H \to \mu\tau_e$ channel, Region I is the signal region in which an isolated muon and an isolated electron are required. In Region III all the analysis selection criteria are applied to the data sample, with the exception that one of the leptons is required to be non-isolated. This creates a region enriched with misidentified leptons. There are two types of events in this region: those with an isolated muon and a non-isolated electron and those with an isolated electron and a non-isolated muon. There are a negligible number of signal events in Region III. Regions II and IV are defined using the same selection criteria as Region I and III, respectively, but same-sign leptons are required instead of opposite sign leptons. The sample in region III is dominated by W+jets and QCD multijets. The misidentified electron background in Region I is estimated by multiplying the event yield in Region III by a factor f_e, where f_e is the ratio of non-isolated to isolated electrons. It is computed in an independent data sample $Z \to \mu\mu + X$, which X is a particle flow object identified as an electron. Background sources of real leptons, such as diboson events, are subtracted from the $Z \to \mu\mu + X$ sample using Monte Carlo simulation. The misidentified muon background is computed in the same way. The technique is validated using the same-sign data from regions II and IV. In Figure 6.2 the observed data yield in Region II is compared to the estimate obtained by scaling the Region IV sample by the measured misidentification rates.

In the $H \to \mu\tau_h$ channel, the τ_h candidate can arise from a misidentified jet coming predominantly from W+jets and QCD multijet events. The misidentification rate f_{τ_h} is measured in
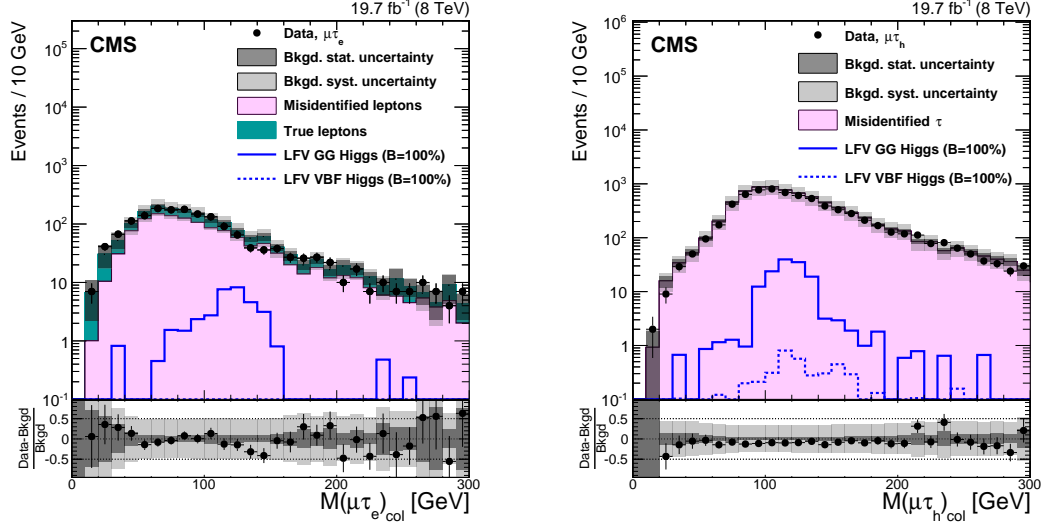
Figure 6.2: Distributions of M_{col} for region II compared to the estimate from scaling the region IV sample by the measured misidentification rates at $\sqrt{s} = 8$ TeV. The bottom panel in each plot shows the fractional difference between the observed data and the estimate. Left: $H \to \mu\tau_e$. Right: $H \to \mu\tau_h$.

$Z \to \mu\mu + X$ events selected in data, where X is an object identified as a τ_h candidate that passes a loose isolation requirement. The factor f_{τ_h} is defined as the fraction of these loosely isolated τ_h candidates that pass the tighter isolation requirement used to define the signal region. The misidentification rate measured in $Z \to \mu\mu + X$ collision data is compared to the measured rate in simulation and is found to be in agreement.

The enriched background regions (III and IV) are defined by requiring the presence of τ_h candidates that pass the looser isolation requirement, but do not pass the tight isolation requirement. The misidentified background yield in the signal region (Region I) is estimated by multiplying the event yield in Region III by a factor $\frac{f_{\tau_h}}{1-f_{\tau_h}}$. The procedure is validated with same-sign $\mu\tau_h$ events (Region II) using Region IV to model the misidentified background, in the same way as for the $H \to \mu\tau_e$ channel. The control region for the $H \to \mu\tau_h$ channel is shown in Figure 6.2.

6.1.5 Additional Backgrounds

The remaining backgrounds play a small role in the analysis and are all estimated with simulation. MADGRAPH5, PYTHIA6, and Tauola are used to simulate diboson events and events with one top quark and one antitop quark (ttbar events). POWHEG 1.0, PYTHIA6, and Tauola are used to simulated single top quark events.

6.2 13 TeV Backgrounds and Datasets

The search is performed on the 2015, $\sqrt{s} = 13$ TeV dataset, comprising $2.3 fb^{-1}$ of data. The trigger selection requires an isolated single muon with $p_T > 20$ GeV in the range $|\eta| < 2.4$.

The LFV and SM Higgs boson samples are generated using POWHEG 1.0, with CT10 parton distribution functions, interfaced to PYTHIA 8.212. The MADGRAPH 5.1.3.30 generator is used for Z+jets, $t\bar{t}$, and diboson production. POWHEG is used for single top production. The POWHEG and MADGRAPH generators are interfaced to PYTHIA for parton showering and hadronization. The misidentified lepton background is modeled using the fake rate method, as discussed in section 6.1.4. Comparisons between data and Monte Carlo in the Region II control region are show in Figure 6.3
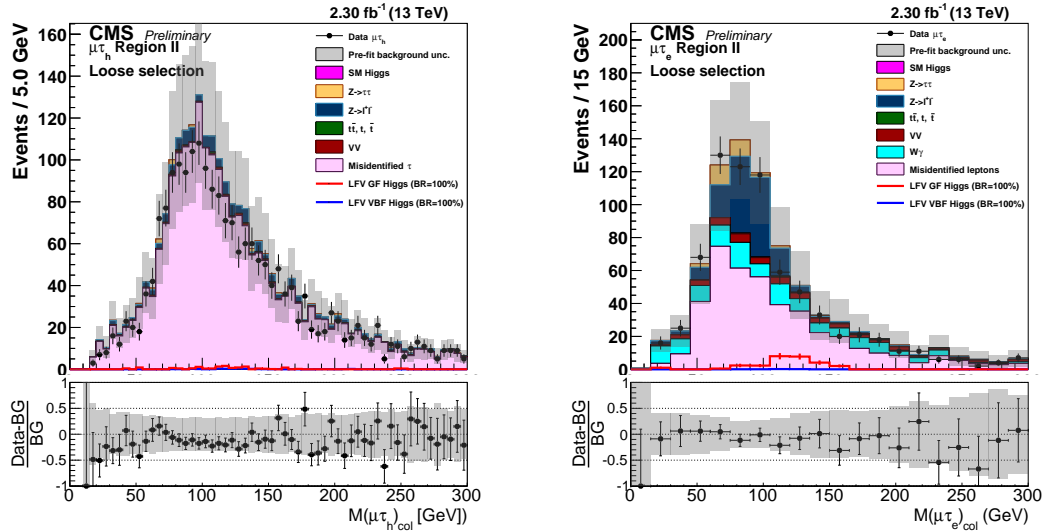
Figure 6.3: Distributions of M_{col} for region II compared to the estimate from scaling the region IV sample by the measured misidentification rates at $\sqrt{s} = 13$ TeV. The bottom panel in each plot shows the fractional difference between the observed data and the estimate. Left: $H \to \mu\tau_e$. Right: $H \to \mu\tau_h$.

6.3 Event Selection

The event selection consists of three steps. First, a loose preselection defines the basic signature. The events are then divided into three categories: 0 jet events, 1 jet events, and 2 jet events. Finally, cuts on relevant kinematic variables are optimized to suppress the backgrounds.

The preselection for the $H \to \mu\tau_e$ channel requires an isolated tight muon (section 5.3)

with $p_T > 25 GeV$, $|\eta| < 2.1$, and an isolated tight election (section 5.2) of opposite charge, with $p_T > 10 GeV$, $|\eta| < 2.3$. The isolation requirement for the muon is $I^\mu_{PF} < 0.12$ at 8 TeV and $I^\mu_{PF} < 0.15$ at 13 TeV. The isolation requirement for the electron is $I^e_{PF} < 0.12$ at 8 TeV, and $I^e_{PF} < 0.10$ at 13 TeV. The effective area corrections for pileup subtraction are applied at 13 TeV. The $H \to \mu\tau_h$ channel requires an isolated tight muon with $p_T > 30 GeV$, $|\eta| < 2.1$, and a tightly isolated hadronic tau (section 5.4.2) of opposite charge, with $p_T > 30$ GeV, $|\eta| < 2.3$. Events with additional leptons are vetoed.

The events are then divided into categories corresponding to the number of jets in the event. Jets are required to pass particle flow identification algorithms (section5.4.1), have $p_T > 30$ GeV and lie within $|\eta| < 4.7$. The zero jet category contains events primarily produced by gluon-gluon fusion. The one jet category is also dominated by gluon-gluon fusion events. The two jet category contains a significant number of vector boson fusion events. Events in the two jet category are required to have the jets separated by a rapidity gap ($\Delta\eta > 3.5$) and to have an invariant dijet mass greater than 550 GeV. In the $H \to \mu\tau_e$ channel events are vetoed if they contain a b-jet. Agreement between data and Monte Carlo simulation at preselection are show in figure 6.4 for $\sqrt{s} = 8$ TeV and in figure 6.5 for $\sqrt{s} = 13$ TeV.

Next, a set of kinematic variables are defined and the criteria for selection is determined by optimizing for $\frac{S}{\sqrt{S+B}}$ for the $\sqrt{s} = 8$ TeV dataset, where S and B are the expected signal and background event yields in the mass window 100 GeV ¡ $M_{collinear} < 150$ GeV. The signal strength is set as $B(H \to \mu\tau) = 10\%$. The 10% value was chosen because it corresponds to the limit from indirect measurments, as mentioned in section 1.2. The criteria for each category, and in each channel, are given in Table 6.2. The optimization was also performed assuming $B(H \to \mu\tau) = 1\%$ and a negligible change in the optimal selection criteria was observed. The variables used are the lepton transverse momenta p_T^ℓ with $\ell = \tau_h, \mu, e$, azimuthal angles between the leptons $\Delta\phi_{\vec{p}_T^{\ell_1} - \vec{p}_T^{\ell_2}}$, azimuthal angle $\Delta\phi_{\vec{p}_T^\ell - \vec{E}_T^{miss}}$, and the transverse mass $M_T^\ell = \sqrt{2p_T^\ell E_T^{miss}(1 - cos\Delta\phi_{\vec{p}_T^\ell - \vec{E}_T^{miss}})}$.

