Title
Search for Excited Tau Leptons in pp Collisions at √? = 13 TeV at the LHC

Author
Welton, Troy

Publication Date
2022

Peer reviewed|Thesis/dissertation

Search for Excited Tau Leptons in pp Collisions at $\sqrt{s} = 13$ TeV at the LHC

By

TROY WELTON
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

PHYSICS

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

John Conway, Chair

Robin Erbacher

Maxwell Chertok

Committee in Charge

2022

Abstract

This dissertation details the context and methods involved in the analysis conducted to search for excited tau leptons (τ^*) using data collected during Run II of the Large Hadron Collider (LHC) by the Compact Muon Solenoid (CMS) experiment. The dataset used corresponds to 138 fb^{-1} of integrated luminosity in proton-proton collisions at 13 TeV center-of-mass energy. The analysis selects events with two tau candidates, at least one of which is a hadronic decay, and a high energy photon. The mass of the excited tau lepton is reconstructed by also accounting for the contribution from the invisible neutrino component. Expected limits are set at 95% confidence level excluding m_{τ^*} up to 4.5 TeV for a compositeness scale equal to the mass of the τ^* and up to 2.75 TeV for a compositeness scale of 10 TeV.

Contents

Chapter 1

Introduction

Scientists have always tried to discern what makes up the world around them. There have been various points in history where certain objects were thought to be the smallest building blocks of matter. From atoms to protons to quarks, new energy frontiers allow probing the compositeness of objects at smaller and smaller length scales. At the compositeness scale Λ for a given object, the object can break apart into its smaller components. Anything which has substructure must have a length scale at which the substructure can be observed. The length scale that may be probed is inverse to the energy scale used. At the compositeness scale, the energy imparted to the components of the object is enough to overwhelm the force holding the object together. The standard model (SM) of particles describes a broad range of observed interactions between particles. It has been tested through decades of experiment and has predicted discoveries such as the Higgs Boson [1]. Currently, leptons and quarks are the limit of the building blocks that have been observed. Leptons exhibit properties of point-like particles in the ways they have been observed them and, therefore, do not seem to be composite. This analysis probes the idea that leptons may have substructure, focusing specifically on the study of τ lepton, the heaviest of the leptons.

Compositeness models [2, 3, 4, 5, 6, 7, 8, 9, 10] aim to explain why quarks and leptons have a similar generational structure, as well as the mass hierarchy of these particles among

other things. The models introduce a unifying building block, called a preon, as a component of all fermions. To not have been observed yet, the compositeness scale (in terms of energy) must be quite large. Therefore, fermions could exchange these preons only via a contact interaction. One of the possible results of this contact interaction is an excited state of a lepton or a quark. This exhibits itself as a higher mass resonance. This excited state would then decay into a normal lepton and a gauge boson via a gauge interaction or into three fermions via another contact interaction.

Chapter 2

Theoretical Foundations

2.1 Natural Units

Most researchers in the field of particle physics use a peculiar system of units in equations and calculations called "natural units", which involves setting the values of $\hbar$ and c to 1, including their units. In this system, the units for mass, energy, and momentum are all identical. This can be seen by thinking about equations like the famous mass-energy equivalence $E = mc^2$ and the relationship between the energy and momentum of a photon $E = pc$. In fact, this reduces the Lorentz invariance from $E^2 = m^2c^4 + p^2c^2$ to $E^2 = m^2 + p^2$. Planck's constant $\hbar$ also comes up frequently in quantum mechanics and when discussing the properties of individual particles, so it is much easier to perform calculations without worrying about these factors. The factors can be put back in by considering the difference between the units of a quantity in another unit system and natural units.

The common unit of energy (and, therefore, of mass and momentum in natural units) is the electron volt (eV). An electron volt is the energy gained by an electron accelerated across a 1V potential. The energy scales relevant to this thesis are on the order of GeV or 10^9 eV. Other units can be derived in natural units in a similar manner to mass and momentum: by looking at standard equations that involve energy. Taking length as an example, $E = \frac{2\pi\hbar c}{\lambda}$

is the equation for the energy of a photon with λ having units of length. Therefore, length in this measuring system has units of eV^{-1}. The cross section is another quantity that will be discussed in more detail in Section 3.2. This has units of area and could be represented in natural units using eV^{-2}. An area of 1 cm^2 would be equivalent to 6.5×10^{11} eV^{-2} by removing two factors of $\hbar c$. The cross sections relevant to this analysis are on the order of 10^{-28} eV^{-2}.

2.2 Early Particle Physics

The field of particle physics, as this area of study is often referenced, is often cited to have begun in 1897 with the discovery of the electron by J.J. Thomson [11]. Thomson observed that cathode rays, which we now know is simply a stream of electrons that have been separated from atoms by the introduction of an electric potential, could be deflected by a magnetic field, implying that they possessed electric charge. Thomson was able to determine the charge-to-mass ratio of these "corpuscles" by using a combination of electric and magnetic fields that would achieve zero deflection of the beam by balancing the electromagnetic force on the particles of the beam in the transverse direction. Thomson believed that the electrons were scattered throughout a positively charged medium that was the rest of the atom. Experiments by Rutherford in 1911 revealed that the positive charge in an atom was concentrated in a very small region at the center of the atom based on the fact that positively charged particles being shot at the screen mostly were deflected at large angles instead of passing straight through, as would be the case if the charge were uniformly distributed. By 1933, protons and neutrons were also understood to exist and, along with electrons, to make up the types of matter that people saw in everyday life.

The next slew of discoveries came with deep inelastic scattering between electrons and protons. When the energy scale of the electron was greater than the mass of the proton, the proton was observed to break apart into constituent pieces that would then create new

particles not previously observed.

2.3 Standard Model

Quantum field theory and group theory form the basis of the Standard Model. Quantum field theory is a framework that combines the classical field theory of field equations with quantum mechanics and special relativity. [12] It is used to allow for the representation of particles as fields with well defined equations and properties. The group theory comes in to represent the properties of particles through inclusion in symmetry groups. Belonging to a symmetry group is expressed as carrying a "charge" of that group. From Noether's theorem, every differentiable symmetry generated by local actions corresponds to a conserved current. These "charges" carried by particles are therefore properties that are conserved in the universe. The standard model can be represented using the following structure:

$$SU(3)_c \times SU(2)_L \times U(1)_Y \tag{2.1}$$

Each of these symmetry groups controls a type of interaction and has associated force carriers, or bosons. $SU(3)_c$ represents the strong force and the associated "color charge", originally named for having three types of charge that were associated with the primary colors of light (red, green, and blue). The bosons associated with this force are the gluons. $SU(3)$ has 8 group generators, a fact which can be derived from the properties of special unitary groups in 3 dimensions. The Gell-Mann matrices are a set of 3×3 matrices that provide a basis for this group and can be used to represent the 8 types of gluons. The $SU(2)_L$ symmetry could be described by the Pauli matrices and produce 3 massless bosons that act only on the left-handed chirality of fermions. This is represented by the weak force. Finally, $U(1)$ could be represented by a single complex number and produce just one massless boson and represents hypercharge. The interesting thing about these symmetries is that they are not observed to be mediated by massless bosons. The reason for this is explained by the

Higgs field.

The key takeaway from the Higgs mechanism is that a gauge transformation can be applied that turns a complex field into a massive scalar field with massive gauge bosons. The Higgs mechanism predicted that there would be massive vector bosons to accompany the weak force. The Higgs mechanism also allowed for the unification of the electromagnetic and weak sectors into the electroweak sector by allowing for the mixing of the weak isospin with the hypercharge. In the end, this theory predicted the existence of the massive W bosons and Z boson and the orthogonal connection between the Z boson and the massless photon. In 1983, the truth of these predictions was substantiated by the discovery of these particles at the Super Proton Synchrotron through the UA1 and UA2 experiments. The existence of the Higgs boson itself was not proven until 2012 [1].

The matter in the universe is subject to the interactions described above. Different types of matter experience different subsets of these interactions. Protons, neutrons and other hadrons are composed of quarks. Quarks experience both the strong force and the electroweak force. The strong force works on such a small scale that quarks are always confined because it would take more energy to separate them by a large distance than to produce more quarks out of the vacuum. Matter in everyday life is composed of up (u) and down (d) quarks, which are the valence quarks in protons (uud) and neutrons (udd). The other quarks exist as sea quarks. The distinction between valence and sea quarks is only relevant for determining the probability that a specific quark would appear in a proton interaction. The probability that a particular quark or a gluon will participate in a proton interaction for two values of Q^2, the four-momentum transfer in the interaction squared, is shown in Figure 2.1. The leptons are the e, μ and τ, along with their companion neutrinos. The leptons do not interact via the strong force. The neutrinos only interact via the weak force, but the other leptons also interact with the electromagnetic force. All of these particles and their interactions can be seen in Figure 2.2.

Currently, leptons and quarks are the limit of the building blocks that have been observed.

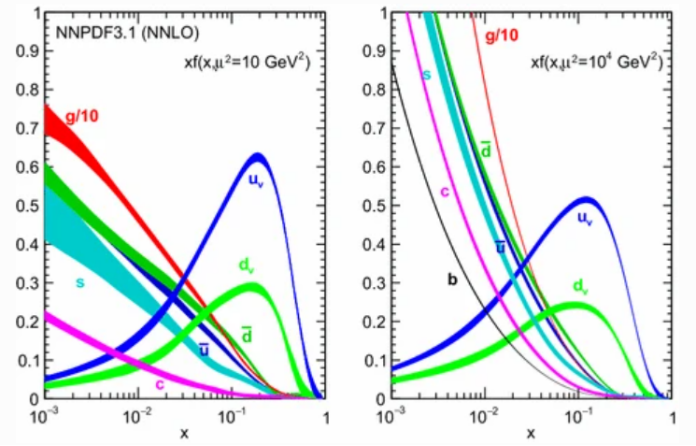

Figure 2.1: The probability that a given product will be involved in a proton interaction depends on the fraction of proton energy involved in the interaction. The parton distribution functions for the proton is shown for two different values of Q^2, the amount of four-momentum transfer squared in the interaction. [13]

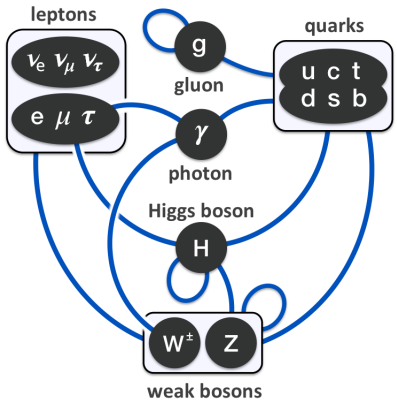

Figure 2.2: The particles of the standard model can be organized into quarks, leptons, and bosons. Quarks and leptons are fermions that are the building blocks of matter and the bosons are the force carriers that lead to interactions between the fermions. [14]

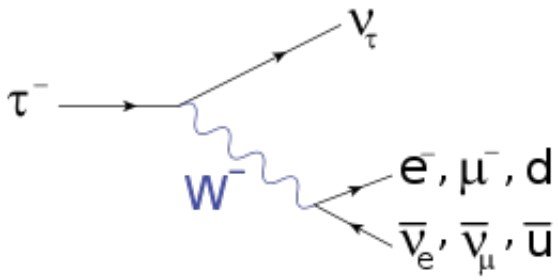

Figure 2.3: Feynman diagram depicting the standard model interaction of a τ decaying via the weak interaction into hadrons or lighter leptons and antineutrinos.

Leptons exhibit properties of point-like particles in the ways we can observe them and, therefore, do not seem to be composite. This analysis probes the substructure present in leptons, if any, and more specifically through the study of τ leptons, the heaviest of the leptons.

This dissertation focuses on τ leptons as an important part of the final state in the search, so it important to understand τ properties in more detail. The τ has a mass of 1.78 GeV and can decay via the weak interaction into a ν_τ in addition to hadrons or lighter leptons and their respective antineutrinos. The Feynman diagram for this interaction is shown in Figure 2.3. A summary of the decay modes of the τ can be seen in Table 2.1. The reason an analysis might separate τ leptons from the other leptons for an analysis is that the τs will decay within the volume of the detector and in so doing, some of the energy will leave undetected with the neutrinos and it will be difficult to reconstruct a mass peak from this. This analysis uses the collinear approximation to navigate around that problem. This technique will be described in more detail in Section 5.1.1.

2.4 Compositeness of Fermions

This analysis uses the formalism presented in [8]. Other sources that define and discuss the theory of compositeness can be found in Ref. [2, 3, 4, 5, 6, 7, 9, 10]. The Lagrangian for the

Decay mode	Branching fraction (%)
Leptonic	35.2
$\tau \rightarrow \mathrm{e}\ \overline{\nu}_{\mathrm{e}}\ \nu_{\tau}$	17.8
$\tau \rightarrow \mu\ \overline{\nu}_{\mu}\ \nu_{\tau}$	17.4
Hadronic	64.8
$\tau \rightarrow h^{-}\nu_{\tau}$	11.5
$\tau \rightarrow h^{-}\pi^{0}\ \nu_{\tau}$	26.0
$\tau \rightarrow h^{-}2\pi^{0}\ \nu_{\tau}$	10.8
$\tau \rightarrow h^{-}h^{+}h^{-}\nu_{\tau}$	9.8
$\tau \rightarrow h^{-}h^{+}h^{-}\pi^{0}\ \nu_{\tau}$	4.8
other	1.9

Table 2.1: List of decay modes for the τ. The final states of interest to the analysis will have at least one hadronic decay, represented here with an end product of h for hadron.

contact interaction is given by

$$\mathcal{L}_{CI} = \frac{g^{*2}}{2\Lambda^{2}} j^{\mu} j_{\mu} \tag{2.2}$$

where g^{*2} is chosen to be 4π, Λ is the compositeness scale and j_{μ} is the fermion current given by

$$j_{\mu} = \eta_{L} \bar{f}_{L} \gamma_{\mu} f_{L} + \eta'_{L} \bar{f}^{*}_{L} \gamma_{\mu} f^{*}_{L} + \eta''_{L} \bar{f}^{*}_{L} \gamma_{\mu} f_{L} + h.c. + (L \rightarrow R) \tag{2.3}$$

where f and f^{*} represent the SM fermions and excited fermions respectively, the η factors are scale factors and the γ_{μ} are the gamma matrices. For simplicity, the η factors are set to one for left-handed fermions and to zero for right-handed fermions. In addition to this contact interaction, the excited lepton may also undergo a gauge-mediated interaction governed by

$$\mathcal{L}_{GM} = \frac{1}{2\Lambda} \bar{f}^{*}_{R} \sigma^{\mu\nu} \left[g_{s} f_{s} \frac{\lambda^{a}}{2} G^{a}_{\mu\nu} + g f \frac{\tau}{2} W_{\mu\nu} + g' f' \frac{Y}{2} B_{\mu\nu} \right] f_{L} + h.c. \tag{2.4}$$

where Λ is the compositeness scale, f^{*}_{R} is the excited fermion, f_{L} is the standard model fermion, $\sigma^{\mu\nu}$ is an anti-symmetric rank two tensor used to represent fermion interactions, λ^{a} are the Gell-Mann matrices, $G^{a}_{\mu\nu}$ represent the SU(3) tensor field, τ represent the Pauli matrices, $W_{\mu\nu}$ represents the $SU(2)$ gauge field, Y is weak hypercharge, $B_{\mu\nu}$ is the U(1) gauge field, the g factors are known factors derivable from the mass of the W and Z bosons and the vacuum expectation value of the Higgs boson, and the f factors in the square brackets are

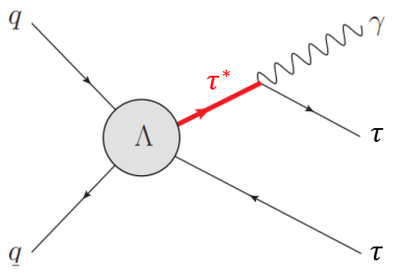

Figure 2.4: Feynman diagram representing excited τ produced in conjunction with a SM τ via contact interaction and its subsequent decay to τ and γ in the final state.

unknown scaling factors relating to the relative strength of the coupling between the bosons and the excited fermions. The first term here represents mediation by a gluon and the second and third terms represent an electroweak process mediated by a photon, W boson, or Z boson. Each of these terms is slightly modified from the standard processes by the inclusion of the excited fermion (f^*) and by each term's individual coupling factor (f_s, f, f').

Putting these possible types of interactions together, we arrive at the possibility of a final state with a photon and two τ leptons, as can be seen in Figure 2.4, which will be the focus of this analysis.

Other analyses have been conducted by CMS at $\sqrt{s} = 13\,\text{TeV}$ in 2016 data for electrons [15] and in 2016 and 2017 data for muons [16]. The limits from that search can be seen in Figure 2.5. Other experiments, such as LEP [17, 18, 19, 20], and ATLAS at $\sqrt{s} = 8\,\text{TeV}$ [21], have searched for excited leptons including τ leptons. None of these searches has found evidence for excited leptons. For τ^*, the most stringent limits set are at 2.5 TeV for a compositeness scale equal to the mass of the τ^*.

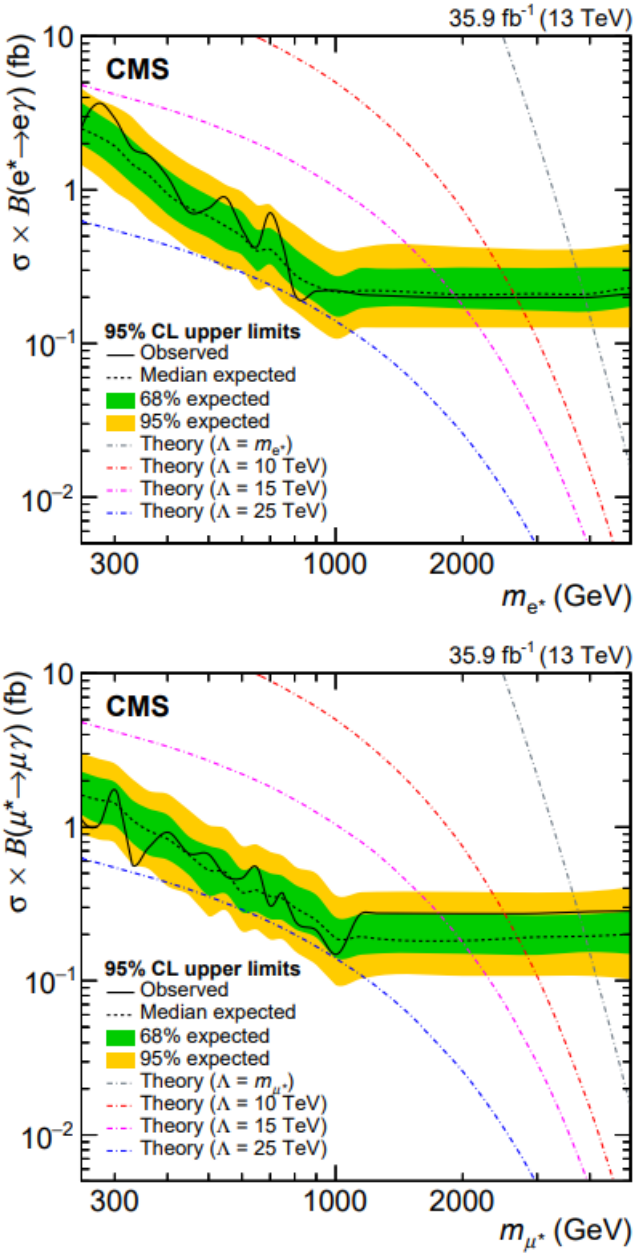

Figure 2.5: The limits from the search for excited electrons and muons are presented here. For the case where the mass of the excited lepton is equal to the compositeness scale, the limit is set at 3.9 TeV (electron) and 3.8 TeV (muon). [15]

Chapter 3

Experimental Apparatus

This chapter discusses the experimental apparatus pertinent to the analysis, including its broader context in collider physics. The chapter begins with a broad overview of collider physics, then moves to the Large Hadron Collider (LHC), the specific collider involved in the experiment, and then ends with an in-depth discussion of the detector setup of the Compact Muon Solenoid (CMS).

3.1 Coordinate System

In CMS, a particular local coordinate system is used to standardize the directions of particles passing through the experimental apparatus. The origin of this coordinate system is placed at the nominal interaction point where the proton beams collide (described in more detail in Section 3.4). The positive z-axis points parallel to the beam pipe counterclockwise when viewed from above. The x-axis points to the center of the LHC and the y-axis points straight up. Using these axes, it is possible to define angular measurements. The azimuthal angle ϕ is measured in the $x - y$ plane, such that $\phi = 0$ is coincident with the positive x-axis and $\phi = \pi/2$ radians is coincident with the positive y-axis. The radial component of this measurement is denoted as r. The $x - y$ plane is also referred to as the transverse plane because it is transverse to the direction of accelerating protons. This plane is very useful for

determining quantities such as momentum, since the initial momentum in this plane is zero, so the final momentum of all particles should balance. The same cannot be said in the z direction because it is the quarks and gluons inside the protons (described in more detail in the next chapter) carrying some unknown fraction of the proton's momentum that actually interact.

Furthermore, an angle can be defined in the $r - z$ plane and this angle is called θ. This angle is measured such that $\theta = 0$ is coincident with the r-axis and $\theta = \pi/2$ is coincident with the positive z-axis. Another variable is generally used in place of θ called pseudorapidity, or η, where $\eta = \ln(\tan(\theta/2))$. This variable proves useful because even particles with very high Lorentz boost in the lab frame, such as the colliding protons, are expected to have uniform distribution of outgoing particles in η on average after the collision.

3.2 Collider Physics

When discussing the search for new forms of matter such as an excited τ lepton, it is important to review the basics of collider physics, since any such particles will only be produced at the high energies achieved by particle colliders. There are two standard ways of constructing a collider experiment. The first is to accelerate a beam of particles into a target and the second is to accelerate two beams of particles and collide them with each other at specific interaction points. It is easier to construct the first kind of collider, but the second kind provides a much greater center-of-mass energy. This energy is the amount available to new particles created in the collisions. It can be related to the new particles' mass and momentum via

$$E^2 = m^2 + \vec{p} \cdot \vec{p}. \tag{3.1}$$

Therefore, in order to discover particles of high mass, it is important to increase the center-of-mass energy to a level where those particles can be produced.

Even with enough energy to produce a given particle, some processes occur more fre-

quently than others, so it is important to produce a large number of collisions in order to be able to observe rare processes of interest. The number of events of interest that are expected to be observed is given by

$$N_{event} = \mathscr{L}\sigma_{event}\epsilon \tag{3.2}$$

where $\mathscr{L}$ is the luminosity, σ_{event} is the event cross section, and ϵ is the detection efficiency.

The luminosity $\mathscr{L}$ is the rate of potential collisions within the detector per unit area and is given by

$$\mathscr{L} = \frac{N^2 f}{A} \tag{3.3}$$

where N is the number of particles contained in one of the beams that are being collided at one time, f is the frequency of the machine, and A is the area of overlap of the two beams during the collision. The beams are organized into bunches of particles that are collided at a specific frequency. Therefore, the number of interesting events can be increased by increasing the rate at which particle bunches that are densely packed pass through each other. To give a concrete example of the luminosity involved here, the total integrated luminosity, or the instantaneous luminosity from above integrated over time, of data collected by CMS is shown in Figure 3.1. The event cross section σ_{event} is the theoretical prediction of the frequency of occurrence of the desired event. The total cross section for all processes at the LHC is approximately 70 mb[1]. The cross section of the events of interest to this analysis are on the order of pb or fb depending on the exact parameters. The efficiency ϵ is a description of the fraction of events that the detector will detect depending on various factors like area coverage and reconstruction efficiency of the final products of the event.

[1]A barn (b) is a unit of area equal to 10^{-24}cm^2, approximately the cross-sectional area of a uranium nucleus

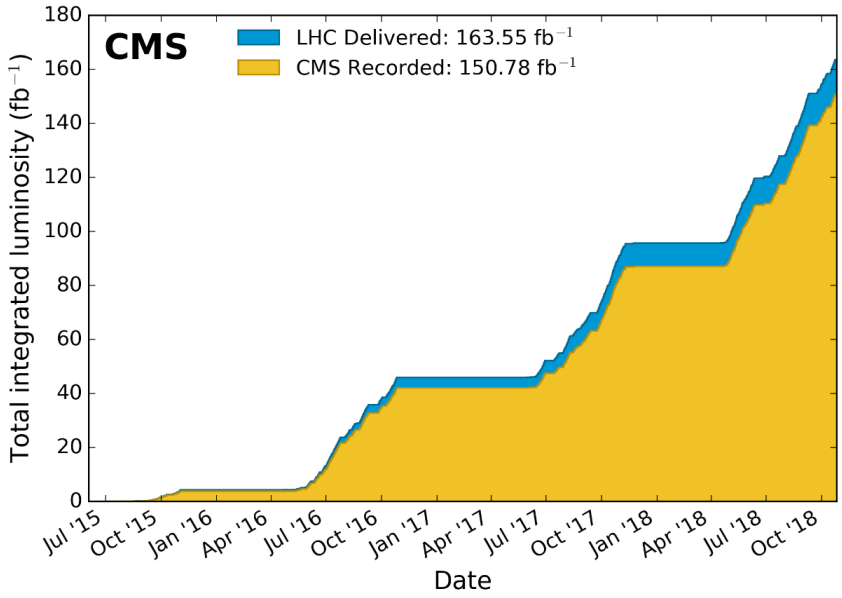

Figure 3.1: The integrated luminosity represents that amount of data collected by CMS. Image is available from CMS public results.[22]

3.3 Large Hadron Collider

The LHC is a synchrotron capable of colliding same-sign particles at various points around its roughly circular ring, including for the CMS experiment. The machine itself is located underground on the border of France and Switzerland at CERN near Geneva with a diameter of 26.7 kilometers. In 2010, the first collisions were performed at 7 TeV center-of-mass energy[2]. At the end of 2011, the beam energy was increased to 4 TeV per beam. The current center-of-mass energy of the collider is 13 TeV and the instantaneous luminosity is 10^{34} cm^{-2}s^{-1}.

The LHC accelerates beams of protons up to a design center-of-mass energy of 7 TeV using RF cavities tuned to 400 MHz. The bunches of protons are kept in a circular trajectory using superconducting dipole magnets and are focused to low cross sectional area by superconducting quadrupole magnets. The bunches of protons are tuned to collide with

[2]The center-of-mass energy is often referred to as $\sqrt{s}$ in reference to the Mandelstam variables.

