Search for vector-like leptons in final states with multiple leptons

A thesis

Submitted in partial fulfillment of the requirements of the degree of

Doctor of Philosophy

by

Anshul Kapoor

Registration ID-20132015

Department of Physics

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH

PUNE-411008, India

2019

Certificate

Certified that the work incorporated in the thesis entitled **"Search for vector-like leptons in final states with multiple leptons"** submitted by **Anshul Kapoor** was carried out by the candidate, under my supervision. The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other university or institution.

Date: 22-August-2019

Dr. Sourabh Dube

(Supervisor)

Declaration

I declare that this written submission represents my ideas in my own words and where others ideas have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Date: 22-August-2019   Anshul Kapoor

                                                                                        Registration ID-20132015

Acknowledgments

This thesis has been possible because of contributions and support from so many people in my professional and personal circle. During good times, and bad, these people have provided an extraordinary amount of support and motivation. I would fail in my duty if I do not acknowledge them.

My supervisor, Prof. Sourabh Dube, has been the most amazing person to learn from and work with. It is only because of your constant motivation that I could scramble through these years. You have not only been an academic supervisor but a friend and a confidant. You are also the nicest person I know apart from being an undoubtedly extraordinary physicist. Thank you for letting me be who I am and to work at my pace. Thank you for always encouraging all my ideas and for motivating me to pursue them. Thank you for having my back at all times. Thank you for believing in me. Thank you, for making me both, a physicist and a better human being. And, thank you so much, for never giving up on me.

So many professors have contributed to my academic and personal growth. Prof. Sunil Somalwar's patient discussions have helped me tremendously as a researcher. Prof. Seema Sharma gave me the much-needed reality checks. Prof. Arun Thalapillil helped review my work and always supported me. Prof. Eva Halkiadakis frequently reviewed my academic progress. Prof. Sunil Mukhi; who I have had the privilege to interact frequently. I am also grateful for all the support from Prof. Paolo Giacomelli, Prof. G.V. Pawan Kumar, Prof. C.V. Dharmadhikari. Thank you all.

Many collaborators have contributed significantly to my academic journey. Lukas, Halil, Sezen, Mathias, Maxi, and Peter. Thank you all. Special thanks to my IISER HEP colleagues, Shubhanshu, Kunal, Angira, Steenu, Arnab, Shubham, Vinay, Irene, Divya, and Aditee, who helped me sail through the academic journey.

Thanks to my closest friends; Akash, Deepak, Vivek, Mehak, Aditi, Kashyap, and Geetanjali. For when I faltered, you were there to pick me up. You are all awesome.

This thesis is dedicated to my mother

Abstract

 The standard model of particle physics stems from the efforts of understanding the ultimate constituents of matter. It has been exceptionally successful in explaining various experimental observations during the last 50 years. Despite being successful, the standard model is incomplete since it fails to answer many open questions. We have failed to develop a successful quantum theory of gravity and connect it to the other three fundamental forces within the framework of the standard model. It is also not understood how the mass of Higgs boson sits at 125 GeV despite it getting significant corrections from all the particles that the Higgs boson couples to. The standard model also does not explain the presence of dark matter in the universe.

Many theoretical models have been proposed to account for the inadequacies in the standard model of particle physics. If there is any truth in these propositions, signs of these might show up in the debris of proton-proton collisions in the Large Hadron Collider (LHC) at CERN. The typical way of testing a new theory is to search for one or more hypothesized particles predicted by the theory. Many such efforts have been carried out, but none have found any new particles. While it is possible that the new particles are extraordinarily massive and hence are out of the reach of current energies of the LHC, it is also possible that we are just not looking for the right particles or in the right manner.

In this thesis, I have presented a search for an SU(2) doublet vector-like lepton extension of the standard model with couplings to the third generation standard model leptons. The last search for such a doublet was carried out by the L3 collaboration in 2001, where a lower bound of $\approx$100 GeV was placed on the mass of these particles. The data sample corresponds to $77\,\mathrm{fb}^{-1}$ of integrated luminosity in pp collisions at $\sqrt{s} = 13$ TeV collected by the CMS experiment at the Large Hadron Collider in 2016 and 2017. This is the first search for a vector-like lepton doublet in any of the LHC experiments.

Events are primarily categorized based on the multiplicity of light leptons and taus.

The missing transverse energy and the scalar sum of transverse momenta of leptons including tau leptons, are used to discriminate the signal model against standard model backgrounds. The observations are consistent with the expectations from the standard model only hypothesis, and the existence of vector-like leptons in the mass range of 150-790 GeV is excluded at 95% confidence level.

I have also presented a technique based on convolutional neural networks to improve the sensitivity of the search by better optimizing signal-to-background discrimination. The observations show that this new method provides significant enhancement to the signal-to-background discrimination and also comfortably beats a conventional neural-network based optimization, in most cases.

List of Figures

List of Tables

Contents

CHAPTER 1

Introduction

The story of human endeavor to understand the working principles of nature goes back to a time when first humans existed on this planet. Humans were given the gift to observe, analyze, and interpret the phenomena they saw happening around. It is incomprehensible how much we have achieved in our quest to understand nature. From walking first time on this planet to sending a spacecraft beyond the edge of the solar system, the story is of an immensely successful enterprise. Over centuries, the scientific method of interpreting and reporting particular phenomena of nature has gone from rudimentary observation to a peer review based rigorous process. Yet, scientific endeavor continues with all its might.

One of the oldest quests of the human enterprise has been to understand nature's most fundamental building blocks. It has led us to some of the most startling discoveries about the universe. Observations like the nature of matter as a particle-wave dual, complexity of the atomic structure, hierarchy in strengths of the four fundamental interactions in nature and presence of more than a few unique fundamental particles, have all led to at-least as many questions as the answers they have provided.

As of today, there is strong evidence that visible matter might be $<5\%$ of the total contents of the universe. As humans, we barely understand visible matter and its properties, and it is overwhelming to think about what we have achieved with just that amount of understanding. We discovered the extraordinary applications of materials; we built cars, rockets, and satellites, even invented computers, and these days, one talks about advancements in artificial intelligence. It is then unimaginable what we can achieve if we understand the rest of the universe.

One such experiment built to understand the inner workings of the universe is the Large Hadron Collider (LHC). It is located beneath the Swiss-French border and runs across a circumference of 27 km. Superconducting magnets steer two proton beams

traveling in opposite directions and bring them to a collision at specific points along the collider. The different points of collisions are also sites for particle detectors including the multipurpose ATLAS and CMS, which detect products of the collision. The proper detection, reconstruction, and analysis of the proton-proton collision debris are what forms the majority of a high energy physicist's research, working on these detectors. The hope is that sometimes one might discover new physical phenomena (new physics) which may fill gaps in our understanding of nature at the most fundamental level.

This thesis describes a search for new physics using data collected by the CMS detector. The CMS detector is a multipurpose particle detector made up of smaller sub-detectors, each designed to identify and measure specific properties of individual particles. When the collision debris hits the different sub-detectors, various kinds of digital signals are recorded, which are then reconstructed to give us insight into what might have happened in the collision.

Each pp beam contains approximately 1.1×10^{11} protons. When two beams are brought to a collision, only rarely do two protons within the beams collide in a head-on fashion which can lead to a high energy collision and result in high energy particles as the products of the collisions. In such a collision *event*, there might occur interesting phenomena, like the production of massive bosons and fermions. Some of these events might then contain signatures of new physics. This thesis describes a search for hypothesized vector-like leptons in the debris of those high energy collisions.

CHAPTER 2

Standard Model and Beyond

Matter around us is made up of spin-$\frac{1}{2}$ fermions, namely quarks, and leptons. These particles can communicate through four kinds of interactions: gravitational, electromagnetic, strong, and weak interactions. Physics of all but the gravitational interaction is embedded into a common framework known as the standard model of particle physics (SM). Developed in the 1970s, the validity of SM has been tested by various experiments [1], and it has stood the test of time. It is thus a theory which has had monumental success in explaining natural phenomena. It explains phenomena from the Planck scale to the scale of the universe.

2.1 Units

In this thesis, I will work with natural units, where both the speed of light and the Planck's constant are set as unity ($\hbar = c = 1$). Mass, momentum, and energy are all expressed in GeV. Cross-sections are expressed in barns and luminosities are expressed in barns^{-1} (1 barn = $10^{-28}\mathrm{m}^2$).

2.2 Lagrangian and Gauge Symmetries

Any physical process understood by a Quantum field theory (QFT) is described via the formalism of Lagrangian mechanics, where the dynamics of a system are described by the Lagrangian $\mathcal{L}(\phi, \partial_\mu \phi)$, where ϕ is the field of interest.

In absence of any other fields, the Lagrangian of a fermionic field, ψ can be written as

$$\mathcal{L}_{free} = i\bar{\psi}\gamma^\mu \partial_\mu \psi. \tag{2.1}$$

The Lagrangian exhibits an interesting property here. If we give the field a complex

3

rotation

$$\psi \to \psi' = e^{-i\lambda} \tag{2.2}$$

the Lagrangian remains invariant. This is referred to as a global U(1) symmetry. The term global here means that the *gauge* λ does not depend on space-time and is a constant. Under a local gauge transformation, say $e^{-i\lambda(x)}$, the Lagrangian won't be invariant anymore.

Although, the extra terms that arise because of local transformation can be absorbed by the Lagrangian, by introduction of a new field A^μ, which transforms as

$$A'^\mu = A^\mu - k\partial_\mu \lambda(x). \tag{2.3}$$

where k is a proportionality constant and $k\partial_\mu \lambda(x)$ signifies any constant gauge transformation.

One could then write the Lagrangian as

$$\mathcal{L} = i\bar{\psi}\gamma^\mu D_\mu \psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.4}$$

where D_μ is the covariant derivative defined as $\partial_\mu + igA_\mu(x)$ and $F_{\mu\nu}$ is the field strength tensor, defined as $\partial_\mu A_\nu - \partial_\nu A_\mu$. The above is the Lagrangian for a massless electron in quantum electrodynamics (QED), where $g = e$. With a little work, it can be seen from the Lagrangian that not only does it exhibit a U(1) global symmetry but also a local gauge symmetry.

Similarly, the Lagrangian of quantum chromodynamics (QCD), for free quarks, can be written as

$$\mathcal{L} = i\bar{\psi}_q \gamma^\mu \partial_\mu \psi_q \tag{2.5}$$

where ψ_q is a triplet of quark fields ψ_q. This Lagrangian is invariant under global $SU(3)_c$ symmetry group, which comprises of 8 different parameters, unlike a single rotation parameter for U(1). To make this a local symmetry, we may write the Lagrangian as

$$\mathcal{L} = i\bar{\psi}_q \gamma^\mu \partial_\mu \psi_q - \frac{1}{2}G^{a,\mu\nu}G^a_{\mu\nu} \tag{2.6}$$

where similarly to U(1), where A is the photon field, G_a corresponds to eight vector

fields that are the gluons. Each gluon will result in a field strength tensor.

A "symmetry" is thus the presence of a transformation that preserves relevant physical properties of a system. More precisely, a system's symmetry transformation (such as particle positions or field value at each space-time point) leaves the action invariant. The action is constructed from the Lagrangian, so symmetry can also be defined as a transformation that leaves the Lagrangian invariant.

As described above, a global symmetry has a parameter that is constant throughout space-time. The symmetry transformation is performed at every point by the same "amount". A local symmetry has a parameter that depends on space-time. Therefore, a global symmetry is a "true" symmetry. It does not reduce the system's degrees of freedom, but only corresponds to the quantities that are conserved. A gauge symmetry is one that relates states that are physically identical. The states it connects are then only different in their mathematical structure. The existence of a gauge symmetry thus reveals redundancy in description: there is more mathematical description than is necessary to describe the physical reality.

Every physical process will have its own Lagrangian that governs the dynamics of the system described by a QFT. Particles are then excitations of those quantum fields. Every Lagrangian can also have one or more associated symmetries.

2.3 Electroweak unification

The electromagnetic interaction can be completely described by the Quantum Electrodynamics (QED) [2]. In QED, the interaction of electrically charged particles happens through the exchange of photons. The gauge field A is gauge invariant under the U(1) symmetry group of electric charge q.

In the 1930s, Enrico Fermi proposed a theory of four-fermion interaction involving a contact force, to explain beta decays [3]. By 1950s, there was an effective QFT of weak interaction known as Quantum Flavour Dynamics (QFD) [4]. The gauge invariance was under the $SU(2)_I$ symmetry group of the weak isospin (I). Later (around the 1960s), QED and QFD were unified into the electroweak (EW) theory [5–7]. The gauge invariance came from a combination of U(1) and SU(2) symmetry groups into one new

symmetry group, $SU(2)_L \otimes U(1)_Y$. This merger of the electromagnetic interaction and weak interaction into one common framework is referred to as the electroweak unification. The electroweak Lagrangian can be written as

$$\mathcal{L}_{EW} = \sum_{j=1}^{3} i\bar{\psi}_j(x)\gamma^\mu D_\mu \psi_j(x) - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W_{\mu\nu}^j W_j^{\mu\nu} \qquad (2.7)$$

- The first term $\sum_{j=1}^{3} i\bar{\psi}_j(x)\gamma^\mu D_\mu \psi_j(x)$ represents the fermion kinetic energy and their gauge field interaction.

- The last two terms represent the gauge field kinetic energy and self-interaction. The three gauge fields of the weak isospin for the SU(2) symmetry group are W^j (j = 1,2,3). B_μ is the massless gauge field of the weak hypercharge Y from the U(1) symmetry group.

- The gauge fields W^j and B_μ can be transformed into mass eigenstates, the $W^\pm$, Z and γ bosons.

2.4   The Standard Model group structure

SM is a gauge-invariant relativistic quantum field theory. Interactions within SM can be described by the gauge group

$$U(1)_Y \otimes SU(2)_L \otimes SU(3)_C \qquad (2.8)$$

The electroweak force, defined by groups U(1) ⊗ SU(2), interacts with the weak hyper-charge (Y) and on left-handed fermions (L). The strong force, defined by the SU(3) group, interacts with particles that have a color charge (C). The interactions are also thus referred to as gauge interactions. Within this theory, lie a variety of fundamental particles, shown in Fig. 2.1.

The elementary particles carry either integer spins or half-integer spins and are referred to as bosons or fermions respectively. All of the fermions (leptons and quarks) participate in one or more of the fundamental interactions in nature. The gauge interactions are mediated by bosons: gluons and photons, respectively, for strong and electromagnetic interactions and the W and Z bosons for the weak interaction. The
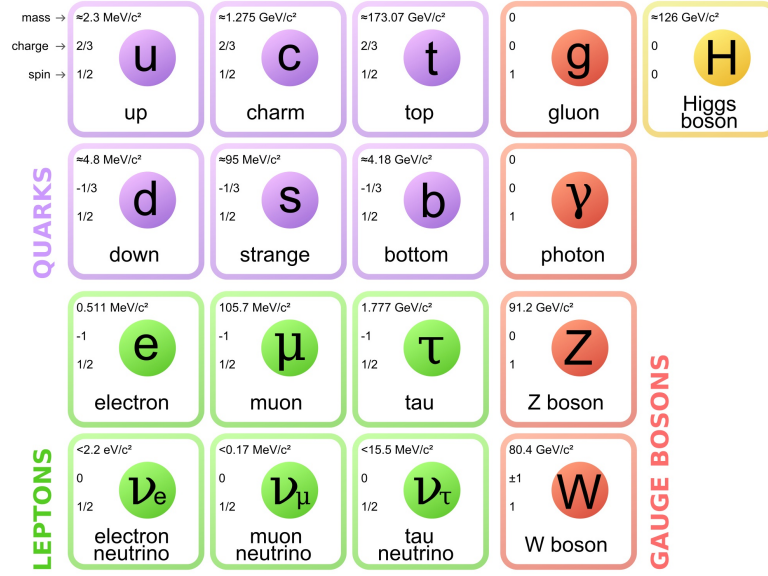
Figure 2.1: The particle content of the standard model of particle physics

Higgs boson H provides mass to fermions through the Higgs mechanism described in Sec 2.6.

2.5   SM Lagrangian

The SM Lagrangian is a combination of four different field sectors, namely quark fields, lepton fields, vector gauge-boson fields, and the scalar Higgs field,

$$\mathcal{L}_{SM} = \mathcal{L}_{gauge} + \mathcal{L}_{leptons} + \mathcal{L}_{quarks} + \mathcal{L}_{scalar} \tag{2.9}$$

where $\mathcal{L}_{gauge}$ describes the dynamics of spin-1 vector bosons. $\mathcal{L}_{quark}$ describes the propagation of free quarks and their interactions with vector gauge fields. $\mathcal{L}_{leptons}$ contains free lepton terms along with their interaction to electroweak fields only since leptons do not interact strongly. $\mathcal{L}_{scalar}$ describes the interaction of the spin-0 Higgs boson with other particles and itself.

2.6   Electroweak symmetry breaking

The electromagnetic and weak forces, as we know them, are two different forces with entirely different properties. Hence the electroweak symmetry was somehow broken. It

was known that local gauge invariance $U(1)_Y \otimes SU(2)_L$ forbids massive gauge bosons and fermions but the force carriers, the W and the Z bosons, were known to be massive. It was thus a mystery as to how the electroweak symmetry was broken and thus how the W and the Z bosons acquired mass.

This was done by introducing a new field with a very specific potential. The new field keeps the full Lagrangian invariant under $SU(2)_L \otimes U(1)_Y$, but it will cause the vacuum to be not invariant under this symmetry.

2.6.1 Abelian example

Let us start with the example of a local U(1) phase symmetry. U(1) is an Abelian gauge symmetry, which means that all elements of the group commute. Consider a complex scalar field $\Phi(x)$ of electric charge q coupled to the electromagnetic field $A^\mu(x)$; the Lagrangian is

$$\mathcal{L}_{KE} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - D_\mu\Phi^* D^\mu\Phi - V(\Phi) \tag{2.10}$$

where $D_\mu = \partial_\mu - ieA_\mu$ and $V(\Phi) = -\mu^2\Phi^*\Phi + \lambda(\Phi^*\Phi)^2$. This Lagrangian remains invariant under the gauge transformations $A_\mu(x) \rightarrow A_\mu(x) - \partial_\mu\eta(x)$ and $\Phi(x) \rightarrow e^{ie\eta(x)}\Phi(x)$.

If $\mu^2 < 0$ then $\Phi = 0$ will be the minimum energy state and the potential preserves the symmetries of the Lagrangian. The theory is then plain QED along with a massless photon and a massive field Φ with mass μ.

However, if $\mu^2 > 0$, the minimum energy state with correspond to a non-zero value. This is referred to as the vacuum expectation value (VEV)

$$\Phi_{VEV} = \sqrt{\frac{\mu^2}{2\lambda}} = \frac{\nu}{\sqrt{2}} \tag{2.11}$$

and the global U(1) symmetry will be spontaneously broken. The Φ can be recast as

$$\Phi = \frac{\nu + h}{\sqrt{2}}e^{i\frac{\chi}{\nu}} \tag{2.12}$$

where h and χ are real scalar fields which have no VEVs. These are referred to as the Higgs and the Goldstone bosons. With this, the Lagrangian in eq. 2.10 describes a theory with a massive photon of mass $e\nu$, a Higgs boson H with mass $\sqrt{2}\mu$, and a

massless Goldstone χ. This mechanism of generating masses for bosons is referred to as the Higgs mechanism.

2.6.2 The Higgs mechanism in the electroweak standard model

The electroweak standard model is based on the $SU(2) \otimes U(1)$ Lagrangian given in eq. 2.9. Thus we need a mechanism that can give masses to the W and Z bosons taking into account the $SU(2) \otimes U(1)$ symmetry. A possible mechanism to explain mass generation without breaking the gauge invariance was given independently by Peter Higgs [8] and, by Robert Brout and Franois Englert [9].

The starting point is the Higgs potential

$$V(\Phi) = -\mu^2 \Phi^* \Phi + \lambda (\Phi^* \Phi)^2 \tag{2.13}$$

and doublet field

$$\Phi = \left\{ \begin{matrix} \Phi^+ \\ \Phi^- \end{matrix} \right\} \tag{2.14}$$

where each component is a complex scalar field. The Higgs potential is symmetric with respect to the origin, but choosing a vacuum state with a finite VEV breaks this symmetry. Inspired from eq. 2.11, the vacuum state is chosen as

$$\Phi_0 = \frac{1}{\sqrt{2}} \left\{ \begin{matrix} 0 \\ \nu \end{matrix} \right\} \tag{2.15}$$

The field is then given by

$$\Phi = \frac{1}{\sqrt{2}} \left\{ \begin{matrix} 0 \\ \nu + H(x) \end{matrix} \right\} \tag{2.16}$$

and the Higgs mass is $m_H = \sqrt{2}\mu$. The coupling of the W and the Z bosons to the Higgs field generate their masses

$$m_W = \frac{1}{2} g \nu \tag{2.17}$$

and

$$m_Z = \frac{m_W}{\cos \theta_W} \tag{2.18}$$

where g is the coupling and θ_W is the weak mixing angle. The Higgs boson has since been found at $m_H = 125.18$ GeV and has added to the robustness of the standard model as a successful theory of nature.

2.7 Problems with Standard Model

While the SM has been extraordinarily successful, it is still known to be incomplete as it can not explain many known observations.

- The most apparent problem is the absence of gravity from the SM. So far, no consistent quantized description of the gravitational interaction has been formulated, and the SM is even believed to be incompatible with the theory of general relativity.

- A variety of astronomical observations prove that only about 5% of the energy content of the universe can be explained by ordinary matter described by the SM [10]. At the moment, there is no way to incorporate dark matter inside the SM except to extend it by the introduction of new symmetries like supersymmetry.

- Another critical aspect is the naturalness (sometimes referred to as hierarchy) problem. The basic argument is that if new physics was to appear first at the Planck scale, then new massive particles (M<<new scales) should exist, coupling to the Higgs field. These new particles will induce loop corrections to the physical mass of Higgs, and these correction terms would be huge. Such corrections should cancel in a fine-tuned manner for the physical mass of the Higgs to be at 125.18 GeV. Since such level of fine-tuning seems quite unnatural for a fundamental theory, this problem is known as the naturalness problem.

2.8 Vector-like leptons

In the last few decades, many extensions of the SM have been proposed to solve some of these problems. Almost all strategies postulate the possibility of new particles in nature. In one class of new particles, there are nonchiral color singlet fermions that couple to

the SM leptons. The term nonchiral implies that the left- and right-handed components of these particles transform identically under gauge symmetries. These particles are thus referred to as vector-like leptons (VLLs). They arise in a wide variety of models ranging from supersymmetry to extra dimensions [11–21]. VLLs are often classified by the SM generation that they will primarily couple to and have lepton numbers identical to that of SM leptons.

Just like the SM leptons, VLLs do not have any color charge, can be electromagnetically charged or neutral, and are classified in SM-like generations. This results in VLL having lepton numbers similar to the SM generation that they couple to. For example, if we add a vector-like electron (e') to the first generation of SM leptons, then this e' would have a corresponding electron number (L_e), and this would be conserved in all its interactions.

One may broadly classify VLL models into two categories: SU(2) singlet or SU(2) doublet models. The distinction between these models from an experimental point of view lies in the fact that the singlet VLL model allows only charged VLLs while the doublet VLL model allows charged and neutral VLLs. For example, only one charged vector-like electron (or muon or tau lepton), and its antiparticle to the first generation is added in the singlet model. On the other hand, the doublet model also adds one neutral vector-like neutrino along with its antiparticle. Therefore, there are two new particles in the singlet model, while the doublet model has four.

In this thesis, an SU(2) doublet VLL extension [22] of the SM has been considered, with couplings to the third generation SM leptons. A search is carried out in final states with multiple charged leptons (e, μ, τ), using proton-proton (pp) collision data collected by the CMS detector at the LHC in 2016 and 2017, in final states with multiple leptons (e, μ, τ). The model considered in this thesis introduces a vector-like τ lepton (τ'^{-}), its antiparticle (τ'^{+}), and the corresponding neutrinos (ν'_τ and $\overline{\nu}'_\tau$). At the LHC, they can be produced in $\tau'^{\pm}\nu'_\tau$, $\tau'^{+}\tau'^{-}$, and $\nu'_\tau\overline{\nu}'_\tau$ channels, with subsequent decays of τ' to Z τ or H τ and of ν'_τ to W τ, where W, Z, and H are the SM W, Z, and Higgs bosons. At tree-level, the τ' and ν'_τ are mass degenerate, whereas higher-order radiative corrections predict $<0.3\%$ relative mass splitting between these two states,

11

for VLL masses greater than 100 GeV. In this paper, τ' and ν'_τ are assumed to be mass degenerate. The mass of the VLL is the only free parameter both in the production cross section and in the branching fraction calculations. The branching fraction for these decays is shown in Fig. 2.2 [22]. The tree-level Feynman diagrams for associated and pair production of the doublet model VLLs are shown in Fig. 2.3 along with possible subsequent decay chains that would result in a multilepton (e, μ, τ, ν) final state. Based on motivation from these decay chains, this search for VLLs is carried out in final states with multiple leptons.
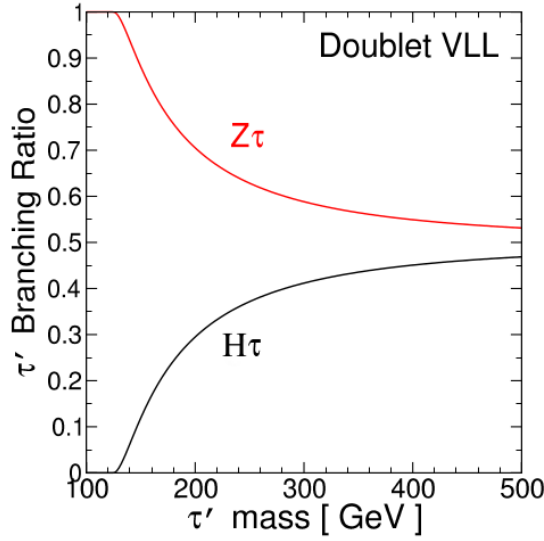
Figure 2.2: The branching ratios for $\tau' \to Z\tau, H\tau$, as a function of $m(\tau')$ for the doublet VLL [22].
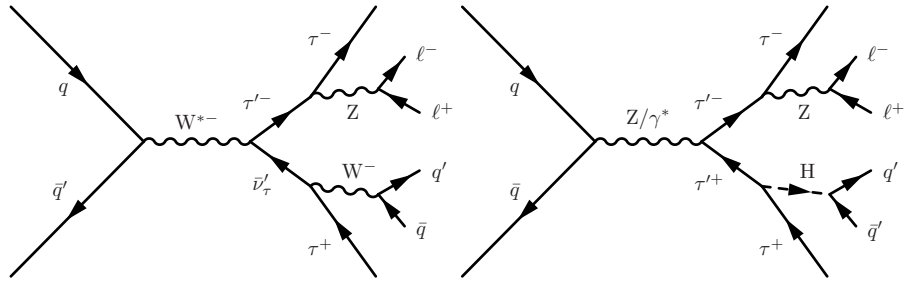
Figure 2.3: Two illustrative leading order Feynman diagrams for associated production of τ' with a ν'_τ (left) and for pair production of τ' (right), and possible subsequent decay chains that result in a multilepton final state.

The L3 collaboration placed a lower bound on additional heavy leptons around 100

GeV [23]. A search for heavy resonances decaying into a Z boson and a lepton has been performed by the ATLAS collaboration in a multilepton final state at $\sqrt{s} = 8$ TeV with 20.8 fb^{-1} of luminosity [24]. This search constrained the singlet VLL model and excluded VLL in the mass range of 114-176 GeV. However, there are no such constraints on the doublet VLL model described in Sec 2.8. Given the existing constraints, this search considered VLLs with masses over 100 GeV.

The most dominant SM backgrounds are the leptonic $t\bar{t}$, and $Z\gamma^*$ (DY) decays with an additional nonprompt (fake) lepton, usually originating from a heavy flavor jet. The other significant backgrounds consist of multiple prompt leptons from WZ, ZZ production and additional conversion leptons from prompt decays, like from DY+jets. These backgrounds are well modeled by Monte Carlo (MC) simulation; however, the backgrounds due to events with fake leptons from hadronic sources are estimated via a data-driven technique.

In the following chapters, I will describe the experimental setup, event simulation, event reconstruction, and the final analysis carried out for this search.

CHAPTER 3

Experimental setup

Particle colliders form the cornerstone of modern-day experimental particle physics. The principle behind the functioning of these colliders is the acceleration of particle beams using electric and magnetic fields. Broadly, a collider can either be linear or circular and can either collide two beams head-on or collide a single beam of particles with a stationary target.

3.1   The Large Hadron Collider

The LHC is a proton-proton collider, located underground, below the France-Switzerland border, with the physical structure running across a circumference of 27 km [25]. The tunnel that is used for the LHC contained the Large Electron-Positron (LEP) Collider until 2000. Around 2001, the LEP project was disassembled to make way for the LHC. The installation of LHC was completed in 2008 with the first collisions being recorded in 2009.

Today, the LHC is a host for proton-proton collisions reaching maximum energy of 13 TeV. The LHC itself is insufficient to accelerate protons to 13 TeV and hence is supported by the super proton synchrotron (SPS) which accelerates the proton beams to 450 GeV. The SPS itself was used to discover the W and Z bosons, back in 1983. A full schematic of the CERN accelerator complex is shown in Fig. 3.1.

At the structural level, the LHC is a two-ring pp collider, with both the rings housed in a single superconducting enclosure. The two rings independently contain proton beams, which only cross from one ring to the other at four locations. Each location houses one large experiment apart from other smaller experiments used for various purposes. The CMS detector is one of those experiments, which is a particle detector composed of many sub-detectors which are used to detect the pp collision debris.
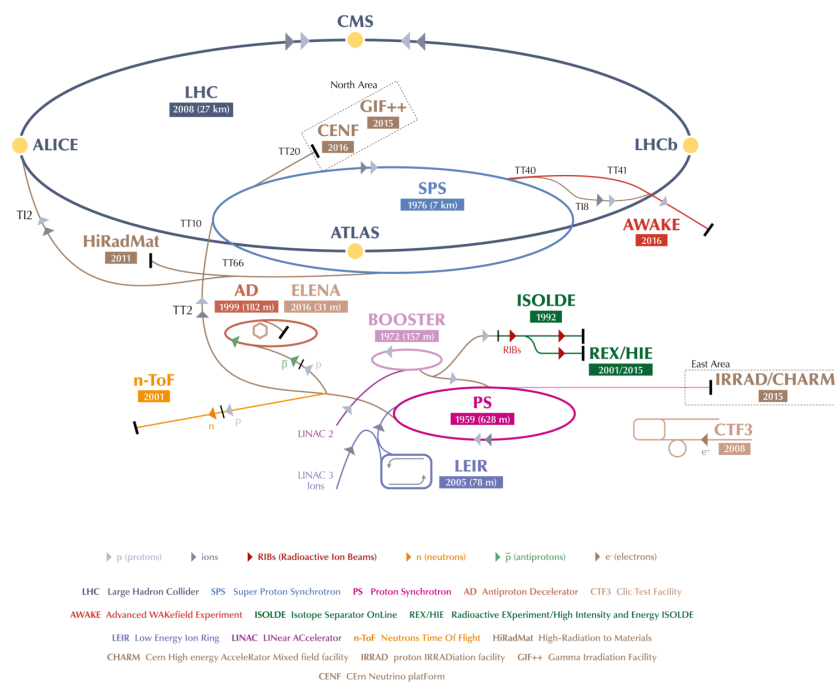
Figure 3.1: A schematic of the CERN accelerator complex showing the SPS and the LHC along with other smaller colliders and experiments. The locations of the CMS, ATLAS, LHCb, and ALICE experiments are also shown. This figure is from Ref [26].

CMS DETECTOR

Total weight   : 14,000 tonnes
Overall diameter : 15.0 m
Overall length   : 28.7 m
Magnetic field   : 3.8 T

STEEL RETURN YOKE
12,500 tonnes

SILICON TRACKERS
Pixel (100x150 μm) ~16m² ~66M channels
Microstrips (80x180 μm) ~200m² ~9.6M channels

SUPERCONDUCTING SOLENOID
Niobium titanium coil carrying ~18,000A

MUON CHAMBERS
Barrel: 250 Drift Tube, 480 Resistive Plate Chambers
Endcaps: 468 Cathode Strip, 432 Resistive Plate Chambers

PRESHOWER
Silicon strips ~16m² ~137,000 channels

FORWARD CALORIMETER
Steel + Quartz fibres ~2,000 Channels

CRYSTAL
ELECTROMAGNETIC
CALORIMETER (ECAL)
~76,000 scintillating PbWO₄ crystals

HADRON CALORIMETER (HCAL)
Brass + Plastic scintillator ~7,000 channels

Figure 3.2: The CMS detector shown along with all the sub-detectors.

3.2 The CMS detector

The CMS detector [27] is a multipurpose multi-detector experiment, where each sub-detector is designed such that it measures specific properties of individual particles. Particles from pp collisions travel through the detector encountering various sub-detectors in their path.

The central feature of the CMS detector is a superconducting solenoid that provides a 3.8 Tesla magnetic field. A silicon pixel and strip tracker, a lead tungsten crystal electromagnetic calorimeter (ECAL) and a brass scintillator hadron calorimeter. Muons are measured in gas-ionization detectors integrated into the steel flux-return yoke outside the solenoid. A visualization of the CMS detector (with labeled sub-systems) is shown in Fig. 3.2.

16

3.2.1   Coordinate System

A right-handed coordinate system is defined, with the z-axis being along the direction of the beam. In the transverse plane, the x-axis points towards the center of the LHC ring. The y-axis is then perpendicular to the x-axis, in the upward direction. The azimuthal angle ϕ is measured with respect to the positive x-axis in the x-y plane and the polar angle θ with respect to the z-axis, where θ is zero at the positive and π at the negative z-axis. The pseudorapidity η is defined as $\eta = -\ln(\tan(\frac{\theta}{2}))$. Differences in η are Lorentz–invariant. The transverse momentum p_T is the momentum perpendicular to the LHC beam axis, and the energy imbalance in the transverse plane is defined as the missing transverse energy p_T^{miss}. Angular distances in η and ϕ between physics objects originating at the interaction point are calculated by $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

3.2.2   Solenoidal magnet

In CMS, a superconducting solenoid magnet is used to measure the transverse momentum p_T of charged particles within the tracker. It provides a homogeneous magnetic field of B = 3.8 Tesla. The solenoid coil has a weight of 220 tons, is 12.48 meters long, 0.31 meters thick and has an inner diameter of 6.3 meters. Its size has been chosen to install all calorimeters system inside the solenoid. This configuration prevents particles from being scattered, absorbed, and showered in the solenoid material before entering the calorimeters. This could have otherwise caused degradation in the resolution of the energy measurements. One exception is the HCAL tail-catcher, which is installed centrally to ensure the measurement of high transverse momentum jets. The magnetic flux is returned through saturated iron, the return yoke. The yoke splits into three rings in the barrel and two endcap disks. The yoke is 15 meters across in diameter, and its axial length (including endcaps) is 21.5 meters. The yoke has a total weight of about 11,400 tons.

3.2.3   Silicon Tracker

After the resultant particles from the pp interaction have crossed the beam pipe, they interact with the silicon-based tracker which is immersed in a solenoidal magnetic field.

Information from various layers of the tracker can be used for track trajectory measurement for the charged particles. The tracker is further divided into an inner pixel and an outer strip detector.