At $\sqrt{s} = 13$ TeV, the 8 TeV optimization is re-used. Due to the lower integrated luminosity the cuts in the 2 jet category are loosened. The dijet mass requirement is loosened from 550 to 200 GeV and the rapidity gap requirement is decreased from $\Delta\eta > 3.5$ to $\Delta\eta > 2.5$. The electron p_T requirement is increased from 10 to 15 GeV to suppress the misidentified electron background.
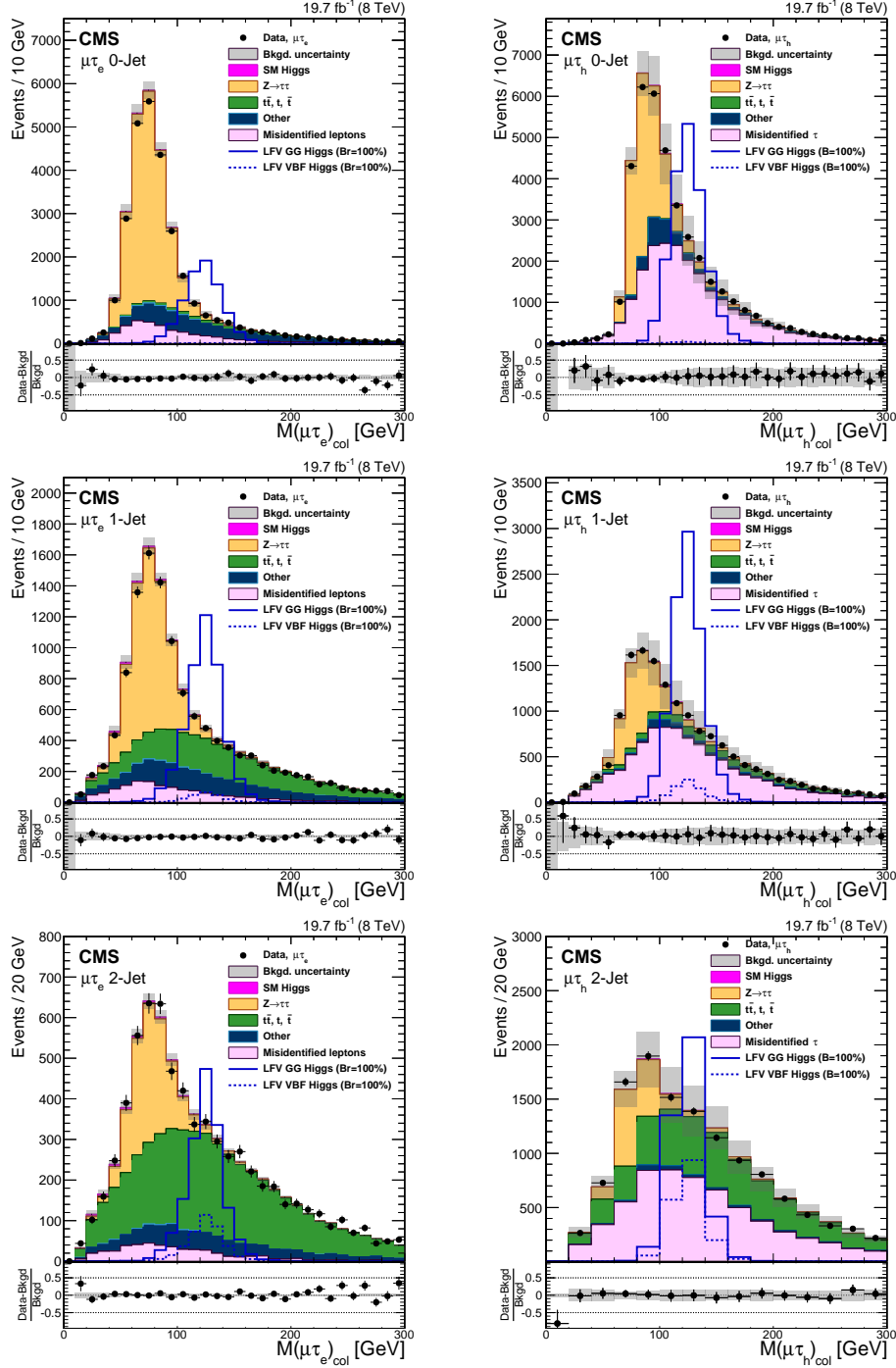
Figure 6.4: Distributions of the collinear mass $M_{collinear}$ at $\sqrt{s} = 8$ TeV, for signal with $B(H \to \mu\tau) = 100\%$ for clarity, and background processes after the loose selection requirements for the LFV $H \to \mu\tau$ candidates for the different channels and categories compared to data. The shaded grey bands indicate the total uncertainty. The bottom panel in each plot shows the fractional difference between the observed data and the total estimated background. Top left: $H \to \mu\tau_e$ 0-jet; top right: $H \to \mu\tau_h$ 0-jet; middle left: $H \to \mu\tau_e$ 1-jet; middle right: $H \to \mu\tau_h$ 1-jet; bottom left: $H \to \mu\tau_e$ 2-jet; bottom right $H \to \mu\tau_h$ 2-jet.

Figure 6.5: Distributions of the collinear mass $M_{collinear}$ at $\sqrt{s} = 8$ TeV, for signal with $B(H \to \mu\tau) = 100\%$ for clarity, and background processes after the loose selection requirements for the LFV $H \to \mu\tau$ candidates for the different channels and categories compared to data. The shaded grey bands indicate the total uncertainty. The bottom panel in each plot shows the fractional difference between the observed data and the total estimated background. Top left: $H \to \mu\tau_e$ 0-jet; top right: $H \to \mu\tau_h$ 0-jet; middle left: $H \to \mu\tau_e$ 1-jet; middle right: $H \to \mu\tau_h$ 1-jet; bottom left: $H \to \mu\tau_e$ 2-jet; bottom right $H \to \mu\tau_h$ 2-jet.

Table 6.2: Selection criteria for the kinematic variables after the loose selection.

Variable	$H \to \mu\tau_e$			$H \to \mu\tau_h$		
[GeV]	0-jet	1-jet	2-jet	0-jet	1-jet	2-jet
$p_T^{\mu} >$	50	45	25	45	35	30
$p_T^{e} >$	10	10	10	-	-	-
$p_T^{\tau_h} >$	-	-	-	35	40	40
$M_T^{e} <$	65	65	25	-	-	-
$M_T^{\mu} >$	50	40	15	-	-	-
$M_T^{\tau_h} <$	-	-	-	50	35	35
[radians]						
$\Delta\phi_{\vec{p}_T^{\mu} - \vec{p}_T^{\tau_h}} >$	-	-	-	2.7	-	-
$\Delta\phi_{\vec{p}_T^{e} - \vec{E}_T^{miss}} <$	0.5	0.5	0.3	-	-	-
$\Delta\phi_{\vec{p}_T^{e} - \vec{p}_T^{\mu}} >$	2.7	1.0	-	-	-	-

6.4 Systematic Uncertainties

6.4.1 8 TeV

To set upper bounds on the signal strength, or determine a signal significance, we use the CL_s method (section 6.5). A binned likelihood is used, based on the distributions of $M_{collinear}$ for the signal and background sources. Systematic uncertainties are represented by nuisance parameters, some of which only affect the background and signal normalizations, while others affect the shape and/or normalization of the M_{col} distributions.

6.4.1.1 Normalization Uncertainties

The uncertainties are summarized in Tables 6.3 and 6.4. The uncertainties in the e and μ selection efficiency (trigger, identification, and isolation) are estimated using the "tag and probe" technique in $Z \to ee, \mu\mu$ data [31]. The identification efficiency of hadronic tau decays is estimated using the "tag and probe" technique in $Z \to \tau\tau$ data [32]. The uncertainty in the $Z \to \tau\tau$ background comes predominantly from the uncertainty in the tau efficiency. The uncertainties in the estimation of the misidentified lepton rate come from the difference in rates measured in different data samples (QCD multijets and W+jets). The uncertainty in the production cross section of the backgrounds that have been estimated by simulation is also included.

There are several uncertainties on the H production cross section, which depend on the production mechanism contribution and the analysis category. They are given in Table 6.4. These affect the LFV H and the SM H background equally, and are treated as 100% correlated. The parton

Table 6.3: Systematic uncertainties in the expected event yield in %. All uncertainties are treated as correlated between the categories, except where there are two numbers. In this case the number denoted with * is treated as uncorrelated between categories and the total uncertainty is the sum in quadrature of the two numbers.

Systematic uncertainty	$H \to \mu\tau_e$			$H \to \mu\tau_h$		
	0-Jet	1-Jet	2-Jets	0-Jet	1-Jet	2-Jets
electron trigger/ID/isolation	3	3	3	NA	NA	NA
muon trigger/ID/isolation	2	2	2	2	2	2
hadronic tau efficiency	NA	NA	NA	9	9	9
luminosity	2.6	2.6	2.6	2.6	2.6	2.6
$Z \to \tau\tau$ background	3+3*	3+5*	3+10*	3+5*	3+5*	3+10*
$Z \to \mu\mu, ee$ background	30	30	30	30	30	30
misidentified μ, e background	40	40	40	NA	NA	NA
misidentified τ_h background	NA	NA	NA	30+10*	30	30
WW, ZZ+jets background	15	15	15	15	15	65
$t\bar{t}$ background	10	10	10+10*	10	10	10+33*
$W + \gamma$ background	100	100	100	NA	NA	NA
b-tagging veto	3	3	3	NA	NA	NA
single top production background	10	10	10	10	10	10

Table 6.4: Theoretical uncertainties in % for Higgs boson production. Anticorrelations arise due to migration of events between the categories and are expressed as negative numbers.

Systematic uncertainty	Gluon-Gluon Fusion			Vector Boson Fusion		
	0-Jets	1-Jets	2-Jets	0-Jet	1-Jet	2-Jets
parton distribution function	+9.7	+9.7	+9.7	+3.6	+3.6	+3.6
renormalization/factorization scale	+8	+10	−30	+4	+1.5	+2
underlying event/parton shower	+4	−5	−10	+10	<1	−1

distribution function (PDF) uncertainty is evaluated b comparing the yields in each category, when spanning the parameter range of a number of different independent PDF sets including CT10 [56], MSTW [55], NNPDF [16], as recommended by PDF4LHC [7]. The scale uncertainty is estimated by varying the renormalization and factorization scales, μ_R and μ_F, up and down by one half or two times the nominal scale ($\mu_R = \mu_F = M_H/2$) under the constraint $0.5 < \mu_F/\mu_R < 2$. [37]. The underlying event and parton shower uncertainty is estimated by using two different PYTHIA tunes, CUET2 and Z2*. Anticorrelations arise due to migration of events between the catories and are expressed as negative numbers

6.4.1.2 Shape Uncertainties

The systematic uncertainties that lead to a change in the shape of the $M_{collinear}$ distribution are summarized in Table 5. In the embedded $Z \to \tau\tau$ $M_{collinear}$ distribution, used to estimate the $Z \to \tau\tau$ background, a 1% shift has been observed with respect to $Z \to \tau\tau$ simulations by

comparing the means of both distributions. This only occurs in the $H \to \mu\tau_e$ channel. The $M_{collinear}$ distribution has been corrected for this effect and a 100% uncertainty on this shift is used as a systematic uncertainty. The jet energy scale has been studied extensively and a standard prescription for corrections ?? is used in all CMS analyses. The overall uncertainty is set using γ+jets events and the most significant uncertainty arises from the photon energy scale. A number of other uncertainties such as jet fragmentation modeling, single pion response, and uncertainties in the pileup corrections are also included. The jet energy scale uncertainties (3-7%) are applied as a function of p_T and η, including all correlations, to all jets in the event, and are also propagated to the missing energy. The unclustered energy scale comes from jets below 10 GeV and PF candidates not within jets. It also has its associated uncertainty which is propagated to the missing transverse energy. The tau energy scale is estimated by comparing $Z \to \tau\tau$ events in data and simulation. An uncertainty of 3% is derived from this comparison. The uncertainty is applied by shifting the p_T of the tau candidates in the event. Finally, the $M_{collinear}$ distributions used in the fit have a statistical uncertainty in each mass bin. This uncertainty is uncorrelated between the bins.

Uncertainties in the shape of the misidentified lepton backgrounds have also been calculated. In the $H \to \mu\tau_e$ channel the misidentified lepton ratios f_μ, f_e are measured and applied in bins of lepton p_T and $|\eta|$. These ratios are all adjusted up or down by one standard deviation and the differences in the shape of the resultant M_{col} distributions are then used as nuisance parameters in the fit. In the $H \to \mu\tau_h$ channel the tau misidentification rate f_τ is found to be approximately flat in p_T and η. To estimate the systematic uncertainty the p_T distribution of f_τ is fit with a linear function and the rate recomputed from the fitted slope and intercept.

6.4.2   13 TeV

The systematic uncertainties used in the 13 TeV analysis are similar to those used in the 8 TeV analysis. They are show in Table 6.5. The 8 TeV analysis was carried out after many standard model studies had been completed, but the same is not true for the 13 TeV analysis. Therefore, a greater uncertainty is applied to the $Z \to \tau\tau$ and the $t\bar{t}$ backgrounds. The understanding of the $W + \gamma$ sample has increased significantly. The dominant theoretical uncertainties are uncertainties in the GGF and VBF Higgs production cross section, which are determined by varying μ_R, μ_F, and PDF sets. The overall theoretical uncertainty is approximated to 10%.

Table 6.5: Systematic uncertainties in the expected event yield in %. All uncertainties are treated as correlated between the categories, except where there are two values. In this case the first value is correlated as above, while the second value (following ⊕) represents an uncorrelated uncertainty for each individual category. The total uncertainty in a given category is the sum in quadrature of the two values.

