Ph.D. DISSERTATION

Test of Neutrino Mass Models at the Large Hadron Collider

대형 강입자 충돌기를 이용한 중성미자 질량 모형 검증.

BY

전시현

FEBRUARY 2023

DEPARTMENT OF PHYSICS AND
ASTRONOMY
COLLEGE OF NATURAL SCIENCE
SEOUL NATIONAL UNIVERSITY

Ph.D. DISSERTATION

Test of Neutrino Mass Models at the Large Hadron Collider

대형 강입자 충돌기를 이용한 중성미자 질량 모형 검증.

BY

전시현

FEBRUARY 2023

DEPARTMENT OF PHYSICS AND
ASTRONOMY
COLLEGE OF NATURAL SCIENCE
SEOUL NATIONAL UNIVERSITY

Test of Neutrino Mass Models at the Large Hadron Collider

대형 강입자 충돌기를 이용한 중성미자 질량 모형 검증.

지도교수 양 운 기

이 논문을 이학박사 학위논문으로 제출함

2023 년 2 월

서울대학교 대학원

물리천문학부

전 시 현

전시현의 이학박사 학위논문을 인준함

2023 년 2 월

위 원 장	유종희
부위원장	양운기
위 원	김선기
위 원	최선호
위 원	이현수

Abstract

Although the Standard Model (SM) has been one of the most successful physics models of all time with the discovery of Higgs boson, there are still unsolved questions that cannot be explained due to conflicting experimental observations not predicted by the SM. Oscillations of the neutrinos is one of the representative example which is a clear evidence for the physics beyond the SM. The fact that neutrinos oscillate can be explained via the difference in mass and flavor eigenstates of neutrinos which in turn requires neutrinos to have nonzero masses. This contradicts to the SM where the mass of neutrinos are assumed to be zero. Many theories have been proposed to describe the mass mechanisms for neutrinos. The seesaw model is one of the most popular explanations that is feasible to explain the origin as well as the smallness of neutrino masses.

In this thesis, two analyses are presented in the context of neutrino mass mechanisms with the LHC experiment. The first analysis is the search for heavy neutrinos in type-I seesaw model using the final state with a single lepton associated with either a large-radius jet or small-radius jets. The proton-proton collision data at a center-of-mass energy of 13 TeV collected with the CMS detector located at the LHC is used for the analysis. With the full Run 2 data of 137 fb^{-1}, the analysis is performed in two folds from different event topologies that are represented by the jet configurations. This is the first attempt to search for heavy neutrinos using the decay mode to the Higgs boson. The second part of this thesis scans the parameter plane defined by the type-II seesaw model using the published LHC results that is the data-driven approach. In particular, the parameter plane which is favored by the anomalous W boson mass report

that the CDF collaboration made recently is analyzed in depth.

Keywords: High Energy Physics, LHC, CMS, Thesis, Neutrino, Heavy Neutrino, Heavy Neutral Lepton, Seesaw Model

Student Number: 2015-22596

Contents

List of Figures

List of Tables

Chapter 1

Preface

Particle physics, as its name suggests, is an effort to understand how the universe functions at very short distances in the basic minimal level constituents of the universe. This field of physics stems from the philosophical question of what is the smallest indivisible unit of matters, defining what the word *atom*, after the Greek word *atomos*, is since ancient Greece. Later in the 19th century, Dalton developed the theory of atoms defining atoms as a set of extremely small particles that form the matter.

This theory later was discovered to be wrong by finding even more smaller particles such as protons and neutrons in the early 1900s. Soon the particle physics field drastically developed with consequent findings of muons, pions, and so on and so forth. With the progress of the field but before the discovery of quarks which are the subparticles of protons and neutrons, a huge multitude of strongly interacting hadrons [1] were found and the term *subnuclear particle zoo* was popularized. Despite new findings of more and more hadrons, simplicity

[1] Hadrons are subatomic level composite particles made of two or more quarks.

and elegance which forms the heart of *good physics theory* were tarnished. Soon later, digging deeper with higher energy resurrected the particle physics by finding quarks and developing the Standard Model (SM) inclusively explaining the particle zoo with great accuracy.

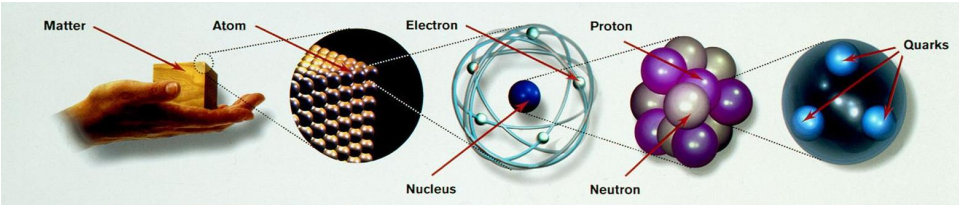

Figure 1.1: Image depicting the substructures of matter ending with the elementary level, quarks. Figure credit : CERN.

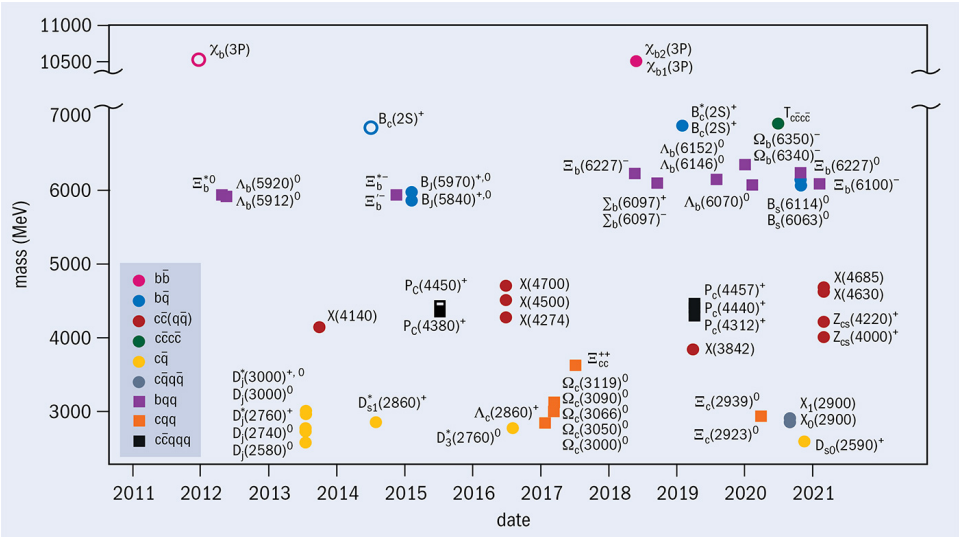

Figure 1.2: 59 hadrons discovered since 2011 by the Large Hadron Collider until the May of 2021. Figure credit : CERN.

During last 100 years, the SM theory was firmly constructed with the final puzzle piece, Higgs boson discovery, at the Large Hadron Collider (LHC). From

the results of the LHC, we know that universe is encoded with seventeen particles in total which are categorized into two sets : (1) Fermion : Constituents of matter that are leptons and quarks (2) Boson : Interaction mdediators that are gauge and scalar bosons. But is this the final remark of the particle physics that we understand everything? The answer is *no* and we still have loads of remaining questions about the universe. Why is there more matters than antimatters in the universe? How can gravitational force be understood in the context of quantum field theory? How can we find the dark matter that accounts for most part of the universe?

This thesis is dedicated to address the intriguing question regarding one of the fundamental particles, ***neutrino***. In particular, it is an attempt to answer the question, *Where does the mass of neutrino comes from?*. A lot of mysteries still remain in the sector of neutrinos, mainly due to its only weak interacting nature which makes it experimentally more difficult to find out what its properties are. But still, within the scope of the LHC using its detectors which include Compact Muon Solenoid (CMS), the neutrino sector can be peeked into with dedicated efforts and will be explained throughout the thesis.

This thesis is organized as following :

- **Chapter 2. Theoretical Backgrounds** : Describe what the SM theory in particle physics is and how Monte-Carlo (MC) event generation takes part in the field.

- **Chapter 3. Experimental Backgrounds** : Briefly introduce how the LHC and the CMS detector are operated and also discuss the Particle-Flow algorithm which is utilized to reconstruct physics objects at the CMS for analysis.

- **Chapter 4. Neutrinos** : Describe how neutrino started to play the role

in particle physics historically and introduce some of the representative works from the CMS related to the topic of neutrino mass mechanism.

- **Chapter 5. Search for Heavy Neutrinos at the CMS** : As for the first main part of the thesis, data analysis is performed using the full Run 2 dataset collected by CMS. Heavy neutrino in type-I seesaw model in single lepton channel is explored for its discovery.

- **Chapter 6. Neutrino Mass Model Reinterpretations with Published LHC Results** : As for the second main part of the thesis, a reinterpretation is performed using the published results from the LHC. In particular, type-II seesaw model is scanned to address the W boson mass anomaly which is from CDF's report in 2021.

- **Chapter 7. Conclusions** : Conclude this thesis with a brief summary of Chapter 4. and 5. and outlooks.

The analysis in Chapter. 5 is a classical analysis approach analyzing data with optimized selections for the BSM signal process (theory-driven), while the reinterpretation in Chapter. 6 is done in opposite direction fully exploiting already published results to exclude new BSM physics (data-driven).

Now, to achieve this goal of answering the question on neutrino masses, obviously I could not do it alone so I have been working as a member of the CMS collaboration since 2016. As the CMS collaboration functions by contributions from thousands of people, I also took roles in two different groups in CMS, Exotic Physics Analysis Group (EXO) and Physics Generator Group (GEN). I have been serving as the level 3 convener of EXO Monte-Carlo & Interpretations subgroup for two years starting from September of 2019 until August of 2021. Also in parallel, I was chosen as the representative CMS contact person for MadGraph_aMC@NLO by GEN which is the most popularly

used MC event generator at the LHC. Since January of 2022, I took the role of HepData contact person of EXO, reviewing and guiding the data preservation efforts ultimately to expedite the reinterpretation activities. In 2020, I was also awarded with the *Marie Sklodowska-Curie Early Stage Researcher Fellowship* to work for the MCnet project. The task was to build the infrastructure in MadGraph_aMC@NLO, interfacing it with Rivet and Contur for user-friendly and easy-approach reinterpretation tools. Finally, my major focus during last seven years (hopefully well explained throughout this thesis) was incorporating my EXO and GEN works in one single theme, analysis (re-)interpretations with BSM theories to explain neutrino mass mechanisms.

Chapter 2

Theoretical Backgrounds

2.1 The Standard Model of Particle Physics

2.1.1 Elementary Particles and Fundamental Forces

The Standard Model (SM) for particle physics is built up on the quantum field theory (QFT), describing the fundamental interactions of leptons, quarks, gauge bosons, and Higgs (H) bosons [1–12]. Specific behaviors of the interactions in strong and electroweak sectors are encapsulated in the $SU_C(3) \times SU_L(2) \times U_Y(1)$ group structure of the SM. This was not something gathered in a short time scale but instead its empirical development started more than 100 years ago from the discovery of the electron in 1897. This soon lead to quantized understanding of the atomic nucleus proposed by Bohr in 1913. Finally in 2012, the discovery of the Higgs boson at the LHC completed our understanding of the universe under the illustration of the SM being one of the most successful physics models of all time.

The constituents that build up the SM as its fundamental particles [1] can

[1] The charge conjugated partners of particles, that are the antiparticles, have identical mass

be divided in two, fermions and bosons. Fermions are particles with half-odd-integer spin which obeys the Fermi-Dirac statistics. Quarks and leptons (including neutrinos) are further categorized for fermions by their charges : leptons have integer charge and quarks have fractional charge. Unlike fermions, bosons are particles with integer spin obeying the Bose-Einstein statistics. Matter is composed of fermions while bosons are responsible for the interactions between the particles acting as the force mediators. The full list of the SM particles and their properties are written in Table. 2.1 for fermions and in Table. 2.2 for bosons.

Particle	**Generation**	**Spin**	**Charge [e]**	**Mass**
Up quark (u)	1	1/2	2/3	2.16 MeV
Down quark (d)	1	1/2	$-1/3$	4.67 MeV
Strange quark (s)	2	1/2	$-1/3$	93.4 MeV
Charm quark (c)	2	1/2	2/3	1.27 GeV
Bottom quark (b)	3	1/2	$-1/3$	4.18 GeV
Top quark (t)	3	1/2	2/3	172.69 GeV
Electron (e)	1	1/2	-1	0.511 MeV
Muon (μ)	2	1/2	-1	105.658 MeV
Tau (τ)	3	1/2	-1	1776.86 MeV
e neutrino (ν_e)	1	1/2	0	~ 0 MeV
μ neutrino (ν_μ)	2	1/2	0	~ 0 MeV
τ neutrino (ν_τ)	3	1/2	0	~ 0 MeV

Table 2.1: Properties of the fermions in the SM.

There are four types of fundamental forces known to date : The gravitational force, the electromagnetic (EM) force, the weak force, and the strong force. The gravitational force has very small scale of interaction compared to the other three fundamental forces and its origin, or the mediator of the force named graviton, is still at a hypothetical stage which still has to be proven. Thus it is not yet fully incorporated as the grand unification theory. However, the other

but carry opposite charges.

Particle	**Spin**	**Charge [e]**	**Mass**
Photon (γ)	1	0	0 MeV
Gluon (g)	1	0	0 MeV
W boson ($W^{\pm}$)	1	± 1	80.377 GeV
Z boson (Z)	1	0	91.1876 GeV
H boson (H)	0	0	125.25 GeV

Table 2.2: Properties of the bosons in the SM.

three forces are enough to explain the phenomena occurring at the high energy physics field, considered in this thesis.

The electromagnetic force occurs between electrically charged particles with exchange of photons as the mediator. The strength of its interaction is known to be the fine structure constant $\alpha_{\mathrm{EM}} \simeq 1/137$ determined by comparing the electron mass with the electrostatic energy between the two electrons separated by one natural unit of length. It further depends on the momentum transfer scale (Q^2) of the interaction and can increase up to $\alpha_{\mathrm{EM}}(Q^2 = m_{\mathrm{Z}}^2) \simeq 1/129$ at $Q^2 = m_{\mathrm{Z}}^2$.

The massless gluons are the mediators of the strong interaction. Its fundamentals are described in the framework of QFT called quantum chromodynamics (QCD). QCD introduces three types of charges that is carried by strong interacting particles, also known as colors : Red (R), green (G), and blue (B). Analogous to electromagnetic interaction where antiparticles carry positive charges instead of negative charges, antiparticles carry anticolors that are denoted as $\bar{R}$, $\bar{G}$, and $\bar{B}$.

Gluons carry eight types of color charges from the combination of colors and anticolors, discarding the $(R\bar{R} + G\bar{G} + B\bar{B})$ color combination as it is colorless and thus becomes only a hypothetical state. Gluon has the self interacting nature which makes such three gluon vertex $g \to gg$ available. Finally, unlike α_{EM},

the strength of QCD interaction α_S decreases with energy. In other words, at high energy (or at short distance scale), strong interaction becomes asymptotically weaker and quarks in these scale can be treated as free particles. On the contrary, at low energy (or at long distance scale), strong interaction becomes stronger so that its effective color charge increases with distance.

The weak force is mediated by W or Z bosons. Charged-current interactions are mediated by the W boson whereas neutral-current interactions are mediated by the Z boson. Weak interaction has several notable characteristics that leads to its special features. First of all, as the name suggests, the interaction strength $\alpha_{\text{weak}} \simeq 10^{-6}$ is much smaller than the electromagnetic and strong interactions. Secondly, the charged current interactions have only been observed with left-handed state fermions and right-handed state antifermions. Following caveat is that numerous experimental evidences [13–23] already indicates that neutrino should have mass above zero. This contradicts to the experimental observations of only left-handed neutrinos while no evidence is shown for right-handed neutrinos where both are needed to form the Yukawa coupling for its mass generation. Such facts lead to many intersting scientific developments in both theoretical and experimental sides which will be further discussed in Chapter. 4.

Later by Glashow, Salam, and Weinberg, the electromagnetic interaction and the weak interaction was unified to the electroweak (EW) theory. It follows from the theory that the ρ *parameter* defined from the mass of W and Z bosons with the electroweak angle (θ_{EW}) forms a special relation. This will be further discussed in Chapter. 6 with its connection to the additional Higgs portal that provides fine tuning corrections.

2.1.2 Open Issues with the Standard Model

Despite the remarkable success of the SM to describe the nature, there still remains several open issues that need to be addressed regarding the SM. In the theoretical aspect, the *hierarchy problem* which concerns the large discrepancy between the mass of H boson around 125 GeV and the other maximum possible energy scale of the theory that can be considered (e.g. the grand unification scale 10^{16} GeV or the Planck scale 10^{19} GeV). As its scale differs by a huge amount, if the discovered H boson is the constituent of the SM, it opens the possibility that the SM is only an effective theory of the true nature at higher energy scale. The Supersymmetry (SUSY) model is introduced in this context, if there is no new physics below the Planck scale, the radiative corrections to the H boson mass is at the enormous Planck scale. However by employing SUSY model, the immense radiative correction cancels out through loop contributions of the SUSY partner particles. Such interpretation is introduced in some studies [24] claiming the need of fine-tuning for *naturalness*, while others criticize that such attempts to solve the mysteries of nature with asethetic reasons are a waste of time [25].

Several tensions derived from measurements that conflict with the predictions of the SM are also left as open issues. One of the famous issues is the long standing anomaly of the magnetic moment of muons which is referred to as muon $g-2$. At a precision of 0.35 ppm, the anomaluos magnetic moment of muon disagrees by 4.2 σ from the SM predictions [26]. Such deviation is not shown for the case of electrons. Additionally, LHCb reported violations of lepton flavor universality [2] evidenced by measuring the branching fractions from B mesons [27, 28]. As aforementioned, neutrinos that are postulated to have zero

[2] The SM expects leptons to behave in a same manner but with differences only arousing from different masses of each flavors.

masses in the SM, are observed with oscillations [13–23] which in turn imply that neutrinos carry nonzero masses. For its possibility to address the matter-antimatter asymmetry of the universe with Majorana nature of neturinos, it is considered as the *smoking gun* for new physics outside the SM coverage which will be discussed in Chapter. 4.

2.2 Monte-Carlo Event Simulation

2.2.1 Factorization Theorem

In order to calculate the cross section of a scattering process in hadron-hadron collisions, information on momentum which a parton carries is the starting point of its anatomy. This is due to the fact that a parton does not singly emerge but instead comes from its parent hadron, carrying fractional momentum of the hadron. This lies in the nonperturbative realm of QCD which is not calculable. However, by introducing the factorization theorem [29], it is possible to calculate by separating the short and long distance effects in QFT. This approach is the key to the success of hadron-hadron collisions for both theory and experiment.

The cross section for a scattering process $ab \to X$ can be written as [30]

$$\sigma = \sum_{a,b} \int dx_a dx_b \int f_a(x_a, \mu_F) f_b(x_b, \mu_F) d\sigma_{ab}(\mu_F, \mu_R). \tag{2.1}$$

This formula consists of two key items which are

- $f(x, \mu_F)$: The parton distribution functions (PDF) [31] encapsulate the information of probability to find partons in a hadron as a function of momentum fraction (x) carried by partons. It contains the information on long distance structure of hadrons.

- $\sigma_{ab}(\mu_F, \mu_R)$: The parton level cross section describes the probability for initial partons a and b to form the final state X. This describes the short distance physics with larger transverse momentum transfers.

Here, we introduce two scales, μ_F and μ_R which are factorization scale and renormalization scale, respectively. The scale μ_F is the energy threshold which sets the boundary between the short and long distance physics. All long distance processes below the threshold are absorbed into the PDF while above are included in the parton level cross section. In order to resolve the logarithmic divergences that rise from loop diagram contributions, the scale μ_R is introduced which eventually controls the QCD coupling constant $\alpha_S(\mu_R)$.

2.2.2 Parton Distribution Function

As a consequence of the factorization theorem, the constituents of proton which are quarks and gluons [3] can be treated independently from the hard scattering process. The constituents of protons are named parton which was first proposed by Richard Feynman, considering any hadron as a composition of multiple point-like constituents. The information of partons and its probability to be found in a hadron are encoded in PDFs [32]. At the LHC, beam energy is in $O(10\ \text{TeV})$ scale and partons carry out some fraction of energy when it gets involved in hard scattering process. Also considering higher order calculations where gluons can even contribute to EW processes, it becomes obviously important to accurately estimate the parton distribution function as it is is the underlying stone for the whole factorization theorem we utilize in high energy physics.

There has been many updates on modern PDFs. Nowadays, not only considering gluons being responsible for the missing momentum of protons and scaling violations, heavy flavor quarks (charm and bottom quarks) that are heavier than the mass of proton are also regarded as a parton that forms a proton in sea quark level. In addition, although much weaker than the QCD

[3] For extreme cases, electrons and photons are sometimes also considered as constituents

interaction, EW interaction is also introduced to PDFs that radiates photons and eventually leptons starting from quarks [33]. Such development lead to diverse physics possibilities at the LHC by considering it as proton-lepton or proton-photon collider, expanding the potential. This will also be explored in this thesis in Chapter. 5.

The determination of PDFs are done by extraction of fitting results from multiple experimental dataset. It is also known as the global QCD fit which depends on three inputs :

- Theoretical base : Structure functions which provides the information on hadron's partonic structure and DGLAP equation [34] describing evolution of PDF under different energy scales. More details on theoretical foundation are explained in Ref. [32].

- Experimental data : A large variety of data is used starting from the first deep inelastic scattering experiments and the latest results from LHC. Most of the low Q^2 phase space is constrained from DIS data while high Q^2 which needs higher energy scale is given constraints from LHC.

- Statistical framework : The PDF that is used in Chapters. 5 and 6 is from the NNPDF collaboration [32]. Every PDF collaboration adopts their own framework for statistical analysis and NNPDF relies on machine learning for parameterization and optimization for fitting the PDF.

The PDFs at are shown as a function of Bjorken scale x for $Q^2 = 10\ \text{GeV}^2$ and $Q^2 = 10^4\ \text{GeV}^2$ as an example in Fig. 2.1. Valence quarks, which give rise to hadron quantum numbers, are either up (u_v) or down (d_v) flavor. It can be easily noticed that u_v is universally double the height when compared to d_v. This can be easily interpreted when considering the proton is composed of two

up and down valence quarks. It is also depicted in Fig. 2.1 that sea quarks are also considered as constituents of protons. These are the flux created from gluon splitting inside protons which becomes more active in larger Q^2.

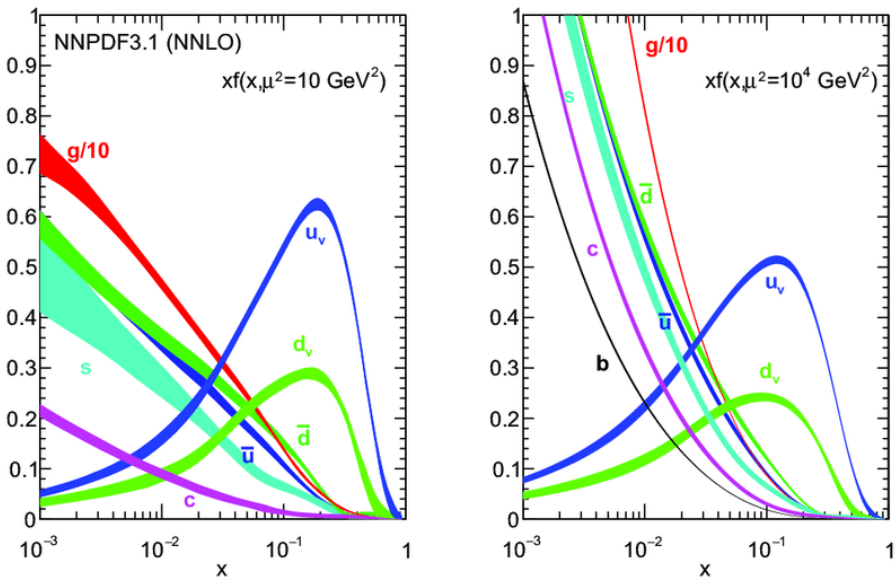

Figure 2.1: PDF at NNLO is shown as an example at factorization scales of 10 GeV^2 (left) and 10^4 GeV^2 (right) extracted from NNPDF collaboration. The figure is taken from Ref. [35].

2.2.3 Hard Process

While the parton content of the proton is determined from PDFs, next step for the MC event simulation is computing the main physics process that are of our interest. Its calculation is done perturbatively, expanded as

$$\sigma = \sigma_{\mathrm{LO}}(\alpha_S^{0+k}) + \sigma_{\mathrm{NLO}}(\alpha_S^{1+k}) + \sigma_{\mathrm{NNLO}}(\alpha_S^{2+k}) + \cdots . \tag{2.2}$$

Leading order (LO), next-to-LO (NLO), and next-to-NLO (NNLO) are represented with σ_{LO}, σ_{NLO}, and σ_{NNLO}, respectively. k denotes order of α_S for the

process's tree-level diagram. One can realize that the orders are corrected for α_S in Eq. 2.2. In order to construct healthy perturbative expansion, the corrections need to be sufficiently small which is the case for LHC collisions when considering the scales of energy, for example, $\alpha_S(m_{\mathrm{Z}}) = 0.1189 \pm 0.0010$. Also EW corrections from photons are considered to be ignorably small compared to α_S and also not employed in most of matrix element level calculators or only available for fixed order calculations due to technical limitations so far.

The cross sections in order X is written as

$$\sigma_{\mathrm{X}} = \int d\Phi_n |\mathcal{M}_{\mathrm{X}}|^2 \tag{2.3}$$

where $\mathcal{M}_{\mathrm{X}} = \sum \mathcal{F}^{\mathrm{X}}_{ab}$ is the matrix element, the sum of all amplitude from Feynman diagrams (and rules), $d\Phi_n$ denotes the phase space elements for particles in the final state. This is a highly multidimensional computation with complex peak structures most of the time coming from resonance productions which adopts the Monte-Carlo (MC) technique to overcome the difficulties.

Most of the matrix element level calculators go through unweighting procedure for every event to assign with same weights per event. This is processed in sequential steps : (1) Estimate a good guessing function of the full integral over the phase space. (2) Generate events at random points based on the guessing function and accept events if it is within the desired phase space integral. If it is not contained in the desired integral, the event is rejected. This is done multiple times at random points until the events stacked altogether is able to represent the desired phase space integral.

Some advantages can be expected by unweighting the events. Most of all when performing systematic uncertainty studies in experiments, without unweighting the events, an event with enormously large weight can induce a huge systematic effect which is not natural and also not a constructive study. This

also connects to experimental intuition that the actual data taken from experiments has equal weight at 1 that later corresponds to the total event count. Another advantage comes from the following simulation step such as the detector simulation which is time consuming and heavy with core usages. In order not to introduce unstable nature of weighted events as discussed above, powerful statistics is needed from MC event generators. But if all events are passed through the following sequential steps, it would lead to a waste of time and resources without much gain in physics.

2.2.4 Parton Shower and Hadronization

The step after the hard scattering process is referred to as the parton shower [36]. Although the quantum electrodynamics (QED) interactions with photons are also considered in this step when the final state particles of hard scattering process carries an electromagnetic charge, the word *parton* is derived from the fact that the QCD interactions are more likely to happen with gluon self interaction and stronger interaction strength. The process goes through multiple radiations of photons and gluons or split into quark pairs until it decreases up to the point where it needs to be taken care by nonperturbative calculations.

The factorization scheme again allows us to describe the parton shower by setting up the soft and collinear boundaries against the hard scattering process. The probabilities of radiating a parton or splitting into partons in this soft and collinear limit can be calculated with logarithmic behavior as

$$d\sigma_{n+1} \simeq d\sigma_n \frac{\alpha_S}{2\pi} \frac{dt}{t} dz d\phi P_{ji}(z, \phi). \tag{2.4}$$

The splitting function [30] describes the n-parton differential cross section transition to $(n+1)$-parton differential cross section through the radiation or split of partons. ϕ represents the azimuthal angles of the parton shower and P_{ji} are the Altarelli-Parisi functions describing the kernels.

Most importantly, t is the evolution variable which can characterize the limits with θ, p_T, or Q which are opening angles, transverse momentum, and virtuality of the parton shower. Different choices of the t parameter results in different computation methods in MC event generators. As can be inferred from Eq. 2.4, the formula introduces singularity when t reaches 0 which implies soft ($p_{\mathrm{T}} \to 0$) and collinear ($\theta \to 0$) limits. This can be easily comprehended from the nature of QCD where gluons can self interact infinitely with low energy emissions ($g \to gg$). Thus, it is necessary to define a scale protecting the divergence from the singularity with low energy scale $O(1\ \mathrm{GeV})$ which is called the hadronization scale.

The hadronization is responsible for converting the multitudes of partons to colorless final states. The first attempt to consider this process with applicable MC modelling was the Feynman-Field model. It starts by considering individual partons recursively applying $q \to q' + hadron$. However this model was strongly frame dependent and not infrared safe. Also the model did not embed the confinement that governs the QCD processes with gluon self-interactions. One commonly used hadronization model is called string model. From the calculations of lattice QCD in the quenched approximation in Fig. 2.2, the confinement can be given with a potential $V(R) \simeq \kappa R$ for large R limit where R is the separation of the two partons [37].

Jet Merging

While the main theoretical calculations are dealt in the first two steps of MC event generation that are hard process and parton shower (including hadronization), one more step would also be worth to discuss. The first hard process is given with a given number of incoming partons and outgoing particles. For example, Drell-Yan (DY) to dilepton process can be considered as $2 \to 2$ pro-

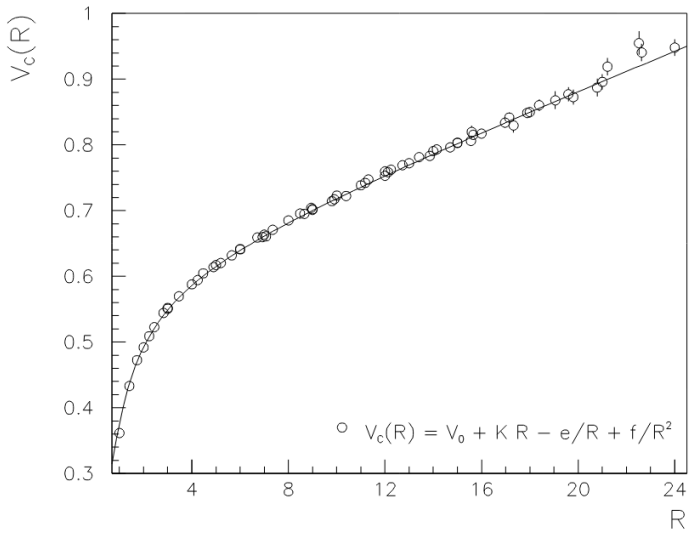

Figure 2.2: The quark potential from quenched lattice QCD shown in lattice units (0.1 fm). The figure is taken from Ref. [37].

cess. But in the real world, especially at the LHC, from QCD interactions of incomping partons, it is very likely to yield jets. If the hard process DY is calculated solely for $2 \to 2$, additional jets are expected to be produced from the parton showers. Although this can be effective for soft QCD emissions that provide small Lorentz boosts to the dilepton system, it would not be feasible to evaluate the hard QCD emissions that could also occur.

So for better modelling of such phase space, jet merging technique was developed [38]. Jet merging is applied where the additional jets are modelled from both hard process and parton shower. Now in matrix element level calculations, instead of $2 \to 2$, matrix elements up to $2 \to 2 + n$ where n is the number of additional jets is calculated and then goes through parton shower step. Matrix element calculations hinder from divergence (e.g. infinitely branching $g \to gg$), and on the other hand, parton shower cannot reach above certain scale. How-

ever, as the jets emerge in both steps, the same jet that hits the same phase space in between can be produced twice which potentially leads to the double counting of events. Thus, jet merging is responsible for splitting the phase space for jets to ensure the correct coverage without double counting. Conceptually, if the jet is soft, below a given artificial jet merging scale, in the event but comes from hard process not parton shower, the event is dropped. Conversely, if the jet is hard above the jet merging scale, events that acquired the jet from parton shower are discarded.

To ensure the jet merging is well modelled, generated MC events are tested with differential jet rate (DJR) distributions. DJR, which roughly is a proxy of the jet's p_{T}, shows the resolution between n-th and $(n+1)$-th jets and is expected to be smooth despite the artificial selection of jet merging scales. As an example, Fig. 2.3 is given with blue and red curves which each represent jets from parton shower and matrix element, while jet merging scale is denoted with the purple dashed line, labeled cutoff. Again, the basic concept of jet merging is accepting the events only if jets above the merging scale is from matrix element and only if jets below the merging scale is from parton shower. Thus the blue and red curves do not appear in each other's sections divided by the merging scale. If the merging scale is poorly chosen the artificial selection of the scale leads to non-continuous jet p_{T} distribution so this needs to be iteratively tuned for proper choices. In addition, as the emission of jets highly depend on the scale of momentum transfer occuring in the event, different physics processes lead to different choices of jet merging scales.

2.2.5 Detector Simulation

In order to construct a realistic MC event, not only the theoretical calculations of a physics process but also a modelling of detector repsponses need to be

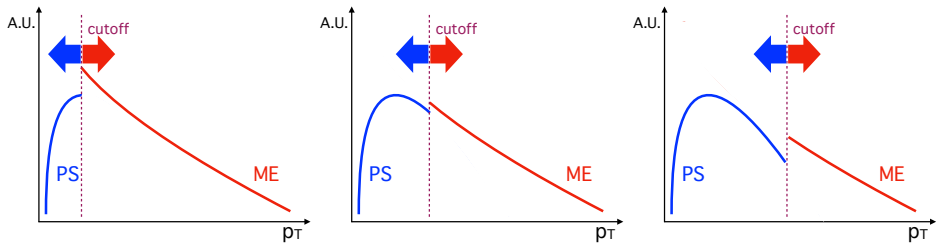

Figure 2.3: Example figures of jet merging scales. Left and right depicts the case where jet merging scale (cutoff) is chosen either too low or too high inducing a sudden jump in the distributions. However with properly tuned choice of the scale as in the middle, the jets are expected to be better modelled.

simulated. This is normally achieved by using the GEANT 4 [39] program which is encoded with a large spectrum of physics models that range from several eV to TeV scale simulating how particles interact when propagating through the detector material. It provides the geometry of how detector is built with its subdetector system and simulates the detector material response of particles. Several detector geometry packages are provided by GEANT 4 which includes the geometry of CMS detector which makes it feasible to simulate MC events to be used for CMS data analysis.

2.2.6 Pileup Events

Dealing with *pileup events* is one of the greatest challenges for proton-proton collision experiments. The deeply inelastic collisions that are usually the processes of interest are contaminated by subordinate proton-proton interactions that can occur either within or across different bunch crossings. Thus, such effect has to be modelled for realistic experimental environment where multiple pileup interactions can be taken into account (See Fig. 2.4 for example). It obviously takes a long time and a lot of computing resources to fully and indi-

vidually simulate such soft QCD interactions for each event. Thus, high energy physics experiments use the premixing method by first constructing a library of minimum bias (MinBias) sample. The MinBias sample is the collection of events that is collected by experiments with the loosest trigger requirements. It is generated from the neutrino gun simulation [4] which effectively has no effect in collider simulations as neutrinos leave no trace even in the hard scattering part. These are processed all the way through detector simulations and form a library of purely pileup interaction events. When the library is prepared, MC simulated events that are of physics interest with detector responses are overlayed (premixed) with randomly selected events in the library to mimic the pileup interactions happening along with the hard process.

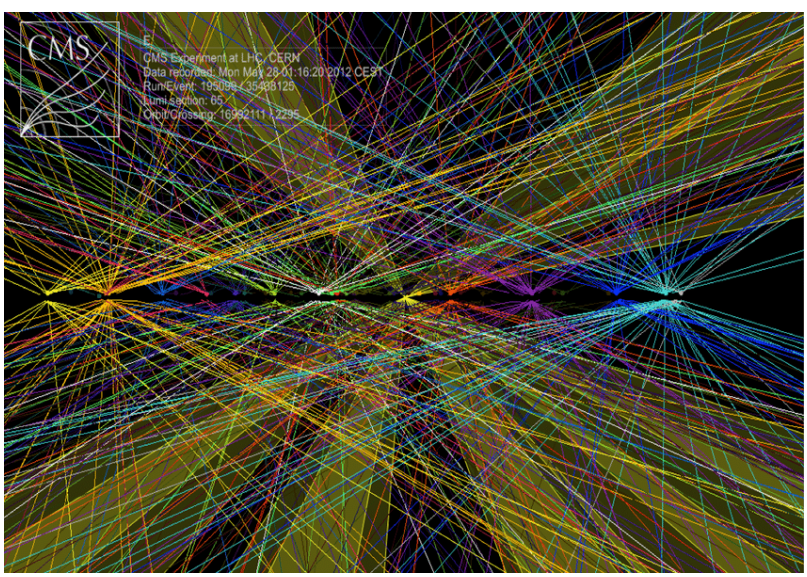

Figure 2.4: An event collected by CMS in 2015, Run/Event number 195099/35438125 showing the reconstructed vertices. Figure credit : CMS.

[4]Event generator for electroweak $pp \to \nu\nu$ process with random neutrino momenta that are back-to-back.

Chapter 3

Experimental Backgrounds

3.1 The Large Hadron Collider (LHC)

The LHC [40, 41] is a multi-purpose particle accelerator designed to conduct a proton-proton (pp) collision experiment, located at the European Organization for Nuclear Research (CERN). Fig. 3.1 shows the whole accelerator complex constructed at CERN. Protons are accelerated through the 26.7 km circumference ring around the border between Switzerland and France with an average depth of 0.1 km. The LHC is designed to collide protons at a center of mass energy $\sqrt{s} = 14$ TeV condition corresponding to a beam energy of 7 TeV per beam [1].

The underground tunnels were originally built for the Large Electron Positron (LEP) collider [42] and the LHC is sharing the investment that has been constructed long time ago. Using the experimental caverns in the underground have several advantages. It is not expensive compared to the high cost of the land

[1] The numbers written here are valid only for proton-proton collisions.