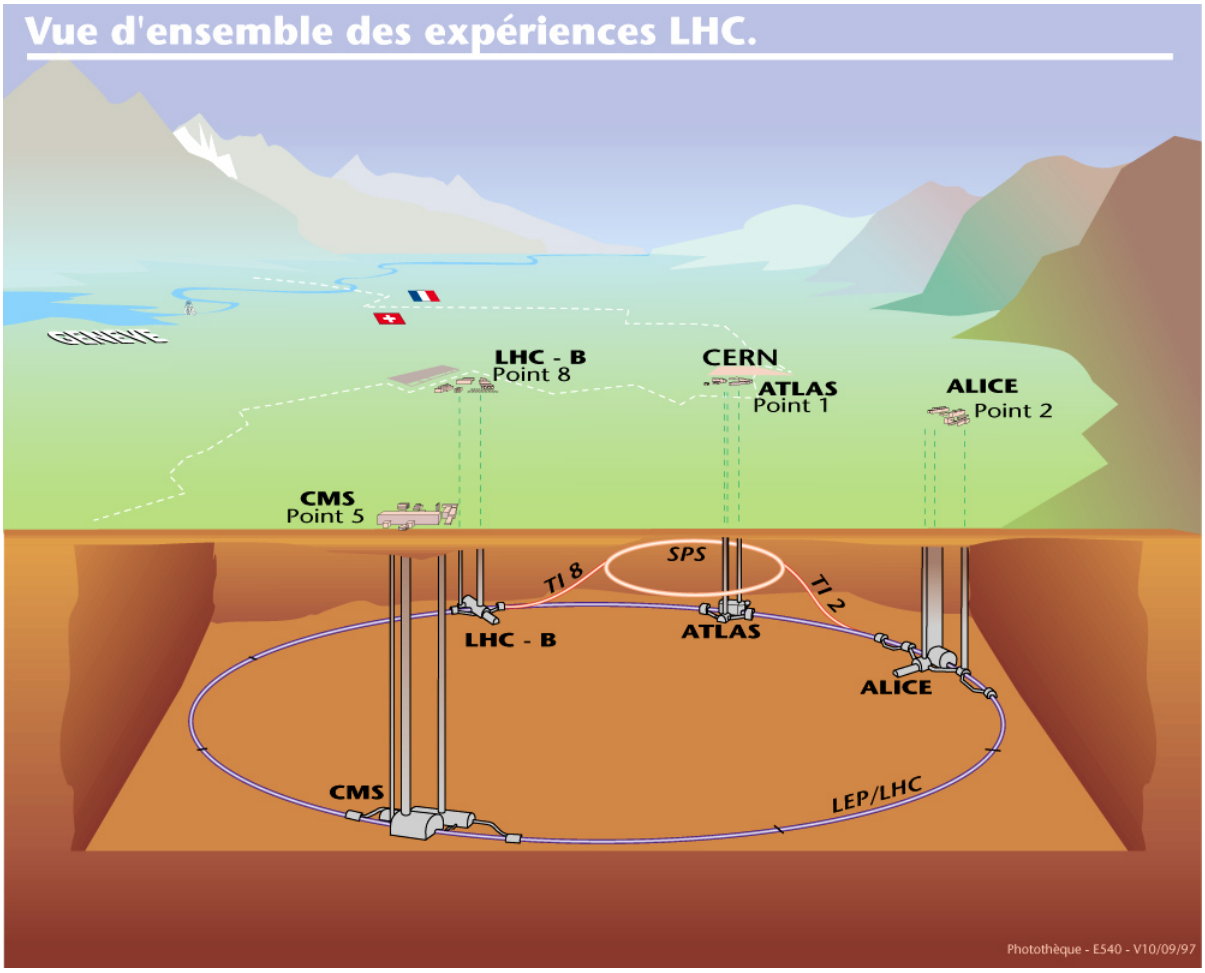

Figure 3.2: The LHC is a large circular underground ring that is 26.7 km in diameter and home to several experiments, such as CMS, ATLAS, ALICE, and LHCb. Image provided courtesy of CERN.[23]

each other at various interaction points throughout the ring every 25 ns. The next section describes the specific detector located at one of those interaction points involved with this analysis.

3.4 Compact Muon Solenoid

The Compact Muon Solenoid (CMS) detector is located at Point 5 of the LHC (described in Section 3.3). CMS has many subdetectors that are responsible for detecting various different types of particles produced in the collisions. Different particles interact in different ways and cannot be detected with one universal method. The following subsections describe the subsystems of the detector starting closest to the interaction point and working outwards. Figure 3.3 shows a slice of the barrel region of CMS and what it looks like in three dimensions. This figure also diagrams the various types of particles relevant to the analysis and which subsystems are responsible for measuring properties of those particles. The solid (dotted) lines represent particles with (no) electric charge.

3.4.1 Pixel Detector

The pixel detector provides "hits", or a pulse above threshold that is associated with charged particles passing through the sensors. The layout can be seen in Figure 3.4. It is composed of 4 layers in the barrel that extend out to 160 mm from the interaction point and 3 rings in the endcap that extend out to 50 cm in the forward and backward directions. The barrel layers are cylindrical and the endcap rings look like turbines because they have tilted blades as opposed to being a flat disk. The tilt of the modules in the rings optimizes radial and azimuthal resolution. Across all the layers and rings there are 1856 individual modules each with 66560 individual pixel sensors. The sensors are 100 microns $\times$ 150 microns and use an n-on-n pn-junction for signal detection that is 280 microns thick. This means that n-type silicon, which is a silicon lattice doped with atoms of an element with one more valence

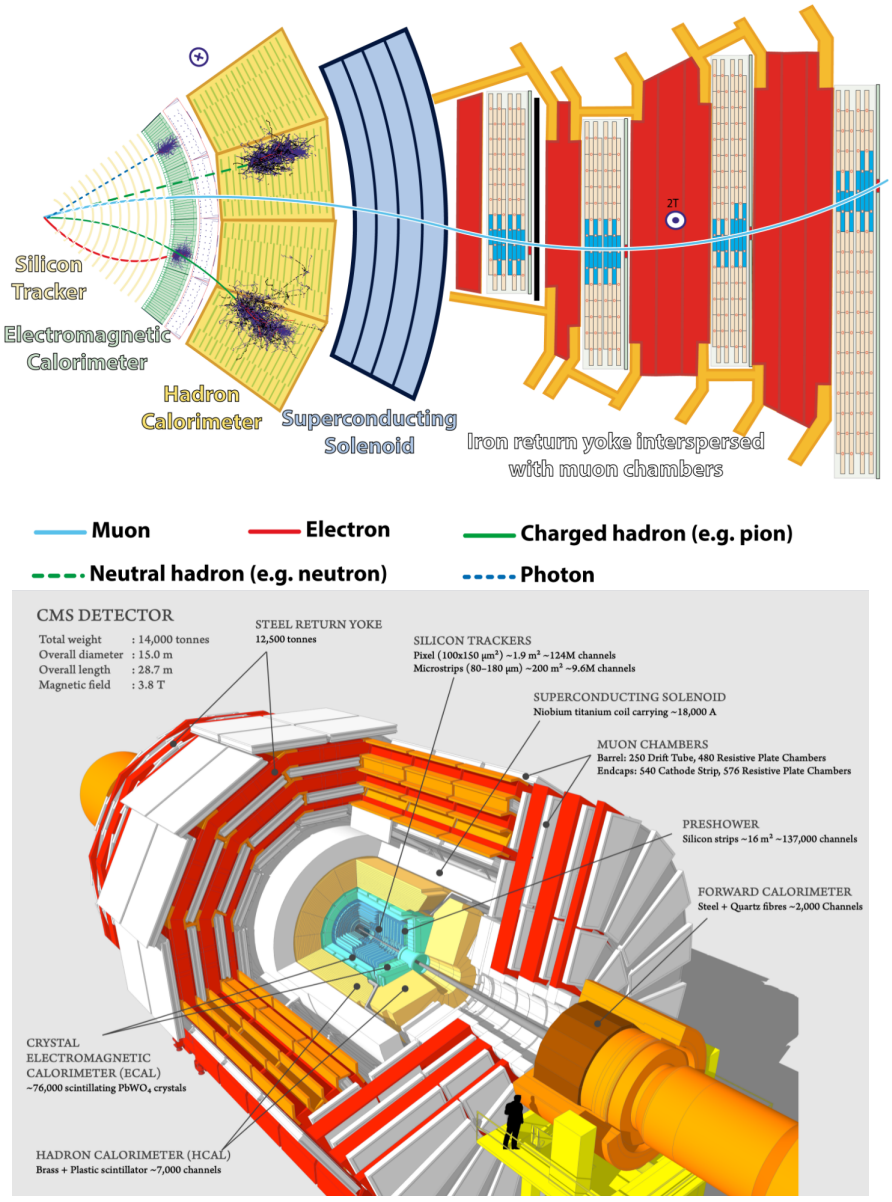

Figure 3.3: (Top) A slice through the CMS experiment [24]. The various subdetectors described in the chapter are shown starting with the innermost on the left moving to the outermost on the right. Example particle signatures are also shown. (Bottom) The CMS experiment in 3D [25]. This diagram demonstrates the sheer size of the experiment and how the subdetectors are arranged cylindrically around the beam pipe.

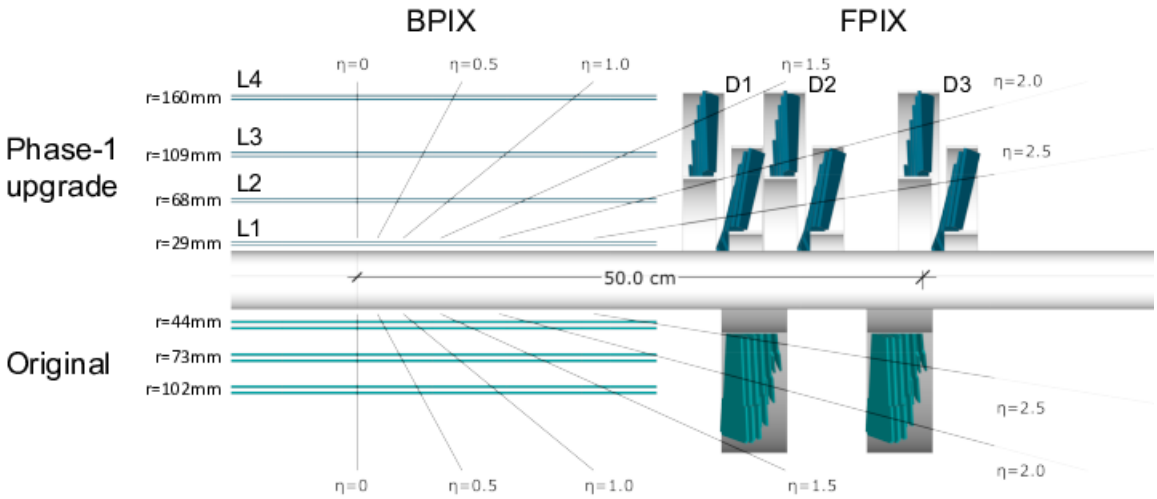

Figure 3.4: The layout of the CMS pixel modules [26]. The top half shows the current system with upgrades and the bottom half shows the original design.

electron, is used as the substrate, or the bulk of the sensor. Then highly doped implants are placed around the charge collecting aluminum strips. To actually create a pn-junction, a highly doped p-type silicon, or silicon doped with atoms with one fewer valence electron, is inserted in the backplane. Then the excess electrons in the n-type substrate migrate to the p-type backplane and combine with the holes (net positive charge that can migrate by accepting an electron and making a net positive charge in the atom it received the electron from) in that region. This creates a region where there are no free charges and this is called the depletion zone. Even though the charges are not free to move around, the original sharing of electrons between the p- and n-type silicon created a net electric field in the depletion zone. A bias voltage can be applied to the junction to increase the size of the depletion zone. The layout of the specific junctions used in the pixels can be seen in Figure 3.5. This can be used to detect ionization of the substrate when a charged particle passes through the material. The liberated electrons will migrate to the edge of the detector under the influence of the electric field. The holes will also migrate to the backplane. The charges also experience Lorentz drift from the magnetic field. Depending on the depth at which the charges are produced, they will experience a different amount of drift. The amount of drift can be larger than the size of an individual pixel. Charge interpolation allows for a final spatial resolution of 15-20 microns. The signal is based on the number of electrons collected at the n+ (highly doped n-type silicon) region guarding the aluminum strips. Signals greater than 1900e (electron charges) are read out as a "hit". The signals are read out by a 250-nm CMOS ASIC and undergo on-chip digitization using an 8-bit ADC. A readout chip (ROC) is responsible for an 80×52 region of pixels. When the ROC receives a message from the token-bit manager (TBM) that an event at a specific timestamp passed a trigger and should be recorded, it adds its hits to the message and passes the message on to the next ROC. When the loop through the ROCs is complete, the TBM sends the data on to the front-end driver (FED) at a rate of 400 Mbit/s. One TBM is responsible for 8 or 16 ROCs, which is either half a module or a whole module. The FED then puts together the hits from the

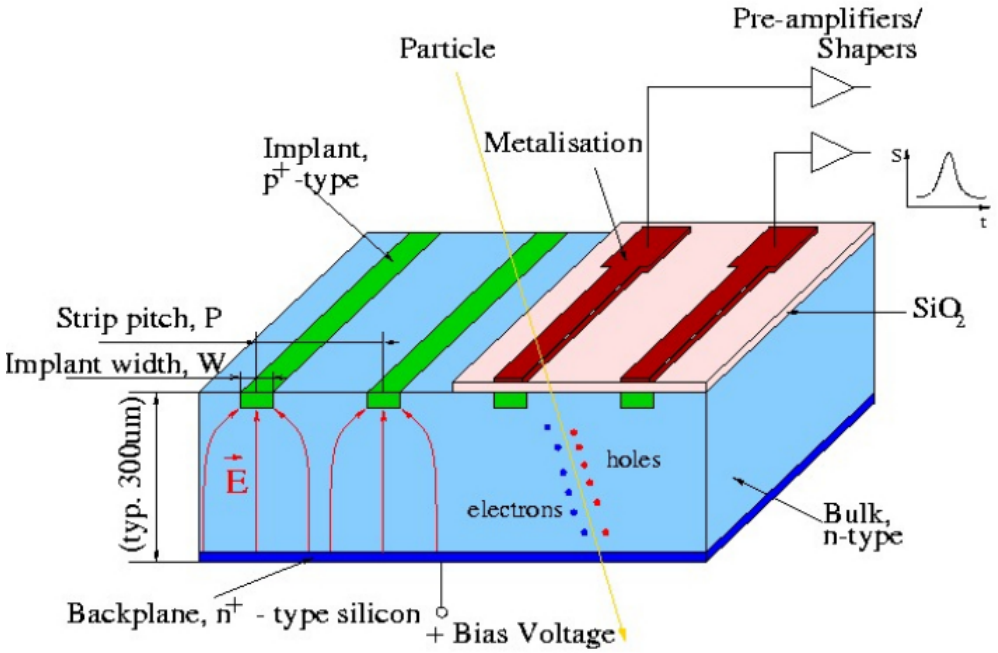

Figure 3.5: The n-on-n pn-junction is used in the pixel sensor to maintain high signal charge at moderate bias voltages up to 600 V [27].

different modules and sends all the data to the central CMS data acquisition system.

3.4.2 Silicon Strip Tracker

The silicon tracker performs a similar function to pixel detector in that it provides hits for charged particles passing through the detector. The main difference is the resolution and the volume of the detector. The detector is arranged in long strips that each have a readout channel instead of individual pixels with a readout channel. Fewer readout channels and greater spacing between the layers keeps the cost from becoming prohibitive by sacrificing some resolution. The sensors themselves are p-on-n type microstrip sensors. The layout of the strip tracking system can be seen in Figure 3.6. The total active area of the silicon tracker varies from 6243.1 mm^2 in the endcap to 17202.4 mm^2 in the outer barrel.

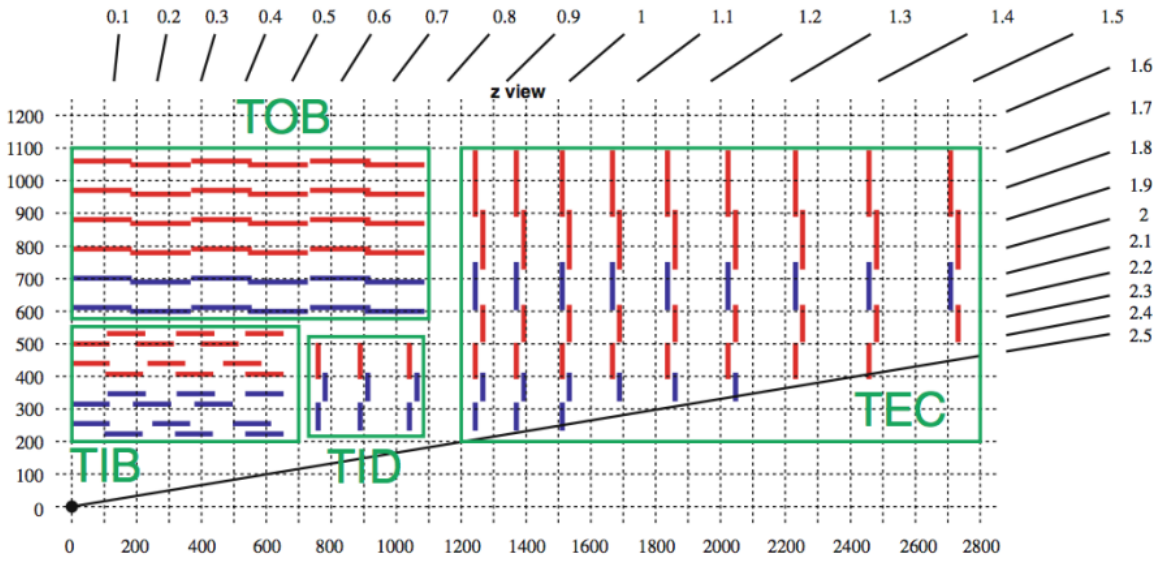

Figure 3.6: The layout of the CMS strip tracking system [28]. Red and blue strips indicate one-sided and two-sided strips respectively.

3.4.3 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) is composed of lead tungstate crystals that scintillate when ionized. This scintillation light is guided along the crystal to a photodetector that is calibrated to report the amount of energy deposited in the crystal by a particle. Particles that interact via the electromagnetic interaction and have relatively low mass will be stopped by the ECAL. These particles include electrons and photons. The electrons undergo bremmstrahlung, or "braking radiation", when they come near a nucleus in a material. This means that they exchange a photon with the nucleus and release a photon with some of their energy. Photons will convert into an electron and a positron (anti-electron) when they come close to a nucleus. Thus there is a cycle of photons produced from the electrons and electrons produced from the photons. Electrons can also deposit energy in the detector by ionizing the material and through Cherenkov radiation. Energy lost to ionization is most significant at low energies. Cherenkov radiation occurs for high energy electrons moving through the crystals at speeds faster than the phase velocity of light in the crystals.[29] This produces

what is called a shower. The scintillation light and Cherenkov radiation from all the various pieces depositing energy in the detector will be used to reconstruct the energy of the original photon or electron. Photons can be distinguished from electrons in this way because they do not produce tracks in the silicon tracking system described above. The very forward (or backward) region has a preshower that has better resolution for detecting the two photons that come from a π_0 and are often merged to look like one photon in the forward direction. The resolution of this detector is given by

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right)^2 + \left(\frac{12\%}{E}\right)^2 + (0.30\%)^2 \tag{3.4}$$

This corresponds to about 0.5% at 120 GeV.

3.4.4 Hadronic Calorimeter

The hadronic calorimeter (HCAL) is important for measuring energy from hadronic particles. These particles organize into jets which are discussed in Section 4.5. This calorimeter is also very important for determining the amount of missing transverse momentum (MET) that can be used to determine the presence of particles that CMS does not detect. The resolution of the detector is $\Delta E/E \approx 100\%/\sqrt{E[GeV]} \oplus 5\%$. The HCAL can be organized into four distinct pieces, shown in Figure 3.7. They are barrel (HB), endcap (HE), outer (HO), and forward (HF).

The barrel ($|\eta| < 1.3$) is made of alternating layers of brass and plastic scintillator that acts as a sampling calorimeter. The use of brass provides an increased number of nuclear interaction lengths, such that particles which experience the strong force undergo a nuclear interaction and deposit their energy in the calorimeter through ionization. The number of interaction lengths experienced by particles traveling through the calorimeter is equal to $\frac{5.82}{\sin\theta}$, where θ is the polar angle. This leads to a maximum of 10.6 interaction lengths at $|\eta| = 1.3$. The plastic scintillators produce light that must be wavelength shifted through secondary

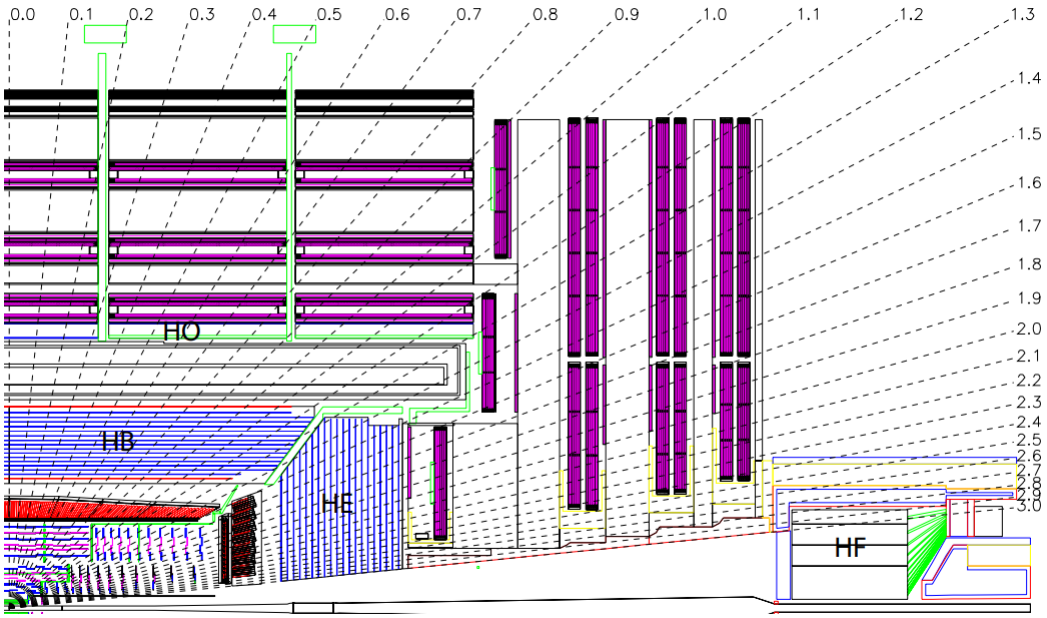

Figure 3.7: The layout of the various parts of the HCAL are shown here, including the barrel (HB), endcap (HE), outer (HO) and forward (HF) [30].

scintillation before being sent via fibres to a hybrid photodiode (HPD). The scintillator is segmented into units of 0.087 $\times$ 0.087 in η and ϕ.

The endcap covers the range $1.3 < |\eta| < 3$. This means that it must be radiation tolerant to experience the high particle flux at these angles and be non-magnetic to operate inside a large magnetic field. The absorber material was chosen to be C26000 cartridge brass. The material is arranged in a staggered geometry to prevent dead zones and to minimize cracks between HB and HE. The granularity of the energy towers in HE is 0.087 $\times$ 0.087 in η and ϕ up to $|\eta| = 1.6$ and 0.17 $\times$ 0.17 after that.

The HO subdetector covers $|\eta| < 1.3$ just as HB, but it is located outside of the solenoid to absorb any showers that could not be contained by the amount of material that particles had already passed through, including $\frac{1.4}{\sin\theta}$ interaction lengths from the solenoid itself. The presence of HO reduces the amount of shower leakage that would otherwise not be measured, which will in turn improve the calculation of missing transverse momentum.

The HF subdetector covers $3.0 < |\eta| < 5.0$ and must withstand incredibly high amounts

of radiation. At $|\eta| = 5$, the detector has experienced 3-4 MGy of radiation over an integrated luminosity of about 150 fb^{-1}. Therefore, quartz fibres were chosen as the active medium for their radiation hardness. These fibres are inserted into steel absorber. The detector measures the Cherenkov radiation produced from the particles through the quartz at phase velocities higher than the speed of light in quartz. This threshold is about 190 keV for electrons. Since the quartz is very radiation-hard, the activation of the absorber from high particle flux has little effect on the measurement quality of the quartz.

3.4.5 Superconducting Solenoid

The superconducting solenoid weighs 220 tons and spans 12.5 m along the length of the detector and cold bore diameter of 6.3 m. The diameter must be so large in order to house the calorimeters and tracking system, which are closer to the interaction point. The magnet is kept at 4.5 K during operation to be able to maintain a large field and stay in the superconducting regime. The solenoid provides a 3.8-T magnetic field at all points within its volume and approximately a 2-T field outside by employing a return yoke. This allows for the determination of the charge and momentum of charged particles by measuring the radius of curvature of the tracks made by charged particles.

3.4.6 Muon Chambers

The muon chambers provide hits that allow the reconstruction of the muon trajectories. The muons are not stopped in the calorimeters because they are much more massive than electrons and aren't moving fast enough to lose energy to ionizing the material in the ECAL and they do not interact via the strong force and so are not stopped in the HCAL either. There are three different types of chambers that are used in the muon system. They are drift tubes (DT), cathode strip chambers (CSC), and resistive plate chambers (RPC). Different chambers are used in different parts of the detector to balance detection efficiency and interaction with the magnetic field with the cost of the chambers.

Drift Tubes

Drift tubes are used in the barrel up to $|\eta| = 1.2$. The drift tubes include an anode wire stretched through the middle of the cell held at 3600V, electrode strips at the top and bottom of the cell held at 1800V, and cathode strips on the sides held at -1200V. The top and bottom of the cell are held at ground potential. When a muon passes through the cell and ionizes the gas (85/15% Ar/CO_2), the drift time provides a location measurement for the passing muon. The drift tube is used in the barrel because it is relatively inexpensive to manufacture and is expected to have good performance since the rate of muons is expected to be low in the barrel region and the magnetic field will not interfere with the detectors in this region since it is mostly contained in the steel yoke interspersed with the muon chambers. The layout of the individual cells of the drift tubes can be seen in Figure 3.9. The drift tubes are organized into 4 layers of tubes that are offset by half a cell from each other to provide redundancy in measurement and to provide better resolution. These sets of layers are called superlayers (SL). Depending on the orientation of the SL, the drift tube will measure either the z position along the beam pipe or the ϕ angle. There are 4 stations in each of the 12 sectors in the barrel of the detector. Each station has two SLs that measure ϕ and the inner three stations also have a SL that measures the z position.

Cathode Strip Chambers

The CSCs are used in the endcap region where the magnetic field is large and nonuniform. The CSCs cover a range of $0.9 \leq |\eta| \leq 2.4$. The actual layout can be seen in Figure 3.10. Each chamber contains alternating layers of anode wires and cathode strips. The chamber is put together in sets of seven panels, four cathode and three anode. This creates six gaps for the gas mixture of 40/50/10% $Ar/CO_2/CF_4$. A passing muon will create an avalanche near the anode wires that then creates an induced charge on the cathode strips. Through interpolation based on the induced charge, precise locations can be determined for the muon. This can be seen in Figure 3.11.