The pixel detector [28] is made up of millions of pixel sensors. Each sensor is a silicon sensor of n+ implant on n bulk and a metalized p-backside. A charged particle deposits charge on the pixel sensors as it passes through the pixel tracker, via the production of electron-hole pairs. An electric current is then used to collect these charges on the surface of the sensor. Each sensor is itself made up of smaller cells and the charge deposited on each of which can be read out individually. On knowing the pixels that have been hit, the particle's trajectory can be deduced. Since the pixel sensor is a 2D tile, knowing locations of all the pixels that have been hit by the traversing charged particles, we can create a three-dimensional picture. The pixel detector sits closest to the interaction point and has to be fine-grained to see the individual hits in a dense environment of many hits.

On their way out of the pixel detector, charged particles encounter the strip detector [29]. It is made up of strip sensors. Since the particle volume inducing radiation damage decreases as we move in radii and the occupancy becomes less, silicon strip detectors can be used. In total, there are ten layers of silicon strips outside of the pixel-detectors in the barrel region. As endcaps, there are first the inner disks (three disks in both endcap regions) and then the outer disks (nine disks).

3.2.3.1 Upgrade of the pixel detector

The pixel detector was upgraded for the 2017 data collection [30]. The plot of R vs Z as a result simulation of charged particles passing through the old and new detector is shown in Fig. 3.3. The new pixel detector has more layers and hence more pixels overall. Instead of 64 million pixel sensors in the old detector, the new detector has 124 million pixels. In the barrel, there are now four layers compared to three layers before. In the endcap, on both sides, there are now three layers instead of two in the old detector.
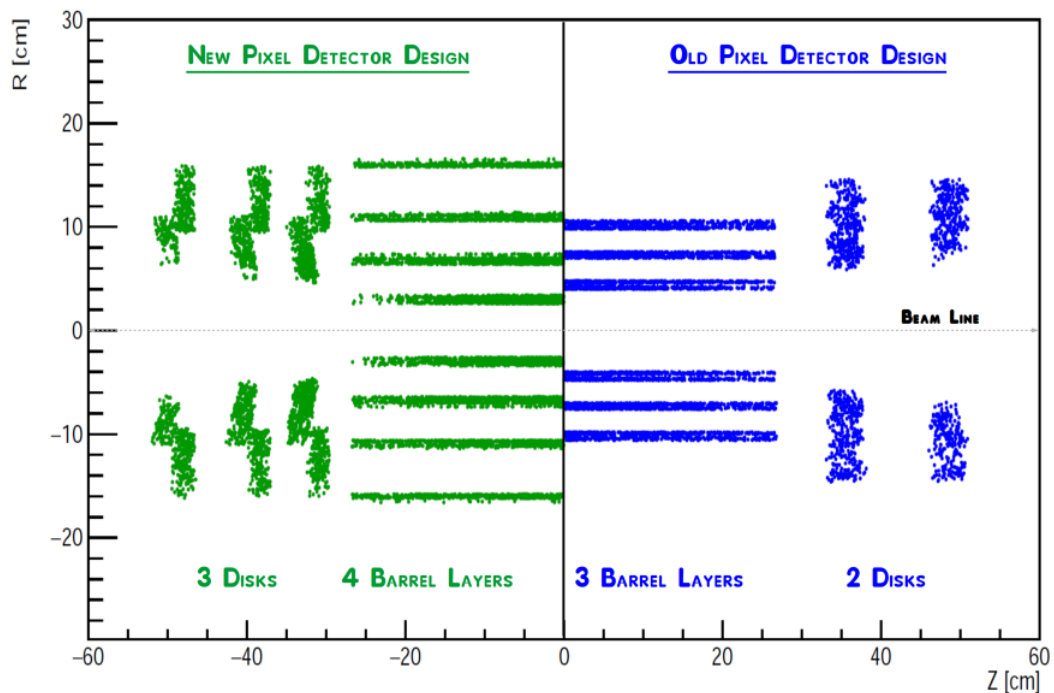
Figure 3.3: Result of simulating the passage of charged particles through the detector. Each charged particle hits multiple layers, and with multiple particles hitting almost all of the detector, an "image" of the pixel detector can be formed. On the right (in blue) are the results for the old pixel detector with 3 barrel layers and 2 disk layers. To the left (in green) are the results with the new pixel detector with a total of 4 barrel layers and 3 disk layers.

3.2.4 Electromagnetic calorimeter

Many studies of physics process are highly dependent on the accurate measurement of the energy of electrons and photons. For this purpose, the CMS uses what is known as the electromagnetic calorimeter (ECAL) [31]. The ECAL consists of a total of 76k scintillating lead tungstate ($PbWO_4$) crystals located in a cylindrical barrel section and the two endcap disks. The primary process through which electrons and photons deposit energy in the ECAL are pair production and bremsstrahlung. An important characteristic of the particles detected by the ECAL is the shape of the developed shower. The shower shape is described by the Moliere radius, the average radius of a cylinder containing approximately 90% of the energy of the electromagnetic shower, which shows the width of the shower, and a radiation length (X_0), is a density-weighted depth of material where a traversing electron reduces its initial energy E_0 to $\frac{1}{e}E_0$ due to radiation, which points to the depth of the electromagnetic shower. The $PbWO_4$ crystals act as scintillators and absorbers simultaneously and have been chosen due to their short radiation length (0.89 cm) and high density of $8.3 \, \mathrm{g/cm^3}$. Multiple avalanche photodiodes (in the barrel) and vacuum phototriodes (in the endcaps) read out the light emitted by the $PbWO_4$ crystals.

3.2.5 Hadronic Calorimeter

For measuring the energy of hadrons (particles made of quarks and gluons), the CMS detector used another kind of calorimeter, known as the HCAL [32]. The particles that reach the HCAL calorimeter are mostly heavier than the ones detected in the ECAL. Thus, a lot more material is needed to contain the shower. For this reason, the HCAL is essentially designed a sampling calorimeter having alternating layers of brass or steel absorbers and tiles of plastic scintillators. The primary process via which particles deposit energy in HCAL is scintillation. Since the charged hadrons start to deposit the part of their energy already in the ECAL, the relative energy resolution of the CMS HCAL is worse than the ECAL. The dead absorber material also contributes to worsening the resolution. The HCAL is hermetic in construction and covers the pseudorapidity range up to $|\eta| < 5.2$. This comes in very handy for the missing transverse energy determination.

3.2.6 Muon System

Since muons have a mass two-hundred times more than that of electrons and because they do not participate in strong interaction, they can pass the tracker system and calorimeters mostly without much energy loss. The primary process through which muons deposit energy in the muon system in ionization. Thus, the muon system [33] constitutes the last sub-detector layers. Three different gaseous detectors are deployed to identify muon candidates and measure momentum. In the barrel region ($\eta < 1.2$), drift tubes (DT) are used along with resistive plate chambers (RPC). In the endcap region (extending till $\eta < 2.4$), cathode strip chambers (CSC) are used along with RPCs. While RPCs can provide can quick information for the Level-1 trigger, the DT and the CSC detectors have a better resolution, and are used to measure the precise position of the muons.

3.2.7 Trigger and Data acquisition

The CMS trigger system has been designed to select only the most interesting collision events. The LHC operates at a bunch-crossing rate of about 40 MHz. Storing information at this rate exceeds the threshold of current technological capabilities. Further, the cross section of standard model processes of interest or theorized new physics interactions results in a rate that is many orders of magnitude below this rate. For this purpose, different trigger streams are designed to select events of interest. For the search described in this thesis, the single muon and single electron trigger streams have been used on the data collected by the CMS detector during part of 2016 and part of 2017.

Any effects due to discrepancies in the trigger efficiency between simulation and data have to be accounted for. This is done by applying scale factor (SF) to the yields from simulation samples, which is defined as a the ratio of the efficiency between data and simulation. These scale factors are measured to be around 1 to 2%.

CHAPTER 4

Event simulation and reconstruction

4.1 Monte Carlo generation

Possibilities of phenomena that can happen at the LHC are given by probabilities dictated by quantum mechanics. To properly simulate these possibilities, event-by-event, Monte Carlo (MC) generators are used. Every simulation starts by first simulating the interaction of two partons from within the incident protons. This interaction is called the "hard scattering". After which the parton shower is simulated during which the incoming and outgoing partons may emit gluons. All the quarks and gluons in the final state will hadronize and form color neutral hadrons. These color neutral hadrons manifest themselves as jets of particles. Lastly, the unstable particles that remain will decay to stable particles eventually, which can be detected by the detector. Each of the final state particles can have various possibilities of momenta, mass, flavor, average life, spin, and other such properties. Various MC techniques are applied to simulate all the possibilities as per their respective probabilities. For this search, MC samples are used to estimate the contribution of all prompt and conversion background processes.

4.2 Full Simulation

The software used for simulation of interaction and the reconstruction of the final state particles in the detector using the best of our knowledge is called the Full Simulation (FullSim). Every detail is simulated using the most intricate knowledge of the process of interest. In CMS, the GEANT4 [34] package is used to simulate particle interaction with the detector. Events are reconstructed using dedicated software.

4.3   Fast Simulation

Fast Simulation (FastSim) [35] software is a faster alternative (about 5 seconds per event) to FullSim (about 100 seconds per event), which economizes the required computing time by loosing only little in performance. This economization makes things like the energy/momentum resolution, reconstruction efficiency, particle identification prone to more discrepancies when compared to standard simulation. However, this is acceptable because the idea is to have a simulation that can model physics processes reasonably well but is faster than FullSim.

For tasks such as systematic studies involving the study of multiple variations of a parameter, each of which might require independent simulation, FastSim is very suitable. It is also widely used in searches for new physics where various new physics hypothesis are tested simultaneously.

4.4   Improvements in FastSim tracking

The simulation of a full collision and further detector interaction involves the proper modeling of various parts of the event chain such as proton-proton interaction physics, detector interaction, and final event reconstruction. The work involved the improvement of the charged particle reconstruction (tracking) module of FastSim so that discrepancies between FastSim and FullSim tracking can be taken care of and FastSim can reproduce the accuracy of the FullSim much better. Various changes were made in FastSim tracking to improve the consistency between FullSim and FastSim by identifying and correcting sources of the discrepancy. All the changes are now part of the FastSim software and are now being used by the CMS collaboration. In the following sections, I have briefly described my contributions to Fastsim tracking.

4.4.1   Removing Sim Dependence

This project aimed to reduce the direct dependence of track reconstruction on the real trajectory of simulated particles (SimTracks). The major dependence was the use of various state properties (p_x, p_y, p_z and x, y, z) of the SimTrack as the initial parameters

of the track. We changed this dependence from SimTrack to the state properties of the track defined at the first hit of a seed, where the seed (a combination of 2 or 3 hits) is the first object reconstructed before full track building is attempted. As of now, as an initial estimate for track building, FastSim uses the initial state properties of the seed (defined at first hit). Since a seed is also a reconstructed object, the Sim dependence was removed. The gain is that FastSim now is synchronized with standard track reconstruction, which is employed for both FullSim and data. This ensures that any effect observed in FullSim and data, because of using information from reconstructed objects, is replicated in FastSim as well.

4.4.2   Realistic seed creation

The seed (for tracking) being used in FastSim used a custom implementation which was very different from FullSim seed creator implementation. I wrote an implementation in FastSim, which inherits from the FullSim seed creator and is synchronized with it. The logical structure of the seed creator can be the responsibility of the tracking group, and in that way, if in the future any change is made in the standard (FullSim) seed creator, it will automatically be applied for FastSim also.

Moreover, for a reconstructed seed to be accepted as a good seed there are a few quality cuts on the d_{xy}, d_z, p_T and minimum hits, which were put in FastSim by hand and the inspiration of those were not clear. Such quality cuts, referred to as global regions, were different in FullSim when compared to FastSim. Since the aim was for FastSim to use the same global regions, the exact global regions were imported in such a way that they are synchronized with the FullSim regions.

The changes described above had a significant effect on FastSim. One of the important results was that the mean hits vs. eta distribution could be reproduced better when compared to older FastSim implementation. See Fig. 4.1 .

4.4.3   Hit pair/triplet checks

Before attempting track reconstructions on the hits left behind by the charged particles, the tracking algorithm checks if, in a combination of hits, the first couple of hits are

24

(a) Mean number of hits vs η before fix

(b) Mean number of hits vs η after fix
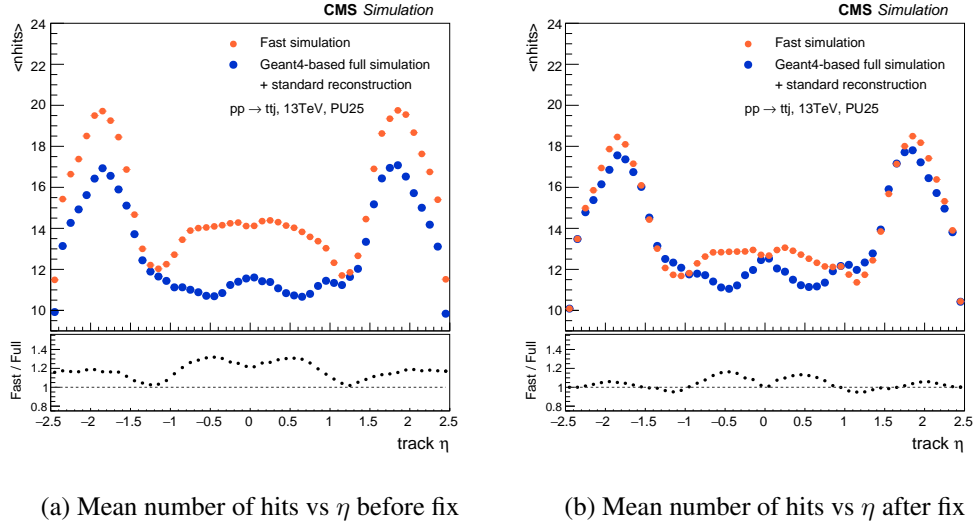
Figure 4.1: Mean hits vs η before and after making the changes described in 4.4.2.

compatible with forming a seed. In FullSim (standard) seeding, a pair/triplet of hits becomes a seed if it passes various criteria. FastSim was missing two important criteria, which were quality checks for two hits (pair) seeds and three hits (triplet) seeds. After inspiration from FullSim, such criteria were also implemented in FastSim. After this change, the efficiency for tracks with low transverse momentum p_T started reproducing FastSim better. See Fig. 4.2

4.5 Detector simulation

There are three components to detector simulation:

- particle interaction with the detector material

- signals from the CMS readout system

- L1 trigger system simulation

Particles that arise from a pp collision will traverse several sub-detectors and also interact with the detector material. The GEANT4 [34] package is used to simulate particle interaction with the detector. This package includes the detector geometry and physics models for particle interactions. The full description of electromagnetic and

(a) Efficiency vs p_T before fix

(b) Efficiency vs p_T after fix
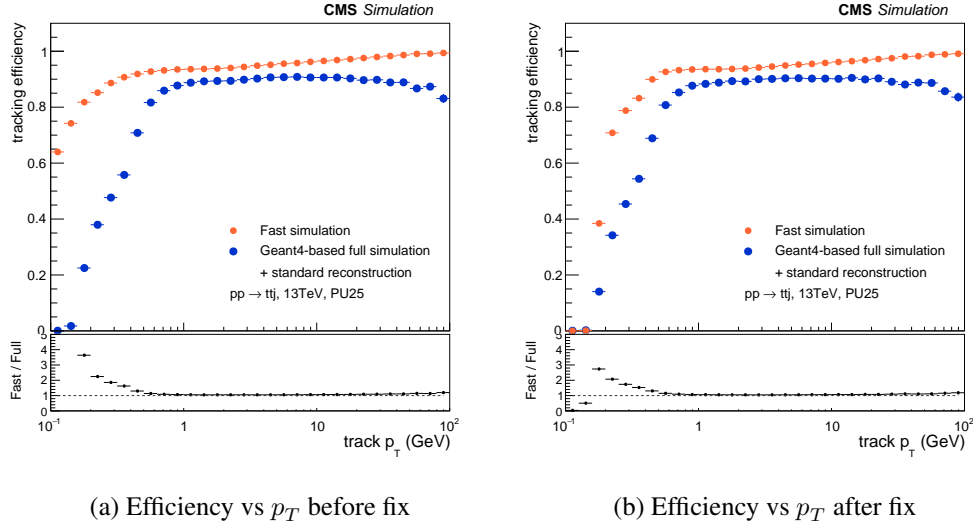
Figure 4.2: Impact on efficiency vs p_T after changes described in 4.4.3.

nuclear interactions in the presence of a magnetic field is contained in the physics models. It can describe interactions in the energy range of a few hundred eVs to a few TeVs. After this, a digitization process is carried out to simulate the response of the electronic readout system.

4.6 Event reconstruction

Particles that emerge from the collision point will first deposit charge in the active material of the inner tracker and then deposit energy in the calorimeter systems. Lastly, particles will encounter the muon detectors. The signatures in the sub-detector vary for each object. A schematic of this is shown in Fig. 4.3.

Every sub-detector, with each containing millions of channels, has to be fully read out first for every proton-proton collision event. The aim is then to backtrace all the digital signal information to the physics of the event. For this purpose, the CMS particle-flow (PF) [37] algorithm is employed. The PF algorithm first processes the signals coming from all the sub-detectors and aggregate them into abstract objects, called physics objects by combining information from all subdetectors to reconstruct and identify individual particles. The output of the PF algorithm will be lots of charged hadrons, neutral hadrons, photons, electrons, and muons in the event. These are referred to as PF objects.
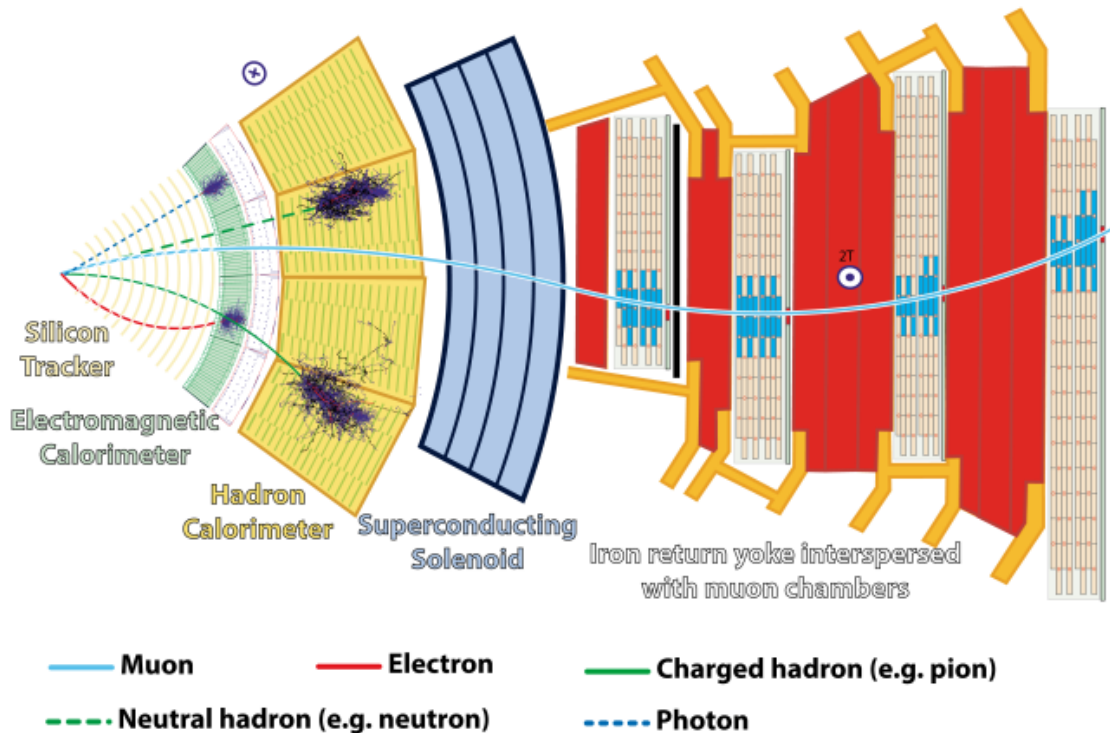
Figure 4.3: A slice showing CMS sub-detectors and how particles interact with them. The figure is from Ref [36]

Electrons will first leave hits in the tracker and then deposit most of their energy in the electromagnetic calorimeter. Muons are minimum ionizing particles (due to their mass) and hence do not deposit much energy in the calorimeter. They are recognized by their signatures in the tracker and the muon system. Hadronically decaying tau leptons and jets can be reconstructed by clustering energy deposits in the calorimeters. Below all steps leading to particle reconstruction are described.

4.6.1 Tracking

Charged particle reconstruction (tracking) is one of the most crucial parts of event reconstruction. Charged particles traversing through the tracker leave several signals in the detector modules. The reconstruction of their trajectories starts from first forming pairs or triplets or quadruplets of hits, using the pixel detector, with the assumption that the track originated from a known vertex. A perfect helix is then fit through reconstructed hit points of the pair/triplet/quadruplet to estimate initial parameters. This process is called seeding. Based on the seed parameters, a Kalman filter [38] proceeds

iteratively through all layers of the tracker system. Each iteration updates the track parameters with each new measurement and finally yields a the trajectory of the track. By using the equation of motion of charged particles in a constant magnetic field, the layers compatible with the initial trajectory are determined.

4.6.2 Vertex reconstruction

The collision of billions of protons leads to several interaction points. It is hence important to reconstruct the correct interaction vertex. The position and the uncertainty of a vertex are computed from a given set of reconstructed tracks. The jets, clustered using the jet finding algorithm [39, 40] with the tracks assigned to the vertex as inputs, are referred to the physics objects. The reconstructed vertex with the largest summed physics-object p_T^2 is taken to be the primary pp interaction vertex. Tracks emerging from the primary collision point are selected by their transverse impact parameter significance with respect to the beamline, the number of hits in silicon pixel and strip detectors and the normalized track χ^2.

4.6.3 Electrons

Both the electrons and the photons may loose energy because of bremsstrahlung and pair conversion in the form of an electromagnetic (EM) shower, because of interaction with the ECAL material. The sum of all energy deposits, including the characteristic EM shower, must be taken into account in order to determine the initial energies of electrons or photons. Due to initial interaction with the tracker, $\approx 50\%$ of the photons will convert into electron-positron pairs, also, $\approx 35\%$ of the electrons will have radiated $\approx 70\%$ of their initial energy before even reaching the ECAL. The particles will thus a shower that is spread in ϕ due to the Lorentz drift that is introduced by the magnetic field. For the barrel and endcap region, different algorithms are employed to form superclusters out of the ECAL energy deposits. The electron candidates are reconstructed using the ECAL superclusters and Gaussian sum filter (GSF) [41] tracks from the silicon tracker [42].

In this thesis, GSF electrons with $p_T > 20$ GeV and $|\eta| < 2.5$ are used. Two (loose

Electron variable	Description
full5x5_sigmaIetaIeta	Average shower width independent of direction
abs(dEtaSeed)	η gap between seed and super-cluster
abs(dPhiIn)	ϕ gap between super-cluster and track at vertex
H/E	Ratio of energy deposited in HCAL to ECAL
relIsoWithEA	Relative isolaton of the electron with effective area correction
abs(1/E-1/p)	Here E and p are the energy and momentum of the electron
expected missing inner hits	Hits expected to be missed after track reconstruction
pass conversion veto	If the electron should pass the photon conversion veto
ESC	Energy of super-cluster

Table 4.1: Variables used to select quality electrons

	Loose	Medium
full5x5_sigmaIetaIeta $<$	0.0112	0.0106
abs(dEtaSeed) $<$	0.00377	0.0032
abs(dPhiIn) $<$	0.0884	0.0547
H/E $<$	0.05+1.16/ESC+0.0324/ESC	0.046+1.16/ESC+0.0324/ESC
relIsoWithEA $<$	0.112+0.506/pT	0.0478+0.506/pT
abs(1/E-1/p) $<$	0.193	0.184
expected missing inner hits $\leq$	1	1
pass conversion veto	Yes	Yes

Table 4.2: Quality criteria for electrons with $|\eta| < 1.479$. The loose and medium selections are used to select two kinds of electrons with slightly different criteria. The exact description of each of these variables is given in Table 4.1.

and medium) additional criteriaare applied based on whether the electron is in the barrel or end-cap. The variables used to select quality electrons are described in Table 4.1 and the exact criteria on these variables are given in Tables 4.2 and 4.3.

4.6.4   Muons

The algorithm used for muon reconstruction combines information of the inner tracker and the muon system There are two ways to do that:

- outside in: muons are reconstructed using only the information in the muon system and then matched with track candidates, also referred to as standalone muons

- inside out: one starts from track candidates and then tries to match the candidates to hits in the muon system

Reconstruction of the standalone muon candidate begins with segments of local track based on measured drift times in the different muon chambers. Initial parameters

	Loose	Medium
full5x5_sigmaIetaIeta <	0.0425	0.0387
abs(dEtaSeed) <	0.00674	0.00632
abs(dPhiIn) <	0.169	0.0394
H/E <	0.0441+2.54/ESC+0.183*/ESC	0.0275+2.52/ESC+0.183*/ESC
relIsoWithEA <	0.108+0.963/pT	0.0658+0.963/pT
abs(1/E-1/p) <	0.111	0.0721
expected missing inner hits ≤	1	1
pass conversion veto	Yes	Yes

Table 4.3: Quality criteria for electrons with $|\eta| > 1.479$. The loose and medium selections are used to select two kinds of electrons with slightly different criteria. The exact description of each of these variables is given in Table 4.1.

are determined using segments in the first muon station. Similar to track reconstruction in the tracker, a Kalman filter based method is used to add the segments of the outer stations successively, till the last segment. The final parameters are calculated at the innermost muon segment by a backward refitting procedure. The energy losses and scattering in the material between the stations is also taken into account.

For the outside-in algorithm standalone muon candidate collection is matched to track candidates by propagating the muon trajectory to the last layer of the tracker. The energy loss and multiple scattering in the detector material, and the dense material of the calorimeter and the solenoid are also accounted for. Finally, using a the track and the standalone muon candidates a global fit is performed.

The inside-in algorithm uses all the track candidates which are reconstructed in the tracker system. The algorithm propagates their trajectories analogously from the outermost tracker layer to the muon system. For a track candidate to be classified as a tracker muon, at least one segment in the muon system has to match the propagated trajectory.

Using the PF muon identification, muon energy deposits in the calorimeters are associated with the muon track, and this information is used to improve the muon identification performance.

Most of the muons, with sufficient momentum, emerging from the collision point are reconstructed both as the global and tracker muons.

PF muons with $p_T > 20$ GeV and $|\eta| < 2.4$, with additional requirements on them

	Loose	Medium
Muon is both tracker and PF muon	Yes (or tracker muon)	Yes
Muon is PF muon	Yes	Yes
Normalized $\chi 2$ of the track associated to the global muon	-	< 3
Tracks matching the standalone muon	-	< 12
Value from kink-finder algorithm	-	< 20
Segment compatibility	-	< 0.303

Table 4.4: Quality criteria for muons.

being wither global or tracker muons, are used in this thesis. Similar to electrons, two (loose and medium), additional criteria are applied. The complete list of criteria is shown in Table 4.4.

4.6.5   Tau reconstruction

Tau leptons can decay into leptons or hadrons with associated neutrinos. The leptonic decay modes can not directly be associated with a τ-decay and are indistinguishable from stable electrons or muons. Hadronically decaying tau leptons have a branching fraction into hadronic single-prong (with one charged π's) decays of about 50% and into three-prong (with three charged π's) of about 15%. The single prong tau leptons can even be further divided into single-prong with and without associated π^0's. To identify hadronic tau leptons two kinds of algorithm are used: the *hadron plus strips* (HPS) [43] and the tau-neural-classifier (TaNC) [44] algorithm. Both algorithms use all the particle flow objects (charged hadrons, photons, neutral hadrons, muons, and electrons) as the basis for reconstructing three different hadronic tau lepton topologies separately. The tau lepton reconstruction starts with the PF jet reconstruction using the $\mathrm{anti} - \mathrm{k_T}$ clustering.

In this thesis, hadronically decaying tau lepton (τ_h) candidates are selected using the HPS algorithm. The algorithm gives particular attention to photon conversion by looking at the bending of electrons and positrons arising from converted photons that emerge from neutral pions, which broadens the calorimeter signatures in the azimuth direction. Electromagnetic particles are iteratively collected in an $\eta \times \phi$ strip ($\Delta\eta \times \Delta\phi = 0.05 \times 0.20$), and the four-momentum vector is recalculated after each new

particle is added, until no further particle can be associated. By combining all the strips with $p_T > 1$ GeV and charged hadrons inside a cone $\Delta R = 2.8$ GeV/$p_T^{\tau_h}$ it is possible to reconstruct the individual τ_h decay modes. Moreover, the four-vector sum of strips and charged hadrons have to be consistent with intermediate resonances (π^0, ρ). The τ_h candidates are required to have $p_T > 20$ GeV and $|\eta| < 2.3$. To remove misidentified τ_h candidates originating from electrons or muons, all candidates are also required to pass some additional discriminators.

4.6.6 Jets

Due to QCD confinement, quarks and gluons produced in the hard scattering process will undergo fragmentation and hadronization processes into collimated streams of color-neutral hadrons. A jet is then defined as a cluster of hadrons inside a pre-defined cone. The clustering algorithm should be stable in two conditions. Firstly, when facing additional soft partons between two jets (infrared safety). Secondly, when facing additional collinear gluons radiated by the hadronic jets (collinear safety).

Common jet clustering algorithms define two distances: d_{ij} between the particles i and j and d_{iB} between particle "i" and the beamline. Objects with a minimum distance d_{ij} get iteratively combined until only particles with $d_{ij} > d_{iB}$ remain. The resulting object is called a jet. Those jets are then removed from the list of particles. This procedure is redone until no particles are left.

In this thesis, anti-k_T algorithm [39, 40] is used with a radius parameter of 0.4. The reconstructed charged and neutral hadron, muon, electron, and photon PF candidates are input to the anti-k_t jet clustering algorithm. The resulting CHS AK4 PF jets are required to have $p_T > 30$ GeV, $|\eta| < 3$. For 2016, jets are required to satisfy the loose working point ID, while for 2017, a tight working point is used since for 2017 the tight working point has $> 99\%$ efficiency. The difference in IDs for different years is since for 2016, the tight working point had lower efficiency at $|\eta| > 2.7$. The exact selection used in the loose and tight working point is given in Table 4.5. The p_T criteria is not applied when using jets for tau lepton fake rate measurements except for a 10 GeV threshold.

	PF Jet ID	Loose	Tight
For $-2.7 <= \eta <= 2.7$	Neutral Hadron Fraction	< 0.99	< 0.90
	Neutral EM Fraction	< 0.99	< 0.90
	Number of Constituents	< 1	< 1
	Muon Fraction	-	-
Additionally for $-2.4 <= \eta <= 2.4$	Charged Hadron Fraction	< 0	< 0
	Charged Multiplicity	< 0	< 0
	Charged EM Fraction	< 0.99	< 0.99
For $2.7 < |\eta| <= 3.0$	Neutral EM Fraction	> 0.01	> 0.01
	Neutral Hadron Fraction	< 0.98	< 0.98
	Number of Neutral Particles	> 2	> 2
For $|\eta| > 3.0$	Neutral EM Fraction	< 0.90	< 0.90
	Number of Neutral Particles	> 10	> 10

Table 4.5: Quality criteria for PF Jets in different η regions for loose and tight working points.

Spurious extra tracks, calorimetric energy depositions to the jet momentum and interactions in the same or nearby bunch crossings (pileup) can contribute to the momentum of the jet. Thus, to account for the remaining neutral pileup particle contributions, the charged particles as identified to originate from pileup vertices are discarded, and an offset correction [45] is applied. Additional jet energy corrections are also applied to account for any nonlinear response of the detectors [46].

4.6.7   Missing transverse momentum

The CMS detector can detect most of the particles except neutrinos, which are very weakly interacting in nature. The way to infer the presence of neutrinos is by imposing conservation of transverse momenta. Since the initial particles exhibit no transverse momentum, the vectorial sum of all particles should vanish, and hence an indirect indication for the presence of neutrinos is the so-called missing transverse energy (p_T^{miss}). It is defined as the magnitude of the negative vectorial sum of all reconstructed stable particles

$$|\vec{p_T^{miss}}| = -\sum_{i=0}^{N} \vec{p_T^i} \qquad (4.1)$$

with $i = i^{th}$ particle.

33

Due to detector conditions during the 2017 data collection, the standard definition as described above was resulting in a poor data-MC agreement in the regions where processes with no real p_T^{miss} dominate. For example, in DY enriched regions (where the mass of the OS dilepton pair is within 15 GeV of the Z mass, i.e., on-Z), a large disagreement was being observed. The primary cause of this disagreement was noise amplification due to the aging of the ECAL. This primarily affected jets with $p_T < 50$ GeV in $2.5 < |\eta| < 3.0$. Hence, while for calculating p_T^{miss} in 2016, the full fiducial volume of the detector is used, for 2017 jets and unclustered PF candidates with $2.5 < |\eta| < 3.0$ and $p_T < 50$ GeV are not considered.

CHAPTER 5

Strategy and analysis

5.1 Event selection

Since the CMS detector collected data independently in 2016 and 2017, the trigger conditions were slightly different since the instantaneous luminosity changed and the detector was upgraded as well. For events collected in 2016, the electron trigger requires at least one electron with $p_T > 27$ GeV, and the muon trigger requires at least one muon with $p_T > 24$ GeV. Similarly, for events collected in 2017, the electron trigger requires at least one electron with $p_T > 35$ GeV and the muon trigger requires at least one muon with $p_T > 27$ GeV. These selections are placed on objects reconstructed at the trigger level. The analysis is done independently for data collected in 2016 and 2017.

As described in Section 2.8, the VLLs couple to the third generation SM leptons with multiple leptons being present in final state after production of the VLL pair. The search is hence carried out in final states with multiple leptons (e, μ, τ). Due to coupling to the third generation SM leptons, there is always a τ in the final state. This τ may decay leptonically or hadronically. When decaying leptonically, the τ is indistinguishable from electrons or muons. Events are thus primarily categorized as those with two light leptons and at least one hadronically decaying tau lepton (2L1T), three light leptons (3L), and four or more light leptons (4L). Depending on whether the two light leptons are of opposite-sign (OS) or same-sign (SS), we have further divisions in the 2L1T channel. Final states with multiple τs are also possible but these have not been tackled in this thesis since they would require a complete overhaul of the background estimation techniques. In all categories, the leading light lepton is required to satisfy a p_T threshold of 28 GeV if it is a muon, and 38 GeV if it is an electron, for the corresponding single lepton triggers to be efficient. All the other trailing leptons are required to satisfy a p_T threshold of 20 GeV.

Lepton candidates arising from pp collisions can be broadly categorized into prompt, nonprompt, and conversion leptons. Leptons that directly come from W, Z and Higgs boson decays are considered as prompt leptons, whereas those originating from semi-leptonic heavy quark decays within jets or other misidentified detector signatures are labeled as misidentified (non-prompt) leptons. A lepton produced when a radiated photon converts to a pair of leptons, is referred to as a conversion lepton.

The misidentified leptons are expected to have significant nearby hadronic activity, which is unlike prompt leptons. An isolation requirement, comparing the p_T of a lepton to the p_T sum of the particles in its neighborhood, is applied to reduce the backgrounds from misidentified leptons. Relative isolation, determined as the scalar p_T sum of photons and charged and neutral hadrons, within a specified ΔR cone around the lepton candidate, normalized to the lepton candidate p_T, is also applied. The ΔR between a particle and the lepton is defined as $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}$, where $\Delta \eta$ is the difference in pseudorapidity, and $\Delta \phi$ is difference in the azimuthal angle (in radians). The relative isolation is required to be less than 7 or 8% within a cone of size $\Delta R = 0.3$ for electrons whose energy deposits are reconstructed in the ECAL barrel ($|\eta| < 1.48$) or in the endcap ($1.48 < |\eta| < 3.00$). For the muons, it is required to be less than 15% within a cone of size $\Delta R = 0.4$ for muons. The τ_h candidates are required to pass an isolation requirement based on a multivariate analysis [47]. All the isolation quantities are corrected for pileup by considering only those charged PF candidates that originated from the primary vertex. The correction is achieved by subtracting a per-event average pileup contribution to the neutral PF components. The misidentified backgrounds (MisID) backgrounds are further reduced by imposing requirements on the longitudinal (d_z), and transverse (d_{xy}) impact parameters of the leptons with respect to the primary vertex in the event. Electrons in the barrel (endcap) must satisfy $|d_z| < 0.1$ (0.2) cm and $|d_{xy}| < 0.05$ (0.1) cm. Muons must satisfy $|d_z| < 0.1$ cm and $|d_{xy}| < 0.05$ cm. For τ_h leptons, we require $|d_z| < 0.2$ cm. The reconstructed leptons are required to lie within the region of pseudorapidity $|\eta| < 2.5$, 2.4, and 2.3 for the electron, muon, and τ_h candidates, respectively.