Systematic uncertainty	$H \to \mu\tau_e$	$H \to \mu\tau_h$
Muon trigger/ID/isolation	3%	3%
Electron trigger/ID/isolation	3%	NA
Hadronic tau efficiency	NA	10%
b-tagging veto	3%	NA
$Z \to \tau\tau$ background	10%⊕5%	10%⊕5%
$Z \to \mu\mu, ee$ background	10%⊕5%	10%⊕5%
Misidentified μ, e background	40%⊕10%	NA
Misidentified τ_h background	NA	30%⊕10%
WW, ZZ background	10%⊕5%	10%⊕5%
$t\bar{t}$ background	20%⊕5%	20%⊕5%
$W + \gamma$ background	10%⊕5%	NA
Single top production background	10%	10%
Jet energy scale	3-20%	3-20%
Hadronic tau energy scale	NA	3%
Misidentified lepton shape	$\pm\sigma$	$\pm\sigma$
Theory uncertainty	10%	10%
Luminosity	2.7%	2.7%

6.5 Statistical Methods

6.5.1 Maximum Likelihood Fit

After selecting events which pass the signal region selections (section 6.3) and defining sources of systematic uncertainty (section 6.4), the next step is to fit the backgrounds to the data. This is done via a binned maximum likelihood fit.[20][33] We assume that the number of events in each bin follows a Poisson distribution, defined as $P(n_i|\mu_i) = \frac{\mu_i^{n_i} e^{-\mu_i}}{n_i!}$, where n_i is the number of events in the ith bin and μ_i is the number of Monte Carlo events in each bin.

We then define the likelihood L as $L = \prod P(n_i|\mu_i)$. The goal of the fit is to determine the μ_i that maximizes the likelihood. This procedure is complicated by the addition of the systematic uncertainties, which are treated as nuisance parameters. For example, the luminosity uncertainty introduces a 2.3% uncertainty. To account for this, each term in the likelihood expression is multiplied by a log-normal distribution. The log-normal distribution is used instead of the Gaussian distribution to prevent the parameter from becoming negative. The lognormal distribution is defined: $N(\mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\log x - \mu)^2}{2\sigma^2}}$. Shape uncertainties, like the jet energy scale and tau energy

scale systematics, are represented by Gaussian distributions. For the jth sample in the ith bin we can write: $\epsilon_{ji} = \epsilon_{ji}^0 + f\frac{\epsilon_{ji}^+ - \epsilon_{ji}^-}{2}$. Here, ϵ_{ji}^0 is the efficiency before the shift, ϵ_{ji}^+ is the efficiency after the scale shift up, and ϵ_{ji}^- is the efficiency after the scale shift down. The morphing parameter f has a gaussian distribution. We can interpolate quadratically for $|f| < 1$ and write $\epsilon_{ji} = \frac{f(f-1)}{2}\epsilon_{ji}^- - (f-1)(f+1)\epsilon_{ji}^0 + \frac{f(f+1)}{2}\epsilon_{ji}^+$ This term is then added as an additional factor in the maximum likelihood formula.

The postfit histograms are shown in Figure 6.6.

6.5.2 Maximum Likelihood Limits

After computing the expected LFV signal branching ratios, the next step is to determine the statistical significance of the result. The profile likelihood $\lambda(\mu)$ is defined as $\lambda(\mu) = \frac{L(\mu,\hat{\hat{\theta}})}{L(\hat{\mu},\hat{\theta})}$ where μ is the hypothesized signal strength defined by $\mu = 0$ as the background hypothesis and $\mu = 1$ the signal hypothesis. [34] The nuisance parameters are represented by θ. The denominator is the maximum value of the likelihood function, where $\hat{\mu}$ and $\hat{\theta}$ take their values that maximize the likelihood. The numerator is a likelihood as a function of μ, where $\hat{\hat{\theta}}$ maximizes the likelihood for a given μ. The profile likelihood can range between 0 and 1. We define our test statistic as $t_\mu = -2\ln\lambda\mu$. The probability distribution of this test statistic is given by $t_\mu = \frac{\mu-\hat{\mu}}{\sigma^2} + \mathcal{O}(\frac{1}{\sqrt{N}})$. [64] Here σ is related to the variance of all the nuisance parameters. We can make the asymptotic approximation that the second term goes to zero as $N \to \infty$. This is known as the asymptotic approximation.

Note that the test statistic is zero for total agreement between the hypothesis and the data and decreases as the data differs more and more from the hypothesis. The probability to observe a given hypothesis μ is defined by $p_\mu = \int_{\mu_{obs}}^\infty -2ln\lambda(\mu)$. The p-value is defined as $p_{\mu=0}$ and is the probability to exclude the background hypothesis. If $p < \alpha$ then we can consider the background hypothesis to be excluded. In high energy physics we define α as 2.87×10^{-7}. It is customary to express the in terms of standard deviations from the mean, assuming a Gaussian probability distribution function (PDF). For a p-value of 2.87×10^{-7} the significance is 5σ.

It is customary to use 95% confidence intervals when setting limits. If $p_\mu < 5\%$ then the hypothesis where the signal strength is μ can be excluded at 95% confidence. However, if the background is very small or fluctuates downward, it is possible to exclude a signal that the analysis is not sensitive to. This can be avoided by using the CL_s method. [61] The p-value for the CL_s method is defined by dividing the p-value for the signal hypothesis by the probability for excluding

the background, which is $(1 - p_{\mu=0})$.

It is useful to compare the observed limit with the expected limit. The expected limit is determined by generating an Asimov dataset [34] which is consistent with the background hypothesis. By definition, the Asimov dataset will produce postfit nuisance parameters that are identical to their expected values. The expected limit is a useful statistic that measures the exclusion power of an analysis. For example, if the expected limit is significantly greater than the signal hypothesis, then the analysis does not have the statistical power to investigate the signal region.

6.6   Results

6.6.1   8 TeV Results

The M_{col} distributions after the fit for signal and background contributions are shown in Fig. 6.6 and the event yields in the mass range $100 < M_{\mathrm{col}} < 150 GeV$ are shown in Table 6.6. The different channels and categories are combined to set a 95% CL_s upper limit on the branching fraction of LFV Higgs decay in the $\mu\tau$ channel, $B(H \to \mu\tau)$.

Table 6.6: Event yields in the signal region, $100 < M_{\mathrm{col}} < 150 GeV$ after fitting for signal and background. The expected contributions are normalized to an integrated luminosity of 19.7 fb^{-1}. The LFV Higgs boson signal is the expected yield for $B(H \to \mu\tau) = 0.84\%$ with the SM Higgs boson cross section.

Sample	$H \to \mu\tau_h$			$H \to \mu\tau_e$		
	0-Jet	1-Jet	2-Jets	0-Jet	1-Jet	2-Jets
misidentified leptons	1770 ± 530	377 ± 114	1.8 ± 1.0	42 ± 17	16 ± 7	1.1 ± 0.7
$Z \to \tau\tau$	187 ± 10	59 ± 4	0.4 ± 0.2	65 ± 3	39 ± 2	1.3 ± 0.2
ZZ, WW	46 ± 8	15 ± 3	0.2 ± 0.2	41 ± 7	22 ± 4	0.7 ± 0.2
$W\gamma$	NA	NA	NA	2 ± 2	2 ± 2	NA
$Z \to ee$ or $\mu\mu$	110 ± 23	20 ± 7	0.1 ± 0.1	1.6 ± 0.7	1.8 ± 0.8	NA
$t\bar{t}$	2.2 ± 0.6	24 ± 3	0.9 ± 0.5	4.8 ± 0.7	30 ± 3	1.8 ± 0.4
$t\bar{t}$	2.2 ± 1.1	13 ± 3	0.5 ± 0.5	1.9 ± 0.2	6.8 ± 0.8	0.2 ± 0.1
SM H background	7.1 ± 1.3	5.3 ± 0.8	1.6 ± 0.5	1.9 ± 0.3	1.6 ± 0.2	0.6 ± 0.1
sum of backgrounds	2125 ± 530	513 ± 114	5.4 ± 1.4	160 ± 19	118 ± 9	5.6 ± 0.9
LFV Higgs boson signal	66 ± 18	30 ± 8	2.9 ± 1.1	23 ± 6	13 ± 3	1.2 ± 0.3
data	2147	511	10	180	128	6

The observed and the median expected 95% CL_s upper limits on the $B(H \to \mu\tau)$ for the H mass at 125 GeV are given for each category in Table 6.7. Combining all the channels, an expected upper limit of $B(H \to \mu\tau) < (0.75 \pm 0.38)\%$ is obtained. The observed upper limit is $B(H \to \mu\tau) < 1.51\%$ which is above the expected limit due to an excess of the observed number of events above the
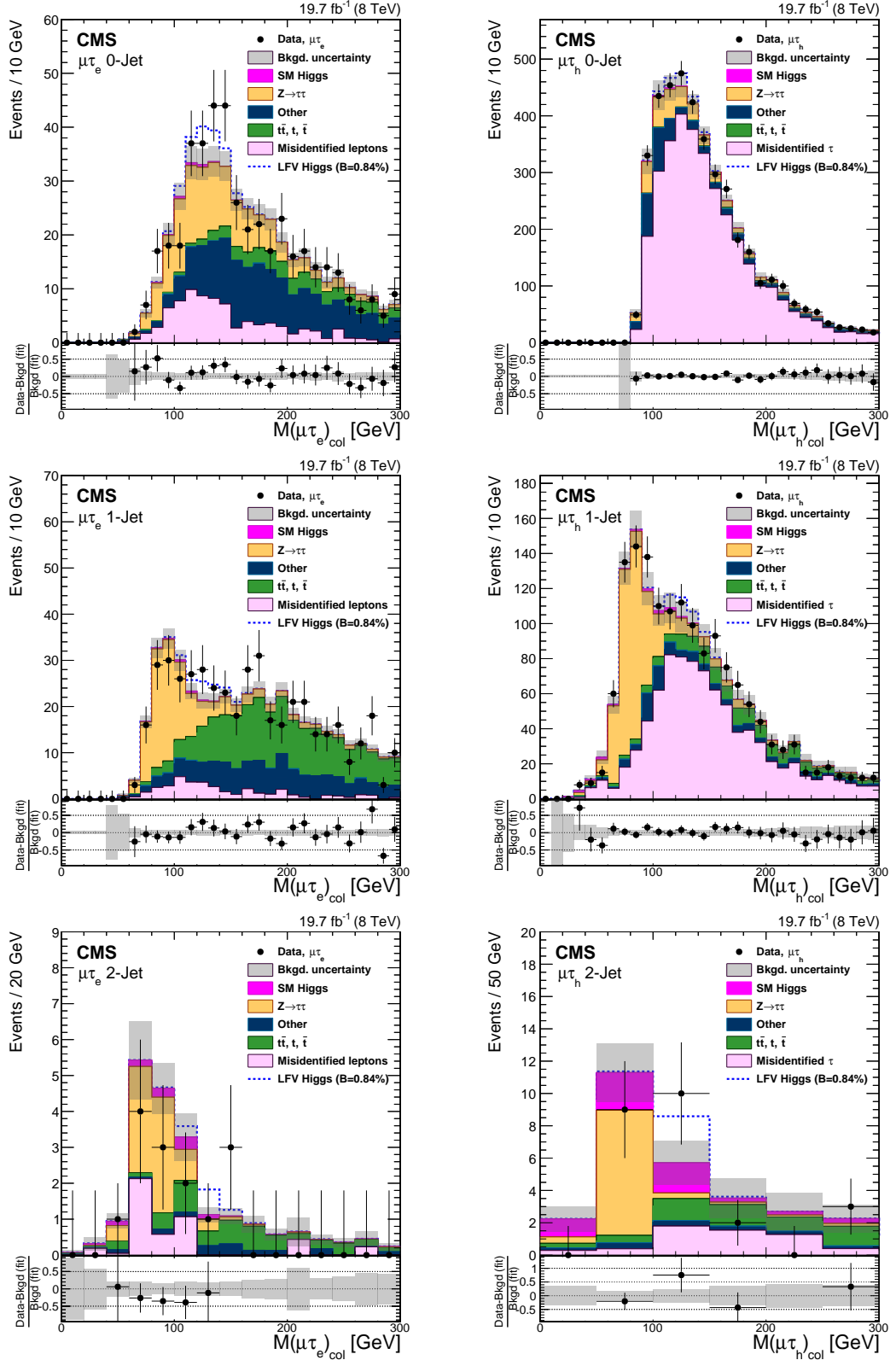
Figure 6.6: Distributions of the collinear mass M_{col} after fitting for signal and background for the LFV $H \to \mu\tau$ candidates in the different channels and categories compared to data. The distribution of the simulated LFV Higgs boson sample is shown for the best fit branching fraction of $B(H \to \mu\tau) = 0.84\%$. The bottom panel in each plot shows the fractional difference between the observed data and the fitted background. Top left: $H \to \mu\tau_e$ 0-jet; top right: $H \to \mu\tau_h$ 0-jet; middle left: $H \to \mu\tau_e$ 1-jet; middle right: $H \to \mu\tau_h$ 1-jet; bottom left: $H \to \mu\tau_e$ 2-jet; bottom right

background prediction. The fit can then be used to estimate the branching fraction if this excess were to be interpreted as a signal. The best fit values for the branching fractions are given in Table 6.7. The limits and best fit branching fractions are also summarized graphically in Fig. 6.7. The combined categories give a best fit of $B(H \to \mu\tau) = (0.84^{+0.39}_{-0.37})\%$. The combined excess is 2.4 standard deviations which corresponds to a p-value of 0.010 at $M_H = 125$ GeV.