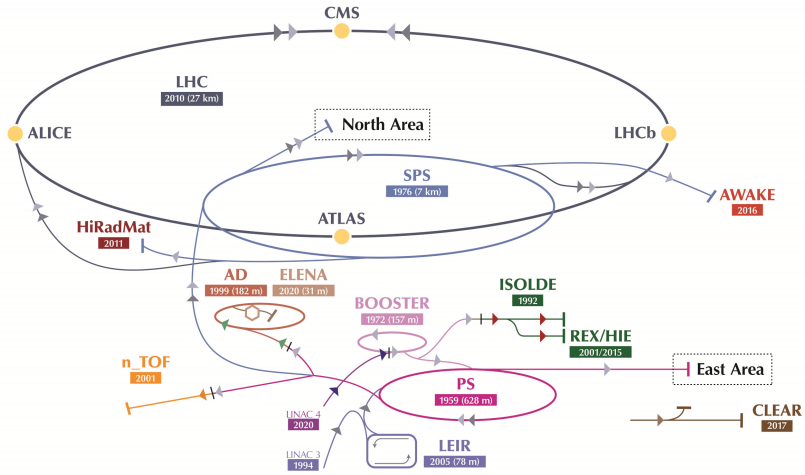

Figure 3.1: The accelerator complex of CERN in 2022. Smaller machines form a chain to boost the particles before being injected to LHC beam pipe. Figure credit : CERN.

above along the Jura mountains. But also it can naturally reduce the cosmic rays which can be considered as a noise for collision experiments due to its high energy and straightness of the tracks while it also serves as an effective shield with the radiations from the collisions at the LHC.

Before the protons are injected through the pipes of the LHC, it goes through the injection chain to acquire the necessary boosts. The boost of protons start from a hydrogen gas, being placed in a strong electric field being taken away their electrons through ionization. The injection chain starts from a linear accelerator named LINAC2 with an energy of 50 MeV. The protons are then fed into the Proton Synchrotron Booster which increases the energy from 50 MeV to 1.4 GeV. Then it goes through the Proton Synchrotron and the Super Proton Synchrotron which further accelerates the protons to reach 25 GeV and 450 GeV, respectively. At last, the beams are now ready to be injected to the LHC ring. 16 radio-frequency (RF) cavities further boost the

protons to the final collision energy 6.5 TeV which is the final beam energy for 13 TeV collsion [2]. With 40 MHz oscillating electric field induced by RF, the proton bunches are separated by 25 ns corresponding to 3564 bunches at the full intensity with 10^{11} protons per beam. With RF, the LHC is able to reach 6.5 TeV after 20 minutes of rotations around the ring.

The very first run of LHC started on 10th of September in 2008. However, after 10 days, magnet quenching incident occurred and caused a serious damage to more than 50 superconducting magnets and vacuum pipes. Thus first operational run happened on 3rd of March in 2010 with 3.5 TeV of proton beam energy. The energy was later elevated to 4.0 TeV in 2012 and was shut down on 13th of February in 2013 for the first upgrade plan *Long Shutdown 1, LS1*. The very first data taking years in 7 and 8 TeV collision energy is called *Run 1*.

After *LS1*, *Run 2* started on 5th of April 2015 after two years of resting with 13 TeV collision energy which was slightly lower than the original goal 14 TeV. This was due to the slow progress of magnet training which is to prevent the quenching effect. In 2016, LHC started to increase the luminosity for proton-proton collisions which was able to reached the design value on 29th of June. This improvement extended further and continued during *Run 2* until 17th of April on 2018. Due to the different proton beam conditions in 2015 and the rest of the years, *Run 2* dataset analyses use usually refer to the data collected in 2016, 2017, and 2018.

After *LS 2*, which was prolonged for quite some time due to COVID-19 crisis, LHC restarted on 22nd of April in 2022 with larger beam energy of 6.8 TeV. This *Run 3* is scheduled until 2026 and this will be the final *Run* before the High-

[2] Beam energy is set to 6.8 TeV for 13.6 TeV collision at the time of this thesis being written

Luminosity LHC (HL-LHC) era [43]. HL-LHC is a project ultimately aiming to take the integrated luminosity by a factor of 10 beyond the design value of LHC by elevating its performance further. This is expected to start from the beginning of 2029 which will allow physicists to measure more rare processes and seek more possibilities for BSM physics.

LHC hosts four different detectors along its beam pipe in Geneva. These detectors are situated at four beam crossing sites. The two main detectors which are designed for general purpose are ATLAS and CMS. These two detectors are designed to perform search for BSM physics and precision measurements of the SM physics. The other two detectors which are ALICE and LHCb are designed to study the quark-gluon plasma in heavy ion collisions and the to study the charge-parity violations and flavor physics with the b quark hadrons, respectively. All of these detectors were intentionally designed and operated by exclusive collaborators which enables mutual cross validations and also induces productive competitions for better science.

3.1.1 Luminosity

The potential of collider physics highly depends on the statistics it can accumulate for interesting physics processes. Measure of this can be expressed through the quantity, luminosity. The number of events N that a particular physics process occurs at a collider for a given interval of time t can be expressed as

$$\frac{dN}{dt} = \sigma \mathcal{L}_{\text{inst}} \tag{3.1}$$

where σ is the cross section of a given physics process. $\mathcal{L}_{\text{inst}}$ is the instantaneous luminosity and can be experimentally given as

$$\mathcal{L}_{\text{inst}} = \frac{n_b N_{p1} N_{p2} f}{A_b} R_{\text{corr}}. \tag{3.2}$$

Here n_b, N_p, f, A_b, and R_{corr} represents the number of colliding proton bunches, the number of protons in a bunch, the frequency of revolution, the area of beam quantifying its narrowness, and additional correction factor to account for the inclination for the beams to overlap at crossing sites, respectively.

The time integration of the instantaneous luminosity, the integrated luminosity $\mathcal{L}_{\text{int}}$ that is accumulated from CMS since the beginning of the LHC in 2010 is shown in Fig. 3.2. The peak value of $\mathcal{L}_{\text{inst}}$ that was recorded by CMS during Run 2 was around 2×10^{34} cm$^{-2}s^{-1}$ which exceeds the design value by a factor of two.

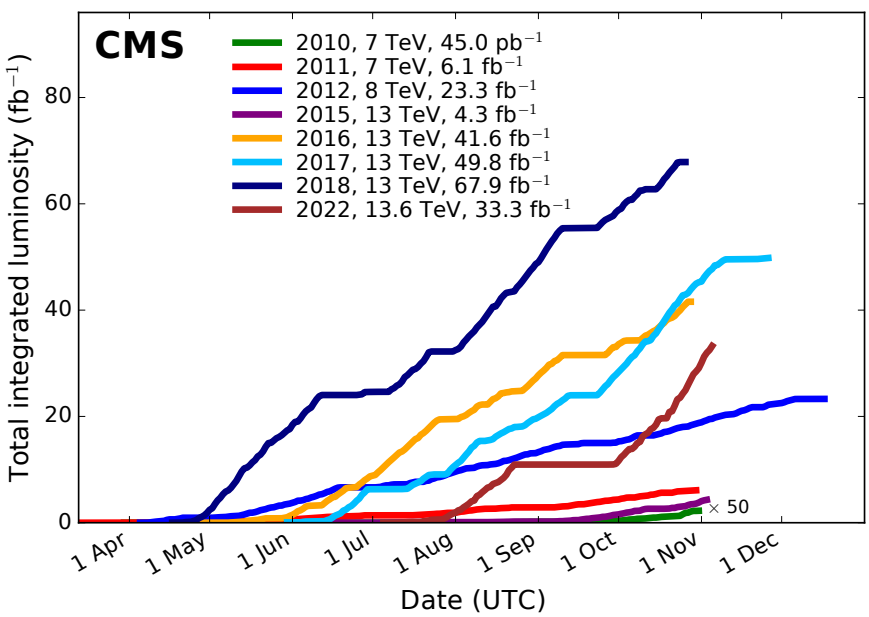

Figure 3.2: Cumulative luminosity versus day delivered to CMS for collisions at nominal center-of-mass energy. Figure credit : CMS.

3.2 The Compact Muon Solenoid (CMS) Detector

The CMS detector [44] is one of the two large general purpose detectors built on the LHC. Its very first goal was to discover the Higgs boson along with ATLAS and provide corroboration of findings for each other. This was achieved on 4th of July in 2012 as both experiments announced the discovery [11, 12]. After the discovery which was the last missing puzzle piece of the SM physics, goals for both ATLAS and CMS extended further : (1) Measuring properties of the Higgs bosons and SM physics processes with precision, (2) Exploring the physics at TeV scale and probe possibilities for undiscovered BSM physics.

The CMS detector can be depicted as an approximate cylindrical shape with the collision point situated in the center and symmetry axis laid along the beam axis. Its most characteristic feature is the central solenoid magnet which provides a large magnetic field of 3.8 T. This enables a precise measurement on transverse component of momentum p_{T} as it bends the track of charged particles. A 3D overview of the CMS detector is in Fig. 3.3. As illustrated, CMS is composed of smaller subdetectors each optimized to measure different types of particles. From innermost to outermost, pixel detector, silicon tracker, electromagnetic calorimeter, hadronic calorimeter, and muon calorimeter are placed.

3.2.1 Coordinate System

In order to describe the particles that emerge in the detector using Lorentz vectors, CMS adopts right-handed non-Euclidean coordinate system. The direction of positive x axis is defined to be the center of LHC ring and y axis points up toward the surface. Then one can find that z axis points to the counter clockwise direction of the ring when looked down from the sky.

The position of a particle can be parameterized in the polar coordinate as

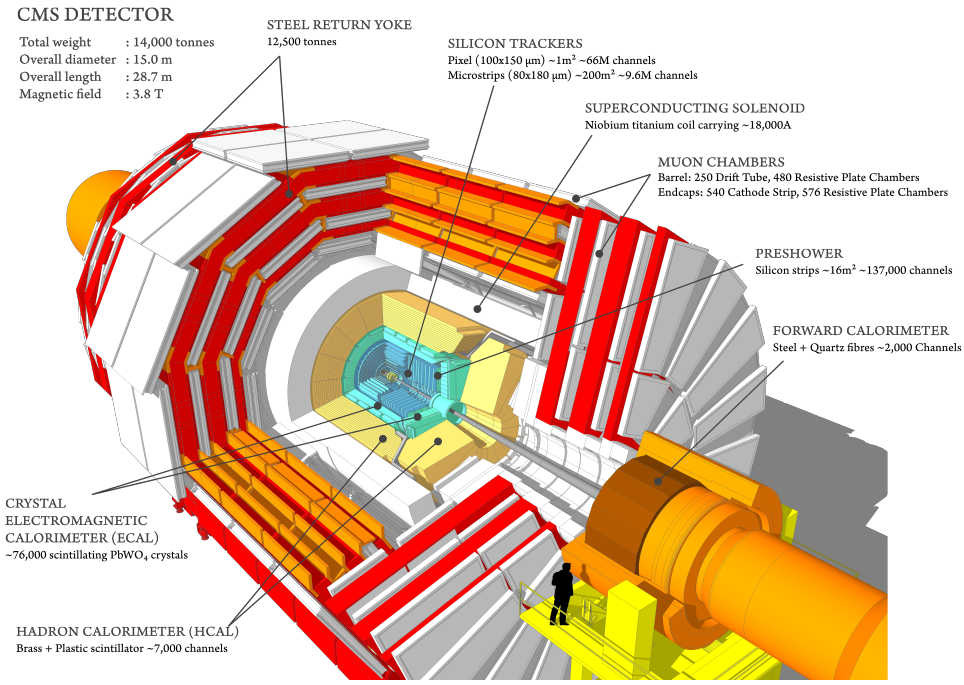

Figure 3.3: 3D overview of the CMS detector showing layers of components. Figure credit : CERN.

illustrated in Fig. 3.4. The azimuthal angle ϕ can be defined in the xy plane as the angle separation from the x axis while the polar angle θ is the angle separation from the z axis. From θ, the pseudorapidity η is further defined as

$$\eta = -\ln[\tan(\frac{\theta}{2})]. \tag{3.3}$$

This is invariant under Lorentz transformations along the longitudinal axis. As both θ and η are Lorentz invariant along the longitudinal axis, the angular separation between two different particles p_1 and p_2 is also Lorentz invariant, which is expressed as

$$\Delta R(p_1, p_2) = \sqrt{(\eta_1 - \eta_2)^2 + (\phi_1 - \phi_2)^2}. \tag{3.4}$$

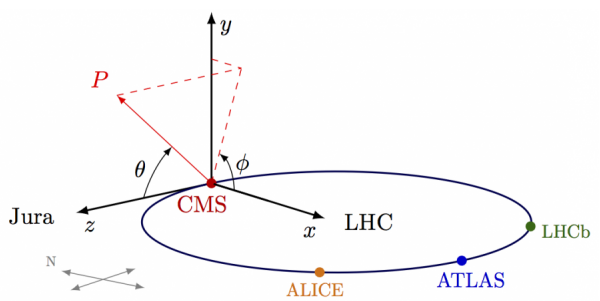

Figure 3.4: The coordinate system of CMS detector. Figure Credit : Izaak Neutelings.

3.2.2 Tracking System

The innermost layer of CMS is instrumented with silicon detectors [45] which is known for their speed and high granularity design. Its active area 200 m^2 is the largest silicon tracker ever built. The main purpose of tracking system is to precisely reconstruct the helical tracks of charged particles through the magnetic field. This system is crucial for two reasons : (1) The reconstructed

tracks serve as the basis of the particle-flow algorithm that will be discussed in Chapter. 3.3. (2) The particle track parameters need to be associated with interaction vertices, identifying pileup interaction points.

The module of tracking system which is closest to the beam axis at $r < 200$ mm is the pixel detector. In the high-multiplicity environment at the LHC, this area needs to deal with the highest particle flux and thus pixels are employed in small sizes, 100×150 μm^2. Going further away from the beam axis, silicon strip detectors are injected as the environment is not that much harsh with decreasing particle flux per area. The lengths of strip detectors are much bigger than those of pixels which is up to 10 cm while the widths vary from 80 to 180 μm. A schematic view of the system is shown in Fig. 3.5. In all layers of the system, the modules all rely on the p-n junction which creates the signal current when charged particles traverse and induces ionization electrons and holes.

3.2.3 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) [46, 47] is built right outside the tracking system, assembled with about $70,000$ lead tungstate ($PbWO_4$) crystals. This was chosen for its high density property (8.3 g/cm^3) and short radiation length (0.89 cm) that can well fit in the space constraints of the CMS detector inside the magnet coil. The length of crystals in barrel (endcap) are 23 (22) cm which corresponds to 25.8 (24.7) radiation lengths. However, because of its low light yield, it requires photodetectors with powerful amplification capability and that can tolerate harsh magnetic field environment. Thus, avalanche photo diodes are glued onto the back of the crystals, converting the light signals from crystals to amplified electric signals.

The crystals are set up in the barrel ($|\eta| < 1.444$) with $61,200$ crystals

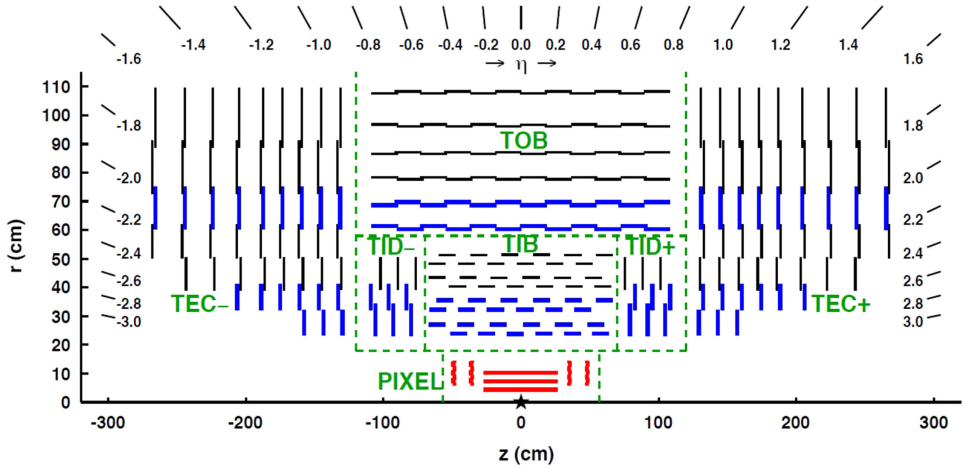

Figure 3.5: The tracking system of CMS detector. It is separated into multiple parts, including the pixel part in the center (PIXEL), tracker inner barrel (TIB), tracker inner disks (TID), tracker outer barrel (TOB), and tracker endcap (TEC±). The + and − signs for TEC corresponds to the z-axis coordinate of the detector. Figure credit : CMS.

and in both ends of the endcap with ($1.566 < |\eta| < 3.0$). The small η gap in between, absent with ECAL system, is for the cabling and supporting structure of the inner located tracking system. Each crystals are placed to face the center interaction point with the cross section 2.2×2.2 cm^2 at $r = 1.29$ m. This is because the Moliere radius of the $PbWO_4$ is 2.2 cm which enables it to contain 90% of energy in average. In front of the ECAL system in both sides of the endcap regions lie preshower sampling calorimeters designed to provide better granularity for the system. Its disk is 20 cm with two layers, composed of lead absorbers and silicon strip readout. Although the main measurement is performed by the $PbWO_4$ crystals, preshower becomes useful when trying to measure the locations of incident particles. Schematic view of the ECAL is shown in Fig. 3.6

The uncertainty in the energy measured with calorimeter is governed by the

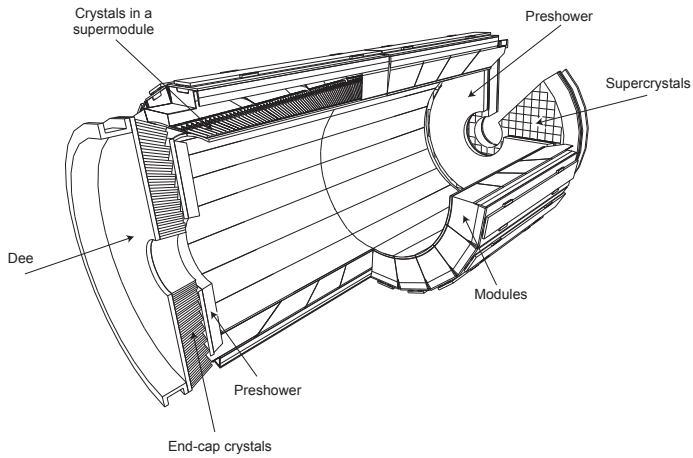

Figure 3.6: Schematic view of the CMS ECAL system. The figure is taken from Ref. [48].

equation [49]

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \tag{3.5}$$

where S represents the stochastic term depending on the number of scintillating photons occur, N is the noise term from the electronics or backgrounds from pileup interactions, and C represents the constant term caused by the non-uniformity of the crystal materials. During the very first CMS commissioning, values were measured to be $S = 2.8\%\ \mathrm{GeV}^{1/2}$, $N = 0.12\%\ \mathrm{GeV}$, and $C = 0.3\%$.

3.2.4 Hadronic Calorimeter

The hadronic calorimeter (HCAL) lies outside the ECAL, split into four regions : the barrel (HB), endcap (HE), forward (HF), and outer (HO) regions (see Fig. 3.7). Unlike the ECAL, HB and HE are sampling calorimeters which cover the intervals $|\eta| < 1.3$ and $1.3 < |\eta| < 3.0$, respectively. designed with alternating layers of absorbing and detecting materials. Brass absorber plates are used as absorbing materials that initiate hadronic showers by inducing an

interaction between the nuclei of the material and the propagating particle. The plastic scintillators which consist of multiple silicon photomultipliers that sample the hadronic showers from brass plates.

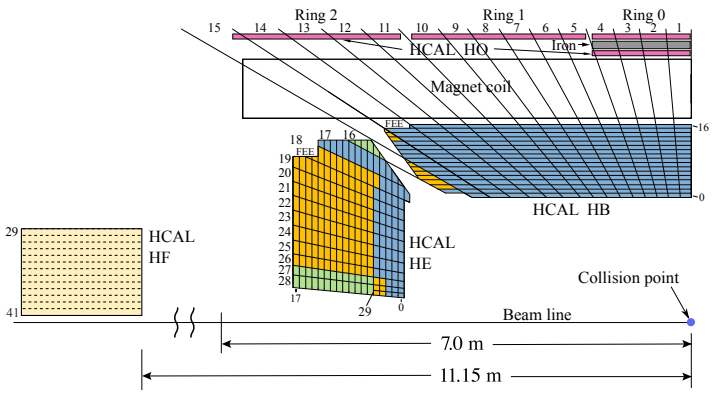

Figure 3.7: Schematic view of the CMS HCAL system. The figure is taken from Ref. [50].

Very close to the beam pipe in the region $3.0 < |\eta| < 5.0$ lies HF which is 11.2 m away from the interaction point. As HF is exposed to harsh radiation environment, it is designed with radiation resilient steel for inducing showers which are captured by quartz fibers as Cherenkov light. Further away outside the magnet coil in the barrel region is inserted with HO in order to compensate for its short absorber length. This addition guarantees the absorption of residual hadronic showers before escaping from the detectors.

As for ECAL, the energy resolution of HCAL is also tested before CMS started taking data. With the energy resolution equation in Eq. 3.5, the values were measured to be $S = 115\%$ GeV$^{1/2}$ and $C = 5.5\%$ with N being negligible. As one can easily notice, HCAL has very poor energy resolution compared to ECAL which makes the PF algorithm more crucial when reconstructing the jets.

3.2.5 Muon System

As the name *Compact Muon Solenoid* already suggests, the muon system is considered to be the most essential part of the detector. It is located at the outermost layer of the CMS detector covering the area or $|\eta| < 2.4$ and $r = 7.4$ m. With its minimally interacting nature, muons with $p_{\mathrm{T}} > 1$ GeV are very much stable and are expected to be the only type of particle that can reach the farthest unlike other particles that stops propagating in the middle of the detector. Thus, the muon system can still function with the delicate technique of detecting ionized free electrons from gas when muons pass through. Three different types of muon detectors, shown in Fig. 3.8, are used to form the whole muon system but they all employ the same technique. The tracks can be inferred from the position and timing information that is given from the free electrons drifting in the electric field to the edges where the read-out electronics are placed at.

Up to $|\eta| < 1.2$, the barrel region of the muon system is instrumented with drift tubes (DT) [52] modules. The system consists of five wheel compartments along the z axis. Each wheels can be further divided into twelve different sectors in ϕ coordinate each corresponding to $\Delta\phi = 30°$ which contains four DT stations. Each DT station is built with eight and four layers that are perpendicular and parallel to the beam pipe, respectively. This allow the muon hits to be recorded in 3D coordinate system that is (r, ϕ, z). At the center of DT module, anode wire is inserted where the ionized electrons drift toward to when muons pass through. The drift time of ionized electrons can maximally be 400 ns which is a bit longer than the bunch crossing gaps. However, given that both muon and background rates are small, DTs are still shown to be useful with great position resolution.

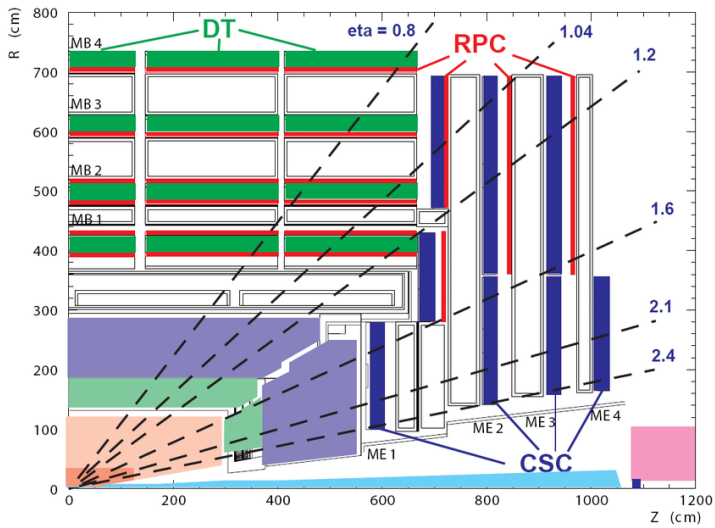

Figure 3.8: Schematic view of the CMS muon system. DT, CSC, and RPC are colored in green, blue, red, respectively. The figure is taken from Ref. [51].

On the other hand, the region in $0.9 < |\eta| < 2.4$ needs a system with shorter drift path as the rate of particles that emerge in this region is much higher. Thus instead of DT, cathode strip chambers (CSC) [53] that has advantage with faster response time and finer segmentation is used in this region. CSC is a multi-wire chamber of trapezoidal shape that has copper cathode strips radially going outwards and anode wires in the orthogonal directions. The position resolutions of DT and CSC are about 100 μm scale for both system depending on the positions.

To correctly assign muons to the corresponding bunch crossings, resistive plate chambers (RPC) [54] with great timing resolution is installed in overall range $|\eta| < 1.6$. RPC is made up of two parallel plates that is filled with gap in between. Traversing muons ionize the gas, emitting electrons sequentially which leads to an avalanche of electrons. Although its position resolution is around

1 cm, 1 ns of timing resolution is a great complementary to other muon systems.

3.2.6 Data Acquisition

From the bunch crossings of proton beams at 25 ns, CMS expects the data to be delivered at the rate of 40 MHz level. With the large number of multiple subsystem, a single event that CMS detector takes amount to a data size about 1 Mb. Combined with the rate, this means that the CMS needs a bandwidth over 40 Tb/s which definitely exceeds the feasible processing rate of modern computing technology. In addition to the unfeasible processing rate, even if it is possible to process all these events, large fraction of it is from soft QCD events which are not so much interesting with regards to CMS detector's purpose. Thus CMS uses a trigger system [55] with hardware and software technologies to identify physics-wise interesting events that are worth being analyzed in depth.

The trigger system starts with Level 1 (L1) trigger which is based on the data acquired from hardware, collected from the calorimeters and the muon system. As the track reconstruction algorithm needs long time to run, the information provided from the tracker system is not used. However, the track reconstruction in the muon system is deployed as it has significantly less number of hits in the system. With all these in account, fast data processing rate is required as the event rate of 40 MHz needs to be reduced to 0.1 MHz level only based on hardware decisions. This decision includes various physics signatures such as single quality muon with $p_{\mathrm{T}} > 22$ GeV, two jets with $p_{\mathrm{T}} > 150$ GeV and $|\eta| < 2.5$, etc.

Events chosen from L1 trigger decisions are passed to the software-based High-Level Trigger (HLT) which performs much detailed filtering of data by reconstructing the physics objects, this time including time consuming track

reconstruction. The events are first identified with the criteria that enabled it to pass the L1 trigger and are further tested with multiple HLT criteria that corresponds to the decisions from L1 trigger. For example, if an event is passed to HLT by containing a single quality muon with $p_{\mathrm{T}} > 22$ GeV from L1, the event is tested with a soup of HLT criteria that necessarily requires a single muon with $p_{\mathrm{T}} > 22$ GeV such as a trigger that needs a HLT level reconstructed muon with good isolation and $p_{\mathrm{T}} > 24$ GeV.

The sequence that L1 and HLT forms is called a trigger path (see Fig. 3.9) and the event is kept in the storage system only if it is accepted by at least one trigger paths. From this step, the event rate decreases to a level of 1 kHz which makes it possible to globally reconstruct the event at a quality in which physics analysis can be performed.

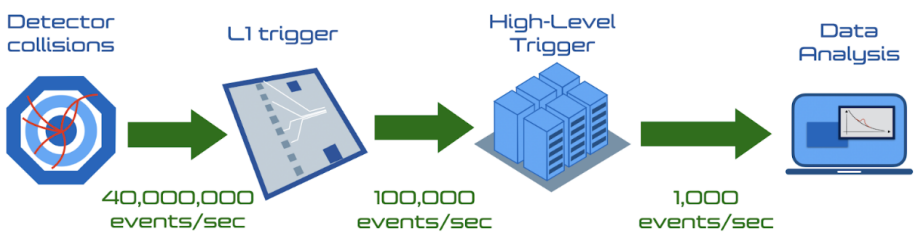

Figure 3.9: Depiction of the data processing flow of the CMS experiment. Among 40 M events occuring per second, about a thousand events are selected by the trigger system for most of the physics analysis purposes. The figure is taken from Ref. [56].

3.3 Particle-Flow Algorithm : Event Reconstruction

To reconstruct the events with physics objects, CMS adopts the Particle-Flow (PF) algorithm [57]. The goal of this algorithm is to collect the information acquired from various subdetectors and combine them to a physics object in

an optimal way. Thus it starts with reconstructing and identifying all stable particles in the event that are electrons, muons, photons, and charged and neutral hadrons which are usually referred as PF candidates. These individual blocks are used as if it came from MC event generator, to build physics objects including the missing transverse energy so that it could be used for physics analysis. Such approach allows to harvest the very best fruits from each detector components and combine them into one physics object.

Schematic illustration of each type of particle's energy deposits and trace of tracks are shown in Fig. 3.10. This already briefly describes how PF candidates are reconstructed. Muons are reconstructed by linking the track formed in tracking system and the detector hits in muon system. Electrons also form tracks while leaving energy deposits in ECAL. Photons on the other hand leave ECAL energy deposits but has no tracks as its charge neutral. Charged hadrons correspond to tracks associated with energy deposit clustered in HCAL. Neutral hadrons only leave energy in HCAL alone.

3.3.1 Tracks and Primary Vertex

PF algorithm starts with reconstructing the tracks of charged particles that have left its trace in the tracking system. The tracks are built based on Kalman Filtering (KF) method [58] that are seeded with two hits in consecutive layers in the pixel detector. From the seeds, KF method extrapolates the tracks to the outer tracker layer one by one which is performed by using the best estimate of a track momentum and its associated bending at the given moment. For each step, energy deposit to materials is also taken into account from the Bethe-Bloch equation and its energy deposit is given with uncertainty when finding the direction to the next layer. For the tracks to be stored, it is required to have at least eight hits in the tracker system (with each hits contributing less than

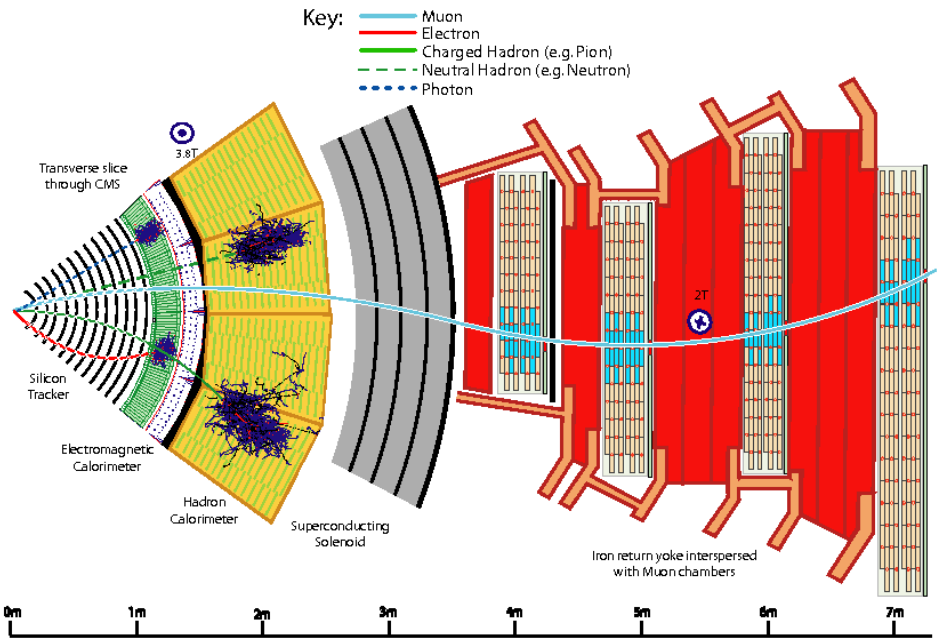

Figure 3.10: The sketch of a specific particle's interactions as it propagates through the CMS detector starting from the interaction point. Cyan, red, and green lines each represent muon, electron, charged hadrons. Green and blue dashed lines each represent neutral hadron and photon. Figure credit : CMS.

30 % of the overall track goodness-of-fit). The tracks are additionally required to be originating within few mm distance from the beam axis and to have $p_{\mathrm{T}} > 0.9$ GeV.

Reconstructed tracks are promptly used to reconstruct the interaction vertices which corresponds to the position where coincidentally interaction occurred, having the same origin of the tracks. Among the vertices in the event, $\sum p_{\mathrm{T}}^2$ maximizing vertex, which is the square sum of AK4 [3] clustered jets from tracks and missing transverse energy of the vertex from the jets, is chosen as the primary vertex. The primary vertex is considered to be the vertex associated with the hard process that is of physics interest. Other interaction vertices are created due to pileup interactions or possibly long-lived particles. Simulated MC samples for the analysis in Chapter. 5 is shown to select the primary vertex more than 99 % of the time correctly.

3.3.2 Muons

Muons are reconstructed using the information on tracks from tracking system, muon segments which are created by fitting the hits across different chambers in muon system, or both. Muon is classified in three different types based on how it is reconstructed :

- Standalone muon : Reconstruction starts from DT and CSC segments as seeds additionally using hits in RPC system. This type of muon solely depends on the peformance of the muon system.

- Tracker muon : Tracks reconstructed from the tracking system are extrapolated in outward direction to the muon system. Compatible segments need to be found in the muon system.

[3] Definition of AK4 will be discussed in Chapter. 3.3.4.

- Global muon : Standalone muons are propagated inwards to find matching tracks from the tracking system. A new fit on combined set of tracks and hits is performed for the global muons.

In addition to muon object reconstruction, for muons with $p_{\mathrm{T}} > 200$ GeV goes through the tune-P algorithm which combines all three types of muon information to calculate the optimal p_{T}. For $p_{\mathrm{T}} < 200$ GeV, almost 99.5 % of the cases favors the p_{T} from the inner tracker decision.

3.3.3 Electrons and Photons

Electrons are reconstructed by combining with the information from tracking system and ECAL energy deposits. The difficulty of electron reconstruction comes from sizable bremssstrahlung photon radiations from electrons within the tracking system. This causes electrons to radiate its energy before propagating to the ECAL system which also hinders the track finding. To find the tracks of electrons, the Gaussian Sum Filter (GSF) algorithm is used which is known to be more suitable than the KF algorithm in case of building the trajectory of radiating particles. In addition, instead of using the ECAL energy deposit in a single cell, supercluster (SC) is built to encompass several clusters in ϕ direction, aggregating neighboring cells to mitigate the energy loss due to photon radiation.

Electrons are then reconstructed with the matching set of GSF tracks and SCs in ECAL. The kinematics of the electron is decided by weighted linear combination from these two information. For electrons with $p_{\mathrm{T}} > 20$ GeV, the measurement from ECAL dominates with negligible contribution from GSF tracks but shows about 2 % improvement in resolution with lower p_{T} electrons. Photons are simultaneously reconstructed along with electrons by requiring the GSF tracks to be absent while being seeded only from ECAL clusters.

3.3.4 Jets

From the self interacting nature of gluons in QCD, particles that can go through strong interactions easily produce a shower while it propagates through the detector. As many of the interesting physics lies in the domain of QCD, it is necessary to reconstruct showered hadrons that originate from quarks and gluons with a proper method. Such task is achieved by utilizing the jet clustering algorithm (see Fig. 3.11 as an example) with PF candidates using the FastJet package [59].

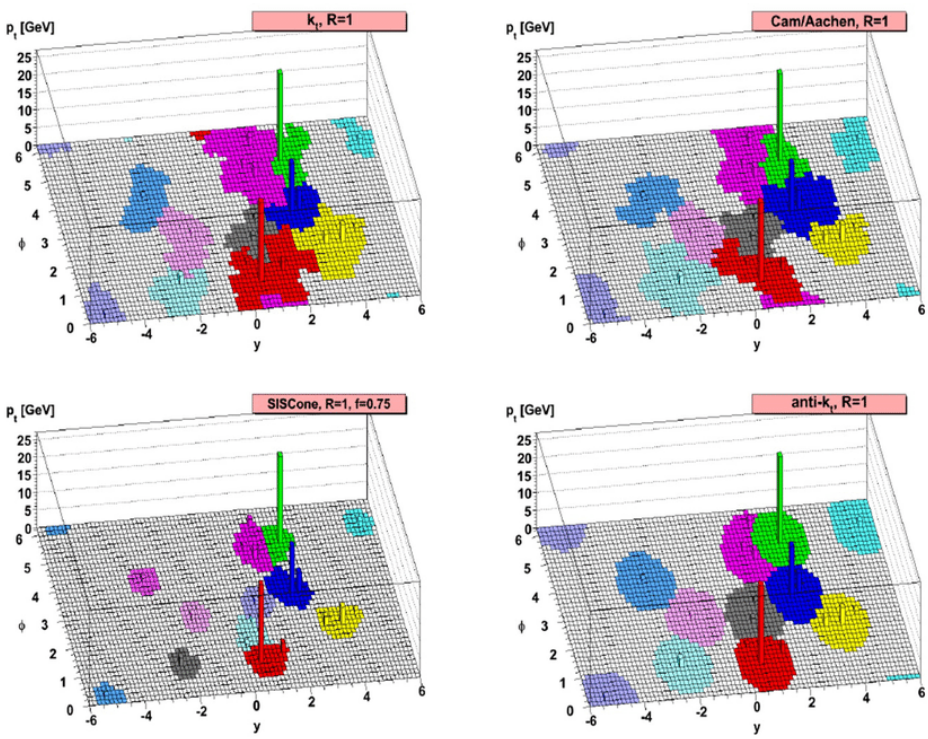

Figure 3.11: An illustrative diagram of the four main jet reconstruction algorithms. The jet reconstructions are performed on the same data with same radius parameter. Note that Anti-k_T algorithm creates conical shape jets with preference for hard radiations. The figure is taken from Ref. [60].

Among many jet clustering algorithms, CMS uses Anti-k_T (AK) algorithm [61] with a distance parameter of $R = 0.4$ for small-radius AK4 and $R = 0.8$ for large-radius AK8 jets. Jet clustering in the past, such as cone iteration algorithm, was shown to be collinear unsafe as a collinear splitting of QCD particle in the event could lead to different jet reconstruction outcomes. Nowadays, including AK algorithm, jet clustering algorithm sequentially clusters two objects into one composite object based on the comparison between weighted distance parameter

$$d_{ij} = \min(k_{Ti}^d, k_{Tj}^d)\frac{\Delta R_{ij}^2}{R^2} \tag{3.6}$$

and $d_{iB} = k_{Ti}^d$ defined as the p_{T} of the object with respect to the beamline. Here, k_T is the p_{T} of the object and ΔR_{ij} is the angular distance between the two objects. Most importantly, the choice of d parameter defines which jet clustering algorithm is adopted. AK algorithm is the case when $d = -2$, which effectively starts clustering objects with the hardest objects in the event, being resilient to soft QCD radiations. The algorithm first computes all d_{ij} and d_{iB} with all possible combinatorics and align them in increasing order. If the smallest among all is d_{iB}, the $i-$th object is defined as a jet and removed from the array. If d_{ij} is the smallest, $i-$th and $j-$th objects are clustered together and the whole calculation repeats until all objects are taken out as jets from the array.