Additionally, isolation and track efficiency scale factors corresponding to the lepton

selections used in this analysis are applied on a per-lepton basis as a function of the lepton flavor, p_T, and η for the background processes estimated via simulated samples (WZ/ZZ/rare) and the signal processes. The scale factors are measured in events having leptons consistent of coming from a Z boson using tag-and-probe method, by accounting for any differences between data and simulation efficiencies and calibrations. The ID and isolation scale factors are in the range of 0.98 to 1.00 for muons, and 0.90 to 1.02 (0.80 to 1.04) for electrons reconstructed in ECAL barrel (endcap). The tracking efficiency scale factors are 0.99-1.00 for muons, whereas they vary between 0.98 and 1.00 (0.90 and 1.31) for ECAL barrel (endcap) electrons. Tau ID scale factors are 0.93 (0.86) for 2016 and 2017.

N_{leptons}	p_T^{miss} (GeV)	CR veto
$\geq 4e/\mu$	<50 >50	2 OSSF on-Z pairs and $p_T^{miss} < 50$ GeV
$3e/\mu$	<150 >150	OSSF on-Z pair and $p_T^{miss} < 100$ GeV, or OSSF below-Z pair and $p_T^{miss} < 50$ GeV, or OSSF below-Z pair and on-Z $m_{3\ell}$
$2e/\mu$ OS (or SS) + $\geq 1\tau_h$	<150 >150	$p_T^{miss} < 50$ GeV

Table 5.1: The signal regions defined in this analysis are shown above. The categorization is done depending on N_{leptons} and p_T^{miss}. In each region, events from control regions (CRs) are vetoed and hence CR veto is also shown above. The on-Z mass window is defined as $76 < m_{\ell\ell} < 106$ GeV, while the below-Z condition is defined as $m_{\ell\ell} < 76$ GeV. Here OSSF refers to an opposite-sign, same-flavor lepton pair

We use the scalar p_T sum of the leptons (denoted as L_T) to discriminate signal from SM backgrounds in all channels. The L_T distribution is divided into 150 GeV bins, each of which is treated as a separate experiment. In the 2L1T and 4L categories that contain more than one τ_h and more than four light-lepton candidates, respectively, only the leading τ_h and the leading four light-leptons are used in the calculation of L_T. In 2L1T and 4L categories, in the presence of more than one tau or more than four light lepton candidates, only the leading tau lepton and the leading four light-leptons are selected and used in the calculation of L_T.

To achieve optimum sensitivity for the signal model, in each of the 2L1T OS/SS,

3L, and 4L categories, the events are divided into low- and high-p_T^{miss} regions. While the 4L category is divided into $p_T^{miss} < 50$ GeV and >50 GeV regions, the 3L and 2L1T (OS, SS) categories are divided into $p_T^{miss} < 150 GeV$ and >150 GeV regions. These categories form the bases of signal regions (SR) that would be sensitive to the presence of a VLL signal. They are complemented by orthogonal control regions (CR) that are expected to be dominantly populated by backgrounds. Additionally, all events with a light-lepton pair invariant mass below 12GeVare vetoed regardless of the flavor and sign of the pair, in order to suppress low mass quarkonia resonances. The SRs are described in Table 5.1, where OSSF refers to an opposite-sign, same-flavor lepton pair. A detailed description of the CRs is given in Sec 5.2.

5.2 Backgrounds

The backgrounds of this search can be divided into three categories. Events from processes such as WZ and ZZ contain multiple prompt leptons, and thus, these backgrounds are classified as prompt backgrounds. The background arising from processes like Z+jets, $t\bar{t}$ and W+jets result in a multilepton final state, only if one or more non-prompt leptons are present. Such sources are referred to as the MisID background. The background arising from processes where conversion can result in additional leptons in the final state is referred to as the conversion background. Simulation samples have been used to estimate the prompt and the conversion backgrounds, while a data-driven method has been developed for estimation of MisID backgrounds.

SM processes yielding 3 or more prompt leptons such as WZ and ZZ, or those with two prompt and additional non-prompt lepton such as Z+jets and ttbar+jets constitute the primary backgrounds for this analysis. Various other processes such as triboson, $t\bar{t}V$, and top or vector boson associated Higgs production can also yield multilepton signatures. These contributions are generally suppressed due to lower production cross sections.

The next-to-leading order (NLO) generation using POWHEGv2 [48–52] is employed for generating WZ and ZZ processes. The Z/γ^*, $Z/\gamma^* + \gamma$, $t\bar{t}$, $t\bar{t} + \gamma$, and triboson processes are generated at NLO using MADGRAPH5_aMC@NLO v5.2.2 [53] and processes with the Higgs boson are generated using POWHEGv2 [54, 55] and the JHUGEN v6.2.8 generator [56–59]. Signal events are generated using MADGRAPH5_aMC@NLO at leading order (LO) precision. PYTHIA8.230 [60] (with tune CUETP8M1 [61] for 2016 samples, and CP5 [62] for 2017 samples) is used for all simulation, the parton showering, fragmentation, and hadronization steps.

All 2016 samples are generated with the same order of the NNPDF3.0 parton distribution function (PDF) [63] as the order of the MC generator. All 2017 samples are generated with the NNPDF3.1 next-to-next-to-leading (NNLO) order PDF [64], irrespective of the order of the MC generator. The GEANT4 toolkit [34] is used to simulate the response of the CMS detector. To account for any differences in the trigger and the lepton identification efficiencies between data and simulation, additional weights

are applied to all background simulation events. Additional minimum bias interactions are superimposed on the primary collision, reweighted in such a way that the frequency distribution of the extra interactions matches that observed in data.

Yields from MC samples for the irreducible diboson backgrounds, WZ$\rightarrow 3\ell\nu$ and ZZ$\rightarrow 4\ell$, are normalized and validated in dedicated control regions in data. The other irreducible rare (triboson, ttV) or Higgs backgrounds are also obtained from MC samples but are normalized with the appropriate theoretical cross sections. For backgrounds that contain misidentified leptons, data-driven methods are utilized.

5.2.1 Dilepton Control Regions

Although the corresponding MC samples play a minor role in the analysis, dileptonic DY+jets and $t\bar{t}$+jets processes have been studied as a cross-check to commission and verify the object selections used in this thesis.

A set of dilepton events enriched in DY($\rightarrow$ ee)+jets and DY($\rightarrow \mu\mu$)+jets processes are created by requiring the mass of opposite sign same flavor dilepton pair to be on-Z and $p_T^{miss} < 50$ GeV. For a selection enriched in $t\bar{t}$+jets, events are used, which fulfill the requirements of having an opposite-sign eμ pair and $S_T > 300$ GeV. Figs. 5.1, 5.2 and 5.3 show the lepton p_T, L_T, and M_{2l} distributions for the DY+jets enriched selections, and the lepton p_T, L_T, and p_T^{miss} distributions for the $t\bar{t}$+jets enriched selection for 2016. Same plots for 2017 are shown in Figs. A.19, A.20 and A.21.

The jet multiplicity is normalized to data to account for higher-order effects that might not be properly estimated by NLO MC. Jet multiplicity distributions are shown in Figs. 5.4 and 5.5 for 2016 and 2017 datarespectively. Additional correction factor as a function of Z p_T (vector sum of lepton p_T) is applied to correct for the incorrect modeling of Z p_T in MC [65]. The correction factors are derived as weights from the dilepton DY region, extracted from the Z p_T distribution in data and MC.

5.2.2 WZ backgrounds

The WZ $\rightarrow 3\ell\nu$ process constitutes one of the major backgrounds in this analysis, and `POWHEG`generated NLO MC samples are used to estimate these contributions for both
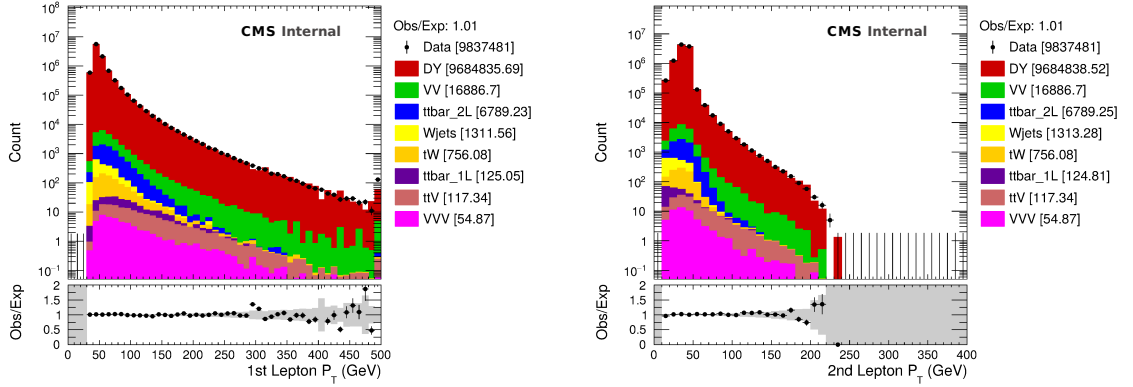
Figure 5.1: Leading (left) and sub-leading (right) electron p_T distributions in the DY→ee enriched dilepton selection in the 2016 data.
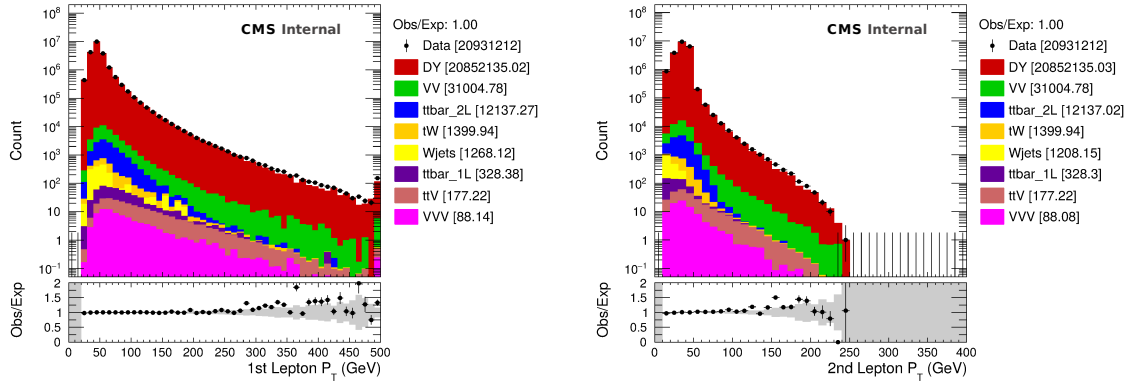
Figure 5.2: Leading (left) and sub-leading (right) muon p_T distributions in the DY→ $\mu\mu$ enriched dilepton selection in the 2016 data.
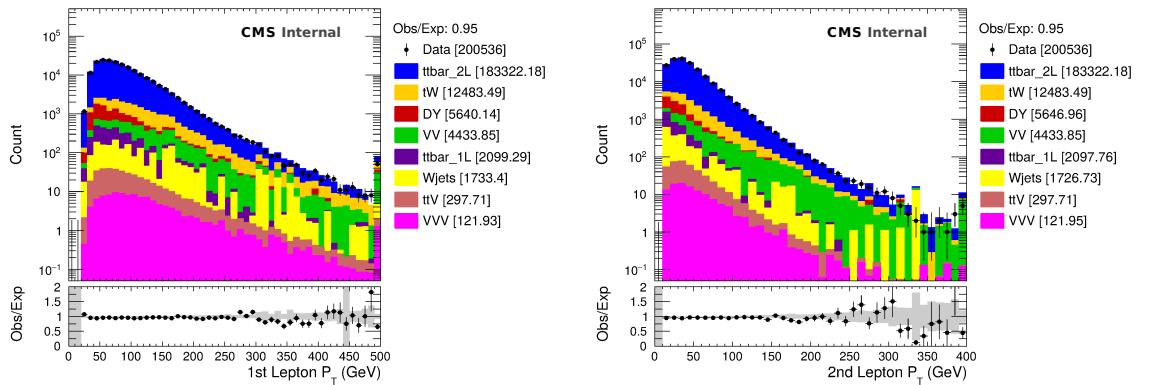
Figure 5.3: Muon p_T (left), electron p_T (right) distributions in the $t\bar{t} \rightarrow e\mu$ enriched dilepton selection in the 2016 data.
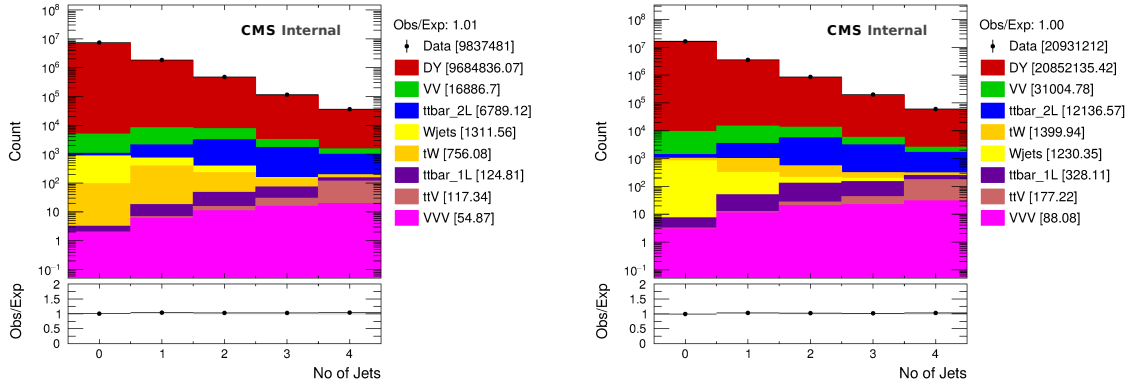
Figure 5.4: Jet Multiplicity distributions for the DY→ee (left) and DY→ $\mu\mu$ (right) enriched dilepton selection in the 2016 data and MC.
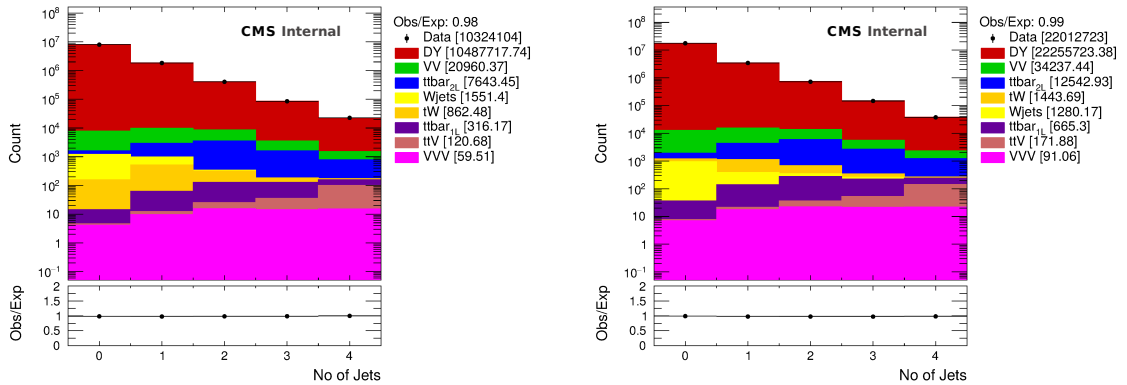
Figure 5.5: Jet Multiplicity distributions for the DY→ee (left) and DY→ $\mu\mu$ (right) enriched dilepton selection in the 2017 data and MC.

2016 and 2017 analysis. A WZ $\to 3\ell\nu$ ($\geq 80\%$) selection of events is created by requiring exactly 3 leptons with an on-Z OSSF pair and 50 GeV $\leq p_T^{miss} \leq$ 100 GeV. The WZ MC normalization scale factor over the NLO cross section is then calculated as the ratio of non-WZ subtracted data events over WZ MC events, yielding a value of 1.14 ± 0.06 (1.07 ± 0.05) for 2016 (2017).This normalization correction factor is then applied to all WZ MC events, leading to a relative normalization uncertainty of 5% (5%) for 2016 (2017). For the WZ enriched selection, the 2016+2017 combination plots for the M_T and L_T distributions are shown on top of Fig. 5.6.

5.2.3 ZZ backgrounds

The ZZ $\to 4\ell$ background dominates $\geq$ four lepton signal regions with at least one on-Z OSSF pair. ZZ contributions are estimated using a Powheg generated NLO MC sample that is normalized to data in a dedicated ZZ-enriched selection of events. This selection is defined by requiring exactly four leptons that form 2 distinct on-Z OSSF pairs and $p_T^{miss} < 50$ GeV, and yields to a set of events 99% pure in ZZ $\to 4\ell$. The ZZ MC normalization scale factor over the NLO cross section is calculated as the ratio of non-ZZ subtracted data events over ZZ MC events, yielding a value of 1.01 ± 0.05 (0.98 ± 0.04) for 2016 (2017), leading to a relative normalization uncertainty of 5% (4%) for 2016 (2017). For the ZZ enriched selection, the 2016+2017 combination plots for the $m_{4\ell}$ and L_T distributions are shown at the bottom of Fig. 5.6.

5.2.4 Conversion backgrounds

In process with two prompt leptons, an internal or external photon conversion might result in 2 additional leptons which might also pass the promptness criteria. If the two conversion leptons were reasonably symmetric in p_T, the standard conversion filter might reject those leptons. This is done by finding the conversion partner of the lepton, traveling parallel to the original electron, within a closest distance of 0.02 cm. In cases where the conversion is asymmetric, one of the leptons might not result in a reconstructed lepton or not pass our quality criteria. In such cases, the final state consists of 3 leptons, and this results in small background component to the three lepton signal
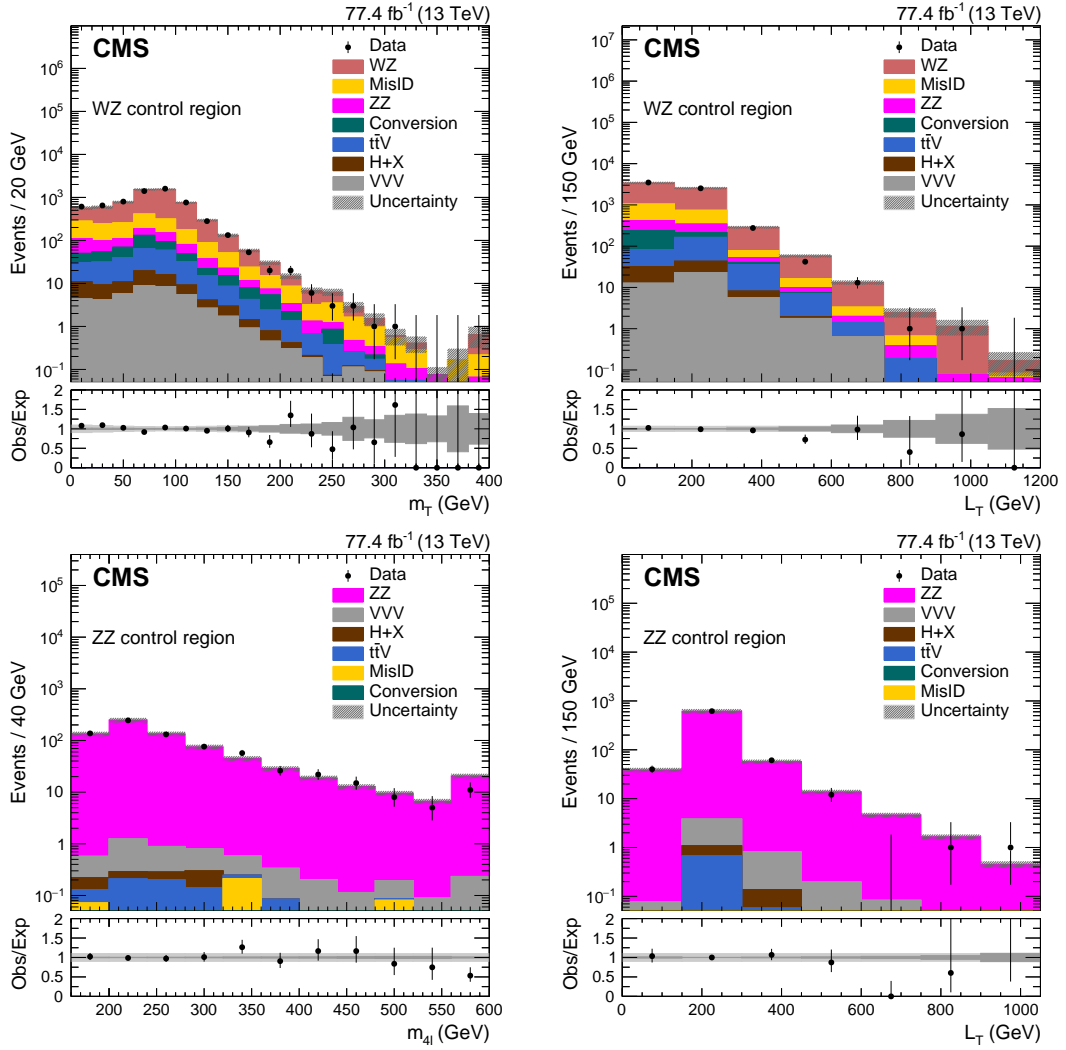
Figure 5.6: The upper row shows the M_T (left) and the L_T (right) distributions in the WZ control region in data and simulation. The WZ control region contains events with three leptons and an OSSF pair with mass on-Z, and $50 < p_T^{miss} < 100$ GeV. The lower row shows the $m_{4\ell}$ (left) and the L_T (right) distributions in the ZZ control region. The ZZ control region contains events with two OSSF lepton pairs, both of which are on-Z, and $p_T^{miss} < 50$ GeV. The total SM background is shown as a stack of all contributing processes. The hatched gray bands in the upper panels represent the total uncertainty in the expected background. The lower panels show the ratios of observed data to the total expected background. In the lower panels, the light gray band represents the combined statistical and systematic uncertainty in the expected background, while the dark gray band represents the statistical uncertainty only. The rightmost bins include the overflow events.

regions. This background is estimated using MC samples. A selection constrained by the invariant mass of the two leading leptons ($M_{2l} \leq 76$ GeV, below-Z), the invariant trilepton mass on-Z and $p_T^{miss} \leq 50$ GeV results in a conversion dominant control

region. The major contribution to this background comes from the Drell-Yan ($Z\gamma^*$) process and taken from a Drell-Yan+jets aMC@NLO generated NLO MC sample. Other processes like $t\bar{t}\gamma$ and WW can also contribute to this background and for their estimation the $t\bar{t}$ and WW simulation samples are used. In case of events with 4 or more leptons, WZ and ZZ can also contribute through conversion, but this is taken care of vetoing events where lepton pairs that come from Z in the signal region.

The overall MC normalization scale factor over the NLO cross section for conversion backgrounds is calculated as the ratio of non-conversion events subtracted data events over conversion events, yielding a value of 0.95 ± 0.11 (0.87 ± 0.10) for 2016 (2017), leading to a relative normalization uncertainty of 11% (11%) for 2016 (2017). This normalization is then applied to all conversions processes in all signal regions.

The L_T distributions for all primary control regions are provided in Fig A.22 and A.23, for 2016 and 2017 independently.

5.2.5 Misidentified Light Lepton Backgrounds

Multilepton background contributions due to misidentified leptons, such as in DY+jets and $t\bar{t}$+jets events, are estimated via a 3-dimensional implementation of a matrix method [66]. Looser selections (loose) leptons are used for the prediction of misidentified lepton background in matrix method.

5.2.5.1 Matrix Method

In its simplest form, matrix method is a data-driven background estimation method that relies on the assumption that the probabilities with which prompt and misidentified leptons pass a tight lepton selection given that they satisfy a loose lepton selection, prompt (p) and fake (f) rates respectively, are universal and can be described as a function of the lepton and event dependent parameters. This assumption allows the measurement of these rates in signal-depleted control regions and then their application to a signal region. In its 3-dimensional form, the matrix method can predict background contributions due to events with up to 3 simultaneous misidentified leptons. Since this analysis relies on the use of isolated single lepton triggers, it is assumed that at least one triggering lepton in 3/4 lepton events is a prompt lepton. In signal regions with 4 or more

leptons, the 3D matrix method is applied on the 2nd, 3rd, and 4th leading p_T leptons, and the leading p_T lepton is assumed to be prompt. A detailed description of the 2D matrix method is provided in Ref. [67], and can be trivially expanded for the 3-lepton case.

5.2.5.2 Determination of Prompt and Fake Rates for Light Leptons

Prompt rates for leptons are measured using a tag-and-probe method in MC and data in dilepton events. For MC measurements a set of events enriched in prompt leptons from $Z \to \ell\ell$ or $t\bar{t}$ decays is created by requiring the reconstructed leptons to be kinematically matched to a generator level prompt lepton ($\Delta R < 0.1$) to be labeled as prompt objects. In data, an on-Z, OSSF pair is required in events with $p_T^{miss} < 50$ GeV. The leading p_T lepton is chosen as the tag and is also required to satisfy the tight lepton selection, and the subleading p_T one is chosen as the probe.

Lepton fake rates are measured inclusively for trilepton events in MC. In data an on-Z, OSSF pair is required with both leptons satisfying the tight lepton selection in events with $p_T^{miss} < 50$ GeV along with a single additional probe lepton satisfying the loose lepton selection. A fake probe lepton enriched selection of events is created in simulated samples by imposing an inverted ΔR requirement among the probe lepton and the prompt generator level leptons in the event ($\Delta R > 0.1$). Additionally, fake probe leptons matching to a generator level photon are also vetoed. The prompt (fake) rate is then defined as the ratio of prompt (fake) probe lepton enriched events as described above where the probe lepton further satisfies the tight lepton selection such that it satisfies the loose lepton selection, such that events in the numerator constitute a subset of those in the denominator. In all rate measurements conducted in data, contributions due to leptons of the undesired origin (prompt leptons for fake rate measurements and vice versa) are estimated and subtracted using MC methods, whereas a simple binomial ratio of tight over loose leptons is taken in MC. These quantities are calculated as given

in Eq. 5.1.

$$p^{\mathrm{MC}} = \frac{\mathrm{N}^{\mathrm{MC}}_{\text{tight prompts}}}{\mathrm{N}^{\mathrm{MC}}_{\text{loose prompts}}}, \qquad p^{\mathrm{data}} = \frac{\mathrm{N}^{\mathrm{data}}_{\text{tight prompts}}}{\mathrm{N}^{\mathrm{data}}_{\text{loose prompts}}},$$

$$f^{\mathrm{MC}} = \frac{\mathrm{N}^{\mathrm{MC}}_{\text{tight fakes}}}{\mathrm{N}^{\mathrm{MC}}_{\text{loose fakes}}}, \qquad f^{\mathrm{data}} = \frac{\mathrm{N}^{\mathrm{data}}_{\text{tight}} - \mathrm{N}^{\mathrm{MC}}_{\text{tight non-fakes}}}{\mathrm{N}^{\mathrm{data}}_{\text{loose}} - \mathrm{N}^{\mathrm{MC}}_{\text{loose non-fakes}}}. \qquad (5.1)$$

The prompt rates and fake rates are parametrized in bins of the lepton p_T, lepton $|\eta|$, whereas fake rates are parametrized in bins of lepton p_T, lepton $|\eta|$ and the p_T of the mother jet of the probe lepton.

Since the fake rate measurement in data is statistically limited and contaminated with prompt background, these fake rates are not used directly but are used to correct the fake rates measured in MC. This procedure is done only for fake rate bins in lepton p_T& lepton $|\eta|$ with a significant non-prompt contribution. A ratio of $k^{\mathrm{f}}_{\mathrm{data}} = f^{\mathrm{data}}/f^{\mathrm{DY\ MC}}$ is calculated and applied as a correction factor back to f^{MC} in the corresponding bins of the p_T of the mother jet of the probe lepton. Prompt rates are corrected in each lepton p_T, lepton $|\eta|$ and number of tracks bin by $k^{\mathrm{p}}_{\mathrm{data}} = p^{\mathrm{data}}/p^{\mathrm{DY\ MC}}$.

Taking into account the variation of the rates in simulated and data enriched DY and only simulated $t\bar{t}$ events, a single prompt, and, a fake rate is defined for each bin as given in Eq. 5.2.

$$p = k^{\mathrm{p}}_{\mathrm{data}} \cdot \frac{p^{\mathrm{DY\ MC}} + p^{t\bar{t}\ \mathrm{MC}}}{2},$$

$$f = k^{\mathrm{f}}_{\mathrm{data}} \cdot \frac{f^{\mathrm{DY\ MC}} + f^{t\bar{t}\ \mathrm{MC}}}{2}. \qquad (5.2)$$

The electron/muon fake rates measured as a function of p_T of the looser electron/muon in bins of electron/muon p_T for barrel and endcap regions are provided from Figs. 5.7 to 5.10. Between 2016 and 2017, there is a systematic increase in fake rate in 2017. The primary reason why fake rate is different is because of a looser ID for 2016 which results in populated side-bands. This results in low fake rate compared to 2017. Thus, a higher fake rate does not necessarily mean more fakes, but in this case, is because of more populated side-bands. The electron/muon prompt rates measured as