Table 6.7: The expected upper limits, observed upper limits and best fit values for the branching fractions for different jet categories for the $H \to \mu\tau$ process. The one standard-deviation probability intervals around the expected limits are shown in parentheses.

	0-Jet (%)	1-Jet (%)	2-Jets (%)
	Expected Limits		
$\mu\tau_e$	<1.32 ($\pm$0.67)	<1.66 ($\pm$0.85)	<3.77 ($\pm$1.92)
$\mu\tau_h$	<2.34 ($\pm$1.19)	<2.07 ($\pm$1.06)	<2.31 ($\pm$1.18)
$\mu\tau$	<0.75 ($\pm$0.38)		
	Observed Limits		
$\mu\tau_e$	<2.04	<2.38	<3.84
$\mu\tau_h$	<2.61	<2.22	<3.68
$\mu\tau$	<1.51		
	Best Fit Branching Fractions		
$\mu\tau_e$	$0.87^{+0.66}_{-0.62}$	$0.81^{+0.85}_{-0.78}$	$0.05^{+1.58}_{-0.97}$
$\mu\tau_h$	$0.41^{+1.20}_{-1.22}$	$0.21^{+1.03}_{-1.09}$	$1.48^{+1.16}_{-0.93}$
$\mu\tau$	$0.84^{+0.39}_{-0.37}$		
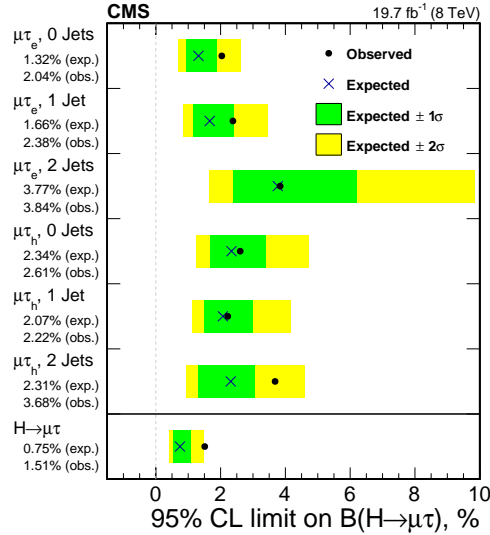
Figure 6.7: 95% CL Upper limits by category for the LFV $H \to \mu\tau$ decays.

6.6.1.1 Limits on lepton-flavour-violating couplings
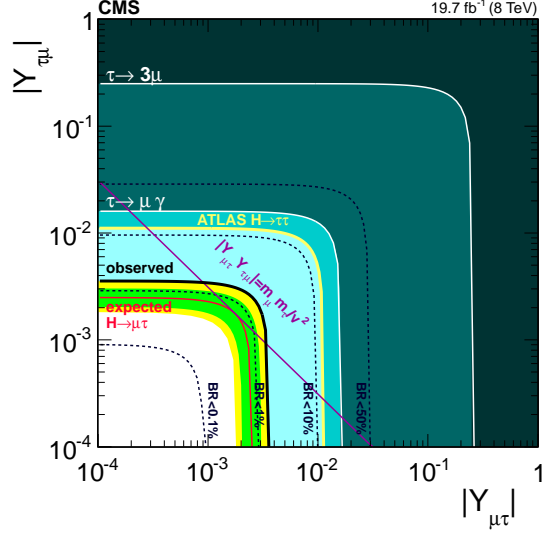
Figure 6.8: [51]Constraints on the flavour-violating Yukawa couplings, $|Y_{\mu\tau}|$ and $|Y_{\tau\mu}|$. The black dashed lines are contours of $B(H \to \mu\tau)$ for reference. The expected limit (red solid line) with one sigma (green) and two sigma (yellow) bands, and observed limit (black solid line) are derived from the limit on $B(H \to \mu\tau)$ from the present analysis. The shaded regions are derived constraints from null searches for $\tau \to 3\mu$ (dark green) and $\tau \to \mu\gamma$ (lighter green). The yellow line is the limit from a theoretical reinterpretation of an ATLAS $H \to \tau\tau$ search [46]. The light blue region indicates the additional parameter space excluded by our result. The purple diagonal line is the theoretical naturalness limit $Y_{ij}Y_{ji} \leq m_i m_j / v^2$.

The constraint on $B(H \to \mu\tau)$ can be interpreted in terms of LFV Yukawa couplings [46]. The LFV decays $H \to e\mu$, $e\tau$, $\mu\tau$ arise at tree level from the assumed flavour-violating Yukawa interactions, $Y_{\ell^\alpha \ell^\beta}$ where ℓ^α, ℓ^β denote the leptons, $\ell^\alpha, \ell^\beta = e, \mu, \tau$ and $\ell^\alpha \neq \ell^\beta$. The decay width $\Gamma(H \to \ell^\alpha \ell^\beta)$ in terms of the Yukawa couplings is given by:

$$\Gamma(H \to \ell^\alpha \ell^\beta) = \frac{m_H}{8\pi}\left(|Y_{\ell^\beta \ell^\alpha}|^2 + |Y_{\ell^\alpha \ell^\beta}|^2\right),$$

and the branching fraction by:

$$B(H \to \ell^\alpha \ell^\beta) = \frac{\Gamma(H \to \ell^\alpha \ell^\beta)}{\Gamma(H \to \ell^\alpha \ell^\beta) + \Gamma_{SM}}.$$

The SM H decay width is assumed to be $\Gamma_{SM} = 4.1\text{MeV}$ [36] for M_H=125GeV. The 95% CL constraint on the Yukawa couplings derived from $B(H \to \mu\tau) < 1.51\%$ and the expression for the branching fraction above is:

$$\sqrt{|Y_{\mu\tau}|^2 + |Y_{\tau\mu}|^2} < 3.6 \times 10^{-3}.$$

Figure ?? compares this result to the constraints from previous indirect measurements.

6.6.2 13 TeV Results

After applying the full selection cuts, a maximum likelihood fit is performed in the M_{col} variable. Each systematic uncertainty is used as a nuisance parameter in the fit. The distributions of the signal and background contributions after the full selection and the fit are shown in Fig. 6.9 and the event yields in the mass range $100 < M_{\mathrm{col}} < 150 GeV$ are shown in Table 6.8. The different channels and categories are combined to set a 95% CL upper limit on the branching fraction of LFV H decay in the $\mu\tau$ channel, $V(H \to \mu\tau)$.

Table 6.8: Event yields in the signal region in the range $100 < M_{\mathrm{col}} < 150 GeV$. The expected contributions are normalized to an integrated luminosity of $2.3 fb^{-1}$. The LFV Higgs boson signal is the expected yield for $B(H \to \mu\tau) = 1\%$ with the SM Higgs boson cross section.

Sample	$H \to \mu\tau_e$			$H \to \mu\tau_h$		
	0-Jet	1-Jet	2-Jets	0-Jet	1-Jet	2-Jets
misidentified leptons	12.2	5.2	2.8	232.3	54.7	4.7
$Z \to \tau\tau$	14.4	10.6	1.7	5.3	2.3	0
ZZ, WW	10.7	4.6	3.2	3.2	2.0	0.3
$W\gamma$	1.2	3.4	0.9	NA	NA	NA
$Z \to ee$ or $\mu\mu$	1.9	2.2	0.3	79.1	11.9	0.1
$t\bar{t}$	1.4	21.8	18.6	1.3	5.4	1.1
t, $\bar{t}$	0.4	4.1	1.7	0.3	2.2	0.2
SM H background	0.4	0.4	0.4	1.1	0.7	0.3
sum of backgrounds	42.6	52.2	29.6	322.5	79.3	6.6
LFV Higgs boson signal	7.1	3.7	1.9	13.8	4.7	1.2
Observed data	33	41	31	315	77	7

6.6.3 Limit computation

The observed and median expected 95% CL upper limits on the $B(H \to \mu\tau)$ for the H mass at 125GeV are given for each category in Table 6.9. Combining all the channels, an expected upper limit of $B(H \to \mu\tau) < (1.62 \pm 0.58)\%$ is obtained. The observed upper limit is $B(H \to \mu\tau) < 1.20\%$. The limits are also summarized graphically in Fig. 6.7.

This observed limit on the branching ratio is slightly tighter than the $B(H \to \mu\tau) < (1.51 \pm 0.83)\%$ limit obtained using the 19.7 fb^{-1} data sample at 8 TeV analyzed in [51]. The 95% CL constraint on the Yukawa couplings derived from $B(H \to \mu\tau) < 1.20\%$ and the expression for the branching fraction above is:

$$\sqrt{|Y_{\mu\tau}|^2 + |Y_{\tau\mu}|^2} < 3.16 \times 10^{-3}.$$
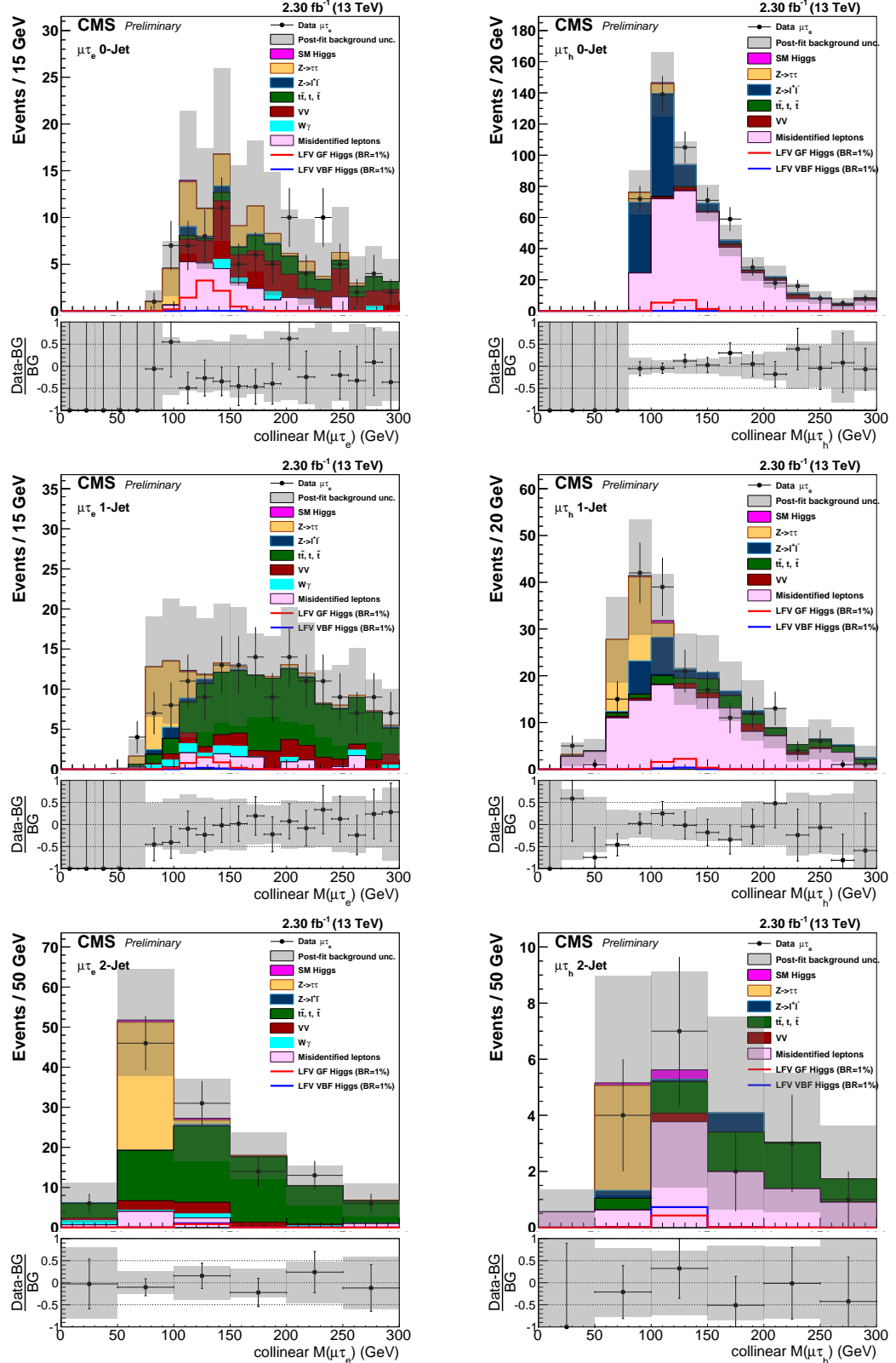
Figure 6.9: Distributions of the collinear mass M_{col} after fitting for signal and background at $\sqrt{s} = 13$ TeV for the LFV $H \to \mu\tau$ candidates in the different channels and categories compared to data. The distribution of the simulated LFV Higgs boson sample is shown for $B(H \to \mu\tau) = 10\%$. The bottom panel in each plot shows the fractional difference between the observed data and the fitted background. Top left: $H \to \mu\tau_e$ 0-jet; top right: $H \to \mu\tau_h$ 0-jet; middle left: $H \to \mu\tau_e$ 1-jet; middle right: $H \to \mu\tau_h$ 1-jet; bottom left: $H \to \mu\tau_e$ 2-jet; bottom right $H \to \mu\tau_h$ 2-jet.