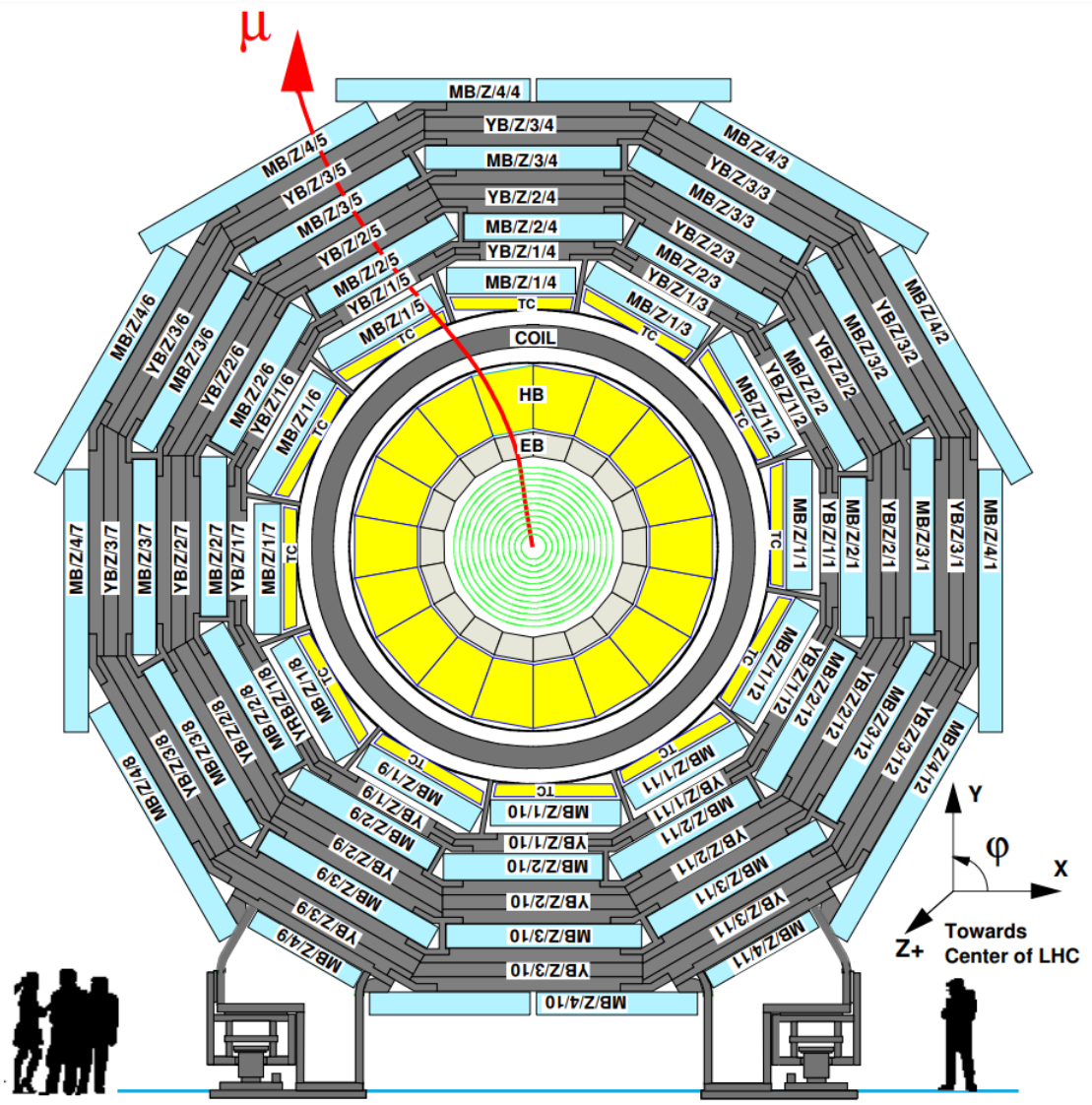

Figure 3.8: There are 12 sectors for the muon system. In addition there are 5 identical wheels that look like the one shown here that cover the length of the barrel [30].

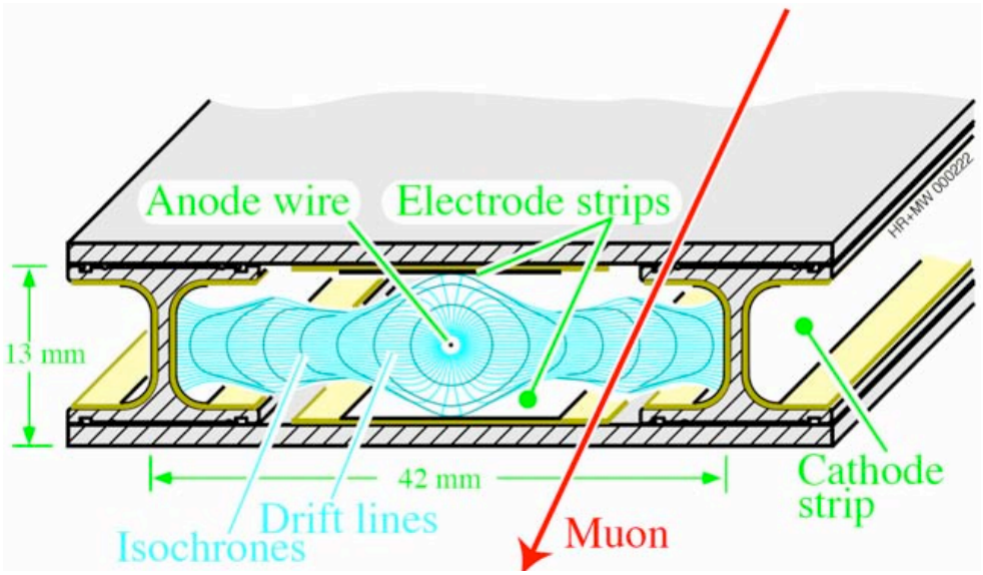

Figure 3.9: The drift tubes include an anode wire stretched through the middle of the cell held at 3600V, electrode strips at the top and bottom of the cell held at 1800V, and cathode strips on the sides held at -1200V. The top and bottom of the cell are held at ground potential. When a muon passes through the cell and ionizes the gas (85/15% Ar/CO_2), the drift time provides a location measurement for the passing muon [30].

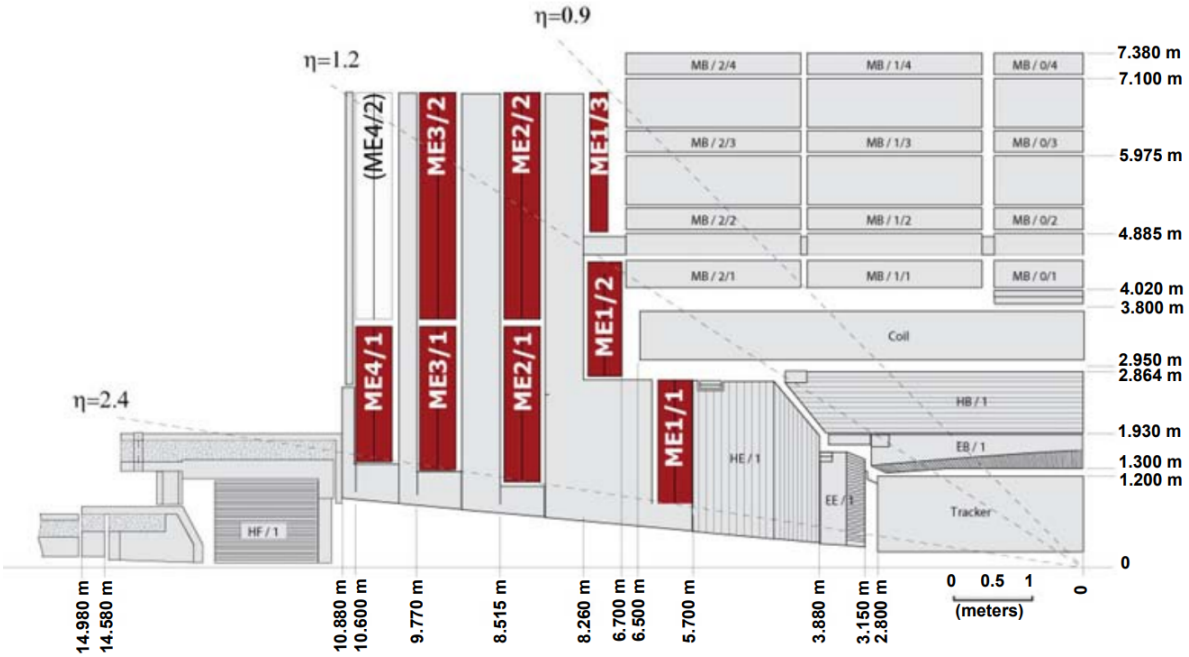

Figure 3.10: A slice through the detector parallel to the beam direction. The CSCs are highlighted in red [30].

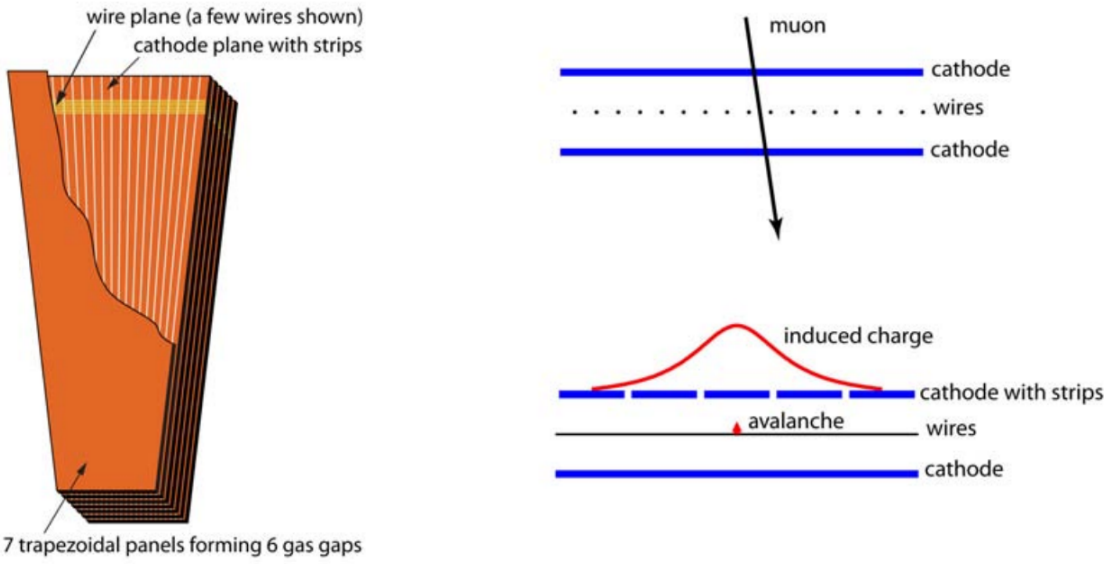

Figure 3.11: The left shows the trapezoidal CSC chambers with alternating layers of cathode strips and anode wires. The right shows the induced charge used for position measurement [30].

Resistive Plate Chambers

The RPCs provide good timing and momentum resolution for muons. Their quick response time allows for an unambiguous determination of the specific bunch crossing in which an interesting muon signature took place. The RPCs cover up to $|\eta| = 1.6$ and provide a good trigger even at the high rates of background at the LHC, which can get up to 1 kHz/cm^2. The RPCs can achieve a timing resolution of 1 ns with bunch crossings occurring every 25 ns. The RPCs themselves are composed of two gaps containing a gas mixture of 96.2% R134a ($C_2H_2F_4$), 3.5% iC_2H_{10}, and 0.3% SF_6 with pick-up readout strips in between. Water vapor is added to the gas mixture until the humidity of the gaps is about 45%.

3.4.7 Triggering and Data Acquisition

The trigger system is responsible for reducing the number of events that occur at the luminosity of the LHC to a number that can be saved and analyzed later. After an event passes the triggering step, it is saved to be processed further by the data acquisition system. The

trigger is organized into two levels, L1 and HLT. The Level-1 (L1) trigger reduces the data rate of 40 MHz (corresponding to the bunch crossing frequency of 25 ns) with 20 pp collisions per bunch crossing to a rate of 100 kHz through the use of programmable hardware such as field-programmable gate arrays (FPGAs), application-specific integrated circuits (ASICs), and look-up tables (LUTs). The FPGAs provide the flexibility to be reprogramed when necessary and the ASICs and LUTs provide speed and better density and radiation resistance requirements when located within the detector volume. The L1 trigger makes its decisions based on coarsely segmented data from the calorimeters and muon systems while holding the high-resolution data in memory in the front-end electronics. The latency allowed between a bunch crossing and the decision about that bunch crossing is 3.2 μs. The full trigger diagram for L1 Accept can be seen in Figure 3.14. The L1 trigger has local, regional, and global levels. There are also different components coming from the calorimeters and the muon system.

For the calorimeter, trigger primitives start as individual energy towers of a 5×5 region of crystals at the local level and then those are passed to the Regional Calorimeter Trigger (RCT). At the regional level, electron/photon candidates and transverse momentum sums are determined, as well as information about the isolation of muons and whether they meet standards for minimally ionizing particles. The electron/photon algorithm determines the tower with the largest energy deposit across the entire ECAL. Isolated and non-isolated candidates are limited to 8 trigger towers clustered together to minimize the impact of pile-up energy deposits. A total of 5 towers in the ϕ direction and 2 in the η direction are allowed. A shape veto is applied to discriminate electron and photon showers from jets. Isolation is then decided by looking at the 5 $\times$ 9 region around the seed tower excluding the footprint of the candidate (see Figure 3.12). Each region of 4 $\times$ 4 trigger towers sends four isolated and four non-isolated candidates to the Global Calorimeter Trigger (GCT). The GCT provides the jets, total transwerse momentum, missing transverse momentum, the scalar tranverse energy sum of all jets known as H_T, and a ranked list of isolated and non-isolated electron/photon candidates. Jets stemming from τs are determined using the

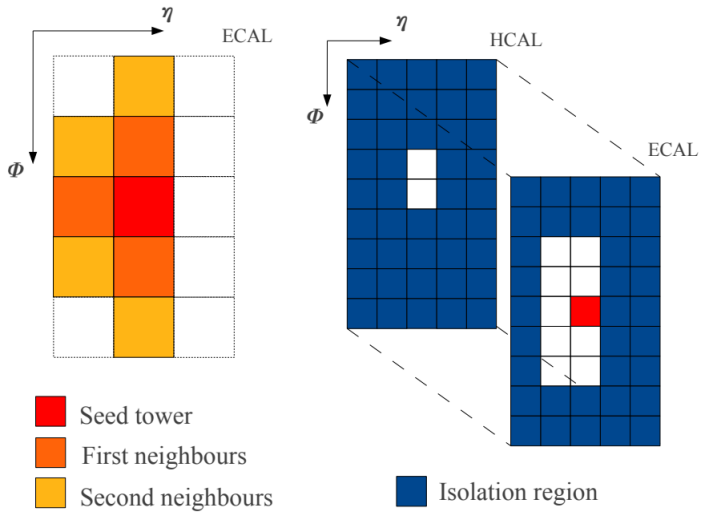

Figure 3.12: A depiction of the electron/photon algorithm. A shape veto is applied to discriminate from jets. The isolation is determined by thresholding the energy in the surrounding region [31].

HPS (hadrons-plus-strips) algorithm. The algorithm starts from the reconstructed jets and looks to reconstruct the energy from neutral pions by clustering the photon and electron constituents of the jets into strips of a dynamic size. The determination of these strip sizes is based on simulation and the desire to contain 95% of the decay products of the τ within the strip between a minimum strip size of 0.05 in $\Delta\eta$ and $\Delta\phi$ and a maximum size of 0.15 in $\Delta\eta$ and 0.3 in $\Delta\phi$. The algorithm makes a strip starting with the highest p_T electron or photon not already associated with a strip. The next highest constituent is then tested to see if it should be added to strip currently being constructed according to a formula shown graphically in Figure 3.13. If added, the strip location is then reweighted by p_T to include the new constituent.

For muons, all three subsystems described in Section 3.4.6 contribute to the trigger. The DTs provide track segments in the ϕ-projection and hit patterns in the η-projection. The CSCs provide 3D track segments. The Regional Muon Trigger (RMT) combines these track segments with the excellent timing of the RPCs to determine muon candidates. The track segments from individual stations are grouped together by extrapolating the trajectory

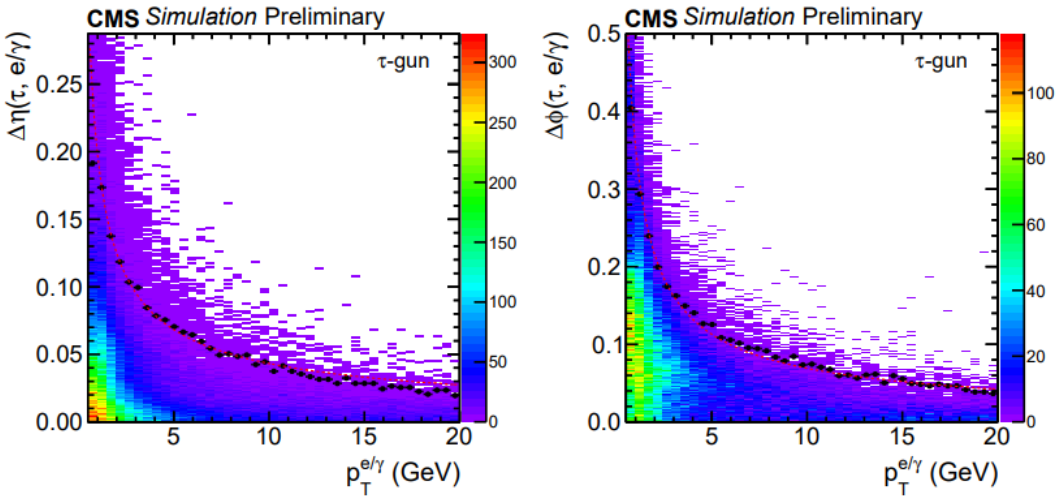

Figure 3.13: Simulation using a τ gun to determine the maximum distance allowed between an existing strip and a potential addition to that strip. The dotted points show the 95th quantile of the data and the line shows the final fitted equation [32].

originating at the vertex and passing through the source segment into another station. If there is a compatible segment in that station based on location and bending angle, then the segments are grouped together. The RMT then sends up to four candidates from the barrel and four fron the endcap on to the Global Muon Trigger (GMT) after determining the tranverse momentum and quality of the reconstructed tracks. The candidates with the highest transverse momentum and best quality tracks the sent on. The GMT then combines the tracks with the isolation information that comes from the calorimeters. The GMT then sends four candidates on to the Global Trigger.

The Global Trigger uses the information from the GCT and GMT to finally accept or reject an event. At this level, things like transverse momentum thresholding is applied to individual objects. A single trigger mask is applied and events passing the mask are accepted. When an accept trigger is reached, that is communicated to the sub-detectors through the Timing, Trigger and Control (TTC) system.

The data is then sent on to the HLT for more processing. The HLT (High Level Trigger) further reduces the rate to 1 kHz. The HLT is based in software. The HLT algorithms have

access to the full readout data and can, therefore, perform calculations similar to those that would be performed in an offline analysis setting on the reduced set of interesting events. The events that still pass the requirements of at least one of the many trigger conditions in the HLT are saved off to disk, assuming the trigger does not need to be prescaled. Prescaling is a process used to take only a fraction of events from a type of trigger that might produce a large volume of events. These HLTs can be used during analysis processing to select for events of particular interest to that analysis, such as a muon that passes a momentum threshold or the presence of a tau lepton meeting certain criteria, for example.

For this analysis, the many different decay modes of the tau leptons produce various possible final states (see Section 2.3). Since the HLT can select events that have specific characteristics such as types of particles in the event, many different triggers are used in the analysis to select the appropriate final state. Muons are the most efficiently reconstructed objects in the detector, so channels involving muons rely on the IsoMu24 or IsoMu27 trigger depending on the year. This trigger passes when there is an isolated muon that has a $p_T \geq 24$ or 27 respectively. The electron trigger is used for the electron+hadronic tau final state and is of the form Ele{p_T}_WPTight_Gsf, where {p_T} can be 25, 27, 32 or 35, depending on what trigger is available. Finally, for the two hadronic tau final state, the trigger is DoubleMediumCombinedIsoPFTau35_Trk1_eta2p1_Reg.

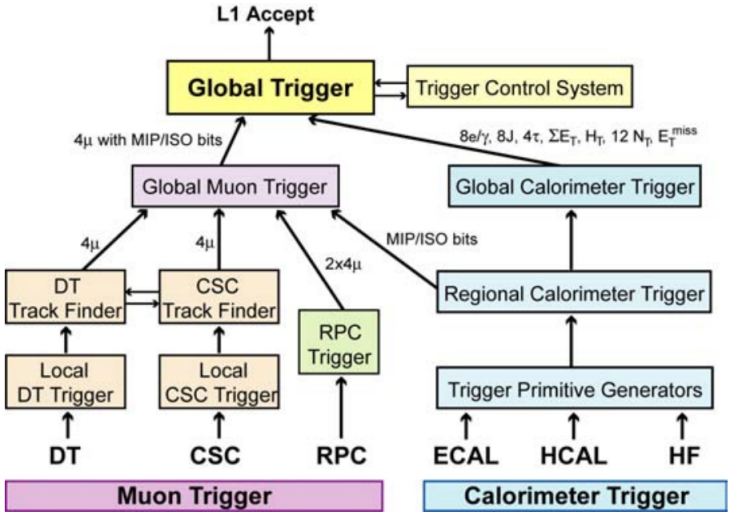

Figure 3.14: The possibilities for accepting an event at the Level-1 Trigger [30].

Chapter 4

Physics Event Reconstruction

With so many subdetectors in the CMS experiment, as described in the last chapter, a great deal of data can be gathered about what types of particles are traversing the detector. However, each subdetector individually would not provide a complete picture of the final-state particles that are the interesting features to analyses. Even after events have triggered the data acquisition system, much work is still required to reconstruct what happened in the event from the recorded data. The Particle Flow (PF) algorithm is central to identifying, reconstructing and calculating kinematic variables for the particles in an event. In previous hadron collider experiments, physics objects were mostly delegated to individual subsystems to identify. [33] However, this system can be greatly improved by looking at different subsystems simultaneously. Extrapolating connections (or lack of connection) between measurements in different subsystems can enhance the resolution of the kinematic measurements of the reconstructed objects. The concept for this algorithm was originally developed for the ALEPH experiment at LEP, which was an e^+ e^- collider. An e^+ e^- collider is a much cleaner environment since there aren't interactions happening that were not part of the hard scattering. With proton colliders, the debris from the parts of the protons that were not involved in the collision create many background signatures that must be separated out. Hadron collider experiments before CMS did not have the required spatial

resolution to resolve the particles of interest. CMS has a finely segmented ECAL, a hermetic HCAL, a strong magnetic field, and a high-resolution tracking system, which all combine to produce better particle identification. Every physics object reconstructed by CMS requires input from multiple detector subsystems. A "link algorithm" [33] is used to evaluate all PF elements simultaneously and identify groups of PF elements that are compatible with a physical object when combined. Elements are excluded from further consideration once they have contributed to a PF object. The following sections describe how various particles and interesting observables are reconstructed.

4.1 Charged particle tracking

In the original design of CMS [34], charged particle tracking was intended to be used to identify energetic isolated muons, energetic isolated τ decays, and secondary vertices related to b quark jets. Its function has since been expanded to include electrons and muons. Track construction is based on a Kalman filter [35] that reconstructs tracks in three stages: first, with initial seed generation from a few hits compatible with the trajectory of a charged particle; second, adding in hits from all tracker layers that match the trajectory; and finally, fitting the track to obtain the origin, p_{T}, and direction of the track. To reduce the overall number of tracks being kept for analysis, tracks must additionally have been seeded with hits from two consecutive layers in the pixel detector and have at least eight hits in total, none of which contributes more than 30% to the track's goodness-of-fit χ^2, and not having more than one layer along the track missing a hit. Finally, tracks are only saved if they originated within 3.5 cm of the beam axis and have $p_{\mathrm{T}} > 0.9\,\mathrm{GeV}$.

The efficiency of track reconstruction and the misreconstruction rate of this process was tested with Monte Carlo simulation. The efficiency is defined as the fraction of simulated tracks that are reconstructed and have at least 50% of the hits simulated for that track and at least 50% of the hits in the reconstructed track belonging to the associated simulated track.

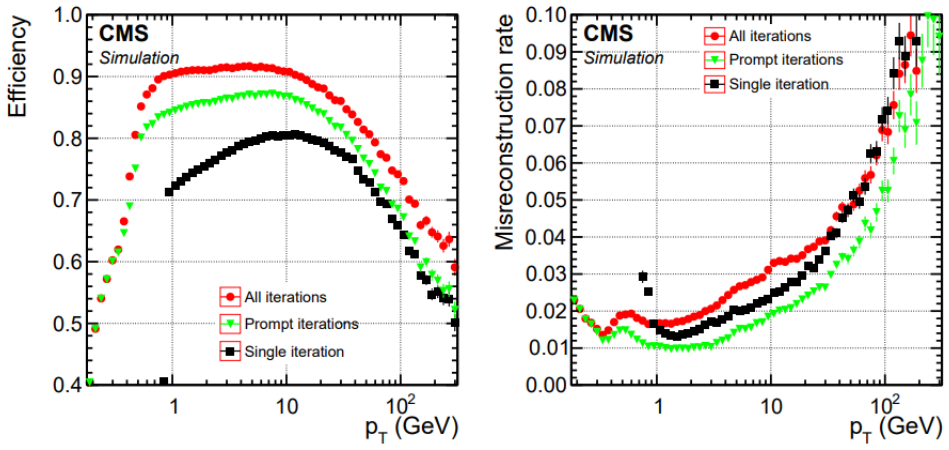

Figure 4.1: The tracking efficiency (left) and misconstruction rates (right) for a single iteration (black), all iterations involving seeds with at least one hit in the pixel detector (green), and all iterations (red). [33]

The misreconstruction rate is the fraction of reconstructed tracks that cannot be associated with a simulated track. The measured efficiency of this process is 99% for isolated muons and 70-80% for charged pions with $p_{\mathrm{T}} > 1\,\mathrm{GeV}$. This level of inefficiency would be unacceptable for the charged pions because those that were missed in tracking would be reconstructed as neutral hadrons that have a biased direction (or not reconstruct them at all). This would cause the PF algorithm to have poorer jet energy and angular resolutions by about 50%. To mitigate this inefficiency, the tracking is run in multiple iterations. The different iterations have different quality criteria applied to them based on what type of track is trying to be recovered in that iteration. All hits that are associated with a track in one iteration are removed from consideration in all future iterations. This process allows the loosening of criteria in later iterations to recover efficiency while maintaining a low misreconstruction rate. The efficiency and misreconstruction rates for this process can be seen for charged hadrons in Figure 4.1.

4.2 Muons

Muons are reconstructed from compatible tracks in the inner tracker and the muon detectors [36]. Tracking for muons can arise in three different pieces. There are standalone muon tracks resulting from fitting the hits in the DTs and CSCs into track segments and then using those as seeds to collect all the hits in the muon system associated with that muon's trajectory. Tracker muons result from tracks from the tracking system that have $p_{\mathrm{T}} > 0.5\,\mathrm{GeV}$ and total momentum greater than 2.5 GeV being extrapolated to the muon system. If at least one muon segment matches to the extrapolated track, the track qualifies as a tracker muon. Matching is performed by using a local coordinate system transverse to the beam direction. If the distance between the extrapolated track and the muon segment is less than 3 cm in the better measured coordinate or if the ratio of this distance to its uncertainty is less than 4, then the two are matched. Global muons are the combination of standalone muons with tracker muons when they have compatible tracks. Global muons improve the momentum resolution at large p_{T}. Matching tracks measured in the inner tracker and the muon detectors results in a relative p_{T} resolution, for muons with p_{T} up to 100 GeV, of 1% in the barrel and 3% in the endcaps, and of better than 7% in the barrel for muons with p_{T} up to 1 TeV [36]. Some muons will not be reconstructed as global muons if they undergo multiple scattering in the steel return yoke, which is common for muons with $p_{\mathrm{T}} < 10\,\mathrm{GeV}$. Even with this, approximately 99% of muons within the acceptance of the muon system are reconstructed as either global muons or tracker muons.