One should be mindful on the fact that jet clustering is performed on all PF candidates, not solely on HCAL energy deposits which is a proxy for QCD particles. This genuinely means that the PF jets could even originate from muons, electrons, or photons. For some even more extreme cases, especially in MC event generators or Delphes which is a fast simulation module, neutrinos or BSM particles that leaves the detector without leaving a trace are also considered as jets as the genuine definition of jet does not come from the detector signals. Instead,

as long as there is energy that can participate in the clustering algorithm, there is a chance that it can be reconstructed as jets. Thus for practical use, jets are normally cleaned up by nearby leptons and the remaining jets are considered as jets originating from quarks or gluons.

3.3.5 Taus

Being the heaviest of leptons, taus have gained great interest at the LHC. This is also connected with the Higgs boson studies, havinga large Yukawa coupling as it has a mass of nearly 1.8 GeV. Taus do not only decay into light leptons and neutrinos but as a matter of fact, nearly 65 % is the branching ratio of taus to hadronic final states. As hadronically decaying taus have charged hadrons that would yield a signal through HCAL energy deposit, it is likely to be reconstructed as jets.

Hadronically decaying taus are reconstructed through the hadrons-plus-strips algorithm performed on PF jets [62]. Hadrons refer to the charged hadrons that hadronically decaying taus carry, likely to be one to three per tau, and strips refer to neutral pion decays which promptly decays into photons but also possible to convert in to e^+e^- pairs as they traverse through the material. Due to the large magnetic field, a spatial separation of the daughter particles of neutral pions are expected to be bound to specific η window while spanned in ϕ direction, forming a strip.

Chapter 4

Neutrinos

Although there are many kinds of fermions which are considered as fundamental particles in the SM, neutrino is outstandingly special and unique compared to others. Neutrinos are the only fermions without electric charges and with mass of exactly 0 GeV in the SM. Due to such properties, neutrinos only interact through weak processes and have the same helicity and chirality. In this section, we will discuss the history of neutrino which was the main driving force of the particle physics field since 1930. Neutrinos provided the chance of rich and deep development for both experiment and theory field.

4.1 History of Neutrinos

4.1.1 Radioactive Beta Decay

The story of neutrinos starts from the conundrum of radioactive beta decays [63]. At the time before neutrinos were not even postulated as a fundamental particle of the universe, beta decay process was only known as a transition $\mathbf{X}(Z, N) \to \mathbf{X}'(Z+1, N-1) + e^-$ where $\mathbf{X}(Z, N)$ denotes the nucleus with Z

protons and N neutrons. Through this process which was yet considered as a two body decay (before neutrinos were predicted), the law of energy conservation requires

$$E_X = E_{X'} + (K_e + m_e c^2) \tag{4.1}$$

which in turn becomes

$$K_e = E_X - E_{X'} - m_e c^2 \simeq (m_X - m_{X'} - m_e)c^2 \tag{4.2}$$

with the assumption that the nucleus before the decay $\mathbf{X}$ is at rest and electron being a much lighter particle than $\mathbf{X}'$. Eq. 4.2 shows that the kinetic energy of the electron K_e would have an uniquely determined value that is derived from the mass differences of nuclei involved in the process.

However surprisingly, the experimental results from Lise Meitner and Otto Hahn in 1911 suggested otherwise. It reported that electrons from the beta decay have a continuous energy spectrum instead of a fixed value at certain point. This lead to many controversial ideas, for example Lise Meitner proposed that this was due to secondary effect which makes electron lose its energy while being disintegrated from the nucelus. However, by Charles Ellis and William Wooster in 1927 from the experiment measuring the radium E (bismuth-210) with calorimetric technique, it was conclusively shown that there is no secondary effect. As shown in Fig. 4.1 the energy spectrum from the measurements give maximum value at 1.05 MeV with 0.39 MeV in average. If what Lise Meitner suggested were to be true, calorimeter system which encapsulates the whole beta decaying source should've been measured in average 1.05 MeV.

After this finding, they both opened up but also rejected the idea of energy non-conservation which violates one of the fundamental laws of physics at the same time as written in Ref. [64] :

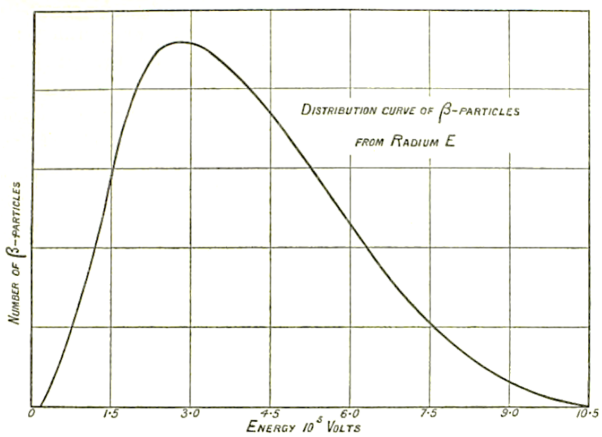

Figure 4.1: Beta decay spectrum measured by Ellis and Wooster from bismuth-210 using calorimeter. The figure is taken from Ref. [63].

> *If we were to consider the energy to be conserved only statistically there would no longer be any difficulty in the continuous spectrum. But an explanation of this type would only be justified when everything else had failed, and although it may be kept in mind as an ultimate possibility, we think it best to disregard it entirely at present.*

This idea of energy non-conservation in beta decays was still pursued by Niels Bohr and a note was sent to Wolfgang Pauli - which was replied with very negative remarks and ultimately never got published [65].

4.1.2 A Desperate Remedy

The famous letter in 1930 [66], opening with *Liebe Radioaktive Damen und Herren* in German which translates to *Dear Radioactive Ladies and Gentlemen* in English, was sent by Wolfgang Pauli to the participants of a workshop in Tübingen (Germany) (See Fig. 4.2). Although a bit informal with excuse of

his absence at the workshop, it had a groundbreaking importance in terms of physics. This was the very first document proposing the existence of a new type of the particle which now is known as neutrino.

Instead of supporting Niels Bohr's idea of energy conservation violating beta decay, Wolfgang Pauli hypothesized a new particle neutron [1] that has not been yet found carrying away the energy during beta decay but undetected. Wolfgang Pauli called this *a desperate remedy to keep the energy conservation law applicable in all physics systems*. Such interpretation was valid for the calorimeter experiment using bismuth-210 by Charles Ellis and William Wooster [64] if this new particle has a larger penetration rate than γ rays and was responsible for the missing energy in the spectrum.

This idea was later picked up in 1934 by Enrico Fermi with his theory of beta decay as the base of the weak interaction we know of today. However, his work was first rejected by the journal *Nature* with negative comments saying *too remote from reality* and in the end had to be submitted to an Italian journal. Although the idea seemed to be *too remote from reality*, experimental results of beta decays from many different atomic nuclei showed beautiful agreements to the calculations Enrico Fermi suggested, making neutrinos more realistic.

4.1.3 Observation of Neutrinos

Frederick Reines and Clyde Cowans were tempted to verify the existence of neutrinos by directly observing (or detecting) it instead of taking indirect hints from energy distribution calculations. One of the main difficulty was of course the tiny cross sections of the processes that involve neutrinos. This could be solved by considering two factors of the experiment : (1) A source that provides an enormous amount of neutrinos. (2) A detector that has abundant amount of

[1] It later changed to the name neutrino after James Chadwick discovered the neutral nuclear particle in 1932.

Original - Photocopie of PLC 0393

Abschrift/15.12.56 PM

Offener Brief an die Gruppe der Radioaktiven bei der Gauvereins-Tagung zu Tübingen.

Abschrift

Physikalisches Institut
der Eidg. Technischen Hochschule
Zürich

Zürich, 4. Des. 1930
Gloriastrasse

Liebe Radioaktive Damen und Herren,

Wie der Ueberbringer dieser Zeilen, den ich huldvollst anzuhören bitte, Ihnen des näheren auseinandersetzen wird, bin ich angesichts der "falschen" Statistik der N- und Li-6 Kerne, sowie des kontinuierlichen beta-Spektrums auf einen verzweifelten Ausweg verfallen um den "Wechselsatz" (1) der Statistik und den Energiesatz zu retten. Nämlich die Möglichkeit, es könnten elektrisch neutrale Teilchen, die ich Neutronen nennen will, in den Kernen existieren, welche den Spin 1/2 haben und das Ausschliessungsprinzip befolgen und sich von Lichtquanten ausserdem noch dadurch unterscheiden, dass sie nicht mit Lichtgeschwindigkeit laufen. Die Masse der Neutronen müsste von derselben Grossenordnung wie die Elektronenmasse sein und jedenfalls nicht grösser als 0,01 Protonenmasse.- Das kontinuierliche beta- Spektrum wäre dann verständlich unter der Annahme, dass beim beta-Zerfall mit dem Elektron jeweils noch ein Neutron emittiert wird, derart, dass die Summe der Energien von Neutron und Elektron konstant ist.

Nun handelt es sich weiter darum, welche Kräfte auf die Neutronen wirken. Das wahrscheinlichste Modell für das Neutron scheint mir aus wellenmechanischen Gründen (näheres weiss der Ueberbringer dieser Zeilen) dieses zu sein, dass das ruhende Neutron ein magnetischer Dipol von einem gewissen Moment μ ist. Die Experimente verlangen wohl, dass die ionisierende Wirkung eines solchen Neutrons nicht grösser sein kann, als die eines gamma-Strahls und darf dann μ wohl nicht grösser sein als $e \cdot (10^{-13}$ cm).

Ich traue mich vorläufig aber nicht, etwas über diese Idee zu publizieren und wende mich erst vertrauensvoll an Euch, liebe Radioaktive, mit der Frage, wie es um den experimentellen Nachweis eines solchen Neutrons stände, wenn dieses ein ebensolches oder etwa 10mal grösseres Durchdringungsvermögen besitzen würde, wie ein gamma-Strahl.

Ich gebe zu, dass mein Ausweg vielleicht von vornherein wenig wahrscheinlich erscheinen wird, weil man die Neutronen, wenn sie existieren, wohl schon längst gesehen hätte. Aber nur wer wagt, gewinnt und der Ernst der Situation beim kontinuierliche beta-Spektrum wird durch einen Ausspruch meines verehrten Vorgängers im Amte, Herrn Debye, beleuchtet, der mir kürzlich in Brüssel gesagt hat: "O, daran soll man am besten gar nicht denken, sowie an die neuen Steuern." Darum soll man jeden Weg zur Rettung ernstlich diskutieren.- Also, liebe Radioaktive, prüfet, und richtet.- Leider kann ich nicht persönlich in Tübingen erscheinen, da ich infolge eines in der Nacht vom 6. zum 7 Des. in Zürich stattfindenden Balles hier unabkömmlich bin.- Mit vielen Grüssen an Euch, sowie an Herrn Back, Euer untertänigster Diener

gez. W. Pauli

Figure 4.2: The letter from Pauli. The figure is taken from Ref. [66].

detector materials. Their first experimental stage in mind was somewhere close to the nuclear fission bomb explosion sites which later they changed their mind to a more peaceful and repeatable area, near the nuclear reactors.

With the inverse beta decay $\overline{\nu_e} + p \rightarrow n + e^+$ process [2] being expected as the reaction in order to be detected, 400 L of water and cadmium chloride ($CdCl_2$) was filled in the tank. It was designed in two layers to first detect the protons from water interacting with neutrinos, creating neutrons along with two photons from positron annihilation that were to be recorded with photomultiplier tubes. The neutrons were captured by Cd nucleus which absorbs the neutrons and then emitting gamma rays to become stable again few microseconds after the positron annihilation. They repeated their measurements also after shutting down the reactor to see if there was a difference in the event rates between turning it on and off. Finally, they were able to measure the cross section 6.3×10^{-44} cm^2 which was predicted to be about 6×10^{-44} cm^2 in 1956 [67].

4.1.4 Solar Neutrinos

Around mid 1950s when Reines and Cowan were executing their experiments, Ray Davis also attempted to detect the neutrinos [68]. His approach was slightly different from the two, utilizing the $Cl - Ar$ detector concept proposed by Bruno Pontecorvo which works from the reaction process that is $\nu + {}^{37}Cl \rightarrow {}^{37}Ar + e^-$. The idea was to count the number of Ar atoms in the detector after separating it (by boiling the liquid) from carbon tetracholoride (CCl_4) to find out how many neutrinos interacted in the detector. This was refined by Ray Davis, using perchloroethylene (C_2Cl_4) and bubbling the He gas through the liquid to flush out the Ar. But his experiment failed to detect the neutrinos, being only sensitive to neutrinos but not antineutrinos as it relies on the reaction

[2] Note that neutrinos were not known to have its antiparticle and flavors at that time yet.

$\nu_e + n \rightarrow e^- + p$ process.

However, Ray Davis did bounce back by facing greater task and challenge that was the solar neutrino detection. He proposed a $380,000$ L detector filled with C_2Cl_4 in the deep underground to avoid contamination from cosmic rays [69]. The proposal was guided by John Bahcall, a theorist, who calculated the neutrino emission rate from the Sun and the capture rate of Cl in the detector. This took place in the Homestake gold mine which was $1,478$ m underground. The technique to observe neutrinos was the same as above, counting Ar atoms which occurred through the process $\nu_e + n \rightarrow e^- + p$ with neutrino energy above the threshold of 0.814 MeV [70].

Although the concept of the experiment was simple as before, and the fact that neutrinos also have its antiparticle and carry flavors were well established by the time the detector was built in 1967, another problem was waiting for Ray Davis and John Bahcall. Davis's measurements were only one third of Bahcall's theoretical prediction. Bahcall checked his calculation while Davis scrutinized his measurements, but both failed to find errors in their own works. This problem was referred to as the *solar neutrino problem*, deficit of solar neutrino observation compared to the solar model prediction, and was left unsolved until 2001.

Again Pontecorvo suggested that the solar neutrino problem was due to the phenomenon now known as neutrino oscillation based on the idea that flavor and mass eigenstates of neutrinos are not identical, allowing neutrinos to transform into other flavors while it propagates. The SNO experiment was able to address this by designing the detector with heavy water that can be sensitive in two ways : (1) Neutrinos with all flavors. (2) Neutrinos with electron flavor. Measuring the deficit of electron flavor neutrinos and the consistent flux of neutrinos of whole flavor simultaneously would suggest the neutrinos oscillation. In 2001,

SNO [20] confirmed that the solar neutrinos oscillate as it propagate to the Earth.

4.2 Neutrino Mass Models

Neutrinos (anti-neutrinos) have only been observed with left-handed (right-handed) helicity state. This fact was well incorporated to Glashow, Weinberg, and Salam's theory of electroweak interactions (the unification of the weak and electromagnetic interaction) when it was formulated with the assumption that neutrinos are massless. However, the observation of neutrino oscillations was a conclusive proof that the theory needs to be tweaked at least with some minimal modifications to generate the Dirac mass of neutrinos in order to distinguish different neutrino mass eigentstates.

Dirac mass of neutrinos cannot be acquired unlike the other charged fermions after the electroweak symmetry breaking as it only exists with left-handed helicity state. Though by brutally postulating neutrinos with right-handed helicity state, assuming it has not yet been found but does exist somewhere, Dirac mass term can be generated to explain the tiny masses of neutrinos ($\leq 1\,\mathrm{eV}$). This automatically leads to the prediction of tiny Yukawa coupling values to be below 10^{-12} which is largely different from those of other fermions. Since there is no theoretical reason for it to be much smaller for neutrinos, such scenario is not so much appealing from experimentalist perspective.

Among many BSM theories that were proposed to explain the nonzero masses of neutrinos, the *seesaw mechanism* is one of the most popular. It was proposed by Weinberg that the effective dimension 5 operator of the form $LLHH$ where L and H are the SM left-handed lepton doublet and Higgs doublet, respectively. Consequently, it was also shown that there are only three possibilities to enforce its ultraviolet completeness at the tree level. These set

of three options are the famously known type-I, type-II, and type-III seesaw mechanisms [71, 72]. As shown in Fig. 4.3, depending on how $LLHH$ is realized, either a fermion singlet N (type-I), a scalar triplet $\Delta = (\Delta^{\pm\pm}, \Delta^{\pm}, \Delta^{0})$, or a fermion triplet $\Sigma = (\Sigma^{+}, \Sigma^{0}, \Sigma^{-})$, gives rise to the SM neutrino mass.

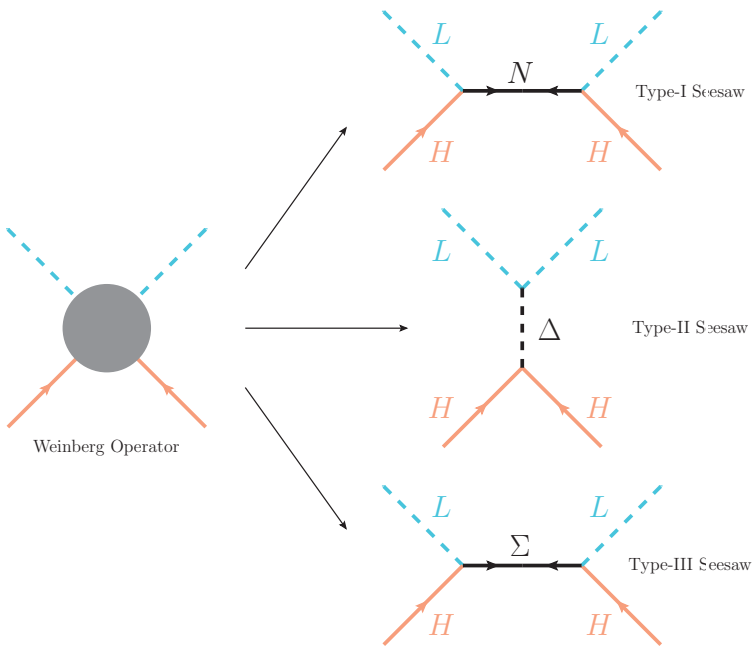

Figure 4.3: Weinberg operator and its opening to three different seesaw models.

4.2.1 Type-I Seesaw Model

Phenomenological implementations of type-I seesaw model [73, 74] is an effective treatment parameterizing the flavor mixing and the mass eigenstates. In this thesis, for simplicity, we only consider one mass eigenstate N that might be accessible by LHC experiments (and if someone or some experiment succeeds to find it, this can be expanded to two or more N eigenstates). Then, a SM

neutrino of flavor ℓ in the interaction basis ν_ℓ^L which is postulated to be left-handed can be expressed as

$$\nu_\ell^L = \sum_{i=1}^{3} U_{\ell i}\nu_i + V_{\ell N} N. \tag{4.3}$$

Here, U is the complex valued mixing matrix and ν_i, where $i = 1, 2, 3$, are mass eigenstates, and similarly, $V_{\ell N}$ denotes the mixing value of ν_ℓ^L and N. Within this formalism, it is not too difficult to consider $|U|^2 \simeq I$ based on the observations from experiments so far while $|V_{\ell N}|^2$ is suppressed to $O(\frac{m_\nu}{m_N})$ scale.

Going further, the interaction Lagrangians for N and EW bosons become

$$\begin{aligned}
\mathcal{L}_{\text{int.}} = &- \frac{g}{\sqrt{2}} W_\mu^+ \sum_\ell \overline{N} V_{\ell N}^* \gamma^\mu P_L \ell^- && \text{(Charged Current)} \\
&- \frac{g}{2\cos\theta_W} Z_\mu \sum_\ell \overline{N} V_{\ell N}^* \gamma^\mu P_L \nu_\ell^L && \text{(Neutral Current)} \\
&- \frac{g}{2M_W} H \sum_\ell \overline{N} V_{\ell N}^* m_{\text{N}} P_L \nu_\ell^L && \text{(Yukawa)} \\
&+ (h.c.).
\end{aligned} \tag{4.4}$$

Major advantage of building the formalism in such mixed mass and flavor basis is the simplicity while indirect constraints on $|V_{\ell N}|$ from electroweak precision measurements allows the ignorance of pair production of N by $|V_{\ell N}|^2 \leq 10^{-4}$. Each line in Eq. 4.4 represents the charged current, neutral current, and Yukawa interactions of N with W, Z, and H bosons. These various production and decay channels provide abundant potential making the type-I seesaw model as a feasible BSM physics at the LHC.

Further important remarks have to be made to justify the claims for studying such phenomenological hypothesis instead of full canonical type-I seesaw model. First of all, by reducing the number of independent model parameters, one can simply report the sensitivity in the m_N and $|V_{\ell N}|^2$ planes. Secondly, such mass and flavor mixed basis description becomes useful as as one can

sum over ν_ℓ^L. This is due to the fact that it can only exist either as a internal mediator or final state that cannot be detected by the LHC experiments.

In this scenario, the smallness of the observed neutrino masses can be explained through the largeness of mass of newly postulated N, where m_N is expressed as $m_\nu \sim y_\nu^2 v^2/m_N$. Here, y_ν is the Yukawa coupling and v is the Higgs vacuum expectation value in the SM. With $v = 1$ and the SM neutrino mass that is as light as $O(1\ \text{eV})$ scale, m_N is required to be around 10^{15} GeV. On the contrary, with minuscule y_ν, m_N can also become very small, near 0 GeV. Thus from this wide spectrum of possible m_N range, by fine tuning the y_ν to be at a scale of the electron Yukawa coupling, m_N could be at $O(1\ \text{TeV})$ scale or lower which makes it feasible to be searched by the LHC experiments.

4.2.2 Type-II Seesaw Model

Type-II seesaw model [75] is a mechanism that is special in a way as it is able to generate neutrino masses without invoking a new sterile heavy neutrino as in type-I seesaw model. Instead, it hypothesizes $SU(2)_L$ scalar triplet in addition to the SM and generates neutrino masses through Yukawa interactions between the SM lepton doublet and the new scalar triplet expressed as

$$\mathcal{L}^Y = \sum_{\alpha,\beta=\text{e},\mu,\tau} Y_{\alpha,\beta} \overline{L_\alpha^c} i\sigma_2 \Delta L_\beta + (h.c.) \tag{4.5}$$

which consequently generates the neutrino mass matrix

$$M_\nu = \sqrt{2} Y_\nu v_\Delta \simeq Y_\nu \frac{\mu v_0^2}{m_\Delta^2}. \tag{4.6}$$

The last approximation is acquired by minimizing the v_Δ in the limit $m_\Delta^2 \gg v^2$ and neglecting the additional perturbative terms. One can easily notice that SM neutrino masses around 0.1 eV could be explained by tuning the three scales introduced by type-II seesaw model : μ, v_0^2, and m_Δ^2.

4.2.3 Left-Right Symmetric Model

The left-right symmetric model (LRSM) [76] is an extension of the SM by a right-handed $SU(2)$ group. This model is an attractive BSM candidate as it is expected to restore the parity symmetry at high energy. In addition, as both type-I and type-II seesaw models can be embedded in this framework, neutrino masses can naturally be obtained.

In the effective treatment of the left-right symmetric field, it consists of SM particles along with newly predicted heavy charged $\mathrm{W}_R^{\pm}$ bosons, neutral Z' bosons, and heavy right-handed Majorana neutrino. In such context, the interaction Lagrangian of the left-right symmetry sector for $\mathrm{W}_R^{\pm}$ can be given as

$$\begin{aligned}\mathcal{L}_{\text{int.}} = &-\frac{g_R}{\sqrt{2}}\sum_{i,j=u,d,..}\overline{u_i}V_{ij}^{\mathrm{CKM}'}\mathrm{W}_{R\mu}^{+}\gamma^{\mu}P_R d_j \quad (\mathrm{W}_R \text{ coupling to quarks})\\ &-\frac{g_R}{\sqrt{2}}\sum_{\ell}\overline{N}V_{\ell N}\mathrm{W}_{R\mu}^{+}\gamma^{\mu}P_R\ell^{-} \quad (\mathrm{W}_R \text{ coupling to leptons})\\ &+(h.c.)\end{aligned} \tag{4.7}$$

where g_R is the right-handed coupling value and CKM' is the CKM matrix in the right-handed sector. For simplicity, $g_R = g$ is usually assumed so that right-handed and left-handed coupling values are the same.

At the LHC, several processes can occur from LRSM, such as $pp \to \mathrm{W}_R \to tb$, $pp \to \mathrm{W}_R \to \ell N$, and $pp \to NN$ when taking Z' interaction into account as well. With decay that is $N \to \ell \mathrm{W}_R^{*} \to \ell qq$ and free parameters m_{W_R} and $m_{\mathrm{Z}'}$, it provides phenomenologically abundant signal processes to be searched for.

4.3 Search for Heavy Neutrinos at the CMS

There has been number of experimental analyses performed to test the neutrino mass models at the LHC. Here we highlight and review some of the selected

studies from the CMS mainly to describe the short historical context during the LHC era on how the analysis that will be presented in Chapter. 5 was constructed. The selected analyses were performed to search for heavy neutrinos in type-I seesaw, left-right symmetric, or composite models using 8 TeV or 13 TeV data collected by CMS.

4.3.1 Type-I Seesaw Model in Dilepton Channel at 8 TeV

A search for heavy neutrinos decaying into a W boson and a lepton was performed in dimuon final state [77] which sequentially lead to the search in dielectron and electron-muon final states [78]. Both searches were performed using $19.7\,\mathrm{fb}^{-1}$ of data collected with the CMS detector at 8 TeV collision energy. The heavy neutrino considered in this analysis is an artifact of the type-I seesaw model and thus it is able to yield lepton number violating signatures from its Majorana nature. Thus, the analysis took the advantage of same-sign dilepton signature with additional jets which has significantly lower level of background events compared to the case of opposite-sign dilepton signature.

The limits were set on $m_N - |V_{\ell N}|^2$ parameter planes separately for all three cases where N mixes with ν_e only, with ν_μ only, and with both ν_e and ν_μ simultaneously [3]. The limits were set on the mass of all signal-like final state objects that is $\ell\ell jj$. The figures in Fig. 4.4 shows the exclusions set by these two analyses, which are all set up to $m_N = 500$ GeV for $|V_{\ell N}|^2$ smaller than 1.0.

4.3.2 Left-Right Symmetric Model in Dilepton Channel at 8 TeV

The process of interest in this analysis was similar to the signal process in Chapter. 4.3.1. However, instead of W boson being the mediator, it is replaced

[3] To be more specific, limits are set on $|V_{\mathrm{e}N}V^*_{\mu N}|^2/(|V_{\mathrm{e}N}|^2+|V_{\mu N}|^2)$ for simultaneous electron and muon flavor mixing scenario.

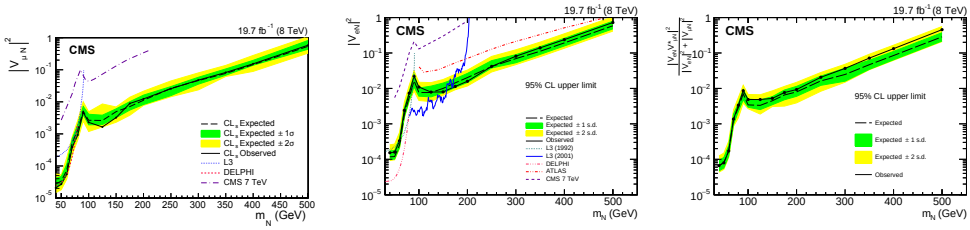

Figure 4.4: Exclusion limits set on $m_N - |V_{N\ell}|^2$ parameter space. From left to right shows the mixing limits set from dimuon, dielectron, and electron-muon channels.

with W_R boson. It effectively considers the scenario in which N only interacts with $SU(2)_R$ gauge bosons.

The analysis was performed in eejj and $\mu\mu jj$ final state [79]. Another notable difference from Chapter 4.3.1 was that this analysis did not exploit the lepton number violating nature. Analysis was instead performed in a charge-blind manner, not imposing any charge requirements to leptons. This way, nonprompt leptons which can contribute to same-sign background becomes ignorable while major background comes from prompt contributions such as Drell-Yan or top pair processes. The limits shown in Fig. 4.5 were set on $m_{W_R} - m_N$ parameter space using $m(\text{ee}jj)$ and $m(\mu\mu jj)$ and excluded $m_{W_R} < 2.87$ (3.00) TeV for scenarios where $m_N = \frac{1}{2}m_{W_R}$ from electron (muon) channel.

Two interesting points should be noted in this analysis. The parameter space this analysis scans on is in 2D, m_N and m_{W_R}, which is hindered by significant amount of signal sample it has to process. Thus analysis chose to study $m_N = \frac{1}{2}m_{W_R}$ scenarios in depth with detector simulations and map the efficiency to other mass grids in the parameter plane using generator level event distributions. Another consequence of this approach was that the parameter space where it has the large mass gap between m_{W_R} and m_N was not fully

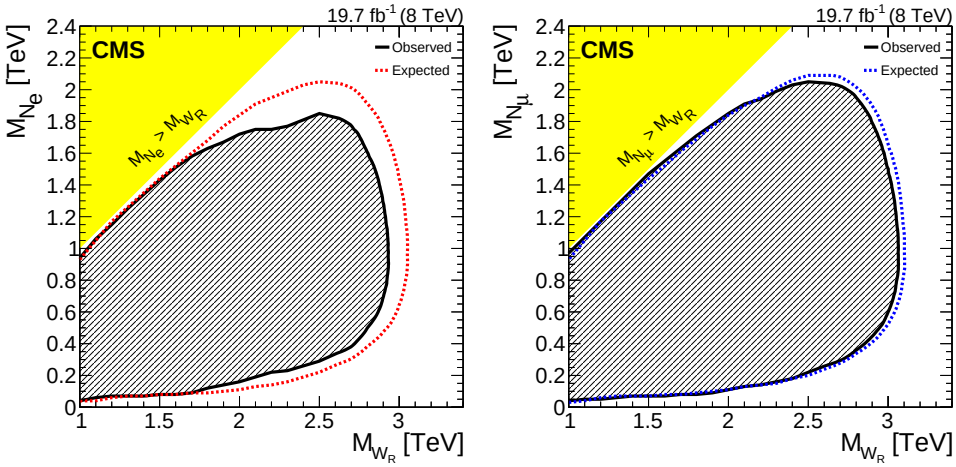

Figure 4.5: Exclusion limits set on $m_{W_R} - m_N$ parameter space. Left and right each shows the exclusion limits set from dimuon and dielectron channels, respectively.

explored. This parameter region $m_{W_R} \gg m_N$ will later be addressed by the full Run 2 analysis described in Chapter. 4.3.5.

Another point to make here is that this analysis showed a local significance of 2.8σ excess near $m(\mathrm{ee}jj) = 2.1$ TeV. However, similar excess was not observant in $m(\mu\mu jj)$ distributions (see Fig. 4.6-4.7). What was more intriguing was that ATLAS also performed the same analysis but using same-sign dilepton events did not show any excess in both electron and muon channel distributions [80]. The analyses from CMS and ATLAS were each public in end of 2014 and middle of 2015. A large number of theorists paid attention to this excess in electron channel while it was missing in muon channel and to the difference between ATLAS and CMS results which left the hint that the excess might be occuring from opposite-sign dilepton event only. Thus the direction of BSM searches in the heavy neutrino searches were very clear on what to do in the beginning of *Run 2* era, addressing the difference in electron and muon channels and

investigating the possible excess in opposite-sign dilepton events.

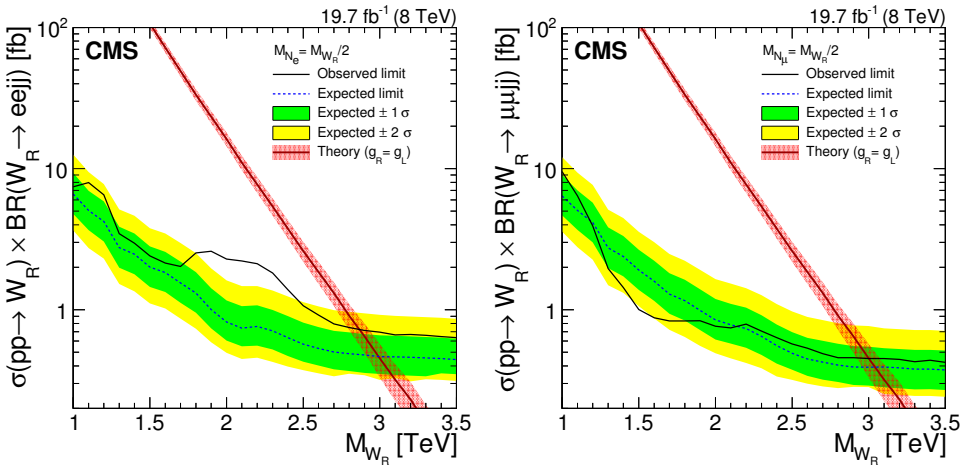

Figure 4.6: Exclusion limits set on $m_{W_R} - m_N$ parameter space. Left and right shows the exclusion limits set from dimuon and dielectron channels, respectively.

4.3.3 Composite Model in Dilepton Channel at 13 TeV

The composite model postulates a hypothetical constituent particle named preon which manifests itself forming an internal substructure at the compositeness scale Λ. In this approach, the contact interaction as well as the gauge interaction also becomes relevant, describing the unknown internal dynamics of compositeness with a prediction of existence of excited states of quarks and leptons. In light of this, heavy neutrino of Majorana type is considered to be the bound state of preons as one of the excited states. In this analysis [81], signal process consists of two leptons and two quarks.

This was a timely study answering the two questions from 8 TeV analysis on LRSM W_R and N searches from CMS and ATLAS :

- Electron channel excess was observed while muon channel did not show any excess. This, with composite model, could be explained if excited

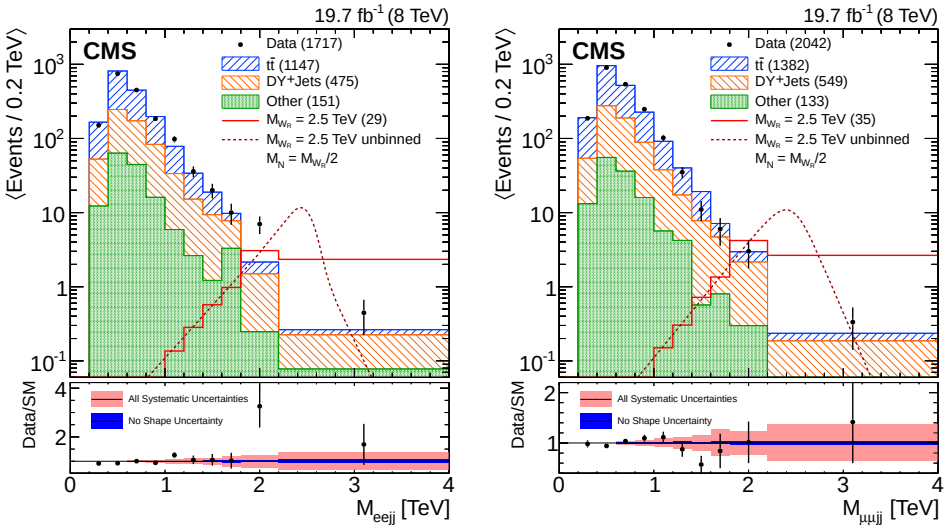

Figure 4.7: Distribution of $m(\ell\ell jj)$ for $(m_{W_R}, m_N) = (2.5, 1.25)$ TeV scenario. Left and right shows the distributions from dimuon and dielectron channels, respectively.

electron bound state is lighter than the excited muon bound state. Thus, muon channel excess could be expected at a larger mass scale.

- Charge blind analysis performed by CMS was observed with an excess while same-sign dilepton analysis result from ATLAS did not show any excess. If composite model forms a doubly charged excited state ($L^{\pm\pm}$), it can emerge as a signal from the process $pp \to e^{\mp}L^{\pm\pm} \to e^{\mp}e^{\pm}\mathrm{W}^{\pm} \to e^{\mp}e^{\pm}qq$ as it shares the same final state that the CMS analysis was considering.

Another fascination in this analysis was that it was the very first analysis to utilize the large-radius jet techniques to search for heavy neutrinos. It provides a huge enhancement to the analysis performance as the two particle-level partons from the signal process are likely to be collimated, being merged into one AK8

jet with relatively large radius $\Delta R = 0.8$.

However, despite many intriguing points this analysis carried, the existence of the composite particles were excluded up to masses of 4.60 (4.70) TeV for $m_N = \Lambda$ scenario from electron (muon) channels using the earliest 13 TeV data corresponding to an integrated luminosity of $2.3\,\mathrm{fb}^{-1}$. The results are shown in Fig. 4.8. Same analysis was performed using the full Run 2 dataset with adding unitarity bound conditions to optimize the analysis in small m_N and large Λ parameter space region. The exclusion contours were expanded, excluding up to 6.0 (6.1) TeV from electron (muon) channels) at the limit of $m_N = \Lambda$.

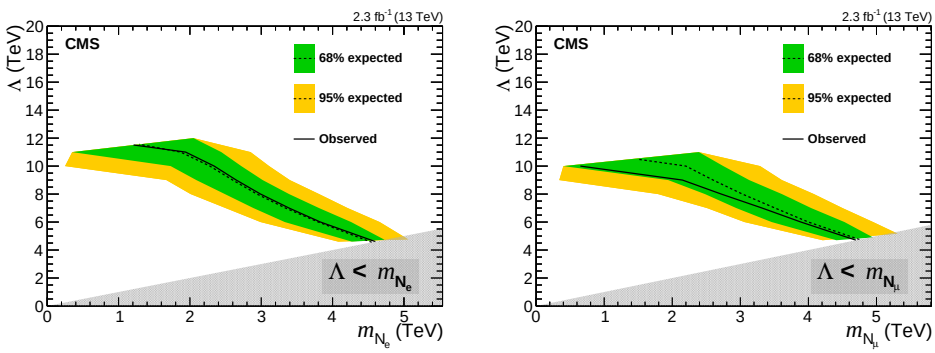

Figure 4.8: Exclusion limits set on $m_N - \Lambda$ parameter space. Left and right shows the exclusion limits set from dimuon and dielectron channels, respectively.

4.3.4 Type-I Seesaw Model in Dilepton Channel at 13 TeV

Same analysis was performed as a continuum from 8 TeV analysis, utilizing the low-level background with same-sign dilepton final state with 2016 13 TeV dataset [82]. However, two notable updates were made in the search.

- Adding more sophisticated signal regions : For large $m_N \gg 80$ GeV, as the composite model analysis shows, induces a collimated W $\rightarrow qq$ signal kinematics which can be greatly enhanced in signal acceptance when using

AK8 jets. Conversely for small $m_N < 80$ GeV, $N \to \ell qq$ becomes a three-body decay as W from N cannot be produced in onshell. This roughly gives $m_N/3$ energy to the lepton and two quarks from the decay that does not collimate unlike for large m_N scenarios. As AK4 jets are $O(20$ GeV) scale, this leads to poor sensitivity. Thus another signal region, using one AK4 jet as a proxy for hadronically decaying W* was also added.