Table 6.9: The observed and expected upper limits for different jet categories for the $H \to \mu\tau$ process. The one standard deviation probability intervals around the expected limits are shown in parentheses.

	Expected limits			
	0-jet (%)	1-jet (%)	2-jets (%)	Combined (%)
$\mu\tau_h$	<4.17	<4.89	<6.41	<2.98
$\mu\tau_e$	<2.24	<4.36	<7.31	<1.96
$\mu\tau$	<1.62 %			
	Observed limits			
	0-jet (%)	1-jet (%)	2-jets (%)	Combined (%)
$\mu\tau_h$	<4.24	<6.35	<7.71	<3.81
$\mu\tau_e$	<1.33	<3.04	<8.99	<1.15
$\mu\tau$	<1.20 %			
	Best fit branching fractions			
	0-jet (%)	1-jet (%)	2-jets (%)	Combined (%)
$\mu\tau_h$	$0.12^{+2.02}_{-1.91}$	$1.70^{+2.41}_{-2.52}$	$1.54^{+3.12}_{-2.71}$	$1.12^{+1.45}_{-1.40}$
$\mu\tau_e$	$-2.11^{+1.30}_{-1.89}$	$-2.18^{+1.99}_{-2.05}$	$2.04^{+2.96}_{-3.31}$	$-1.81^{+1.07}_{-1.32}$
$\mu\tau$	$-0.76^{+0.81}_{-0.84}$%			
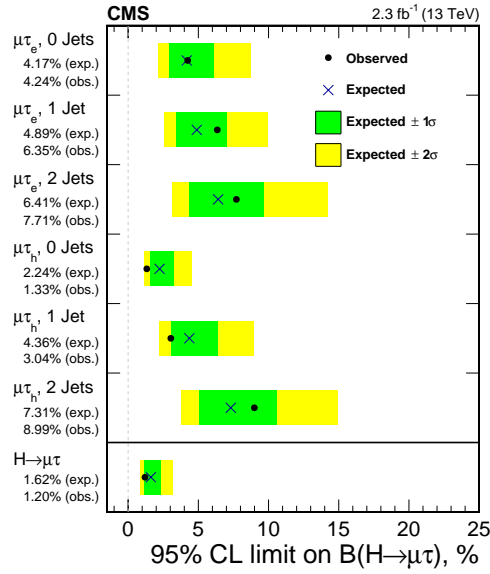
Figure 6.10: 95% CL Upper limits by category for the LFV $H \to \mu\tau$ decays.

Chapter 7

W+Jets Differential Cross Section Measurement

7.1 Backgrounds and Datasets

7.1.1 Data

This measurement utilizes data collected at $\sqrt{s} = 13$ TeV in 2015 at the LHC. The trigger selection requires an isolated single muon with $p_T > 20$ GeV in the range $|\eta| < 2.4$. The integrated luminosity used is $2.5 fb^{-1}$. This is 0.2 fb^{-1} more than the 13 TeV LFV dataset. The reason for this discrepancy is that LFV events must come from a time period in which the forward hadronic calorimeter is certified to be working correctly. The jet pull in VBF Higgs events requires the use of the forward calorimeter system to identify high $|\eta|$ jets, but this requirement is not necessary for a W+jets measurement.

7.1.2 Monte Carlo

Background processes with similar final state signatures as W+jets are Z+jets, $t\bar{t}$, single top, diboson +jets, and QCD multijets. The QCD multijet background is estimated using a data-driven method, as discussed in section 7.1.3. The other backgrounds are all estimated using Monte Carlo. W+jets and Z+jets samples are generated with MADGRAPH5_AMC@NLO. The $t\bar{t}$ background is generated using POWHEG. Single top backgrounds are generated using MADGRAPH5_AMC@NLO or POWHEG, depending on the channel. The single top simulation is generated with a mixture of

POWHEG and MadGraph5_aMC@NLO, depending on the decay channel. Dibison simulation is generated using POWHEG for WW and PYTHIA8 for WZ and ZZ. The CUETP8M1[52] tune scenario is used in PYTHIA8 for parton showering and hadronization. MCFM is used the evaluate the Monte Carlo cross sections at NLO.

The experimentally measured W+jets differential cross sections are compared to two matrix element calculations. The leading order calculation is generated by MADGRAPH5 interfaced with PYTHIA8 for parton showering and hadronization. The CTEQ6L1 parton distribution functions are used [54], and the matrix element calculation is matched to the parton showering using the MLM scheme [11][12] The MLM matching algorithm is as follows. After parton showering in pythia, as described in section 4.1, the anti-k_t jet cone algorithm is applied to the showered partons. The jet closest in (η, ϕ) to the hardest parton is selected. If the distance between the jet and the parton is less than 1.5 times the cone of radius R used in the jet clustering algorithm, then the jet and the parton are matched. Each hard parton must be matched to a jet. If a hard parton in an event is not matched to a jet, then the event is rejected. The NLO calculation is computed with MadGraph5_aMC@NLO interfaced with PYTHIA8 in the CUETP8M1 tune. The FXFX merging scheme [39] is used with a merging scale parameter of 30 GeV. The NNPDF 3.0 NLO PDF [17] is used for the matrix element calculation, while the NNPDF 2.3 LO [16][15] is used for the parton showering and hadronization. This gives NLO accuracy for 0,1, and 2 jet events and LO accuracy for 3 and 4 jet events.

7.1.3   QCD Background

The QCD background is estimated using a data driven method. Four regions are defined. Region A, the signal region, requires an isolated muon and a $M_T > 50$ GeV cut. Region B requires an isolated muon and a $M_T < 50$ GeV, minimizing the W+Jets contribution. Regions C and D are analogous to regions A and B, but with an inverted muon isolation requirement, which maximizes the QCD contribution.

In region A, the W+jets Monte Carlo is scaled by a factor f_W such that the number of W+jets and background Monte Carlo events exactly matches the number of data events in region A. This same scale factor is applied to W+jets in regions B,C, and D as well.

In region C, a histogram that determines the QCD shape is created by subtracting the Monte Carlo histograms from the data histograms.

The muon isolation fake rate $f_{B/D}$ is calculated by dividing the difference between the data and Monte Carlo event counts in region B by the difference between the data and Monte Carlo event counts in region D. This fake rate factor is applied to the QCD histogram in region C to give a QCD estimate in the signal region.

However, the scale factor for W+jets computed in region A at the beginning of the procedure did not take the QCD background into account. The calculation is repeated 11 additional times, at which point f_W and $f_{B/D}$ will have stabilized.

7.2   Event Selection

As mentioned in section 2.2, W bosons decay to hadrons or leptons. This analysis focuses on the muonic decay channel, where the W boson decays to a muon and its associated neutrino. The muon system of CMS, as discussed in section 3.2.5, enables particularly high muon efficiency and accurate energy resolution. Complicated scenarios such as hadronic tau decays or superclusters of e/γ deposits are avoided by choosing the muonic channel. Muons are required to have $p_T > 25$ GeV, $|\eta| < 2.4$, and $I_{PF}^{\mu} < 0.15$. Jets are identified using the anti-k_T algorithm (section 5.4.1), with $p_T > 30$ GeV and $|y| < 2.4$. Jets are required to be separated by $\Delta R > 0.4$ from muon candidates. Jets that originate from pileup interactions, rather than the hard scattering event that produces the W boson, are referred to as pileup jets. The contamination from pileup jets is reduced by requiring jets to originate from the same vertex as the muon candidate. This is accomplished by passing vertex and jet shower shape variables into a boosted decision tree which returns a discriminator for pileup jets [2]. An additional cut is placed on the discriminator so that W+jets events exhibit a minimal dependence on the pileup.

The main background to W+jets production at high jet multiplicities is $t\bar{t}$ production. The top quarks will decay into their lighter partners, bottom quarks, which will produce b jets. The $t\bar{t}$ contamination is reduced by applying a b-jet veto, as discussed in section 5.4.1. The yield differences between table 7.1 and table 7.2 show the effect of the b-jet veto. At high jet multiplicities the $t\bar{t}$ contribution is reduced by approximately a factor of six while the $W \rightarrow \mu\nu$ signal remains at 80%-90% of its pre-veto yield.

Events are vetoed if they contain additional muons with $p_T > 15$ GeV and $|\eta| < 2.4$. The transverse mass (M_T) between the muon and the missing transverse energy is defined as $M_T(\mu, E_T^{miss}) = \sqrt{2p_T^{\mu} E_T^{miss}(1 - cos\Delta\phi)}$, where $\Delta\phi$ is the difference in azimuthal angle between

Table 7.1: Number of events in data and simulation as a function of the exclusive jet multiplicity before the implementation of b tag veto. The diboson samples (WW, WZ, and ZZ) are represented by VV.

	$N_{\text{jets}} = 0$	$N_{\text{jets}} = 1$	$N_{\text{jets}} = 2$	$N_{\text{jets}} = 3$	$N_{\text{jets}} = 4$	$N_{\text{jets}} = 5$	$N_{\text{jets}} = 6$
VV	4041	3122	1336	382	92	19	3
QCD	411822	109220	23304	4343	639	126	0
Single top	4341	13379	11519	5437	1956	612	176
DYJets	624220	94046	21916	4805	1036	209	59
$t\bar{t}$	1951	12400	34151	46371	31871	13838	5147
$W \rightarrow \mu\nu$	14401433	2004189	433512	84961	16012	3438	369
TOTAL	15447808	2236356	525738	146299	51606	18242	5754
Data	16105074	2204919	526117	153912	55339	18580	5483

Table 7.2: Number of events in data and simulation as a function of the exclusive jet multiplicity after the implementation of b tag veto. The diboson samples (WW, WZ, and ZZ) are represented by VV.

	$N_{\text{jets}} = 0$	$N_{\text{jets}} = 1$	$N_{\text{jets}} = 2$	$N_{\text{jets}} = 3$	$N_{\text{jets}} = 4$	$N_{\text{jets}} = 5$	$N_{\text{jets}} = 6$
VV	3961	2917	1189	325	76	15	2
QCD	393632	94835	18662	3526	465	263	0
Single top	3426	6850	4747	1884	558	145	38
DYJets	617130	89398	19782	4278	887	180	45
$t\bar{t}$	1323	5178	9765	9989	5850	2332	825
$W \rightarrow \mu\nu$	14250375	1914920	400201	76319	13153	2995	310
TOTAL	15269847	2114098	454346	96321	20989	5930	1220
Data	15962756	2096555	454919	101282	24903	6319	1515

the muon momentum and the $\vec{E}_T^{miss}$ vector. In W+jets events the muon and the neutrino tend to be emitted back-to-back, which maximizes M_T with respect to $\Delta\phi$. Events are required to be in the W transverse mass peak region, which is defined by $M_T > 50$ GeV.

7.3   Data-Simulation Comparisons

After applying selections, the data and simulation agreement is compared for several kinematic observables. These are show in figures 7.1 through 7.5.