After tracks have been established, muons are the first objects to be identified by the PF algorithm and their signatures are then removed from further consideration. Additional track fit and matching quality criteria suppress the misidentification of hadronic showers that punch through the calorimeters and reach the muon system. Isolated global muons are only accepted if the sum of the p_{T} of the inner tracking tracks and the calorimeter energy deposits within ΔR of 0.3 is less than 10% of the muon p_{T}. For nonisolated global muons, the tight muon selection [36] is applied and there must be three track segments in the muon

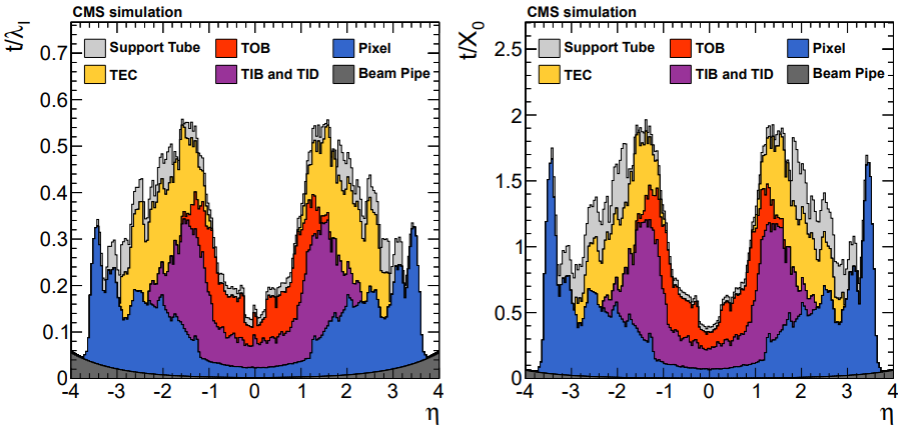

Figure 4.2: The amount of material of the tracking system is shown in terms of nuclear interaction lengths (left) and radiation lengths (right) as a function of pseudorapidity. The probability of an interaction leads to a loss of efficiency in the tracking. [33]

system or calorimeter deposits consistent with the muon hypothesis. Standalone muons may also be selected if they have a high-quality track associated with many hits in the muon system.

4.3 Electrons

Electrons are reconstructed by geometrically matching charged-particle tracks from the tracking system with clusters of energy deposited in the ECAL. [37] The electron momentum is estimated by combining the energy measurement in the ECAL with the momentum measurement in the tracker. The momentum resolution for electrons with $p_{\mathrm{T}} \approx 45\,\mathrm{GeV}$ from $\mathrm{Z} \to \mathrm{ee}$ decays ranges from 1.7 to 4.5%. It is generally better in the barrel region than in the endcaps, and also depends on the bremsstrahlung energy emitted by the electron as it traverses the material in front of the ECAL. Figure 4.2 shows on the right-hand side the number of radiation lengths that an electron would have to traverse before even reaching the ECAL. Since a large fraction of electrons will emit bremsstrahlung photons before reaching the ECAL, the energy of the electron is measured in what is called a supercluster of ECAL clusters that have a small range in η and a large range in ϕ to account for bending of the

trajectory from the magnetic field in the azimuthal direction. This reconstruction method assumes that the electron is well isolated from other particles, which does not hold for electrons in jets. To recuperate efficiency for these types of electrons, as well as for low-p_{T} electrons that are bent significantly by the magnetic field, electrons are also seeded in the PF algorithm with charged-particle tracks, as described in the previous section. The later iterations of the tracking can recover these electrons with their reduced requirements for the tracks, and all tracks with $p_{\mathrm{T}} > 2\,\mathrm{GeV}$ are used to seed the electrons.

Electrons can be distinguished from charged hadrons because they are much more likely to radiate in the tracker. When they don't radiate very much in the tracker, they can still be distinguished from charged hadrons because they will have a track with a well behaved χ^2 and a corresponding ECAL energy deposit. These kinds of electron seeds are required to have a momentum measured from the track and an energy deposit in the ECAL close to unity to be considered. When the electrons do radiate, the pattern recognition may not be able to piece together the whole track or the fit may have a large χ^2. These kinds of situations are fit again with a Gaussian-sum filter (GSF), which is better suited to dealing with sudden loss of energy (as by the emission of a photon) than a Kalman filter. Figure 4.3 shows how efficiently electrons can be reconstructed through this process and how often charged hadrons are confused for electrons.

4.4 Photons

Photons are reconstructed similarly to electrons, but they lack the charged tracks from tracking system, so the energy is just determined by the shower in the ECAL. However, the tracker-based seeding as described for the electrons is also useful for photons because the electrons and positrons resulting from conversions in the tracker material can be correctly associated with the parent photon using a dedicated conversion finder. A prompt photon can then be distinguished from a bremsstrahlung photon by summing the momenta of the

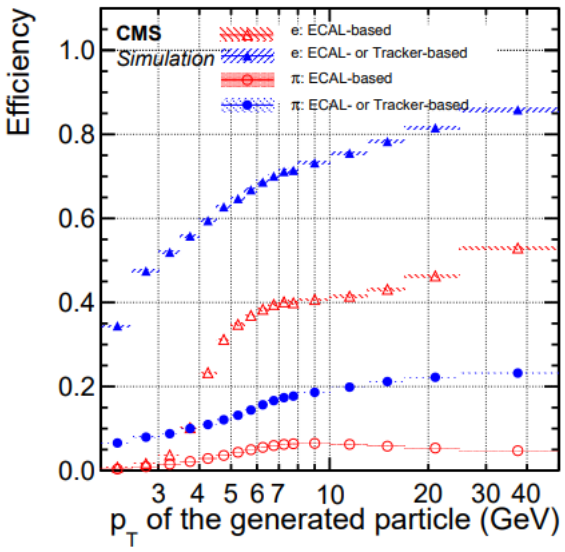

Figure 4.3: The efficiency for electrons and charged hadrons to give rise to an electron seed with and without the tracker-based seeding added for a sample of b quark jets with $80 < p_{\mathrm{T}} < 170\,\mathrm{GeV}$ and a semileptonic decay.

conversion tracks and seeing if the resulting trajectory coincides with a track tangent of a different GSF track. If it does, it will be linked to that track as part of the electron shower. In the end, an isolated photon candidate is seeded from an ECAL supercluster with transverse energy greater than 10 GeV that does not have a link to a GSF track and is isolated from other tracks and calorimeter clusters. Additionally, the HCAL cells within 0.15 in the (η, ϕ) plane must not have more than 10% of the ECAL supercluster energy when summed together. The energy finally assigned to the photon is the energy of the ECAL clusters corrected by analytical functions of E and η to account for energy not correctly associated to the object.

4.5 Jets

After the muons, electrons, and photons have been reconstructed and their signatures have been masked by the PF algorithm so as not to be used in later reconstructions, the particles that remain to reconstruct are the hadrons. The process of hadronization leads to many

hadrons being created in close proximity to each other. The signature of these hadrons is energy clusters in the ECAL and HCAL. These hadrons are clustered in jets using the anti-k_{T} algorithm [38] with a distance parameter of 0.4, as implemented in the FastJet package [39]. Jets are composite objects made up of several particles, hence the momentum is determined as the vectorial sum of all particle momenta. Additional pp interactions within the same or nearby bunch crossings can contribute additional tracks and calorimetric energy depositions, increasing the apparent jet momentum. To mitigate the effect of the charged particle contribution from pileup (uninteresting products not involved in the hard scatter) on reconstructed jets, a charged hadron subtraction technique is employed, which removes the energy of charged hadrons not originating from the PV [33]. In addition, the impact of neutral pileup particles in jets is mitigated by an event-by-event jet-area-based correction of the jet four-momenta [40, 41, 42]. Aside from pileup contamination removal, additional quality criteria are applied to each jet to remove those potentially mismeasured because of instrumental effects or reconstruction failures [43].

4.6 Missing Transverse Momentum (MET)

The vector $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is defined as the negative vector p_{T} sum of all the PF candidates in an event, and its magnitude is denoted as $p_{\mathrm{T}}^{\mathrm{miss}}$ [44]. The $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is computed from the PF candidates weighted by their probability to originate from the PV, and is modified to account for corrections to the energy scale of the reconstructed jets in the event. Jet energy measurements have a resolution of approximately 15%, so that dominates the uncertainty on the $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ measurement. After applying these corrections, this observable is referred to as the Type-I Missing Energy Transverse (or Type-I MET). However, this name is a misnomer because one of the major components of this variable is that it is a *vectorial* sum of components. This dissertation will, therefore, use the more modern term of $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ for the vector and $p_{\mathrm{T}}^{\mathrm{miss}}$ for its magnitude to reference the quantity of momentum on which it is based.

4.7 Hadronically decaying Taus

Hadronically decaying τ lepton candidates (τ_{h}) are reconstructed from jets using the hadrons-plus-strips algorithm [45], which combines one or three tracks with energy deposits in the calorimeters, to identify the corresponding one- or three-prong τ lepton decay modes. The main τ decay modes can be seen in Table 2.1. Here the number of prongs refers to the number of charged hadrons involved in the decay. Neutral pions from τ lepton decay are reconstructed as strips with variable size in η-ϕ from reconstructed electrons and photons, where the ϕ is azimuthal angle in radians and the strip size varies as a function of the p_{T} of the electron or photon candidate. The decay of a τ lepton differs from generic hadronization in the multiplicity, isolation and collimation of its decay products. Because the products are the result of a decay, the momentum of the original τ determines the boost of the decay products, which leads to the more characteristic features.

Chapter 5

Background and Signal Modeling

Models can be used to create artificial events that ideally have the same properties as data events collected through collisions. Understanding the properties of a particular process being modelled is important for the success of the model. This chapter details important properties of the signal used for the analysis and describes how the simulated events used in the analysis (both signal and background) were generated.

5.1 Signal Properties

Signal events are simulated using PYTHIA [46] with a compositeness scale of $\Lambda = 10\,\text{TeV}$ and τ^* masses ranging from 175 GeV to 5000 GeV. Other small datasets were also generated for a few compositeness scales equal to the mass of the τ^* and values for f and f' other than 1 that showed no significant deviations in the properties described in this section, and therefore, all limits at different compositeness scales are simply calculated from the limits found at $\Lambda = 10\,\text{TeV}$.

5.1.1 Collinear Approximation

As mentioned in Section 2.4, the τ leptons produced in the signal events decay into various final states, including electrons, muons, and hadrons (pions). Each decay also includes at least one neutrino. The neutrino is not detected by any of the subdetectors of CMS and its momentum contributes to the missing tranverse momentum $p_{\mathrm{T}}^{\mathrm{miss}}$, also called MET (for missing energy transverse). If it is assumed that all the missing momentum in an event comes from the τ neutrinos, then the MET is simply a sum of those momenta. The collinear approximation states that given a sufficiently high Lorentz boost (or momentum) of the τ lepton, the resulting neutrino's trajectory will be essentially parallel to the original τ direction. The simulation was checked for the validity of this assumption. The ΔR between the τ and its daughter neutrino is well within the reconstruction cone of the τ, so the approximation can be used. To further give a sense for the benefit of applying this approximation in our analysis, we turn to the properties of the individual τs in the event. Because there are two essentially indistinguishable final state τs, an ambiguity arises over which one to pair with the photon. The proximity of the correct pair is not uniform for all mass points. The momentum of the τs is also not sufficient to narrow down the pairing. Therefore, the correct pairing can only be obtained by making both pairings and selecting events that are sufficiently close to the correct mass in one of the pairings. The magnitude of each neutrino transverse momentum is calculated according to the following equations:

$$\nu_0 = |\mathrm{MET}|\frac{\sin(\phi_M - \phi_1)}{\sin(\phi_0 - \phi_1)} \tag{5.1}$$

$$\nu_1 = |\mathrm{MET}|\frac{\sin(\phi_0 - \phi_M)}{\sin(\phi_0 - \phi_1)} \tag{5.2}$$

The subscripts 0,1, and M represent each of the two leptons and the MET, respectively. These equations are simply the result of solving the linear system of equations to decompose the MET along the two tau directions. The four-momenta of the neutrino (assuming a

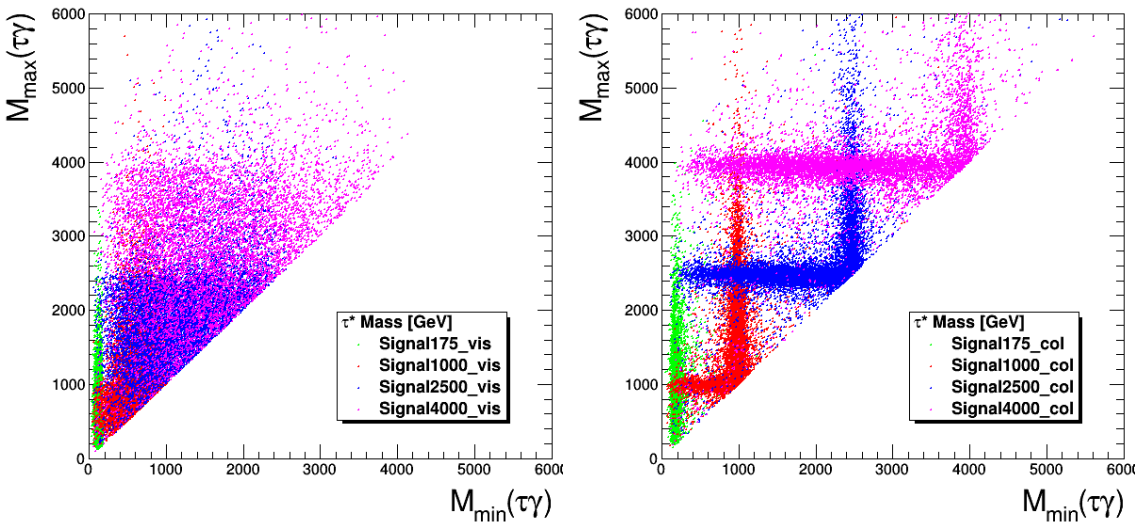

Figure 5.1: 2D mass plot of both pairings of the photon with a τ before (left) and after (right) the collinear approximation is applied in the e+τ_{h} channel.

massless neutrino) is added to its respective tau when applying the collinear approximation. The two mass combinations are then obtained by adding each modified tau four-momentum to the photon four-momentum and finding the mass from the equation $\mathrm{m}^2 = \mathrm{E}^2 - \vec{p} \cdot \vec{p}$. The 2-dimensional mass plot obtained from doing this is shown in Figure 5.1. The need for the collinear approximation is apparent from these graphs based on the much more sharply defined signal shape after applying the approximation and restoring the missing momentum. Another example of a graphical representation of this can be seen in Figure 5.2, where the difference between the simulated momentum and the reconstructed momentum is plotted. After applying the approximation, that distribution is centered around 0 and more sharply peaked, allowing for a more accurate and precise reconstruction of the τ^* mass.

5.1.2 Important Signal Properties

In making the some of the decisions in Section 6, the signal properties informed the method of making the decision. For example, the signal events have a high-p_{T} photon. Other important properties include the expectation of an oppositely charged τ pair. From simulation, we expect less than 10% of our signal events to be misreconstructed as same-sign events. We

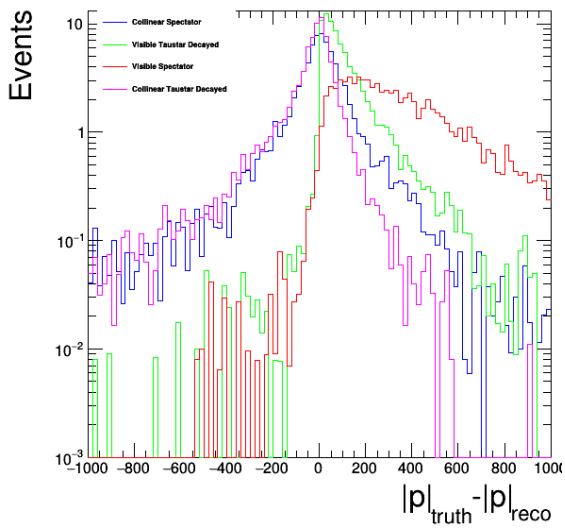

Figure 5.2: Momentum difference between simulated τs and their reconstructed counterparts for a τ^* mass of 175 GeV in the $e + \tau_h$ channel, with and without the collinear approximation applied. The corrected curves are more sharply peaked and centered around zero, given a better representation of the τ momentum.

also use b tagging to discriminate from some of our backgrounds. When choosing a working point for eliminating b tags, we considered both the medium and the tight working points. The final choice came down to maintaining as much relative signal as possible.

The final key feature to discuss is the selection of the binning of the 2D mass plane. Because the signal has such a distinct shape in this plane, we divide the plane up into bins based on that shape, as shown in Figure 5.3. Bin 2 is then taken as the Accept region and all other bins are combined into a Reject region. The width of Bin 2 is a free parameter to choose and we choose 10% of the hypothesis mass to either side of the central mass based on the behavior of the signal. The acceptance rate for choosing different bin widths is shown in Figure 5.4. A balance must be struck between acceptance of signal vs. acceptance of other backgrounds. Backgrounds are discussed in more detail in Section 6.9.

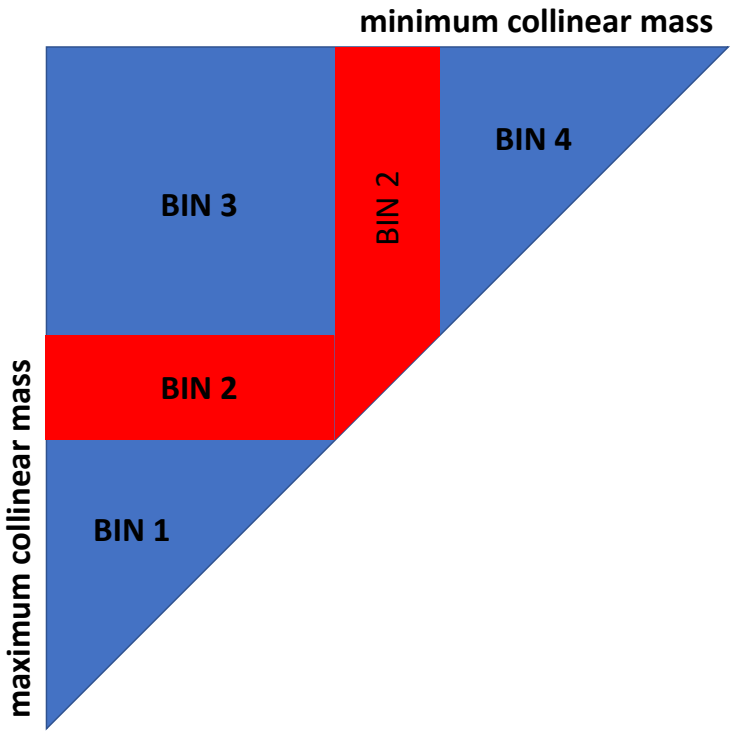

Figure 5.3: Binning scheme used to increase signal sensitivity.

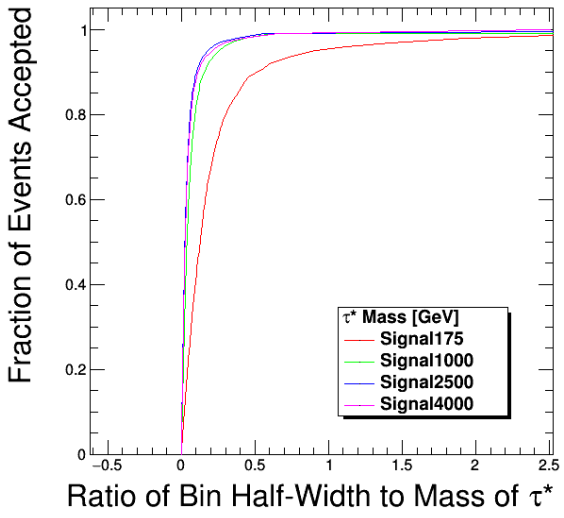

Figure 5.4: Scan of acceptance rate for various bin widths in the e + τ_h channel.

5.2 Data and simulated samples

The data corresponds to 138 fb^{-1} collected between 2016 and 2018 and recorded by the CMS experiment in pp collisions at $\sqrt{s} = 13\,\text{TeV}$. Monte Carlo simulation is used to model the Standard Model backgrounds that contain real τs in the final state and other leptons that are misreconstructed as τ_{h}. Backgrounds involving fake taus arising from jets are estimated using data-driven methods, as described in Section 6.9.

MadGraph5_aMC@NLO [47] 2.6.5 at LO using the MLM [48] jet matching scheme is used to generate Drell-Yan samples and W + jets samples. $\text{Z}\gamma$ samples are generated using MadGraph5_aMC@NLO 2.6.5 at NLO and using the FxFx [49] jet matching scheme. Samples of $\text{t}\bar{\text{t}}$ and single top s-channel samples are generated using POWHEG 2.0 [50, 51, 52]. Single top t-channel samples are produced using POWHEG 2.0 in combination with MadSpin [53], a package provided as part of the MadGraph5_aMC@NLO framework. WW, ZZ, and WZ samples are produced using PYTHIA 8.240.

Parton distribution functions (PDFs) from the NNPDF3.1 next-to-next-to-leading-order (NNLO) set [13] are used for all background and signal samples. PYTHIA 8.240 is used to perform parton showering, fragmentation and hadronization in all simulated samples with the CP5 [54] event tune. The simulation of the response of the CMS detector to incoming particles is performed using the GEANT4 toolkit [55]. Additional inelastic pp interactions from the same or nearby bunch crossings (pileup) are simulated and incorporated in the MC samples.

Table 5.1: Cross section for various signal sample masses.

Mass of τ^* [GeV]	Cross section [pb]
175	2.90E-02
250	2.16E-02
375	1.56E-02
500	1.18E-02
625	9.14E-03
750	7.07E-03
1000	4.26E-03
1250	2.55E-03
1500	1.53E-03
1750	9.07E-04
2000	5.37E-04
2500	1.87E-04
3000	7.51E-05
3500	2.72E-05
4000	9.88E-06
4500	3.63E-06
5000	1.36E-06

Chapter 6

Event Selection

The reconstruction and identification of individual particles in an event is based on the particle-flow (PF) algorithm [33], with an optimized combination of information from the various elements of the CMS detector. In each event, the candidate vertex with the largest total physics-object p_{T}^2 is taken to be the primary pp interaction vertex (PV). The energy of photons is obtained from the ECAL measurement. The energy of electrons is determined from the electron momentum at the PV as determined by the tracker, the energy of the corresponding ECAL cluster, and the energy sum of all bremsstrahlung photons spatially compatible with originating from the electron track. The momentum of muons is determined from the curvature of the corresponding track, and the energy is obtained from the momentum. The energy of charged hadrons is determined from a combination of their momentum measured in the tracker and the matching ECAL and HCAL energy deposits, corrected for the response function of the calorimeters to hadronic showers. Finally, the energy of neutral hadrons is obtained from the corresponding corrected ECAL and HCAL energies.

6.1 Electrons

Electron identification is achieved using a multivariate approach (MVA) which combines multiple parameters related to the electron reconstruction (e.g. isolation) into a single vari-

able to discriminate signal (prompt) from background (non-prompt or unmatched) electrons. Selected electrons are those which pass a cut value corresponding to 90% signal efficiency. A scale factor, further discussed in Section 7.1, is applied to the simulation to correct the MVA tagging efficiency compared to that in data. The electron is required to have a minimum p_T of 20 GeV and fall within the silicon tracker acceptance of $|\eta| < 2.5$. The electron is required to not fall within the small gap region between the ECAL endcap and barrel ($1.444 < |\eta_{sc}| < 1.566$). For the Z veto, discussed later in this section, we select electrons using the loose working point of the MVA-based tagger without isolation variables included in the training.

6.2 Photons

Photons are reconstructed similarly to electrons, but they lack the charged tracks from tracking system, so the energy is just determined by the shower in the ECAL. Photon identification is achieved using a multi-variate approach which combines multiple parameters related to the photon reconstruction into a single variable to discriminate signal from background. Selected photons are those which pass a cut value corresponding to 90% signal efficiency. In addition to the MVA identification, additional selection is applied to the photon candidate to discriminate against electrons. There are two options: a "conversion safe electron veto" to be used when the analysis is not sensitive to photons faked by electrons and a "pixel seed veto" to be used when the analysis is sensitive to photons faked by electrons. For our analysis, we choose the former for the channels with no electrons (μ+μ, μ+τ_{h}, τ_{h}+τ_{h}) and the latter for the channels with electrons (e+e, e+μ, e+τ_{h}). Two scale factors, further discussed in Section 7.1, are applied to the simulation to correct for differences in tagging efficiency for both the MVA identification and electron veto compared to that in data. The photon pT cut is set at 20 GeV in the e+e, μ+μ and e+μ final states as this is the lowest p_{T} value for which POG-provided scale factors are provided for `Photon_mvaID_WP90`. The photon p_{T} cut

is increased to 100, 100, 75 GeV in the μ+τ_{h}, e+τ_{h}, τ_{h}+τ_{h} final states to have an improved signal to background ratio. The lower threshold of 75 GeV is used in the τ_{h}+τ_{h} final state to increase event statistics in our signal and sideband regions. In order to discriminate from electrons (which produces a reconstructed track), the photon is required to fall within the silicon tracker acceptance of $|\eta| < 2.5$. The photon is additionally required to not fall within a gap region between the barrel and endcap electromagnetic calorimeters with no coverage.

6.3 Jets

Jet energy corrections are derived from simulation studies so that the average measured energy of jets matches that of particle level jets. In situ measurements of the p_{T} balance in dijet, photon+jet, leptonically decaying Z +jet, and multijet events are used to determine any residual differences between the jet energy scale in data and in simulation, and appropriate corrections are made to the jet p_{T} [42].

The reconstructed jets originating from b hadrons are identified using the tight working point of the DeepJet b tagging algorithm [56]. This working point has an identification efficiency of 60–75% for b quark jets, depending on jet p_{T} and η, and a misidentification probability of about 0.1% for identifying light jets as b jets. Selected jets are required to have a minimum p_{T} of 20 GeV. In order to fully reconstruct the tracks produced in the decays of b hadrons, the jets are required to fall within the silicon tracker acceptance of $|\eta| < 2.5$. Finally, the qualifying jets must lie outside a cone of $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.4$ around a selected muon, electron, or τ_{h} candidate, where $\Delta\phi$ is the ϕ angle between the jet and lepton. A scale factor, further discussed in Section 7.1, is applied to the simulation to correct for differences in the DeepTau tagging efficiency compared to that in data.

Once these jets are selected, any event containing such a jet is vetoed to reduce background levels.

6.4 Missing Transverse Momentum (MET)

The vector $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is defined as the negative vector p_{T} sum of all the PF candidates in an event, and its magnitude is denoted as $p_{\mathrm{T}}^{\mathrm{miss}}$ [44]. The pileup-per-particle identification algorithm [57] is applied to reduce the pileup dependence of the $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ observable. The $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is computed from the PF candidates weighted by their probability to originate from the PV, and is modified to account for corrections to the energy scale of the reconstructed jets in the event.