- Adding γ-induced vector boson fusion (VBF) process (Feynman Diagram is depicted in Fig. 4.9 : As described in Ref. [74], this process is available when considering the photon effectively emitted from the PDF level. For large m_N scenarios, γ-induced process becomes dominant with larger cross sections, and also yielding very similar signal-like objects in the event (two isolated leptons with two particle-level quarks with additional quark going to the forward eta direction).

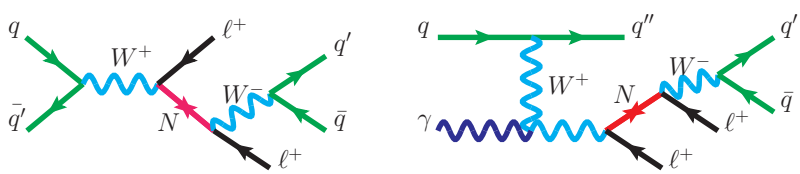

Figure 4.9: Feynman diagrams of type-I seesaw model's same-sign dilepton channel. Left is the typical charged current DY channel and right is the γ induced VBF channel. Both channels yield two leptons and channel on the right yields one additional particle-level quark.

The two improvements above for the analysis leveled up in both phenomenological and experimental aspects, excluding limits were set on 1240, 1430, and 1600 GeV of heavy neutrino mass for e, μ, and e$-\mu$ mixing scenarios at mixing $|V_{\ell N}|^2 = 1.0$ as shown in Fig. 4.10. This was the most restrictive limits set for m_N greater than 430 GeV at the time.

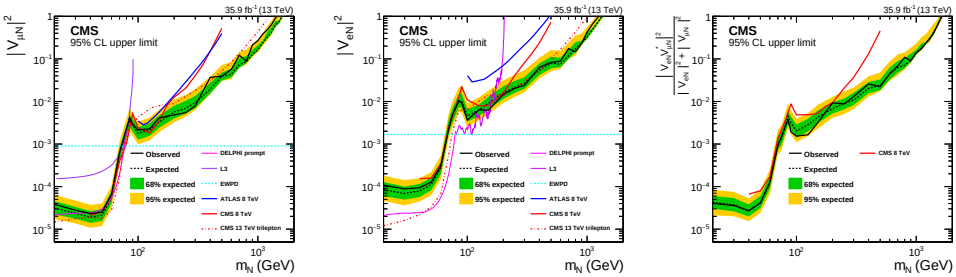

Figure 4.10: Exclusion limits set on $m_N - |V_{N\ell}|^2$ parameter space. From left to right shows the mixing limits set from dimuon, dielectron, and electron-muon channels.

4.3.5 Left-Right Symmetric Model in Dilepton Channel at 13 TeV

The search as an extension from 8 TeV was performed by CMS using the same approach with 13 TeV dataset collected in 2016 [83]. Two isolated leptons with two AK4 jets were selected as signal-like final state objects and the exclusion limits were set by mapping from the $m_N = \frac{1}{2}m_{W_R}$ sensitivity to the other grids in $m_N - m_{W_R}$ parameter space. As it evidently shows in Fig. 4.11, weakness in the $m_{W_R} \gg m_N$ parameter space region.

A huge progress has been made when the analysis was done with the full Run 2 dataset from CMS [84] (see Fig. 4.12). The major difficulties in the unexplored $m_{W_R} \gg m_N$ phase space was due to collimation of N decaying particles that are a lepton and two quarks in particle-level. Continuing the efforts of utilizing AK8 jets to mitigate the difficulties from composite model and type-I seesaw model analyses, an AK8 jet was used to identify the AK8 jet with three-body substructure (ℓqq) that originates from N.

The analysis introduced a new type of variable called lepton-subjet-fraction (LSF) [85] that can be used as an pseudo-isolation parameter for harsh environment where lepton cannot be fully isolated from nearby QCD activities.

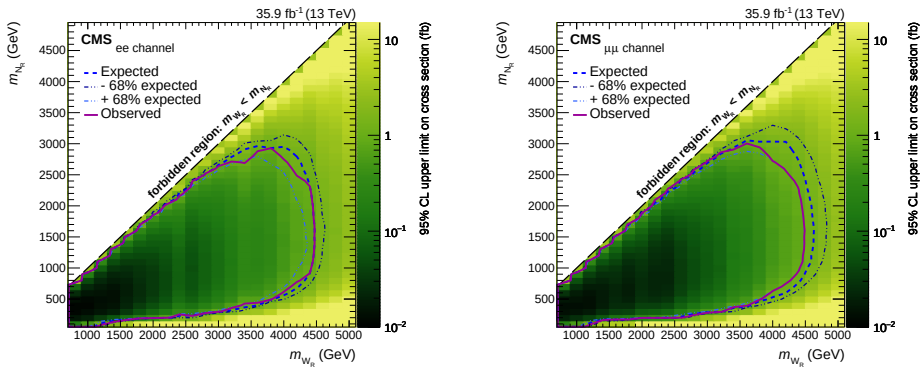

Figure 4.11: Exclusion limits set on $m_{W_R} - m_N$ parameter space. Left and right each shows the exclusion limits set from dimuon and dielectron channels, respectively.

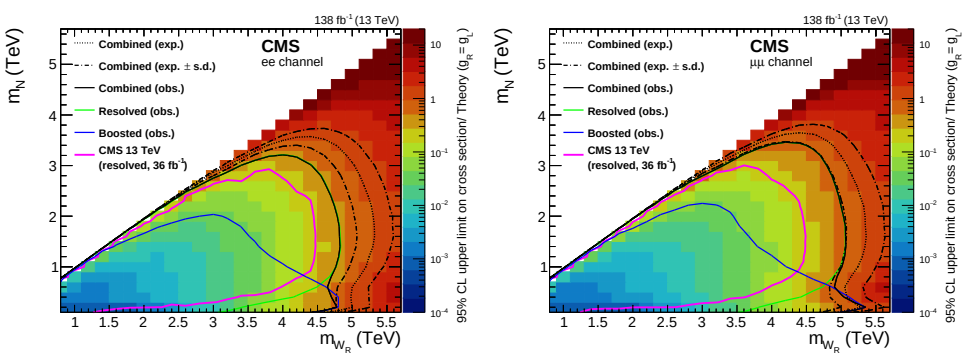

Figure 4.12: Exclusion limits set on $m_{W_R} - m_N$ parameter space. Left and right each shows the exclusion limits set from dimuon and dielectron channels, respectively.

LSF first disassembles the AK8 jet into a given number (n) subjets. After n subjets are formulated, the p_{T} of lepton inside the AK8 jet and the p_{T} of subjet that is closest to the lepton is compared which is the value of LSF parameter. If LSF is close to 1.0, the AK8 jet is likely to be incorporating the particle-level lepton along with two quarks. Giving a requirement on this variable for AK8 jets greatly improved the exclusion limits where $m_N < 0.5$ TeV that was not fully explored in the past.

4.3.6 Overall Review and Outlook

The search for heavy neutrino analysis has been one of the hot topics in Exotica (EXO) physics group in CMS. In particular, Jets+X subgroup of EXO has been leading efforts for heavy neutrino searches with advanced usages of jets as a physics object as described above since Run 1 at 8 TeV until Run 2 at 13 TeV. Such effort now spread out to long-lived analyses such as Ref. [86], searching for heavy neutrinos in slightly different parameter space where m_N is small having longer life time with displaced features in signal processes.

This thesis was also developed in the same context, as one of the extensions of efforts carried out in CMS, by exploring more advanced usage of jets, in particular, AK8 jets with recently developed technique that is ParticleNet tagger [87]. This gives a huge discriminating power against background events by identifying AK8 jets that originate from resonance with various type of jet substructures. More details will be described in Chapter. 5 with deeper details.

Chapter 5

Search for Heavy Neutrino at the CMS

5.1 Introduction

This search concentrates on the single lepton signatures with large missing transverse momentum. In addition, the heavy neutrino also yield W, Z, or H boson when it decays as shown in Feynman diagrams (see Fig. 5.1). The cases where the bosons decay into to two light or heavy flavor quarks are considered, which in the end yields either two AK4 jets or one AK8 jet depending on the scale of the Lorentz boost of the bosons. The analysis is performed using the full data collected during Run 2 at $\sqrt{s} = 13$ TeV with the CMS detector. As aforementioned, the final state that is targeted consists of a single lepton, large intrinsic missing transverse momentum from neutrinos, and either two AK4 jets or one AK8 jet. As the analysis is mainly motivated to search for heavy neutrinos using the AK8 jet as explained in Chapter. 4.3.6, the mass scenarios of N above 500 GeV are considered which possibly provide create enough Lorentz

boost for the bosons decayed from N.

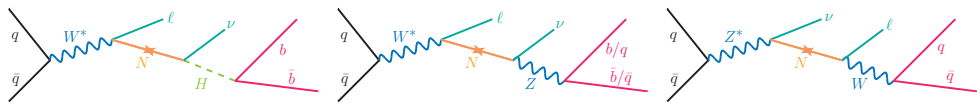

Figure 5.1: Heavy N Feynman diagrams with single lepton final state in particle level.

Searches for heavy neutrinos in the same type-I seesaw model was done in various channels at the LHC. The most recent search was from CMS using the full Run 2 data in the same-sign WW scattering channel [88]. The search was performed in same-sign dimuon channel associated with two forward direction AK4 jets, excluding m_N up to 23 TeV for $|V_{\mu N}|^2 = 1$ scenario and setting the best limits in the $m_N - |V_{\mu N}|^2$ parameter plane so far for $m_N > 650$ GeV. It is important to note that heavy neutrinos in this process is produced indirectly, involved as a mediator through the t-channel Feynman diagram. This is the main reason how it was able to provide the sensitivity even for m_N above the collision energy at the LHC. Other searches in a more relatively similar context to this analysis, searches for heavy neutrinos from direct production, have been done in two and three lepton final states using the 2016 data at $\sqrt{s} = 13$ TeV collected from CMS [82, 89]. Both studies explored the sensitivity in $m_N - |V_{\ell N}|^2$ parameter plane for both e and μ mixing N scenarios.

Here, it is worth noting that this is the first search for heavy neutrinos in the type-I seesaw model using the H boson decay modes. As the coupling strength between the H boson and a fermion is proportional to the mass of fermion, this channel provides the opportunity to probe the Yukawa coupling of neutrinos. In addition, this is the first search at the LHC to perform in the single lepton channel since the DELPHI and L3 collaborations at LEP [90, 91].

As these searches from the LEP have constraints from the collision energy that was maximally at 200 GeV, this analysis provides the first chance to seek heavy neutrinos with mass above 500 GeV with the same final state particles.

5.2 CMS Dataset and Triggers

5.2.1 Data

This analysis uses the pp collision data collected by the CMS detector during Run 2 corresponding to an integrated luminosity of 137.4 fb^{-1} in total. Data collected in 2016 is split into two dataset to better describe the saturation effects in APV readout chip that happened until 13th of August. The drain speed was found out to be slower than what was anticipated which was later changed with the fix of VFP parameter for faster recovery. 2016 dataset before the fix is called preVFP while the dataset after the fix is called postVFP. Each corresponds to 19.52 fb^{-1} and 16.81 fb^{-1}, summing up to 36.3 fb^{-1} in total for the whole 2016. Other years, 2017 and 2018 dataset corresponds to 41.48 fb^{-1} and 59.83 fb^{-1}. The quality luminosity sections that are enlisted in *Golden JSON*, indicating when all subdetectors were operational during the data taking, are only considered in this analysis.

5.2.2 MC Event Samples

Signal MC Event Samples

The signal MC events are generated using NNPDF 3.1 LUXQED NNLO PDF [32] at NLO using MADGRAPH_AMC@NLO [92, 93] and then passed to PYTHIA 8 for showering and hadronization modellings. This is an intentional choice to incorporate the γ induced VBF production channel to other channels easily which will be described in detail below. The UFO model card used for generation is available in Feynrules database [73, 74]. The model is an extension of

the SM by adding three singlet fermions, also known as heavy neutrinos, which the mixing with the SM neutrinos are controlled by the mixing matrix, written as below :

$$V_{\ell N} = \begin{pmatrix} V_{eN_1} & V_{eN_2} & V_{eN_3} \\ V_{\mu N_1} & V_{\mu N_2} & V_{\mu N_3} \\ V_{\tau N_1} & V_{\tau N_2} & V_{\tau N_3} \end{pmatrix}. \tag{5.1}$$

As can be observed in Eq. 5.1, mixing elements are split into three flavor compositions for three heavy neutrinos, separately. The oscillation mechanism of SM neutrinos suggests that there should be at least two types of heavy neutrinos. But since we have not yet found any direct evidence for the heavy neutrinos, it is also sensible to first only concentrate the study on the lightest possible heavy neutrino scenario. Thus, we restrict ourselves to phenomenological type-I seesaw model studies by turning on the entries of first column of the matrix only (electron-only or muon-only mixing N_1 search, ignoring additional heavy neutrinos that are N_2 or N_3).

Despite the fact that the study is bound to phenomenological type-I seesaw model, unitarity cannot be violated for neutrino oscillations which indicates that the values of mixing elements should be kept below 1.0. For the signal MC events $|V|^2 = 0.0001$ is kept while varying the mass of m_N between 500 to 1500 GeV. This allows the signals to be generated under narrow width approximation with plausible assumption obtained from electroweak precision data [94]. The neutrino mixings are only turned on for one lepton flavor only consequently prohibiting lepton flavor violation [1]. The samples for muon and electron channels are generated separately as below :

- μ channel : $|V_{\mu N_1}|^2 = 0.0001$, $|V_{eN_1}|^2 = 0.0$,

[1] Lepton number violation is allowed, however with the CMS detector, $\not{E}_T$ cannot be used to distinguish neutrinos and antineutrinos.

- e channel : $|V_{eN_1}|^2 = 0.0001$, $|V_{\mu N_1}|^2 = 0.0$,
- The mixing matrix entries not specified are set to 0.0.

We consider three different production channels. Charged current (CC) DY channel is N production channel through offshell W process in the SM, considered as the dominant production channel due to largeness of W production cross section. γ induced vector boson fusion channel which considers the initial state parton to be a photon is added as it shows larger contributions to the total cross section as the mass of N gets larger. Another production channel that has not yet been probed for N searches at the LHC is the neutral current (NC) DY channel where N is produced through offshell Z process in the SM. We add this to our search as it yields the same final state objects with similar kinematics. Thus the signal processes can be categorized into three N production modes as below, written with MadGraph_aMC@NLO syntaxes :

- Charged Current Drell-Yan

```
define ll = e+ e- mu+ mu-
define vv = ve ve~ vm vm~
generate p p > ll n1 [QCD]
```

- γ induced Vector Boson Fusion

```
define ll = e+ e- mu+ mu-
define vv = ve ve~ vm vm~
generate p a > ll n1 j [QCD]
add process a p > ll n1 j [QCD]
```

- Neutral Current Drell-Yan

```
define ll = e+ e- mu+ mu-

define vv = ve ve~ vm vm~

define ww = w+ w-

generate p p > vv n1 [QCD]
```

For the resonant N decays of the signal processes, MadSpin [95] is used with spin correlations properly taken into account. It's also worth noting that branching ratio for H boson is inaccurate due to the technical limitations of LO decay treatment. This computes H boson's branching ratio to bottom quark pair to be around 80 % which is quite far from the known level of precision at 58.9 %. Thus, signal samples are all separately generated for individual processes to account for correct production cross section and branching ratios. Considered decay channels of N are triggered in MadSpin with syntaxes written as below :

- W to light quarks [2] : `decay n1 > ww ll, ww > j j`
- Z to light quarks : `decay n1 > z vv, z > j j`
- Z to bottom quarks : `decay n1 > z vv, z > b b~`
- H to bottom quarks : `decay n1 > h vv, h > b b~`

Combinations of the N production channel and N decay channels which include exactly one lepton are the considered signal processes in the analysis. The signal processes are categorized into seven processed based on the combinations listed as below:

- CC DY $\otimes$ Z to light quarks

[2] We consider charm quarks as light quarks in this analysis.

- CC DY $\otimes$ Z to bottom quarks
- CC DY $\otimes$ H to bottom quarks
- NC DY $\otimes$ W to light quarks
- γ induced VBF $\otimes$ Z to light quarks
- γ induced VBF $\otimes$ Z to bottom quarks
- γ induced VBF $\otimes$ H to bottom quarks

Fig. 5.2 shows the cross sections of the different N production channels as a function of m_N, ranging from 500 to 3000 GeV. Red, blue, and green lines are the summed up cross sections yielding $\ell\nu q\bar{q}$ or $\ell\nu b\bar{b}$ final states from CC DY, NC DY, and γ induced VBF processes. Several points can be raised, first of all, γ induced VBF process becomes the dominant process for m_N above 900 GeV. Secondly, CC DY and NC DY show similar trends in cross sections as the production mechanism of N is similar ($pp \to \mathrm{W}^*$ and $pp \to \mathrm{Z}^*$). Although the production cross section of offshell W is much larger than offshell Z, the branching ratio of Z $\to \nu N$ being bigger than W $\to \ell N$ allows the NC DY channel to provide comparable sensitivity. As shown in 5.2, all three processes give contributions throughout the m_N range that cannot be ignored. Thus, all three processes are worth considering for the analysis.

Background MC Event Samples

Analysis purely relies on MC simulated events for background predictions in all control and signal regions. Thus the physics processes that yield one lepton from hard scattering with considerable cross sections can be considered as backgrounds. Background MC event samples are categorized into four : Single V,

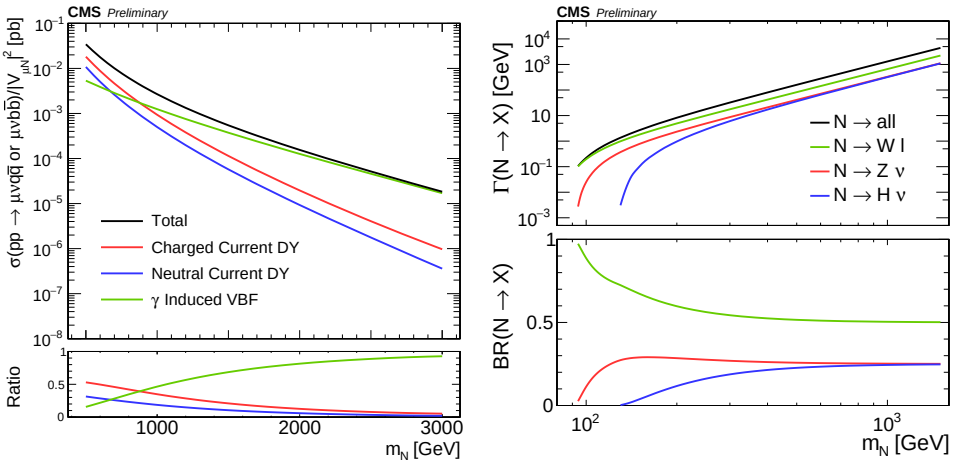

Figure 5.2: On the left is the NLO cross sections for CC DY, NC DY, and γ induced VBF processes with single lepton final state. Lower panel shows the fraction of each processes to the total sum. On the right is the total or partial decay widths of heavy neutrino. Lower panel shows the branching ratio to W, Z, and H bosons.

VV + X, Top Pair + X, Single Top. The configurations for the samples are written in detail below :

- Single V
  - DY : MiNNLO method [96] is used to calculate DY process at NNLO precision in Powheg [97]. It is interfaced with Pythia 8 for showering and hadronization. In addition, Photos [98] is added to simulate the QED part of the showering process instead of Pythia 8. The sample is split into three depending on its final state : ee, $\mu\mu$, and $\tau\tau$.
  - W : Sherpa [99] is responsible for the whole theoretical calculations that are hard scattering, showering, and hadronization. Up to two partons are added at NLO and additional three or four partons are

added at LO. The sample is biased with the module that populates the tail distribution in hadronic p_{T} sum. Thus events are not fully unweighted, which means that each event carries different unequal weights.

- VV + X
  - WW, WZ, ZZ : Diboson samples are all generated using Pythia 8 at LO that is responsible for the whole theoretical calculations. The decay modes of the bosons are inclusive.
  - WWW, WWZ, WZZ, ZZZ : MadGraph_aMC@NLO is used to calculate the hard scattering for triboson production at NLO. The inclusive decay of bosons are handled with Pythia 8 which is also used for the rest of theoretical calculations.
  - WH : WH samples are separated into two by the charge of the W bosons. W and H bosons in hard scattering are produced using Powheg with MiNLO method [100] at NLO. The W and H bosons are forced to decay leptonically with Powheg and into bottom quark pair using Pythia 8, respectively. The rest of theoretical calculations are treated by Pythia 8.
- Top Pair + X
  - Top pair : Powheg with MiNLO method is used for the hard scattering calculations at NLO. This is later interfaced to Pythia 8 for showering and hadronization step calculations. The samples are split into three to account for different decay modes : Fully leptonic, semi-leptonic, and fully hadronic.

- Top pair + W : Hard scattering is calculated at NLO using MadGraph_aMC@NLO with the decays handled with MadSpin module. One additional parton is added in matrix element level which thus goes through FxFx jet merging [101] setup in Pythia 8 which also takes care of the rest of calculations. The decays of W boson that does not originate from top quarks decide the splitting of the sample : Leptonically and hadronically decaying W boson.
- Top pair + Z : The samples are generated in the same way as top pair + W process except that there is no additional partons added in matrix element and thus does not need FxFx jet merging. The samples are split into two depending on the Z boson : Pair of leptons and neutrinos are simultaneously generated in one sample and pair of quarks are separated.

- **Single Top**
  - Single top (s-channel) : The sample targets the process which W boson mediates the top and bottom quark production in s-channel. It is generated using MadGraph_aMC@NLO at NLO where the decays of top quark is forced to decay into leptonic channel using MadSpin. Showering and hadronization steps are processed using Pythia 8.
  - Single top (t-channel) : The t-channel single top sample is a mixture of Powheg and MadGraph_aMC@NLO. The hard scattering up to top quark production is handled by Powheg but the rest of the top decays are interfaced to MadGraph_aMC@NLO in order to use the MadSpin module for resonance decays. Rest of the calculations are treated by Pythia 8.

– Single top + W : Top quark associated with W boson process is generated using POWHEG for hard scattering and the rest calculations are done by PYTHIA 8. The samples are split in two layers, first the top and antitop samples are separately generated. As top quark also yields a W boson with a bottom quark, with two bosons, samples are split into two, fully hadronic and not fully hadronic decays. In the analysis we only use the not fully hadronic decaying sample.

5.2.3 Trigger Selection

SingleMuon dataset is used to search for N which mixes with μ with events which are triggered by isolated leptons that have p_{T} greater than 24 GeV, 27 GeV, and 24 GeV for 2016, 2017, and 2018, respectively. For N search that mixes with e, *SingleElectron dataset* is used for 2016 and 2017 while *EGamma dataset* is used for 2018. Multiple triggers are combined for e channel analysis. First of all, lowest p_{T} requiring triggers fire the events with isolated leptons having p_{T} greater than 27 GeV, 35 GeV, and 32 GeV for 2016, 2017, and 2018 respectively. Photon triggers which are triggered by events that contains a photon with p_{T} above 175 (200) GeV for 2016 (2017 and 2018) are also used [3]. Lastly, for all three years, triggers with looser identification criteria for electron with p_{T} threshold at 115 GeV are added to further enhance the acceptance of the event. The triggers used for the analysis are shown in Table. 5.1.

5.3 Object Definition

5.3.1 Leptons : Electron and Muon

Due to the limitations of the detector and high energy beam conditions, analyses with one lepton and jet(s) as a kinematic signature is prone to backgrounds that

[3] Electrons have a huge chance to be reconstructed as a photon as the main difference of the two comes from either fit quality or existence of the tracks.

Year	μ channel	echannel
2016preVFP	HLT_IsoMu24	HLT_Ele27_WPTight_Gsf or HLT_Photon175 or HLT_Ele115_CaloIdVT_GsfTrkIdT
2016postVFP	HLT_IsoMu24	HLT_Ele27_WPTight_Gsf or HLT_Photon175 or HLT_Ele115_CaloIdVT_GsfTrkIdT
2017	HLT_IsoMu27	HLT_Ele35_WPTight_Gsf or HLT_Photon200 or HLT_Ele115_CaloIdVT_GsfTrkIdT
2018	HLT_IsoMu24	HLT_Ele32_WPTight_Gsf or HLT_Photon200 or HLT_Ele115_CaloIdVT_GsfTrkIdT

Table 5.1: Name of triggers used for the analysis.

consist of *nonprompt leptons* that emerge from signals that are misidentified as leptons. These includes leptons which originate from the hadron decays or jets that are misidentified as leptons are classified as nonprompt leptons. On the other hand, the word *prompt leptons* refers to leptons that decays from electroweak gauge bosons or leptonically decaying taus.

Relative isolation (I_{Rel}) is an important discriminating variable for lepton identification, distinguinshing the prompt and nonprompt leptons, for better analysis performance in both electron and muon channels. It is defined as

$$I_{\mathrm{Rel}} = \frac{1}{p_{\mathrm{T}}}[\sum p_{\mathrm{T}}^{\mathrm{CH}\in\mathrm{PV}} + \max(0, \sum p_{\mathrm{T}}^{\mathrm{NH}} + \sum p_{\mathrm{T}}^{\gamma} - p_{\mathrm{T}}^{\mathrm{PU}})]. \tag{5.2}$$

The calculation is performed by summing over the transverse momenta of charged hadrons (after removing the components that are not associated to the primary vertex), neutral hadrons, and photons, denoted as $\sum p_{\mathrm{T}}^{\mathrm{CH}\in\mathrm{PV}}$, $\sum p_{\mathrm{T}}^{\mathrm{NH}}$, and $\sum p_{\mathrm{T}}^{\gamma}$ within a $\Delta R = 0.4$ cone around the leptons, respectively. $p_{\mathrm{T}}^{\mathrm{PU}}$ is an additional correction factor that estimates the contributions coming from pileup effect which is differently defined for electrons and muons. $p_{\mathrm{T}}^{\mathrm{PU}} = \rho A_{\mathrm{eff}}$ is used to mitigate the pileup effect for electrons by scaling the average pileup energy density (ρ) to the effective area (A_{eff}) to account for geometrical area correction of the cone. For muons, $p_{\mathrm{T}}^{\mathrm{PU}} = 0.5 \sum p_{\mathrm{T}}^{\mathrm{CH}\notin\mathrm{PV}}$ is used, subtracting half of the sum of the transverse momenta of charged hadrons that are not from the primary vertex. QCD processes have relatively large I_{Rel} values due to hadronic activity near the leptons.

The energy deposits in ECAL energy and the GSF tracks are associated to reconstruct electrons. Geometric designs of the calorimemter and the tracker bounds the electron acceptance to $|\eta| < 2.5$ with minimum requirement to the transverse momentum of $p_{\mathrm{T}} > 10$ GeV. The energy of electrons in MC samples are corrected for different detector response between MC simulation and

actual data by scaling and smearing. The different reconstruction performance of electrons is corrected with the scale factor measured using Z $\rightarrow$ ee events with tag-and-probe method. Multivariate analysis selection criteria is trained including the relative isolation values of electrons and is used for electron identification. Two working points, WP80 and WP90 which are defined as working points with selection efficiency 80 % and 90 % respectively, are used in the analysis. The former is to identify prompt electron that is likely to be from signal processes while the WP90, relatively looser criteria, is used to reject events with additional electrons.

Muons are required to be within the acceptance $|\eta| < 2.4$ with minimum requirement to the transverse momentum at $p_{\mathrm{T}} > 10$ GeV. The reconstruction combines the track information from silicon tracker and muon system. Rochester correction is applied to correct the scale and resolution of muon's transverse momentum scale and resolution. The changes in transverse momentum due to the correction are propagated to $\not{E}_{\mathrm{T}}$. Muon identification is required to suffice the standard tight working point and two isolation criteria are further applied. Relative isolation $I_{\mathrm{Rel}} < 0.15$ is applied to capture the prompt muon and events containing additional muons with $I_{\mathrm{Rel}} < 0.6$ are discared from the analysis.

5.3.2 Jets : AK4 and AK8 Jets

Jets are reconstructed using the anti-k_T algorithm, clustering PF candidates within the distance of $\Delta R = 0.8$ (AK8) or $\Delta R = 0.4$ (AK4). For the AK4 jets, contamination from pileup is mitigated by applying Charged Hadron Subtraction (CHS) algorithm [102]. CHS algorithm removes the tracks that are not associated to the primary vertex when reconstructing the jets. Pileup Per Particle Identification (PUPPI) algorithm is applied to AK8 jets which give different weights to particle flow candidates prior to jet clustering which are likely to be

from pileup interactions, making the object to be more resilient to harsh pileup conditions.

AK8 jets passing the standard tight working point with $p_{\mathrm{T}} > 200$ GeV and $|\eta| < 2.4$ are used in the analysis. If an AK8 jet is close to the leptons by a distance of $\Delta R < 0.4$, the AK8 jet is considered as an energy clustering of leptons and thus discarded. Similarly, AK4 jets are required to satisfy the standard tight working point but with loose lepton veto requirements in order to pick up only jets that are likely to be originating from QCD. $p_{\mathrm{T}} > 30$ GeV and $|\eta| < 2.4$ are the initial requirements for the AK4 jets and those are further checked for the overlaps with other objects in the event. An AK4 jet is discarded if it has an overlap with either a lepton or an AK8 jet by a distance of $\Delta R < 0.4$ and $\Delta R < 1.2$, respectively.

As the analysis heavily relies on the substructure of the AK8 jet and its masses to identify the resonances, soft drop algorithm [103] is used to calculate the mass of AK8 jets. This algorithm removes the wide-angle soft radiations in a jet to mitigate the contamination from initial state radiations, underlying events, and pileup interactions. Soft drop mass of AK8 jets are required to be above 50 GeV.

AK8 Jet Tagging : ParticleNet

ParticleNet tagger [87] is used to identify the different substructure characteristics of an AK8 jet. ParticleNet is an advanced neural network architecture fed by various low-level and point-like PF candidates. Each AK8 jets are assigned with mass-decorrelated ParticleNet scores indicating how much the AK8 jet is likely to be originating from a resonance to bottom quark pair ($pnet^{\mathrm{raw}}_{\mathrm{MD}}(Xbb)$), charm quark pair ($pnet^{\mathrm{raw}}_{\mathrm{MD}}(Xcc)$), light quark pair ($pnet^{\mathrm{raw}}_{\mathrm{MD}}(Xqq)$), or originating from QCD interactions rather than a resonance decay ($pnet^{\mathrm{raw}}_{\mathrm{MD}}(QCD)$), etc.

classifying various decay modes.

$$pnet_{\mathrm{MD}}(Xbb) = \frac{pnet^{\mathrm{raw}}_{\mathrm{MD}}(Xbb)}{pnet^{\mathrm{raw}}_{\mathrm{MD}}(QCD) + pnet^{\mathrm{raw}}_{\mathrm{MD}}(Xbb)} \tag{5.3}$$

$$pnet_{\mathrm{MD}}(Xqq) = \frac{pnet^{\mathrm{raw}}_{\mathrm{MD}}(Xqq) + pnet^{\mathrm{raw}}_{\mathrm{MD}}(Xcc)}{pnet^{\mathrm{raw}}_{\mathrm{MD}}(QCD) + pnet^{\mathrm{raw}}_{\mathrm{MD}}(Xqq) + pnet^{\mathrm{raw}}_{\mathrm{MD}}(Xcc)}. \tag{5.4}$$

ParticleNet is the key for securing sensitivity in the merged signal regions. Fig. 5.3 shows how signal and background processes have different ParticleNet score distributions. The distributions are shown after merged preselection requirements are met for the events (details are explained in Chapter. 5.4.3). The processes are all normalized to same area, dashed and solid lines represent the signal and background processes. In particular, $X \to bb$ and $X \to qq$ in the figure are the collection of all the processes that yields particle-level bottom quark pair and non-bottom quark pair, respectively. As shown in Fig. 5.3, Xbb score shows relatively better signal-background discrimination power when compared to Xqq score. This is because many known processes for background also innately include $X \to qq$ thus hindering the discrimination. From this, we give higher priority to Xbb score when defining the signal regions.

AK4 Jet Tagging : B-Tagging

To identify AK4 jets originating from a bottom quark, multiclass flavor tagging algorithm, DeepJet [104], is used. It makes use of low-level features as much as possible from the multitude of jet constituents rather than utilizing the few well identified and reconstructed higher level features. This algorithm assigns to each jet the likelihood that it contains a bottom quark. A jet is identified as a b-tagged jet if it passes the medium working point.

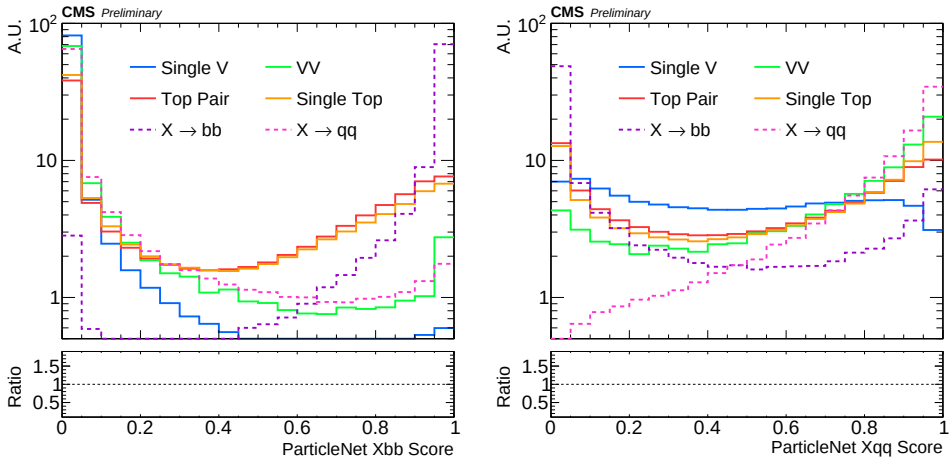

Figure 5.3: ParticleNet score distributions for Xbb (left) and Xqq (right) after preselection.

5.3.3 Missing Transverse Momentum

Missing transverse momentum ($\not{E}_{\mathrm{T}}$) [105] is a proxy for the particles that leave the detector without any trace such as neutrinos or possibly BSM particles with vanishingly small interactions. PFMET which is typically used in most of the analyses so far is defined as the negative of the sum of PF candidate's transverse momentum. Evolved from this approach, PUPPIMET was developed to mitigate the impacts from pileup interactions [106]. The pileup interactions mostly are isotropical and thus the projection of $\not{E}_{\mathrm{T}}$ in any direction is 0 GeV for both PFMET and PUPPIMET. However when comparing the resolutions of $\not{E}_{\mathrm{T}}$ for PFMET and PUPPIMET, PFMET carries bigger degrade in resolution compared to PUPPIMET in the presence of pileup. Thus, we use PUPPIMET for $\not{E}_{\mathrm{T}}$ as a proxy of the SM neutrinos that are present in the signal processes. Jet energy correction is propagated to $\not{E}_{\mathrm{T}}$ which is called type-I correction of $\not{E}_{\mathrm{T}}$.

5.4 Analysis Strategy

5.4.1 Background Estimation

Source of background processes are the ones that are not reducible such as W or top pair processes can innately give out one lepton with missing transverse momentum from the neutrino. It becomes more important for the analysis to properly control top pair process as it can also mimic the boosted signature with AK8 jet with the W boson gaining a huge Lorentz boost while decaying from the top quark (W process can yield AK8 jets only through initial-state radiation which is not that much likely too happen). Thus, we study the background modellings of W and top pair processes in depth for the analysis.

5.4.2 Signal Process Event Kinematics

Isolated single lepton triggers with minimum p_{T} threshold across all years are used for both electron and muon channel analyses. As we focus on the searches for N masses above 500 GeV, the signal lepton in the event is typically boosted. The p_{T} distribution of leptons in the signal processes as can be seen in Fig. 5.4, varies depending on the mass, however generally are well over the trigger's thresholds which is ~ 30 GeV and ~ 40 GeV for muon and electron, respectively.

We refer to the bosons that decay from N as X^{gen} in generator level and X^{reco} in reconstruction level. The mass gap between N and the bosons decayed from N, which is X^{gen}, can create a Lorentz boost to X^{gen} which eventually decays into either pair of light or bottom quarks. This provides the chance for an event to have one large-radius jet that incorporates the two quarks with small ΔR. This is more likely to occur for the larger m_N scenarios as the increment in mass gaps between N and X^{gen} leads to larger Lorentz boost of X^{gen}. The study of this effect using the signal samples with $N \to \mathrm{H}\nu \to b\bar{b}\nu$ decay chain

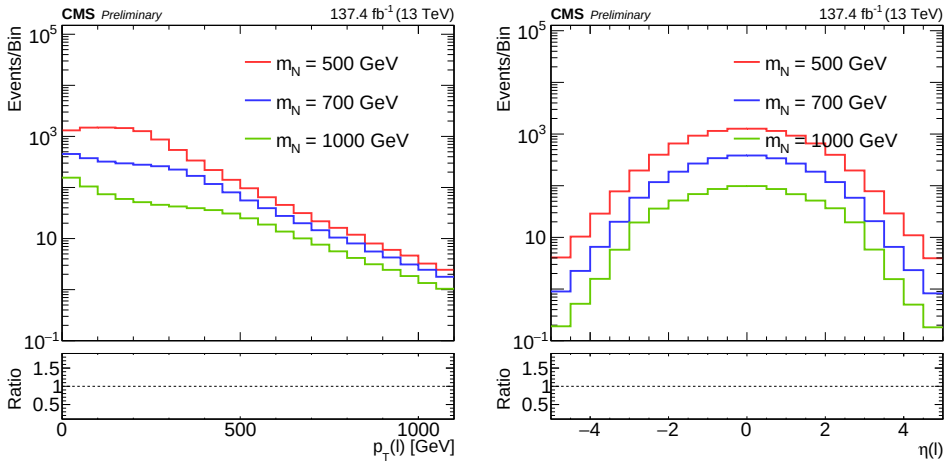

Figure 5.4: p_{T} (left) and η (right) distributions of leptons in signal processes.

is shown in Fig. 5.5. As m_N increases, the bottom quark pair becomes more collimated ($\Delta R(b\bar{b}) \to 0$). In light of this, we consider both an AK8 jet and AK4 jets to exploit the full sensitivity with merged and resolved event topologies, respectively, and accordingly reconstruct X^{reco} using either types of jets.

5.4.3 Preselection

This analysis searches for heavy neutrinos, in the scenarios where its mass is above 500 GeV. As shown in Feynman diagrams in Fig. 5.1, signal events contain a well isolated lepton and two quarks in particle level. In addition, as neutrinos escape from the detector without leaving a trace, large missing transverse momentum is expected. Event selections for preselection can be first categorized in lepton flavors, electron and muon channels. Isolated single electron (muon) HLT is required to be fired with the presence of one tight electron (muon) that has p_{T} above 50 GeV in the event. Events with additional leptons that satisfy loose isolation criteria with p_{T} above 10 GeV is vetoed.