The event yields for data and simulation processes in each bin of exclusive jet multiplicity are listed in Table 7.2.

At high jet multiplicites, the W+jets signal is less dominant and the accuracy of the largest background, $t\bar{t}$, becomes more important. A $t\bar{t}$ enriched control sample was created by removing the b-jet veto and requiring two or more b-tagged jets. The data-simulation agreement was found to be with 10% up to the exclusive jet multiplicity of 5 and the inclusive jet multiplicity of 4.
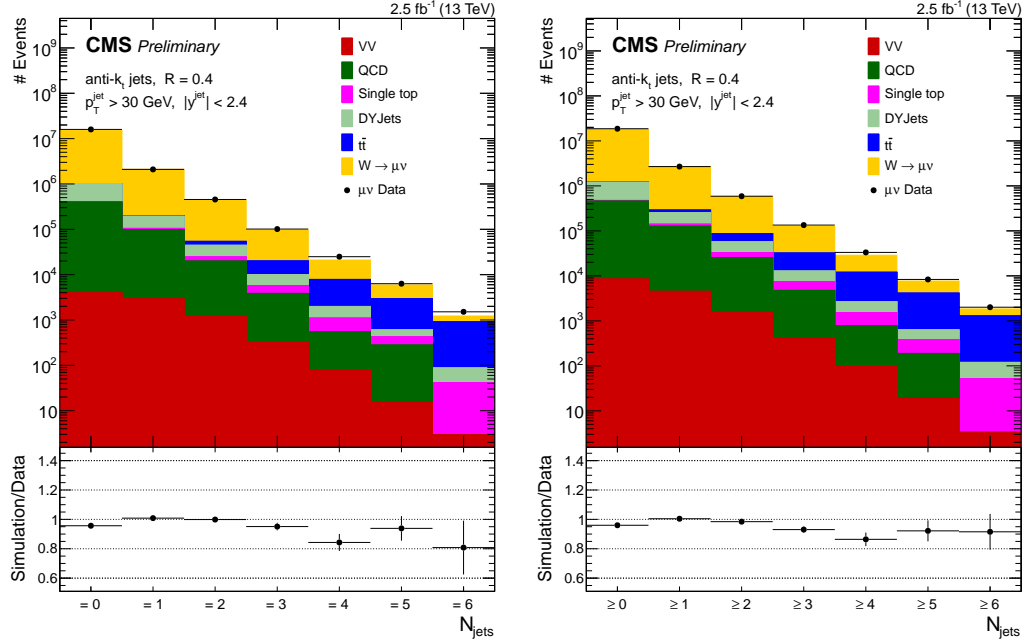
Figure 7.1: Data to simulation comparison of exclusive (left) and inclusive (right) jet multiplicity. QCD background is estimated using a data-driven method. The diboson samples (WW, WZ, and ZZ) are represented by VV.
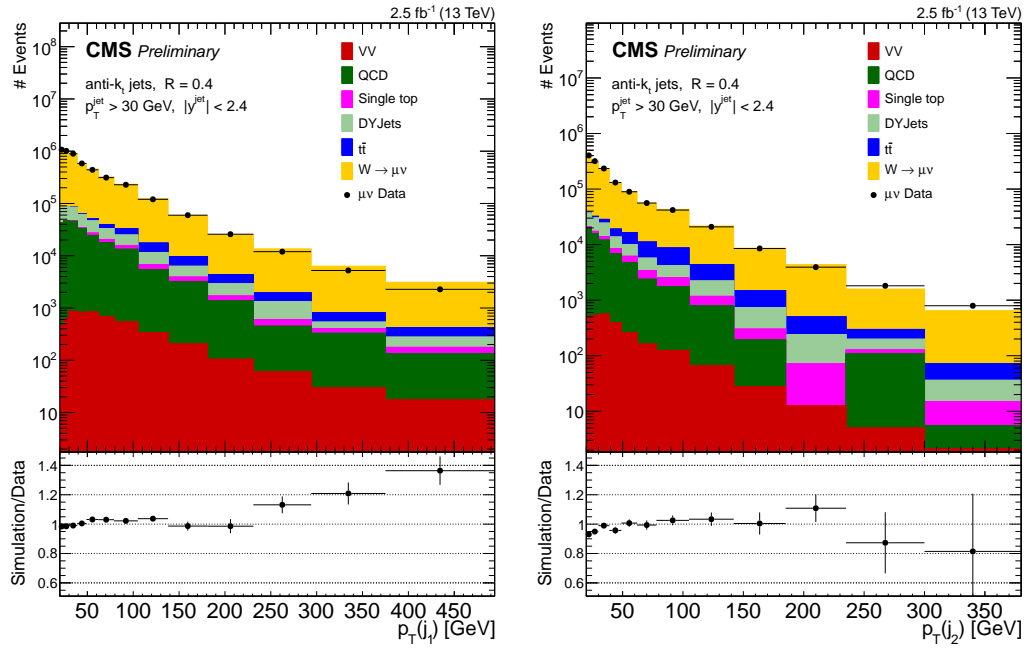
Figure 7.2: Data to simulation comparison of 1st (left) and 2nd (right) jet p_T. QCD background is estimated using a data-driven method. The diboson samples (WW, WZ, and ZZ) are represented by VV.
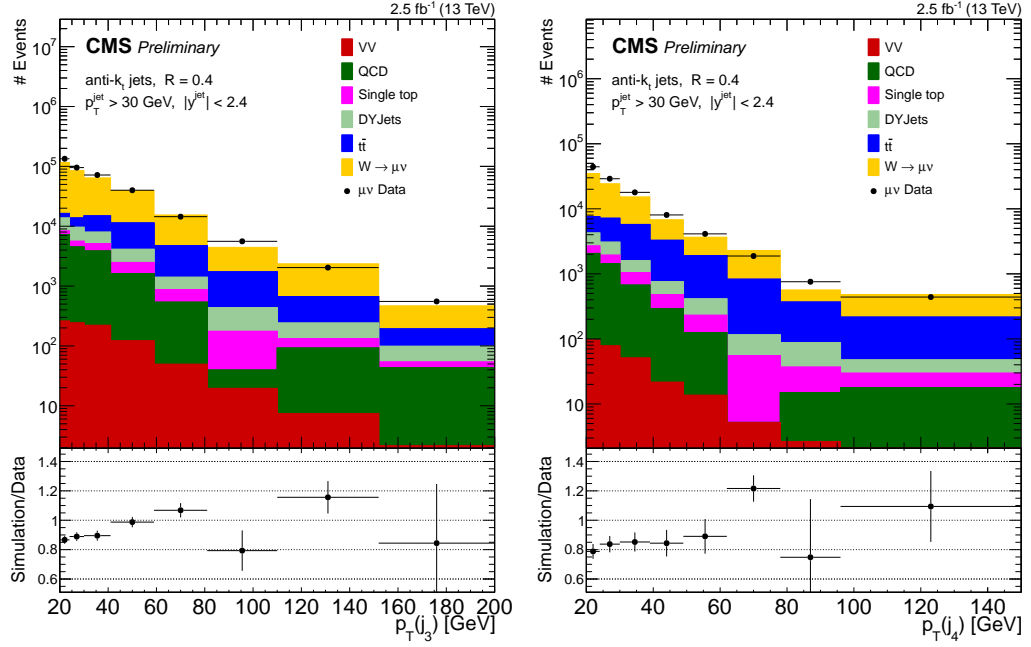
Figure 7.3: Data to simulation comparison of 3rd (left) and 4th jet p_T. QCD background is estimated using a data-driven method. The diboson samples (WW, WZ, and ZZ) are represented by VV.
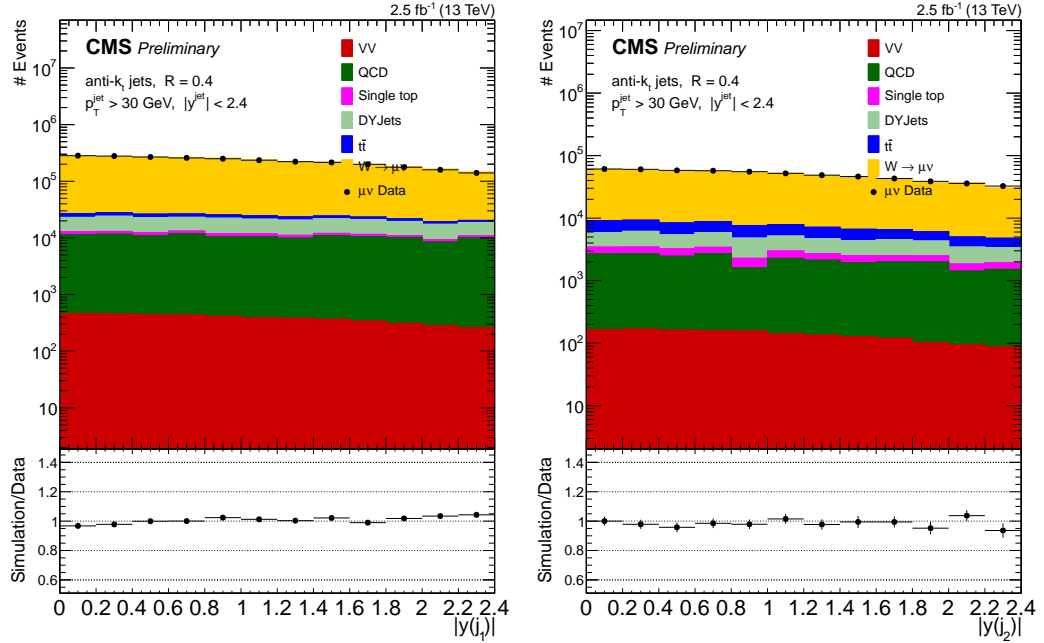
Figure 7.4: Data to simulation comparison of 1st (left) and 2nd jet $|y|$. QCD background is estimated using a data-driven method. The diboson samples (WW, WZ, and ZZ) are represented by VV.
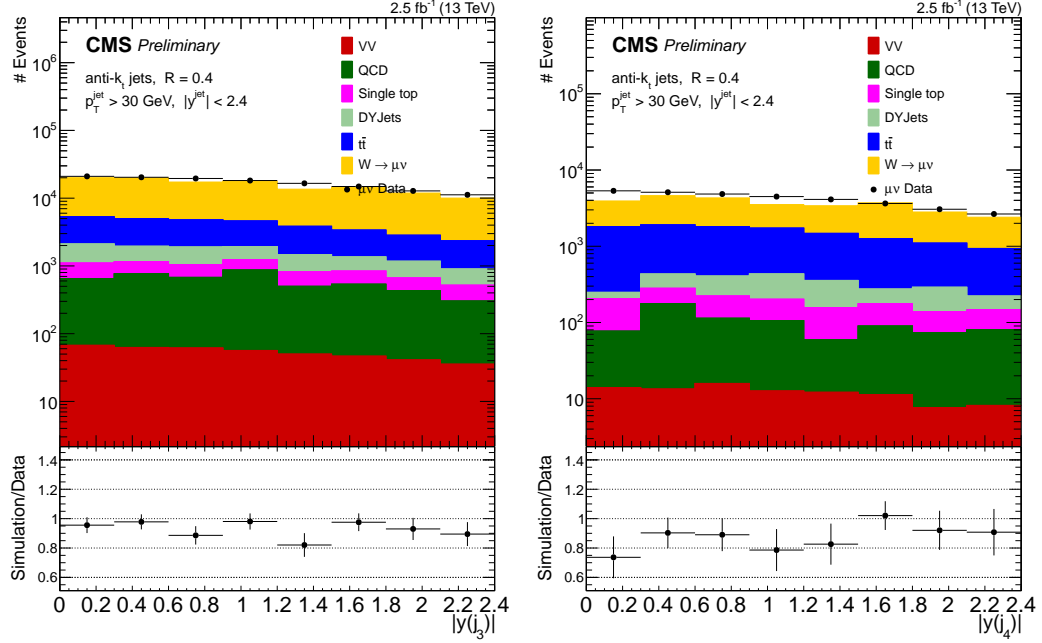
Figure 7.5: Data to simulation comparison of 3rd (left) and 4th jet $|y|$. QCD background is estimated using a data-driven method. The diboson samples (WW, WZ, and ZZ) are represented by VV.

7.4 Detector Unfolding

The CMS detector, like any experimental instrument, is not infallible. The detector response can cause a measurement to deviate from its true value. For example, for a given event with a leading jet p_T in the 40-45 GeV bin, there is a probability that the jet could have had a p_T in the 35-40 GeV bin or in the 45-50 GeV bin. The detector response can cause the measured value to deviate from its true value. A response matrix is used to transform the measured values to the true values.