6.5 Muons

Muons are reconstructed from compatible tracks in the inner tracker and the muon detectors [36]. Additional track fit and matching quality criteria suppress the misidentification of hadronic showers that punch through the calorimeters and reach the muon system. Matching tracks measured in the inner tracker and the muon detectors results in a relative p_{T} resolution, for muons with p_{T} up to 100 GeV, of 1% in the barrel and 3% in the endcaps, and of better than 7% in the barrel for muons with p_{T} up to 1 TeV [36]. Muon identification is achieved using a cut-based approach with selection applied to parameters related to the muon track and detector geometry. Selected muons are those which pass a set of cuts denoting the tight working point. Muon isolation is characterized with a momenta sum of particle flow candidates within a cone of size $\Delta R = 0.4$ about the muon axis. Selected muons are those which pass the tight working point corresponding to 95% signal efficiency. Two scale factors, further discussed in Section 7.1, are applied to the simulation to correct the tagging efficiency for both the identification and isolation variables compared to that in data. The muon is required to have a minimum p_{T} of 15 GeV and fall within the muon detector acceptance of $|\eta| < 2.4$. For the Z veto, discussed later in this section, we select muons using the loose working point of the cut-based ID and no isolation cuts.

6.6 Hadronically decaying Taus

Hadronically decaying tau lepton identification uses a neural network named DeepTau [58] to discriminate hadronically decaying tau leptons from those faked by jets, electrons, or muons. For discrimination against jets the signal region utilizes the tight, tight, and medium working points for the e+τ_{h}, μ+τ_{h}, and τ_{h}+τ_{h} channels respectively. The medium working point is used in the τ_{h}+τ_{h} final state to increase statistics. For discrimination against muons we use the tight working point for all channels. For discrimination against electrons we use the tight, very-loose, and very-loose working points for the e+τ_{h}, μ+τ_{h}, and τ_{h}+τ_{h} final states respectively. The very-loose working point for discrimination against electrons (instead of tight) is used in the μ+τ_{h}, and τ_{h}+τ_{h} final states as electrons faking τ_{h} is not a concern and we benefit from increased statistics. Three separate scale factors, further discussed in Section 7.1, are applied to simulation to correct the tagging efficiency for each of these 3 classifiers compared to that in data. Reconstruction of τ_{h} candidates includes special "2-prong" τ_{h} decays which aim to reconstruct (true) 3-prong decays in which only 2 tracks were reconstructed – these are not considered in this analysis. The τ_{h} is required to have a minimum p_{T} of 20 GeV and fall within the acceptance of $|\eta| < 2.3$, well within the silicon tracker such that isolation sums about the central τ_{h} axis may be adequately calculated.

6.7 Triggers

An overview of the CMS trigger system can be found in Ref. [59]. A single electron trigger requiring the presence of an electron with an offline p_{T} of at least 29 (34) GeV for 2016 (2017 and 2018) is used to select events for the e+τ_{h} and the e+e final states. Additionally, for those channels we also accept a single photon trigger requiring a photon of 175 (200) GeV for 2016 (2017 and 2018) because that was shown to increase the efficiency of these types of final states passing the selection. A single isolated muon trigger requiring a muon with offline p_{T} of 26 (29) GeV for 2016 and 2018 (2017) is used to populate the μ+μ and μ+τ_{h}

final states. A suite of triggers requiring an electron and a muon are used for the e+μ final state. The p_{T} thresholds of the triggers vary between 8 and 23 GeV for the electron and between 8 and 23 for the muon depending on which particle had a higher p_{T} in the event. For high trigger efficiency, the higher p_{T} object is required to have an offline p_{T} of 25 GeV and the lower p_{T} electron (muon) must have an offline p_{T} of 20 (15) GeV. A ditau trigger requiring the presence of two τ_{h} leptons with offline p_{T} of 40 GeV and $|\eta| < 2.1$ is used for the τ_{h}+τ_{h} final state.

The trigger is used two different ways in simulation depending on the final state. In the e+τ_{h}, μ+τ_{h}, τ_{h}+τ_{h} and e+μ final states, the HLT decision is used to select simulated events and the remaining yields are corrected using scale factors, further discussed in Section 7.1. In the e+e and μ+μ final states, the HLT decision is not used to select simulated events but all simulated events are weighted by the trigger efficiency on an event-by-event basis. As there are two objects that could prompt the trigger to fire, the probability for at least one lepton passes the trigger is $p_1p_2 + p_1(1-p_2) + (1-p_1)p_2$. Because of the combinatorics, proper application of trigger scale factors is complicated and requires knowledge of the trigger efficiency in data, the trigger efficiency in simulation, and the scale factors to correct the simulated efficiency to match that in data. Direct application of the trigger efficiency, as calculated in data, simplifies the procedure. The trigger efficiency is calculated in data and parameterized in p_{T} and η of the lepton. The applied uncertainty related to this method is seen in Section 7.1.

6.8 Event Selection Criteria

All reconstructed objects are required to be separated by $\Delta R \geq 0.4$. For a given channel, the appropriate final state τ objects are selected from the objects identified above along with a photon. Additionally, the leptons are required to have opposite electric charge must have a combined visible mass greater than 100 GeV in the signal channels (ones containing

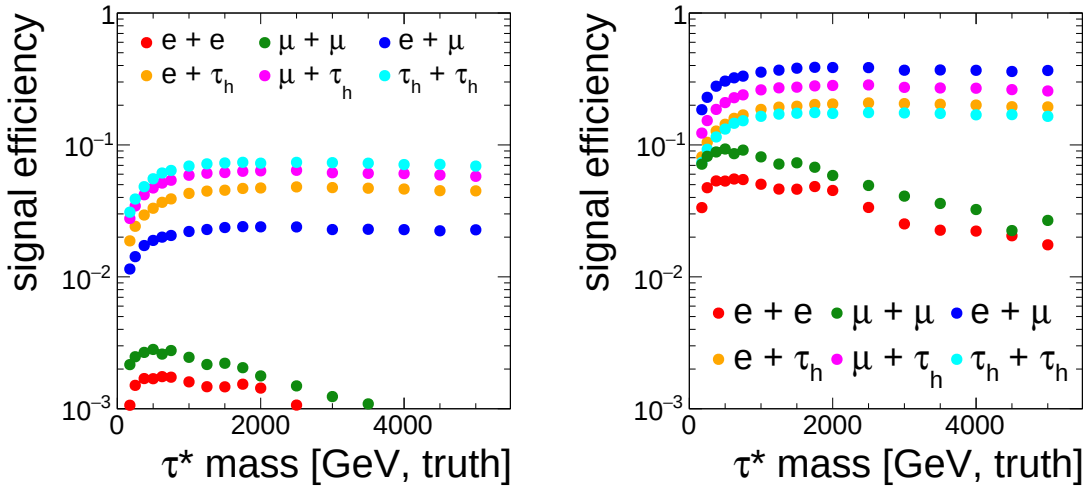

Figure 6.1: Signal efficiencies after all event selection. The left plot shows that relative to inclusive $\tau\tau$ decays, the right plot is normalized to the branching fraction of its particular final state.

τ_h) to eliminate background from Drell-Yan events. Visible mass is specified because the τ_h will have some invisible decay products. In the control regions with two identical leptons (e+e and μ+μ), the combined mass must be between 50 and 140 GeV since these regions are hoping to constrain the normalization on Drell-Yan events and, therefore, are hoping to include as many as possible. An addition veto is applied in the signal channels that looks for non-isolated electrons and muons that would reproduce a mass consistent with the Z boson (between 50 and 140 GeV) when paired with one of the selected isolated leptons. For this identification we use an MVA-based discriminator for non-isolated electrons with $p_T > 20$ GeV and a cut-based discriminator for muons with $p_T > 15$ GeV. To suppress top quark events in the signal channels, we veto events that have tight b tagged jets. Additional filters are applied to check that the subdetectors were working properly for the data taking period and that there are no spurious contributions to p_T^{miss}. Final states are kept orthogonal by requiring exactly the number of electrons or muons passing the cuts that are required for that final states. No requirement is placed on the number of τ_h to maintain efficiency. Figures 6.1 and 6.2 show the expected signal efficiencies and yields after the above event selection is applied in the signal region A.

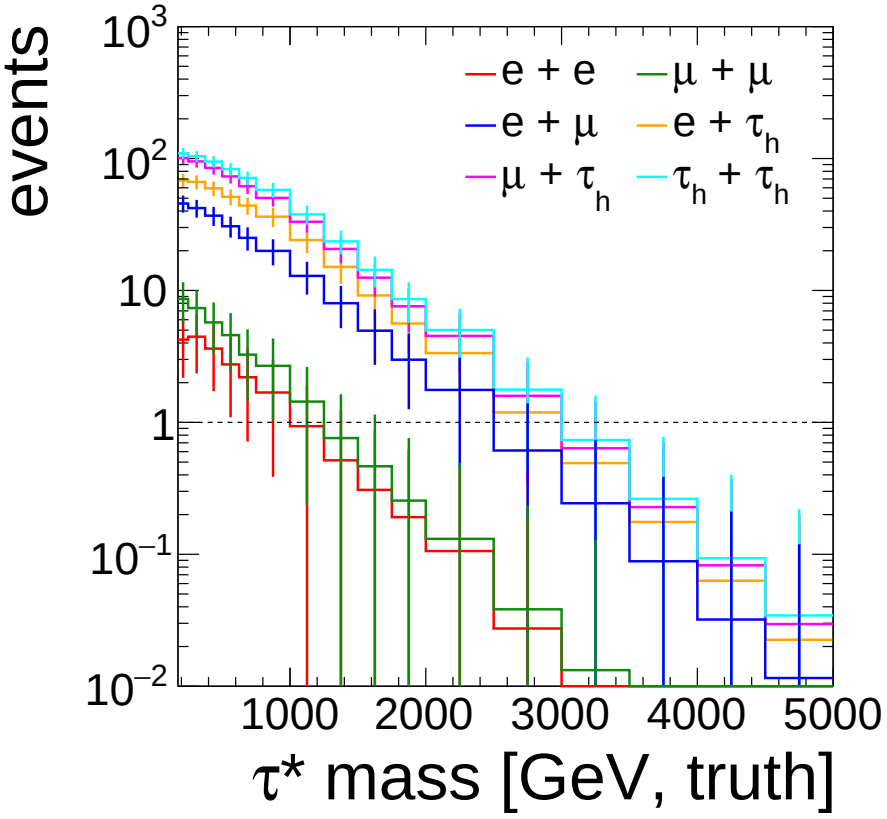

Figure 6.2: Expected signal yields after all event selection. Reference cross sections are found in Section 5.2

6.9 Background Estimation

As mentioned in Section 5.2, the backgrounds are categorized according to the source of the reconstructed τ_{h} candidate. When simulated events meet the criterion of having reconstructed τ_{h} candidates matched to a lepton in the initial hard scattering, those events contribute to the MC estimate of background processes. Simulated events in which one or more of the reconstructed τ_{h} candidates is matched to a jet or unmatched are discarded because those events are estimated by a data-driven approach described in Section 6.9.2. The expected background yield taken from simulation is corrected by defining additional control regions in data, as detailed in Section 6.9.1.

6.9.1 Control regions to correct MC yields

To correct the normalization of the background taken from simulation, three control regions are defined from the final states that do not include τ_{h}. These regions are included in the statistical analysis and thus help constrain the normalization of the SM processes in the signal regions. These regions make no use of τ_{h} reconstruction but are sensitive to most other nuisance parameters which simultaneously affect the signal region. Two Drell-Yan enriched control regions are defined by requiring the presence of two oppposite-sign same-flavour leptons (e or μ) and a photon of $p_{\mathrm{T}} > 20\,\mathrm{GeV}$. As in the signal region, these channels make use of the relevant single-lepton trigger and is briefly discussed in Section 6. A single control region enriched in WW and $\mathrm{t\bar{t}}$ is defined by requiring the presence of one opposite-sign electron-muon pair and a photon of $p_{\mathrm{T}} > 20\,\mathrm{GeV}$. The floating normalization of these control regions is implemented as three multiplicative scale factors affecting the individual normalizations of the Drell-Yan, $\mathrm{t\bar{t}}$, and diboson processes.

To help distinguish the WW and $\mathrm{t\bar{t}}$ processes from each other, the fit is performed using histograms of the number of reconstructed (AK4) jets with $p_{\mathrm{T}} \geq 30\,\mathrm{GeV}$. This procedure is also implemented in the low-p_{T} photon sideband used for the corrections to the ABCD

background estimation described in Section 6.9.2.

6.9.2 ABCD estimation

The background arising from events in which a jet is misreconstructed as a τ_{h} is estimated with what is called an ABCD approach. Control and sideband regions are defined by inverting requirements on a) the relative charge of the selected leptons and b) the isolation of the τ_{h}. These sideband regions allow for calculation of transfer factors which are used to scale the control region yields to produce and estimate of the background in the signal region.

The background estimation method in e+τ_{h} and μ+τ_{h} final states are identical. A control region (B) is defined by inverting the isolation of the τ_{h} to create a region enriched in events in which a jet fakes a τ_{h}. This region as well as the signal region are parameterized by the binning scheme shown in Figure 5.3. The anti-isolated τ_{h} candidate must satisfy the loosest working point of the DeepTau anti-jet discriminator but fail the tight working point. Sideband regions (C, D) are defined by requiring the charge of both leptons to have the same sign. These regions are defined as simply a total count of all events in this region. These regions are summarized in Figure 6.3. The background estimate in the τ_{h}+τ_{h} final state follows similarly, but the presence of two τ_{h} presents the complication of inverting the isolation on one or both τ_{h}. We have chosen to define the control and sideband regions in this final state to consist of events with at least one anti-isolated τ_{h}, using the medium working point of the DeepTau anti-jet discriminator for the nominal selection.

Before performing the estimate, the yields observed in these three regions (B, C, D) are corrected by subtracting the event yields from backgrounds taken from simulation to ensure that the ABCD method captures only those events in which a jet fakes a τ_{h}. The transfer factor is calculated by dividing the observed yields in the C region by that of the D region. The transfer factor is calculated separately for the 1 and 3-prong τ_{h} decay modes. In the τ_{h}+τ_{h} final state, the decay mode of the sub-leading τ_{h} is used. Each event in Control Region B is then scaled by the transfer factor appropriate for the decay mode of the τ_{h}.

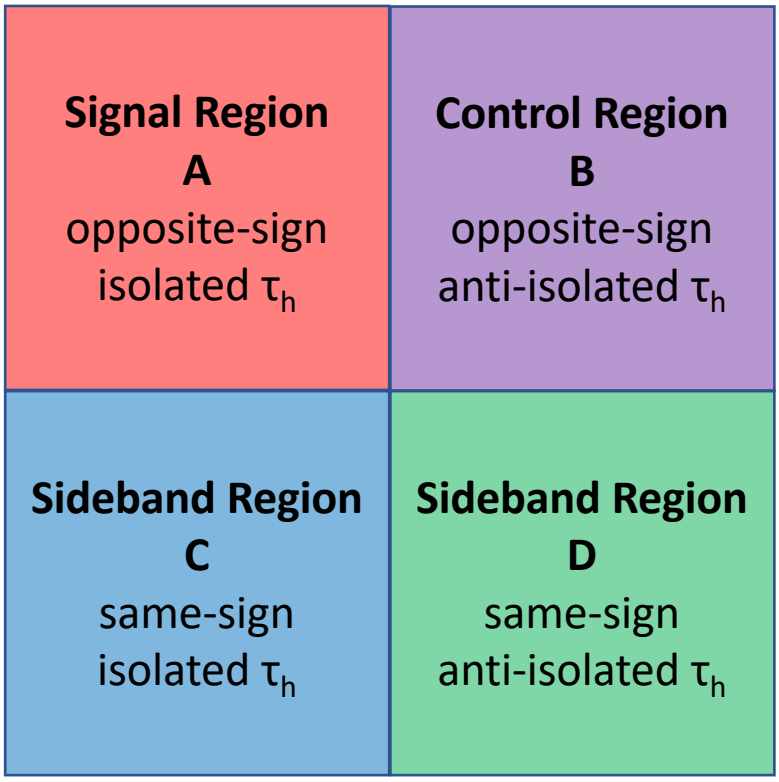

Figure 6.3: Definitions of control and sideband regions used in the jet background estimation. The $\tau_{\mathrm{h}}+\tau_{\mathrm{h}}$ final state requires at least one anti-isolated τ_{h} for the B and D regions.

Chapter 7

Data Fitting and Results

7.1 Systematic Uncertainties

Systematic uncertainties arise from corrections applied to background and signal simulation. These corrections are often represented by scale factors applied to the simulation to deal with mismodeling of the simulation with respect to the data. Such scale factors arise to correct things such as a Level-1 prefiring issue for ECAL and muons; the anti-e, anti-μ, and anti-jet DeepTau discriminator; trigger efficiencies; the pileup distribution; reconstruction, identification and isolation efficiencies for muons; reconstruction and MVA identification efficiencies for electrons; MVA identification, electron veto and pixel seed veto efficiency for photons; the ditau trigger HLT decision; and b tagging efficiencies. Additionally there is an uncertainty associated with the luminosity for all years of data taking. Unless otherwise stated, all nuisance parameters across different years are taken as uncorrelated. Table 7.1 summarizes the size of these uncertainties. More detail is about the scale factors in the following paragraphs.

The systematic uncertainties enter into the maximum likelihood fit as nuisance parameters. The variance of these parameters as they affect the final optimization is estimated by taking a shape uncertainty obtained by taking the nominal value for the scale factor and

varying it up and down by one standard deviation. Those new values are then applied to the appropriate samples and the results form the up and down histograms for the fit. Separate nuisance parameters are maintained for individual data-taking years and are treated as uncorrelated. Additionally, there a problem arose due to saturation in the preamplifier of the APV chips in early 2016. The setting were eventually changed to prevent the saturation, but this resulted in two independent data-taking periods in 2016 that would also require separate nuisance parameters in the fit.

Some parameters only affect certain channels or affect different channels in different ways. Often the scale factors are applied for individual objects being present in an event, so multiple final state objects of the same kind in an event can change the significance of the uncertainty.

Most scale factors are adequately described above, but a few require more description. For the anti-jet discriminator in the DeepTau algorithm, scale factors used by CMS are divided either by p_{T} or by the decay mode of the τ_{h}, which is described by the number of charged and neutral hadrons involved in the decay. Scale factors organized via decay mode are only available starting at $p_{\mathrm{T}} > 40\,\mathrm{GeV}$, so the p_{T}-based scale factors are used below that range.

7.2 Maximum Likelihood Fit

A maximum likelihood fit is performed simultaneously for all regions referenced in Section 6.9. The individual Poisson-distibuted probability density functions for all bins and all samples are included. Each hypothetical mass point listed in Table 5.1 is profiled separately because only one mass of the excited τ would be expected to exist at a time. This fit is performed using the asymptotic formula for likelihood tests as described in Ref. [60]. These methods are far less computationally intensive and, in the limit of large datasets, the results would be the same. The fit is optimized for sensitivity using an Asimov dataset, which is a

Table 7.1: Summary of systematic uncertainties. Values are in % of inclusive background yield and rounded to the nearest whole number. For the τ_h containing columns the second number is the value for signal ($m_{\tau^*} = 1750\,\text{GeV}$).

	e+e	μ+μ	e+μ	e+τ_h	μ+τ_h	τ_h+τ_h
luminosity	2	2	2	2, 2	2,2	2,2
L1PreFiringWeight ECAL	0	0	0	1,0	0,0	0,0
L1PreFiringWeight Muon	-	0	1	-	1,0	-
DeepTau anti-e	-	-	-	0,0	0,0	0,0
DeepTau anti-μ	-	-	-	0,0	0,0	0,0
DeepTau anti-jet	-	-	-	6,7	6,7	13,13
τ_h+τ_h trigger	-	-	-	-	-	10,13
pile-up	1	0	0	2,1	1,1	2,1
μ reco	-	7	5	-	1,1	-
μ id	-	0	0	-	0,0	-
μ isolation	-	1	0	-	0,0	-
μ trigger	-	-	-	-	0,0	-
$\mu\mu$ trigger	-	0	-	-	-	-
e reco	1	-	1	1,1	-	-
e id	7	-	3	3,3	-	-
γ id	2	2	1	1,2	1,2	1,2
γ pixel-veto	1	-	0	0, 0	-	0
γ csev	-	0	-	-	0,0	0,0
e trigger	-	-	-	1, 1	-	-
ee trigger	0	-	-	-	-	-
eμ trigger	-	-	1	-	-	-
btag light	-	-	-	0,0	0,0	-
btag light corr	-	-	-	0,0	0,0	-
btag bc	-	-	-	1,0	1,0	-
btag bc corr	-	-	-	1,0	1,0	-

dataset in which the expectation value for all parameters are used to generate the dataset. This dataset is profiled using the distribution of events expected to fall into the regions from Section 6.9 while additionally dividing the events into those that fall into the L-shaped window on the 2D mass plane of $\tau + \gamma$ mass combinations, as seen in Figure 5.3 as the red Bin 2. The other bins in that figure are combined into one Reject bin for statistical reasons when performing the maximum likelihood fit. The amount of parameter space belonging to each of the Bins 1, 2, 3 and 4 depends on the mass of the τ^*. As previously stated in Section 5.1.1, the width of bin 2 is equal to 10 % of the τ^* mass, and it is centered around the τ^*

mass in each dimension. Therefore, Bin 1 can be quite small at low masses and Bins 3 and 4 may have a paucity of events at very high masses. Combining these 3 bins into one Reject bin reduces problems with fitting from bins containing no events.

The systematic uncertainties are included as nuisance parameters and bin-by-bin methods for the uncertainties introduced by Barlow and Beeston [61] are employed. In fact, this approach takes it one step farther by combining all the independent sources of uncertainty into a single overall parameter, as described in [62]. Additionally, many of the parameters are also subject to shape uncertainty, or template morphing. This is also handled by methods described in Ref. [62]. The maximum likelihood method allows these parameters to float and the final value selected for these parameters in the fit is the one which maximizes the overall likelihood by also taking into account the probability of the parameter taking that value.

7.3 Results

In the end, the useful quantity to extract is the signal rate observed in the data. Given the ambiguous nature of pairing of the photon with the correct tau, it can also be difficult to extract an understanding of the results through plots. The following is an attempt to distill that information. The final dataset in the signal region was not available at the time of the completion of this dissertation, so what will be shown are the various cross checks that would give us confidence on our ability to see the signal when the data are available. All limits shown here are expected limits based on the Asimov dataset.

As described in the previous section, the distributions are submitted to the maximum likelihood fit as histograms separated into Accept and Reject bins in each region. The individual regions are divided according to the presence of a high- or low-p_{T} photon, the region of the ABCD method, and the number of prongs in the decay of the hadronic τ in the event. Figures 7.1-7.9 show the event totals for three different masses of the τ^* (175, 1000

and 2500 GeV) for each of the signal final states. The figures do not show the data for the A region of high-photon-p_{T} because that is where the bulk of the signal events are expected to fall. To comply with the current blinding policy of the analysis, that data is withheld at this stage. These plots show the level of agreement for all regions between data and number of expected events from Monte Carlo and data-driven prediction.

To focus more specifically on the A region, the canonical 2D plot can be reduced to one dimension by just selecting the pairing of τ with γ that is closer to the τ^* mass under hypothesis and looking for the difference between the reconstructed mass and the hypothesis mass. Figures 7.10-7.18 show the results of looking at that. Even though the data isn't shown for the high-photon-p_{T} region, the plots show that signal peaks in the center bin, which is equivalent to the Accept region described previously. The number of signal events expected would clearly show over the background events present. The number of expected signal events dwindle as the mass of the τ^* increases, such that for the last few plots at 2500 GeV, fewer than one signal event is expected.

Figure 7.19 shows where the limit falls at the 95 % confidence level using the Asimov dataset to generate the expected limit for the data. The solid theory curve represents the compositeness scale $\Lambda = 10$ TeV, which is the same scale used to produce the previous plots. The limit in such a case is approximately 2750 GeV, or 2.75 TeV. The dashed theory curve represents another natural scale to choose, making the scale equal to the hypothesis mass of the τ^*. Doing so increases the limit, since it should be easier to observe events that could be produced at a lower energy scale. The limit in this case is closer to 4500 GeV, or 4.5 TeV. Previous analyses that looked at excited leptons have also placed limits on mass of an excited lepton. The CMS analysis looking at excited electrons and muons places the limit at 3.9 and 3.8 TeV, respectively, for a compositeness scale equal to the mass of the excited lepton. Previous searches for τ^* have set the limit at 2.5 TeV with the compositeness scale equal to the mass of the τ^*.

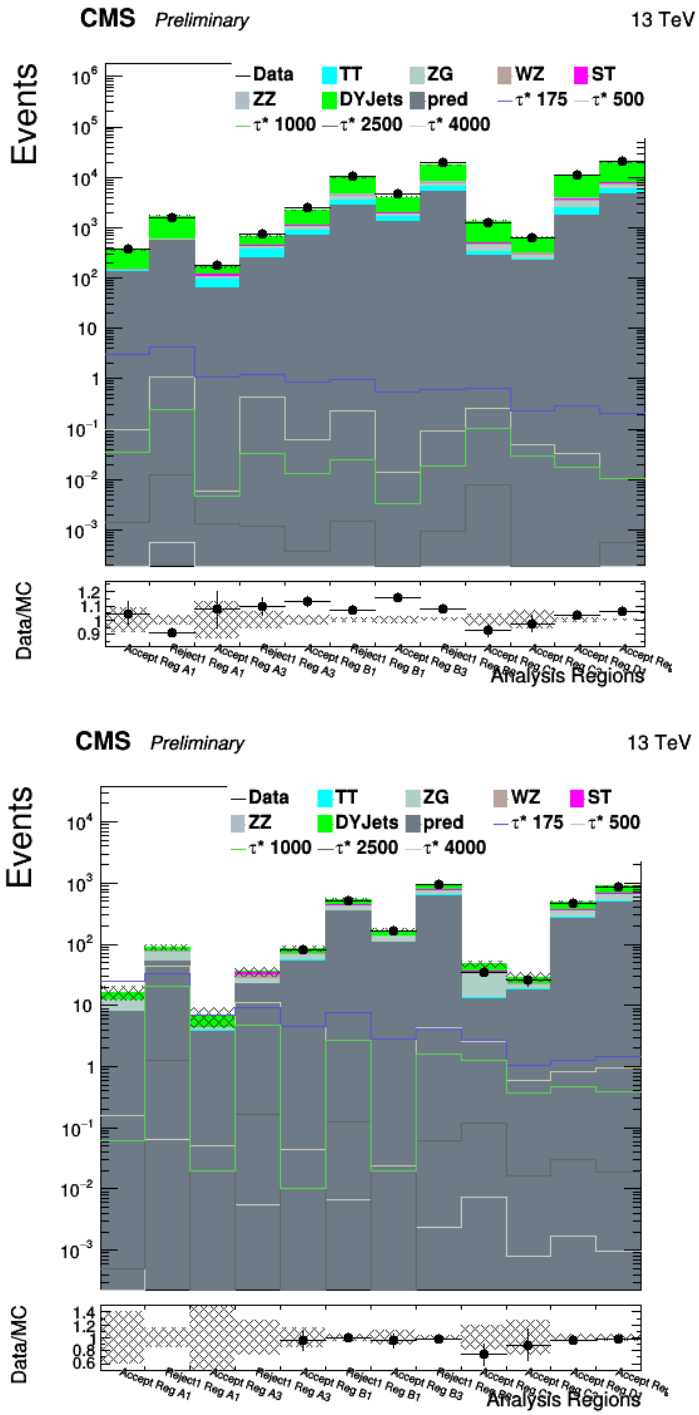

Figure 7.1: Final event totals for the e + τ_{h} channel for low-photon-p_{T} sideband (top) and high-photon-p_{T} region (bottom). Bins labeled Accept (Reject) represent events in (outside) the L-shaped window on the 2D mass plane of $\tau + \gamma$ for a τ^* mass of 2500 GeV. Bins are also separated by ABCD region and by prong. The A region of the high-photon-p_{T} region is blinded.