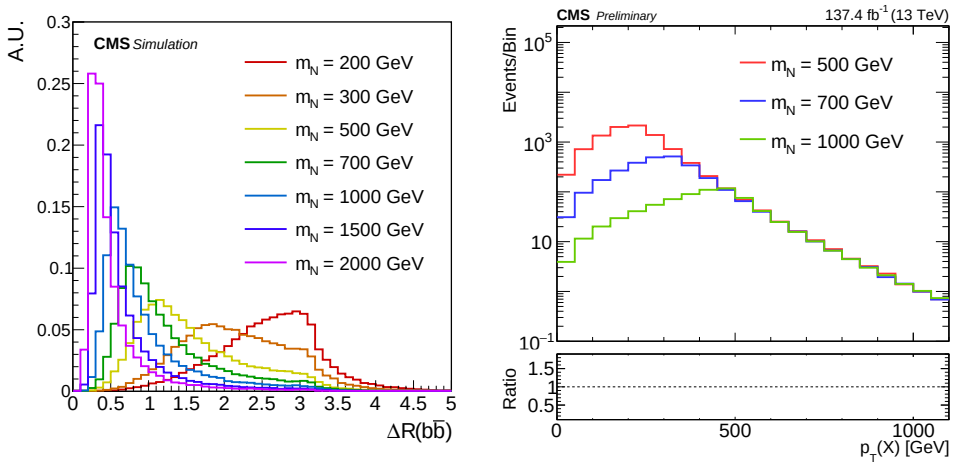

Figure 5.5: $\Delta R(b\bar{b})$ and $p_{\mathrm{T}}(X^{\mathrm{gen}})$ distributions in signal processes.

- **Electron preselection**
  - Single electron HLT is triggered.
  - Exactly one tight electron with $p_{\mathrm{T}} > 50$ GeV and $|\eta| < 2.1$ is in the event.
  - No additional loose leptons ($p_{\mathrm{T}} > 10$ GeV) are present in the event.
- **Muon preselection**
  - Single muon HLT is triggered.
  - Exactly one tight muon with $p_{\mathrm{T}} > 50$ GeV and $|\eta| < 2.1$ is in the event.
  - No additional loose leptons ($p_{\mathrm{T}} > 10$ GeV) are present in the event.

The event regions are further categorized into two, merged and resolved regions based on the topology of jets in the event. Two quarks from the boson can formulate one merged AK8 jet (J) or two resolved AK4 jets (j) depending

on the mass of N that could induce Lorentz boost of the boson denoted as X^{gen}. If J exists in the event, it first falls into the merged region category and X^{reco} is taken from the four-momentum of J. If there is no J in the event, at least two j is required and the two leading j are used to reconstruct X^{reco} which then belongs to the resolved region.

Minimum 100 GeV cut is applied on $\not{E}_{\text{T}}$ as signal events intrinsically carry largely boosted SM neutrinos. QCD background process is mostly suppressed at this stage with this cut. We define the variable $m_T^{\text{eff}} = m_T(\ell,\ \not{E}_{\text{T}},\ X^{\text{reco}})$, the transverse mass of all signal objects, as a proxy of the offshell W boson (or Z boson for neutral current DY production process). Since our interest is the scenario where m_N is above 500 GeV, $m_T^{\text{eff}} > 450$ GeV is given based on the fact that offshell W (Z) mass is likely to be above 500 GeV that is the minimal m_N we consider in the analysis.

The topology preselections are categorized into two as below. p denotes the Lorentz four-momentum of the particle.

- **Merged preselection**
  - $\not{E}_{\text{T}} > 100$ GeV.
  - An AK8 jet is present in the event.
  - The leading AK8 jet is used to reconstruct the boson from N decay, $p(X^{\text{reco}}) = p(J)$.
- **Resolved preselection**
  - $\not{E}_{\text{T}} > 100$ GeV.
  - There is no AK8 jet in the event but at least two AK4 jets are present in the event.

– Two leading AK4 jets are used to reconstruct the boson from N decay, $p(X^{\text{reco}}) = p(j_1) + p(j_2)$.

Results are separately analyzed for electron and muon channels with events passing electron preselection and muon preselection, respectively. Events that are considered in the analysis satisfies one of the four conditions below :

- **Electron merged preselection** : Events that fulfill both **electron preselection** and **merged preselection** requirements.
- **Electron resolved preselection** : Events that fulfill both **electron preselection** and **resolved preselection** requirements.
- **Muon merged preselection** : Events that fulfill both **muon preselection** and **merged preselection** requirements.
- **Muon resolved preselection** : Events that fulfill both **muon preselection** and **resolved preselection** requirements.

From the event selections after the preselection requirements, we further categorize the event selections into control and signal regions by utilizing the signal process's event kinematics such as number of b-tagged jets, $m(X^{\text{reco}})$, $m_T(\ell,\ \not{E}_{\text{T}})$, and $\Delta\phi(\ell,\ \not{E}_{\text{T}})$. The distributions of the variables described above as well as the lepton p_{T} and $\not{E}_{\text{T}}$ distributions for full Run 2 are shown in Figs. 5.6-5.13, overlayed with $m_N = 700$ GeV scenario from all signal processes.

5.4.4 Signal Region

Depending on the jet topology preselection, X^{reco} is reconstructed differently either using an AK8 jet or two AK4 jets. X^{reco} serves as the proxy of the boson (could be either W, Z, or H bosons) that is the daughter particle of N. Thus, $m(X^{\text{reco}})$ is required to lie within the mass window between 65 GeV and

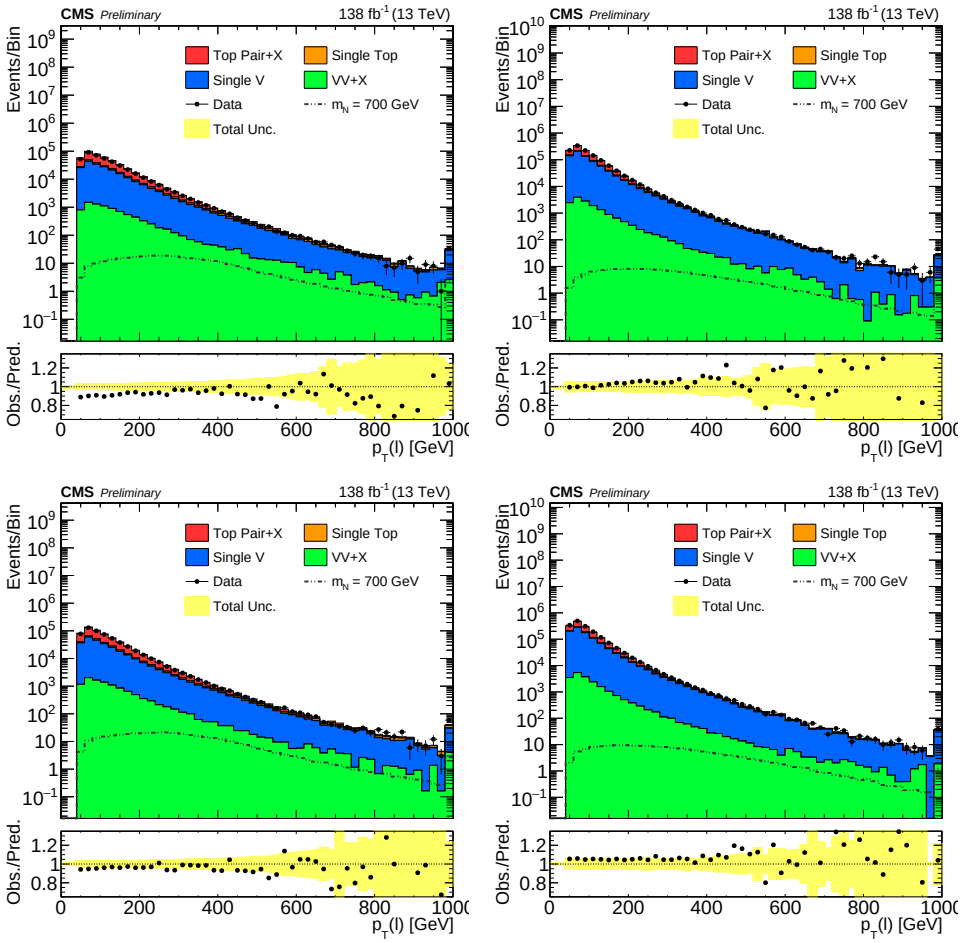

Figure 5.6: Transverse momentum of the lepton distributions for full Run 2 after preselection. The distributions are separately shown for electron merged (top left), resolved (top right), and muon merged (bottom left), resolved (bottom right) topologies.

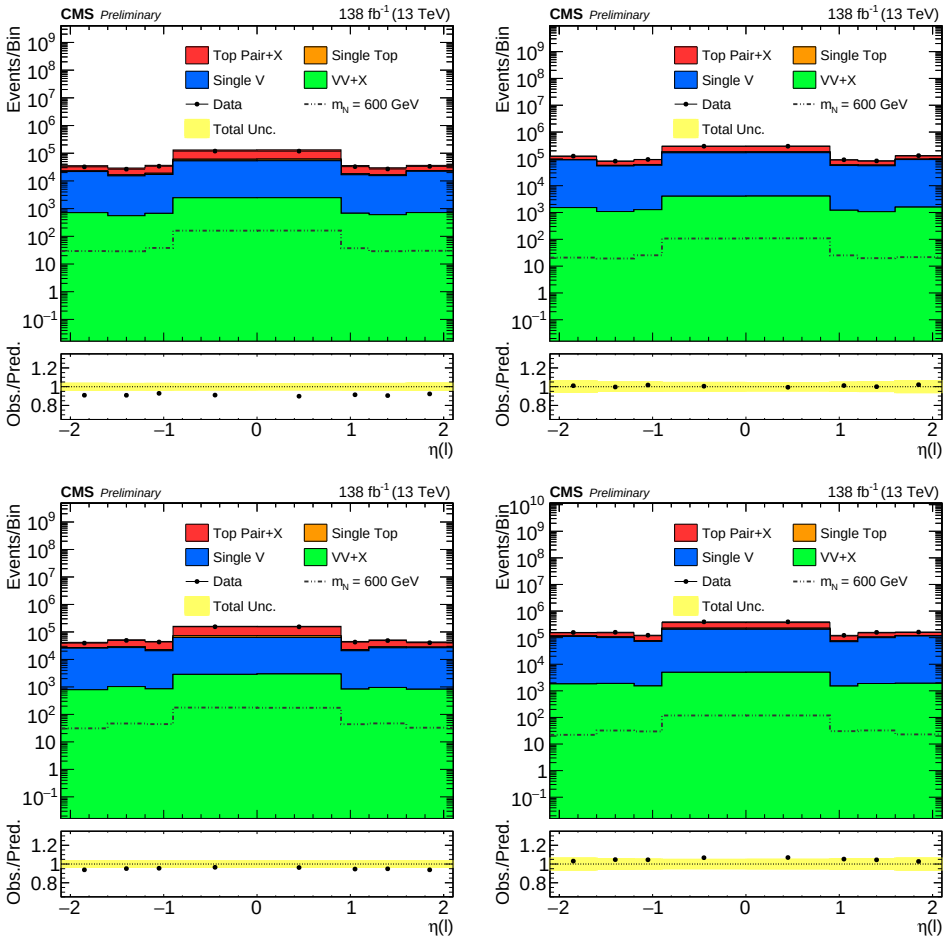

Figure 5.7: Pseudorapidity of the lepton distributions for full Run 2 after preselection. The distributions are separately shown for electron merged (top left), resolved (top right), and muon merged (bottom left), resolved (bottom right) topologies.

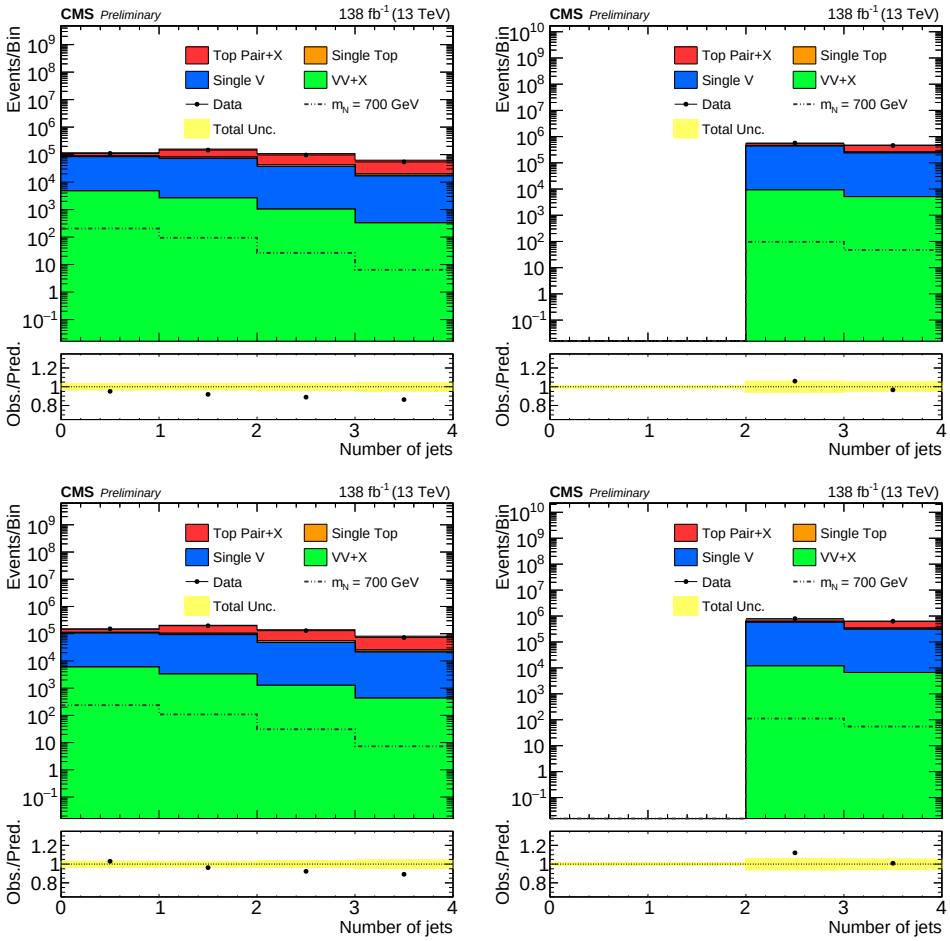

Figure 5.8: Number of jet distributions for full Run 2 after preselection. The distributions are separately shown for electron merged (top left), resolved (top right), and muon merged (bottom left), resolved (bottom right) topologies.

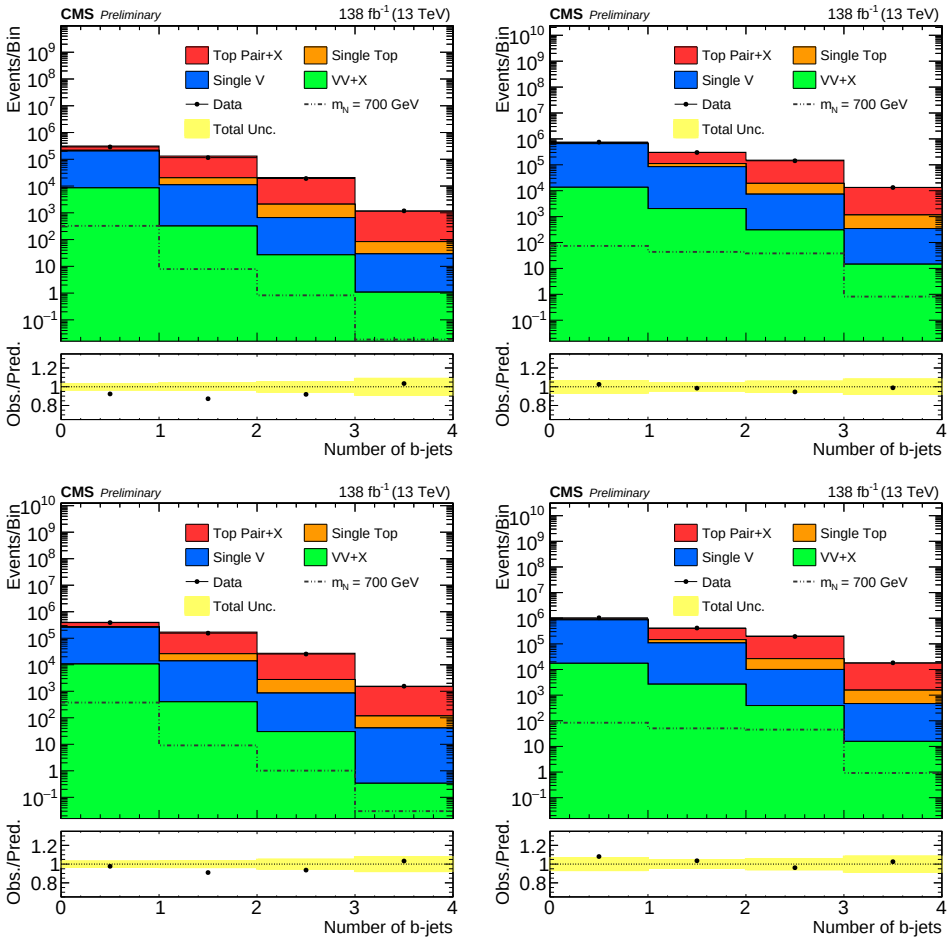

Figure 5.9: Number of b-tagged jet distributions for full Run 2 after preselection. The distributions are separately shown for electron merged (top left), resolved (top right), and muon merged (bottom left), resolved (bottom right) topologies.

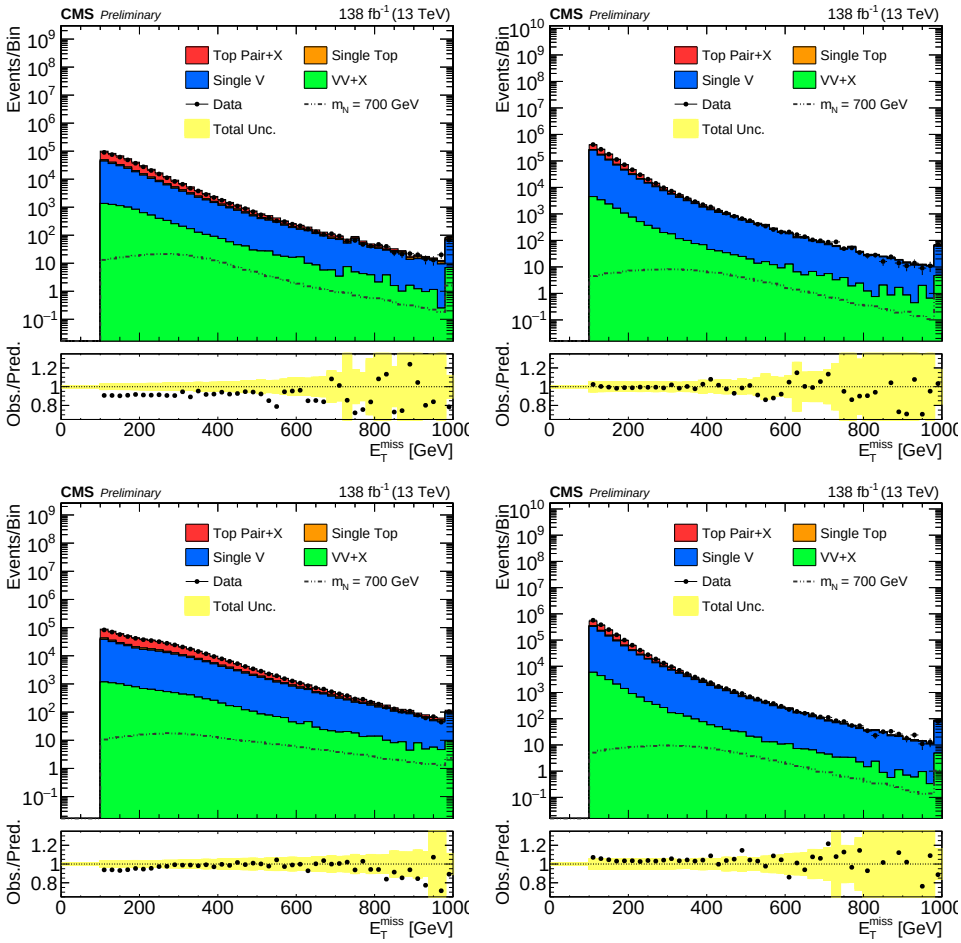

Figure 5.10: Missing transverse energy distributions for full Run 2 after preselection. The distributions are separately shown for electron merged (top left), resolved (top right), and muon merged (bottom left), resolved (bottom right) topologies.

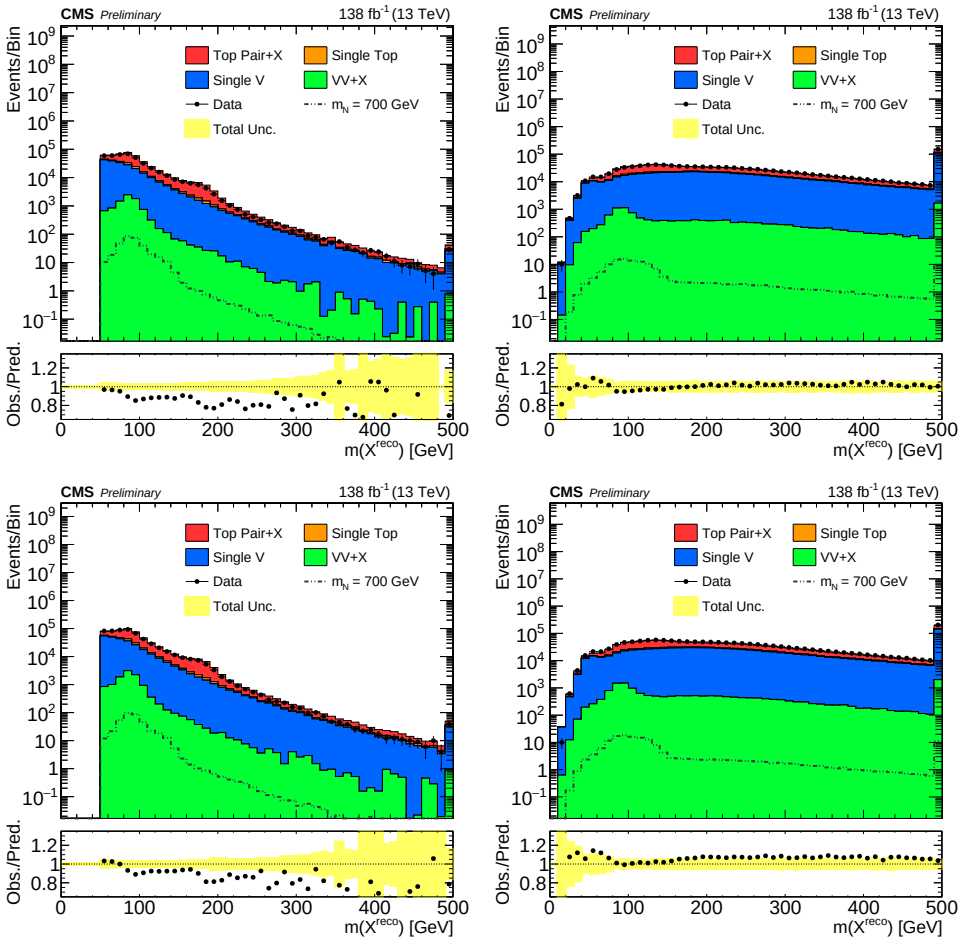

Figure 5.11: Reconstructed mass of boson from N decay distributions for full Run 2 after preselection. The distributions are separately shown for electron merged (top left), resolved (top right), and muon merged (bottom left), resolved (bottom right) topologies.

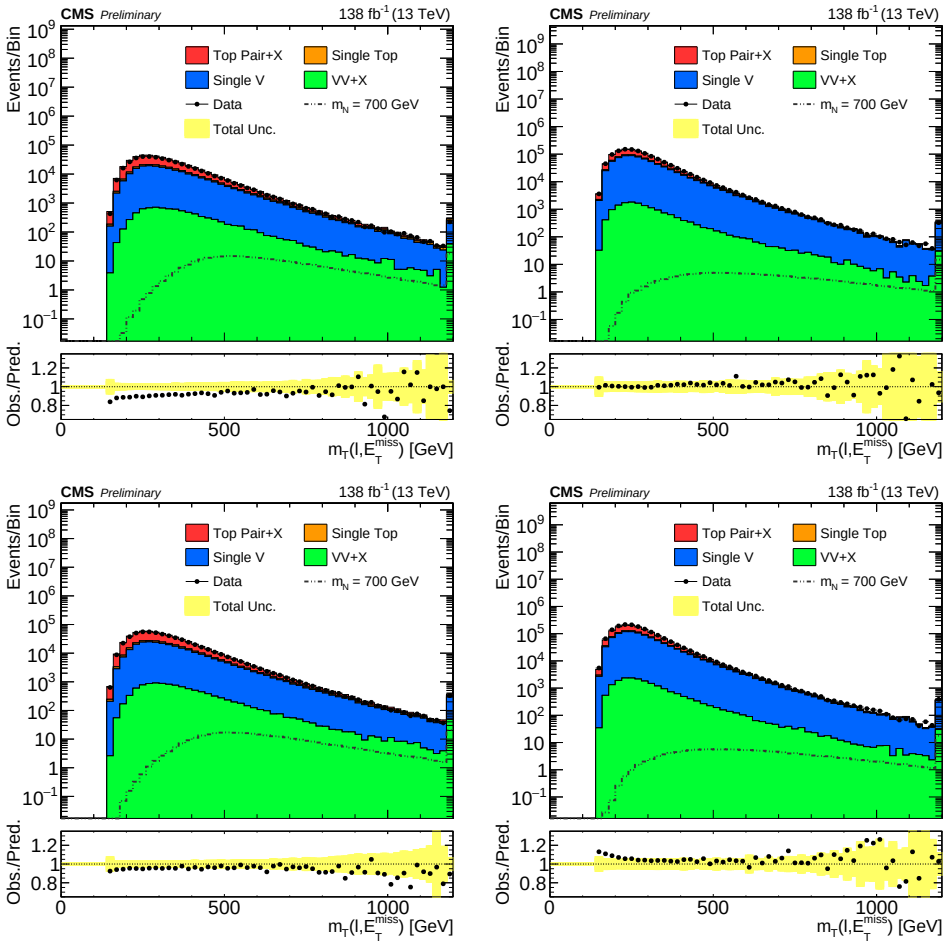

Figure 5.12: Transverse mass of lepton and missing transverse energy distributions for full Run 2 after preselection. The distributions are separately shown for electron merged (top left), resolved (top right), and muon merged (bottom left), resolved (bottom right) topologies.

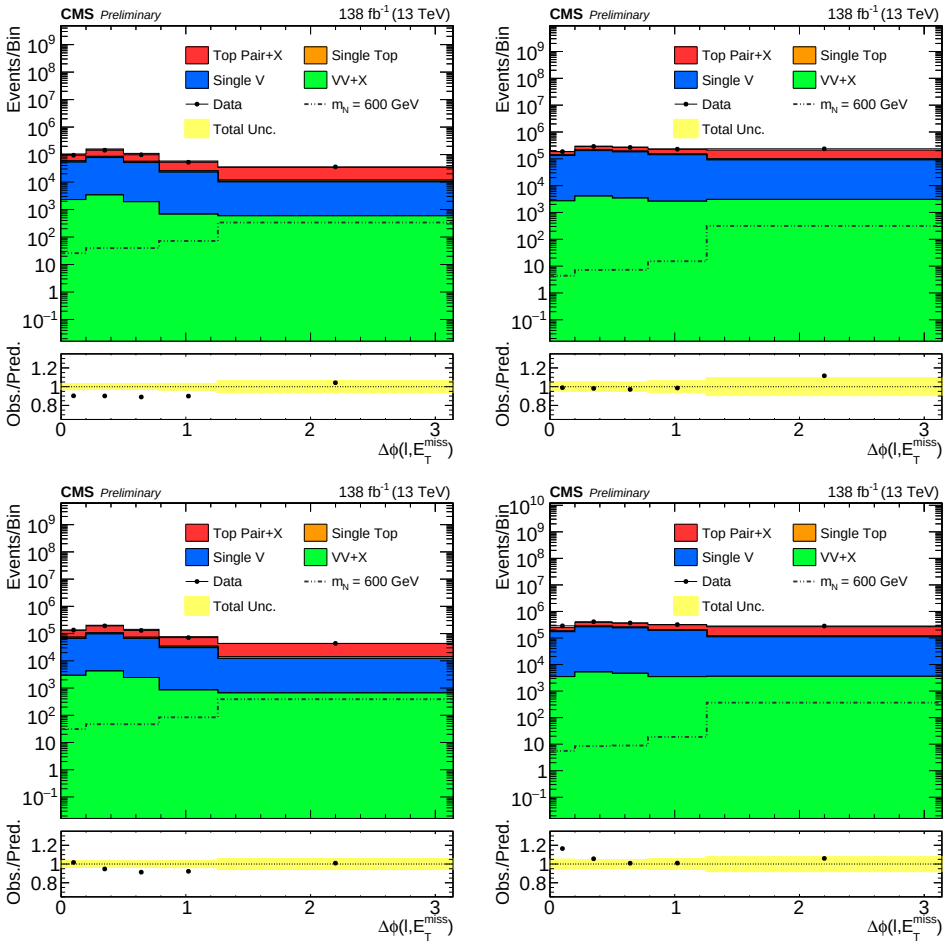

Figure 5.13: Azimuthal angle difference of lepton and missing transverse energy distributions for full Run 2 after preselection. The distributions are separately shown for electron merged (top left), resolved (top right), and muon merged (bottom left), resolved (bottom right) topologies.

145 GeV, inclusively considering all possible bosons from N. To denote the different jet modes of X^{reco} which will be explained below, the term $X^{\text{reco}}_{\alpha\beta}$ will be used. α and β are either b or q depending on the status of J or two j. For merged region where AK8 jet is used as a proxy for X, $m(X_{\text{reco}})$ is taken from the soft drop mass.

The major portion of the background events are from W and top pair processes. To suppress the background events with onshell W boson which includes those two processes, the transverse mass of a lepton and $\not{E}_{\text{T}}$, $m_T(\ell,\ \not{E}_{\text{T}})$, has to be greater than 250 GeV. In addition, the difference in ϕ between a lepton and $\not{E}_{\text{T}}$, $|\Delta\phi(\ell,\ \not{E}_{\text{T}})|$ at least needs to be 0.4π. These are expected to be effective discriminators for signal processes the background W boson and top pair processes. As both background processes include onshell a leptonically decaying W boson, applying $m_T(\ell,\ \not{E}_{\text{T}}) > 250$ GeV removes most of the events whereas $m_T(\ell,\ \not{E}_{\text{T}})$ in signal processes are not bound to the onshell W boson mass scale.

In order to meet the jet requirements for the background W boson process, it needs to have a large initial state radiations which in turn form jets. Accordingly, W boson is Lorentz boosted in the other direction, decaying into a lepton and a neutrino ($\not{E}_{\text{T}}$) with collimation. Similar logic also holds for the top pair process as well. Consider a semi-leptonic decaying top pair event with a lepton and $\not{E}_{\text{T}}$ that has small $|\Delta\phi(\ell,\ \not{E}_{\text{T}})|$. This small $|\Delta\phi(\ell,\ \not{E}_{\text{T}})|$ is due to enoughly large Lorentz boost of leptonically decaying W boson that is from a top quark which also can imply that the top quark is boosted. As top pair event yields back-to-back top quarks, this implies that the top quark on the other side is likely to be Lorentz boosted which becomes easier to make W boson reconstructed as an AK8 jet. However, the signal processes mostly have large $|\Delta\phi(\ell,\ \not{E}_{\text{T}})|$ as the lepton and the N which oscillates back to SM neutrino propagates into opposite directions.

For the signal region of the merged topology, events with the presence of additional b-tagged jet is discarded as two bottom quarks (for $X \to bb$ signal processes) are assumed to be incorporated, forming one AK8 jet. For further optimization of the merged signal region, ParticleNet tagger scores of the AK8 jet is used to further extract the signal events. The ParticleNet's mass decorrelated version of Xbb tagger ($pnet_{\mathrm{MD}}(Xbb)$) implies the AK8 jet is likely to be originating from resonance decaying into two b quarks. If the AK8 jet's $pnet_{\mathrm{MD}}(Xbb)$ is above 0.94, it falls into *MergedXbb* signal region. If an event fails the *MergedXbb* signal region, $pnet_{\mathrm{MD}}(Xqq)$ which is used to identify an AK8 jet that comes from resonance decaying into two light quarks is checked. If $pnet_{\mathrm{MD}}(Xqq) > 0.82$, the event belongs to *MergedXqq* signal region.

The signal region where events show resolved topology resembles that of the merged topology. The number of b-tagged AK4 jets among the two leading AK4 jets are checked to classify the resolved signal region : *ResolvedXbb*, *ResolvedXbq*, and *ResolvedXqq* which are the cases for 2, 1, and 0 b-tagged AK4 jets among the two leading AK4 jets in the event, resepectively.

The signal regions for both merged and resolved toplogies can be summarized as below :

- **Merged signal region**
  - Passes **merged preselection**.
  - $m_T(\ell, \not{E}_{\mathrm{T}}) > 250$ GeV.
  - $|\Delta\phi(\ell, \not{E}_{\mathrm{T}})| > 0.4 \times \pi$.
  - No b-tagged AK4 jet is contained in the event.
  - **MergedXbb** : $pnet_{\mathrm{MD}}(Xbb) > 0.94$.
  - **MergedXqq** : Fails **MergedXbb** but $pnet_{\mathrm{MD}}(Xbb) > 0.82$.

- **Resolved signal region**
  - Passes **resolved preselection**.
  - $m_T(\ell, \not{E}_{\mathrm{T}}) > 250$ GeV.
  - $|\Delta\phi(\ell, \not{E}_{\mathrm{T}})| > 0.4 \times \pi$.
  - **ResolvedXbb** : Both of the two leading AK4 jets are b-tagged.
  - **ResolvedXbq** : One of the two leading AK4 jets are b-tagged.
  - **ResolvedXqq** : None of the two leading AK4 jets are b-tagged.

The distributions of m_T^{eff} are plotted in $[450, 600, 750, 900, 900^+]$ GeV bins for merged and $[450, 600, 750, 750^+]$ GeV bins for resolved cases. This shape of m_T^{eff} is later used when performing the maximum likelihood fit with systematic uncertainties treated as nuisance parameters.

5.4.5 Control Region

Control regions are desinged for two major purposes. Firstly, to validate the contributions of major background processes in the signal regions. Thus W and top pair processes dominant control regions are separately defined. Secondly, to validate the modeling of the variables that are related to $\not{E}_{\mathrm{T}}$ in MC simulated events as it is a physics object which is intrinsically created through neutrinos but also a collective effect other mismeasurements and bias from the detector. Thus we invert the $m_T(\ell, \not{E}_{\mathrm{T}})$ and $|\Delta\phi(\ell, \not{E}_{\mathrm{T}})|$ requirements from the signal regions.

- **Merged control region for backgrounds**
  - W **process dominant**
    - Passes **merged preselection**.

    * $m(X_{\text{reco}}) < 65$ GeV or $m(X_{\text{reco}}) > 145$ GeV.
    * Event contains no b-tagged AK4 jet.
  - **Top pair process dominant**
    * Passes **merged preselection**.
    * $65 < m(X_{\text{reco}}) < 145$ GeV.
    * Event contains at least one b-tagged AK4 jet.

- **Resolved control region for backgrounds**
  - W **process dominant**
    * Passes **resolved preselection**.
    * $m(X_{\text{reco}}) < 65$ GeV or $m(X_{\text{reco}}) > 145$ GeV.
    * Event contains no b-tagged AK4 jet.
  - **Top pair process dominant**
    * Passes **resolved preselection**.
    * $m(X_{\text{reco}}) < 65$ GeV or $m(X_{\text{reco}}) > 145$ GeV.
    * Event contains at least one b-tagged AK4 jet.

As can be seen above, control regions for background estimations are defined in a way to have at least one criteria that makes it mutually orthogonal to other control regions and also to signal regions.

Control regions dedicated to check the modeling of $\not{E}_{\text{T}}$ has $m_T(\ell, \not{E}_{\text{T}})$ and $\Delta\phi(\ell, \not{E}_{\text{T}})$ requirements inverted from the signal regions.

- **Control region for $\not{E}_{\text{T}}$ modeling**
  - $m_T(\ell, \not{E}_{\text{T}}) \leq 250$ GeV or $|\Delta\phi(\ell, \not{E}_{\text{T}})| \leq 0.4 \times \pi$.
  - **MergedXbb** : Passes every other requirements in signal region's **MergedXbb** selection.

- **MergedXbb** : Passes every other requirements in signal region's **MergedXqq** selection.
- **ResolvedXbb** : Passes every other requirements in signal region's **ResolvedXbb** selection.
- **ResolvedXbq** : Passes every other requirements in signal region's **ResolvedXbq** selection.
- **ResolvedXqq** : Passes every other requirements in signal region's **ResolvedXqq** selection.

The distributions of m_T^{eff} are plotted in $[450, 600, 750, 900, 900^+]$ GeV bins for merged and $[450, 600, 750, 750^+]$ GeV bins for resolved cases. This shape of m_T^{eff} is later used when performing the maximum likelihood fit with systematic uncertainties treated as nuisance parameters.

5.5 Systematic Uncertainties

Two types of uncertainties determine the quality of the high energy physics result. Statistical uncertainty usually refers to the the stochastic fluctuations that comes from the finite resources that an experiment can provide which are background and signal MC simulated events, and observed data. From the experimentalist perspective, it's usually more important to control systematic uncertainty (and it is also somewhat related to how well experimentalist performed the analysis). Systematic uncertainty is an artifact of the apparatus that experiment uses (that by nature cannot be impeccable) or some of the assumptions that is made by the analyzers. The systematic uncertainties that are considered in the analysis for both signal and background will be discussed in the following sections in detail. Table. 5.2 shows how the systematic uncertainties are handled for signal and background MC simulated events. Systematic

uncertainty prescriptions basically follow the recommendations from CMS.