A Bayesian unfolding method [35] is used for W+Jets. The ith generated event Gen_i results in a measured event $Reco_i$. The probability of a generated event to be observed in the ith bin is $P(Gen_i)$. The probability that a reco event in bin j is due to a generated event in bin i is $P(Reco|Gen_i)$. From those relations, Bayes' theorem [20] is used to obtain $P(Gen_i|Reco_j) = \frac{P(Reco_j|Gen_i)P(Gen_i)}{\Sigma_{l=1}^{n_{bins}} P(Reco_j|Gen_l)P(Gen_l)}$

The probability to observe a reconstructed event in bin j given an generator level event in bin i is given by $P(Reco_j|Gen_i)$. This probability is represented by a matrix, defined as the response matrix (R_{ji}). The response matrix is calculated by using Monte Carlo (Section 4) to compare

generator level events to events reconstructed with detector simulation. It can be visualized for a particular variable by plotting the number of generator level events versus the number of reco level events. After calculating the response matrix, Bayes' theorem is used to obtain the probability that a reconstructed event in bin j was due to a generator level event in bin i. This probability distribution, $P(Gen_i|Reco_j)$, is referred to as the smearing matrix, S_{ij}.

The number of true events observed in bin i is defined as $\hat{n}(i) = \frac{1}{\epsilon_i}\Sigma_{j=1}^{n_{bins}}n_{obs}(j)S_{ij}$, where ϵ_i is the efficiency of observing an event that was generated in bin i and is defined by $\epsilon_i = \Sigma_{j=1}^{n_bins}R_{ji}$.

Given the above formulas, the number of true events in each bin can be calculated. The only unknown quantity is $P(Gen_i)$, the probability to observe a generator level event in bin i. This probability is determined by an iterative χ^2 fit. First, $P(Gen_i)$ is estimated from the Monte Carlo distribution. This allows a simple estimation of the true number of events: $\hat{n}_0 = P(Gen_i)N_{obs}$, where N_{obs} is the total number of observed events. Then use Bayes' Theorem to calculate $\hat{n}(i)$ and $\hat{P}(Gen_i)$. The third step is to calculate the χ^2 distribution between $\hat{n}(i)$ and $\hat{n}_0(i)$. The iterative procedure is then repeated, with $\hat{n}(i)$ and $\hat{P}(Gen_i)$ used in the first step. The iterative procedure is repeated at least four times and is concluded when $\chi^2/\nu < \frac{1}{\sqrt{2}}$.

In the generator level simulation, the events are required to pass the same event selection used in the reconstruction level, including the requirements on muon p_T and $|\eta|$, jet p_T and $|y|$, and M_T. The generator level jets are built in the same way as the reconstructed jets, via the the anti-k_T algorithm. The generator level muons are "dressed" by recombining the bare generator level muons with all of the radiated photons within a cone of radius 0.1 to accout for final state radiation. The M_T of the the W Boson is calculated using the dressed muon and the neutrino. The response object is then constructed from the response matrix, the reconstructed level distribution, and the corresponding generator level distribution.

7.5   Systematic Uncertainties

The dominant source of systematic uncertainty is the jet energy scale uncertainty. This uncertainty is equal to 1.4% for a jet multiplicity of 1 and increases with the number of reconstructed jets. Uncertainties in the jet energy scale are propagated to the calculation of E_T^{miss}.

The background cross sections are varied within their theoretical uncertainties. The cross section of $t\bar{t}$, which is the largest background contribution, is varied by 10%. The additional back-

grounds are simultaneously varied up and down by 7% (ZZ, WZ), 6% (WW and tW single top), and 4% (Z+jets, s and t-channel single top).

A systematic uncertainty associated with the generator used to build the unfolding response matrix is computed by weighting the simulation to agree with the data in p_T, $|y|$, and H_T distributions, and then building a reweighted response to unfold the data. The difference between the nominal results and the results unfolded using the reweighted response matrix is taken as the systematic uncertainty associated with the unfolding response matrix.

A systematic uncertainty associated with the jet energy resolution (JER) scale factors is computed by varying the scale factors up and down by 1σ. The resulting uncertainty is on the order of 1%.

The uncertainty in the pileup modeling is computed by varying the inelastic proton-proton cross section within its uncertainty of $\pm5\%$. The resulting uncertainty is on the order of 1%.

The data-to-simulation correction factors of the b tagging efficiencies are varied up and down by 1σ. The entire analysis is performed with these variations and the final unfolded results are compared to the results of the unshifted analysis. The effect on the measured cross section varies between 0.4%-11% depending on the jet multiplicity as shown in table 7.4.

The uncertainties of the muon trigger, identification, and isolation efficiencies are summed in quadrature to give an overall systematic uncertainty on the data-to-simulation muon scale factors. The uncertainty is computed to be 1.23%.

The uncertainty on the integrated luminosity measurement is estimated as 4.6%.

7.5.1   Theoretical Uncertainties

As shown in figure 7.7, the measured W+Jets cross section is compared to the results from the MadGraph5_aMC@NLO FXFX event generator. The shaded blue band corresponding to the theoretical uncertainties on the MadGraph5_aMC@NLO FXFX event generator is computed as described below.

The factorization (μ_f) and renormalization scales (μ_r) in QCD control the cut-offs for infrared divergences and ultraviolet divergences. By varying these scales in Monte Carlo and measuring the shifts in event acceptance, systematic uncertainties for W+Jets production can be obtained.

These scales affect QCD radiation, which will effect parton showering and jet creation. Therefore, the effects of varying these scales are jet bin dependant. However, we cannot simply vary

the scales in exclusive jet bins because the uncertainties will be underestimated [43] [63] To properly calculate uncertainties in exclusive jet bins, the exclusive uncertainties must be defined in terms of the inclusive uncertainties. This is accomplished by defining the exclusive cross section in terms of inclusive cross sections: $\sigma_N = \sigma_{\geq N} - \sigma_{\geq N+1}$

The uncertainty in the exclusive cross section is then calculated by adding the inclusive cross section uncertainties in equation (7.1).

$$\Delta_N^2 = \Delta_{\geq N}^2 + \Delta_{\geq N+1}^2 \tag{7.1}$$

7.5.2   Methodology

The analysis cuts are applied at generator level in the W+Jets MadGraph5_aMC@NLO sample. The muon is required to have originated from a W boson decay. Jets are required to have $p_T > 30$ and $|\eta| < 2.4$. Muons are required to have $p_T > 25$ GeV and $|\eta| < 2.4$. An M_T cut of the form $M_T(\mu, E_T^{miss}) > 50$ GeV is applied, and the muon and the jet are required to be separated by $\Delta R(\mu, jet) > 0.4$. Starting from a default value of $\mu_r = \mu_f = 1$, the scales are varied by a factor of two. The theoretical uncertainty, which is defined in terms of event acceptance, is equal to $\frac{|A_{\mu=2} - A_{\mu=1/2}|}{2 \times A_{\mu=1}}$ The results are in Table 7.3. Note that adding the inclusive uncertainties in quadrature increases the uncertainty, especially in the zero jets bin.

Note that in Figure 7.6 the exclusive acceptance difference in the first bin ($N_{jets} = 0$) is much smaller than the acceptance differences in the neighboring inclusive bins of $N_{jets} \geq 0$ and $N_{jets} \geq 1$, which represent the uncertainties at the boundaries of the $N_{jets} = 0$ region. The seemingly small uncertainty in the $N_{jets} = 0$ region is not in fact caused by small changes within the region but by cancellation of large migrations of events into and out of the region. A smaller version of this effect can be seen in all of the other exclusive jet bins. It is therefore justified to quantify the theoretical uncertainty of an exclusive jet bin by the migrations of events into and out of the bin, rather than the absolute change of the number of events in the jet bin. The uncertainties at the boundaries of the jet bin are represented by the inclusive uncertainties, which leads us to equation.

Table 7.3: Comparison of exclusive uncertainties in jet bins. The direct scale variation method means that the uncertainties have been computed by observing the difference in yields after varying μ_r and μ_f. The via inclusive method adds the inclusive uncertainties according to 7.1.

$N_{jets} =$	Inclusive(%)	Direct Scale (%)	Via Inclusive (%)
0	0.1	1.1	6.2
1	6.2	5.5	10.8
2	8.9	8.2	14.7
3	11.8	11.7	16.8
4	11.9	10.9	19.5
5	15.5	16.4	19.8
6	12.3	9.9	23.7
7	20.3	19.2	30.6
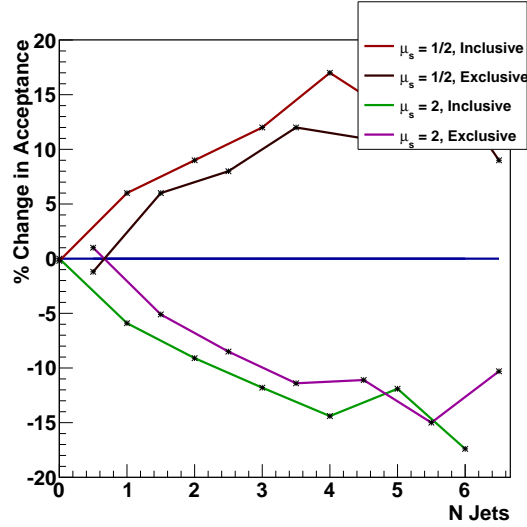
Figure 7.6: Changes in acceptance for inclusive and exclusive jets, as a function of number of jets. The points in the middle of the bins are the exclusive variations and the points at the edges of the bins are the inclusive variations. The inclusive variations show the changes in acceptance at the boundaries of the exclusive jet bins.

7.6 Results

The cross sections of jet multiplicities are measured up to 5 jets. The measured W+jets differential cross section distributions are compared with the predictions of the MADGRAPH5_AMC@NLO FXFX and MADGRAPH5 LO MLM event generators. As shown in figure 7.7, the measured data and predictions are generally agreement within uncertainties. MADGRAPH5 appears to underestimate the measured cross section up to a multiplicity of 4. Tables 7.4 and 7.5 show the measured cross sections and associated uncertainties.

Table 7.4: Differential cross section in exclusive jet multiplicity and break down of the systematic uncertainties for the muon decay channel.

					Exclusive jet multiplicity							
N_{jets}	$\frac{d\sigma}{dN_{jets}}$[pb]	Tot. Unc [%]	stat [%]	Simulation stat [%]	JES [%]	JER [%]	PU [%]	XSEC [%]	Lumi [%]	BtagSF [%]	Unf [%]	Eff [%]
= 0	7.43e+03	5.3	0.064	0.15	0.15	0.092	1.7	0.17	4.9	0.40	0.0	0.53
= 1	867.	5.6	0.28	0.69	1.4	0.43	1.3	0.25	5.2	0.82	0.0	0.56
= 2	197.	8.8	0.76	1.8	6.4	0.31	1.5	0.58	5.3	1.7	0.0	0.57
= 3	44.8	13.	1.9	4.2	8.7	0.74	3.3	1.9	5.9	2.8	0.0	0.64
= 4	12.8	18.	4.3	9.4	11.	0.42	0.12	4.9	7.2	5.1	0.0	0.78
= 5	1.76	41.	14.	27.	21.	0.12	2.1	11.	11.	11.	0.0	1.1
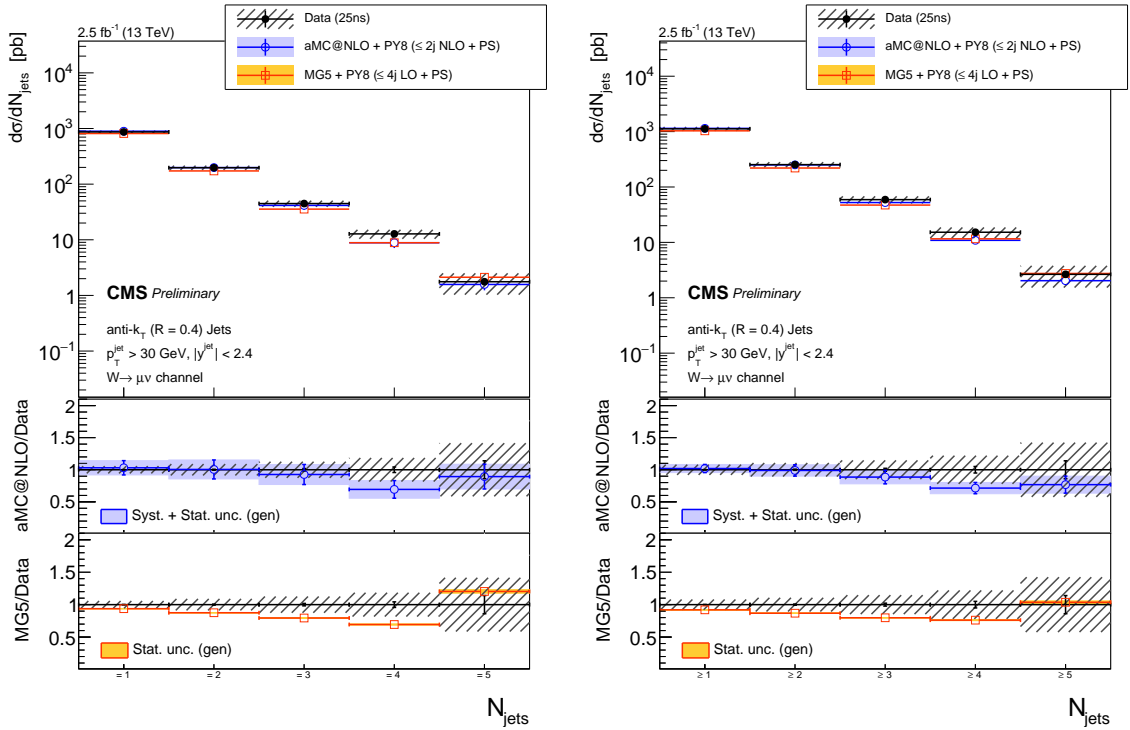
Figure 7.7: The differential cross section measurement for the exclusive and inclusive jet multiplicities, compared to the predictions of MADGRAPH5_AMC@NLO and MADGRAPH5. The black circular markers with the grey hatched band represent the unfolded data measurement and its total experimental uncertainty. The blue shaded band around the MADGRAPH5_AMC@NLO prediction represents its theoretical uncertainty including both statistical and systematical uncertainties. MADGRAPH5 is shown only with its statistical uncertainty. The lower panels show the ratios of the prediction to the unfolded data.