a function of p_T of the looser electron/muon in bins of electron/muon p_T for barrel and endcap regions are provided from Figs. 5.11 to 5.14. As can be seen from the figures, the electron and muon prompt rates are measured to be $\gtrsim 80\%$, whereas the fake rates vary in the range of 5-25% for both lepton flavors. This is also expected since prompt leptons (that come from bosons) have a very high chance of passing tighter quality criteria, where misidentified leptons have a lower chance of doing the same.
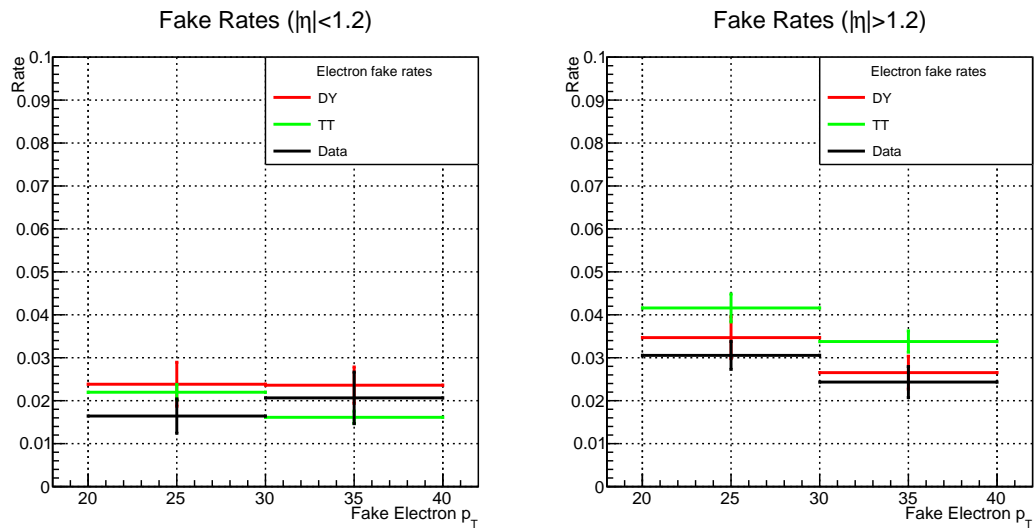
Figure 5.7: Electron fake rates in 2016 data and MC.
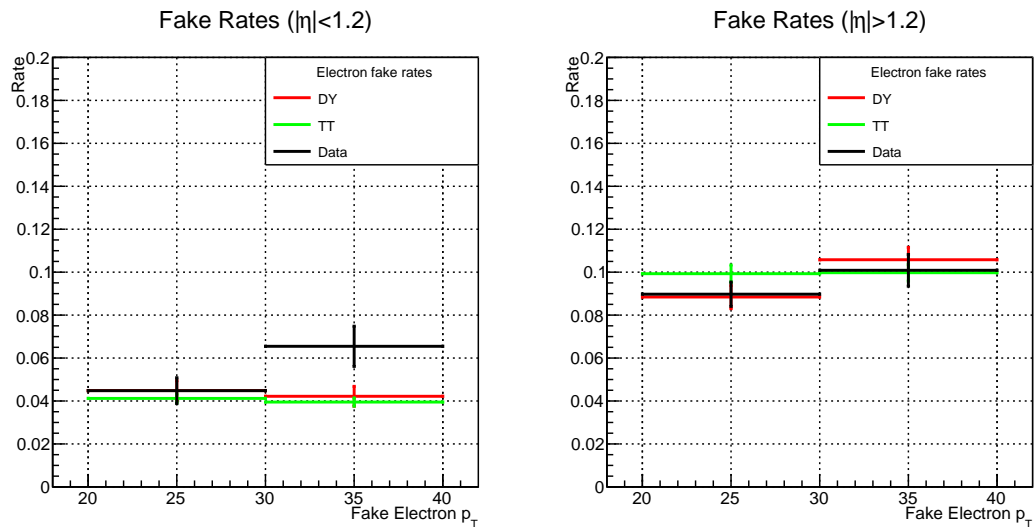
Figure 5.8: Electron fake rates in 2017 data and MC.
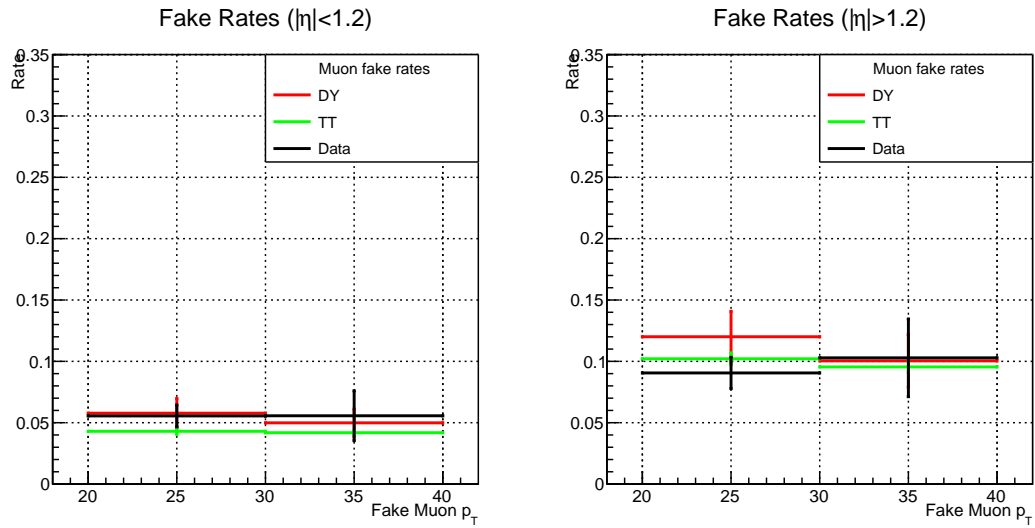
Figure 5.9: Muon fake rates in 2016 data and MC.
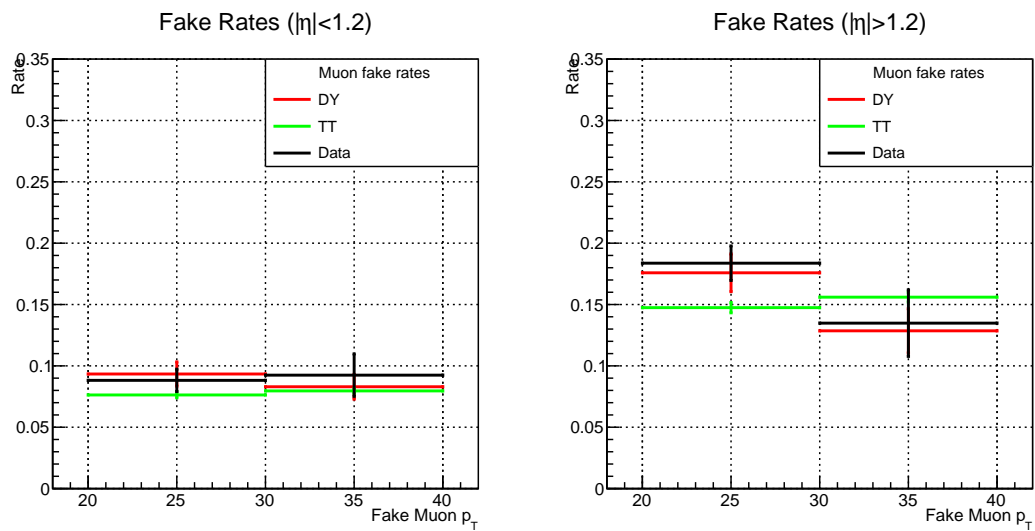
Figure 5.10: Muon fake rates in 2017 data and MC.
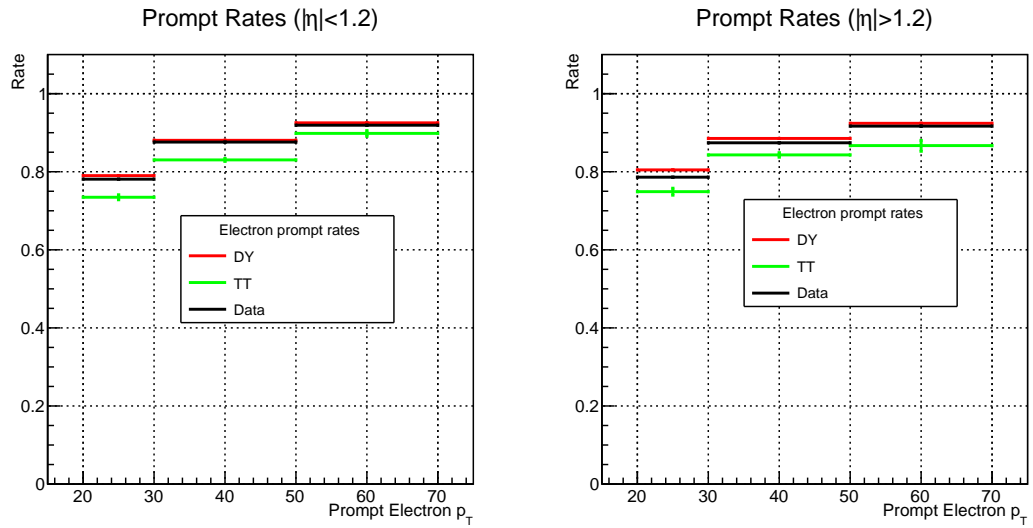
Figure 5.11: Electron prompt rates in 2016 data and MC.
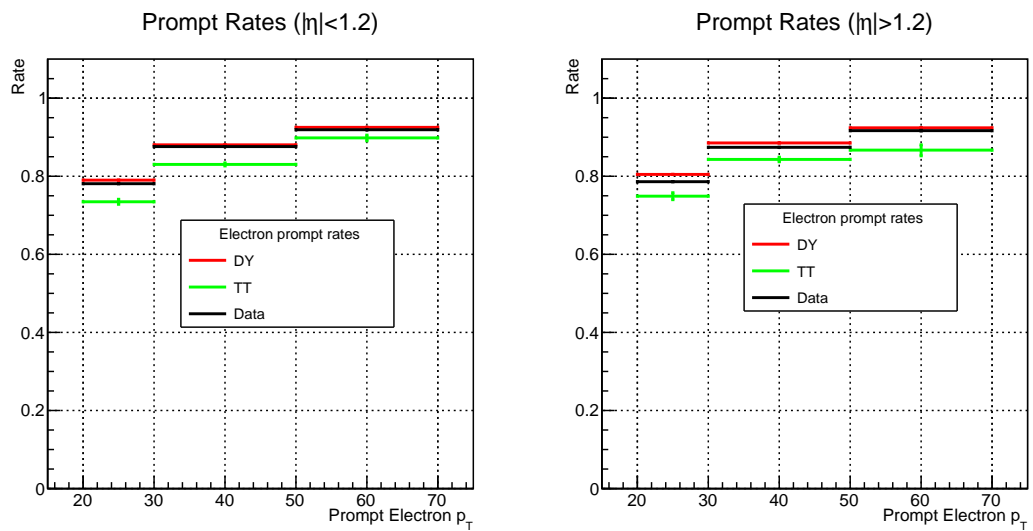
Figure 5.12: Electron prompt rates in 2017 data and MC.
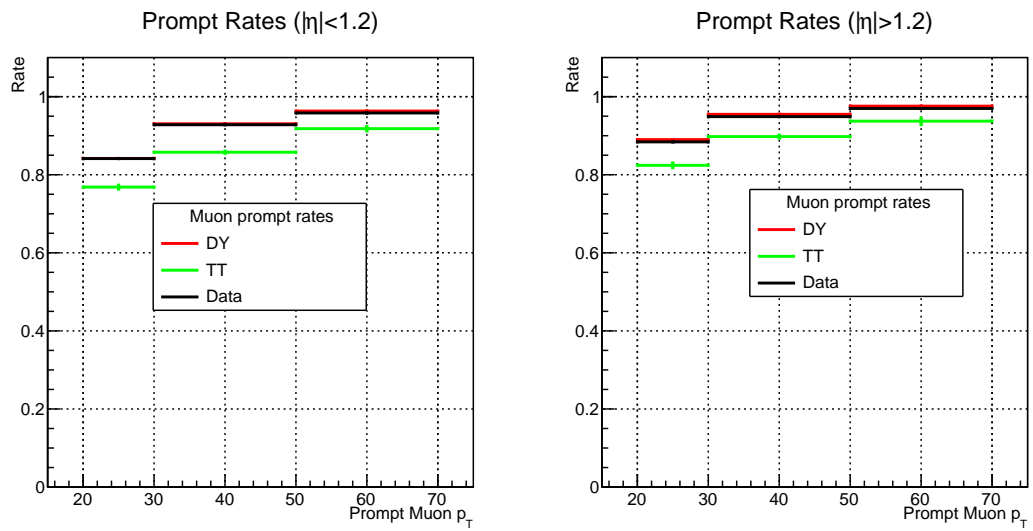
Figure 5.13: Muon prompt rates in 2016 data and MC.
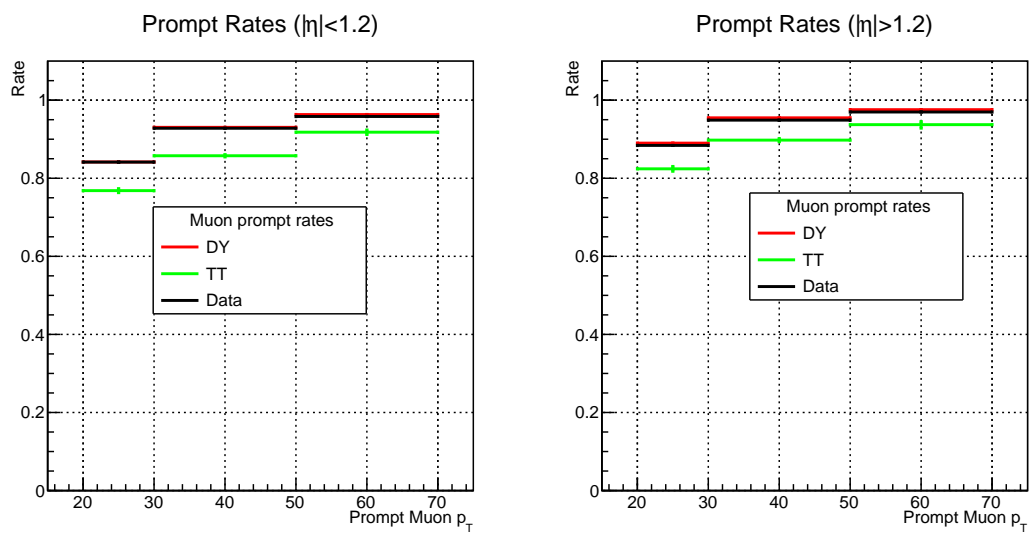
Figure 5.14: Muon prompt rates in 2017 data and MC.

The DY and $t\bar{t}$ samples are specifically chosen as they constitute the major background processes with fake leptons in this analysis and yet differ significantly in terms of event hadronic activity and kinematics.

The individual simulation-based prompt and fake rates are verified in closure tests performed in the simulated DY+jets and $t\bar{t}$+jets samples where at least one misidentified lepton is required as a part of event selection. The self-closure of the matrix method is shown in Figs. 5.15 and 5.16 for 2016 and 2017, respectively.
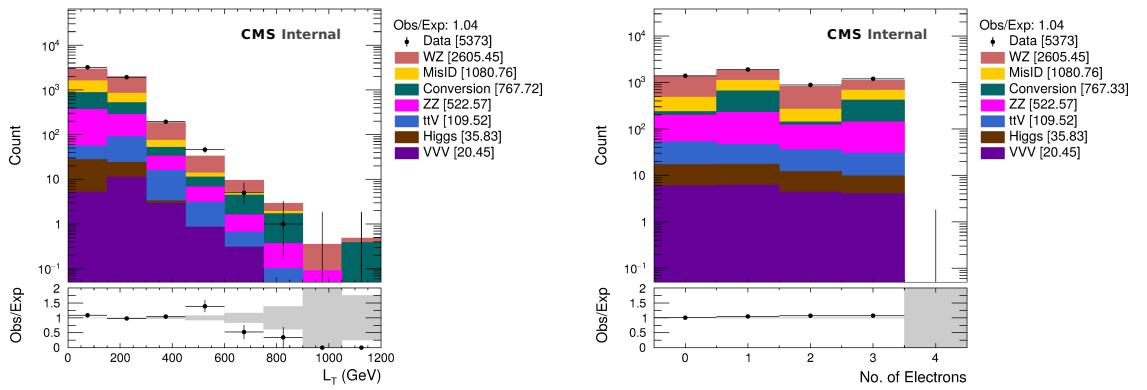
Figure 5.15: Self-closure tests in DY enriched 2016 data for the L_T and electron multiplicity distributions.
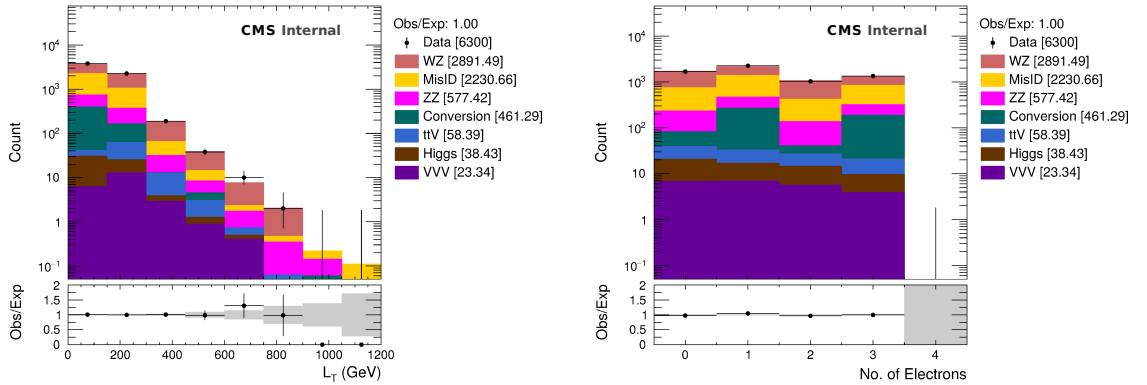
Figure 5.16: Self-closure tests in DY enriched 2017 data for the L_T and electron multiplicity distributions.

5.2.6 Misidentified tau lepton backgrounds

Events with exactly two light leptons with an additional one or more τ_h are classified by dilepton pair being SS or OS. In the regions with two light-leptons and a τ_h, the most dominant background comes from processes with two prompt light leptons and a fake τ_h, and fully leptonic decays of Z and W boson in DY + jets and $t\bar{t}$ + jets are the most important processes resulting in this final state. Other smaller backgrounds are from fully leptonic decays of diboson events (WZ, ZZ) and some rare SM processes like ttV and Higgs.

In events with an OS pair + τ_h, the τ_h is most likely the misidentified lepton while for SS pair + τ_h this need not be true.

5.2.6.1 Tau fake and prompt rates

Tau lepton fake rates are measured in events with two opposite sign, same flavor leptons satisfying tight lepton selection with an additional loose τ_h. Light lepton pair invariant mass is required to be on-Z. A $p_T^{miss} < 50$ GeV requirement is also imposed to create a signal depleted, DY + jets enriched selection of events to measure and validate τ_h fake rates.

The fake rates are defined as the ratio of fake τ_h enriched events described above where τ_h also satisfies a tight τ_h selection given that it satisfies loose τ_h selection. Fake rates are measured separately for τ_h's reconstructed with one charged hadron (1 prong) and with three charged hadrons (3 prongs) in the barrel ($|\eta| < 1.46$) and endcap ($|\eta| > 1.56$) region of the detector. Then the τ_h p_T and p_T of the mother jet (AK4 PF jet matched to τ_h) are used for fake rate parametrization. The τ_h fake rates are measured as a function of mother jet p_T in regions with τ_h $p_T < 30, 30-50$ and $50-100$ GeV. For τ_h $p_T > 100$ GeV inclusive fake rates are measured. The mother jet p_T parameterization is used both to minimize the spread of fake rates in DY vs. $t\bar{t}$ events, and also to capture a particular recoil effect which is most visible in DY + 1 jet processes where the jet is misidentified as a fake τ_h. Since τ_h isolation is defined by an MVA that is based on absolute (rather than relative) isolation quantities, it is observed that the τ_h candidates with a significant mismatch between τ_h p_T and the mother jet p_T are more likely to fail

53

tighter isolation requirements. This effect is always present but most visible in DY + 1 jet events where the recoil of the Z boson is correlated with the mother jet p_T of the fake τ_h object, and it affects the τ_h fake rate irrespective of the τ_h p_T. A mother jet p_T parametrization is used to capture this effect of mismatch between τ_h p_T and τ_h-jet p_T. Tau fake rates as a function of the mother jet p_T as measured in the data, DY+jets and $t\bar{t}$ simulation samples are provided in Fig 5.17-5.27 for both 2016 and 2017 data and MC. The data fake rates are shown in each bin for illustrative purposes. The actual correction is applied as an inclusive correction independent of tau lepton mother jet p_T.

Figure 5.17: Tau fake rates as a function of the mother jet p_T, for 1 prong tau leptons with $20 < p_T < 30$ GeV in 2016 data and MC. The lower bound comes about due to the minimum p_T of the selected reconstructed tau leptons as mentioned in Section 4.6.5.

An additional correction factor to τ_h fake rates is derived from accounting for the hadronic activity dependence of rates. The fake rates are measured as a function of the number of high purity tracks in the events and divided that with the average τ_h fake rate. This then gives us a correction to τ_h fake rates as a function of the number of tracks. The plots for these correction factors for the 2016 and 2017 version of the matrix method are shown in Fig. 5.28.

The prompt rates for tau leptons are measured in simulated DY + jets samples. It is required that a loose tau lepton candidate matches to a gen tau lepton within a $\Delta R < 0.2$ and measure rate at which it also satisfies the tight working point. For 2017 analysis tau
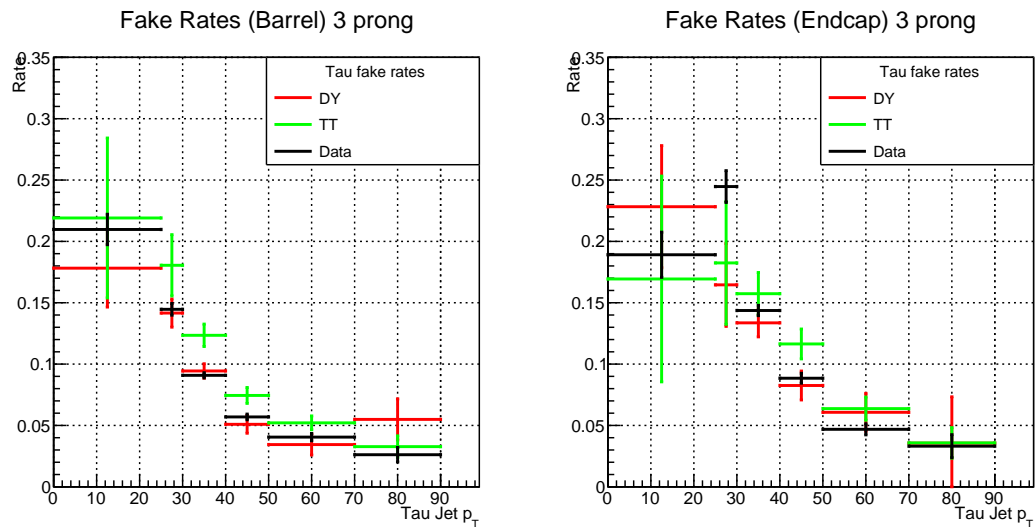
Figure 5.18: Tau fake rates as a function of the mother jet p_T, for 3 prong tau leptons with $20 < p_T < 30$ GeV in 2016 data and MC. The lower bound on the p_T of the tau lepton comes about due to the minimum p_T of the selected reconstructed tau leptons as mentioned in Section 4.6.5.
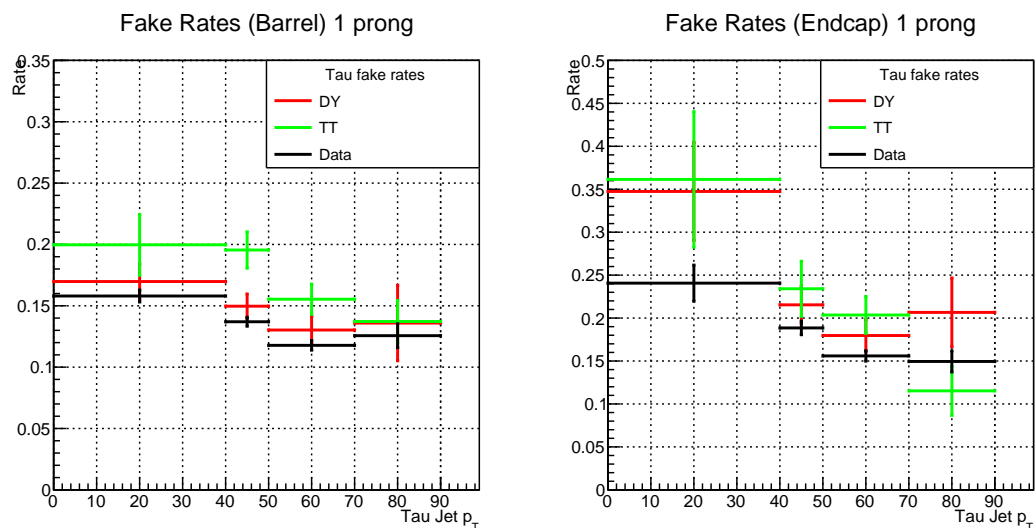
Figure 5.19: Tau fake rates as a function of the mother jet p_T, for 1 prong tau leptons with $30 < p_T < 50$ GeV in 2016 data and MC.
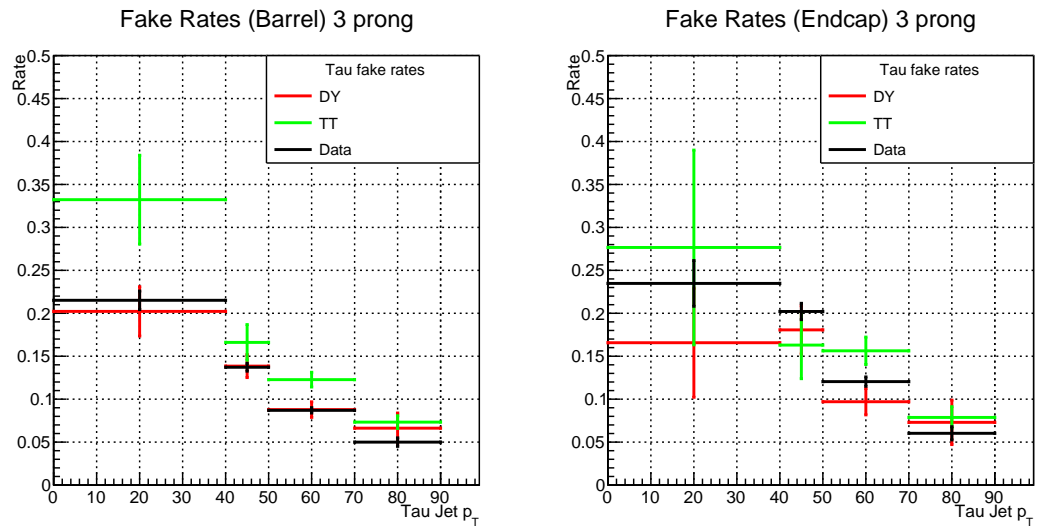
Figure 5.20: Tau fake rates as a function of the mother jet p_T, for 3 prong tau leptons with $30 < p_T < 50$ GeV in 2016 data and MC.
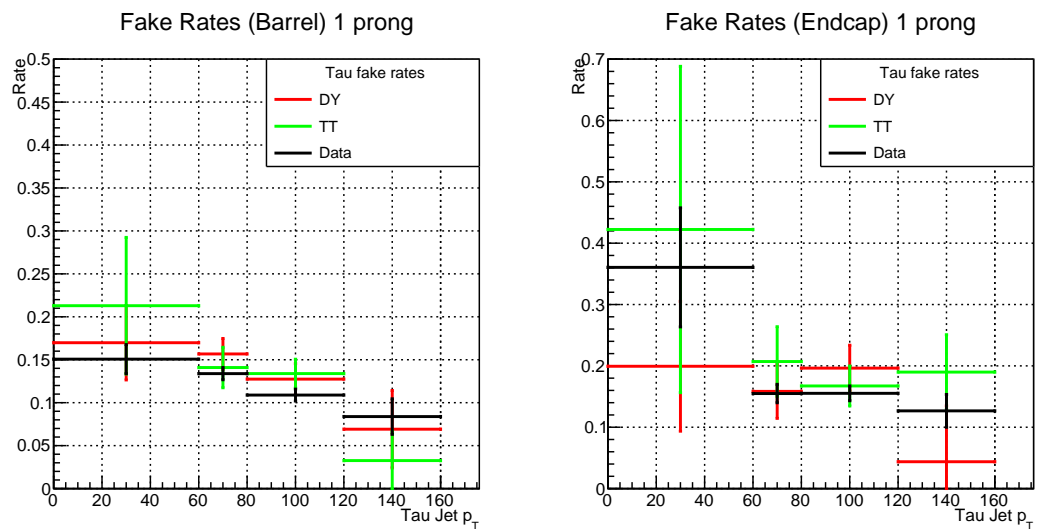
Figure 5.21: Tau fake rates as a function of the mother jet p_T, for 1 prong tau leptons with $50 < p_T < 100$ GeV in 2016 data and MC.
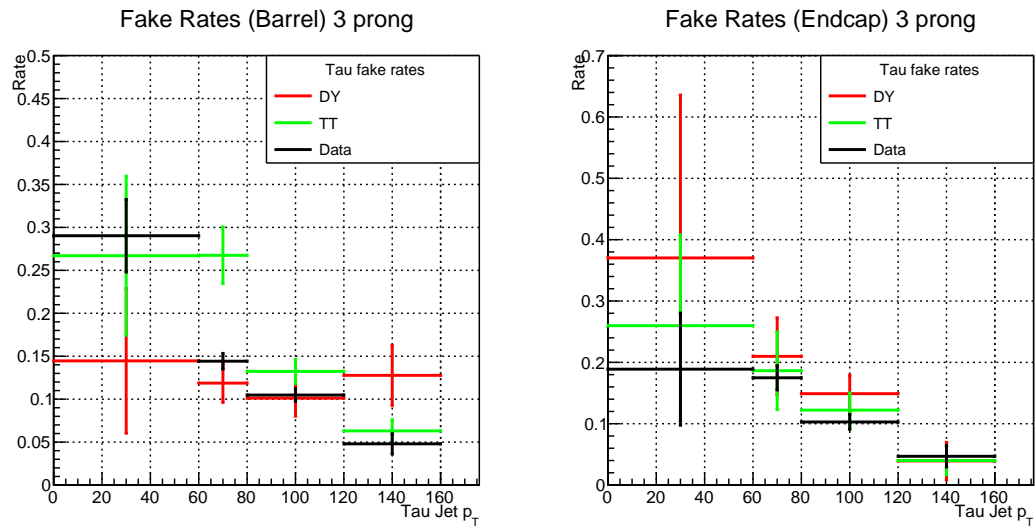
Figure 5.22: Tau fake rates as a function of the mother jet p_T, for 3 prong tau leptons with $50 < p_T < 100$ GeV in 2016 data and MC.
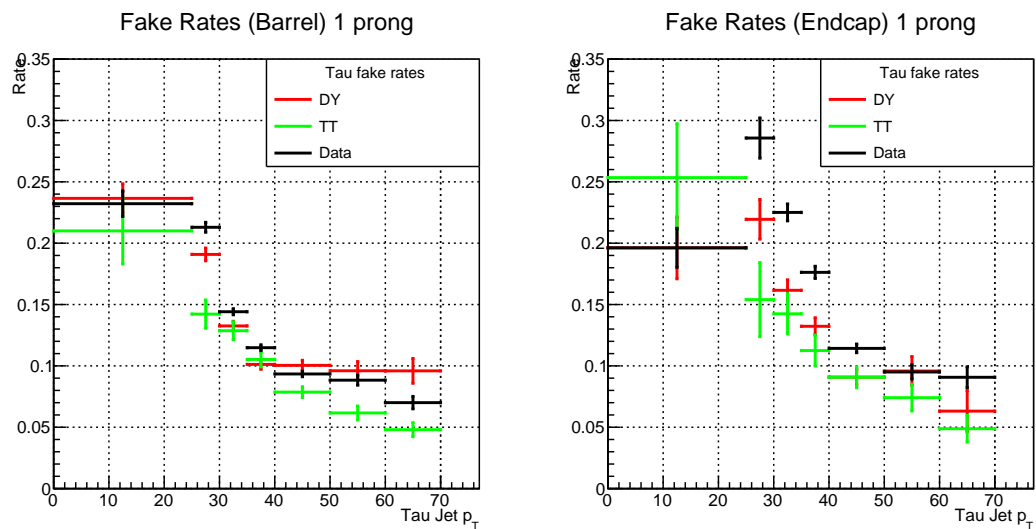
Figure 5.23: Tau fake rates as a function of the mother jet p_T, for 1 prong tau leptons with $p_T < 30$ GeV in 2017 data and MC.
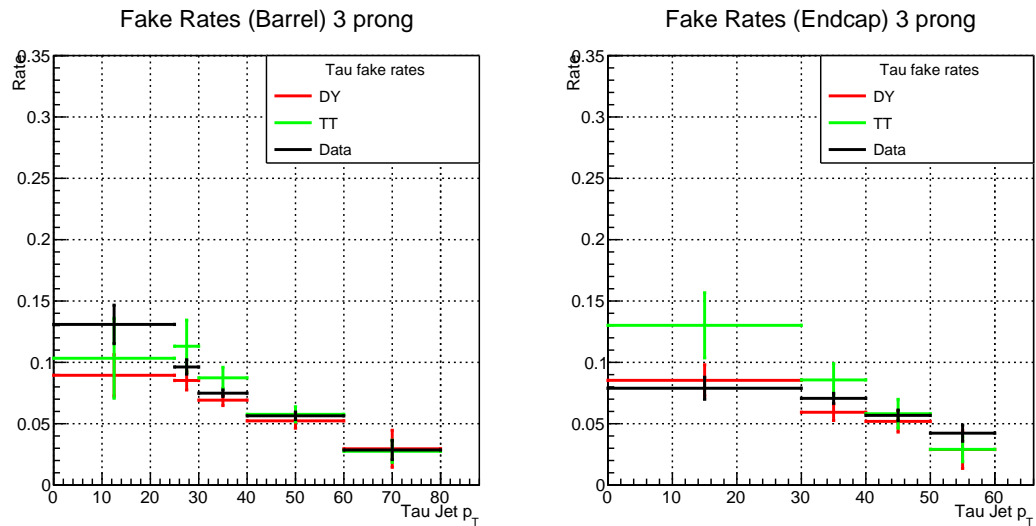
Figure 5.24: Tau fake rates as a function of the mother jet p_T, for 3 prong tau leptons with $p_T < 30$ GeV in 2017 data and MC.
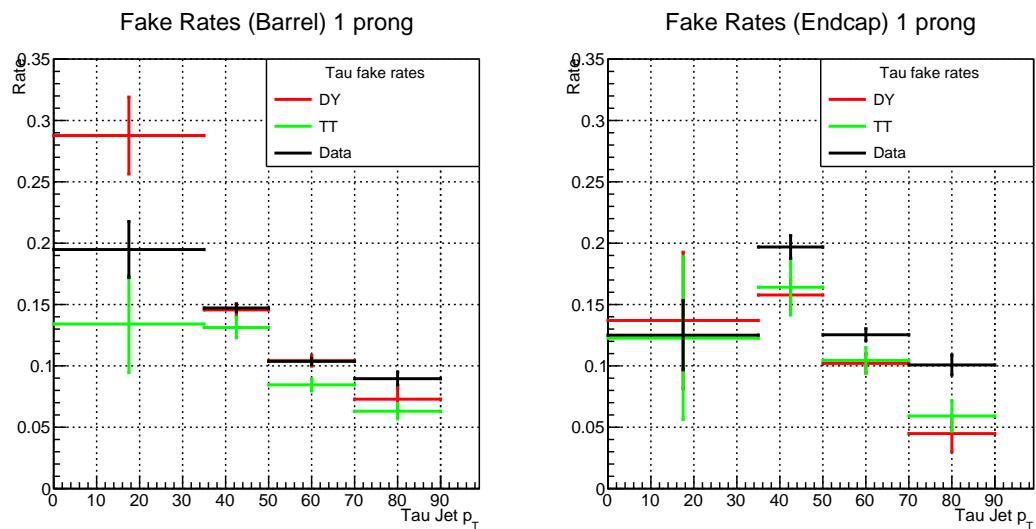
Figure 5.25: Tau fake rates as a function of the mother jet p_T, for 1 prong tau leptons with $30 < p_T < 50$ GeV in 2017 data and MC.
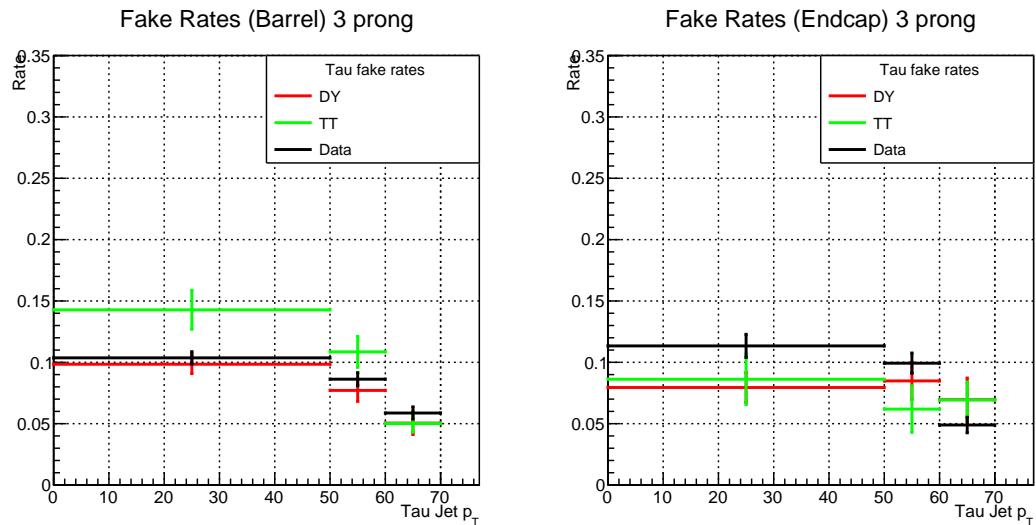
Figure 5.26: Tau fake rates as a function of the mother jet p_T, for 3 prong tau leptons with $30 < p_T < 50$ GeV in 2017 data and MC.
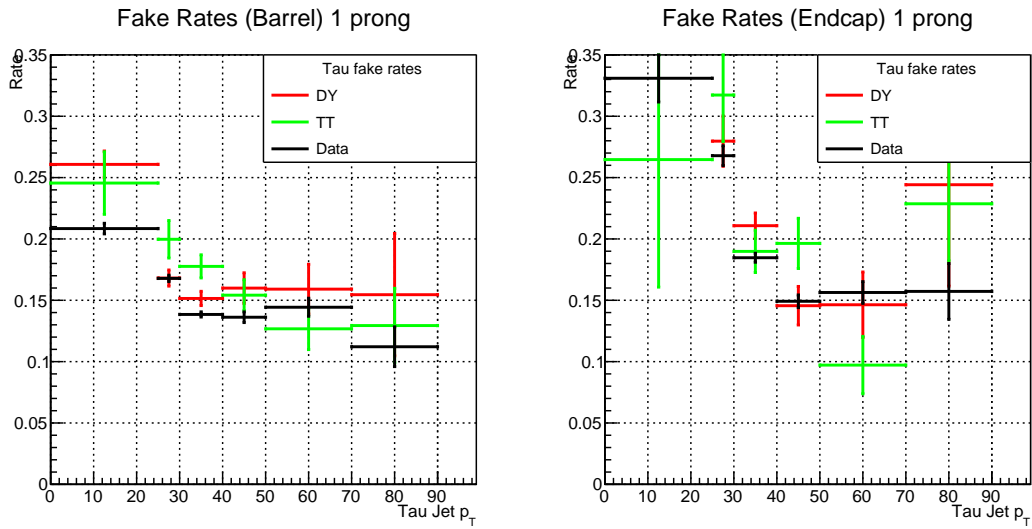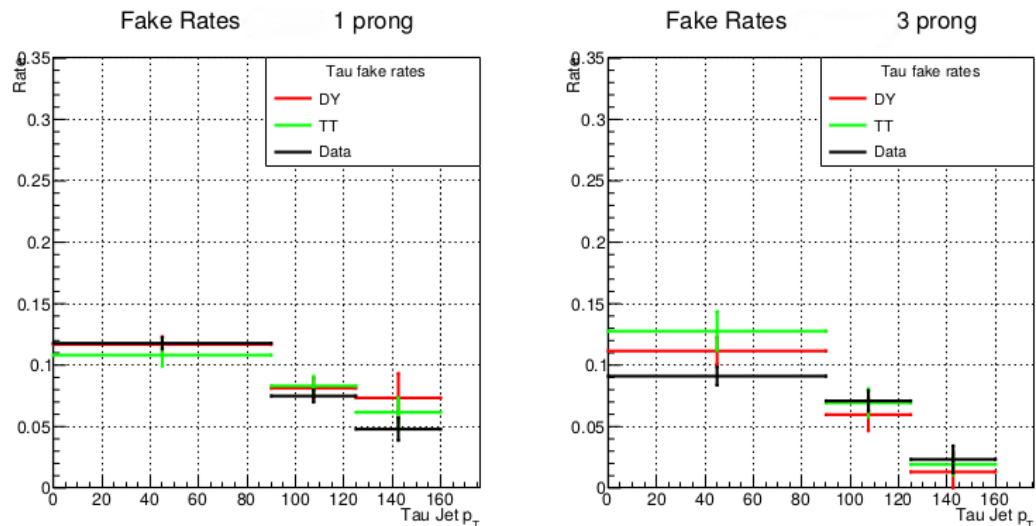

Figure 5.27: Tau fake rates as a function of the mother jet p_T, for all tau leptons with $50 < p_T < 100$ GeV in 2017 data and MC. For 2017 data, due to low statistics, rates for both 1-prong and 3-prong tau leptons with $50 < p_T < 100$ GeV are measured independent of detector region.
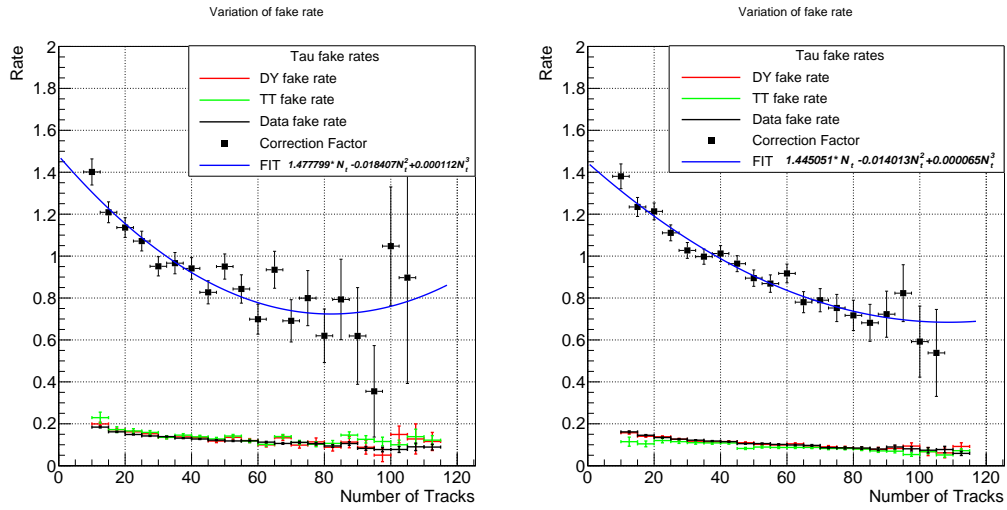
Figure 5.28: Fake rate correction factor as a function of number of high purity tracks in the event extracted from 2016 (left) and 2017 (right) data and MC. This correction factor is used to account for the hadronic activity dependence of tau lepton fake rate. A polynomial is fit to extract the correction factor.

lepton p_T based prompt rates are measured, which are roughly between $50 - 55\%$. For the 2016 analysis, a 61% flat rate is sufficient. Fig. 5.29 demonstrates the 2016+2017 agreement between the expected background and the observed data yields, as a function of the dilepton mass and L_T, in a signal-depleted 2L1T (OS) selection.
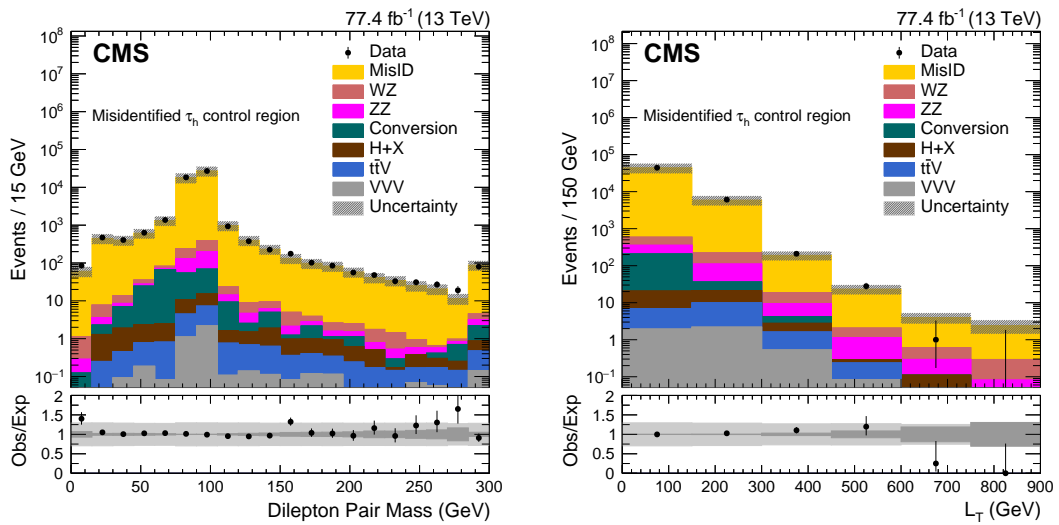
Figure 5.29: The dilepton mass (left) and the L_T (right) distributions in data and simulation in a misidentified τ_h control region. This control region contains 2L1T (OS) events with $p_T^{miss} < 50$ GeV. The total SM background is shown as a stack of all contributing processes. The hatched gray bands in the upper panels represent the total uncertainty in the expected background. The lower panels show the ratios of observed data to the total expected background. In the lower panels, the light gray band represents the combined statistical and systematic uncertainty in the expected background, while the dark gray band represents the statistical uncertainty only. The rightmost bins include the overflow events.