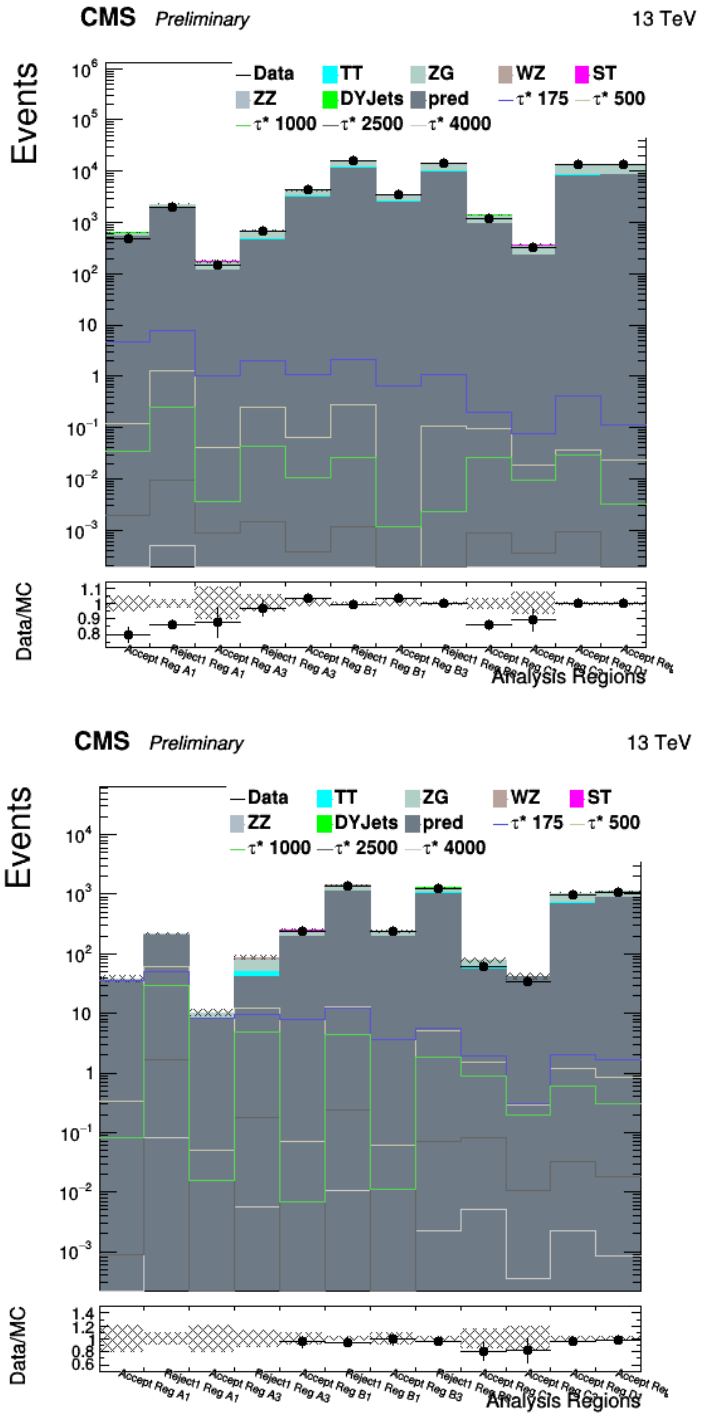

Figure 7.2: Final event totals for the $\mu + \tau_h$ channel for low-photon-p_T sideband (top) and high-photon-p_T region (bottom). Bins labeled Accept (Reject) represent events in (outside) the L-shaped window on the 2D mass plane of $\tau + \gamma$ for a τ^* mass of 2500 GeV. Bins are also separated by ABCD region and by prong. The A region of the high-photon-p_T region is blinded.

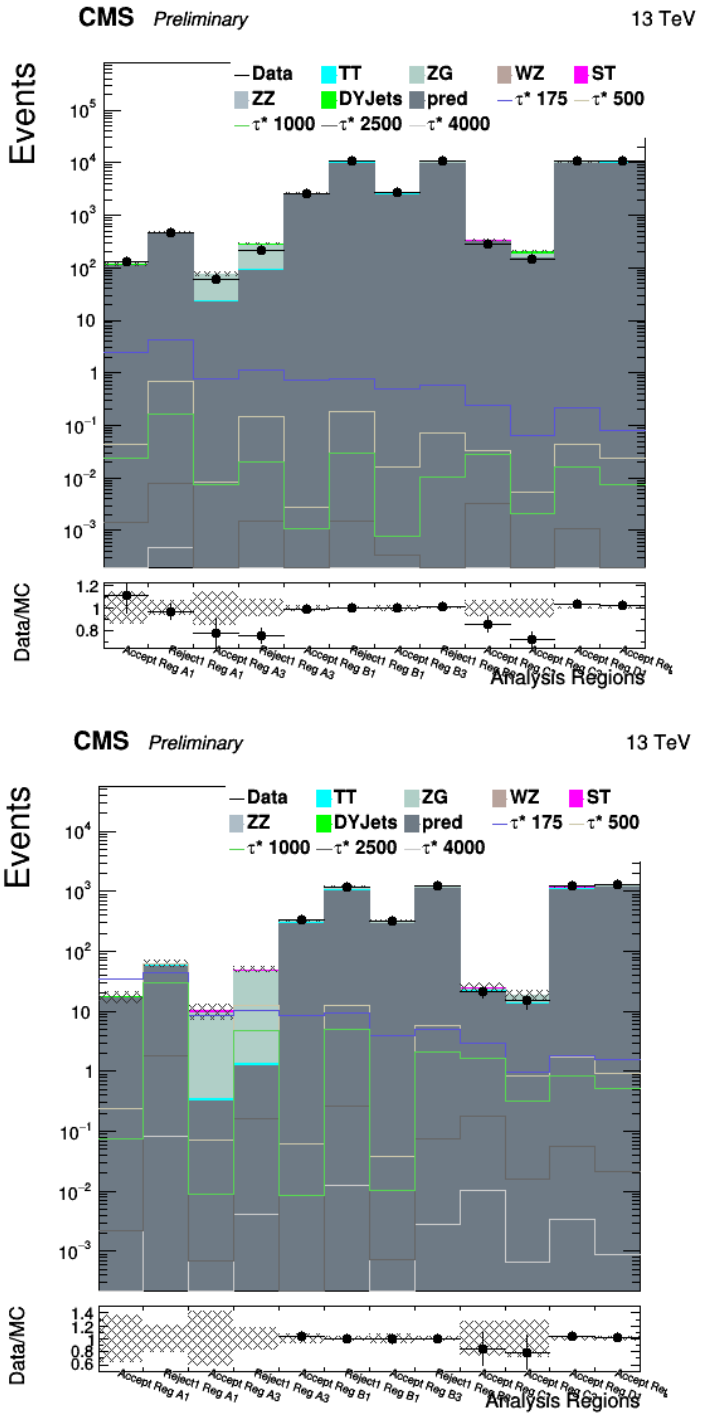

Figure 7.3: Final event totals for the $\tau_h + \tau_h$ channel for low-photon-p_T sideband (top) and high-photon-p_T region (bottom). Bins labeled Accept (Reject) represent events in (outside) the L-shaped window on the 2D mass plane of $\tau + \gamma$ for a τ^* mass of 2500 GeV.Bins are also separated by ABCD region and by prong. The A region of the high-photon-p_T region is blinded.

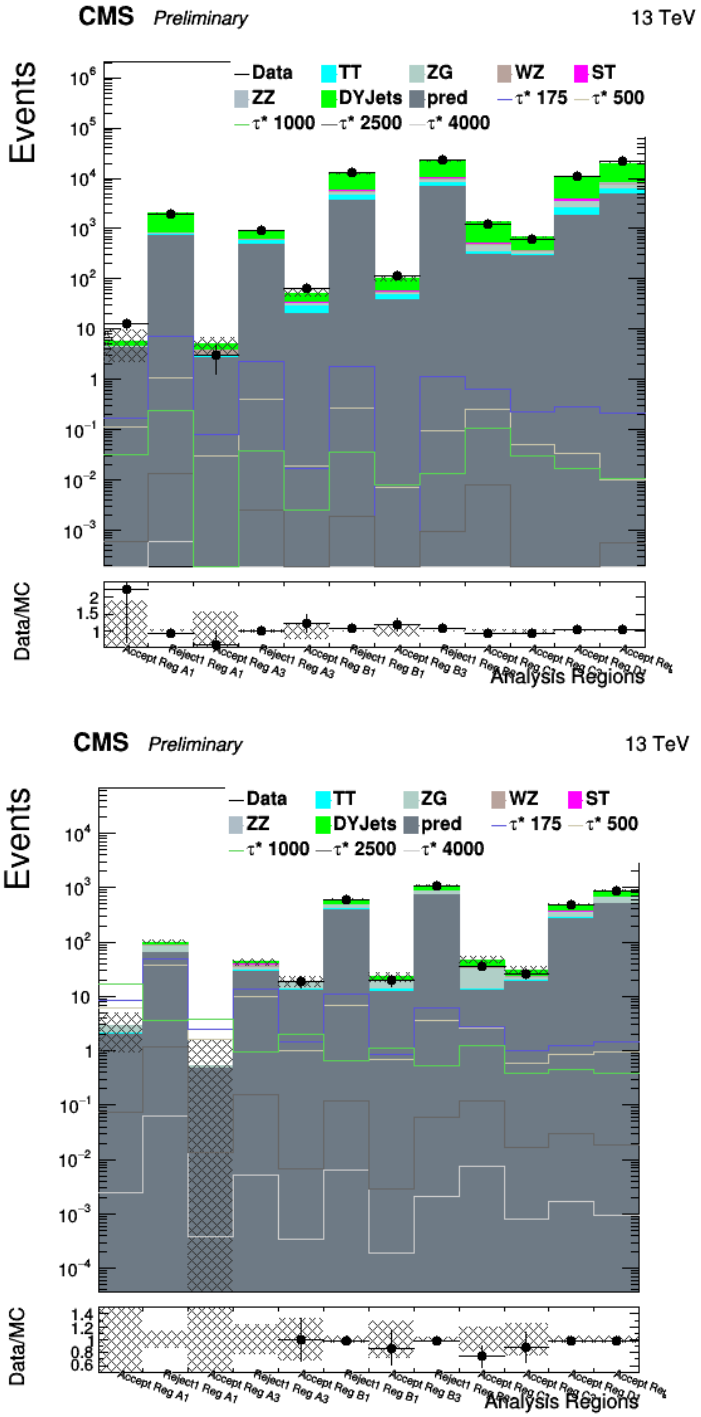

Figure 7.4: Final event totals for the e + τ_h channel for low-photon-p_T sideband (top) and high-photon-p_T region (bottom). Bins labeled Accept (Reject) represent events in (outside) the L-shaped window on the 2D mass plane of $\tau + \gamma$ for a τ^* mass of 1000 GeV. Bins are also separated by ABCD region and by prong. The A region of the high-photon-p_T region is blinded.

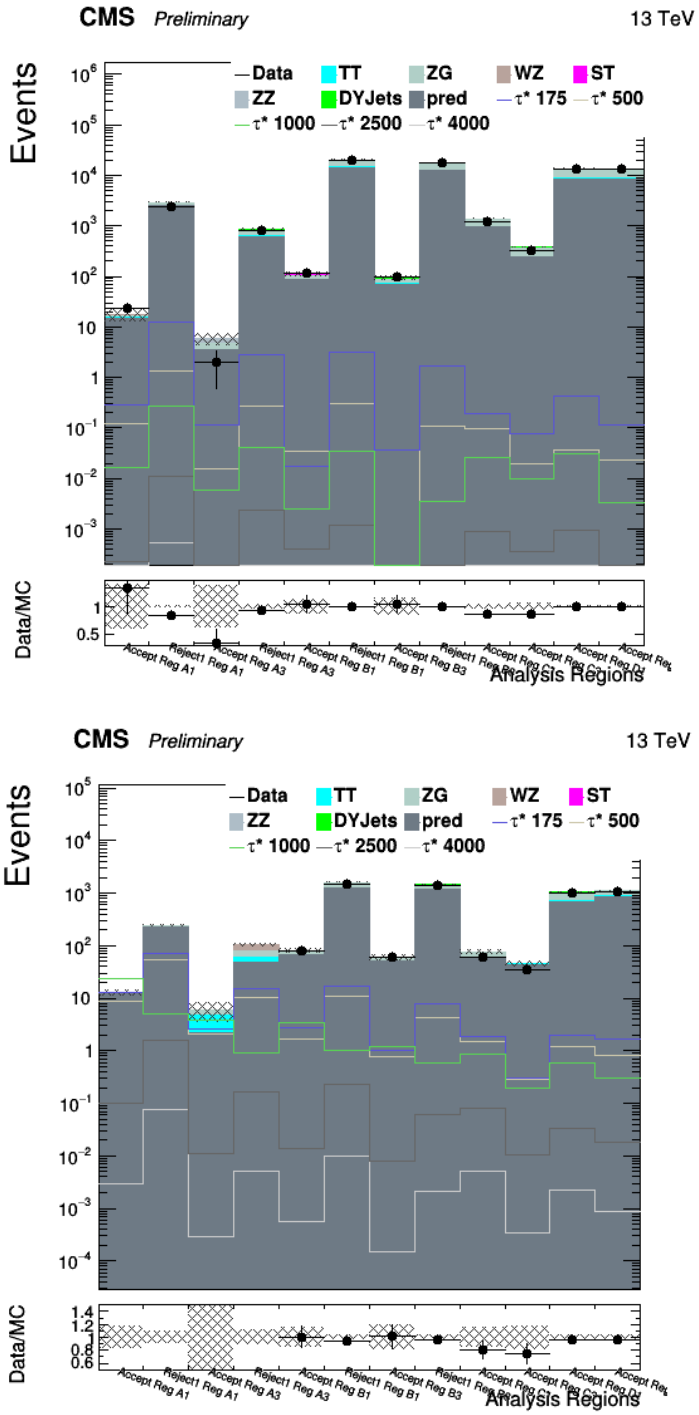

Figure 7.5: Final event totals for the $\mu + \tau_h$ channel for low-photon-p_T sideband (top) and high-photon-p_T region (bottom). Bins labeled Accept (Reject) represent events in (outside) the L-shaped window on the 2D mass plane of $\tau + \gamma$ for a τ^* mass of 1000 GeV. Bins are also separated by ABCD region and by prong. The A region of the high-photon-p_T region is blinded.

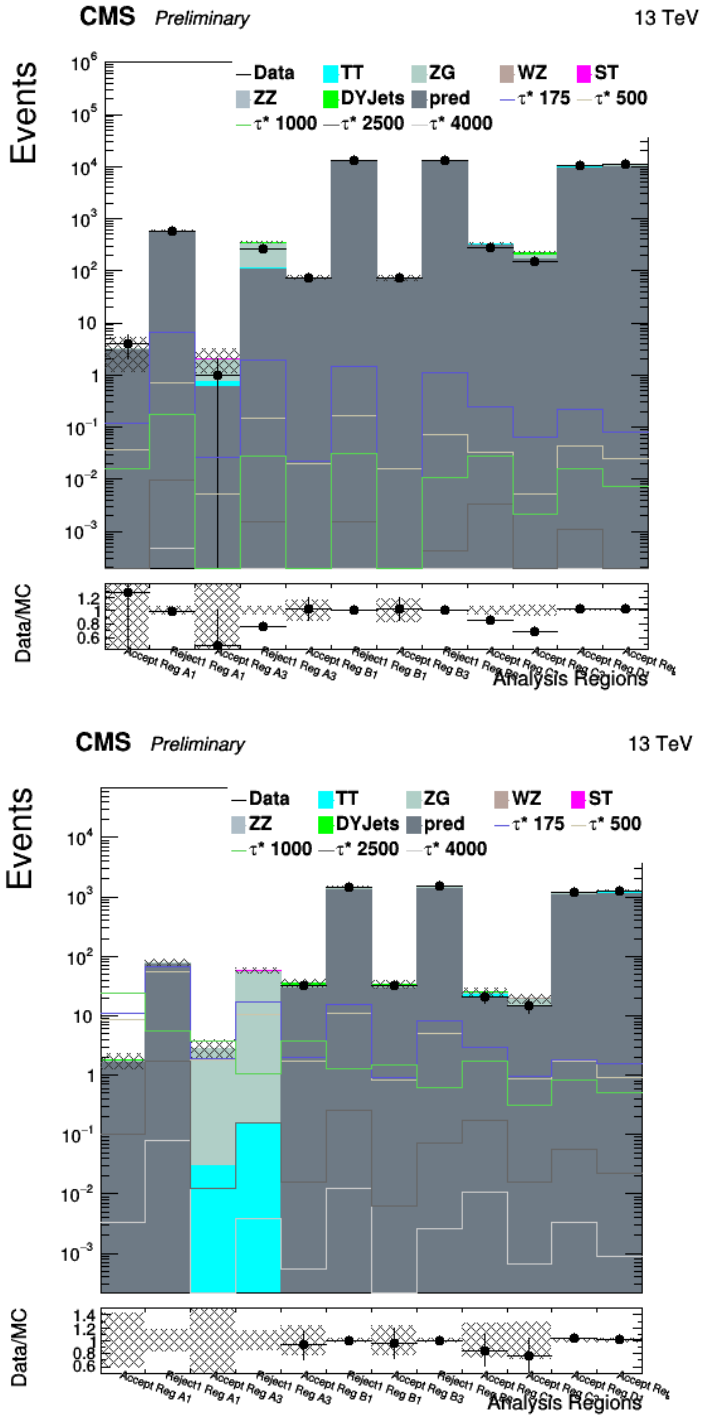

Figure 7.6: Final event totals for the $\tau_h + \tau_h$ channel for low-photon-p_T sideband (top) and high-photon-p_T region (bottom). Bins labeled Accept (Reject) represent events in (outside) the L-shaped window on the 2D mass plane of $\tau + \gamma$ for a τ^* mass of 2500 GeV.Bins are also separated by ABCD region and by prong. The A region of the high-photon-p_T region is blinded.

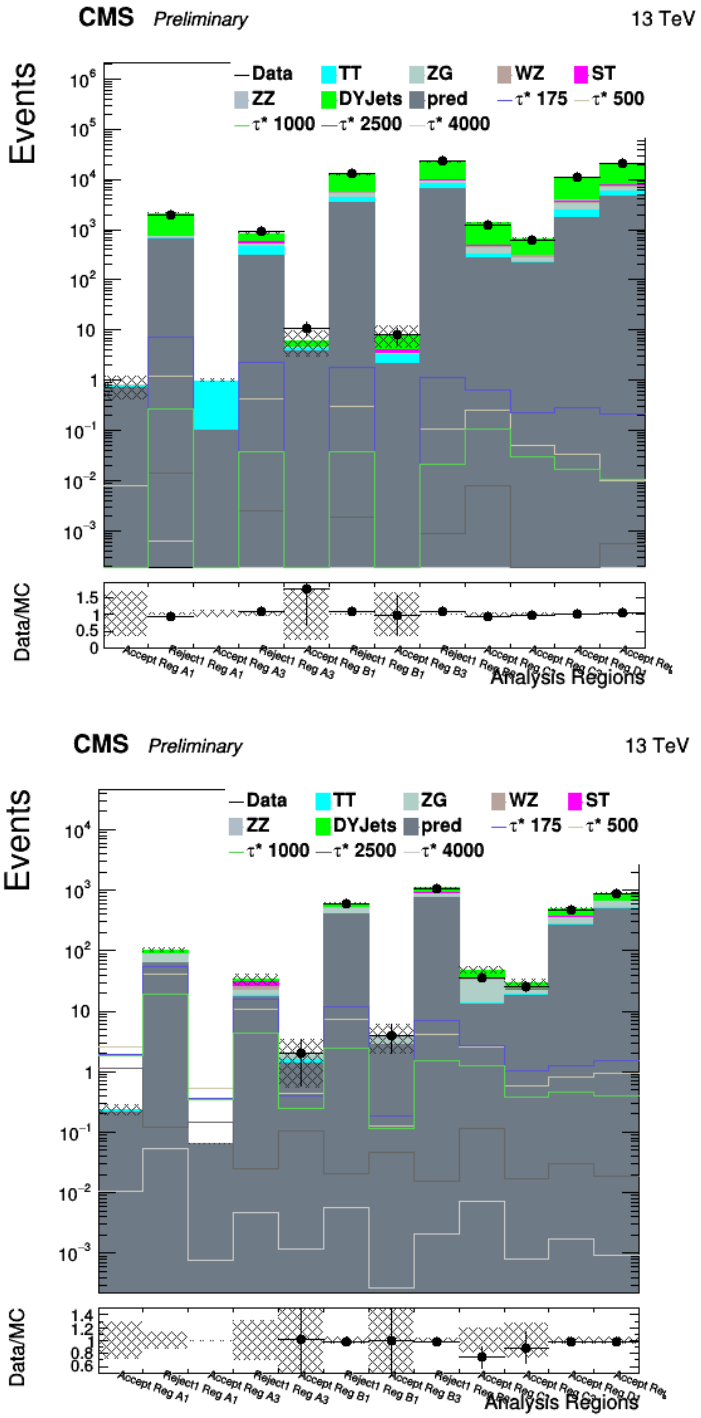

Figure 7.7: Final event totals for the e + τ_h channel for low-photon-p_T sideband (top) and high-photon-p_T region (bottom). Bins labeled Accept (Reject) represent events in (outside) the L-shaped window on the 2D mass plane of $\tau + \gamma$ for a τ^* mass of 2500 GeV. Bins are also separated by ABCD region and by prong. The A region of the high-photon-p_T region is blinded.

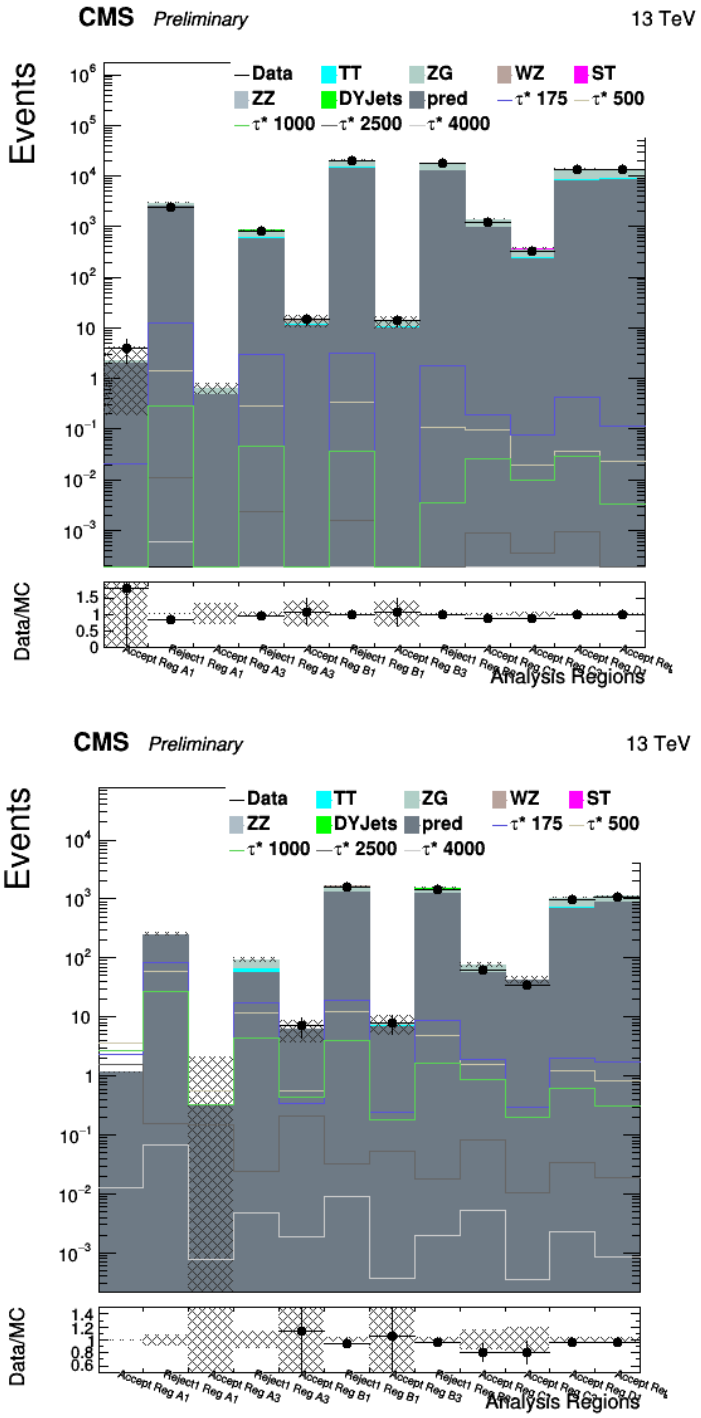

Figure 7.8: Final event totals for the $\mu + \tau_h$ channel for low-photon-p_T sideband (top) and high-photon-p_T region (bottom). Bins labeled Accept (Reject) represent events in (outside) the L-shaped window on the 2D mass plane of $\tau + \gamma$ for a τ^* mass of 2500 GeV. Bins are also separated by ABCD region and by prong. The A region of the high-photon-p_T region is blinded.