Table 7.5: Differential cross section in inclusive jet multiplicity and break down of the systematic uncertainties for the muon decay channel.

Inclusive jet multiplicity

N_{jets}	$\frac{d\sigma}{dN_{\text{jets}}}$[pb]	Tot. Unc [%]	stat [%]	MC stat [%]	JES [%]	JER [%]	PU [%]	XSEC [%]	Lumi [%]	BtagSF [%]	LER [%]	Unf [%]	Eff [%]
≥ 0	8.59e+03	5.3	0.056	0.15	0.53	0.024	1.7	0.19	4.9	0.48	0.0	0.54	
≥ 1	1.12e+03	7.9	0.32	0.55	5.6	0.15	1.3	0.42	5.2	1.1	0.0	0.57	
≥ 2	253.	11.	0.90	1.4	7.9	0.49	2.9	1.1	5.6	2.2	0.0	0.61	
≥ 3	59.1	15.	2.3	3.2	11.	1.1	2.5	3.1	6.5	4.0	0.0	0.71	
≥ 4	15.2	22.	5.1	8.0	15.	1.6	0.29	6.8	8.1	7.1	0.0	0.88	
≥ 5	2.65	42.	14.	18.	26.	2.3	1.5	14.	12.	15.	0.0	1.3	

Chapter 8

Conclusions

8.1 Summary

8.2 Future Outlook

Bibliography

[1] Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV. Technical Report CMS-PAS-PFT-10-002, CERN, Geneva, 2010.

[2] Pileup Jet Identification. Technical Report CMS-PAS-JME-13-005, CERN, Geneva, 2013.

[3] Tau ID Performance Plots. Apr 2014.

[4] Wolfgang Adam, R Frhwirth, Are Strandlie, and T Todor. Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC. Technical Report CMS-NOTE-2005-001, CERN, Geneva, Jan 2005.

[5] Q. R. Ahmad et al. Measurement of the rate of $\nu_e + d \rightarrow p + p + e^-$ interactions produced by ^8B solar neutrinos at the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 87:071301, 2001.

[6] Q. R. Ahmad et al. Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 89:011301, 2002.

[7] Sergey Alekhin et al. The PDF4LHC Working Group Interim Report. 2011.

[8] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06:043, 2010.

[9] John Allison et al. Geant4 developments and applications. *IEEE Trans. Nucl. Sci.*, 53:270, 2006.

[10] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014.

[11] Johan Alwall, Simon de Visscher, and Fabio Maltoni. QCD radiation in the production of heavy colored particles at the LHC. *JHEP*, 02:017, 2009.

[12] Johan Alwall et al. Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions. *Eur. Phys. J.*, C53:473–500, 2008.

[13] J R Andersen et al. Handbook of LHC Higgs Cross Sections: 3. Higgs Properties. 2013.

[14] Nima Arkani-Hamed, Savas Dimopoulos, and G. R. Dvali. The Hierarchy problem and new dimensions at a millimeter. *Phys. Lett.*, B429:263–272, 1998.

[15] Richard D. Ball, Valerio Bertone, Francesco Cerutti, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Jose I. Latorre, Juan Rojo, and Maria Ubiali. Impact of Heavy Quark Masses on Parton Distributions and LHC Phenomenology. *Nucl. Phys.*, B849:296–363, 2011.

[16] Richard D. Ball, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Jose I. Latorre, Juan Rojo, and Maria Ubiali. A first unbiased global NLO determination of parton distributions and their uncertainties. *Nucl. Phys.*, B838:136–206, 2010.

[17] Richard D. Ball et al. Parton distributions for the LHC Run II. *JHEP*, 04:040, 2015.

[18] Vernon D. Barger and Roger J. N. Phillips. *Collider Physics (Frontiers in Physics)*. Addison Wesley Publishing Company, 1987.

[19] G. W. Bennett et al. Measurement of the negative muon anomalous magnetic moment to 0.7 ppm. *Phys. Rev. Lett.*, 92:161802, 2004.

[20] Philip Bevington and D. Keith Robinson. *Data Reduction and Error Analysis for the Physical Sciences*. McGraw-Hill Education, 3rd edition, 7 2002.

[21] G. C. Branco, P. M. Ferreira, L. Lavoura, M. N. Rebelo, Marc Sher, and Joao P. Silva. Theory and phenomenology of two-Higgs-doublet models. *Phys. Rept.*, 516:1–102, 2012.

[22] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Anti-k(t) jet clustering algorithm. *JHEP*, 04:063, 2008.

[23] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet User Manual. *Eur. Phys. J.*, C72:1896, 2012.

[24] John M. Campbell and R. Keith Ellis. An Update on vector boson pair production at hadron colliders. *Phys. Rev.*, D60:113006, 1999.

[25] John M. Campbell, R. Keith Ellis, and Walter T. Giele. A Multi-Threaded Version of MCFM. *Eur. Phys. J.*, C75(6):246, 2015.

[26] John M. Campbell, R. Keith Ellis, and Ciaran Williams. Vector boson pair production at the LHC. *JHEP*, 07:018, 2011.

[27] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett.*, B716:30–61, 2012.

[28] Serguei Chatrchyan et al. Identification of b-quark jets with the CMS experiment. *JINST*, 8:P04013, 2013.

[29] Serguei Chatrchyan et al. The performance of the CMS muon detector in proton-proton collisions at sqrt(s) = 7 TeV at the LHC. *JINST*, 8:P11002, 2013.

[30] Serguei Chatrchyan et al. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *JINST*, 9(10):P10009, 2014.

[31] Serguei Chatrchyan et al. Measurement of inclusive W and Z boson production cross sections in pp collisions at $\sqrt{s} = 8$ TeV. *Phys. Rev. Lett.*, 112:191802, 2014.

[32] CMS Collaboration. Performance of -lepton reconstruction and identification in cms. *Journal of Instrumentation*, 7(01):P01001, 2012.

[33] J. S. Conway. Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra. In *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN,Geneva, Switzerland 17-20 January 2011*, 2011.

[34] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J.*, C71:1554, 2011. [Erratum: Eur. Phys. J.C73,2501(2013)].

[35] G. D'Agostini. A multidimensional unfolding method based on bayes' theorem. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 362(23):487 – 498, 1995.

[36] A. Denner, S. Heinemeyer, I. Puljak, D. Rebuzzi, and M. Spira. Standard Model Higgs-Boson Branching Ratios with Uncertainties. *Eur. Phys. J.*, C71:1753, 2011.

[37] S. Dittmaier et al. Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables. 2011.

[38] R. Keith Ellis, I. Hinchliffe, M. Soldate, and J. J. van der Bij. Higgs Decay to tau+ tau-: A Possible Signature of Intermediate Mass Higgs Bosons at the SSC. *Nucl. Phys.*, B297:221–243, 1988.

[39] Rikkert Frederix and Stefano Frixione. Merging meets matching in MC@NLO. *JHEP*, 12:061, 2012.

[40] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP*, 11:070, 2007.

[41] R. Fruhwirth. Application of Kalman filtering to track and vertex fitting. *Nucl. Instrum. Meth.*, A262:444–450, 1987.

[42] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81:1562–1567, 1998.

[43] Shireen Gangal and Frank J. Tackmann. Next-to-leading-order uncertainties in Higgs+2 jets from gluon fusion. *Phys. Rev.*, D87(9):093008, 2013.

[44] D.J. Griffiths. *Introduction to Electrodynamics*. Prentice Hall, 1999.

[45] F. Halzen and Alan D. Martin. Quarks and leptons: An introductory course in modern particle physics. 1984. New York, Usa: Wiley (1984) 396p.

[46] Roni Harnik, Joachim Kopp, and Jure Zupan. Flavor Violating Higgs Decays. *JHEP*, 03:026, 2013.

[47] Andreas Hocker et al. TMVA - Toolkit for Multivariate Data Analysis. *PoS*, ACAT:040, 2007.

[48] Vardan Khachatryan et al. Differential cross section measurements for the production of a W boson in association with jets in protonproton collisions at $\sqrt{s} = 7$ TeV. *Phys. Lett.*, B741:12–37, 2015.

[49] Vardan Khachatryan et al. Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at s = 8 TeV. *JINST*, 10(06):P06005, 2015.

[50] Vardan Khachatryan et al. Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8$ TeV. *JINST*, 10(02):P02006, 2015.

[51] Vardan Khachatryan et al. Search for Lepton-Flavour-Violating Decays of the Higgs Boson. *Phys. Lett.*, B749:337–362, 2015.

[52] Vardan Khachatryan et al. Event generator tunes obtained from underlying event and multi-parton scattering measurements. *Eur. Phys. J.*, C76(3):155, 2016.

[53] Henning Kirschenmann. Jet performance in CMS. Technical Report CMS-CR-2013-325, CERN, Geneva, Oct 2013.

[54] Aneesh V. Manohar and Wouter J. Waalewijn. What is Double Parton Scattering? *Phys. Lett.*, B713:196–201, 2012.

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

[56] Pavel M. Nadolsky, Hung-Liang Lai, Qing-Hong Cao, Joey Huston, Jon Pumplin, Daniel Stump, Wu-Ki Tung, and C. P. Yuan. Implications of CTEQ global analysis for collider observables. *Phys. Rev.*, D78:013004, 2008.

[57] Paolo Nason. A New method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004.

[58] K. A. Olive et al. Review of Particle Physics. *Chin. Phys.*, C38:090001, 2014.

[59] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, Pavel M. Nadolsky, and W. K. Tung. New generation of parton distributions with uncertainties from global QCD analysis. *JHEP*, 07:012, 2002.

[60] Lisa Randall and Raman Sundrum. A Large mass hierarchy from a small extra dimension. *Phys. Rev. Lett.*, 83:3370–3373, 1999.

[61] A L Read. Modified frequentist analysis of search results (the CL_s method). (CERN-OPEN-2000-205), 2000.

[62] Torbjrn Sjstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An Introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159–177, 2015.

[63] Iain W. Stewart and Frank J. Tackmann. Theory Uncertainties for Higgs and Other Searches Using Jet Bins. *Phys. Rev.*, D85:034011, 2012.

[64] A. Wald. *Tests of Statistical Hypotheses Concerning Several Parameters when the Number of Observations is Large*. American Mathematical Society, 1943.

[65] Zbigniew Was. TAUOLA for simulation of tau decay and production: perspectives for precision low energy and LHC applications. *Nucl. Phys. Proc. Suppl.*, 218:249–255, 2011.