5.3 Systematics

5.3.1 Systematic uncertainties on misidentified lepton background

The systematic uncertainties on the matrix method prediction are estimated by varying the fake/prompt rates within respective uncertainties and observe their effect on the background predictions. Tau fake rate uncertainty comes from various sources which include DY-$t\bar{t}$ difference, data MC correction factors, and $|\eta|$, NTrk and p_T correction factors as explained in the previous subsection. DY-$t\bar{t}$ difference is generallyquitelow (5-10%) and correction factor uncertainty vary between 5-30%. All these uncertainties are added in quadrature to estimate the τ_h fake rate uncertainty. The τ_h lepton fake rate is then varied with these uncertainties to get the upper and lower one sigma variations of matrix method predictions. The impact of varying rates rates on LT distribution in different signal regions is provided from figure Fig 5.31 to 5.34. Overall, a 20-40% change in the background yields is observed across signal regions bins, and this is assigned as the systematic uncertainty on misidentified lepton background yield.
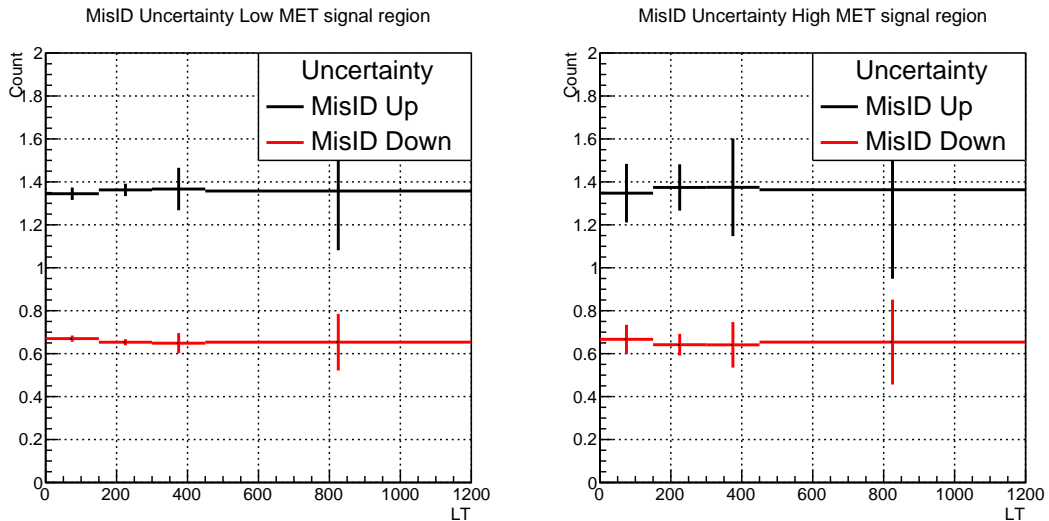
Figure 5.30: Variations of fake rates in the trilepton signal regions done with 2016 data: low p_T^{miss} (left plot) and high p_T^{miss} (right plot).

Closure tests are performed in DY + jets and $t\bar{t}$ + jets MC sample to make sure our method works as expected. For closure tests, MC samples are treated as data to make predictions for the fake contribution and the compared to actual events seen in the MC.
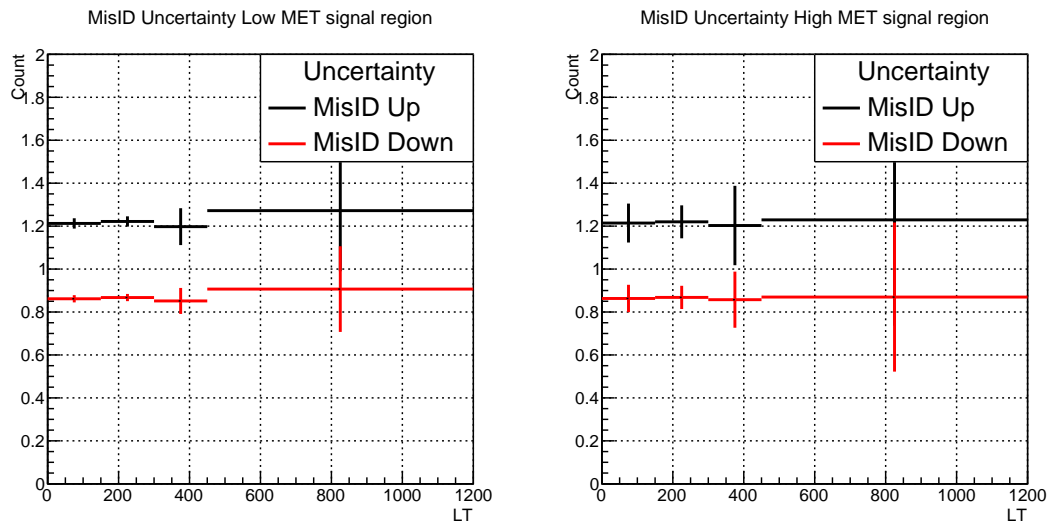
Figure 5.31: Variations of fake rates in the trilepton signal regions done with 2017 data: low p_T^{miss} (left plot) and high p_T^{miss} (right plot).
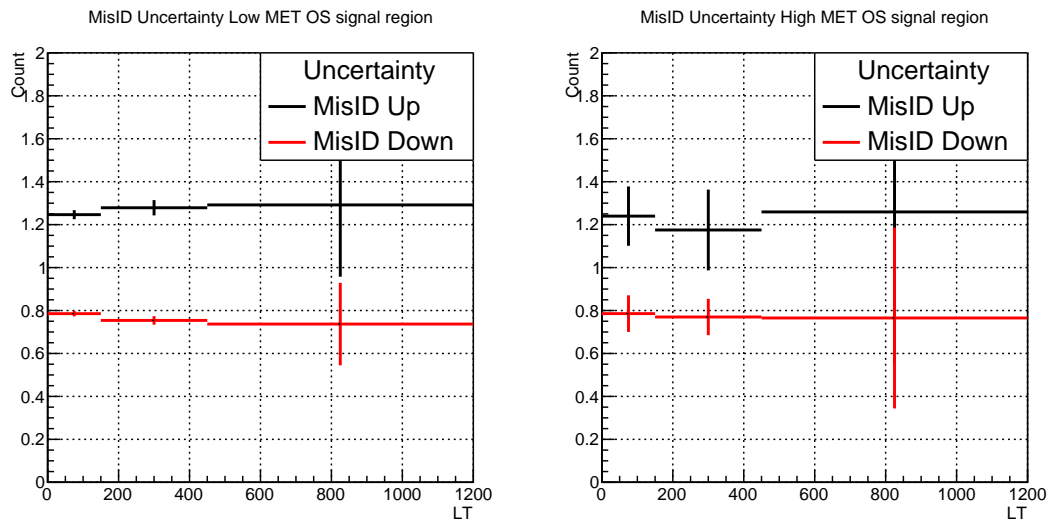
Figure 5.32: Variations of fake rates in the 2L(OS)1T signal regions done with 2016 data: low p_T^{miss} (left plot) and high p_T^{miss} (right plot).

63
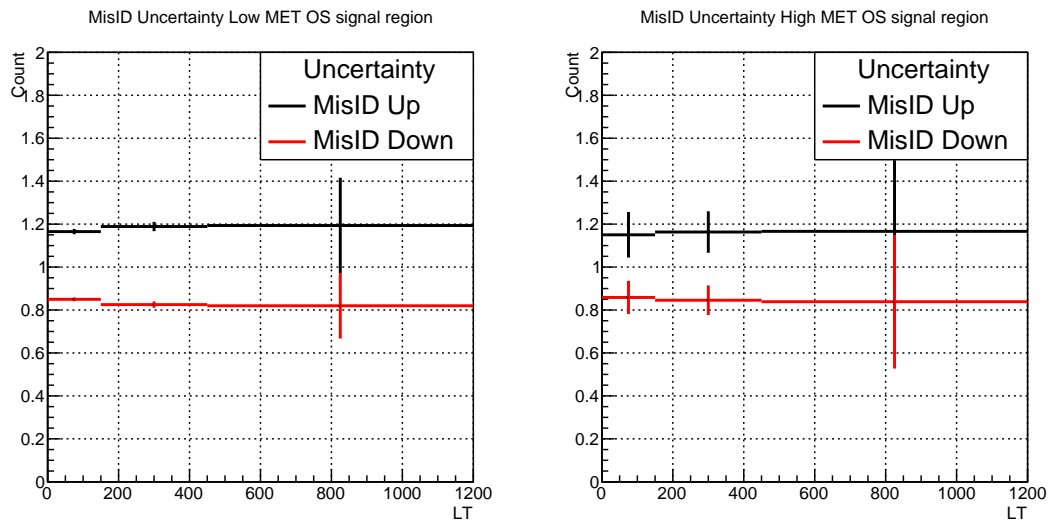
Figure 5.33: Variations of fake rates in the 2L(OS)1T signal regions done with 2017 data: low p_T^{miss} (left plot) and high p_T^{miss} (right plot).
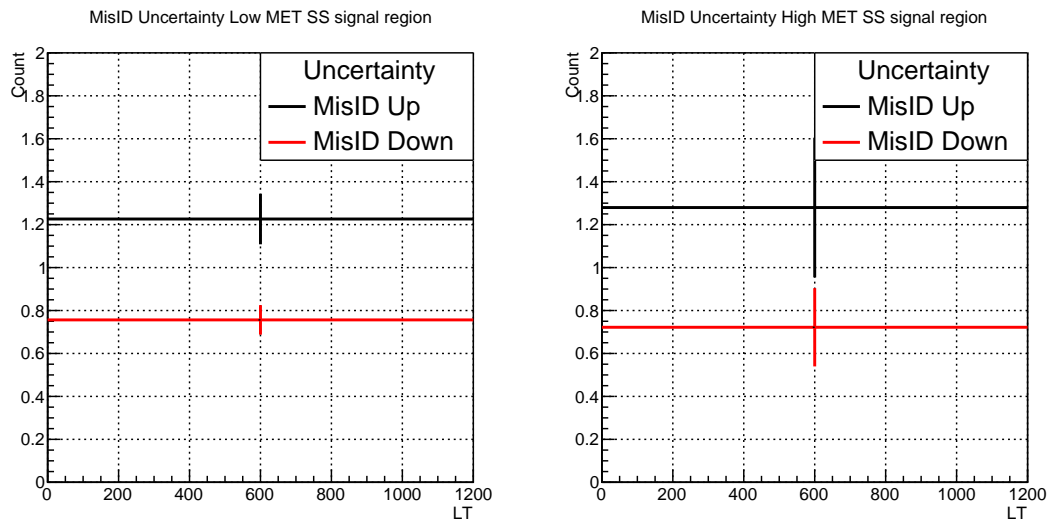
Figure 5.34: Variations of fake rates in the 2L(SS)1T signal regions done with 2016 data: low p_T^{miss} (left plot) and high p_T^{miss} (right plot).
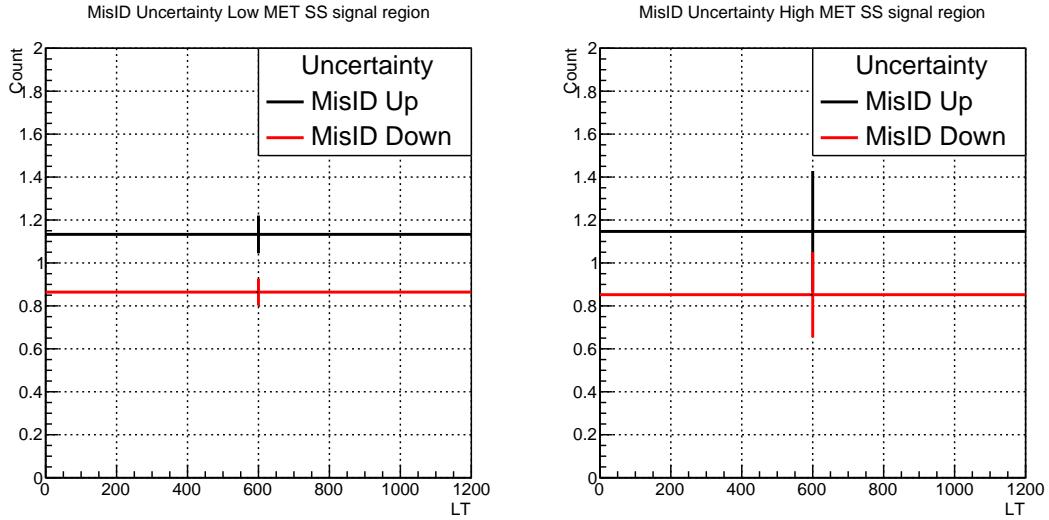
Figure 5.35: Variations of fake rates in the 2L(SS)1T signal regions done with 2017 data: low p_T^{miss} (left plot) and high p_T^{miss} (right plot).

An excellent agreement is seen overall, and predictions in all the regions agree well within 20%. A few deviations larger than 20% are seen in tails where the statistics are a limiting factor.

5.3.2 Other uncertainties

Since the WZ and ZZ backgrounds are normalized in dedicated control regions, a relative uncertainty arises due to this normalization. It is 5(5)% on the WZ(ZZ) background, in analysis with 2016 data. For 2017 analysis, the uncertainty is 5(4)% on the WZ(ZZ) background. Since background contributions due to all-prompt lepton events are estimated via simulated samples, several different sources of systematic uncertainty are considered to account for differences between MC and data events.

The uncertainties on the muon and electron identification, isolation efficiency scale factors are 2%, the τ_h identification and isolation uncertainties are 3% which are applied per-lepton in each event. The trigger efficiency scale factors have an overall <2-3% impact per event. The jet, unclustered energy, and lepton energy (or momentum) scale uncertainties as well as jet and lepton resolution uncertainties are applied at the per-object level, where the corresponding object p_Ts are varied up and down within the recommended uncertainty range. The impact of these variations on the kinematic

quantities is then observed. Additionally, a luminosity uncertainty of 2.6% is applied to the MC based rare background estimates as well as the signal yields since these are not normalized to data but the theoretical cross sections. For the sub-dominant background processes such as $t\bar{t}V$, triboson, or associated Higgs production, a 50% systematic uncertainty is applied on the theoretical normalization cross sections to cover both PDF as well as re-normalization and factorization scale uncertainties. Additionally, the recommended luminosity uncertainty of 2.3% [68] is applied to the MC based rare background estimates as well as the signal yields since these are not normalized to data but the theoretical cross-sections. The pileup reweighing uncertainty is evaluated by varying the minimum bias cross section used in the reweighing procedure up and down by 5% and is applied to all MC based backgrounds. The impact on MC yield from varying pileup cross section up and down by 5% for 2016 and 2017 is shown in Fig. 5.36.

The uncertainties on the muon and electron identification, isolation efficiency scale factors are $\sim$ 2% and applied per-lepton in each event. The uncertainties on the MC based background yields from jet energy scale, electron energy scale, and muon energy scale in 2016 and 2017 are shown in Figs. 5.37 and 5.38, respectively. The trigger efficiency scale factors have an overall 2-8% impact per event, and the impact is shown in Figs. 5.39 and 5.40.

A summary of the uncertainty sources considered in this analysis and the corresponding typical variations in the affected processes is provided in Table 5.2.
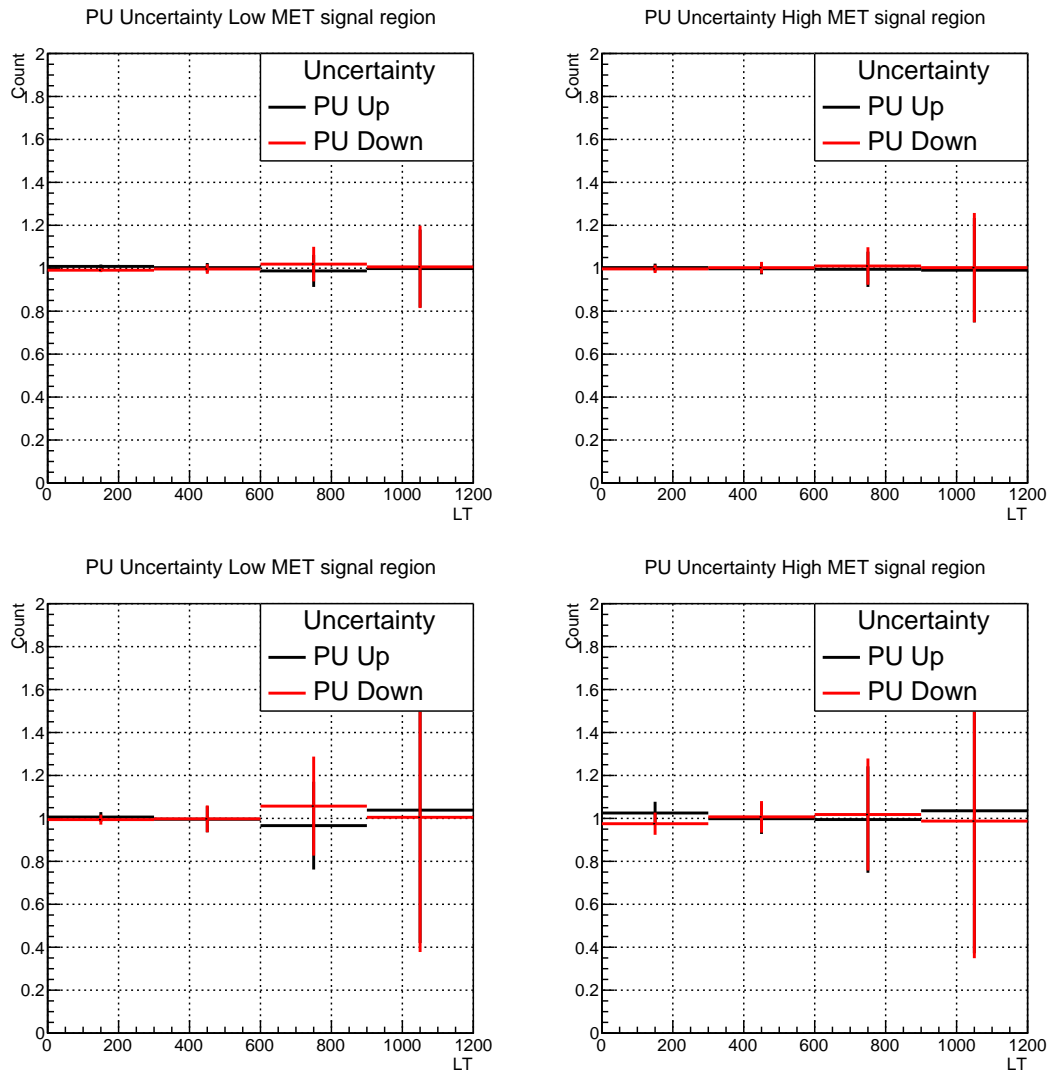
Figure 5.36: Impact of varying pileup xsec up and down by 5% on L_T distribution in 3L 2016 (top) and 2017 (bottom) MC.
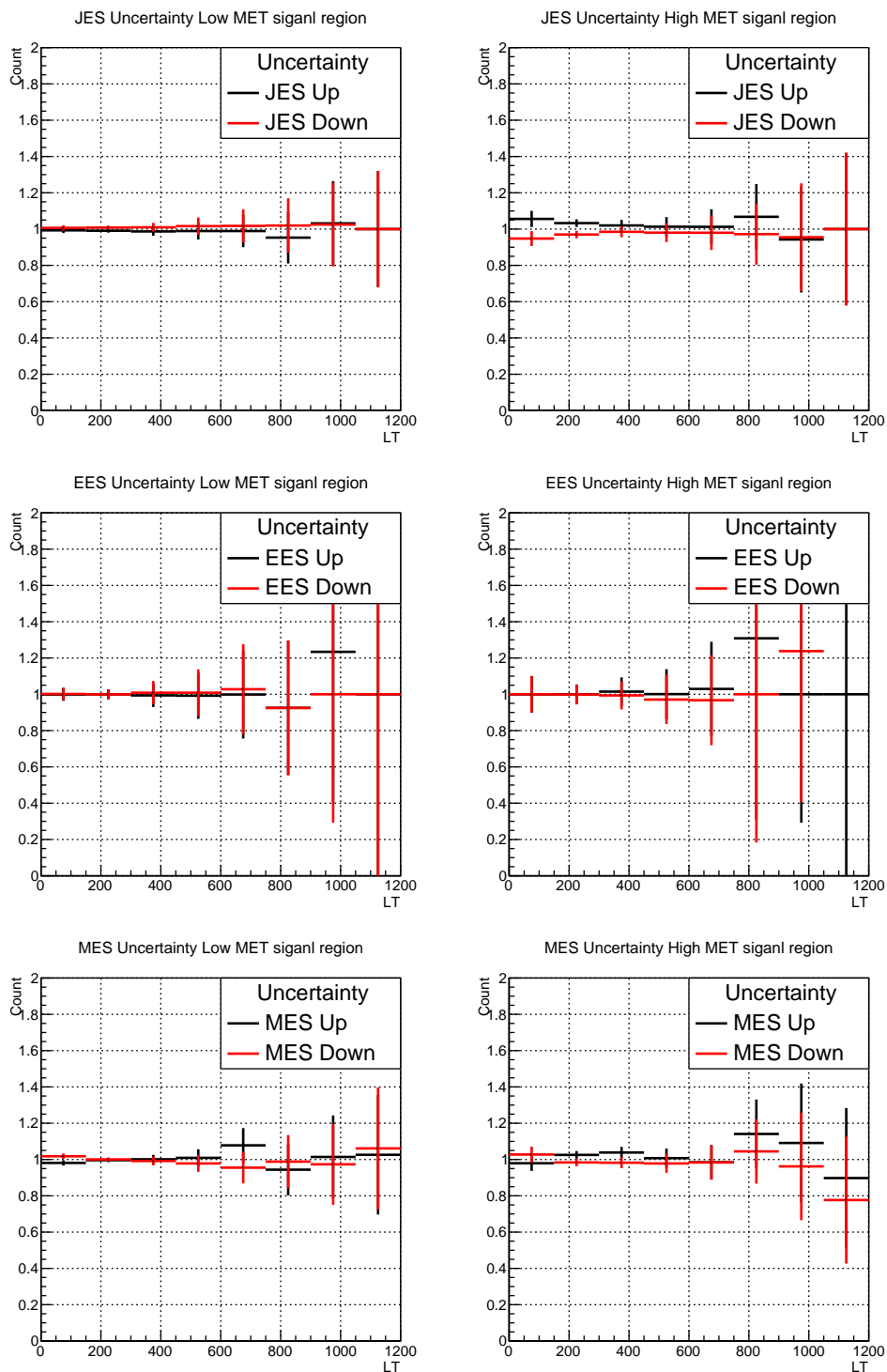
Figure 5.37: Impact of varying different uncertainties up and down on L_T distribution in 3L and 2L1T channels in 2016 MC.
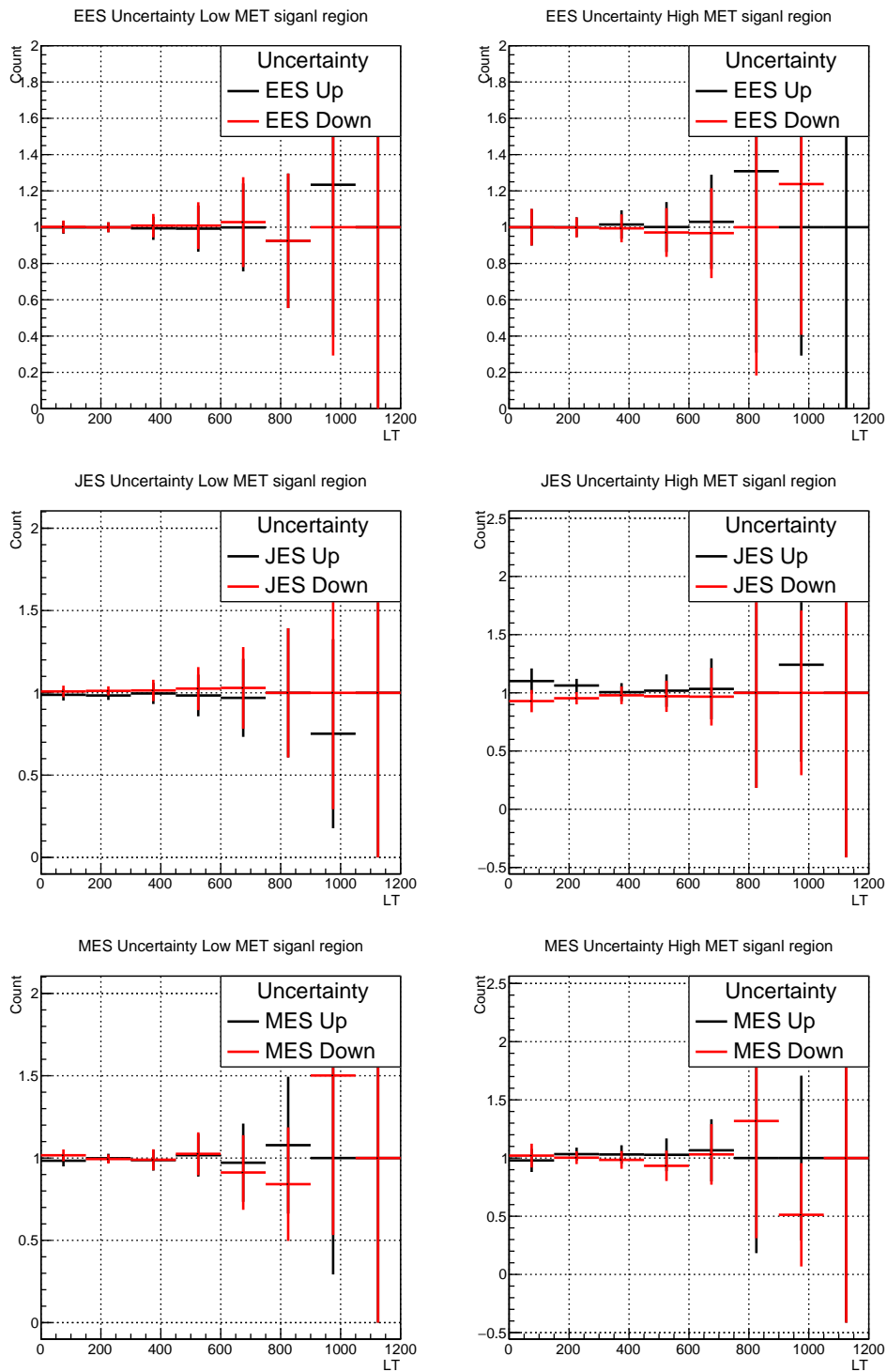
Figure 5.38: Impact of varying different uncertainties up and down on L_T distribution in 3L and 2L1T channels in 2017 MC.
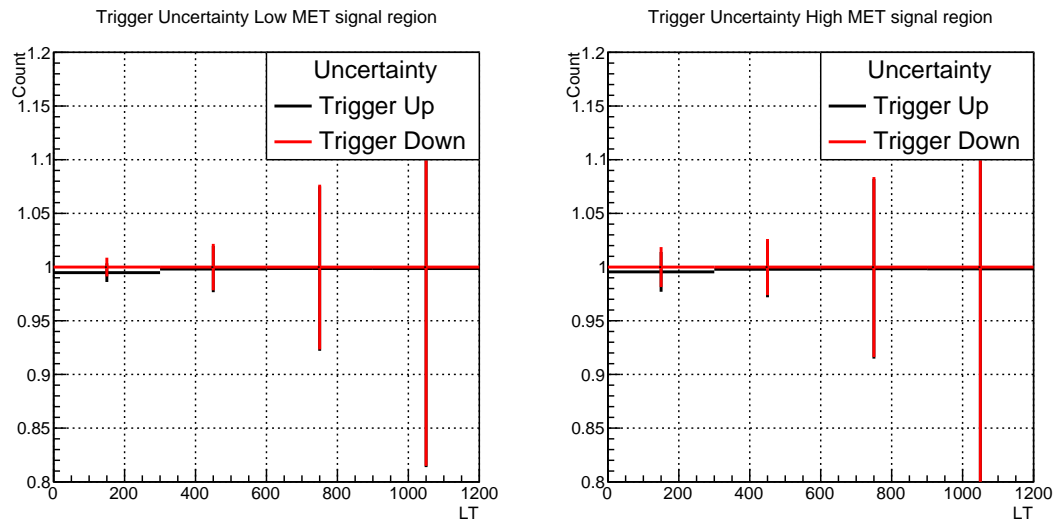
Figure 5.39: Impact of varying trigger efficiency up and down by 2% on L_T distribution in 3L 2016 MC. This a conservative variation irrespective of which trigger is fired.
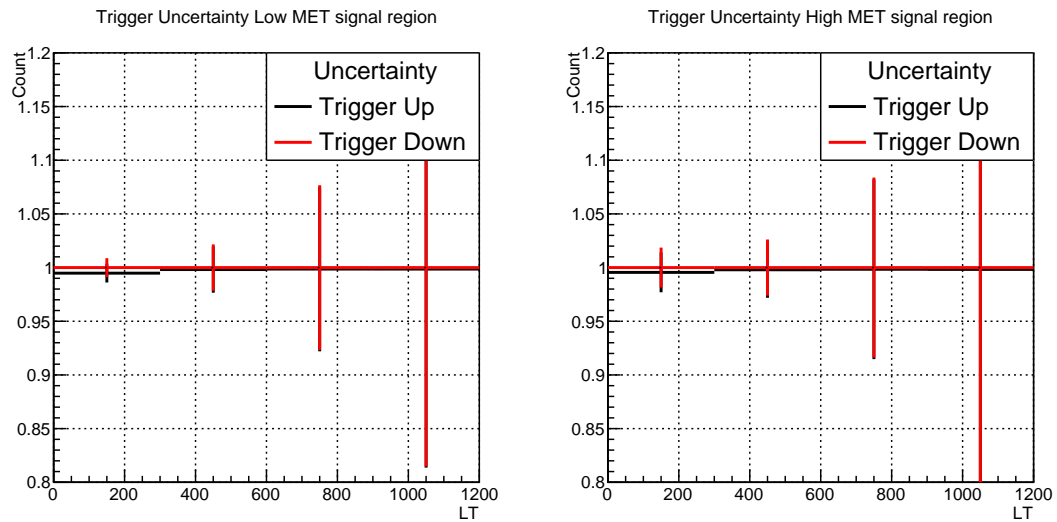
Figure 5.40: Impact of varying trigger efficiency up and down by 2% on L_T distribution in 3L 2017 MC. This a conservative variation irrespective of which trigger is fired.