Sources	**Correlation**	**Shape**	**Signal**	**Background**
Integrated Luminosity	Correlated	No	Yes	Yes
Jet E Scale	Correlated	Yes	Yes	Yes
Jet E Resolution	Uncorrelated	Yes	Yes	Yes
Jet Mass Scale	Correlated	Yes	Yes	Yes
Unclustered E	Correlated	Yes	Yes	Yes
Muon E Scale	Correlated	Yes	Yes	Yes
Muon Identification	Correlated	Yes	Yes	Yes
Muon Isolation	Correlated	Yes	Yes	Yes
Muon Reconstruction	Correlated	Yes	Yes	Yes
Muon Trigger	Uncorrleated	Yes	Yes	Yes
Electron E Scale	Correlated	Yes	Yes	Yes
Electron E Resolution	Correlated	Yes	Yes	Yes
Electron ID	Correlated	Yes	Yes	Yes
Electron Reconstruction	Correlated	Yes	Yes	Yes
Electron Trigger	Uncorrelated	Yes	Yes	Yes
Pileup Weight	Correlated	Yes	Yes	Yes
Prefire Weight	Correlated	Yes	Yes	Yes
B-Tagging Corr. Heavy	Correlated	Yes	Yes	Yes
B-Tagging Corr. Light	Correlated	Yes	Yes	Yes
B-Tagging Uncorr. Heavy	Uncorrelated	Yes	Yes	Yes
B-Tagging Uncorr. Light	Uncorrelated	Yes	Yes	Yes
ParticleNet	Uncorrelated	Yes	Yes	Yes
PDF Replicas	Correlated	No	Yes	No
μ_F, μ_R Scales	Correlated	No	Yes	No

Table 5.2: Sources of systematic uncertainties and its year-by-year correlation treatments.

5.5.1 Systematic Uncertainties from Objects

Differences can be expected between data and MC simulated events for physics objects in the analysis and this needs to be taken into account as systematic uncertainties.

Leptons

Discrepancies in the reconstruction, identification, and trigger efficiencies between data and MC simulation are corrected by applying the scale factors to MC simulated events. This is measured using the tag-and-probe method, calculating the efficiencies in Z boson events by requiring two leptons. The scale factors depend on p_{T} and η, and $\pm 1\ \sigma$ variations are applied to take the difference in the yield as the systematic uncertainty.

Systematic uncertainty on lepton momentum scale is computed by varying the momentum of the leptons according to their assigned uncertainties. Muons are applied with Rochester muon momentum correction [107], removing bias from detector misalignments or magnetic fields. Rochester muon momentum correction is derived on two sequential steps : (1) Derive the p_{T} correction factors split into charge, η, ϕ bins by matching the mean value to the ideal $\mathrm{Z} \to \mu^+\mu^-$ MC sample that has no alignment effects. (2) The correction factors in η, ϕ bins are scaled so that the mean value of the dimuon invariant mass is matched to the ideal MC sample. Electrons also utilize the Z boson for its energy corrections. The bins are split into η and R_9 bins, where R_9 is the total energy deposit in the 3×3 crystal array in ECAL for a given supercluster, to derive the correction factors for the scale and resolution of electron energy by matching MC to data. The variation in momentum of leptons is later propagated accordingly to $\not{E}_{\mathrm{T}}$ for recalculation.

Jets

Jets, as a composite object from multitude of particle flow objects, innately carries the complexity of detector responses. In order to correctly describe the detector response of the jet in MC simulated events, corrections to the jet energy are applied in two ways : (1) Jet energy scale, shifting the overall mo-

mentum estimations in both data and MC. (2) Jet energy resolution, scaling the momentum resolution (width) in MC to match the distributions in data.

Jet energy corrections are applied in several consecutive steps. At the first stage, pileup contributions are subtracted by estimating the pileup offset by comparing the samples with and without pileup overlay. For the next step, assuming MC response from particle level jets to reconstructed jets is same in data, jet response correction is applied to both MC and data. These two steps attribute to the jet energy scale corrections. Finally, after applying the jet energy scale correction measured from the first two steps, jet energy resolution is corrected using the dijet and γ/Z+jet events. In this step, an overall shift in momentum is produced but it only widens the distribution as jet energy resolution in data is shown to be worse than that in MC.

In a same manner, jet mass scale for AK8 jets in MC also need to be corrected to properly describe the softdrop mass in data. This is an important factor of systematic uncertainties as the softdrop mass is the key variable when associating AK8 jets with hadronically decaying bosons. The correction scale factors for softdrop mass are measured using W boson tagged AK8 jets in top pair selected events.

Systematic uncertainties on jet energy scale and resolution corrections are separately treated by scaling or smearing the jet momentum with $\pm 1\ \sigma$ variations and analyzing how the shape of observables change. The varied amount of jet energy is propagated to $\not{E}_{\mathrm{T}}$ for recalculation.

B-Tagged Jets

The jets that originate from b quarks are identified using the DeepJet algorithm. To mitigate the difference in performance of the algorithm between in data and MC simulated events, additional weights are applied. Systematic uncertainty

of this weight is given in a relatively detailed manner, by first separating the yearly correlated and uncorrelated sources. It is further broken down into the cases where heavy flavor jets (b or c quarks) are identified as b-tagged jets and light flavor jets (gluon, u, d, c, or s quarks) are mistakenly identified as b-tagged jets, resulting in eight categories of systematic uncertainties in total.

5.5.2 Systematic Uncertainties from Detector Effects

Integrated Luminosity

The expected numbers of MC simulated events are subject to the integrated luminosity during the data taking period. Thus systematic uncertainties on the integrated luminosities are given 1.2 %, 2.3 %, and 2.5 % for 2016, 2017, and 2018, respectively.

Pileup Modeling

To correct the different description of pileups between data and MC simulated events, MC simulated events are weighted in number of pileups in the event. The cross section of minimum bias events is given 69.2 mb at 13 TeV and $\pm 4.6\%$ variations on the cross section is given to calculate the systematic uncertainties for the weight factor due to pileups.

Missing Transverse Momentum Modeling

PF candidates that are not associated to any of the physics objects of the event are left out as the unclustered energy. It is usually expected to be isotropic given that such particle flow elements can emerge in any random direction. We vary the estimated energy of such PF candidates up and down and estimate its effect to $\not{E}_{\mathrm{T}}$ in MC simulated events.

Prefiring Effect

Due to the radiations and the loss of transparency in ECAL crystals, a gradual timing shift was created which makes two consecutive bunch crossings to fire which is automatically vetoed by L1 triggers. This effect becomes significant for all types of physics objects in $|\eta| > 2.0$ region. To take this effect that occurred in data into account, MC simulated events are weighted and correspondingly systematic uncertainty is given to the weight factors.

5.5.3 Systematic Uncertainties from Theoretical Effects

Signal Modeling

Systematic uncertainties on signal modeling is evaluated by estimating the difference in acceptance of signal MC events in signal regions. We follow the procedures that is recommended by the PDF4LHC [108], varying factorization and renormalization scales as well as the PDF replicas accompanied to the nominal NNPDF 3.1 LUXQED NNLO PDF used for signal sample generation. As systematic uncertainty on scales and PDF are found to be small, conservatively 0.5 % and 0.1 % are given, respectively.

Top p_{T} Modeling

The distribution of top quark p_{T} in top pair events was found to be significantly softer in the actual data when compared to that of MC simulated event prediction calculated at NLO accuracy for matrix elements convoluted to parton showers. Although studies with higher fixed order corrections showed better descriptions, the discrepancy has not been fully resolved. To account for the discrepancy, top quark p_{T} is reweighted for the top pair process MC simulated events based on the measurements of the top p_{T} from CMS. Systematic uncertainty is assigned by taking difference of the distribution when reweighting is

not applied to the MC events.

5.6 Results

5.6.1 Effective Transverse Mass Distributions

$m_T^{\text{eff}} = m_T(\ell, \not{E}_{\text{T}}, X^{\text{reco}})$ distributions are shown for merged and resolved selections in their control and signal regions. This distribution is used to perform the maximum-likelihood fit when extracting the final sensitivity. In case of merged selection, m_T^{eff} is binned in four bins, whereas resolved selection's m_T^{eff} is binned in three bins. The difference of binning is due to the fact that resolved signal regions have lesser number of events from MC expectations.

- Merged m_T^{eff} bins : $[450, 600, 750, 900, 900^+]$ GeV
- Resolved m_T^{eff} bins : $[450, 600, 750, 750^+]$ GeV

For the merged selections, two signal regions for Xbb and Xqq tagged categories, two control regions for validations of W and top pair processes, and two control regions for validations of $\not{E}_{\text{T}}$ modelings for different jet tagged categories are studied. In a similar manner, for the resolved selections, three signal regions for Xbb, Xbq, and Xqq jet tagged categories, two control regions for validations of W and top pair processes, and three control regions $\not{E}_{\text{T}}$ modelings for different jet tagged categories are studied. The distributions of m_T^{eff} are shown in Figs. 5.14-5.27 overlayed with $m_N = 700$ GeV scenario from all signal processes.

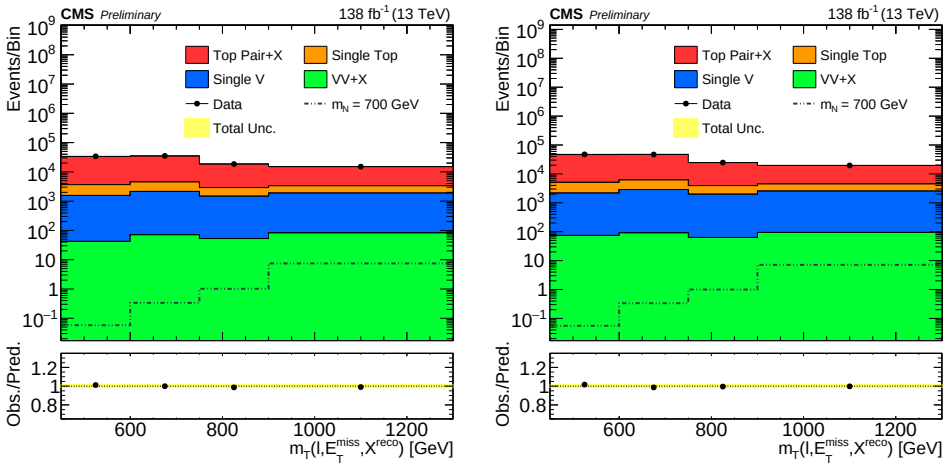

Figure 5.14: m_T^{eff} distributions for full Run 2 after top pair process dominant control region selections in electron (left) and muon (right) channels with merged event topology.

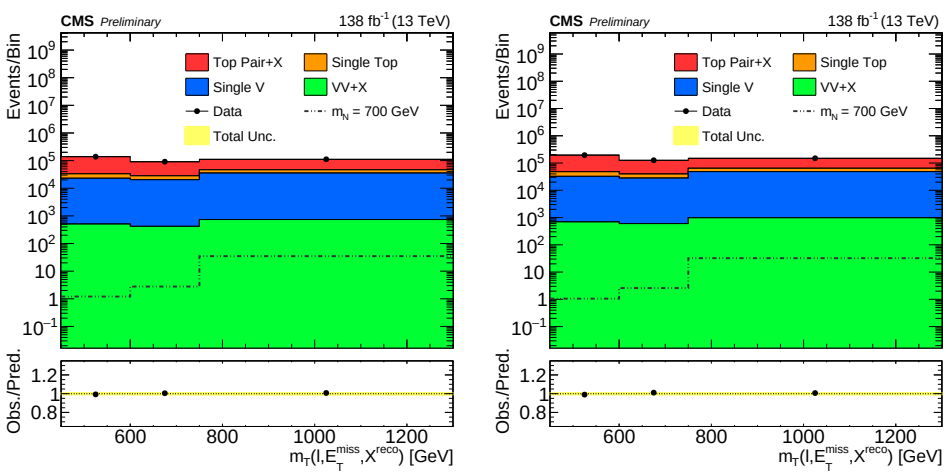

Figure 5.15: m_T^{eff} distributions for full Run 2 after top pair process dominant control region selections in electron (left) and muon (right) channels with resolved event topology.

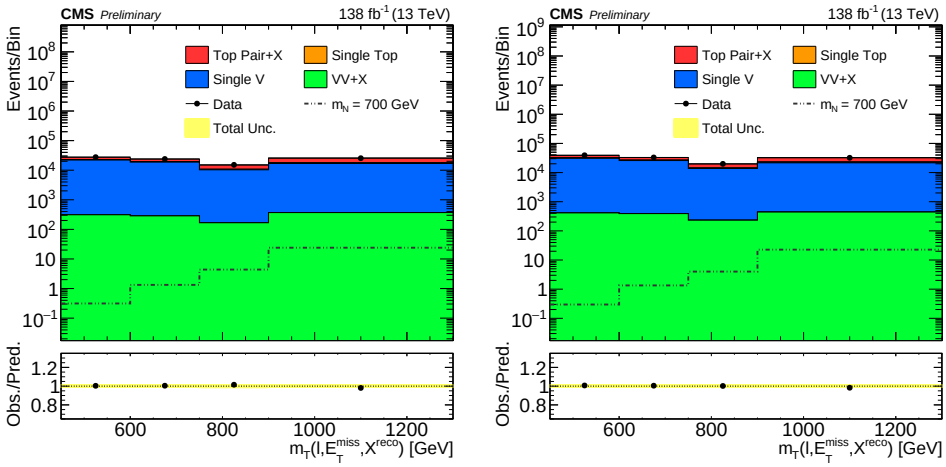

Figure 5.16: m_T^{eff} distributions for full Run 2 after W process dominant control region selections in electron (left) and muon (right) channels with merged event topology.

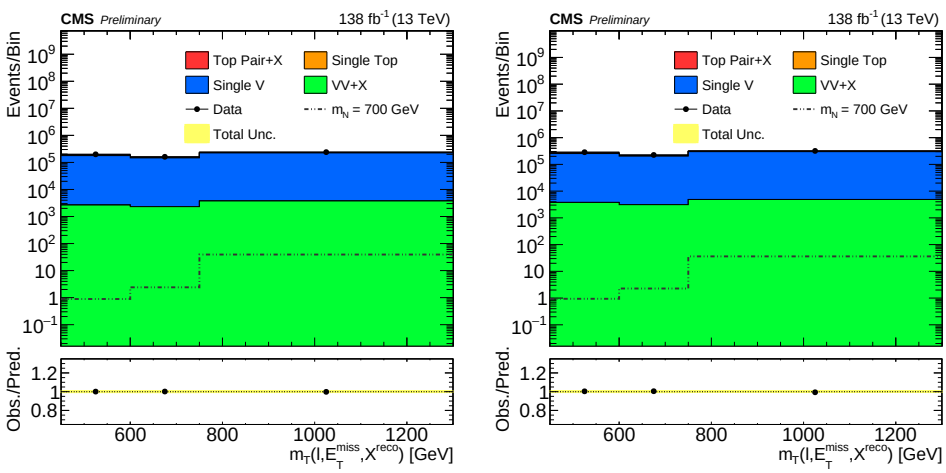

Figure 5.17: m_T^{eff} distributions for full Run 2 after W process dominant control region selections in electron (left) and muon (right) channels with resolved event topology.

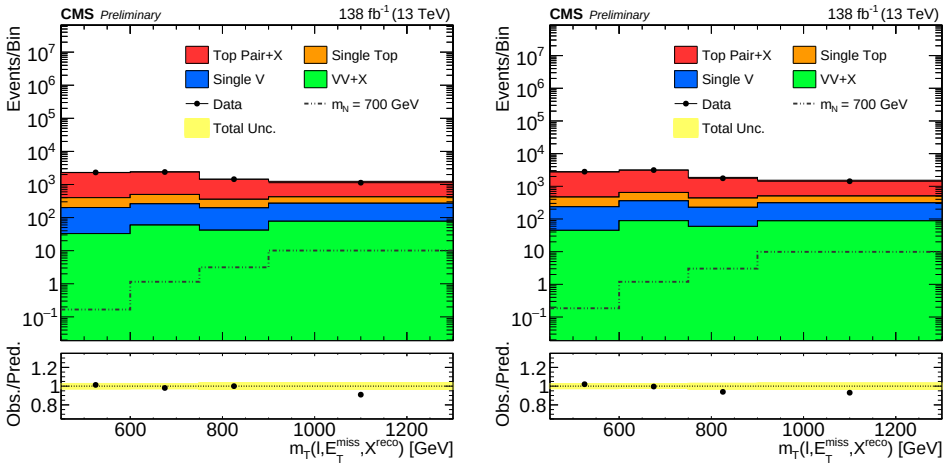

Figure 5.18: m_T^{eff} distributions for full Run 2 after Xbb $(\ell, \not{E}_{\text{T}})$ modeling control region selections in electron (left) and muon (right) channels with merged event topology.

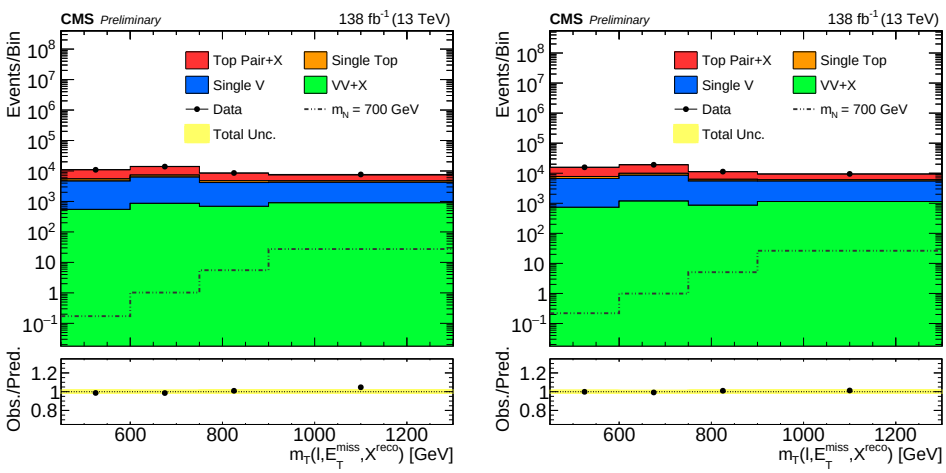

Figure 5.19: m_T^{eff} distributions for full Run 2 after Xqq $(\ell, \not{E}_{\text{T}})$ modeling control region selections in electron (left) and muon (right) channels with merged event topology.

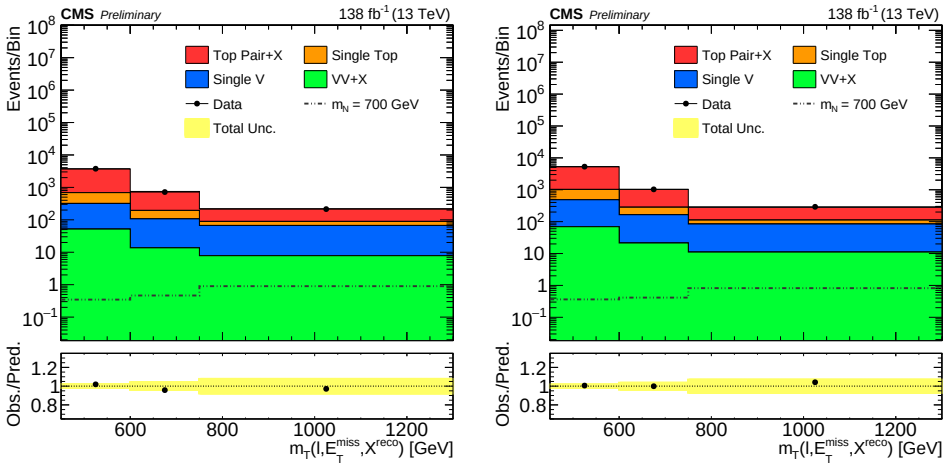

Figure 5.20: m_T^{eff} distributions for full Run 2 after Xbb ($\ell, \not{E}_{\text{T}}$) modeling control region selections in electron (left) and muon (right) channels with resolved event topology.

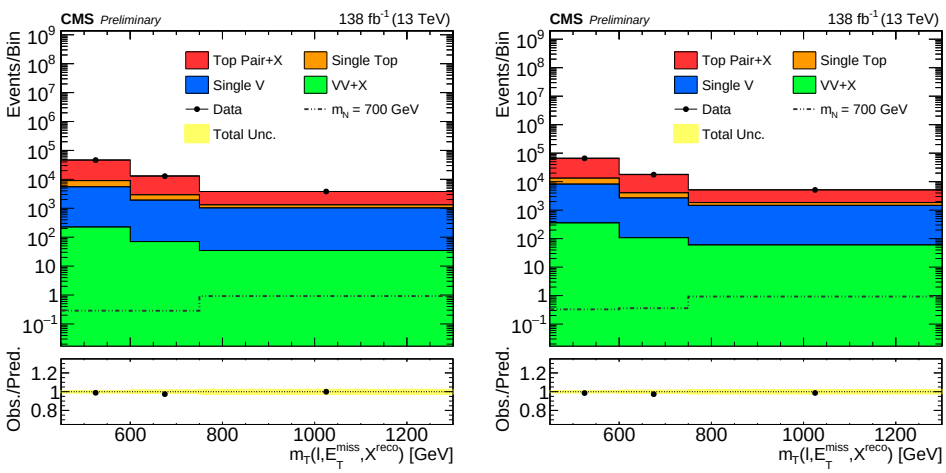

Figure 5.21: m_T^{eff} distributions for full Run 2 after Xbq ($\ell, \not{E}_{\text{T}}$) modeling control region selections in electron (left) and muon (right) channels with resolved event topology.

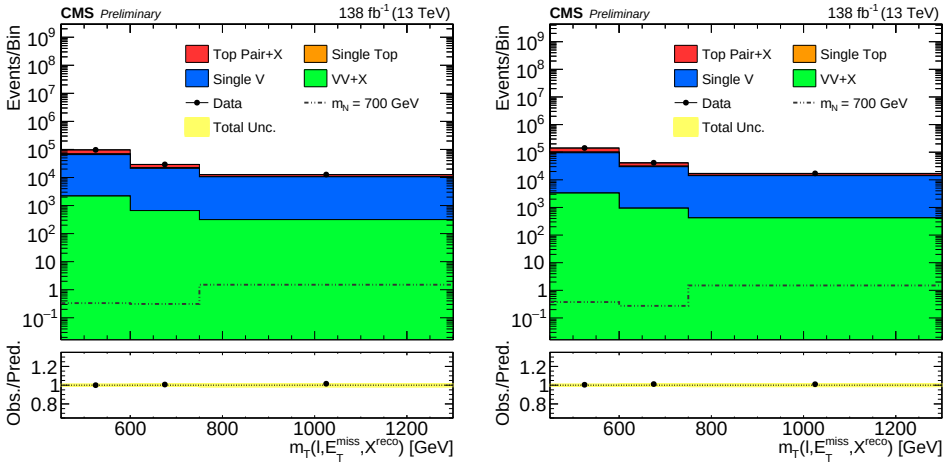

Figure 5.22: m_T^{eff} distributions for full Run 2 after Xqq ($\ell, \not{E}_{\text{T}}$) modeling control region selections in electron (left) and muon (right) channels with resolved event topology.

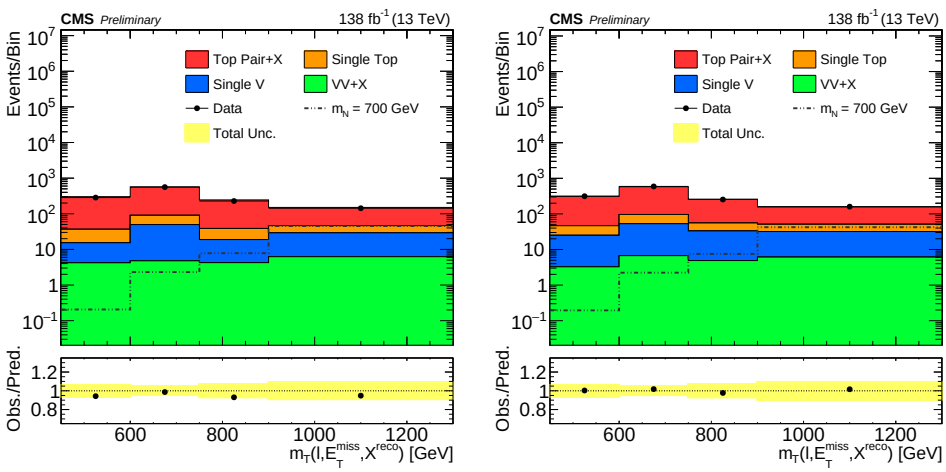

Figure 5.23: m_T^{eff} distributions for full Run 2 after signal region merged Xbb selections in electron (left) and muon (right) channels.

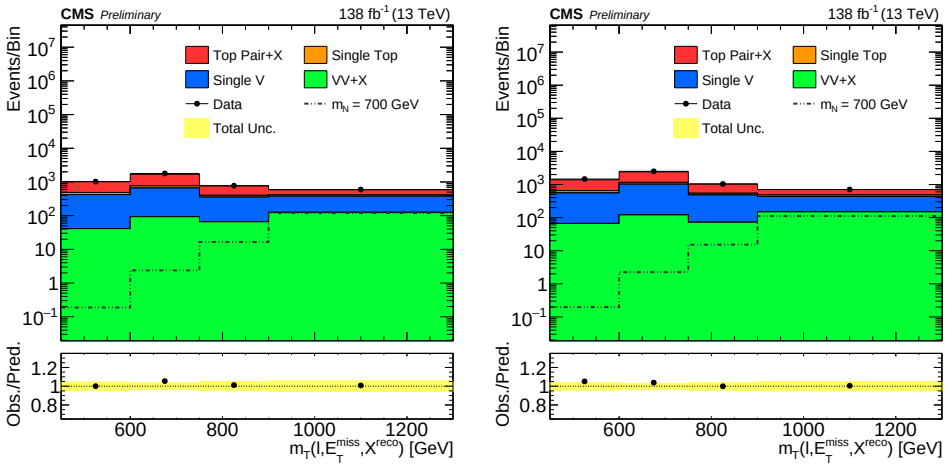

Figure 5.24: m_T^{eff} distributions for full Run 2 after signal region merged Xqq selections in electron (left) and muon (right) channels.

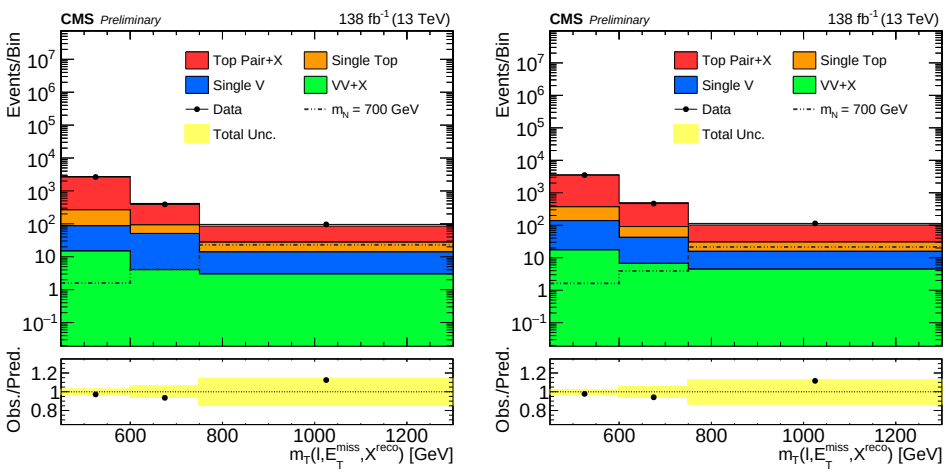

Figure 5.25: m_T^{eff} distributions for full Run 2 after signal region resolved Xbb selections in electron (left) and muon (right) channels.

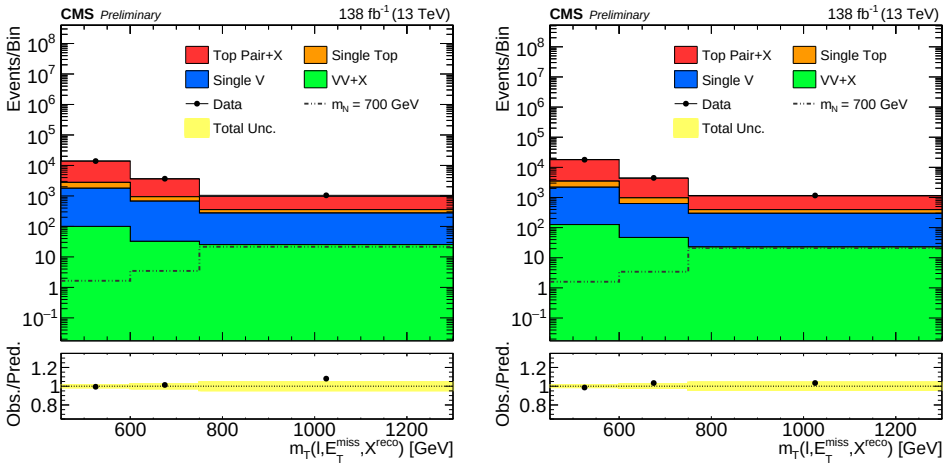

Figure 5.26: m_T^{eff} distributions for full Run 2 after signal region resolved Xbq selections in electron (left) and muon (right) channels.

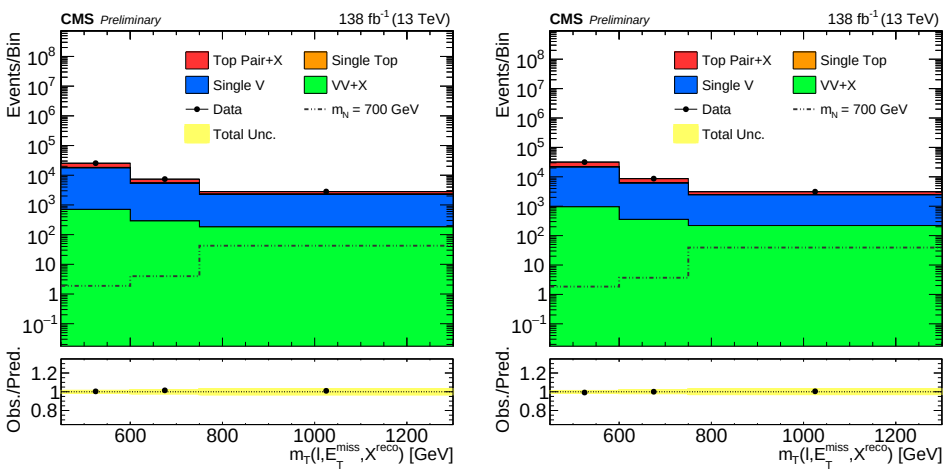

Figure 5.27: m_T^{eff} distributions for full Run 2 after signal region resolved Xqq selections in electron (left) and muon (right) channels.

5.6.2 Statistical Method

The result of an analysis needs to be performed with a statistical method to be interpreted with the postulated BSM physics. This basically compares how much the observed data are consistent between two hypotheses where one correspond to the *background-only* (null) hypothesis and the other corresponds to the *signal-plus-background* (alternate) hypothesis. Roughly speaking, statistical method infers how much the observed data is compatible with the background-only hypothesis and incompatible with the signal-plus-background hypothesis.

The likelihood function ($\mathcal{L}$) is defined with the signal strength (μ) which is the parameter of interest that differentiates the two hypotheses, and in this analysis represents the $|V_{\ell N}|^2$ value. In particular, the background-only hypothesis is obtained by assuming $\mu = 0$, effectively yielding 0 signal events. The likelihood function can be written as

$$\mathcal{L}(\text{data}|\mu,\ \theta) = \prod_{i\in\text{bins}} \frac{(\mu s_i(\theta) + b_i(\theta))^{n_i}}{n_i!} \exp\left(-\mu s_i(\theta) - b_i(\theta)\right) \prod_{j\in\text{nuisances}} p_j(\tilde{\theta}_j|\theta_j). \tag{5.5}$$

In the above equation, n_i, s_i, and b_i each represents the observed data, signal, and background event yields in the $i-$th bin. These values depend on the nuisance parameters (θ) which correspond to uncorrelated sources of uncertainty that the analysis needs to consider. Lastly, Gaussian distribution is chosen to represent the degree of belief that a value (θ_j) is compatible with our chosen value $\tilde{\theta}_j = 0$, which is denoted with p_j.

From the likelihood above, we can construct the test statistic ($\tilde{q}_\mu$) as

$$\tilde{q}_\mu = -2\ \ln \frac{\mathcal{L}(\text{data}|\mu\ ,\hat{\theta}_\mu)}{\mathcal{L}(\text{data}|\hat{\mu}\ ,\hat{\theta}_{\hat{\mu}})} \tag{5.6}$$

where $\hat{\theta}_\mu$ is the set of nuisance parameters maximizing the likelihood for a given signal strength μ. The $\hat{\mu}$ and $\hat{\theta}_{\hat{\mu}}$ are called maximum likelihood estimates that

are values of parameters that globally maximizes the likelihood.

CL_s method [109, 110] is adopted to derive the exclusion limits and computationally it can be written as

$$CL_s(\mu) = \frac{P(\tilde{q}_\mu \geq \tilde{q}_\mu^{\text{obs}}|\text{alternate})}{P(\tilde{q}_\mu \geq \tilde{q}_\mu^{\text{obs}}|\text{null})}. \tag{5.7}$$

Numerator and denominator each represents the probability to find the test statistic $\tilde{q}_\mu$ to be larger than the observed value $\tilde{q}_\mu^{\text{obs}}$ under alterante and null hypotheses, respectively. At a given signal strength μ, $CL_s(\mu)$ infers that the signal hypothesis is excluded at a confidence level of α when $CL_s(\mu) < 1 - \alpha$.

5.6.3 Interpretation

Binned maximum likelihood fit is performed with all control regions and signal regions defined in Chapter. 5.4 to extract the signal strength using Asymptotic method calculations. In addition, we introduce additional rate parameters which renormalizes W and top pair processes based on the observed and expected yields to remedy the residual data and MC simulated event discrepancies (see Fig. 5.28). The limits are set on the square of N mixing element as a function of mass of N, on a $m_N - |V_{\ell N}|^2$ plane, as shown in Fig. 5.29 and Fig. 5.30 for electron and muon, respectively. Blue and red dashed lines each represent sensitivities acquired from resolved and merged event topologies and black dashed line shows the sensitivity by combining the two. For $|V_{\ell N}|^2 = 1.0$ scenario, we exclude m_N up to 940 GeV and 906 GeV for muon and electron channels, respectively.

The sensitivity reaches with same $|V_{\ell N}|^2 = 1.0$ scenario are relatively poor compared to the previously reported results from direct N production analyses which is 1430 (1240) GeV in muon (electron) case as shown in Fig. 5.31. One of the main reason is due to the fact that the dominant background processes for

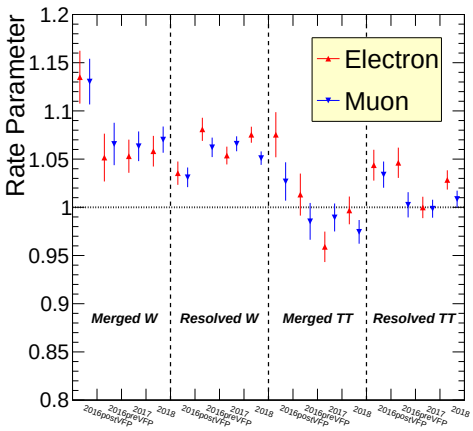

Figure 5.28: Rate parameters for W and top pair processes derived from its dedicated control regions in different years.

this analysis are W and top pair processes which are irreducible and inevitably degrades the performance of the analysis. On the other hand, other results from Ref. [82] and Ref. [89] which are depicted in brown and purple solid lines, respectively, take advantage of the lepton number violating signatures[4].

5.6.4 Summary

Search for heavy neutrinos in type-I seesaw model using single lepton final state is performed using full Run-II data at $\sqrt{s} = 13$ TeV collected with the CMS detector. The sensitivity is checked on a 2D plane parameterized with the mass of heavy neutrino and the mixing of heavy neutrino with SM neutrino (with electron or muon flavors).

The analysis is designed in a way to deal with two different cases which are merged and resolved event topologies. For the larger mass scenarios of heavy

[4]Ref. [89] includes both lepton number violating and conserving final states in the analysis. For example, both $\mu^+\mu^+e^-$ and $\mu^+\mu^-e^-$ final states are both used to extract the mixing limits in muon channel. In other words, these analyses take benefit from lower background level which is mostly composed of reducible nonprompt background contributions.

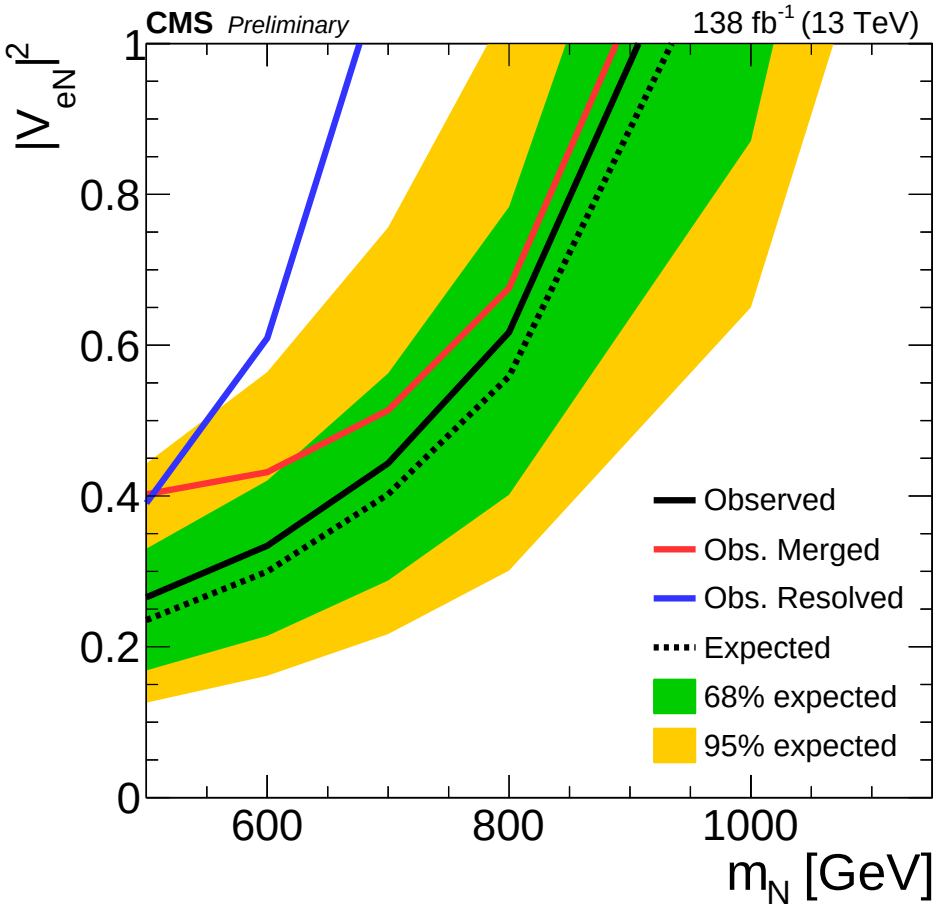

Figure 5.29: Exclusion limits on $m_N - |V_{eN}|^2$ parameter space.