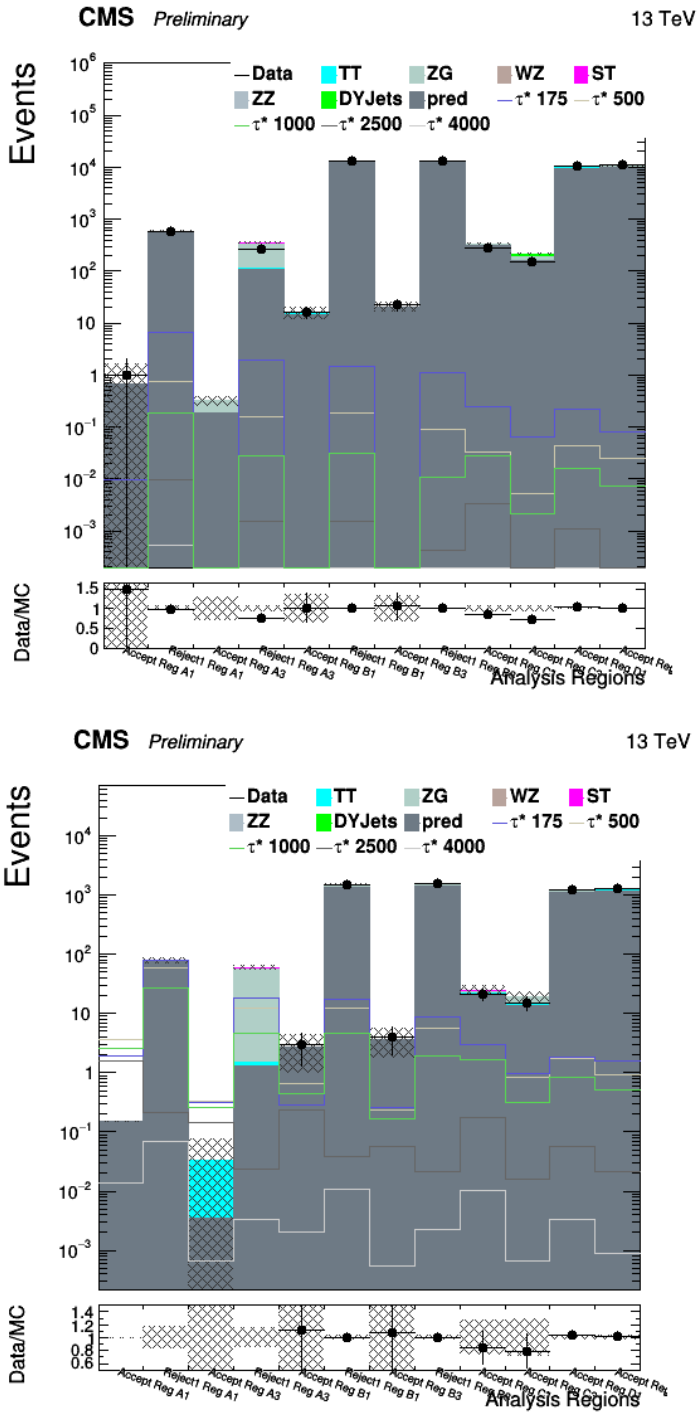

Figure 7.9: Final event totals for the $\tau_h + \tau_h$ channel for low-photon-p_T sideband (top) and high-photon-p_T region (bottom). Bins labeled Accept (Reject) represent events in (outside) the L-shaped window on the 2D mass plane of $\tau + \gamma$ for a τ^* mass of 2500 GeV.Bins are also separated by ABCD region and by prong. The A region of the high-photon-p_T region is blinded.

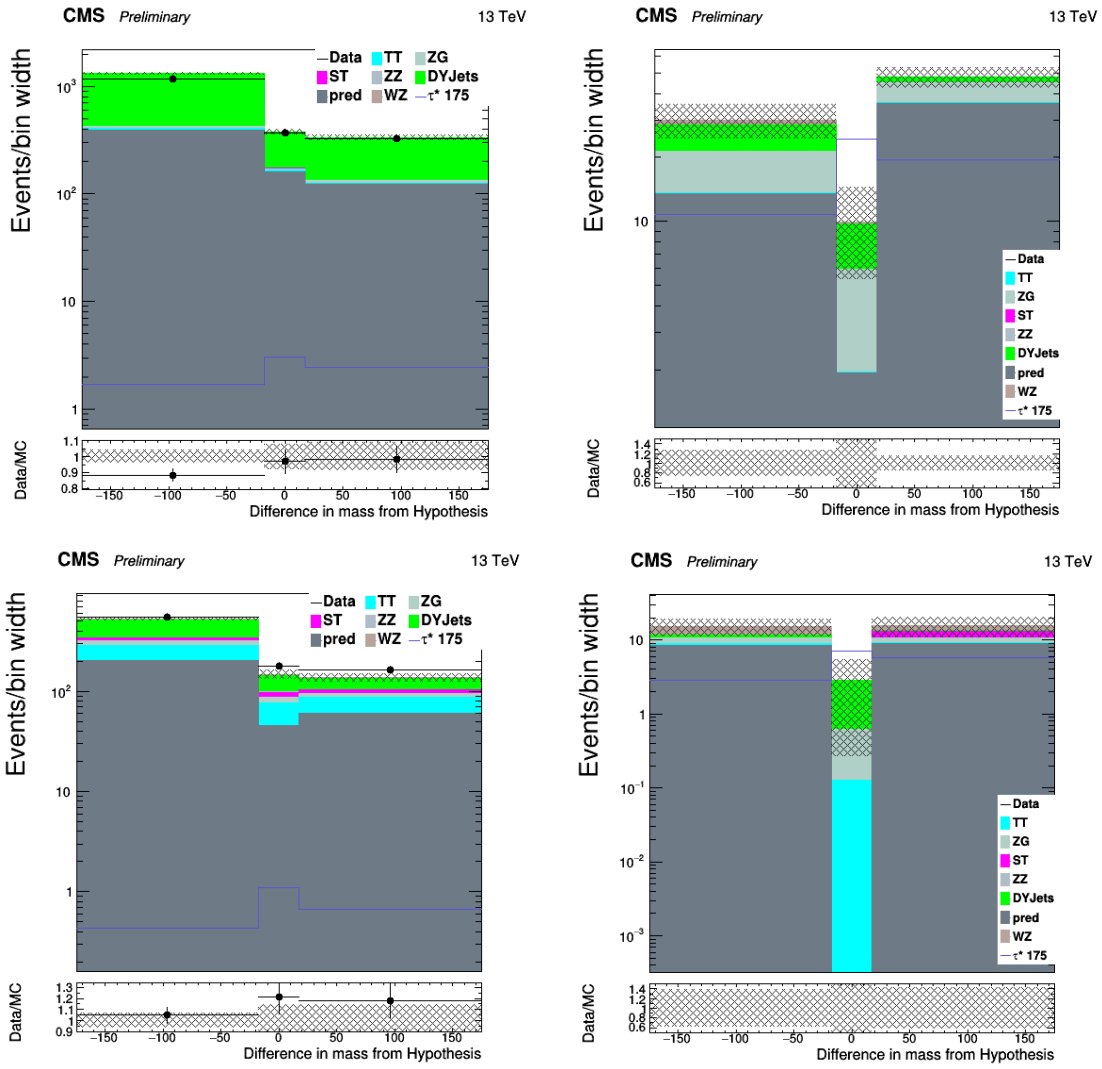

Figure 7.10: Final event totals for the e + τ_{h} channel in Region A of the ABCD method. The L-shaped window on the 2D mass plane of $\tau + \gamma$ is reduced to one dimension by taking the closer reconstructed mass to 175 GeV of the two possible pairings. The four plots show the results for one-prong τ_{h} decays (top row) and three-prong τ_{h} decays (bottom row) as well as low-photon-p_{T} sideband (left) and the high-photon-p_{T} region (right). Event totals in the A region of the high-photon-p_{T} region are masked to stay consistent with the current blinding policy of the analysis.

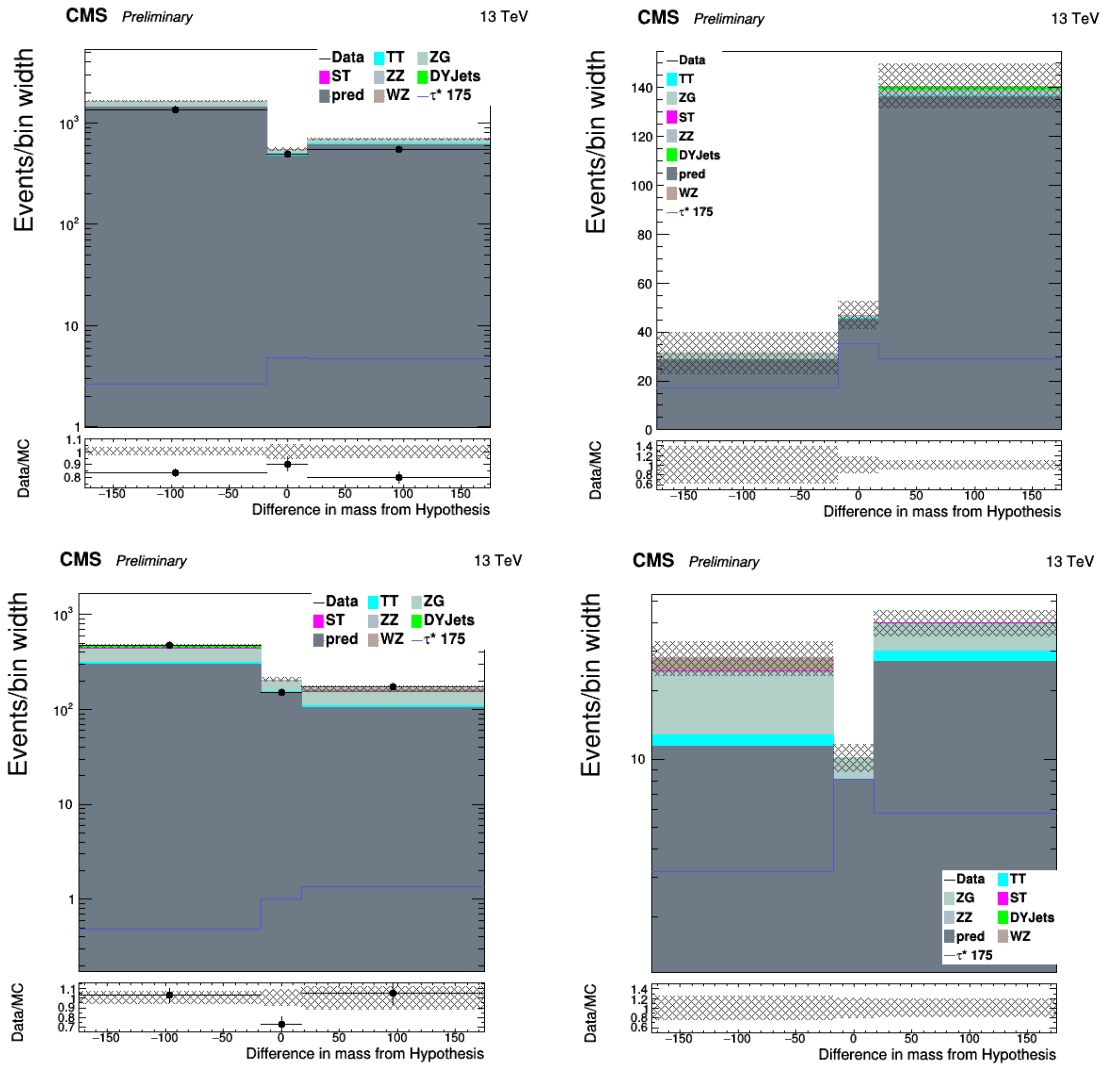

Figure 7.11: Final event totals for the $\mu + \tau_{\mathrm{h}}$ channel in Region A of the ABCD method. The L-shaped window on the 2D mass plane of $\tau + \gamma$ is reduced to one dimension by taking the closer reconstructed mass to 175 GeV of the two possible pairings. The four plots show the results for one-prong τ_{h} decays (top row) and three-prong τ_{h} decays (bottom row) as well as low-photon-p_{T} sideband (left) and the high-photon-p_{T} region (right). Event totals in the A region of the high-photon-p_{T} region are masked to stay consistent with the current blinding policy of the analysis.

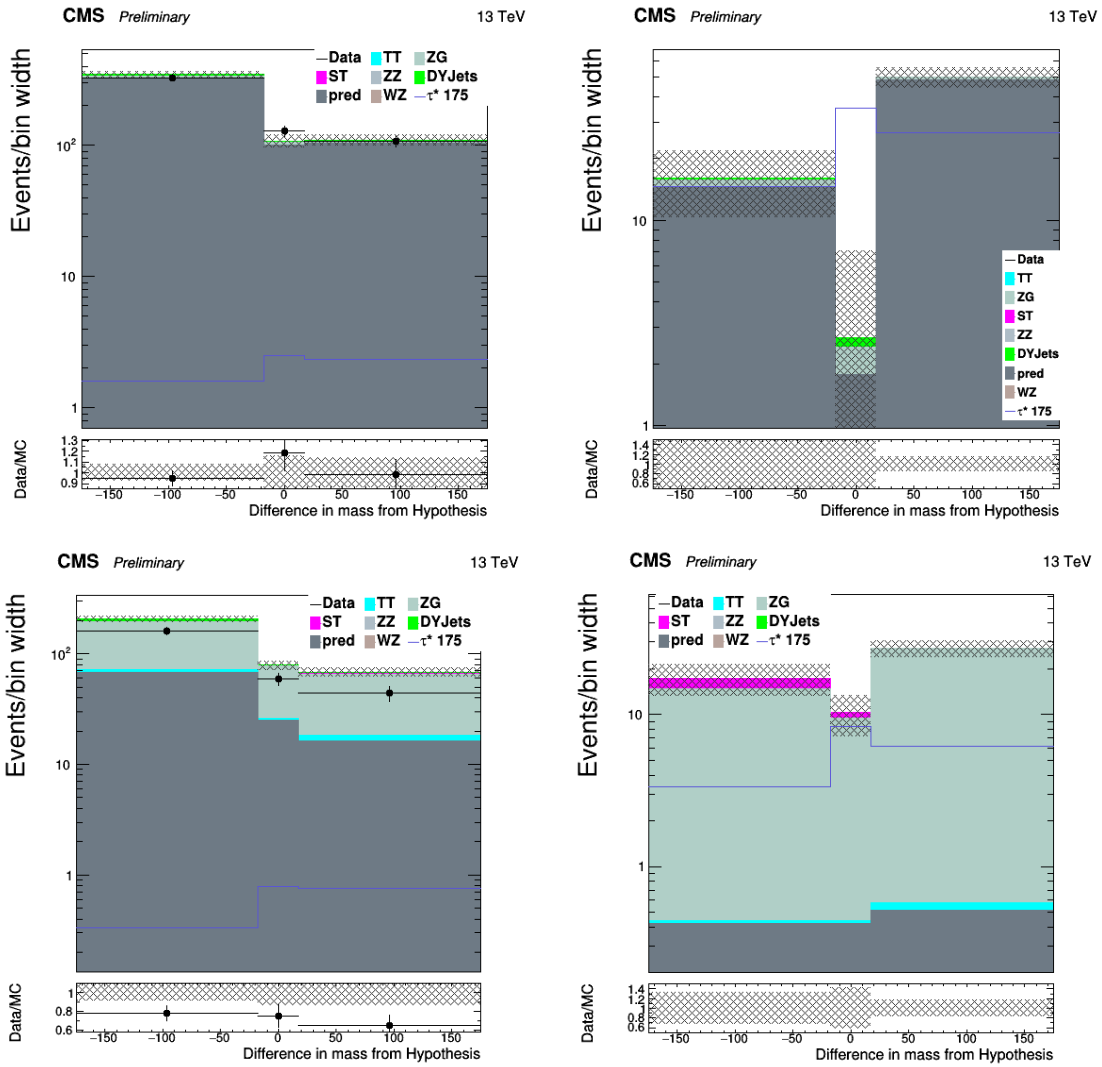

Figure 7.12: Final event totals for the $\tau_h + \tau_h$ channel in Region A of the ABCD method. The L-shaped window on the 2D mass plane of $\tau + \gamma$ is reduced to one dimension by taking the closer reconstructed mass to 175 GeV of the two possible pairings. The four plots show the results for one-prong τ_h decays (top row) and three-prong τ_h decays (bottom row) as well as low-photon-p_T sideband (left) and the high-photon-p_T region (right). Event totals in the A region of the high-photon-p_T region are masked to stay consistent with the current blinding policy of the analysis.

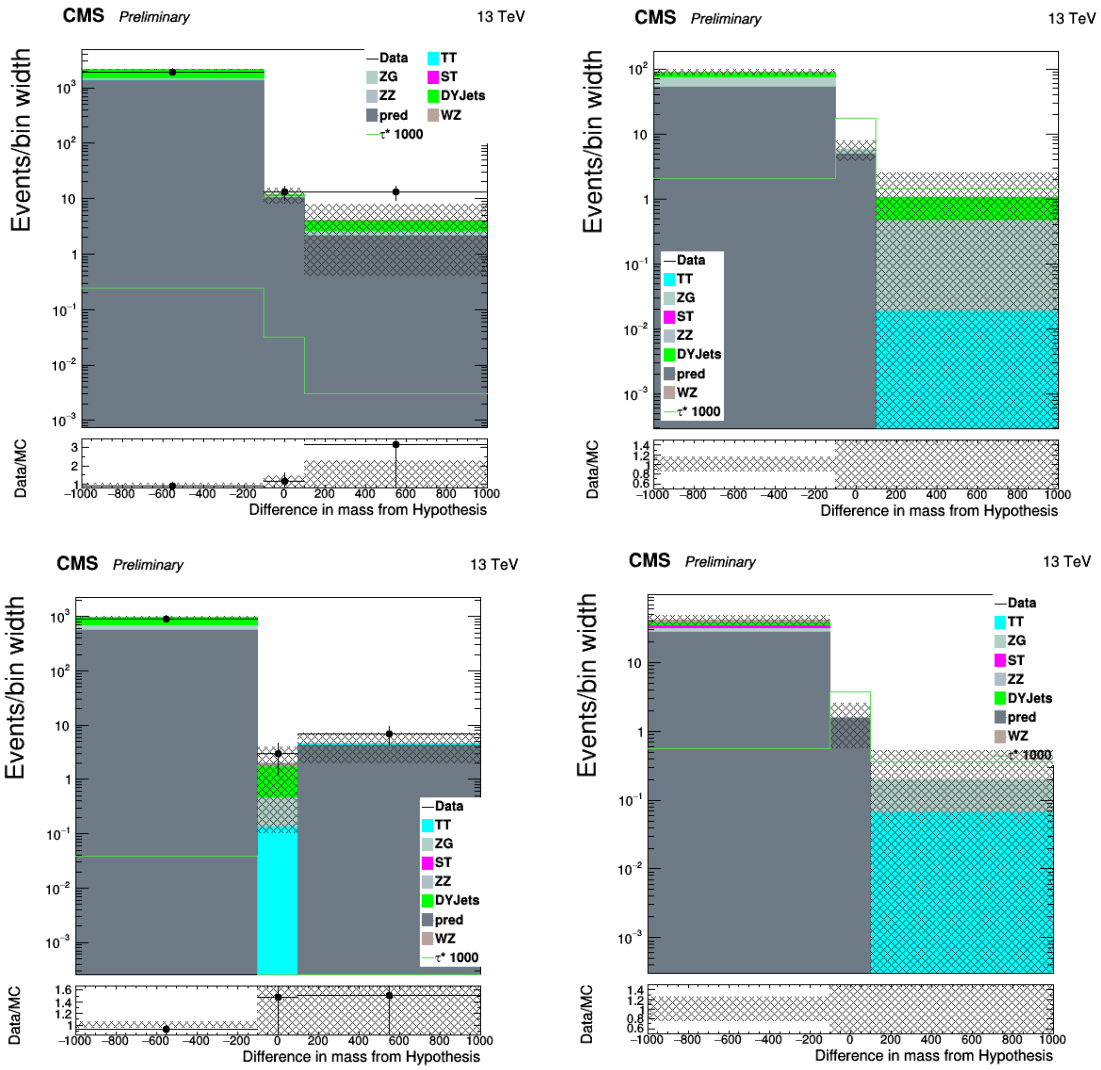

Figure 7.13: Final event totals for the e + τ_{h} channel in Region A of the ABCD method. The L-shaped window on the 2D mass plane of $\tau + \gamma$ is reduced to one dimension by taking the closer reconstructed mass to 1000 GeV of the two possible pairings. The four plots show the results for one-prong τ_{h} decays (top row) and three-prong τ_{h} decays (bottom row) as well as low-photon-p_{T} sideband (left) and the high-photon-p_{T} region (right). Event totals in the A region of the high-photon-p_{T} region are masked to stay consistent with the current blinding policy of the analysis.

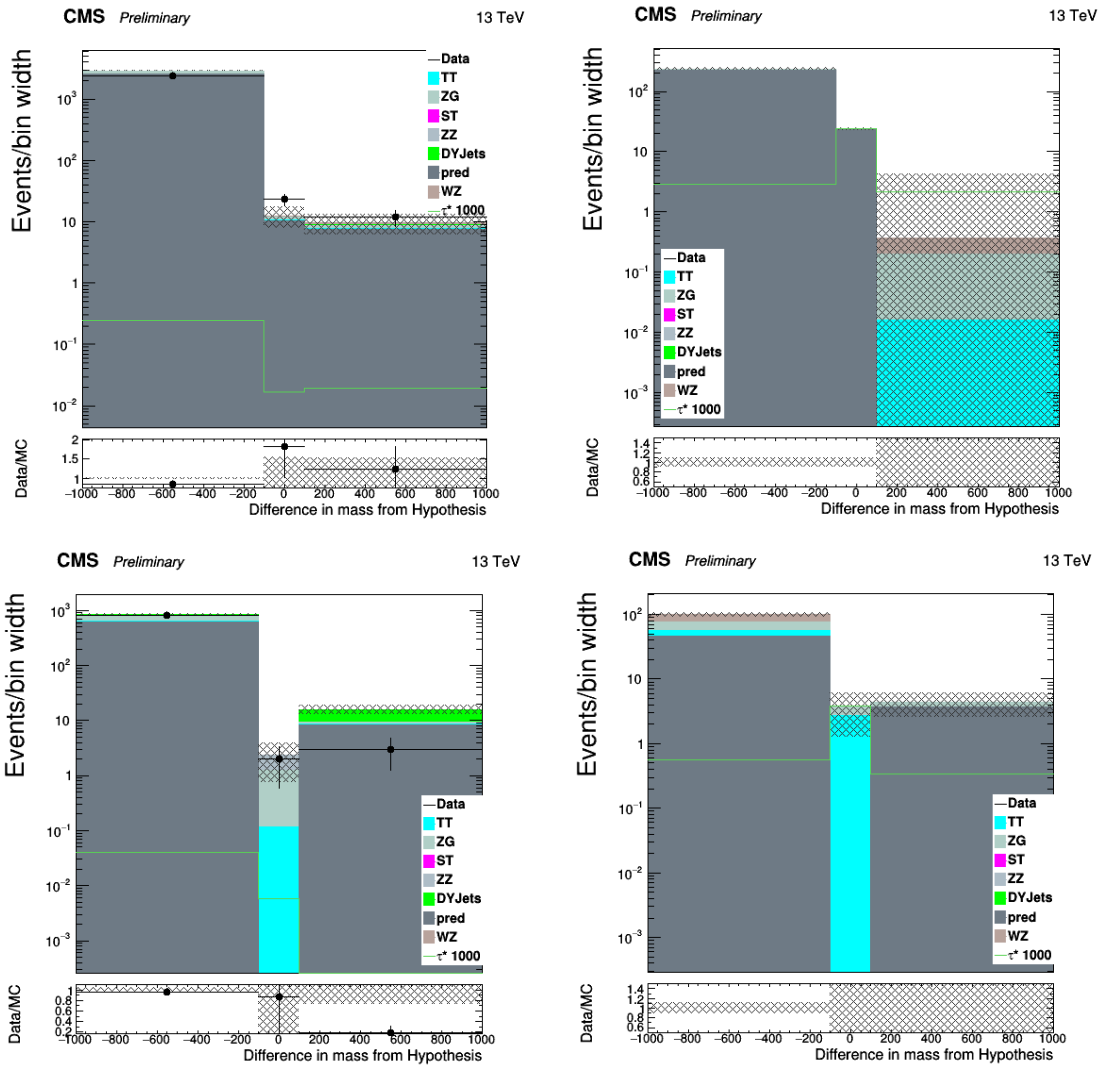

Figure 7.14: Final event totals for the $\mu + \tau_{\mathrm{h}}$ channel in Region A of the ABCD method. The L-shaped window on the 2D mass plane of $\tau + \gamma$ is reduced to one dimension by taking the closer reconstructed mass to 1000 GeV of the two possible pairings. The four plots show the results for one-prong τ_{h} decays (top row) and three-prong τ_{h} decays (bottom row) as well as low-photon-p_{T} sideband (left) and the high-photon-p_{T} region (right). Event totals in the A region of the high-photon-p_{T} region are masked to stay consistent with the current blinding policy of the analysis.

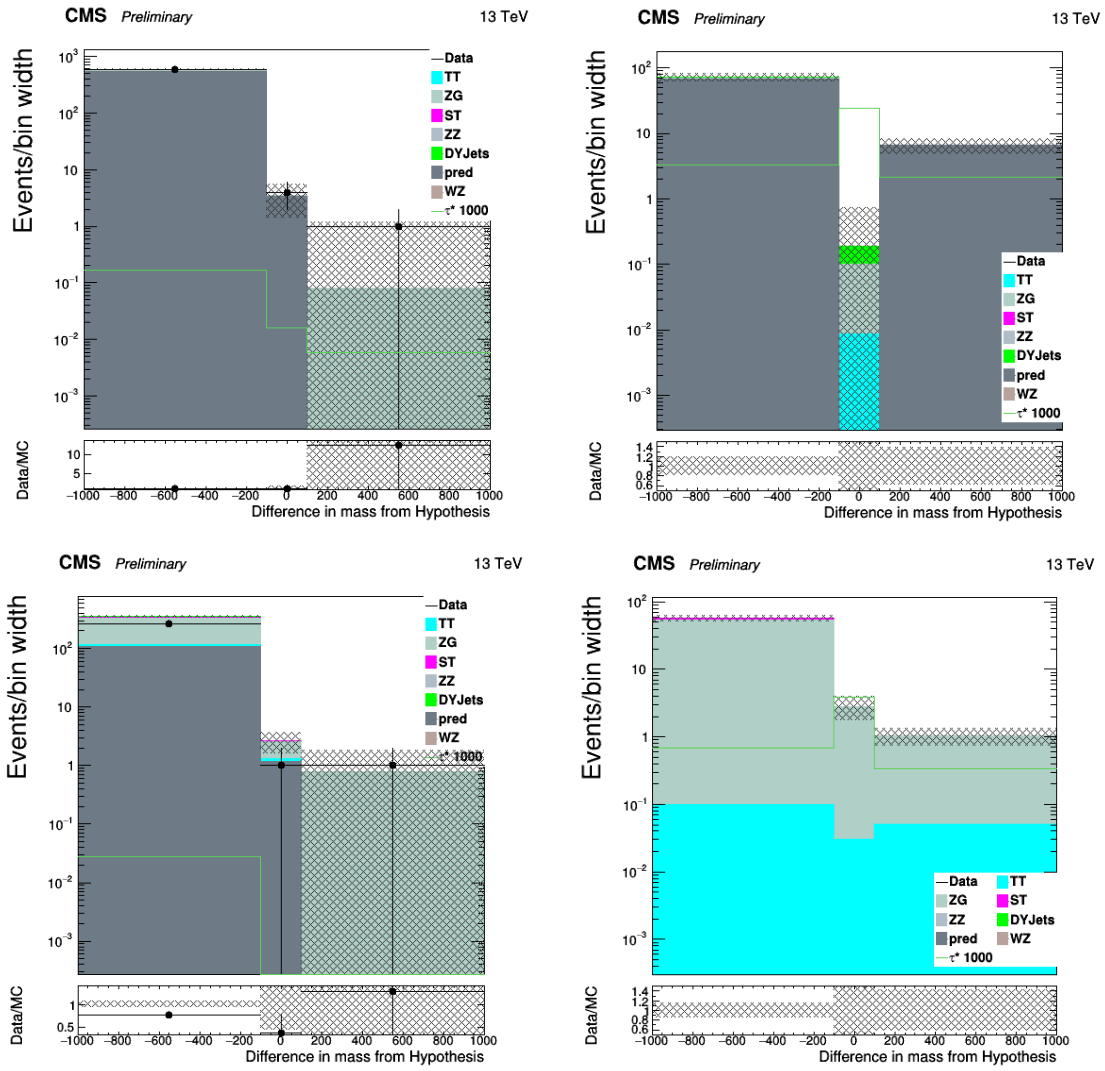

Figure 7.15: Final event totals for the $\tau_{\mathrm{h}} + \tau_{\mathrm{h}}$ channel in Region A of the ABCD method. The L-shaped window on the 2D mass plane of $\tau + \gamma$ is reduced to one dimension by taking the closer reconstructed mass to 1000 GeV of the two possible pairings. The four plots show the results for one-prong τ_{h} decays (top row) and three-prong τ_{h} decays (bottom row) as well as low-photon-p_{T} sideband (left) and the high-photon-p_{T} region (right). Event totals in the A region of the high-photon-p_{T} region are masked to stay consistent with the current blinding policy of the analysis.