70

Table 5.2: The sources of systematic uncertainty and the typical variations (%) observed in the affected background and signal yields in the analysis. All sources of uncertainty are considered as correlated between the 2016 and 2017 data analyses except for the lepton identification and isolation, the single lepton trigger, and the integrated luminosity. The label ALL is defined as WZ, ZZ, Rare ($t\bar{t}V$, VVV, Higgs boson), and Signal processes.

Source of uncertainty	Typical variations (%)	Processes
MisID background	20–35	-
Rare background normalization	50	-
Conversion background normalization	11	-
WZ background normalization	5	-
ZZ background normalization	4–5	-
Lepton identification & isolation	6–8	ALL
Single lepton trigger	<3	ALL
Electron energy scale and resolution	2–5	ALL
Muon momentum scale and resolution	2–10	ALL
Hadronic τ lepton energy scale	<5	ALL
Jet energy scale	5–10	ALL
Unclustered energy scale	1–10	ALL
Integrated luminosity	2.3–2.5	Rare/Signal
Pileup modeling	<4	ALL

5.4 Results

The L_T distributions for the 4L and 3L SRs (2016+2017) are shown in Fig. 5.41, while those for various 2L1T SRs are shown in Fig. 5.42. The individual 2016 and 2017 plots are provided in Appendix A. No significant discrepancies are observed between the background predictions and the observed data. Limits are set on the combined cross section for associated ($\tau'\nu'_\tau$) and pair ($\tau'\tau'/\nu'_\tau\nu'_\tau$) production of VLLs.

To obtain upper limits on the signal cross section at 95% confidence level (CL), a modified frequentist approach is used with a test statistic based on the profile likelihood in the asymptotic approximation and the CLs criterion [69–71]. A linear interpolation of the expected event yields is used between the simulated signal samples in the limit calculations. Systematic uncertainties are incorporated into the likelihood as nuisance parameters with log-normal probability distributions, while statistical uncertainties are modeled with gamma functions. Channel-wise observed limits for VLLs at 95% confidence level with 2016 and 2017 data are shown in Fig. 5.43 and Fig. 5.44. The observed limits are within two standard deviations of the expected limits from the background-only hypothesis. Because of the preferential coupling of VLLs to τ leptons, the major contribution to these results comes from the 2L1T SRs. The measurements in these channels alone exclude VLLs in the mass range 120–740 GeV. On combining all the 4L, 3L, and 2L1T SRs, with the hypothesis of an SU(2) mass degenerate VLL doublet with couplings to the third generation SM leptons, the VLLs with a mass in the range of 120–790 GeV are excluded at 95% CL as shown in Fig. 5.45.
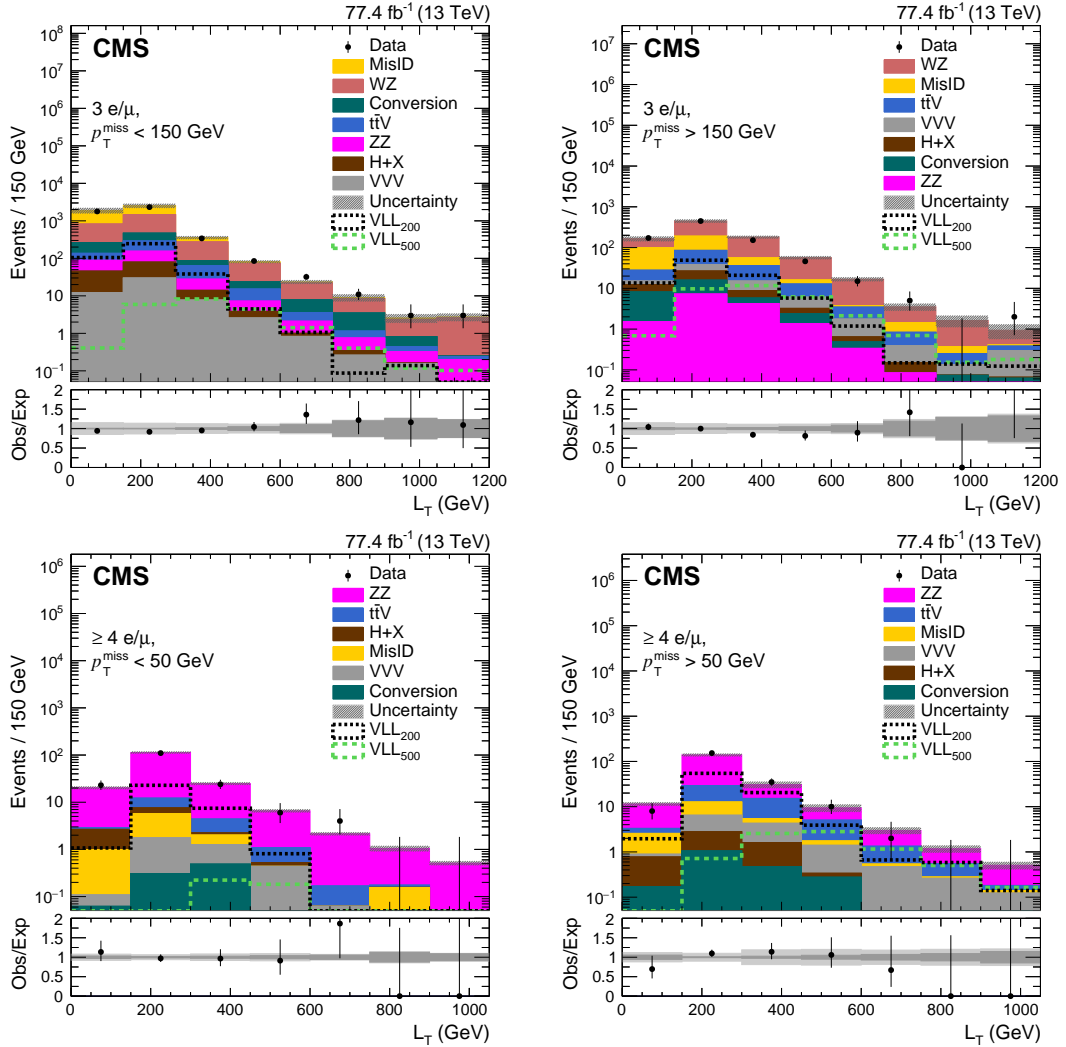
Figure 5.41: The L_T distributions for the 3L signal regions with $p_T^{miss} < 150$ GeV (upper left) and $p_T^{miss} > 150$ GeV (upper right), and for the 4L signal regions with $p_T^{miss} < 50$ GeV (lower left) and $p_T^{miss} > 50$ GeV (lower right). The total SM background is shown as a stack of all contributing processes. The predictions for VLL signal models (the sum of all production and decay modes) with $m_{\tau'/\nu'} = 200$ and 500 GeV are shown as dashed lines. The hatched gray bands in the upper panels represent the total uncertainty in the expected background. The lower panels show the ratios of observed data to the total expected background. In the lower panels, the light gray band represents the combined statistical and systematic uncertainty in the expected background, while the dark gray band represents the statistical uncertainty only. The rightmost bins include the overflow events.
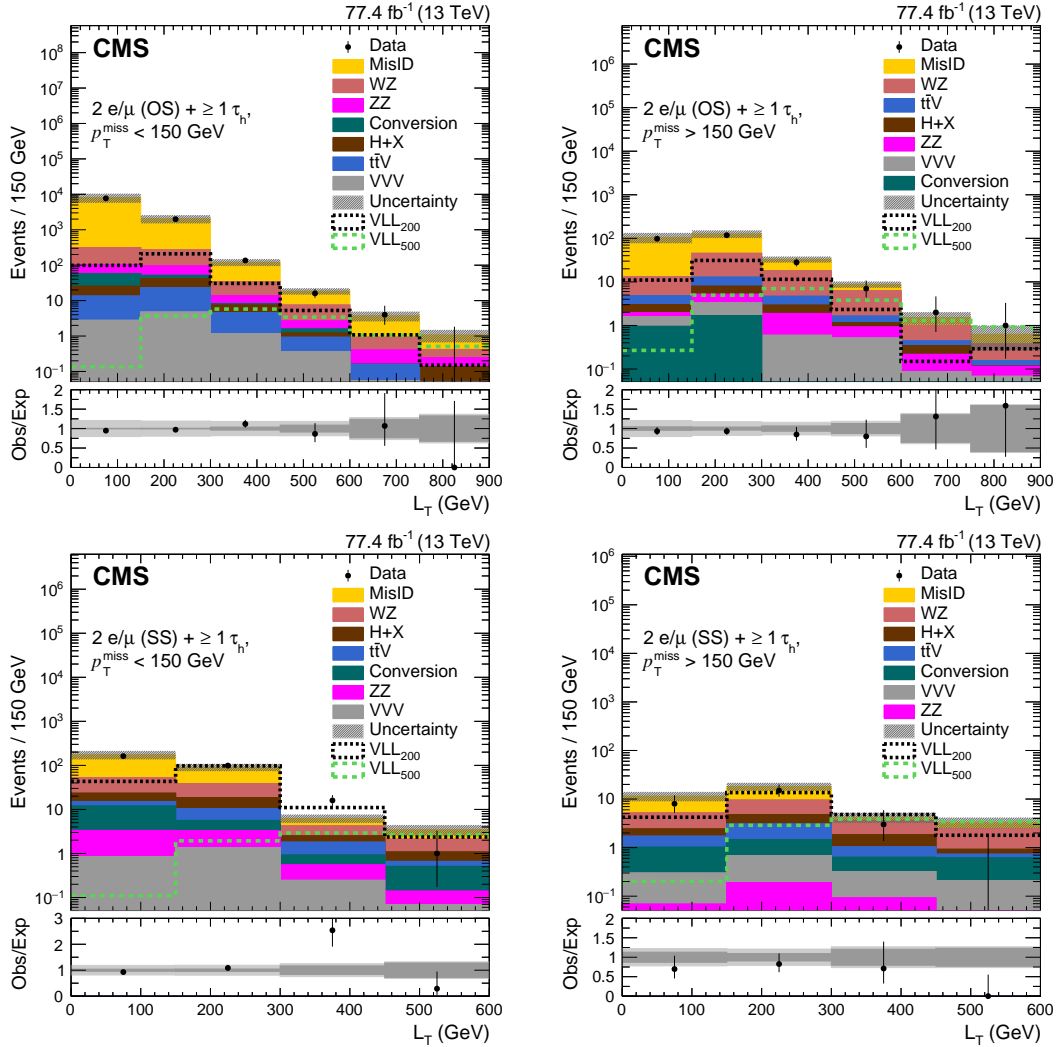
Figure 5.42: The L_T distributions for the 2L1T OS signal regions with $p_T^{miss} < 150$ GeV (upper left) and $p_T^{miss} > 150$ GeV (upper right), and for the 2L1T SS signal regions with $p_T^{miss} < 150$ GeV (lower left) and $p_T^{miss} > 150$ GeV (lower right). The total SM background is shown as a stack of all contributing processes. The predictions for VLL signal models (sum of all production and decay modes) with $m_{\tau'/\nu'} = 200$ and 500 GeV are also shown as dashed lines. The hatched gray bands in the upper panels represent the total uncertainty in the expected background. The lower panels show the ratios of observed data to the total expected background. In the lower panels, the light gray band represents the combined statistical and systematic uncertainty in the expected background, while the dark gray band represents the statistical uncertainty only. The rightmost bins include the overflow events.
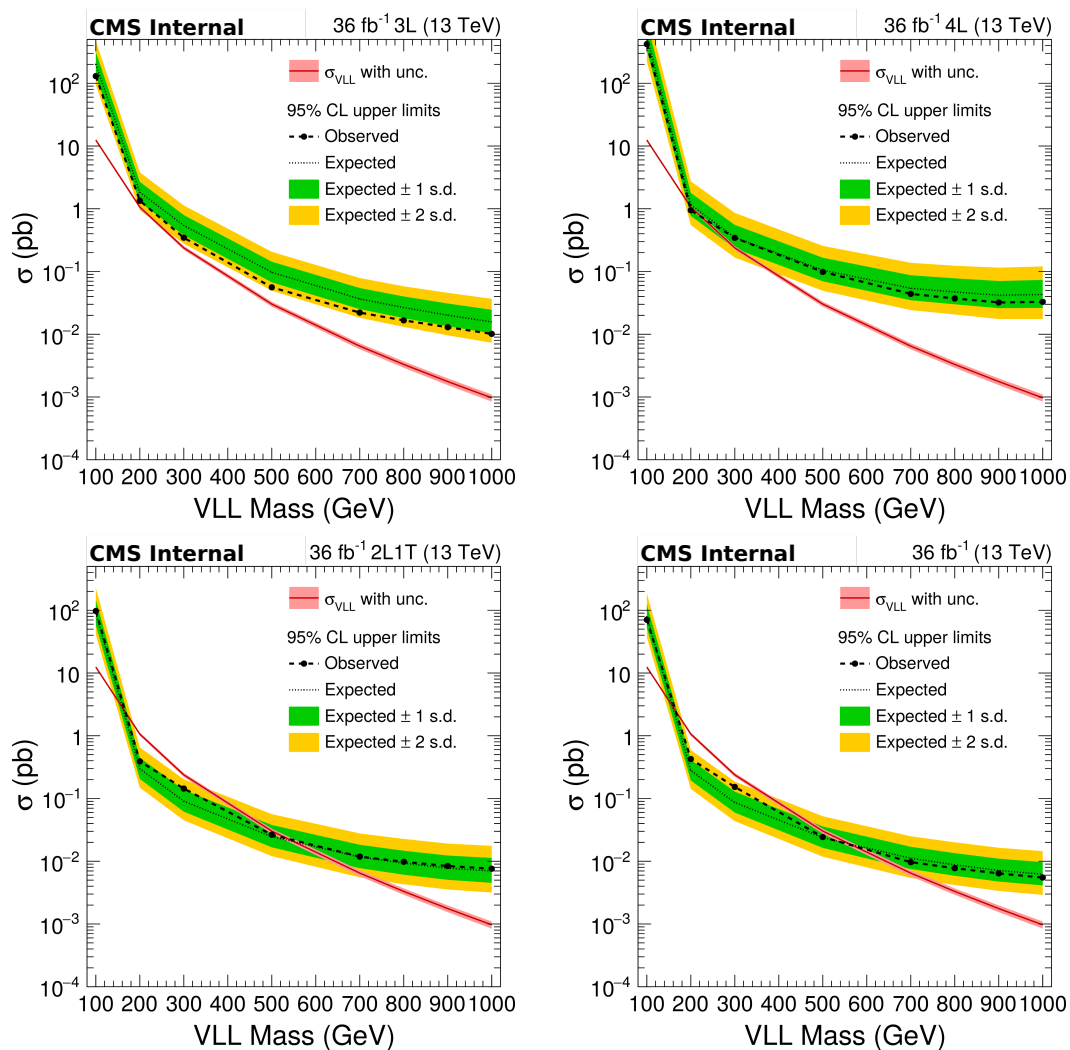
Figure 5.43: Observed limits for 3L, 4L, 2L1T and combined channels in 2016 analysis.
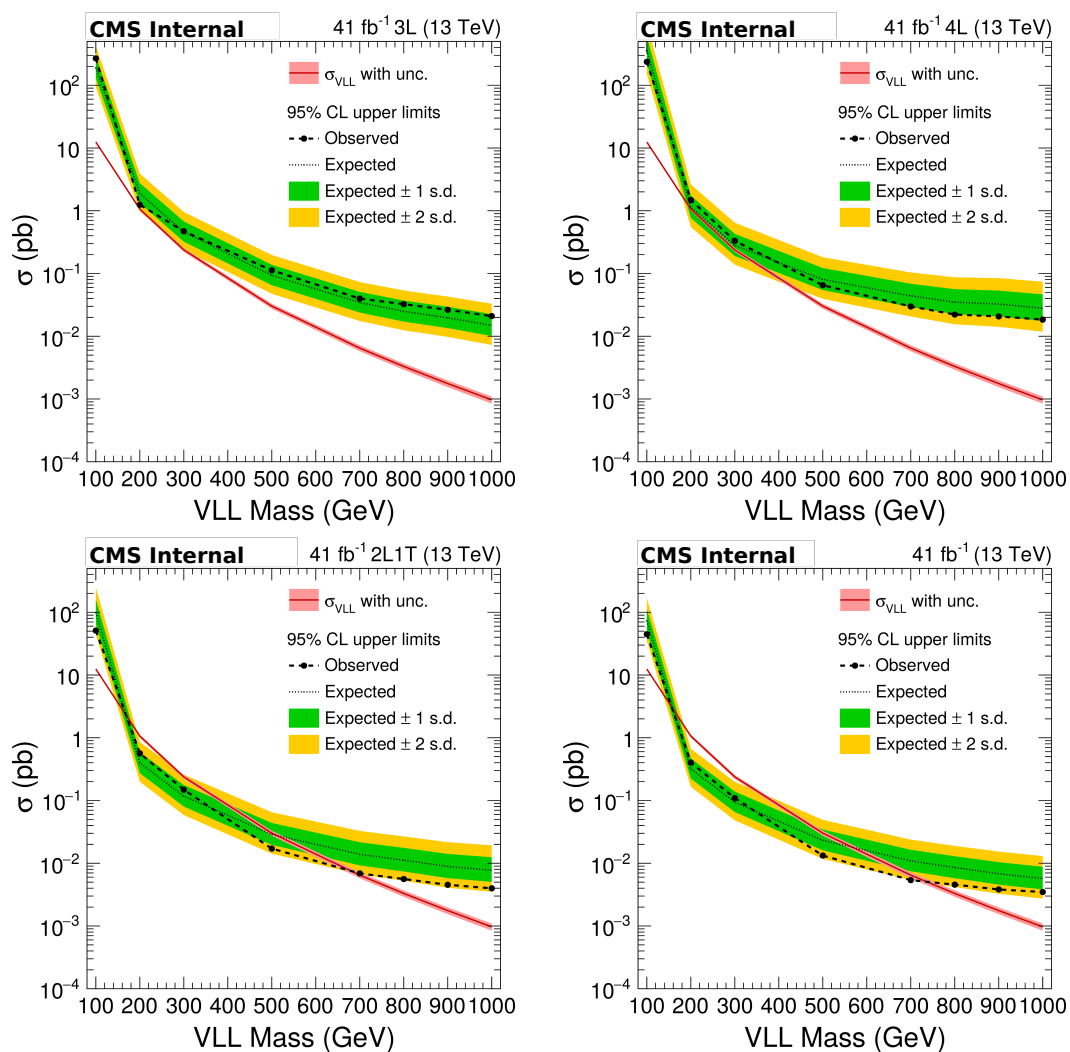
Figure 5.44: Observed limits for 3L, 4L, 2L1T and combined channels in 2017 analysis.
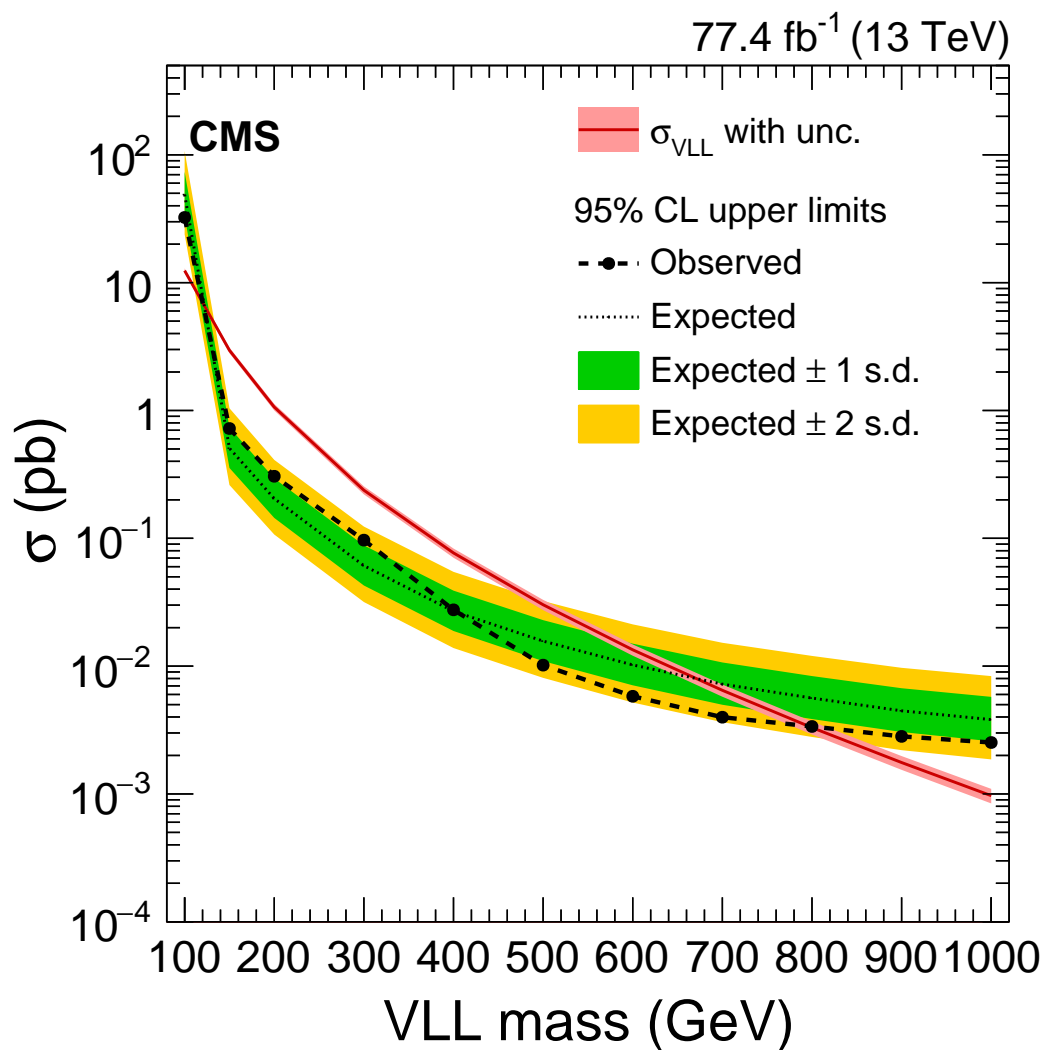
Figure 5.45: The 95% confidence level upper limits on the total cross section for associated ($\tau'^{\pm}\nu'_\tau$) and pair ($\tau'^{+}\tau'^{-}/\nu'_\tau\overline{\nu}'_\tau$) production of VLLs. The theoretical prediction for the production cross section of a vector-like lepton doublet coupling to the third generation SM leptons is also shown. The observed (expected) exclusion limit on the masses of VLLs is in the range of 120–790 (120–680)GeV.

CHAPTER 6

Event matrix assisted signal to background optimization for new physics searches

6.1 Introduction

The use of machine learning (ML) for searches of new physics at particle colliders is not a new idea. Many searches, including the 2012 discovery of the Higgs boson, have used ML to optimize the signal to background separation [72]. Conventional ML algorithms like BDT and artificial neural networks have been used extensively for searches of new physics [73–75].

In this chapter, a new method is introduced to distinguish new physics from the standard model background using convolutional neural networks as applied to the search for vector-like leptons. The results of the method have been demonstrated in the $2e/\mu(OS)+ \geq 1\tau_h$ signal region of the search for vector-like leptons. To evaluate the results, only simulated signal and background events have been used.

6.2 Convolutional neural networks

A deep learning neural network can learn to map a set of inputs to outputs from the training data. Convolutional neural networks (CNNs) are a variation of deep neural networks which are generally employed to analyze visual imagery [76]. Images can be thought of higher dimensional vectors which are generally described by the variation of light intensity along a grid of pixels. In simple words, a matrix of numbers, where each number represents light intensity at a specific x-y coordinate. See for example Fig. 6.1.

A CNN takes in various such matrices for both the signal of interest and the background that needs to be separated. Multiple 2-dimensional convolutions are then performed on the input matrices (one at a time). The first such list of convolutions forms the first layer of the network. The output from this layer is another matrix formed as a
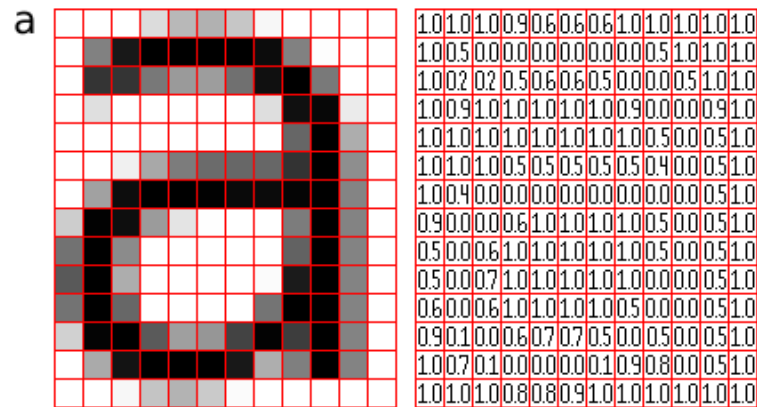
Figure 6.1: A 12x14 pixel image of the english alphabet **a** (left) along with it's normalized light intensity matrix (right).

result of various convolutions. This intermediate new matrix can then act an input for a new layer of convolutions. In the end, a multi-layer CNN gives out the probability of an initial input matrix to belong to either the signal class or the background class.

6.2.1 Design of a convolutional neural network

The core building block of a CNN is a convolutional layer. The exact mathematical description of the convolutional layers is tunable. Apart from this, CNN also has two more kinds of layers: a pooling layer, and fully connected layer.

The convolution layer conducts a dot product between two matrices, where one matrix is the set of learnable parameters otherwise recognized as a kernel, and the other matrix is a restricted portion of the input matrix. The kernel will slide across the height and width of the image, producing an image representation of that receptive region. This will create a two-dimensional representation of the image known as an activation map. The activation map then gives the response of the kernel at each spatial position of the image. The size of the slide of the kernel is called a stride. Fig. 6.2 shows an example of how this is done.

For the first convolutional layer, the input matrix is the one that the user provides. The other hidden convolutional layers take intermediate new matrices as inputs which are the outputs of previous convolutional layers.

The pooling layer will replace the output of the network at specific locations by a
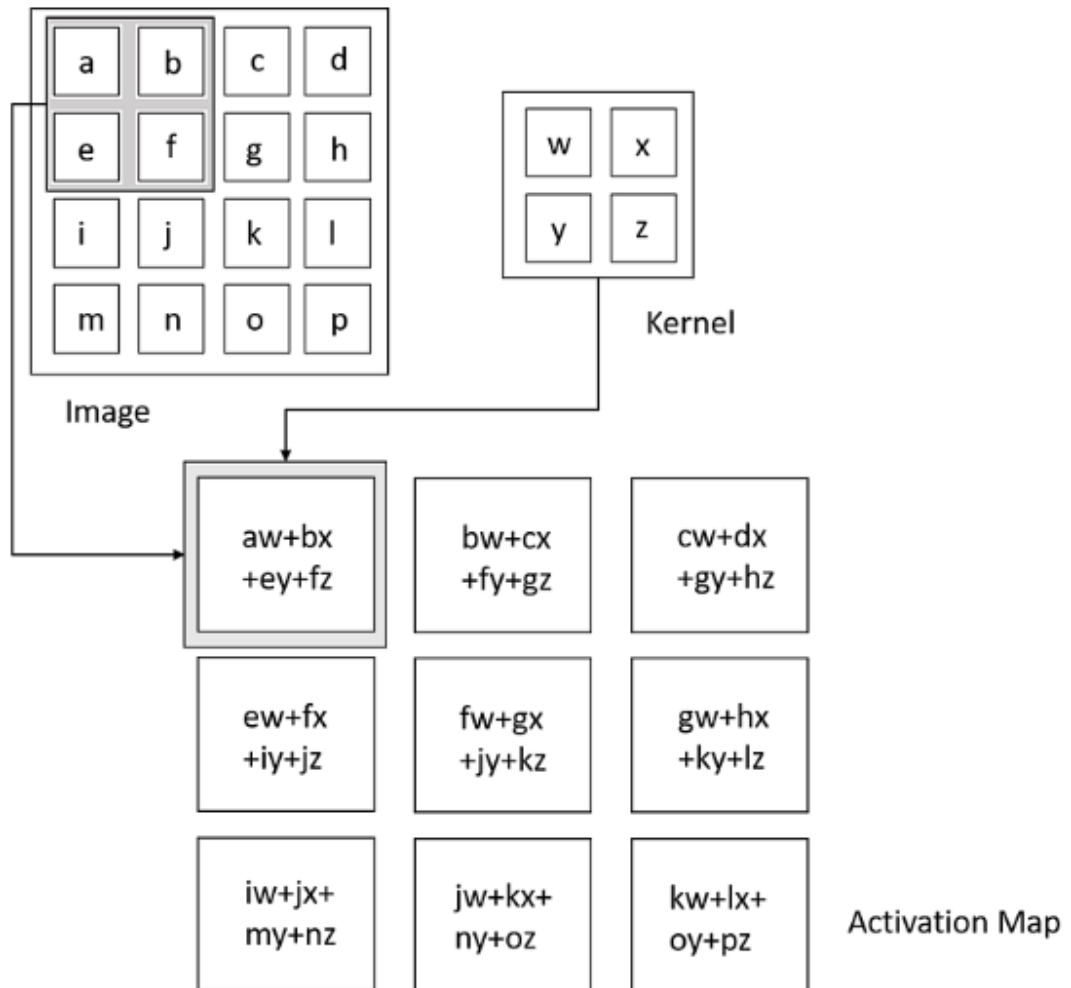
Figure 6.2: The process of convolution, where the dot product of a kernel is taken with sections of the input image, resulting in an activation map. One convolutional layer may contain multiple such kernels, in turn perform multiple such convolutions.
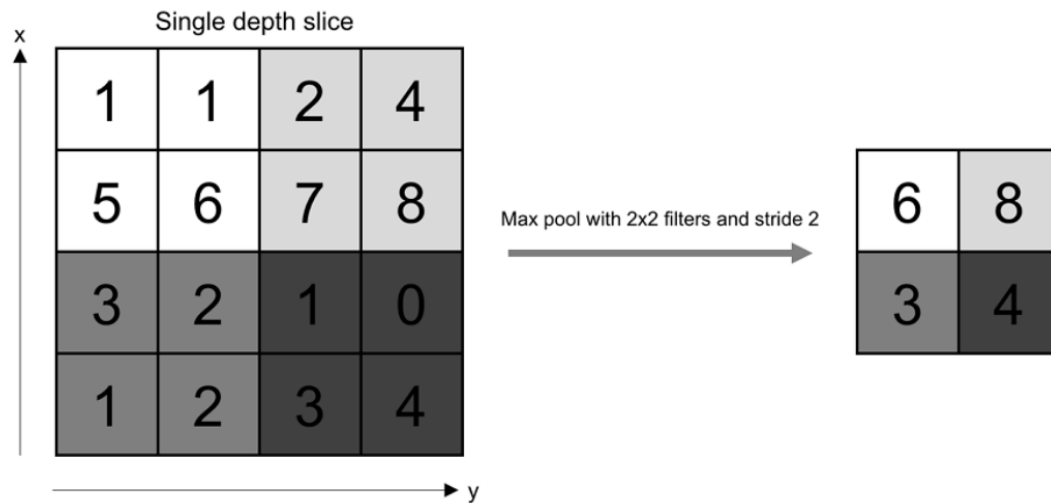
Figure 6.3: The operation of pooling.

summary representation of nearby outputs. This will help in reducing the spatial size of the representation, which decreases the required amount of computation and weights. The pooling operation is processed on every slice of the representation individually. Fig. 6.3 shows an example of how this is done.

To increase the non-linearity of the network without affecting input matrices, an activation layer is also used [77].

6.3 Terminology of a neural network

6.3.1 Training

The first phase is referred to as the training phase, where kernels (weights) are tuned to optimize signal-background separation. The underlying problem is an optimization problem, wherein a specific loss function is minimized when CNN (DNN) predicts the correct output. The output in our case is the probability of an input matrix belonging to either signal or background.

6.3.2 Testing

Once the kernels (weights) have been tuned and fixed, the CNN (DNN) is ready to be tested on an unseen independent dataset.

6.3.3 Overtraining

If a neural network is overtrained, it works very well on the training data but on an independent testing data. The main challenge of training a neural network is then to train it long enough to learn the mapping, but not so long that it overfits the training data. A good neural network works well on both the training and testing data.

6.3.4 Epochs

One epoch is when an entire dataset is passed forward and backward through the neural network once. Since I am limited by statistics, each time a full dataset is passed through the neural network, the loss function is further minimized and the network is better optimized to separate signal from background. If the testing and training data are properly sampled, such that each of them is statistically consistent, the factor that plays a major role in overtraining is the number of epochs.

6.3.5 Loss function

Neural networks are trained using stochastic gradient descent. It requires that you choose a loss function when designing and configuring your model. We may seek to maximize or minimize an objective function by searching for a solution that has the highest or lowest score, respectively. The function that is to be minimized or maximized is called the objective function or criterion. When minimizing, it may also be called the loss function. For this report, a binary cross-entropy loss function [78] has been used. It is better suited to act as loss function when we are trying to optimize a neural network to identify the class of the input when there are just two possible classes.

6.3.6 Architecture of neural networks

Neural networks have many design parameters, like the number of layers, the number of nodes per layers, number of kernels (for a CNN), etc., that can be chosen and then fixed for the problem of interest. For this report, I have worked with the neural networks described in Table 6.1.

CNN	4 convolution layers, each accompanied by a relu activation layer First (second) layer with 16 (32) kernels of size 3x3 with a stride of 1 Third (fourth) layer with 64 (128) kernels of size 2x2 with a stride of 1 Padding is same for all layers A max pooling layer of size 2x2 after the 4 convolution layers The output of max pooling is fed to fully connected layer with 1000 nodes 20% of the nodes are randomly dropped to avoid overtraining Softmax activation, returns the probability of initial input being signal or background
DNN	4 hidden layers, each accompanied by a relu activation layer Each hidden layer contains 100 nodes 20% of the nodes from the final layer are randomly dropped to avoid overtraining Softmax activation, returns the probability of initial input being signal or background
ANN	1 hidden layers, accompanied by a relu activation layer The hidden layer contains 100 nodes 20% of the nodes from the final layer are randomly dropped to avoid overtraining Softmax activation, returns the probability of initial input being signal or background

Table 6.1: Three different neural network architectures used in this report

6.4   Event matrix based search for vector-like leptons

This report presents a CNN based observable as an alternative to L_T for distinguishing the signal from background for the VLL search as described in earlier chapters of this thesis. It's applicability in the $2e/\mu(OS)+ \geq 1\tau_h$ region has been demonstrated. The following is used as the design of the input matrix to the CNN:

$$
\begin{bmatrix}
p_{T(l1)} & |\Delta\phi_{(l1-l2)}| & M_{T(l1-p_T^{miss})} & M_{l1-l2} \\
p_{T(l2)} & |\Delta\phi_{(l2-\tau_h 1)}| & L_T & p_{T(l1-l2)} \\
p_{T(\tau 1)} & |\Delta\phi_{(l1-\tau_h 1)}| & dR_{l1-l2} & M_{T(\tau-p_T^{miss})} \\
p_T^{miss} & |\Delta\phi_{(p_T^{miss}-\tau_h 1)}| & N_{tracks} & HT
\end{bmatrix}
\tag{6.1}
$$

Each simulated event is then described by a matrix, referred to as the event matrix. The classification task then is to recognize patterns between signal event matrices from background event matrices.

6.4.1 Designing the event matrix

The design of the event matrix is based on contrast rows/columns. Rows/columns are designed such that contrast between cells can give information about the kinematics (an inturn the physics) of the event. The first column will capture patterns between transverse momenta of the three leptons in the event and their scale difference from the p_T^{miss} in the event. The second column will capture contrast differences in $\Delta\phi$ between various objects. The left bottom cells give an essence of hadronic activity in the event and also the relative contrast between the number of charged particles and total jet momenta scale. The other cells carry information about total mass and momenta scale. Once cell also carries information on the ΔR between the two light leptons in the $\eta - \phi$ space.