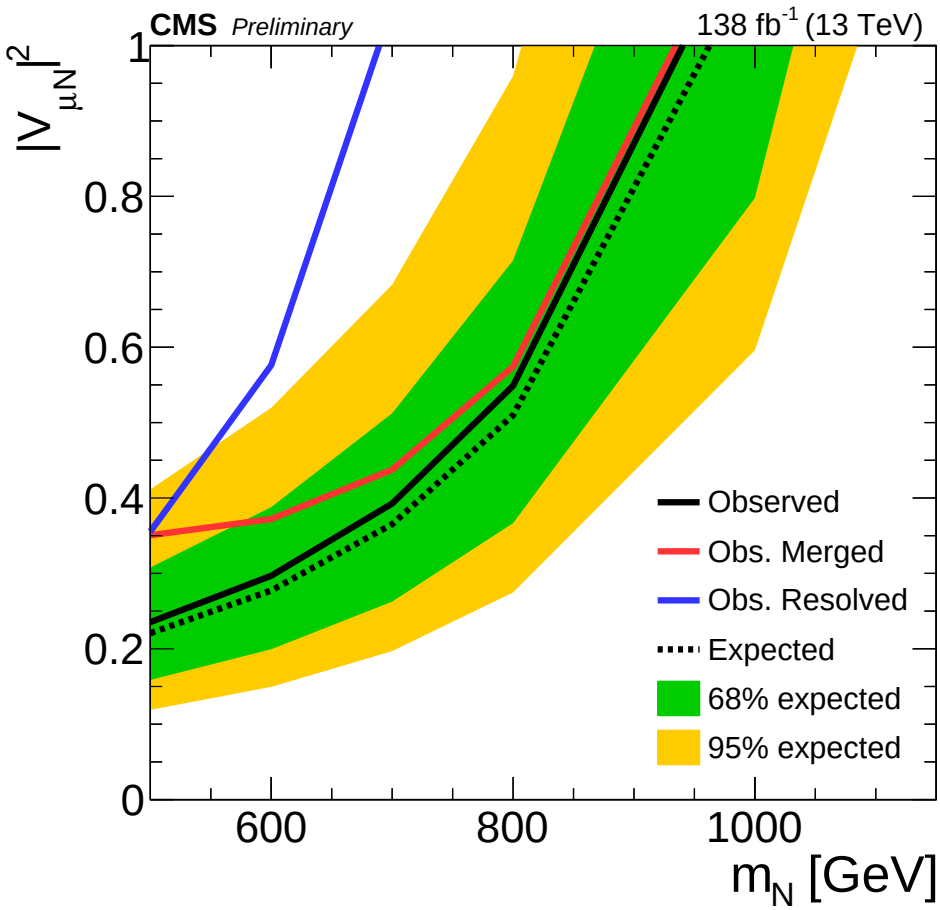

Figure 5.30: Exclusion limits on $m_N - |V_{\mu N}|^2$ parameter space.

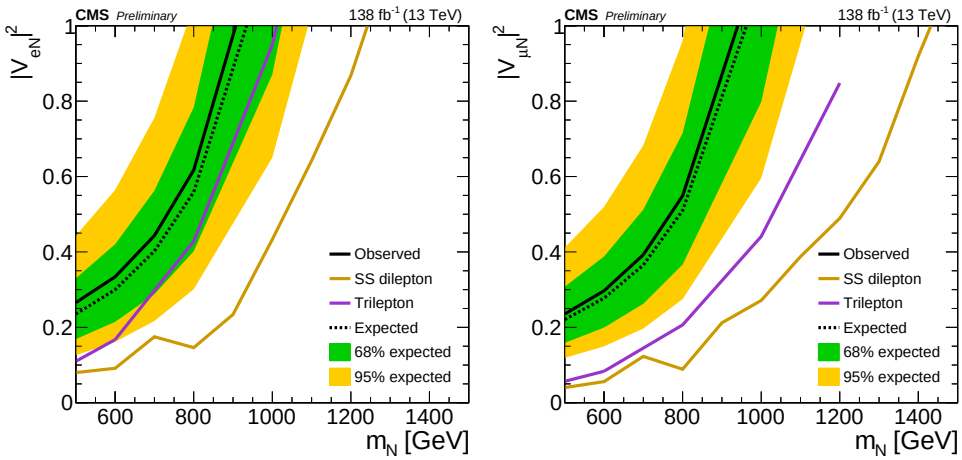

Figure 5.31: Exclusion limits on $m_N - |V_{eN}|^2$ (left) and $m_N - |V_{\mu N}|^2$ (right) parameter space. *SS dilepton* and *trilepton* refers to the observed limit results from Ref. [82] and Ref. [89], respectively.

neutrinos, bosons that decay from heavy neutrinos are likely to be Lorentz boosted. This then creates a collimation of the light or bottom quark pairs which are daughter particles of bosons which can be reconstructed by a single large-radius (AK8) jet. For the smaller mass scenarios of heavy neutrino, boson is less likely to be Lorentz boosted and instead create resolved light or bottom quark pair. Such cases can be reconstructed by using two small-radius (AK4) jets. Thus, either the large-radius jet or a pair of small-radius jets can be used as the proxy for bosons decaying from heavy neutrinos. In addition, a single lepton and large transverse missing energy are required to pick up signal-like events.

Binned maximum likelihood fit to data is performed simultaneously in all signal and control region event categories using the transverse mass of all signal objects which are lepton, missing transverse energy, and the proxies for boson (a large-radius jet or small-radius jets). The sensitivity of type-I seesaw model

are analyzed from heavy neutrino mass ranges 500 GeV to 1500 GeV. In the scenario where mixing of heavy neutrino with SM neutrino is 1.0, the observed 95 % CL_s limit reaches up to 940 GeV and 906 GeV for muon and electron flavors, respectively.

This is the very first search to use the Higgs decay mode of heavy neutrinos and also the first search at the LHC to explore the sensitivity of heavy neutrinos in single lepton final state. Probing Higgs decay mode is an important task for all fermions as its interaction strength with Higgs directly relates to how massive a fermion is. From this fact accompanied by the discovery of heavy neutrinos in type-I seesaw model, it becomes possible to give constraints to the SM neutrino mass ranges. In addition, despite the weakness of this analysis that cannot benefit from lepton number conserving signal processes unlike the other previous searches, it is shown that it still gives comparable limits which can be a complementary search.

Chapter 6

Neutrino Mass Model Reinterpretations with Published LHC Results

6.1 Reinterpretation

Although mathematically self-consistent and beautifully formed, it is already known that the SM cannot fully describe the universe. One example is the aforementioned experimental observation of neutrino oscillations in Chapter. 4 which violates zero mass hypothesis of neutrinos. However, although a vast pool of the BSM scenarios have been studied and tested against the LHC data, nothing in particular has been found which we can consider as a key to the BSM. The biggest problem particle physicists are now encountering is that this is the first time when there is no clear guidance or direction to take since the revolutions of the field starting from the early 20th century. The LEP, being the most powerful lepton accelerator ever built served its purpose by precisely measuring a lot of quantities of the SM, in particular the masses of W and

Z bosons. Later with the Tevatron, top quark was found and completed the third generation of quarks in the SM. Most recently, the LHC was built with an anticipation to find the Higgs boson and it was indeed found by the CMS and ATLAS collaborations.

Although current status can be thought of as a failure of the BSM tests we are performing with the LHC, this definitely is not the only way to look at it. There is a simple psychological test called *Is the glass half full or half empty?* which is used to infer whether a person is optimistic or pessimistic. Imagine there are three cups as shown in Fig. 6.1. There won't be too much argument if someone says *the glass on the left is empty* and *the glass on the right is filled*. However, depending on one's philosophical view the glass in the middle could be described in both terms *half-filled glass* and *half-emptied glass*. In the same spirit, what the LHC has achieved so far can be perceived as a huge success of the SM even in the extreme phase space and for the exotic event signatures rather than failed attempts to find the BSM physics.

Figure 6.1: Three glasses.

Here, we can realize how reinterpretation [111] can be beneficial to the development of BSM physics programs that are carried out within and beyond the scope of the LHC. It is needless to say there can't be a single favorable BSM scenario as long as it incorporates the SM and explain the anomalous experimental results that cannot be described with the SM. Therefore, every BSM theories have their own merits and reasons to be tested. But the problem is that many of the BSM scenarios have numerous parameters while we experimantalists can only focus on some subsets for practical reasons such as limited people power and time to investigate every possibilities. This is where reinterpretation can

be a powerful approach, devising a concrete method to preserve the scientific results and utilizing it wisely to exploit its full potential. In addition, it is natural to consider a paradigm shift from theory-driven approach to data-driven approach especially given that the promised dataset from the HL-LHC runs would be even remarkably larger compared to what we have accumulated so far.

Inverting the chain of workflow of experimentalists who sought the BSM physics can be considered as the methodology of reinterpretation. The workflow of experimentalists can be illustrated as shown in Fig. 6.2.

1. Take a benchmark BSM scenario to test against the SM and experimental data.

2. Extract the BSM signals by utilizing its characteristics to obtain the best sensitivity (by removing the SM backgrounds as much as possible).

3. Write a scientific research paper with details of the procedure.

To describe in more detail, experimentalists usually select some BSM Lagrangian of their interest from theorists. With available MC event generators, by importing the provided Lagrangians, both SM and BSM MC events are generated. The stacked SM and BSM MC events are then compared to the actual data obtained with the detectors at LHC to calculate the sensitivities. This finally turns into a scientific result by writing, which we often use the term *analysis*.

If we invert the chain as in Fig. 6.3, the workflow in the opposite direction can be described as

1. Take an analysis to use as the benchmark studies for reinterpretation.

2. Recast the results in the analysis with available preserved data.

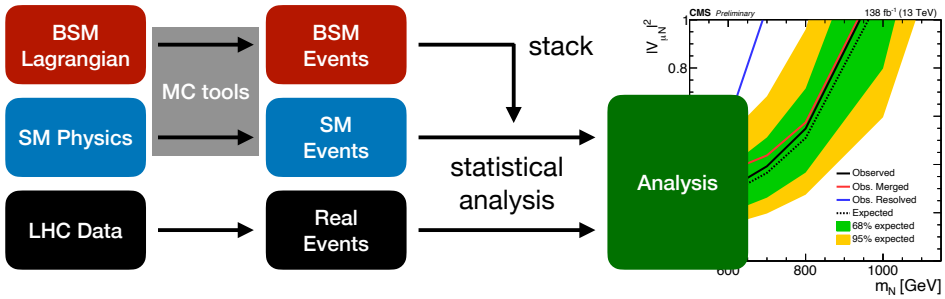

Figure 6.2: Typical workflow of BSM search analysis at the LHC. The BSM events are stacked to the SM events and statistical analysis is performed, checking deviations/agreements to the data acquired from the detectors.

3. Extract the sensitivity of a new BSM scenario that has not been tested yet using the recasted results.

For example, consider a case where one founds a new interesting BSM scenario that is accessible by the LHC experiment. For most of cases, it won't be difficult to find another existing analysis that shares similar event topologies with common final object although not exact. Then instead of going through time consuming and person power requiring usual workflow, one can try to recast the analysis through available frameworks such as CheckMate [112], MadAnalysis [113], SModelS [114], etc. If the recasting is successful, with generated MC events for the new BSM scenario of interest, one can stack this to the SM event yields provided and preserved by the recasted analysis to compute the sensitivity against the provided LHC event yields.

One important remark that should be noted here is that the workflow of reinterpretation can only be functional when the results of analysis paper is well preserved in an understandable format. This is usually done through the HepData [115] database. HepData is a web-based open-access repository created to collect multitude of experimental particle physics results in an unified

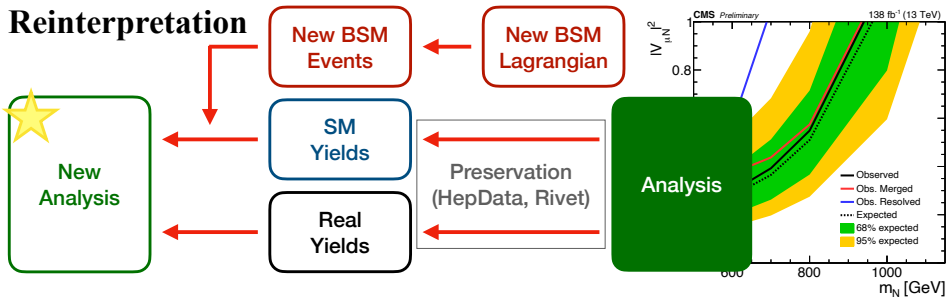

Figure 6.3: Inverting the workflow starting from already completed analysis from experiments. The new BSM events are stacked to the SM yields and statistical analysis is performed, checking deviations/agreements to the data yields that are well preserved.

format. Physics analyses from the CMS and ATLAS collaborations nowadays are required to at least provide the HepData materials such as tabulatization of the histograms. Some cases supply even more information with additional materials such as an inputs for MC event generator to generate the signal MC events or the built likelihoods that were used in the analysis to extract the final sensitivity.

Another way of preserving is through the Rivet [116] implementation. Rivet is a framework developed mainly to validate the MC event generators with analyses that are *unfolded*. Through the unfolding procedure, biases and inefficiencies introduced due to detector effects are corrected. This is advantageous in a sense that we can estimate what truly happened at the fundamental interaction point which does not depend on which detector was used to capture the event. The database of Rivet is built up with analyses from various experiments (but all corrected with detector effects of experiments) encoded in a C^{++} format which can be executed with MC event generator outputs after parton shower and hadronization have been processed. Thus, the inputs do not have

to go through the detector simulation such as DELPHES or GEANT 4 toolkits.

6.1.1 Tools for Reinterpretation

There are various software tools for reinterpretation on the market. Here, we will only discuss three tools among them and internal efforts from CMS and ATLAS, briefly introducing and comparing them in order to illustrate the different methodology and philosophy that drives the motivation for each of the tools. The details for most of the tools are explained in Ref. [111] and while we will discuss the tools SMODELS and MADANALYSIS, CMS and ATLAS studies for the phenomenological Minimal Supersymmetric Standard Model (pMSSM) and ATLAS developements for RECAST which is a framework to preserve the ATLAS analyses. CONTUR will be described separately later on in this chapter as it will be the framework utilized for the studies in the thesis.

The most specialized feature of SMODELS [114] is that it does not employ MC event generations unlike most of the other tools. SMODELS database is built on the LHC results that have been analyzed with Simplified Model Spectra (SMS) framework. The SMS framework is a powerful description to understand the experimental limits with small number of parameters, providing a chance to easily interpret the BSM searches with multiple parameters such as SUSY. The overall workflow of SMODELS is shown in Fig. 6.4. SMODELS starts from decomposing the event topologies of a newly proposed BSM with $\mathcal{Z}_2$ symmetry, $P \to P' + SM$ where P and P' are particles that arise from the postulation of the BSM. Every event topologies corresponding to particular Feynman diagrams have their weights computed by its production cross sections and branching ratios that are given by the new BSM of interest. Finally, SMODELS combines event topologies with common final state particles and compare upper limits of the cross sections stored in the database. This method has advantage over other

methods in terms of computing time and resources as it does not require the actual MC event generation but only computation for cross sections. An easily foreseen disadvantage of this method is that a hard assumption is required that event topologies in the database and the signal processes that match the SModelS database are pretty much similar. This is up to the users to check if it is true and thus the results should be taken in a conservative manner.

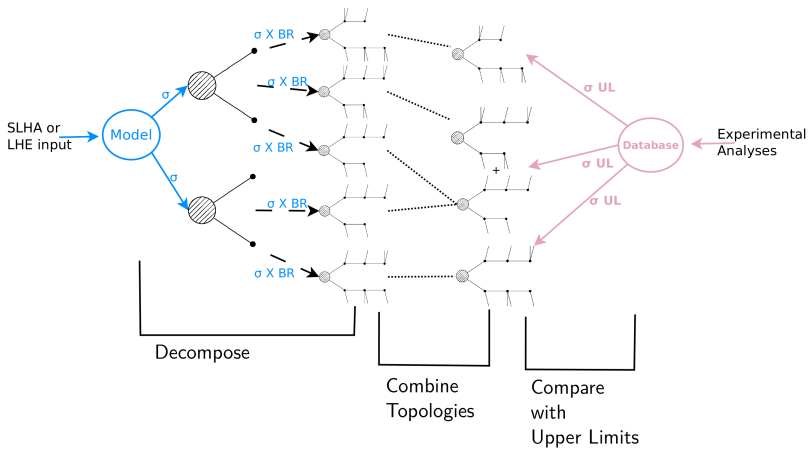

Figure 6.4: Schematic view of SModelS workflow. The figure is taken from Ref. [114]

MadAnalysis [113] is in some sense more full scale framework which is very similar to the workflow that experimentalists are more familiar with. Delphes or its own smearing function is used to emulate detector effects for MC event samples that went through parton showering and hadronization are injected to MadAnalysis. All analyses that are recasted through MadAnalysis framework are available in Public Analysis Database (PAD) which serves as its database for reinterpretation studies. PAD mostly relies on the BSM analysis results from the LHC which do not correct for the detector effects through unfolding method. The recasting consists of two parts : (1) Implementing the

analyzer code in a C^{++} format based on the reading of analysis that one wants to recast. (2) Validating it with materials such as cutflow tables [1] mostly provided in HepData, ensuring each steps of object or event selections in the original analysis and the recasted work shows reasonable agreement. If recasting is successful, the recasted work is registered to PAD and ultimately can be used for reinterpretation studies.

There are also several other attempts carried on internally within CMS and ATLAS collaborations which are the legacy analyses collecting all the efforts for SUSY searches, namely the pMSSM studies [117, 118]. As SUSY model is composed of more than 120 parameters by theory, experimental searches usually bound themselves to MSSM where it is reduced to substantially smaller number of 19 parameters. Thus, it becomes very much inspiring to explore the other remaining 80 % SUSY parameters although it would likely take huge amount of time and effort to make it happen. pMSSM study is built on this interest and motivation, scanning full potential of SUSY analyses. It begins with identifying the parameters and the parameter spaces that have not been covered by the existing analyses. Once the list of analyses and parameters are identified, pMSSM analyzers generate the signal MC samples accordingly and contact the corresponding analysis analyzers for cooperation, that is running the same analysis code with the newly delivered signal MC samples. After multiple iterations with SUSY analyzers, pMSSM assembles the delivered results which are the yield tables in signal regions and perform a full sensitivity study. As it needs long time and large computing resources to make the full simulated MC events as the other typical CMS or ATLAS analyses do, pMSSM study is usually accompanied by the development of fast simulation techniques [119].

[1] A cutflow table is a table of event yields, providing information of how much fraction of events is kept after each event selection requirements in the analysis.

Recently, ATLAS has further developed a framework for reinterpretation called RECAST which can mitigate one of the major disadvantages that is innate in the pMSSM studies. pMSSM first of all needs the support from the whole collaboration, most importantly from the SUSY analyzers of the SUSY analysis that are considered to be included in the pMSSM study. However in many occassions, this is the ideal case and things are not as easy as expected. For example, analyzers might have left physics, might have lost their computing environment or resources that they used for the analysis, and so on and so forth. These all are potential threats to the pMSSM, leading to a stall for unpredictable amount of time and poor performance of the results. In light of this, RECAST framework was built to facilitate the preservation and reinterpretation through a more stable channel, having a centrally maintained infrastructure by the ATLAS collaboration so that studies like pMSSM no longer has to rely on individual analyzers whom might be missing in the future for various reasons. Interesting example work done with RECAST is in Ref. [120] where ATLAS reinterpreted the mono-Higgs analysis targeting a model which predicts DM production in association with a dark Higgs boson decaying into pair of bottom quarks, sharing the same final state objects.

6.1.2 Constraints on New Theories Using RIVET (CONTUR)

Recall the analogy of *half-filled* or *half-emptied* glass of water, despite yet failed attempt to find the direct evidence of the BSM at the LHC, understanding of the SM has prospered thanks to the results from the LHC. In light of this, CONTUR [121] aims to test the BSM physics using the well measured SM results that are preserved in the RIVET package [116]. Underlying premise here is that if the BSM in consideration were to be true, it would have additional Lagrangian terms that would likely distort the SM measurements showing some deviations

from the measurements already although the effect could be small. That being said, the goal of CONTUR is to answer the question *How much a proposed BSM physics is compatible with published LHC results.*

CONTUR is special in a way that it neither utilizes the signal topology decomposed cross sections (SMODELS) nor the recasted workflow of the analysis (MADANALYSIS) which are based on the BSM efforts from CMS and ATLAS. Instead, as aforementioned, CONTUR exploits the full potential of the SM efforts by making use of the well preserved data in RIVET. As RIVET preservation is done after unfolding procedure from the experiment collaborations, particle level studies without introducing any detector effects become available.

More details on the computational setup for CONTUR in between each step are in Ref. [121] while here only the main stream workflow of it is explained in four different steps as below :

1. Defining model parameters and processes : Define which parameters in the proposed BSM physics to be scanned as well as the physics processes that involves the BSM particle of interest.

2. Calculating observables : Generating MC simulated events for the proposed BSM physics with parameter grids defined previously. The output in HepMC format is fed into RIVET to analyze contributions from the proposed BSM physics in the measurements from LHC.

3. Evaluating the likelihood : Measurements from LHC are grouped into orthogonal pools and from each pools, single bin that provides the largest CL_s is picked up. The bins from pools are later combined assuming uncorrelated.

4. Visualization of parameter space : External tools built on MATPLOTLIB

aids the visual understanding of results in 2D. It also visualizes the distribution of observables that provide the sensitivity.

CONTUR uses a χ^2 test statistic from the ratio of the likelihoods that a given measurement was obtained under competing assumptions, signal+background against background only. There are two different modes that differently defines the *competing assumptions* which is the background as explained below :

- *SM as background* mode : Sets the MC predictions under SM assumption as background.
- *Data as background* mode : Sets the observed data as background.

SM as background mode is the default mode that CONTUR adopts. This mode has a drawback that it can only be performed when the analysis in RIVET that we scan on also provides the MC predictions which is not always the case. However, it makes a reasonable assumption that data we observe is not necessarily bound to the SM.

There also exists several limitations in CONTUR. One important note to make is that it cannot be used as a tool for discovery of the BSM physics. Most of the analyses that RIVET contains lean toward the measurements that are shown to have a certain level of agreement with the SM. Thus it cannot identify the BSM physics favored parameter or phase spaces which can be more sensitive to the BSM physics. More fundamental reasons for limitations stem from the incomplete information that experiments provide such as the correlation between different bins for the systematic uncertainties which degrades the sensitivity that CONTUR provides.

Obviously if the analysis is not preserved in RIVET or even if is preserved but not in a usable way are the limiting cases for CONTUR. Some cases of limiting RIVET analyses are

- Analysis on ratio measurements : One of the powerful method of particle-level measurements involving neutrinos is by measuring the ratio of $\ell\ell$ with jets (as a proxy for DY$\rightarrow \ell\ell$) to $\not{E}_{\mathrm{T}}$ with jets. Such approach brings greater precision by the cancellation of several systematic uncertainties. But as $\ell\ell$ production in the denominator is hard-coded, the changes in production of $\ell\ell$ due to proposed BSM physics is not treated properly.

- Data driven estimation of backgrounds : H $\rightarrow \gamma\gamma$ channel was the key to success of discovering H boson at the LHC as it provides a clean final state with a sharp invariant mass peak at a great precision. The fit to the mass of diphoton continuum was used for background estimation for the H boson discovery. Although this was a powerful methodology, proposed BSM physics introducing a nonresonant diphoton production would have a wrong estimate of the sensitivity.

- Analysis with neutrino implications in particle level : For measurements of the WZ diboson production which can be enhanced in performance by calculating the event kinematics with neutrino. However, this implication makes it difficult to evaluate the impact of $\not{E}_{\mathrm{T}}$ which can be crucial for some of the BSM physics such as the dark matter.

6.2 Reinterpretation with Published LHC Results

6.2.1 Type-I Seesaw Model

Analysis on type-I seesaw model using CONTUR proceeds as follows. The Lagrangian of the model is provided in the FeynRules database [122] which is imported into MADGRAPH_AMC@NLO for the matrix element calculations. The processes that can yield N through charged or neutral current DY channel are considered at 7, 8, and 13 TeV center-of-mass energy. The decay of N to

H boson is not considered due to miscalculation of the H boson decay width. The generated hard process events are then passed to PYTHIA 8 for further leading-logarithmic calculations, showering and hadronization, which in turn gives out the files in HEPMC format. RIVET is fed with the HEPMC files so that their contributions to the LHC measurements can be evaluated. Finally, CL is calculated at which parameter point of the BSM physics is disfavored and can possibly be excluded by the LHC measurements. This last procedure is done by CONTUR.

Here, three physics cases related to type-I seesaw model are introduced : 2D scan performed on $m_N - |V_{\ell N}|^2$ parameter planes for each lepton flavors ($\ell = \mathrm{e}, \mu, \tau$). Newly developed MADGRAPH_AMC@NLO v3.4.0 which provides an user-friendly interface with RIVET and CONTUR is used for the scan. Matrix element calculations use the type-I seesaw model Lagrangian imported from UFO model which are convoluted with NNPDF 3.0 LO PDF [123]. The events in LHE format is passed to PYTHIA 8 for showering and hadronization simulations. By default, the tune is set to Monash 2013 tune while multiparton interactions are disabled for faster simulation. The outputs from PYTHIA 8 then runs on RIVET analyses to check the type-I seesaw model's contribution to LHC Run I and Run II measurements. Ultimately, CONTUR is used to compute the final sensitivity of the model.

Both CC DY and NC DY channels which yields heavy neutrino in association with a lepton or a neutrino are inclusively considered. Heavy neutrino further decays into three fermions which is via W or Z bosons. Due to LO calculations of the H width away from the known measured values which potentially might cause wrong interpretation of results, H boson decay channel of heavy neutrinos is discarded. The steering syntaxes in MADGRAPH_AMC@NLO for $m_N - |V_{\mu N}|^2$ parameter scan as an example is below :

```
import model SM_HeavyN_Gen3Mass_NLO
define mm = mu+ mu- vm vm~
define ee = e+ e- ve ve~
define tt = ta+ ta- vt vt~
define vv = ve vm vt ve~ vm~ vt~
define ll = e- mu- ta- e+ mu+ ta+
define qq = u d c s u~ d~ c~ s~
define ff = ll vv qq
generate p p > mm n1, n1 > mm ff ff
```

The mass of heavy neutrino are scanned in $[10 - 1000]$ GeV ranges for muon and electron cases while for tau case, it scans smaller mass range $[5 - 500]$ GeV both with eleven logarithmic steps. Square values of mixings are scanned in range $[0.01 - 1.0]$ with eleven logarithmic steps.

The CONTUR scan results on type-I seesaw model are shown in Fig. 6.5. It can be easily noticed that the excluding parameter space is significantly smaller for tau only mixing scenario. Two points can be inferred from the results : (1) Most of the SM physics programs at the LHC usually consider both muon and electron measurements when carrying out a measurement analysis in the EW sector. (2) The dominant analysis pools in tau results are mostly similar to those of electron or muon results which shows the sensitivity is driven from measurements using light leptons not hadronically decaying taus.

In the future, LHC would provide more and more results especially during the HL-LHC era. As there is no huge jump in the beam energy compared to Run II, LHC would likely benefit on better precision from growing statistics. However, as evident in the CONTUR scans on Run I and Run II results, there is a huge imbalance in the leptonic sector which does not yet fully exploit the hadronically decaying tau objects. As a lot of BSM ideas, especially involving additional Higgs sectors, have great expected sensitivity in tau channels, exploring the measurements with such objects that could eventually be preserved

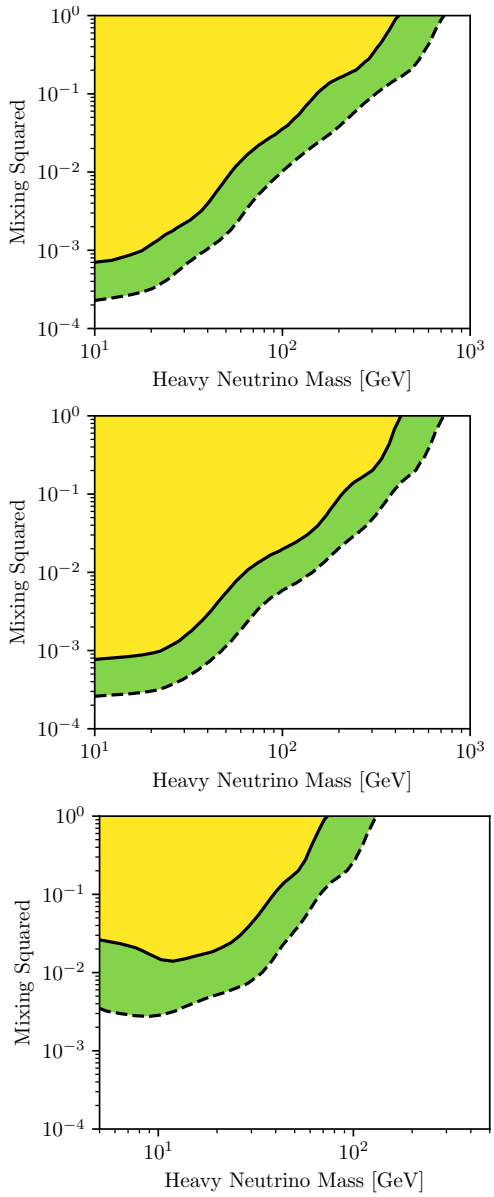

Figure 6.5: Exclusion limits set on $m_N - |V_{N\ell}|^2$ plane. From the top to the bottom are exclusion limits for electron, muon, and tau only mixing scenarios, respectively. Yellow area enclosed with solid line and green area enclosed with dashed line each shows 95 % and 68 % CL exclusions, respectively.

through RIVET would be one of the big goals that can fill in the that LHC is yet currently missing.

6.2.2 Type-II Seesaw Model

W Boson Mass Measurements

With the growing precision level of the EW observables in SM, measurement of W boson mass has become much more important over the last few decades [124]. W boson mass was preliminarily measured by UA1 and UA2 at SppS at CERN after its very first discovery [125, 126]. Later, both CDF and D0 measured m_{W} using the Tevatron Run 1 data which were later combined to be 80456 ± 59 MeV [127]. In parallel, LEP II had its energy increased above the Z boson mass allowing the production of W boson pairs. The turn-on spectrum of cross sections as a function of the collision energy at LEP II near the m_{WW} threshold and the direct measurements from fully and semi-leptonic decays of WW process were able to yield the value of 80376 ± 33 MeV [128]. CDF performed further precise measurement which combined the results from CDF and D0 that resulted in 80387 ± 16 MeV [129].

Later at the LHC, ATLAS performed the first m_{W} measurement using the data recorded in 2011 at $\sqrt{s} = 7$ TeV corresponding to an integrated luminosity of 4.6 fb^{-1} [130]. By a thorough calibration procedure based on the Z boson studies, detector response to electrons, muons, and the recoil was precisely modelled. By combining the m_{W} measured from p_{T} of lepton and transverse mass distributions, in electron and muon decay modes, the combination yielded 80370 ± 19 MeV along with a measurement of $m_{\mathrm{W}^+} - m_{\mathrm{W}^-}$ being -29 ± 28 MeV. LHCb reported their first measurement of m_{W}, 80354 ± 31 MeV, at $\sqrt{s} = 13$ TeV [131] from a simultaneous fit of the q/p_{T} distributions from W boson event candidates and ϕ^* distributions from Z boson event candidates both in

muon channels.

Presently accepted m_{W} values from LEP II and Tevatron combination and the world average by taking the measurement from ATLAS [2] at $\sqrt{s} = 7$ TeV are as below :

$$m_{\mathrm{W}}^{\mathrm{LEP\ II\ +\ Tev.}} = 80385 \pm 15 \ \ \mathrm{MeV}, \tag{6.1}$$

$$m_{\mathrm{W}}^{\mathrm{World}} = 80379 \pm 12 \ \ \mathrm{MeV}. \tag{6.2}$$

Both of the values showed some tensions, being not totally consistent with the SM prediction, $m_{\mathrm{W}}^{\mathrm{SM}} = 80357 \pm 6$ MeV [132]. This tension was enhanced further by the report from the CDF's latest measurement result in 2022 which was $80387 \pm 12(\mathrm{stat}) \pm 15(\mathrm{syst})$ MeV which deviates from $m_{\mathrm{W}}^{\mathrm{SM}}$ by 7σ [133].

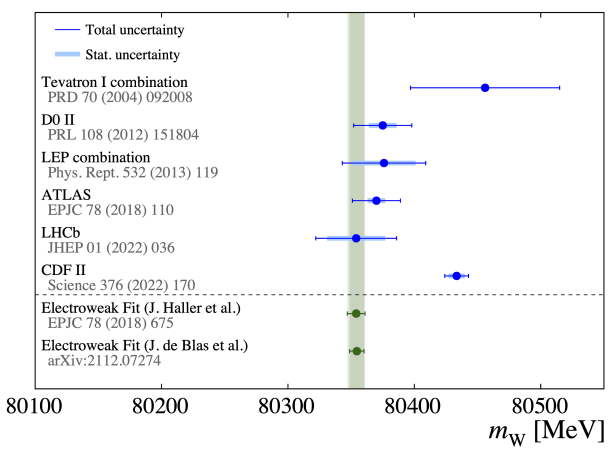

Figure 6.6: The measurements of W boson mass from various experiments. The dark and light blue represents total and statistical uncertainties, respectively. The SM prediction is in green.

Several suggestions have been made to account for this deviation from $m_{\mathrm{W}}^{\mathrm{SM}}$ under the SM context, such as with improvements in PDF [134] or perturbative

[2] LHCb's measurement is not included and CMS has not yet reported the measurement studies

matrix element calculations [135, 136]. However, these studies so far show that this discrepancy cannot be fully covered. Alternative SM explanations from the width of W boson (Γ_{W}) rather than the mass itself is yet unexplored. Since the measurement from CDF was from the one parameter fit of the W boson's transverse mass [3], a shift in Γ_{W} might be the actual tension invoking reason which can be revealed from a two parameter fit in $(m_{\mathrm{W}}, \Gamma_{\mathrm{W}})$ parameter space. Despite all attempts to explain this from the perspectives of SM physics, it is needless to say that many were craving for the BSM physics that can resolved the matter.

W **Boson Mass Corrections from Type-II Seesaw Model**

The ρ *parameter* defined as

$$\rho \equiv \frac{m_{\mathrm{W}}^2}{m_Z^2 \cos^2 \theta_{\mathrm{W}}}. \tag{6.3}$$

is a great measure of the SM precision tests. At tree level, $\rho_{\mathrm{tree}} = 1$ is expected whereas the deviation from this provided a indirect hints of yet undiscovered physics known to date. With the corrections from top quark and H boson, correction to ρ was acquired as

$$\Delta\rho \simeq \frac{3G_F}{8\sqrt{2}\pi^2}\left(m_t^2 - m_{\mathrm{Z}}^2 \sin^2\theta_{\mathrm{W}} \ln \frac{m_{\mathrm{H}}^2}{m_{\mathrm{W}}^2}\right). \tag{6.4}$$

Such precision test was able to provide rough estimates on the masses of top quark and H boson from LEP experiment results [137] although it did not have enough energy to produce the particles directly. From the global fit, the measured value up to date is known to be $\rho = 1.00038 \pm 0.00020$ [138] from multiple experiments [4]. In light of this, if the BSM physics proposals predict

[3] Fit from the p_{T} and $\not{E}_{\mathrm{T}}$ was also reported in the same paper which still is consistent with the result from transverse mass parameter fit.

[4] The most recent CDF measurement is not included.

values of ρ parameter that is too far from unity can already be considered excluded.

The deviation of ρ can be formulated in terms of oblique parameters, S, T, and U, introduced by M.E.Peskin and T.Takeuchi in 1990 [139]. When a triplet scalar field $\hat{\Delta}$ is added to the SM, ρ_{tree} is modified by a small correction

$$\rho_{\text{tree}} = 1 + \alpha_{\text{EM}} T_{\text{tree}} = 1 - \frac{2v_\Delta^2}{v^2 + 2v_\Delta^2}. \tag{6.5}$$

This manifests a larger tension making the situations even worse when taking CDF result into account by giving an opposite sign to the correction term T. However, under the assumption demanding $v_\Delta \ll v$, tree level contribution can be ignored and acquire dominant one-loop level contribution that ultimately leads to a shift in the mass of W boson. In terms of the oblique parameters, it is expressed as

$$m_{\text{W}} \simeq m_{\text{W}}^{\text{SM}} \left(1 - \frac{\alpha_{\text{EM}}}{4(1-2s_{\text{W}}^2)} (S - 2(1-s_{\text{W}}^2)T) + \frac{\alpha_{\text{EM}}}{8s_{\text{W}}^2} U \right). \tag{6.6}$$

With U parameter given 0 [5], pull from the CDF result finds the best fit value $(S, T) = (0.17, 0.27)$ with $T \simeq 0.18 + 0.65S$ relation.

This can be further investigated with the type-II seesaw model which explains the neutrino mass generating mechanism through the additional Higgs portal. Additional Higgs bosons (triplet scalars) are Δ^0, $\Delta^\pm$, and $\Delta^{\pm\pm}$ bosons are introduced by the type-II seesaw model. In the limit of small β and masses of additional Higgs bosons at $O(\text{ TeV})$ scale, S, T, and U parameters are given as

$$S \simeq -\frac{(2 - 4s_{\text{W}}^2 + 5s_{\text{W}}^4) m_{\text{Z}}^2}{30\pi M_{\Delta^0}^2} + \beta \frac{v^2}{6\pi M_{\Delta^0}^2}, \tag{6.7}$$

$$T \simeq \frac{v^2 \beta^2}{192\pi^2 \alpha_{\text{EM}} M_{\Delta^0}^2}, \tag{6.8}$$

[5] U parameter contribution is negligible compared to S and T parameters in type-II seesaw model context that will be described below.

$$U \simeq \frac{(2 - 4s_{\mathrm{W}}^2 + 5s_{\mathrm{W}}^4)m_{\mathrm{Z}}^2 - 2m_{\mathrm{W}}^2}{30\pi M_{\Delta^0}^2}. \tag{6.9}$$

It can be noticed that U parameter is suppressed compared to S and T parameters. In addition, as $T \simeq 0.18 + 0.65S$ relation should be satisfied, positive S and T parameters need $\beta > 0$. With the well-known sum rule that type-II seesaw model obeys [75],

$$M_{\Delta^0}^2 - M_{\Delta^\pm}^2 = M_{\Delta^\pm}^2 - M_{\Delta^{\pm\pm}}^2 = \frac{\beta}{4}v^2, \tag{6.10}$$

a nonzero value of $\beta > 0$ and the spectrum of $v_\Delta \ll v$ provides phenomenologically rich environment with different decay modes. Those are leptonic, cascade, and gauge boson decay modes of the additional Higgs bosons which also leads to very different event kinematics even with the same final state objects in LHC experiments. An easy example of this would be $pp \to \Delta^{++}\Delta^{--} \to 2\ell^+2\ell^-$ in leptonic decay mode sector and $pp \to \Delta^{++}\Delta^{--} \to 2\mathrm{W}^+2\mathrm{W}^- \to 2\ell^+2\ell^- + \not{E}_{\mathrm{T}}$ in gauge boson decay mode sector.