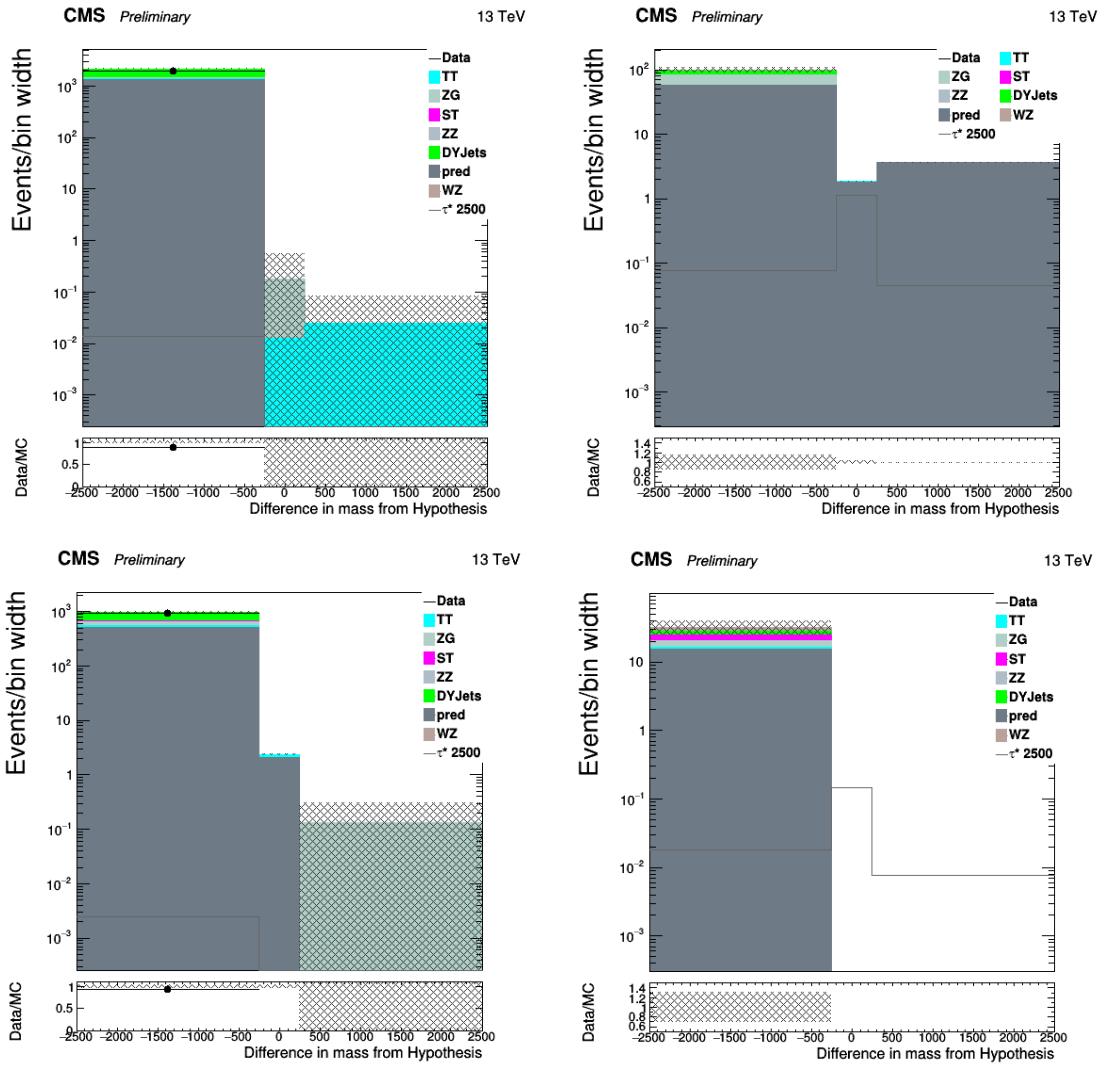

Figure 7.16: Final event totals for the e + τ_{h} channel in Region A of the ABCD method. The L-shaped window on the 2D mass plane of $\tau + \gamma$ is reduced to one dimension by taking the closer reconstructed mass to 2500 GeV of the two possible pairings. The four plots show the results for one-prong τ_{h} decays (top row) and three-prong τ_{h} decays (bottom row) as well as low-photon-p_{T} sideband (left) and the high-photon-p_{T} region (right). Event totals in the A region of the high-photon-p_{T} region are masked to stay consistent with the current blinding policy of the analysis.

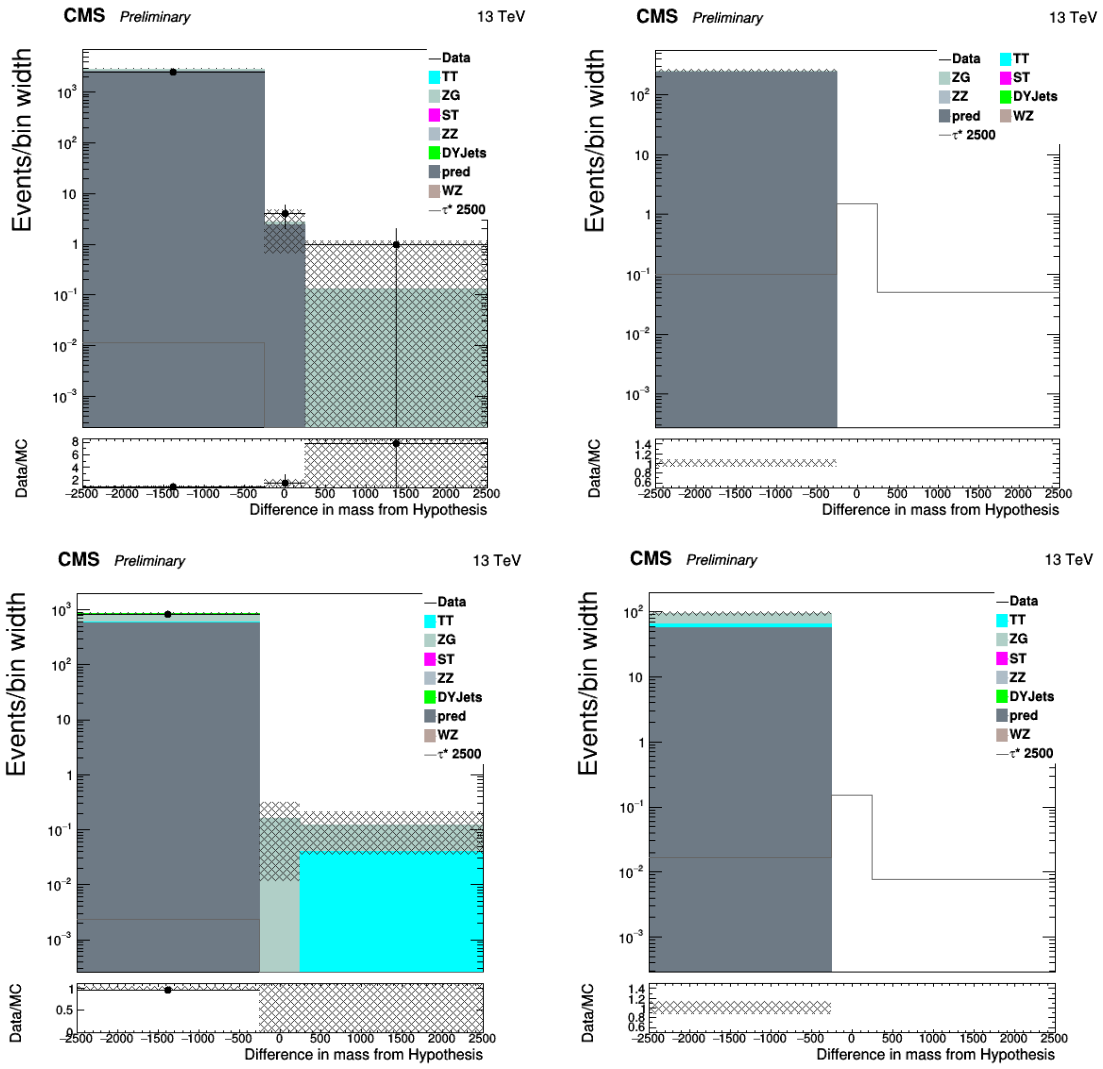

Figure 7.17: Final event totals for the $\mu + \tau_{\mathrm{h}}$ channel in Region A of the ABCD method. The L-shaped window on the 2D mass plane of $\tau + \gamma$ is reduced to one dimension by taking the closer reconstructed mass to 2500 GeV of the two possible pairings. The four plots show the results for one-prong τ_{h} decays (top row) and three-prong τ_{h} decays (bottom row) as well as low-photon-p_{T} sideband (left) and the high-photon-p_{T} region (right). Event totals in the A region of the high-photon-p_{T} region are masked to stay consistent with the current blinding policy of the analysis.

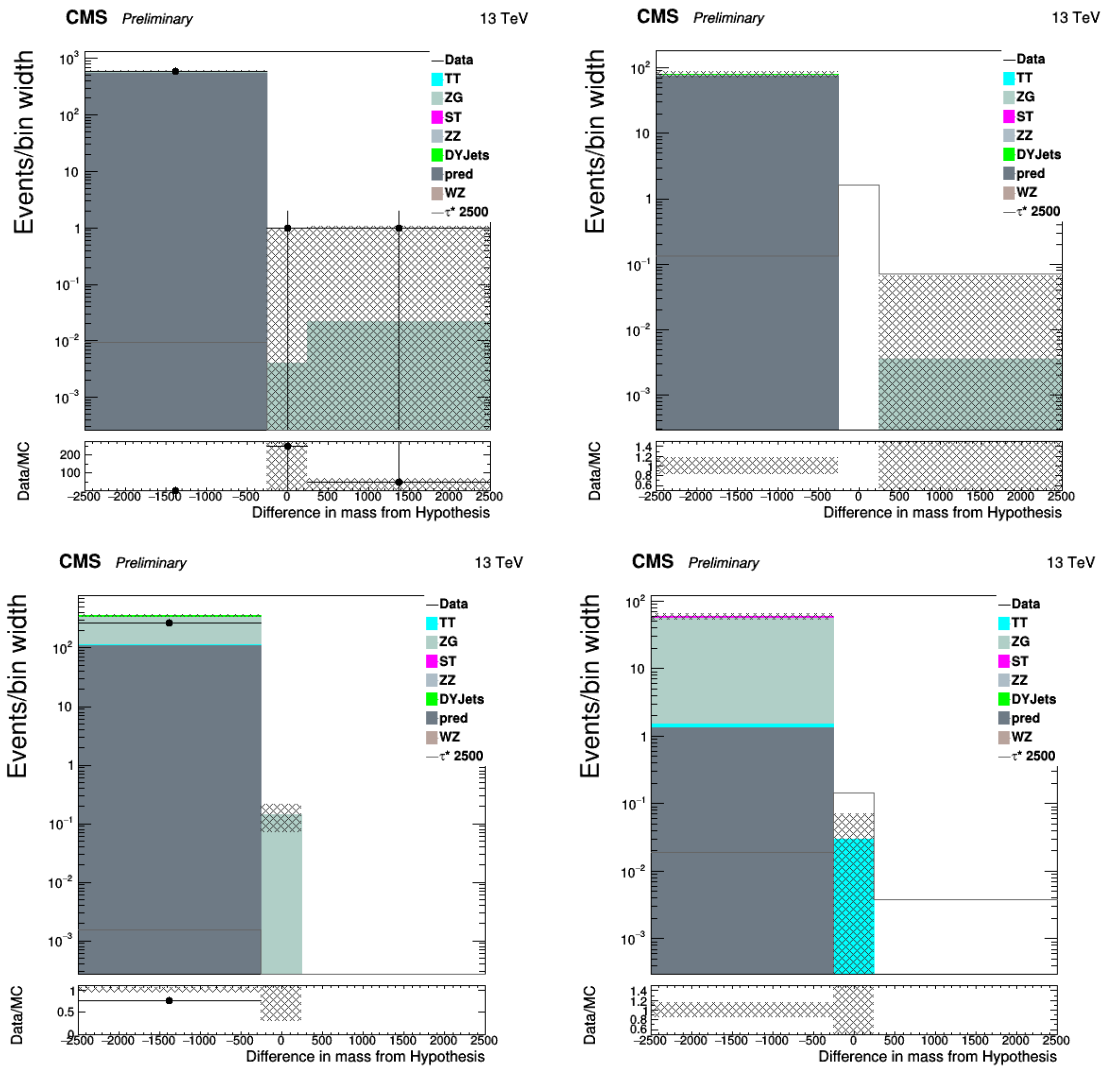

Figure 7.18: Final event totals for the $\tau_{\mathrm{h}} + \tau_{\mathrm{h}}$ channel in Region A of the ABCD method. The L-shaped window on the 2D mass plane of $\tau + \gamma$ is reduced to one dimension by taking the closer reconstructed mass to 2500 GeV of the two possible pairings. The four plots show the results for one-prong τ_{h} decays (top row) and three-prong τ_{h} decays (bottom row) as well as low-photon-p_{T} sideband (left) and the high-photon-p_{T} region (right). Event totals in the A region of the high-photon-p_{T} region are masked to stay consistent with the current blinding policy of the analysis.

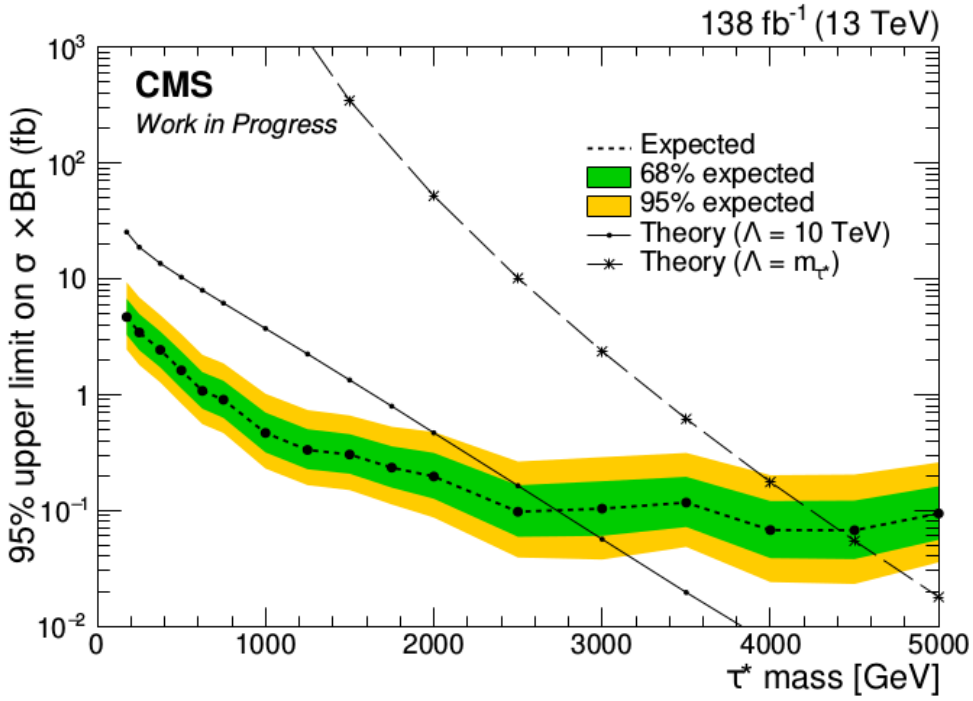

Figure 7.19: Expected limits on τ^* mass when performing a simultaneous fit of all three signal channels (e + τ_{h}, μ + τ_{h}, τ_{h} + τ_{h}) in the high- and low-photon-p_{T} regions and the non-hadronic (e + e, e + μ, μ + μ) low-photon-p_{T} regions. For the case where the compositeness scale $\Lambda = 10\,\mathrm{TeV}$, this analysis would exclude masses below around 2750 GeV (2.75 TeV) at the 95 % confidence level. For $\Lambda = \mathrm{m}_{\tau^*}$, or the scale equal to the resonant mass of the τ^*, the exclusion would rise to around 4500 GeV (4.5 TeV) at the 95 % confidence level.

Chapter 8

Summary and Conclusions

The Standard Model has produced several important predictions that have been verified through experiment. Many questions still linger about what the Standard Model does not have the tools to answer. The theory of compositeness of quarks and leptons attempts to provide an answer to some of those unaddressed questions. One such particle that would be produced in the event of compositeness is the excited τ, or τ^*.

This analysis, and this dissertation, search for the presence of such a particle using the data collected by the CMS experiment at the LHC. These events span three years of data collection (138 fb^{-1}) from 2016-2018 at a center-of-mass energy of 13 TeV. This increased energy of Run II allows analyses to push the envelope even farther in search of new particles, and if failing to find them, the limits that can be set on their properties and their existence. This dissertation presents the expected findings of an analysis whose compositeness scale was unattainable previously. The simulation of both signal at various mass points and various background processes compared to the data collected shows the effectiveness of the analysis techniques employed here to search the τ^*. In the end, the limits for a τ^* can be set at around 2750 GeV once the unblinding of the analysis has taken place. For direct comparison with previous limits set for electrons and muons whose limits from other previous analysis is 3.8 or 3.9 TeV when $\Lambda = \text{m}_{\tau^*}$, the limit from this analysis is expected to be 4.5 TeV.

Bibliography

[1] S. Chatrchyan, V. Khachatryan, A.M. Sirunyan, A. Tumasyan, W. Adam, E. Aguilo, T. Bergauer, M. Dragicevic, J. Erö, C. Fabjan, and et al. Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc. *Physics Letters B*, 716(1):30â€“61, Sep 2012.

[2] Jogesh C. Pati, Abdus Salam, and J. A. Strathdee. Are quarks composite? *Phys. Lett. B*, 59:265, 1975.

[3] H. Terazawa, Masaki Yasuè, Keiichi Akama, and Masaki Hayshi. Observable effects of the possible substructure of leptons and quarks. *Phys. Lett. B*, 112:387, 1982.

[4] M. Abolins, B. Blumenfeld, E. Eichten, H. Kagan, K. Lane, J. Leveille, D. Pellett, M. Peskin, and J. Wiss. Testing the Compositeness of Quarks and Leptons. In *Elementary Particles and Future Facilities (Snowmass 1982)*, page 274, 1982. eConf C8206282.

[5] E. Eichten, Kenneth D. Lane, and Michael E. Peskin. New Tests for Quark and Lepton Substructure. *Phys. Rev. Lett.*, 50:811, 1983.

[6] H. Harari. Composite models for quarks and leptons. *Physics Reports*, 104:159, 1984.

[7] K. D. Lane, F. E. Paige, T. Skwarnicki, and W. J. Womersley. Simulations of supercollider physics. *Phys. Rept.*, 278:291, 1997.

[8] U. Baur, M. Spira, and P. M. Zerwas. Excited quark and lepton production at hadron colliders. *Phys. Rev. D*, 42:815, 1990.

[9] O. W. Greenberg and C. A. Nelson. Composite models of leptons. *Phys. Rev. D*, 10:2567, 1974.

[10] O. W. Greenberg and Joseph Sucher. A quantum structure dynamic model of quarks, leptons, weak vector bosons, and Higgs mesons. *Phys. Lett. B*, 99:339, 1981.

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

[12] M.E. Peskin and D.V. Schroeder. *An Introduction To Quantum Field Theory*. Frontiers in Physics. Avalon Publishing, 1995.

[13] Richard D. Ball et al. Parton distributions from high-precision collider data. *Eur. Phys. J. C*, 77:663, 2017.

[14] Eric Drexler. Elementary particle interactions in the standard model, 2014. Distributed under a CC0 1.0 license.

[15] CMS Collaboration, A. M. Sirunyan, A. Tumasyan, W. Adam, F. Ambrogi, E. Asilar, T. Bergauer, J. Brandstetter, M. Dragicevic, J. Erö, and et al. Search for excited leptons in $\ell\ell\gamma$ final states in proton-proton collisions at $\sqrt{\mathrm{s}} = 13$ TeV. *Journal of High Energy Physics*, 2019(4), Apr 2019.

[16] CMS Collaboration, A. M. Sirunyan, A. Tumasyan, W. Adam, F. Ambrogi, E. Asilar, T. Bergauer, J. Brandstetter, M. Dragicevic, J. Erö, and et al. Search for an excited lepton that decays via a contact interaction to a lepton and two jets in proton-proton collisions at $\sqrt{\mathrm{s}} = 13$ TeV. *Journal of High Energy Physics*, 2020(5), May 2020.

[17] D. Buskulic et al. Search for excited leptons at 130–140 GeV. *Phys. Lett. B*, 385:445, 1996.

[18] P. Abreu et al. Search for composite and exotic fermions at LEP 2. *Eur. Phys. J. C*, 8:41, 1999.

[19] G. Abbiendi et al. Search for unstable heavy and excited leptons at LEP 2. *Eur. Phys. J. C*, 14:73, 2000.

[20] P. Achard et al. Search for excited leptons at LEP. *Phys. Lett. B*, 568:23, 2003.

[21] G. Aad, B. Abbott, J. Abdallah, S. Abdel Khalek, O. Abdinov, R. Aben, B. Abi, M. Abolins, O. S. AbouZeid, and et al. Search for new phenomena in events with three or more charged leptons in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *Journal of High Energy Physics*, 2015(8), Aug 2015.

[22] CMS Collaboration. Picture taken from the public plots available at https://cmslumi.web.cern.ch/publicplots/multiYear/int_lumi_allcumulative_pp_run2.png.

[23] Jean-Luc Caron. Overall view of LHC experiments.. Vue d'ensemble des experiences du LHC. AC Collection. Legacy of AC. Pictures from 1992 to 2002., May 1998.

[24] David Barney. CMS Detector Slice. CMS Collection., Jan 2016.

[25] Tai Sakuma. Cutaway diagrams of CMS detector. May 2019.

[26] W. et al. Adam. The CMS Phase-1 Pixel Detector Upgrade. *JINST*, 16:P02027. 84 p, Dec 2020.

[27] Matteo Centis Vignali. Silicon Sensors for the Upgrades of the CMS Pixel Detector, 2015. presented 2 Dec 2015.

[28] Ivan Shvetsov. Status of the CMS silicon strip detector. Technical report, CERN, Geneva, Jun 2022.

[29] N. Akchurin, L. Berntzon, A. Cardini, R. Ferrari, G. Gaudio, J. Hauptman, H. Kim, L. La Rotonda, M. Livan, E. Meoni, H. Paar, A. Penzo, D. Pinci, A. Policicchio, S. Popescu, G. Susinno, Y. Roh, W. Vandelli, and R. Wigmans. Contributions of cherenkov light to the signals from lead tungstate crystals. *Nuclear Instruments and*

Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 582(2):474–483, 2007.

[30] The CMS Collaboration and S Chatrchyan et al. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08004–S08004, Aug 2008.

[31] Jean-Baptiste Sauvan et al. Performance and upgrade of the CMS electron and photon trigger for run 2. *Journal of Physics: Conference Series*, 587:012021, Feb 2015.

[32] Performance of reconstruction and identification of tau leptons in their decays to hadrons and tau neutrino in LHC Run-2. Technical report, CERN, Geneva, 2016.

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

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

[35] Wolfgang Adam, Boris Mangano, Thomas Speer, and Teddy Todorov. Track Reconstruction in the CMS tracker. Technical report, CERN, Geneva, 2006.

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

[37] Albert M Sirunyan et al. Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC. *JINST*, 16:P05014, 2021.

[38] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_{T} jet clustering algorithm. *JHEP*, 04:063, 2008.

[39] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. Fastjet user manual. *Eur. Phys. J. C*, 72:1896, 2012.

[40] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The catchment area of jets. *JHEP*, 04:005, 2008.

[41] Matteo Cacciari and Gavin P. Salam. Pileup subtraction using jet areas. *Phys. Lett. B*, 659:119, 2008.

[42] Vardan Khachatryan et al. Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV. *JINST*, 12:P02014, 2017.

[43] CMS Collaboration. Jet algorithms performance in 13 TeV data. CMS Physics Analysis Summary CMS-PAS-JME-16-003, 2017.

[44] Albert M Sirunyan et al. Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector. *JINST*, 14:P07004, 2019.

[45] A. M. Sirunyan et al. Performance of reconstruction and identification of τ leptons decaying to hadrons and ν_τ in pp collisions at $\sqrt{s} = 13$ TeV. *JINST*, 13:P10005, 2018.

[46] Torbjörn Sjöstrand, 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, 2015.

[47] 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.

[48] Stefan Hoeche, Frank Krauss, Nils Lavesson, Leif Lonnblad, Michelangelo Mangano, Andreas Schaelicke, and Steffen Schumann. Matching parton showers and matrix elements. 2006.

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

[50] Paolo Nason. A new method for combining NLO QCD with shower monte carlo algorithms. *JHEP*, 11:040, 2004.

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

[52] 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.

[53] Pierre Artoisenet, Rikkert Frederix, Olivier Mattelaer, and Robbert Rietkerk. Automatic spin-entangled decays of heavy resonances in monte carlo simulations. *Journal of High Energy Physics*, 2013(3), Mar 2013.

[54] Albert M Sirunyan et al. Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements. *Eur. Phys. J. C*, 80:4, 2020.

[55] S. Agostinelli et al. GEANT4—a simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250, 2003.

[56] E. Bols, J. Kieseler, M. Verzetti, M. Stoye, and A. Stakia. Jet flavour classification using DeepJet. *Journal of Instrumentation*, 15(12):P12012–P12012, dec 2020.

[57] Daniele Bertolini, Philip Harris, Matthew Low, and Nhan Tran. Pileup per particle identification. *JHEP*, 10:059, 2014.

[58] Albert M Sirunyan et al. Identification of hadronic tau lepton decays using a deep neural network. Submitted to *JINST*, 2022.

[59] V. Khachatryan, A.M. Sirunyan, A. Tumasyan, W. Adam, E. Asilar, T. Bergauer, J. Brandstetter, E. Brondolin, M. Dragicevic, J. Ero, and et al. The cms trigger system. *Journal of Instrumentation*, 12(01):P01020–P01020, Jan 2017.

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

[61] Roger Barlow and Christine Beeston. Fitting using finite monte carlo samples. *Computer Physics Communications*, 77(2):219–228, 1993.

[62] J. S. Conway. Incorporating nuisance parameters in likelihoods for multisource spectra, 2011.