6.4.2 Scaled event matrix

A CNN is conventionally used for image classification purposes. The whole design of a CNN (kernels, convolution, max pooling, etc.) is based on the idea of learning features in visual imagery. Hence we convert our event matrices to a format that is suited to be handled by the CNN. For doing this, I normalize elements of our event matrices such that each cell is then a grayscale integer intensity from 0 to 256. For this, I scale each element using this equation

$$m \mapsto |(\frac{m - r_{\min}}{r_{\max} - r_{\min}} \times 256) - 256|$$
(6.2)

where m maps to [0,256], such that the maximum m has the lowest intensity and the least m has the highest intensity. This is done so that a white pixel in the image means a very high intensity and black means low intensity. The image is then a grayscale image with each pixel having an intensity from 0 to 256. The resultant matrix is then a scaled event matrix (SEM). The $r_{\min}$ and $r_{\max}$ for any element are hardcoded. Table 6.2 provides these hardcoded values for each element of the event matrix. Figs. 6.4 and 6.5 show some SEMs for a few random signal and background events.

Variable	minimum	maximum	Variable	minimum	maximum
$p_{T(l1)}$	20 GeV	300 GeV	$M_{T(l1-p_T^{miss})}$	0	300
$p_{T(l2)}$	20 GeV	150 GeV	L_T	0	900
$p_{T(\tau_h 1)}$	20 GeV	160 GeV	$dR_{(l1-l2)}$	0.4	5
p_T^{miss}	50 GeV	150 GeV	N_{tracks}	0	1000
	150 GeV	350 GeV			
$|\Delta\phi_{(l1-l2)}|$	0	6	$M_{(l1-l2)}$	0	300
$|\Delta\phi_{(l2-\tau_h 1)}|$	0	6	$p_{(l1-l2)}$	0	300
$|\Delta\phi_{(l1-\tau_h 1)}|$	0	6	$M_{T(\tau_h-p_T^{miss})}$	0	250
$|\Delta\phi_{(p_T^{miss}-\tau_h 1)}|$	0	6	HT	0	800

Table 6.2: Hard coded minimum and maximum values to be used in designing the scaled event matrix. The numbers are chosen so as to cover most of the phase space. Any value above the maximum is truncated at the maximum.

$$\begin{bmatrix} 194 & 27 & 172 & 203 \\ 207 & 117 & 210 & 153 \\ 238 & 166 & 222 & 149 \\ 159 & 197 & 161 & 105 \end{bmatrix} \begin{bmatrix} 233 & 176 & 126 & 195 \\ 212 & 225 & 220 & 212 \\ 220 & 207 & 173 & 191 \\ 198 & 225 & 72 & 167 \end{bmatrix} \begin{bmatrix} 143 & 252 & 8 & 195 \\ 189 & 119 & 161 & 87 \\ 44 & 123 & 234 & 199 \\ 252 & 239 & 38 & 233 \end{bmatrix}$$

Figure 6.4: Three randomly picked SEM's from 500 GeV VLL production

$$\begin{bmatrix} 162 & 20 & 16 & 151 \\ 108 & 161 & 170 & 83 \\ 135 & 115 & 215 & 217 \\ 241 & 242 & 200 & 231 \end{bmatrix} \begin{bmatrix} 146 & 230 & 0 & 211 \\ 210 & 118 & 191 & 105 \\ 208 & 145 & 240 & 208 \\ 226 & 234 & 31 & 245 \end{bmatrix} \begin{bmatrix} 101 & 237 & 0 & 171 \\ 152 & 110 & 174 & 37 \\ 243 & 129 & 231 & 252 \\ 221 & 253 & 125 & 188 \end{bmatrix}$$

Figure 6.5: Three randomly picked SEM's from SM production.

6.4.3   Proportionality Weights

For this report, each kind of neural network is trained to identify patterns between a specific mass point of VLL signal and SM background as a whole. This means our network is blind to whether a background event is coming from WZ or ZZ SM production. It only cares whether an event belongs to signal of interest or the SM background. But since technically events come from independent samples for each SM process like WZ, ZZ, ttZ, etc, input events need to be appropriately weighted such that for each phase space, the least important background has the lowest weights per event and vice versa. These are the same weights that scale the total events from an independent simulation sample to required luminosity. This way even though the total number of events from a particular simulation sample might be much larger than the other, the proportionality weights make sure that those events are counted proportional to their importance. These proportionality weights are then factored into the loss function, such that contribution of minimizing the loss from a more important event is more than a less important event.

6.4.4   Oversampling

Since the raw number of signal events are generally much lower than the total background events available, I oversample the signal events to be equivalent to total background events available. To avoid this resulting in an overtraining, proportionality weights of each signal event are correspondingly reduced so that each signal event now contributes lesser to the minimization of the loss, such that minimum of the loss function remains the same.

6.4.5   Creating the training and testing datasets

The total events (signal and background) are randomly divided into four datasets with equal proportions (25% events each). One such mixture of signal and background forms the training dataset. The other three form independent testing datasets (testing dataset 1,2,3). These are used to check for any overtraining that might have happened. It is to note that since 25% of events have been used for training, only the rest 75% is actually available for any further study of the final results from the neural network.

6.5 Results

This section describes the performance of SEM based signal to background optimization for signal regions with $2e/\mu(OS)+ \geq 1\tau_h$ in low $p_T^{miss}(50 < p_T^{miss} < 150)$ and high $p_T^{miss}(p_T^{miss} > 150)$ for 200 and 500 GeV VLL signal mass hypothesis.

6.5.1 Training and testing the CNN in low p_T^{miss} region

The following section contains results from training and testing of the neural network in low p_T^{miss} signal regions for two signal hypothesis, namely a 200 GeV VLL hypothesis and 500 GeV VLL hypothesis.

200 GeV VLL vs SM background

Fig. 6.6 shows the L_T distribution in various signal and background testing samples and the result of training and testing of a CNN with the architecture as described in Table 6.1, for a 200 GeV VLL signal hypothesis in low p_T^{miss} region. Fig. 6.7 shows the distribution of output classifier for the three independent testing datasets of signal and background.

Figure 6.6: The L_T distribution for three different independent test samples (left) and output classifier distribution from the CNN for the training and testing sample, shown for 200 GeV VLL signal and SM background.

Figure 6.7: The output classifier distribution from the CNN for three independent testing and the training sample, shown for 200 GeV VLL signal (left) and SM background (right).

500 GeV VLL vs SM background

Fig. 6.8 shows the L_T distribution in various signal and background testing samples and the result of training and testing of a CNN with the architecture as described in Table 6.1, for a 500 GeV VLL signal hypothesis in low p_T^{miss} region. Fig. 6.9 shows the distribution of output classifier for the three independent testing datasets of signal and background.
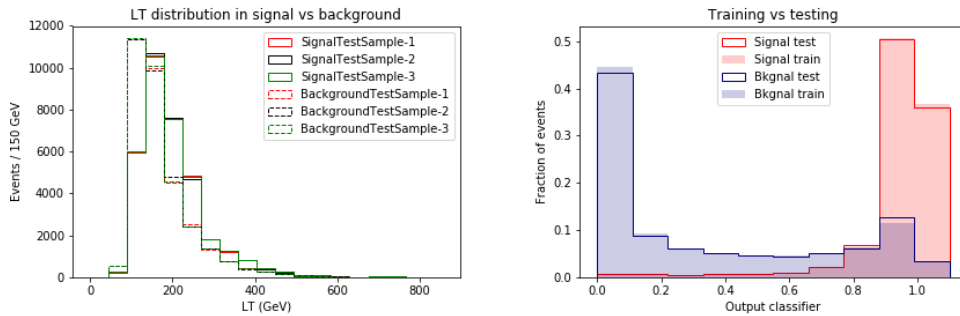

Figure 6.8: The L_T distribution for three different independent test samples (left) and output classifier distribution from the CNN for the training and testing sample, shown for 500 GeV VLL signal and SM background.

Figure 6.9: The output classifier distribution from the CNN for three independent testing and the training sample, shown for 500 GeV VLL signal (left) and SM background (right).

6.5.2 Training and testing the CNN in high p_T^{miss} region

The following section contains results from training and testing of the neural network in high p_T^{miss} signal regions for two signal hypothesis, namely a 200 GeV VLL hypothesis and 500 GeV VLL hypothesis.

200 GeV VLL vs SM background

Fig. 6.10 shows the L_T distribution in various signal and background testing samples and the result of training and testing of a CNN with the architecture as described in Table 6.1, for a 200 GeV VLL signal hypothesis in high p_T^{miss} region. Fig. 6.11 shows the distribution of output classifier for the three independent testing datasets of signal and background.

Figure 6.10: The L_T distribution for three different independent test samples (left) and output classifier distribution from the CNN for the training and testing sample, shown for 200 GeV VLL signal and SM background.
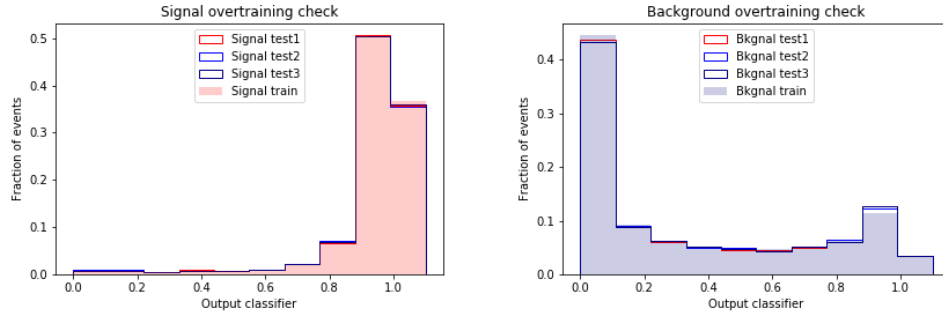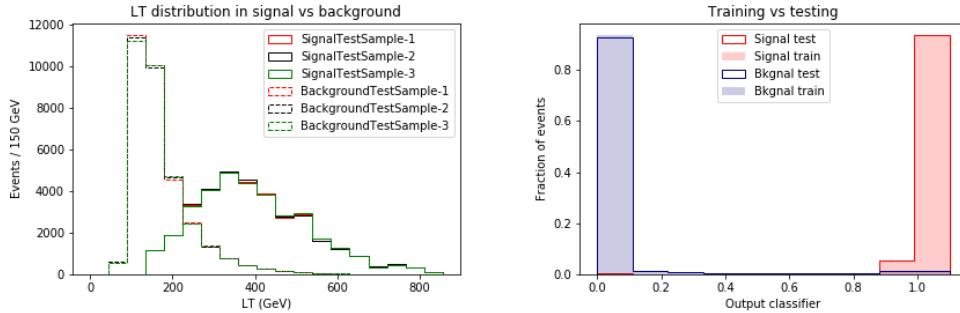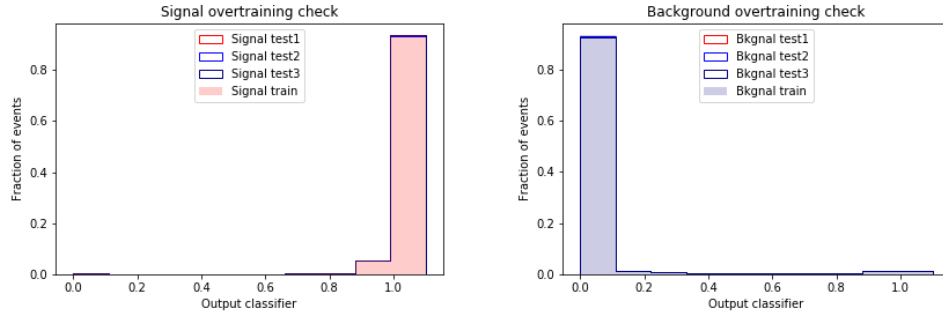
Figure 6.11: The output classifier distribution from the CNN for three independent testing and the training sample, shown for 200 GeV VLL signal (left) and SM background (right).

500 GeV VLL vs SM background

Fig. 6.12 shows the L_T distribution in various signal and background testing samples and the result of training and testing of a CNN with the architecture as described in Table 6.1, for a 500 GeV VLL signal hypothesis in high p_T^{miss} region. Fig. 6.13 shows the distribution of output classifier for the three independent testing datasets of signal and background.
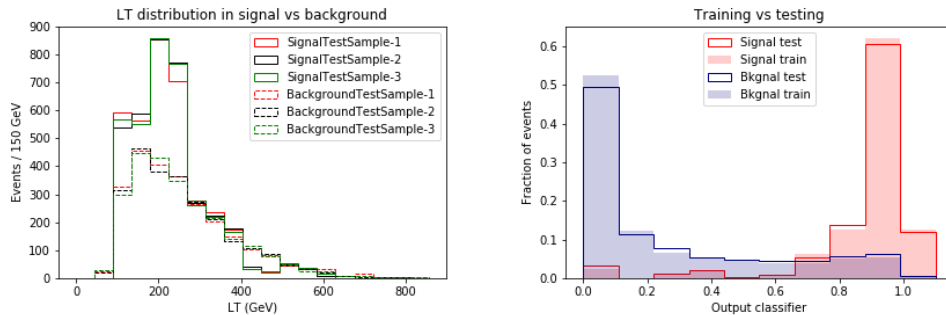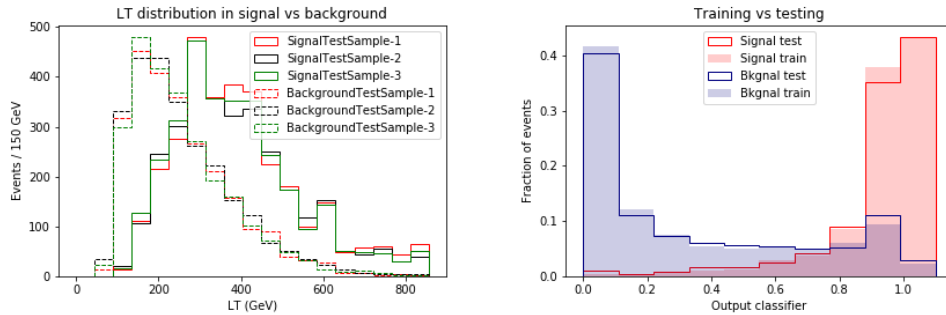

Figure 6.12: The L_T distribution for three different independent test samples (left) and output classifier distribution from the CNN for the training and testing sample, shown for 500 GeV VLL signal and SM background.
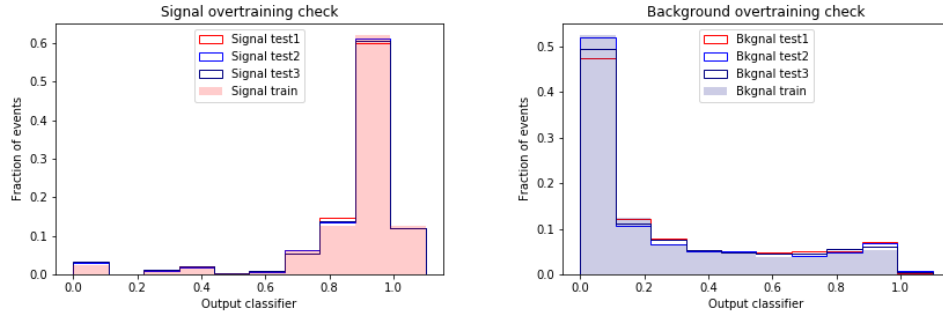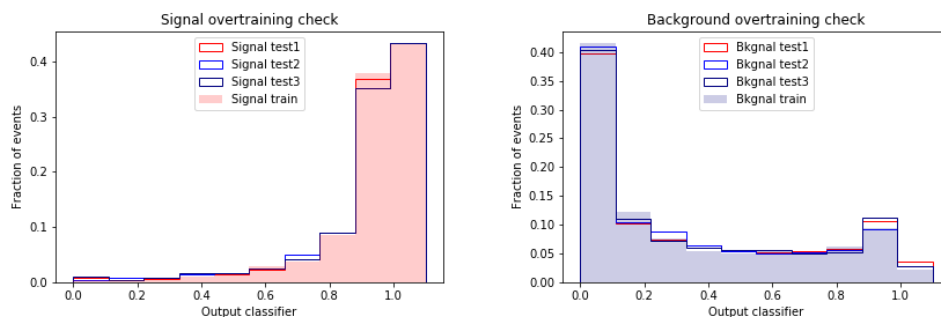
Figure 6.13: The output classifier distribution from the CNN for three independent testing and the training sample, shown for 500 GeV VLL signal (left) and SM background (right).

Performance of CNN vs. DNN, ANN and L_T

This report uses ROC curves to demonstrate the performance of neural networks and L_T for separating signal and background. Before looking into these curves, one needs to be familiar with the terms described in Table 6.3.

	Is a signal event	Is a background event
Algorithm says it is a signal event	True positive	False Postive
Algorithm says it is a background event	False negative	True negative

Table 6.3: Terms used in a ROC curve

Fig. 6.14 show the performance of the CNN, DNN, ANN with the architecture as described in Table 6.1 along with the performance of L_T, in low and high p_T^{miss} signal regions for the 200 GeV VLL hypothesis and 500 GeV VLL hypothesis.
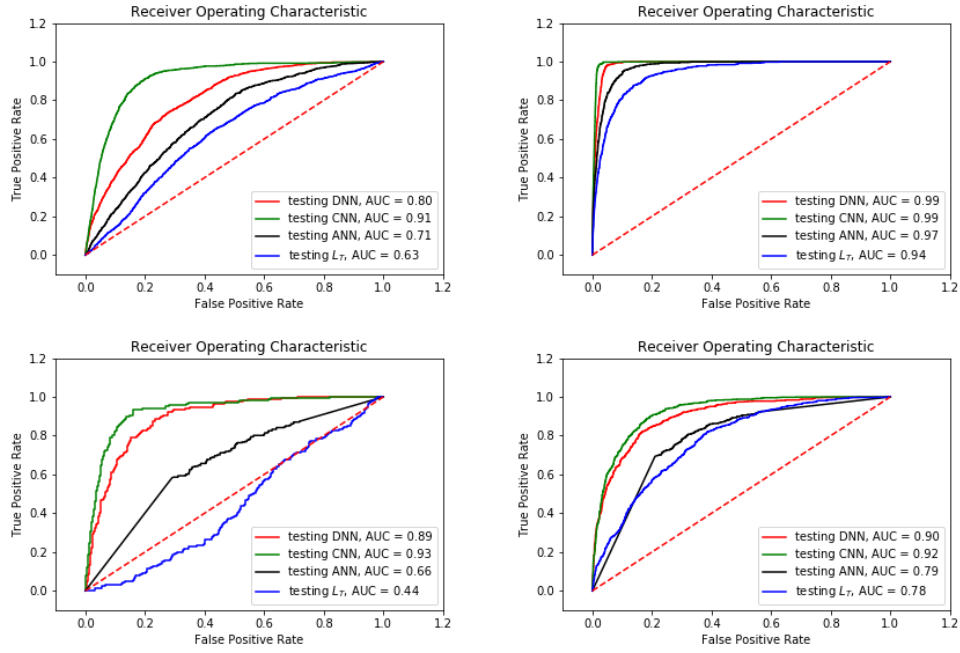
Figure 6.14: ROC curves for all four signal to background optimization regions. Each figure contains ROC curves for testing the performance of L_T, an artificial neural network, a deep neural network and an SEM based convolutional neural network. The low p_T^{miss} 200 GeV (500 GeV) VLL region is shown top-left (top-right). The high p_T^{miss} 200 GeV (500 GeV) VLL region is shown bottom-left (bottom-right).

6.5.3   Additional studies on the event matrix description

All the studies above were done using SEMs arranged in the form as shown in matrix 6.1. It is thus interesting to study what happens when the form of the matrix is changed. For this a study was carried out in low p_T^{miss} region for VLL 200 GeV vs SM background by changing the form of the matrix. In one case, the positions of elements in the first two rows were randomly scrambled to arrive at the new form. In the second case, all the element positions were randomly scrambled. The performance of these new forms compared to the original form is shown is Fig. 6.15. The studies suggest that scrambling the elements of the first two rows has minimal effect on performance, while scrambling all the rows significantly affects the performance of the SEM based method.

92
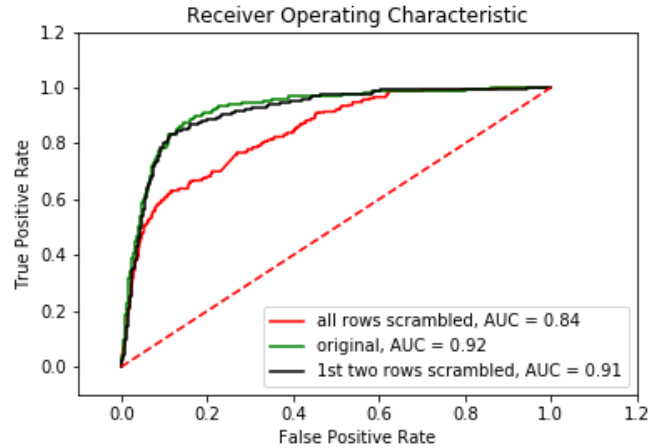
Figure 6.15: ROC curves for performance of the SEM based method for two different matrix forms along with original form is shown in the low p_T^{miss} region for VLL 200 GeV vs SM background.

6.6   Conclusion

Performance of a scaled event matrix based convolutional neural network (CNN) was compared to an artificial neural network (ANN), a deep neural network (DNN) and the L_T based cut and count strategy. In all signal regions that are investigated, for both the 200 GeV VLL hypothesis and 500 GeV VLL hypothesis, a CNN beats both a DNN and an ANN with their architectures described in Table 6.1. In most cases, any of the neural networks beats L_T as a signal to background discriminating variable. It is also seen that the form the input matrix can have a effect on the performance of the CNN.

CHAPTER 7

Summary

This thesis presents a search for vector-like leptons coupled to the third generation standard model leptons has been performed in several multilepton final states using $77\ \text{fb}^{-1}$ of proton-proton collision data at a center-of-mass energy of 13 TeV, collected by the CMS experiment in 2016 and 2017. We use the scalar p_T sum of the leptons (denoted as L_T) to discriminate signal from SM backgrounds in all channels. The L_T distribution is divided into 150 GeV bins, each of which is treated as a separate experiment. Analyses of the 2016 and 2017 data are done separately, and results from both the analyses are individually looked at. This is the first search for a vector-like lepton doublet in any of the LHC experiments. No significant deviations of the data from the standard model predictions are observed.

For further analysis of the observations, two hypotheses are considered: background only hypothesis and the signal plus background hypothesis. meaning, the observed events are composed of only background and background along with signal, respectively. The exclusion of one hypothesis from the other implies the observations agree better with the other hypothesis. To obtain signal cross section upper limits at 95% confidence level (CL) a modified frequentist approach is used with a test statistic based on the profile likelihood in the asymptotic approximation and the CLs criterion [69–71]. Limits are set on the combined cross section for associated ($\tau'\nu'_\tau$) and pair ($\tau'\tau'/\nu'_\tau\nu'_\tau$) production of VLLs. On combining all the signal regions, with the hypothesis of an SU(2) mass degenerate VLL doublet with couplings to the third generation SM leptons, VLLs with mass in the range of 120–790 GeV are excluded at 95% CL. These are the best limits yet on the production of a vector-like lepton doublet, coupling to the third generation standard model leptons.

A machine learning extension to this search is also presented that shows the possibility of significant enhancement in the sensitivity of the search. This was a simulation

based study. Future analyses can employ this extension and demonstrate the increase in sensitivity using data, and in turn increase LHC's sensitivity to searching for vector-like leptons. The premise of this new technique is general, and while the performance is only shown in the context of the search for vector-like leptons, this can also be applied to other searches for new physics at the LHC.

APPENDIX A

A.1 Additional checks on the misidentified tau lepton background yield

The fake rates for PGt_h are measured in 2L1T OS, $p_T^{miss} < 50$ GeV, on-Z region. Figure A.1 and A.6 shows the invariant mass distribution of light leptons in the events 2L1T OS and SS, $p_T^{miss} < 50$ GeV regions for 2016 and 2017 analyses. A closure test is done in data in the measurement region (OS on-Z) as given in left plot of Figure A.4. As an independent test of the background methods we test our predictions in OS, $p_T^{miss} < 50$ GeV regions with on-Z requirement reversed (off-Z) as given in right plot of Figure A.4. Overall, Figure A.1 to Figure A.8 all control and validation regions for τ_h fake background estimation. The 2016+2017 combination plots are shown in Figure 5.29. Since the 2L1T OS/SS, $p_T^{miss} < 50$ GeV regions are used to measure and validate our fake rates, these are excluded from the signal regions.
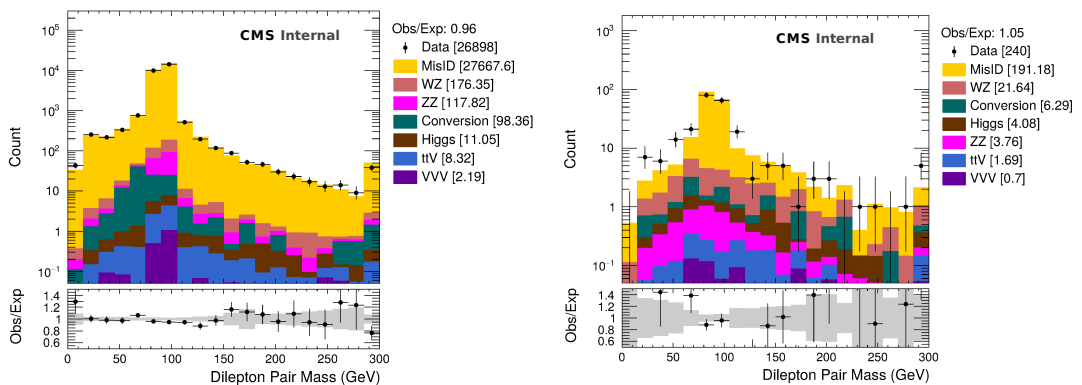
Figure A.1: Dilepton Mass distribution in 2L1T (OS and SS), $p_T^{miss} < 50$ GeV region in 2016 data.

96

Figure A.2: L_T distribution in 2L1T (OS and SS), $p_T^{miss} < 50$ GeV region in 2016 data.
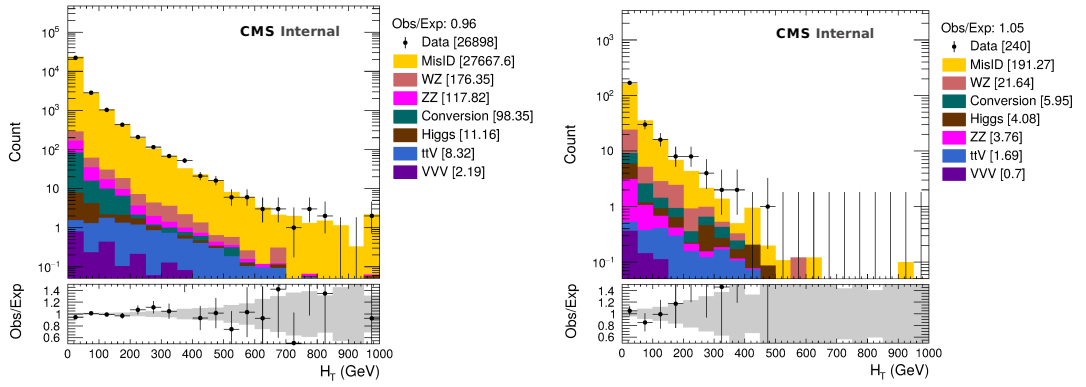
Figure A.3: H_T distribution in 2L1T (OS and SS), $p_T^{miss} < 50$ GeV region in 2016 data.
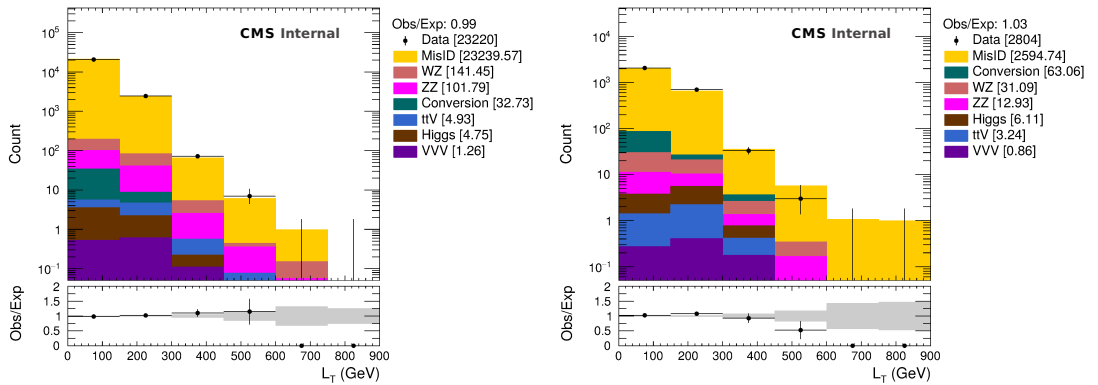
Figure A.4: L_T distribution in 2L1T (OnZ and OffZ), $p_T^{miss} < 50$ GeV region in 2016 data.

Figure A.5: L_T distribution in 2L1T (OnZ and OffZ), $p_T^{miss} < 50$ GeV region in 2017 data.
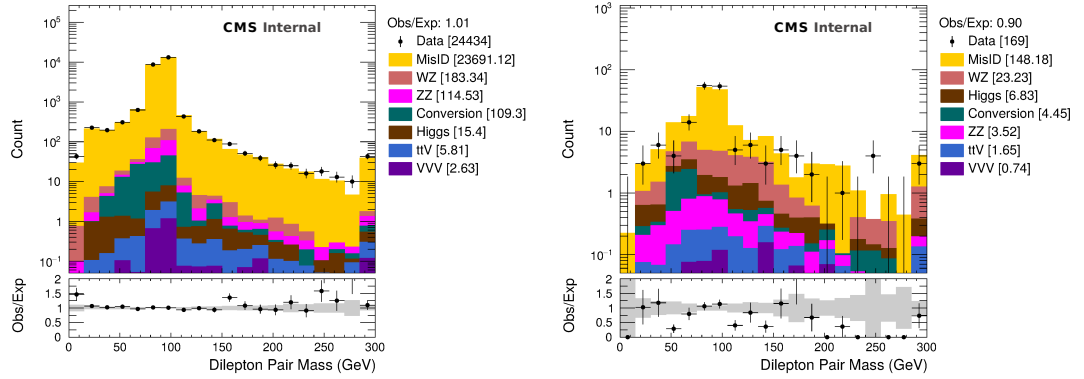
Figure A.6: Dilepton Mass distribution in 2L1T (OS and SS), $p_T^{miss} < 50$ GeV region in 2017 data.
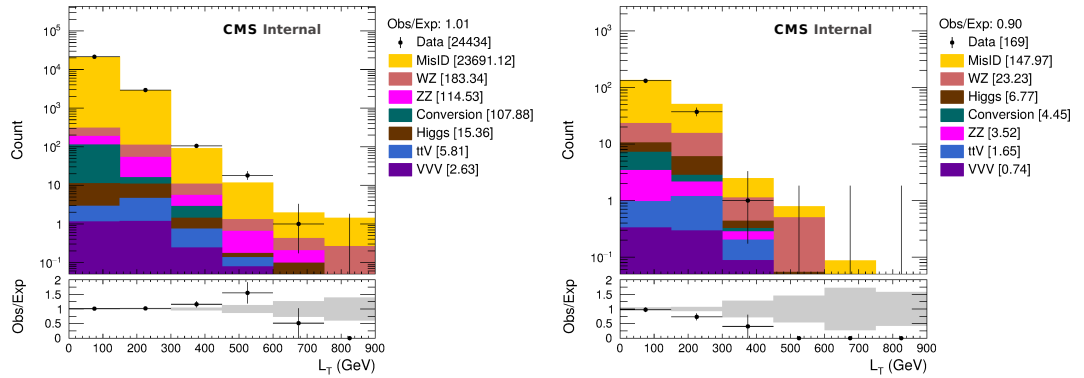
Figure A.7: L_T distribution in 2L1T (OS and SS), $p_T^{miss} < 50$ GeV region in 2017 data.
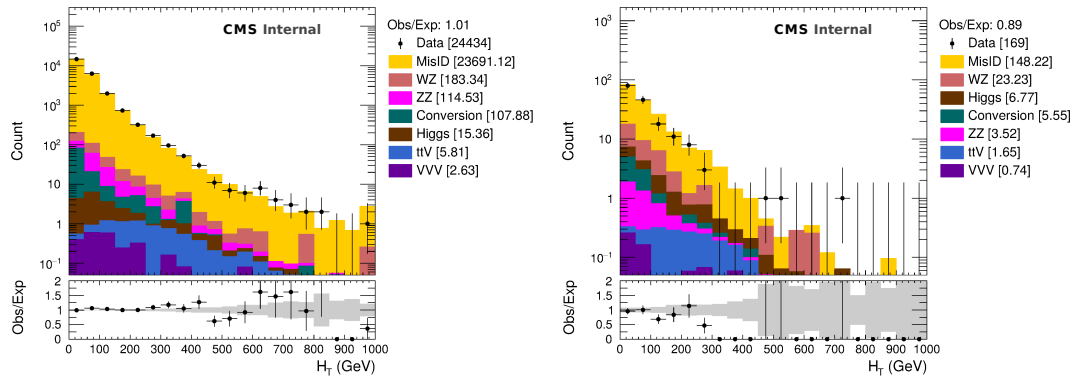
Figure A.8: H_T distribution in 2L1T (OS and SS), $p_T^{miss} < 50$ GeV region in 2017 data.

A.2 Single Lepton Trigger Efficiency

Measurements are done using Tag and Probe method in which dilepton events are selected with a single lepton trigger. Tag object is required to pass our tight object definition and match to an HLT level object. Probe object also passes the tight object definition, is opposite in charge with the Tag and lies outside the cone of radius $\Delta R = 0.4$ centered around the Tag object. Probe matching to another HLT level object gives the desired trigger efficiency. For matching purposes, $\Delta R < 0.2$ is used.
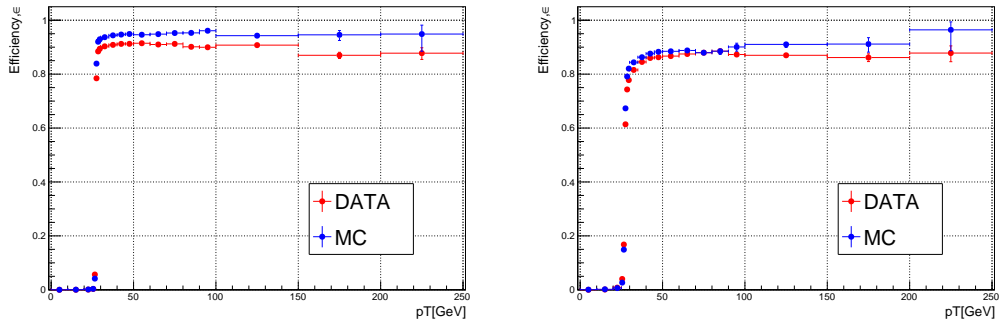
Figure A.9: Trigger efficiency in Data and MC as a function of p_T of tight muon objects in two broad η bins : barrel which is for $|\eta| \leq 1.2$ (left) and endcap which is for $|\eta| > 1.2$ (right). These are obtained in dimuon events collected by HLT_IsoMu27 trigger path.

Figure A.10: Trigger efficiency in Data and MC as a function p_T of tight electron objects in two broad η bins : barrel which is for $|\eta| \leq 1.479$ (left) and endcap which is for $|\eta| > 1.479$ (right). These are obtained in dielectron events collected by HLT_Ele35_WPTight_Gsf trigger path.

100

A.3   Signal region plots in 2016 and 2017

Figure A.11: The L_T distribution for events with three light leptons with $p_T^{miss} < 150$ GeV (left) and with $p_T^{miss} > 150$ GeV (right) for 2016 data and MC. The to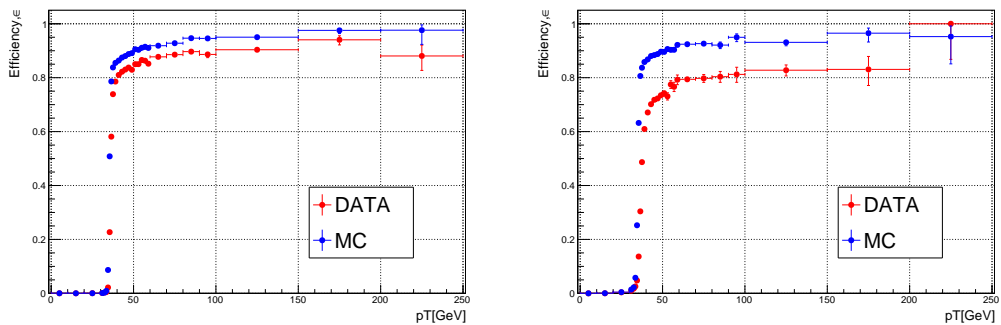tal SM background is shown as a stack of all contributing processes. The predictions for signal models with $m_{\tau/\nu} = 200$ GeV and $m_{\tau/\nu} = 500$ GeV (sum of all production and decay modes) are also shown as solid lines. The hatched gray band in the upper panel and the dark and light gray bands in the lower panel represent the total, statistical, and systematic uncertainties on the expected background, respectively.

Figure A.12: The L_T distribution for events with four or more light leptons with $p_T^{miss} < 50$ GeV (left) and with $p_T^{miss} > 50$ GeV (right) for 2016 data and MC. The total SM background is shown as a stack of all contributing processes. The predictions for signal models with $m_{\tau/\nu} = 200$ GeV and $m_{\tau/\nu} = 500$ GeV (sum of all production and decay modes) are also shown as solid lines. The hatched gray band in the upper panel and the dark and light gray bands in the lower panel represent the total, statistical, and systematic uncertainties on the expected background, respectively.

101

Figure A.13: The L_T distribution for events with two opposite charge light leptons and a hadronically decaying τ lepton with $50 < p_T^{miss} < 150$ GeV (left) and with $p_T^{miss} > 150$ GeV (right) for 2016 data and MC. The total SM background is shown as a stack of all contributing processes. The predictions for signal models with $m_{\tau/\nu} = 200$ GeV and $m_{\tau/\nu} = 500$ GeV (sum of all production and decay modes) are also shown as solid lines. The hatched gray band in the upper panel, and the dark and light gray bands in the lower panel represent the total, statistical, and systematic uncertainties on the expected background, respectively.

Figure A.14: The L_T distribution for events with two same charge light leptons and a hadronically decaying τ lepton with $50 < p_T^{miss} < 150$ GeV (left) and with $p_T^{miss} > 150$ GeV (right) for 2016 data and MC. The total SM background is shown as a stack of all contributing processes. The predictions for signal models with $m_{\tau/\nu} = 200$ GeV and $m_{\tau/\nu} = 500$ GeV (sum of all production and decay modes) are also shown as solid lines. The hatched gray band in the upper panel and the dark and light gray bands in the lower panel represent the total, statistical, and systematic uncertainties on the expected background, respectively.
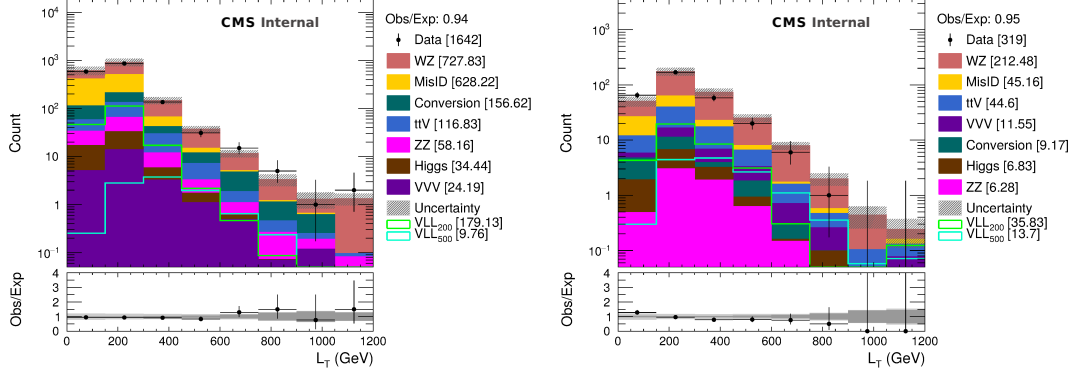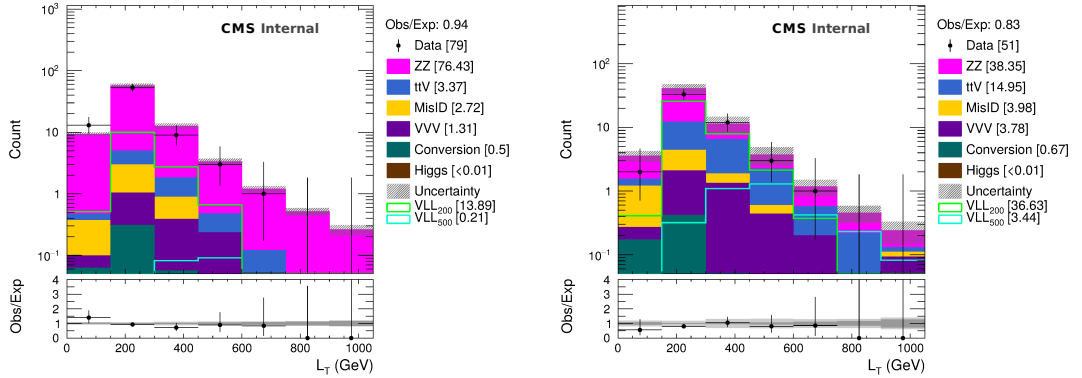
Figure A.15: The L_T distribution for events with three light leptons with $p_T^{miss} < 150$ GeV (left) and with $p_T^{miss} > 150$ GeV (right) for 2017 data and MC. The total SM background is shown as a stack of all contributing processes. The predictions for signal models with $m_{\tau/\nu} = 200$ GeV and $m_{\tau/\nu} = 500$ GeV (sum of all production and decay modes) are also shown as solid lines. The hatched gray band in the upper panel and the dark and light gray bands in the lower panel represent the total, statistical, and systematic uncertainties on the expected background, respectively.
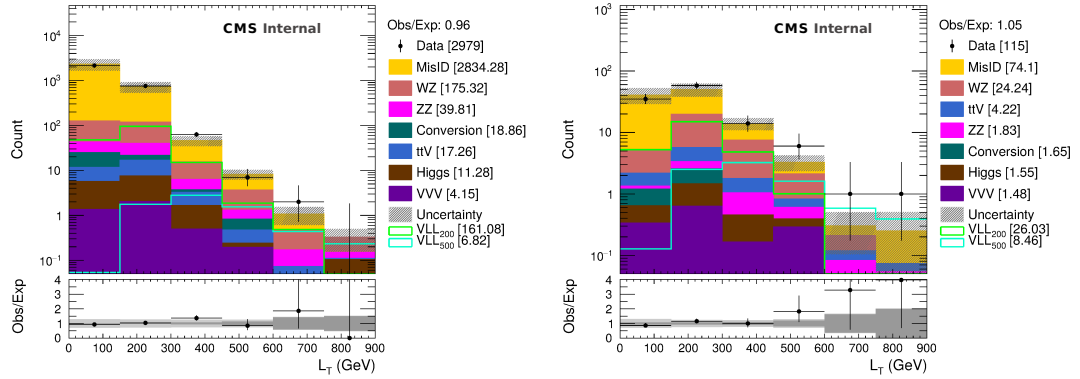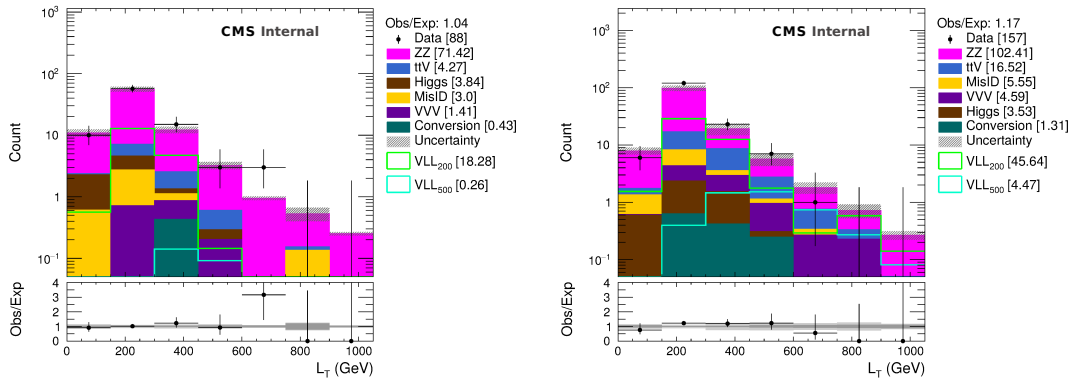

Figure A.16: The L_T distribution for events with four or more light leptons with $p_T^{miss} < 50$ GeV (left) and with $p_T^{miss} > 50$ GeV (right) for 2017 data and MC. The total SM background is shown as a stack of all contributing processes. The predictions for signal models with $m_{\tau/\nu} = 200$ GeV and $m_{\tau/\nu} = 500$ GeV (sum of all production and decay modes) are also shown as solid lines. The hatched gray band in the upper panel and the dark and light gray bands in the lower panel represent the total, statistical, and systematic uncertainties on the expected background, respectively.
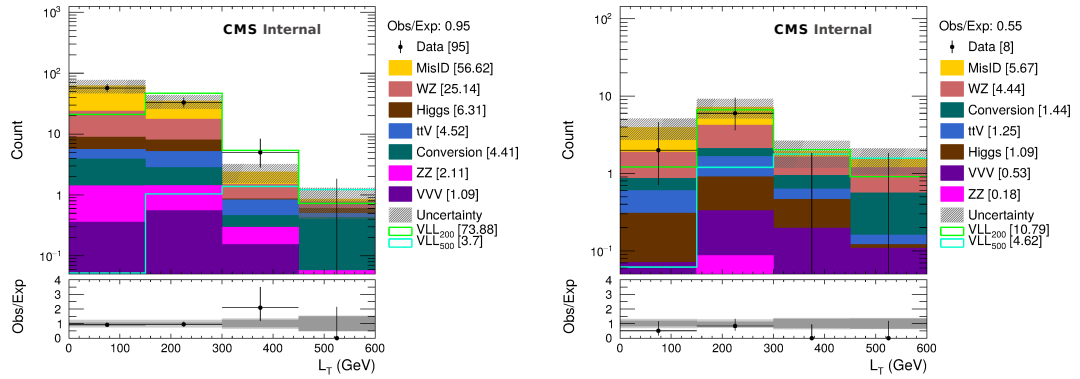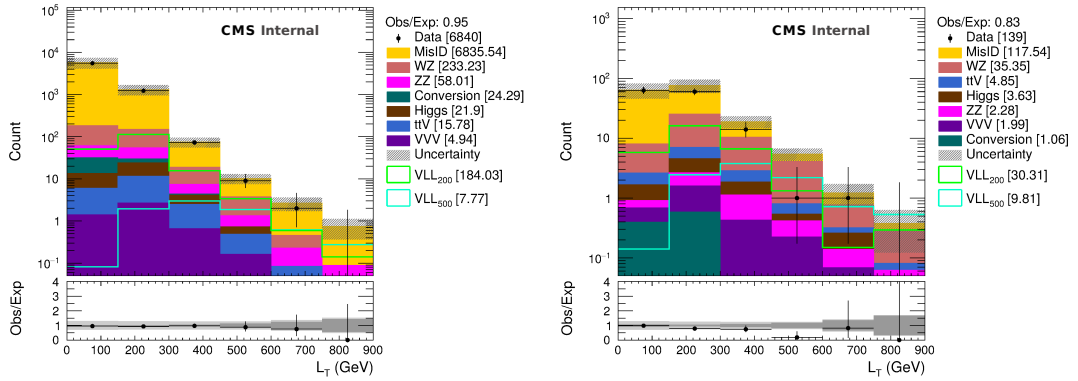
Figure A.17: The L_T distribution for events with two opposite charge light leptons and a hadronically decaying τ lepton with $50 < p_T^{miss} < 150$ GeV (left) and with $p_T^{miss} > 150$ GeV (right) for 2017 data and MC. The total SM background is shown as a stack of all contributing processes. The predictions for signal models with $m_{\tau/\nu} = 200$ GeV and $m_{\tau/\nu} = 500$ GeV (sum of all production and decay modes) are also shown as solid lines. The hatched gray band in the upper panel and the dark and light gray bands in the lower panel represent the total, statistical, and systematic uncertainties on the expected background, respectively.
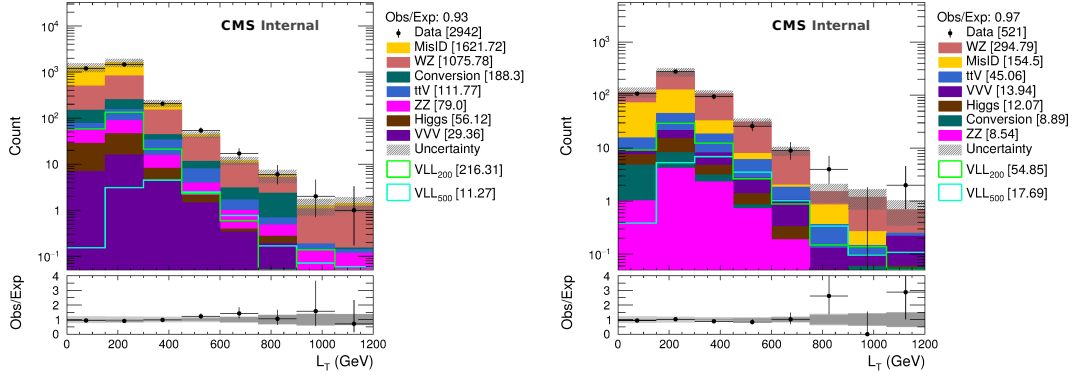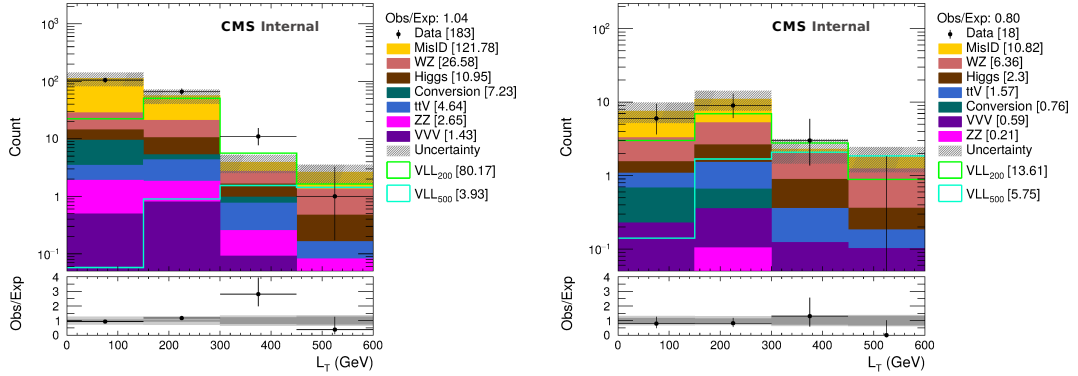

Figure A.18: The L_T distribution for events with two same charge light leptons and a hadronically decaying τ lepton with $50 < p_T^{miss} < 150$ GeV (left) and with $p_T^{miss} > 150$ GeV (right) for 2017 data and MC. The total SM background is shown as a stack of all contributing processes. The predictions for signal models with $m_{\tau/\nu} = 200$ GeV and $m_{\tau/\nu} = 500$ GeV (sum of all production and decay modes) are also shown as solid lines. The hatched gray band in the upper panel, and the dark and light gray bands in the lower panel represent the total, statistical, and systematic uncertainties on the expected background, respectively.

A.4   Dilepton Control Regions for 2017 analysis
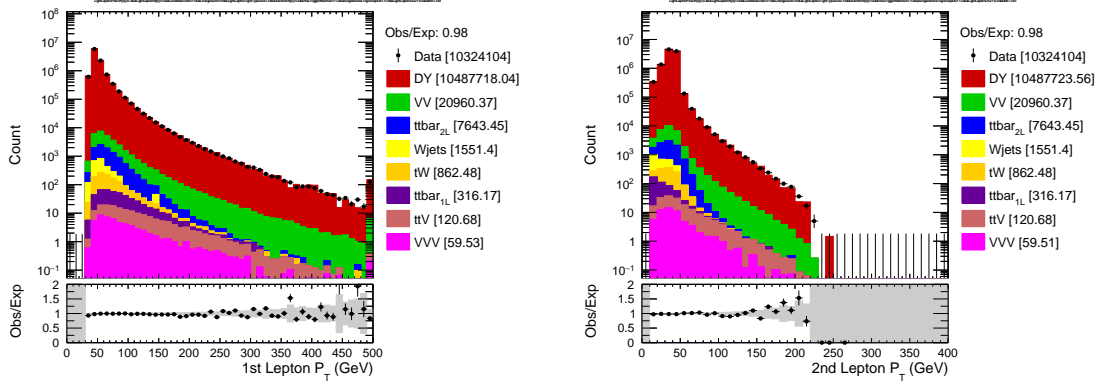
Figure A.19: Leading (left) and subleading (right) electron p_T distributions in the DY$\rightarrow$ee enriched dilepton selection in the 2017 data.
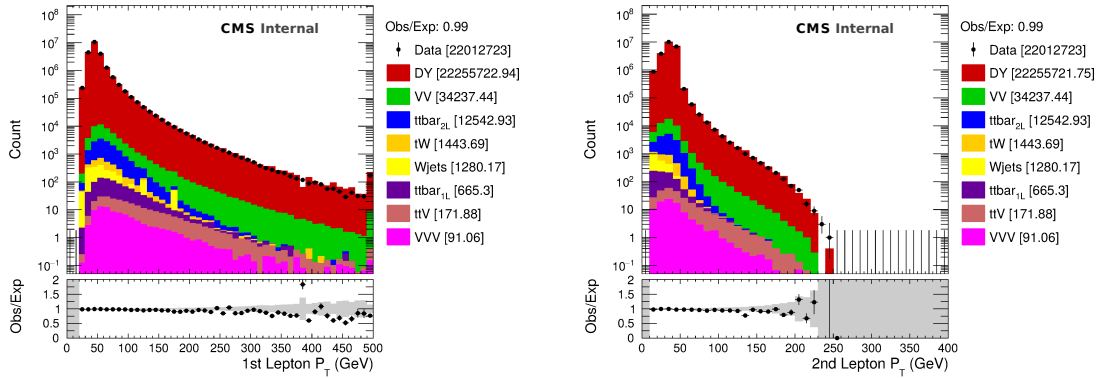
Figure A.20: Leading (left) and sub-leading (right) muon p_T distributions in the DY$\rightarrow$ $\mu\mu$ enriched dilepton selection in the 2017 data.
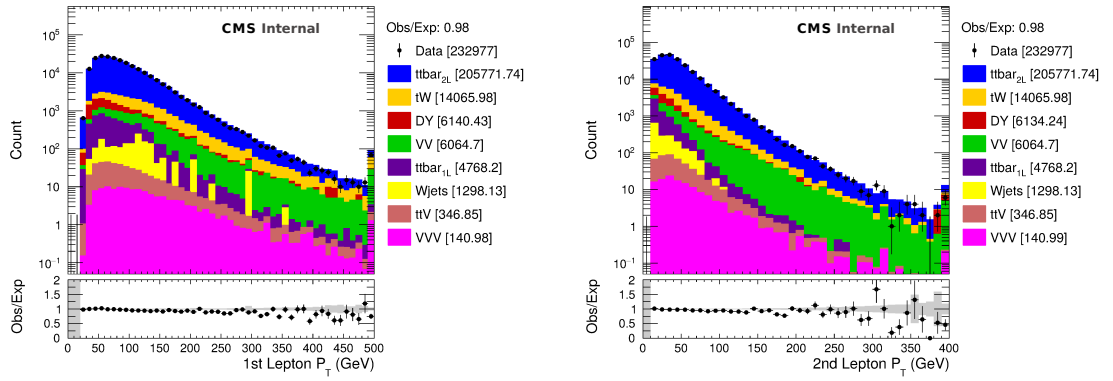
Figure A.21: Muon p_T (left), electron p_T (right) distributions in the $t\bar{t} \to e\mu$ enriched dilepton selection in the 2017 data.

A.5 The L_T distributions in all primary control regions, independently for 2016 and 2017
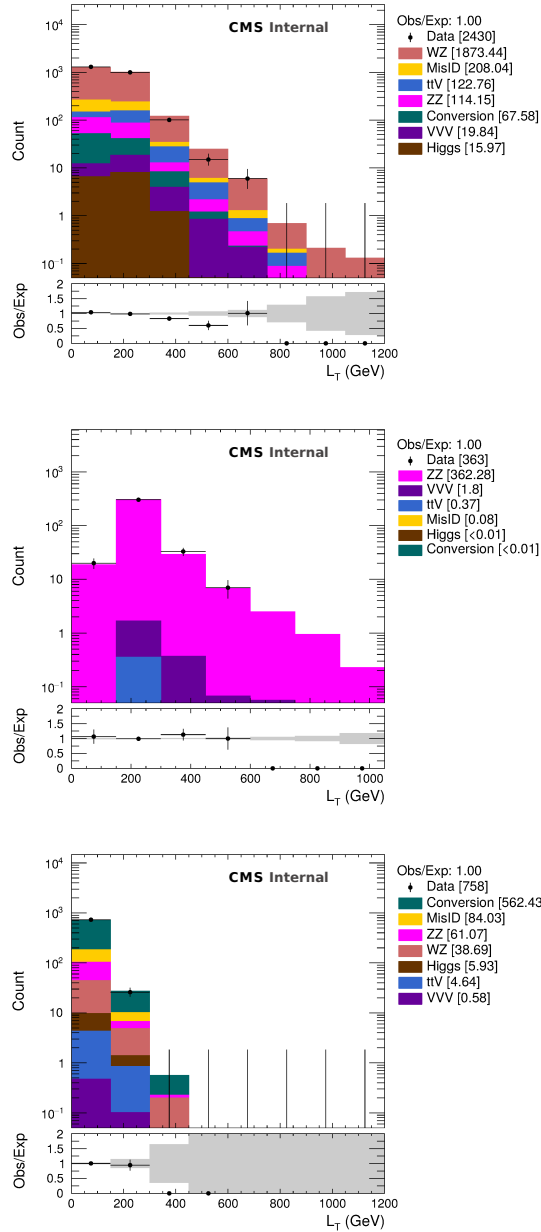
Figure A.22: The L_T distributions in WZ(top), ZZ(center), conversion(bottom) enriched region with 2016 data and MC.

107
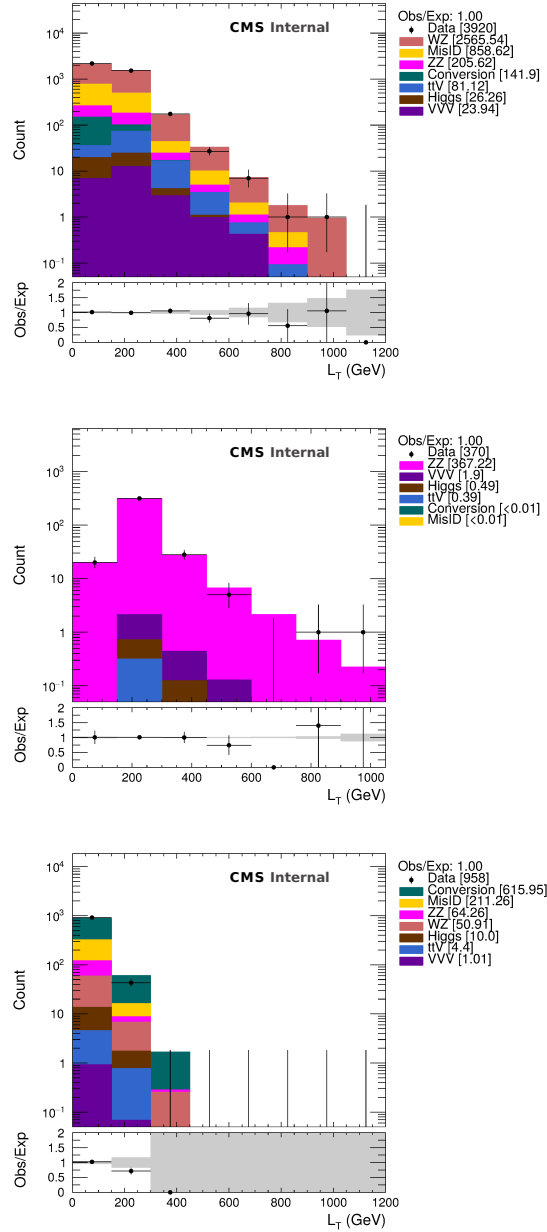
Figure A.23: The L_T distributions in WZ(top), ZZ(center), conversion(bottom) en-riched region with 2017 data and MC.

Bibliography

[1] M. Tanabashi et al. Review of particle physics. *Phys. Rev. D*, 98:030001, 2018.

[2] C. Cohen-Tannoudji, J. Dupont-Roc, and Gilbert Grunberg. *PHOTONS AND ATOMS. INTRODUCTION TO QUANTUM ELECTRODYNAMICS. (IN FRENCH)*. 1987.

[3] Enrico Fermi. Tentativo di una teoria dei raggi β. *Il Nuovo Cimento (1924-1942)*, 11(1):1, Sep 2008.

[4] H. Fritzsch. Quantum Flavor Dynamics. *Stud. Nat. Sci.*, 12:275–322, 1977.

[5] Abdus Salam. Weak and Electromagnetic Interactions. *Conf. Proc.*, C680519:367–377, 1968.

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

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

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

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

[10] P. A. R. Ade et al. Planck 2013 results. I. Overview of products and scientific results. *Astron. Astrophys.*, 571:A1, 2014.

[11] Takeo Moroi and Yasuhiro Okada. Radiative corrections to Higgs masses in the supersymmetric model with an extra family and antifamily. *Mod. Phys. Lett. A*, 7:187, 1992.

[12] K. S. Babu, Ilia Gogoladze, Mansoor Ur Rehman, and Qaisar Shafi. Higgs boson mass, sparticle spectrum and little hierarchy problem in extended MSSM. *Phys. Rev. D*, 78:055017, 2008.

[13] Stephen P. Martin. Extra vector-like matter and the lightest Higgs scalar boson mass in low-energy supersymmetry. *Phys. Rev. D*, 81:035004, 2010.

[14] Motoi Endo, Koichi Hamaguchi, Sho Iwamoto, and Norimi Yokozaki. Higgs mass and muon anomalous magnetic moment in supersymmetric models with vector-like matters. *Phys. Rev. D*, 84:075017, 2011.

[15] Marc Sher. Charged leptons with nanosecond lifetimes. *Phys. Rev. D*, 52:3136, 1995.

[16] Scott D. Thomas and James D. Wells. Phenomenology of massive vectorlike doublet leptons. *Phys. Rev. Lett.*, 81:34, 1998.

[17] Paul H. Frampton, P. Q. Hung, and Marc Sher. Quarks and leptons beyond the third generation. *Phys. Rept.*, 330:263, 2000.

[18] Radovan Dermisek. Insensitive unification of gauge couplings. *Phys. Lett. B*, 713:469, 2012.

[19] James Halverson, Nicholas Orlofsky, and Aaron Pierce. Vectorlike leptons as the tip of the dark matter iceberg. *Phys. Rev. D*, 90:015002, 2014.

[20] Radovan Dermisek and Aditi Raval. Explanation of the muon g-2 anomaly with vectorlike leptons and its implications for Higgs decays. *Phys. Rev. D*, 88:013017, 2013.

[21] Radovan Dermisek. Unification of gauge couplings in the standard model with extra vectorlike families. *Phys. Rev. D*, 87:055008, 2013.

[22] Nilanjana Kumar and Stephen P. Martin. Vectorlike leptons at the Large Hadron Collider. *Phys. Rev. D*, 92:115018, 2015.

[23] P. Achard et al. Search for heavy neutral and charged leptons in e^+e^- annihilation at LEP. *Phys. Lett. B*, 517:75, 2001.

[24] Georges Aad et al. Search for heavy lepton resonances decaying to a Z boson and a lepton in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *JHEP*, 09:108, 2015.

[25] Lyndon Evans and Philip Bryant. LHC machine. *Journal of Instrumentation*, 3(08):S08001–S08001, aug 2008.

[26] Esma Mobs. The CERN accelerator complex. Complexe des acclrateurs du CERN. Jul 2016. General Photo.

[27] S. Chatrchyan et al. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008.

[28] W. Erdmann. The CMS pixel detector. *Int. J. Mod. Phys.*, A25:1315–1337, 2010. [,41(2010)].

[29] S. Schael. The CMS silicon strip detector - mechanical structure and alignment system. *Nucl. Instrum. Meth.*, A511:52–57, 2003.

[30] Hannsjrg Weber. The phase-1 upgrade of the CMS pixel detector. In *Proceedings, 2016 IEEE Nuclear Science Symposium and Medical Imaging Conference: NSS/MIC 2016: Strasbourg, France*, page 8069719, 2016.

[31] CERN. Geneva. LHC Experiments Committee. The CMS electromagnetic calorimeter project. 1997.

[32] *The CMS hadron calorimeter project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997. The following files are from ¡a href=.

[33] P. Giacomelli. The cms muon detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 478(1):147 – 152, 2002. Proceedings of the ninth Int.Conf. on Instrumentation.

[34] S. Agostinelli et al. Geant4–a simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250, 2003.

[35] S. Abdullin, P. Azzi, F. Beaudette, P. Janot, and A. Perrotta. The fast simulation of the CMS detector at LHC. *J. Phys. Conf. Ser.*, 331:032049, 2011.

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

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

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

[39] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *JHEP*, 04:063, 2008.

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

[41] W Adam, R Frühwirth, A Strandlie, and T Todorov. Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC. *J. Phys. G*, 31:N9, 2005.

[42] Vardan Khachatryan et al. Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV. *JINST*, 10:P06005, 2015.

[43] Vardan Khachatryan et al. Reconstruction and identification of τ lepton decays to hadrons and ν_τ at CMS. *JINST*, 11:P01019, 2016.

[44] E.K. Friis. Tau reconstruction and identification at the compact muon solenoid. *Nuclear Physics B - Proceedings Supplements*, 218(1):256 – 261, 2011. Proceedings of the Eleventh International Workshop on Tau Lepton Physics.

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

[46] CMS Collaboration. Jet energy scale and resolution performance with 13 TeV data collected by CMS in 2016. CMS Detector Performance Report CMS-DP-2018-028, CERN, 2018.

[47] 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.

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

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

[50] 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.

[51] Tom Melia, Paolo Nason, Raoul Rontsch, and Giulia Zanderighi. W^+W^-, WZ and ZZ production in the POWHEG BOX. *JHEP*, 11:078, 2011.

[52] Paolo Nason and Giulia Zanderighi. W^+W^-, WZ, and ZZ production in the POWHEG-BOX-V2. *Eur. Phys. J. C*, 74:2702, 2014.

[53] 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.

[54] E. Bagnaschi, G. Degrassi, P. Slavich, and A. Vicini. Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM. *JHEP*, 02:088, 2012.

[55] Michael Klasen, Karol Kovarik, Paolo Nason, and Carole Weydert. Associated production of charged Higgs bosons and top quarks with POWHEG. *Eur. Phys. J. C*, 72:2088, 2012.

[56] Yanyan Gao, Andrei V. Gritsan, Zijin Guo, Kirill Melnikov, Markus Schulze, and Nhan V. Tran. Spin determination of single-produced resonances at hadron colliders. *Phys. Rev. D*, 81:075022, 2010.

[57] Sara Bolognesi, Yanyan Gao, Andrei V. Gritsan, Kirill Melnikov, Markus Schulze, Nhan V. Tran, and Andrew Whitbeck. On the spin and parity of a single-produced resonance at the LHC. *Phys. Rev. D*, 86:095031, 2012.

[58] Ian Anderson, Sara Bolognesi, Fabrizio Caola, Yanyan Gao, Andrei V. Gritsan, Christopher B. Martin, Kirill Melnikov, Markus Schulze, Nhan V. Tran, Andrew Whitbeck, and Yaofu Zhou. Constraining anomalous HVV interactions at proton and lepton colliders. *Phys. Rev. D*, 89:035007, 2014.

[59] Andrei V. Gritsan, Raoul Röntsch, Markus Schulze, and Meng Xiao. Constraining anomalous Higgs boson couplings to the heavy flavor fermions using matrix element techniques. *Phys. Rev. D*, 94:055023, 2016.

[60] 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.

[61] Vardan Khachatryan et al. Event generator tunes obtained from underlying event and multiparton scattering measurements. *Eur. Phys. J. C*, 76:155, 2016.

[62] Albert M Sirunyan et al. Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements. Submitted to *Eur. Phys. J. C*, 2019.

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

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

[65] Measurements of inclusive and differential Z boson production cross sections in pp collisions at $\sqrt{s} = 13$ TeV. Technical Report CMS-PAS-SMP-15-011, CERN, Geneva, 2016.

[66] Serguei Chatrchyan et al. Measurement of the $t\bar{t}$ production cross section in the dilepton channel in pp collisions at $\sqrt{s} = 7$ TeV. *JHEP*, 11:067, 2012.

[67] Vardan Khachatryan et al. Search for Third-Generation Scalar Leptoquarks in the tτ Channel in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV. *JHEP*, 07:042, 2015. [Erratum: JHEP11,056(2016)].

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

[69] Thomas Junk. Confidence level computation for combining searches with small statistics. *Nucl. Instrum. Meth. A*, 434:435, 1999.

[70] Alexander L. Read. Presentation of search results: The CL_s technique. *J. Phys. G*, 28:2693, 2002.

[71] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J. C*, 71:1554, 2011.

[72] Serguei Chatrchyan et al. Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC. *Phys. Lett.*, B716:30–61, 2012.

[73] Albert M. Sirunyan et al. Search for the Higgs boson decaying to two muons in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Phys. Rev. Lett.*, 122(2):021801, 2019.

[74] A. M. Sirunyan et al. Search for new physics in events with a leptonically decaying Z boson and a large transverse momentum imbalance in proton?proton collisions at $\sqrt{s} = 13$ TeV. *Eur. Phys. J.*, C78(4):291, 2018.

[75] Dan Guest, Kyle Cranmer, and Daniel Whiteson. Deep learning and its application to lhc physics. *Annual Review of Nuclear and Particle Science*, 68(1):161–181, 2018.

[76] N. Aloysius and M. Geetha. A review on deep convolutional neural networks. In *2017 International Conference on Communication and Signal Processing (ICCSP)*, pages 0588–0592, April 2017.

[77] Maxwell Stinchcombe and Halbert White. Universal approximation using feedforward networks with non-sigmoid hidden layer activation functions. pages 613–617, 12 1989. IJCNN International Joint Conference on Neural Networks ; Conference date: 18-06-1989 Through 22-06-1989.

[78] Katarzyna Janocha and Wojciech Marian Czarnecki. On loss functions for deep neural networks in classification. *CoRR*, abs/1702.05659, 2017.