Scans Using Contur

At the LHC, several analyses have been performed to search for additional Higgs bosons directly but usually focusing on mass degenerate scenarios. With small v_Δ, leptonic decays are enhanced for charged scalars, $\Delta^{\pm\pm} \to \ell^\pm\ell^\pm$ and $\Delta^\pm \to \ell^\pm\nu$ [140]. Searches for charged scalars in leptonic decay modes are considered golden channels as it exemplifies the problem with possibility of being interpreted in numerous scenarios. ATLAS excludes $M_{\Delta^{\pm\pm}} < 1080$ GeV at 95 % CL with Run 2 data [141]. For larger v_Δ scenarios, bosonic decays are more involved, $\Delta^{\pm\pm} \to \mathrm{W}^{(*)}\mathrm{W}^{(*)}$ and $\Delta^\pm \to \mathrm{W}^{(*)}\mathrm{Z}^{(*)}$. Using fully leptonic and semi-leptonic final states, $M_{\Delta^{\pm\pm}}$ in the range $200 - 350$ GeV are excluded by ATLAS at 95 % CL with Run 2 data [142, 143]. In addition, a recasting of Run

1 results show that $M_{\Delta^{\pm\pm}} < 84$ GeV is excluded [144–146] leaving $84-200$ GeV not yet excluded by any of the experiments directly.

Briefly turning back to the investigations on S and T parameters and values of $\beta > 0$, it can be concluded that CDF's result favors the mass hierarchy $M_{\Delta^{\pm\pm}} < M_{\Delta^{\pm}} < M_{\Delta^{0}}$. This can be further translated into best fit values [147, 148]

$$(M_{\Delta^{\pm\pm}}, M_{\Delta^{\pm}} - M_{\Delta^{\pm\pm}}) \simeq (95.5, 72.5) \text{ GeV}. \tag{6.11}$$

Intriguingly, the masses of charged scalars preferred values by incorporating type-II seesaw model and the CDF measurement lies within the range where it has not been excluded by the results from experiments up to date for large v_Δ scenarios. Thus, by using the recently developed tool package that interfaces MadGraph_aMC@NLO and Contur, an investigation on this favored parameter space is performed.

Benchmark points for the type-II seesaw model parameters scanned with Contur are set as below :

- $M_{\Delta^{\pm\pm}} \in [60, 400]$ GeV : Mass ranges are set to cover the allowed doubly charged scalar mass region $84-200$ GeV and the best fit point 95.5 GeV.

- $\Delta_M \equiv M_{\Delta^{\pm}} - M_{\Delta^{\pm\pm}} \in [35, 155]$ GeV : The best fit point at 72.5 GeV and $\beta > 0$ dependence on S and T parameters (Eq. 6.7, 6.8) is covered. This in turn implies $M_{\Delta^{\pm}} > M_{\Delta^{\pm\pm}}$ mass hierarchy in type-II seesaw model.

- $pp \to \Delta^{\pm\pm}\Delta^{\mp\mp}$ and $pp \to \Delta^{\pm\pm}\Delta^{\mp} \to \Delta^{\pm\pm}\Delta^{\mp\mp}\mathrm{W}^{(*)}$ channels : Processes accompanied by $\Delta^{\pm\pm}$ is only considered. From the assumption above, $M_{\Delta^{\pm}} > M_{\Delta^{\pm\pm}}$, cascade decay of singly charged scalar ultimately yields four doubly charged scalars.

- $\Delta^{\pm\pm} \to \mathrm{W}^{\pm(*)}\mathrm{W}^{\pm(*)}$ decays ($v_\Delta = 1$ GeV) : Larger v_Δ reduces the

Yukawa couplings to leptons and inversely enhances the bosonic decays of charged scalars. Smaller v_Δ values lead to smaller widths and hence longer life time of the charged scalars. Long-lived signatures results in ambiguity of interpreting from the prompt signatures in RIVET, $v_\Delta = 1$ GeV is set to ensure the promptness.

- Other free parameters : The CP and Majorana phases are set to 0 and the mass of lightest SM neutrino is set to 0.01 eV. In addition, normal hierarchy of the SM neutrino masses assumed.

The steering commands in MADGRAPH_AMC@NLO is below :

```
define lepton = e+ e- mu+ mu- ta+ ta- ve vm vt
define quark = u d c s u~ d~ c~ s~
define fermion = lepton quark
generate p p > d++ d--
add process p p > d++ d-, d- > d-- fermion fermion
add process p p > d+ d--, d+ > d++ fermion fermion
```

As can be noticed from the commands above, $\Delta^{\pm\pm}$ is not decayed using MADEVENT promptly. MADSPIN is interfaced to decay $\Delta^{\pm\pm}$, which is a module that generate resonance decay events separately and later appended to the events produced from MADEVENT with proper Lorentz boosts under narrow-width approximation. This approach is in particular helpful for reducing the computing time. Below is the commands to launch MADSPIN for $\Delta^{\pm\pm}$ decays :

```
set spinmode none
decay d++ > w+ > fermion fermion fermion fermion
decay d-- > w- > fermion fermion fermion fermion
```

Here, the `spinmode none` setting in MADSPIN does not consider the spin correlations when computing the decay kinematics although as the particle that goes through the decay with this module is spinless particless makes it still safe to be used. Also, to account for symmetry factor computations that cannot

be validly computed while utilizing MadSpin, the decay widths of $\Delta^{\pm\pm}$, the widths are computed externally using the command : `generate d++ > w+ > fermion fermion fermion fermion`.

The Contur scan results are shown in Fig. 6.7 overlayed on the $M_{\Delta^{\pm\pm}} - \Delta_M$ parameter space. As discussed above, doubly charged scalars in mass ranges $84 - 200$ GeV are not yet excluded from direct searches by ATLAS or CMS and the best fit value which is indicated with a black asterisk (*) lie at the point $(95.5, 72.5)$ GeV. Also shown in the figure is the 95% expected exclusion (dotted), 95% observed exclusion (solid), and 68% observed exclusion (dashed) limits with the CL_s method. The left side of the lines are the parameter space which is excluded from Contur.

Most of the exclusion especially in the high $M_{\Delta^{\pm\pm}}$ region is driven by the four-lepton cross section measurement using 139 fb^{-1} data at $\sqrt{s} = 13$ TeV collected with ATLAS [149]. Other contributions for relatively smaller mass phase spaces are from the search for triboson WWW production at 8 TeV [150] and WW + jet measurement at 13 TeV [151] which corresponds to 20 fb^{-1} and 139 fb^{-1} of data, respectively, analyzed by ATLAS.

Several interesting phenomenological effects can be addressed from both observed and expected exclusions. First looking at the region where both $m_{\Delta^{\pm\pm}}$ and Δ_M are small, the contributions from $\Delta^{\pm\pm}\Delta^{\mp\mp}$ pair production process and $\Delta^{\pm\pm}\Delta^{\mp}$ associated production process are expected to give roughly comparable contributions. This follows from the the study in [75] that tells $\Delta^{\pm\pm}\Delta^{\mp}$ associated production process have about $O(2\times)$ larger cross section than that of $\Delta^{\pm\pm}\Delta^{\mp\mp}$ pair production process. This means that the two production channels of interest can give effectively same cross sections in this region. From the fluctuations of observed limits for $\Delta_M \leq 100$ GeV it can be inferred that two processes interplay on sensitivity.

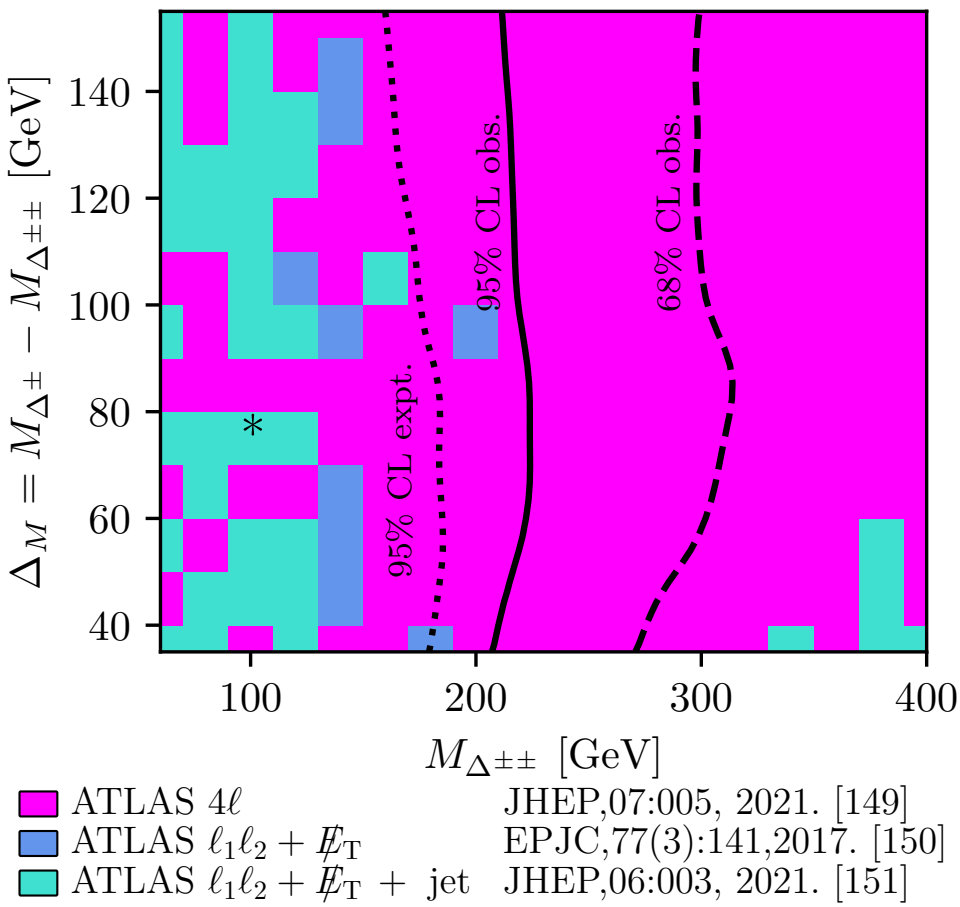

Figure 6.7: Exclusion limits on $M_{\Delta^{\pm\pm}} - \Delta_M$ parameter space.

When the mass splitting is increased, there occurs a competition in the associated production process : (1) The overall momentum transfer scale of the event increases and this limits the available phase space, ultimately yielding smaller cross sections. (2) The branching ratio of $\Delta^{\pm} \to \Delta^{\pm\pm} f \bar{f}'$ becomes larger due to more onshell W boson involvement which is $\Delta^{\pm} \to \Delta^{\pm\pm} \mathrm{W}^{\mp}$. The first factor eventually makes associated production process a negligible contribution to the signals. As shown in Fig. 6.7, observed limits for $\Delta_M \geq 100$ GeV starts to from a straight line because there only pair production channel practically becomes useful for exclusions.

Finally, it can be found that the observed limits give more constraints on $M_{\Delta^{\pm\pm}}$ by $30-50$ GeV than the expected limits. This can be accounted from two factors : (1) The main sensitivity driving 4 lepton cross section measurement analysis is statistically limited. (2) The SM prediction from Sherpa MC event generator [99] already lies above the measured data in the most signal sensitive bin of the measurement. The expected limits are set with the assumption that background yields exactly match with the MC predictions but also assigned with the uncertainty from the measurement. This measurement is mainly limited from lack of statistics which hinders the sensitivity.

The distributions showing the SM prediction and measured data is presented in Fig. 6.8 as an example. It is the distribution of the highest-mass dilepton pair in 4 lepton events after discarding the events that does not pass $m(\ell\ell\ell\ell) > 2 \times m_{\mathrm{Z}}$. The distributions are acquired through the Rivet implementations of Ref. [149]. Green line shows the SM prediction and the black marker shows the measured data. As observant from the figure, SM prediction already overpredicts the measured data which makes the signal prediction added values in blue line even deviate more. Signal prediction is given with $(M_{\Delta^{\pm\pm}}, M_{\Delta^{\pm}}) = (180, 255)$ GeV assumption as a representative example mass

point.

6.3 Results

It is astonishing to find that type-II seesaw model which has been popular neutrino mass model can also be an explanation for the difference between the W boson mass measurement from CDF and the SM expectation. Recent studies Refs. [147, 152, 153] point out that triplet scalars with $O(100)$ GeV masses can accommodate the deviation and this is evidently well within the energy reach of LHC. However due to its complex decay structure, model's full potential couldn't be fully covered by experimental searches. This left some uncovered window in masses of triplet scalars which happen to include the best fit value of oblique parameters.

From this study, we were able to fully exploit the complex decay structure of triplet scalars, cascade decays $\Delta^{\pm} \to \Delta^{\pm\pm}\mathrm{W}^{\mp(*)}$ and bosonic decays $\Delta^{\pm\pm} \to \mathrm{W}^{\pm}\mathrm{W}^{\pm}$. This is because the newly developed MadGraph_aMCatNLO-CONTUR interface can take the generated signal MC events up to particle level without detector simulations for reinterpretation purposes which is much faster approach than going through detector simulations in experiments and some other reinterpretation tools. As a result, we found that by assuming prompt decays of $\Delta^{\pm\pm}$ excludes the best fit point and the 1σ region in $m_{\Delta^{\pm\pm}} - \Delta_M$ parameter space can be fully excluded by LHC data at 95 % CL by LHC. Most of the exclusions are driven by the four lepton measurement by ATLAS which was designed to test the precision level of all physics processes that contributes to four lepton final state (e.g. single Z boson, $\mathrm{ZZ}^{(*)}$ diboson including H boson mediation, and other complex mix of interference terms) occurring at the LHC.

As for the last remark, this work shows the powerfulness of such reinterpretation methodology as it can timely address arising questions for BSM physics

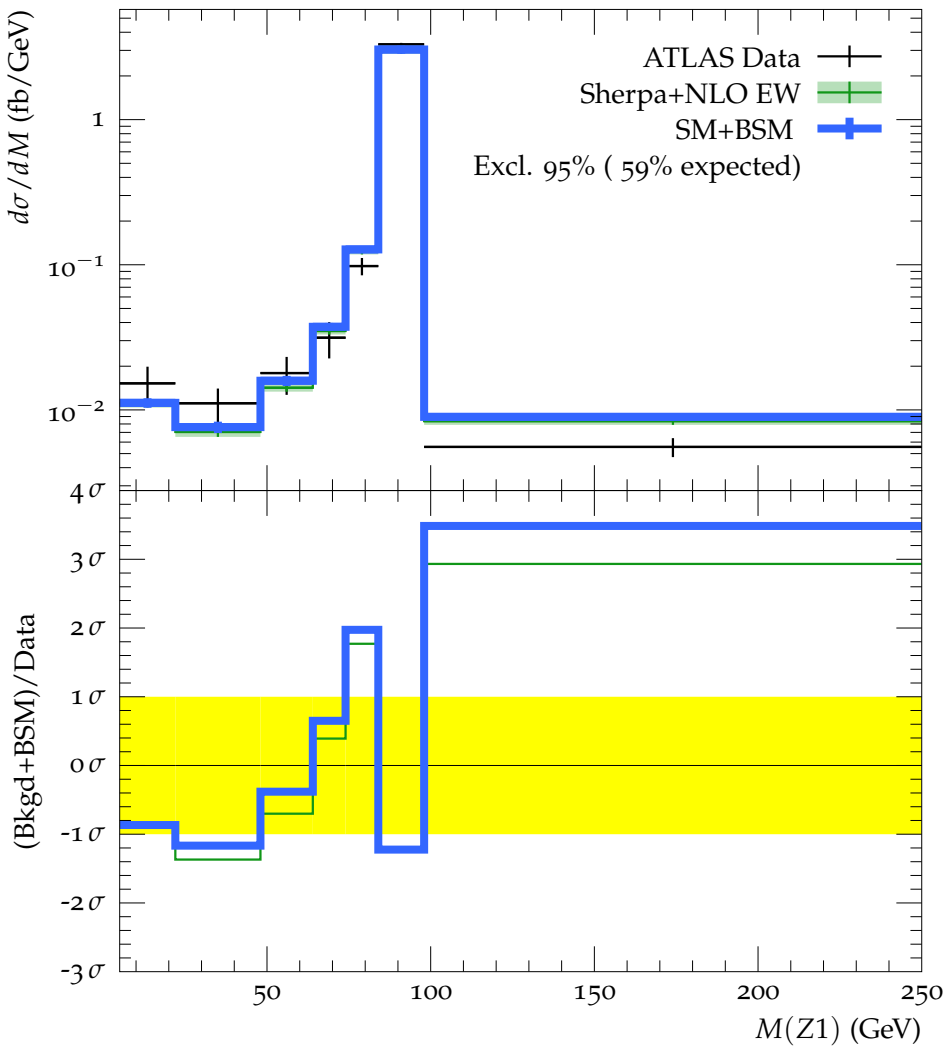

Figure 6.8: Upper panel shows the representative differential cross section of highest mass dilepton pair distribution from four lepton measurement. Lower panel shows the bin-by-bin significance of expected theory yields relative to data with combined uncertainty.

proposals. In addition, it provides the chance to probe BSM physics with multitude of parameters that cannot be fully covered by experiments. However, this at the same time demands experiments to execute more precise measurements and ultimately preserve them to build up the database in order to be used for reinterpretations. This also shows the direction of physics programs during the HL-LHC era which will benefit from statistics providing much higher level of precision.

Chapter 7

Conclusions

7.1 Summary

In this thesis, we tested two neutrino mass models which realizes the Weinberg operator at the tree level with fermion singlet or scalar triplet. The first model is referred to as *type-I seesaw model*, introducing a BSM particle that is the *right-handed Majorana heavy neutrino N* while the second model, *type-II seesaw model* introduces scalar triplet $\Delta = (\Delta^{\pm\pm}, \Delta^{\pm}, \Delta^{0})$.

The first analysis is performed with a popular approach that is *theory-driven*, optimizing the event selections to extract the signal processes as much as possible while keeping the background processes suppressed. The signal processes, under the type-I seesaw model context, consist of a light lepton, neutrino, and a boson that decay into two quarks in particle-level. The key to this analysis was utilizing the boosted signatures with large-radius jets that is used to reconstruct the boson as a whole with collimated signatures. For complementary purposes, resolved signatures where the boson is not so much boosted and

end up yielding two small-radius jets in the event is also explored. We use the full Run 2 dataset collected by the CMS detector at the Large Hadron Collider that corresponds to an integrated luminosity of 137.4 fb^{-1} From this analysis, heavy neutrino masses above 500 GeV, up to 940 and 906 GeV from muon and electron channel analyses, respectively. This analysis was the first attempt to search for heavy neutrinos in single lepton final state at the LHC and the very first in the world to search for heavy neutrinos using its H boson decay mode.

The second part of this thesis is in an opposite direction that is *data-driven* which is also known as *reinterpretation*. Instead of optimizing the analysis for a particular BSM theory of interest, we exploit the full potential of the LHC using the results that have already been published that are mostly measurements of the data corrected with detector effects. To demonstrate this, we use CONTUR tool interfaced with MADGRAPH_AMC@NLO MC event generator that runs over the analyses stored in RIVET and take the event yields from HEPDATA for extracting sensitivity. One important feature of CONTUR is that it relies on the measurement results that are stored in RIVET, reversing the data to particle-level which can be directly compared with MC event generator outputs without detector simulation. With CONTUR, we investigate one of the timely interesting W boson mass anomaly reported by CDF which can be explained through type-II seesaw model. The parameter points preferred by the CDF report in the $M_{\Delta^{\pm\pm}} - (M_{\Delta^{\pm}} - M_{\Delta^{\pm\pm}})$ parameter space including its $\pm 1\ \sigma$ bands have been fully excluded.

7.2 Outlook

For the type-I seesaw model analysis, it would be interesting to seek a possibility of performing a charge dependent analysis. The human knowledge up-to-date on PDFs is not so much complete for the high momentum fraction carrying

partons. However if it is complete, such analysis that is dominated by charge neutral processes (e.g. top pair process which gives exactly same amount of positively and negatively charged leptons) can be highly beneficial as the signal processes are not charge neutral (e.g. W^{+*} and W^{-*} depend a lot on PDF as the LHC is a proton-proton collider machine).

Reinterpretation of type-I seesaw model has been briefly discussed in Chapter. 6.2.1 as well. One obstacle was found that heavy neutrinos mixing with taus had relatively poor exclusions in the $m_N - |V_{N\ell}|^2$ plane compared to other lepton flavors. This is because compared to other lepton flavors, measurements with taus are either less performed or less preserved at the LHC. Although it lacks accuracy as being one of the most difficult object to reconstruct, taus can bring the most fruitful physics results as being the most massive lepton which directly connect to the Higgs physics.

As mentioned above, this thesis tests two popular and famous neutrino mass models using two different analysis approaches. It is true that theory-driven approach can be more optimized to particular signal processes and thus has better performance in terms of its result. However, data-driven approach has its own unique strength that is takes much lesser time and smaller effort with fully exploiting the potential of the LHC making use of all the published results. As we recently started the Run 3 at the LHC and also approaching the era of HL-LHC pushing the LHC and its detectors around the ring to its high-end performance, setting the physics programs would be one of the most important tasks. In light of this thesis, we can set two major goals during HL-LHC for hunting BSM physics : (1) Searching for BSM scenarios with more extreme phase spaces. (2) Measuring SM processes with much better statistics and in higher level of precision. Ultimately, the efforts for preservation of results, either in exotic phase space or conventional reaches of the SM, will lead to more

prosperous particle physics.

Bibliography

[1] E. D. Bloom et al. "High-Energy Inelastic $e - p$ Scattering at 6° and 10°". In: *Phys. Rev. Lett.* 23 (16 Oct. 1969), pp. 930–934. DOI: 10.1103/PhysRevLett.23.930. URL: https://link.aps.org/doi/10.1103/PhysRevLett.23.930.

[2] M. Breidenbach et al. "Observed Behavior of Highly Inelastic Electron-Proton Scattering". In: *Phys. Rev. Lett.* 23 (16 Oct. 1969), pp. 935–939. DOI: 10.1103/PhysRevLett.23.935. URL: https://link.aps.org/doi/10.1103/PhysRevLett.23.935.

[3] J. -E. Augustin et al. "Discovery of a Narrow Resonance in e^+e^- Annihilation". In: *Phys. Rev. Lett.* 33 (23 Dec. 1974), pp. 1406–1408. DOI: 10.1103/PhysRevLett.33.1406. URL: https://link.aps.org/doi/10.1103/PhysRevLett.33.1406.

[4] J. J. Aubert et al. "Experimental Observation of a Heavy Particle J". In: *Phys. Rev. Lett.* 33 (23 Dec. 1974), pp. 1404–1406. DOI: 10.1103/PhysRevLett.33.1404. URL: https://link.aps.org/doi/10.1103/PhysRevLett.33.1404.

[5] S. W. Herb et al. "Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions". In: *Phys. Rev. Lett.* 39 (5 Aug. 1977), pp. 252–255. DOI: 10.1103/PhysRevLett.39.252. URL: https://link.aps.org/doi/10.1103/PhysRevLett.39.252.

[6] F. Abe et al. "Observation of Top Quark Production in $\overline{p}p$ Collisions with the Collider Detector at Fermilab". In: *Phys. Rev. Lett.* 74 (14 Apr. 1995), pp. 2626–2631. DOI: 10.1103/PhysRevLett.74.2626. URL: https://link.aps.org/doi/10.1103/PhysRevLett.74.2626.

[7] J. J. Thomson M.A. F.R.S. "XL. Cathode Rays". In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 44.269 (1897), pp. 293–316. DOI: 10.1080/14786449708621070. eprint: https://doi.org/10.1080/14786449708621070. URL: https://doi.org/10.1080/14786449708621070.

[8] G. Danby et al. "Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos". In: *Phys. Rev. Lett.* 9 (1 July 1962), pp. 36–44. DOI: 10.1103/PhysRevLett.9.36. URL: https://link.aps.org/doi/10.1103/PhysRevLett.9.36.

[9] M. L. Perl et al. "Evidence for anomalous lepton production in e^+e^- annihilation". In: *Phys. Rev. Lett.* 35.22 (Dec. 1975), pp. 1489–1492. DOI: 10.1103/PhysRevLett.35.1489.

[10] DONUT Collaboration et al. "Observation of tau neutrino interactions". In: *Physics Letters B* 504.3 (Apr. 2001), pp. 218–224. DOI: 10.1016/S0370-2693(01)00307-0. arXiv: hep-ex/0012035 [hep-ex].

[11] G. Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Physics Letters B* 716.1 (2012), pp. 1–29. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2012.08.020. URL: http://www.sciencedirect.com/science/article/pii/S037026931200857X.

[12] S. Chatrchyan et al. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Physics Letters B* 716.1 (2012), pp. 30–61. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2012.08.021. URL: http://www.sciencedirect.com/science/article/pii/S0370269312008581.

[13] Raymond Davis, Don S. Harmer, and Kenneth C. Hoffman. "Search for Neutrinos from the Sun". In: *Phys. Rev. Lett.* 20 (21 May 1968), pp. 1205–1209. DOI: 10.1103/PhysRevLett.20.1205. URL: https://link.aps.org/doi/10.1103/PhysRevLett.20.1205.

[14] K. S. Hirata et al. "Real-time, directional measurement of ^{8}B solar neutrinos in the Kamiokande II detector". In: *Phys. Rev. D* 44 (8 Oct. 1991), pp. 2241–2260. DOI: 10.1103/PhysRevD.44.2241. URL: https://link.aps.org/doi/10.1103/PhysRevD.44.2241.

[15] Y. Fukuda et al. "Solar Neutrino Data Covering Solar Cycle 22". In: *Phys. Rev. Lett.* 77 (9 Aug. 1996), pp. 1683–1686. DOI: 10.1103/PhysRevLett.77.1683. URL: https://link.aps.org/doi/10.1103/PhysRevLett.77.1683.

[16] Y. Fukuda et al. "Measurements of the Solar Neutrino Flux from Super-Kamiokande's First 300 Days". In: *Phys. Rev. Lett.* 81 (6 Aug. 1998), pp. 1158–1162. DOI: 10.1103/PhysRevLett.81.1158. URL: https://link.aps.org/doi/10.1103/PhysRevLett.81.1158.

[17] Y. Fukuda et al. "Measurement of the Solar Neutrino Energy Spectrum Using Neutrino-Electron Scattering". In: *Phys. Rev. Lett.* 82 (12 Mar. 1999), pp. 2430–2434. DOI: 10.1103/PhysRevLett.82.2430. URL: https://link.aps.org/doi/10.1103/PhysRevLett.82.2430.

[18] M. Altmann et al. "Complete results for five years of GNO solar neutrino observations". In: *Physics Letters B* 616.3 (2005), pp. 174–190. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2005.04.068. URL: http://www.sciencedirect.com/science/article/pii/S0370269305005149.

[19] J. N. Abdurashitov et al. "Measurement of the solar neutrino capture rate with gallium metal. III. Results for the 2002–2007 data-taking period". In: *Phys. Rev. C* 80 (1 July 2009), p. 015807. DOI: 10.1103/PhysRevC.80.015807. URL: https://link.aps.org/doi/10.1103/PhysRevC.80.015807.

[20] Q. R. Ahmad et al. "Measurement of the Rate of $\nu_e + d \to p + p + e^-$ Interactions Produced by ^8B Solar Neutrinos at the Sudbury Neutrino Observatory". In: *Phys. Rev. Lett.* 87 (7 July 2001), p. 071301. DOI: 10.1103/PhysRevLett.87.071301. URL: https://link.aps.org/doi/10.1103/PhysRevLett.87.071301.

[21] Y. Fukuda et al. "Evidence for Oscillation of Atmospheric Neutrinos". In: *Phys. Rev. Lett.* 81 (8 Aug. 1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562. URL: https://link.aps.org/doi/10.1103/PhysRevLett.81.1562.

[22] Y. Abe et al. "Indication of Reactor $\overline{\nu}_e$ Disappearance in the Double Chooz Experiment". In: *Phys. Rev. Lett.* 108 (13 Mar. 2012), p. 131801. DOI: 10.1103/PhysRevLett.108.131801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.108.131801.

[23] F. P. An et al. "Observation of Electron-Antineutrino Disappearance at Daya Bay". In: *Phys. Rev. Lett.* 108 (17 Apr. 2012), p. 171803. DOI: 10.1103/PhysRevLett.108.171803. URL: https://link.aps.org/doi/10.1103/PhysRevLett.108.171803.

[24] Gerard 't Hooft. "Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking". In: *NATO Sci. Ser. B* 59 (1980). Ed. by Gerard 't Hooft et al., pp. 135–157. DOI: 10.1007/978-1-4684-7571-5_9.

[25] Sabine Hossenfelder. "Screams for explanation: finetuning and naturalness in the foundations of physics". In: *Synthese* 198.Suppl 16 (2021), pp. 3727–3745. DOI: 10.1007/s11229-019-02377-5. arXiv: 1801.02176 [physics.hist-ph].

[26] B. Abi et al. "Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm". In: *Phys. Rev. Lett.* 126.14 (2021), p. 141801. DOI: 10.1103/PhysRevLett.126.141801. arXiv: 2104.03281 [hep-ex].

[27] Roel Aaij et al. "Measurement of the ratio of branching fractions $\mathcal{B}(\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \to D^{*+}\mu^-\bar{\nu}_\mu)$". In: *Phys. Rev. Lett.* 115.11 (2015). [Erratum: Phys.Rev.Lett. 115, 159901 (2015)], p. 111803. DOI: 10.1103/PhysRevLett.115.111803. arXiv: 1506.08614 [hep-ex].

[28] R. Aaij et al. "Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays". In: *JHEP* 08 (2017), p. 055. DOI: 10.1007/JHEP08(2017)055. arXiv: 1705.05802 [hep-ex].

[29] John C. Collins, Davison E. Soper, and George F. Sterman. "Factorization of Hard Processes in QCD". In: *Adv. Ser. Direct. High Energy Phys.* 5 (1989), pp. 1–91. DOI: 10.1142/9789814503266_0001. arXiv: hep-ph/0409313.

[30] Tilman Plehn. *LHC Phenomenology for Physics Hunters*. 2008. eprint: arXiv:0810.2281.

[31] Davison E. Soper. "Parton distribution functions". In: *Nucl. Phys. B Proc. Suppl.* 53 (1997). Ed. by C. Bernard et al., pp. 69–80. DOI: 10.1016/S0920-5632(96)00600-7. arXiv: hep-lat/9609018.

[32] Richard D. Ball et al. "Parton distributions from high-precision collider data". In: *The European Physical Journal C* 77.10 (Oct. 2017). DOI: 10.1140/epjc/s10052-017-5199-5. URL: https://doi.org/10.1140%5C%2Fepjc%5C%2Fs10052-017-5199-5.

[33] Aneesh Manohar et al. "How bright is the proton? A precise determination of the photon parton distribution function". In: *Phys. Rev. Lett.* 117.24 (2016), p. 242002. DOI: 10.1103/PhysRevLett.117.242002. arXiv: 1607.04266 [hep-ph].

[34] Mark Thomson. *Modern particle physics*. New York: Cambridge University Press, 2013. ISBN: 978-1-107-03426-6. DOI: 10.1017/CBO9781139525367.

[35] Luigi Del Debbio. "Parton distributions in the LHC era". In: *EPJ Web of Conferences* 175 (2018), p. 01006. ISSN: 2100-014X. DOI: 10.1051/epjconf/201817501006. (Visited on 01/04/2023).

[36] Zoltán Nagy and Davison E. Soper. "What is a parton shower?" In: *Physical Review D* 98.1 (July 2018). DOI: 10.1103/physrevd.98.014034. URL: https://doi.org/10.1103%5C%2Fphysrevd.98.014034.

[37] Stephen R. Sharpe. *Phenomenology from the Lattice*. 1994. DOI: 10.48550/ARXIV.HEP-PH/9412243. URL: https://arxiv.org/abs/hep-ph/9412243.

[38] Nils Lavesson and Leif Lonnblad. "Merging parton showers and matrix elements: Back to basics". In: *JHEP* 04 (2008), p. 085. DOI: 10.1088/1126-6708/2008/04/085. arXiv: 0712.2966 [hep-ph].

[39] S. Agostinelli et al. "GEANT4–a simulation toolkit". In: *Nucl. Instrum. Meth. A* 506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[40] Michael Benedikt et al. *LHC Design Report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. DOI: 10.5170/CERN-2004-003-V-3. URL: https://cds.cern.ch/record/823808.

[41] Lyndon R Evans and Philip Bryant. "LHC Machine". In: *JINST* 3 (2008). This report is an abridged version of the LHC Design Report (CERN-2004-003), S08001. 164 p. DOI: 10.1088/1748-0221/3/08/S08001. URL: https://cds.cern.ch/record/1129806.

[42] Carlo Wyss. *LEP design report, v.3: LEP2*. Vol. 1-2 publ. in 1983-84. Geneva: CERN, 1996. URL: https://cds.cern.ch/record/314187.

[43] G. Apollinari et al. "High Luminosity Large Hadron Collider HL-LHC". In: *CERN Yellow Rep.* 5 (2015). Ed. by G Apollinari et al., pp. 1–19. DOI: 10.5170/CERN-2015-005.1. arXiv: 1705.08830 [physics.acc-ph].

[44] G L Bayatian et al. *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*. Technical Design Report CMS. There is an error on cover due to a technical problem for some items. Geneva: CERN, 2006. URL: https://cds.cern.ch/record/922757.

[45] A Dominguez et al. *CMS Technical Design Report for the Pixel Detector Upgrade*. Tech. rep. CERN-LHCC-2012-016. CMS-TDR-11. Sept. 2012. URL: https://cds.cern.ch/record/1481838.

[46] "The construction of the CMS ECAL: avalanche photodiodes (APDs). La construction du calorimetre electromagnetique de CMS: APDs". CMS Collection. Oct. 2005. URL: http://cds.cern.ch/record/929105.

[47] R.M. Brown. "Avalanche photodiodes and vacuum phototriodes for the CMS electromagnetic calorimeter". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 572.1 (2007). Frontier Detectors for Frontier Physics, pp. 26–28. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2006.10.160. URL: http://www.sciencedirect.com/science/article/pii/S0168900206019693.

[48] R.M. Brown. "The CMS electromagnetic Calorimeter". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 572.1 (2007). Frontier Detectors for Frontier Physics, pp. 29–32. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2006.10.240. URL: https://www.sciencedirect.com/science/article/pii/S016890020601970X.

[49] "The CMS Collaboration". In: *Journal of Physics G: Nuclear and Particle Physics* 35.10 (Oct. 2008), p. 109805. ISSN: 0954-3899, 1361-6471. DOI: 10.1088/0954-3899/35/10/109805. URL: https://iopscience.iop.org/article/10.1088/0954-3899/35/10/109805 (visited on 01/04/2023).

[50] CMS Collaboration. "Calibration of the CMS hadron calorimeters using proton-proton collision data at $\sqrt{s} = 13\,\mathrm{TeV}$". In: *Journal of Instrumentation* 15.05 (May 2020), P05002. DOI: 10.1088/1748-0221/15/05/P05002. URL: https://dx.doi.org/10.1088/1748-0221/15/05/P05002.

[51] Julia Bauer and Thomas Muller. "Prospects for the Observation of Electroweak Top Quark Production with the CMS Experiment". In: (Jan. 2023).

[52] S Chatrchyan et al. "Performance of the CMS Drift Tube Chambers with Cosmic Rays. Performance of the CMS Drift Tube Chambers with Cosmic Rays". In: *JINST* 5.arXiv:0911.4855. CMS-CFT-09-012 (Nov. 2009), T03015 . 45 p. DOI: 10.1088/1748-0221/5/03/T03015. URL: https://cds.cern.ch/record/1223944.

[53] CMS Collaboration. "Performance of the CMS cathode strip chambers with cosmic rays". In: *Journal of Instrumentation* 5.03 (Mar. 2010), T03018–T03018. DOI: 10.1088/1748-0221/5/03/t03018. URL: https://doi.org/10.1088%5C%2F1748-0221%5C%2F5%5C%2F03%5C%2Ft03018.

[54] Seong-Kwan Park et al. "CMS endcap RPC gas gap production for upgrade". In: *Journal of Instrumentation* 7 (Nov. 2012), P11013. DOI: 10.1088/1748-0221/7/11/P11013.

[55] A.M. Sirunyan et al. "Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Journal of Instrumentation* 15.10 (Oct. 2020), P10017–P10017. DOI: 10.1088/1748-0221/15/10/p10017. URL: https://doi.org/10.1088%2F1748-0221%2F15%2F10%2Fp10017.

[56] Ekaterina Govorkova et al. "LHC physics dataset for unsupervised New Physics detection at 40 MHz". In: *Sci. Data* 9 (2022), p. 118. DOI: 10.1038/s41597-022-01187-8. arXiv: 2107.02157 [physics.data-an].

[57] A. M. Sirunyan et al. "Particle-flow reconstruction and global event description with the CMS detector". In: *JINST* 12 (2017), P10003. DOI: 10.1088/1748-0221/12/10/P10003. arXiv: 1706.04965 [physics.ins-det].

[58] Wolfgang Adam et al. *Track Reconstruction in the CMS tracker*. Tech. rep. CMS-NOTE-2006-041. Geneva: CERN, Dec. 2006. URL: https://cds.cern.ch/record/934067.

[59] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "FastJet User Manual". In: *Eur. Phys. J. C* 72 (2012), p. 1896. DOI: 10.1140/epjc/s10052-012-1896-2. arXiv: 1111.6097 [hep-ph].

[60] Ryan Atkin. "Review of jet reconstruction algorithms". In: *Journal of Physics: Conference Series* 645 (Oct. 2015), p. 012008. ISSN: 1742-6588, 1742-6596. DOI: 10.1088/1742-6596/645/1/012008. URL: https://iopscience.iop.org/article/10.1088/1742-6596/645/1/012008 (visited on 01/04/2023).

[61] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. "The anti-kt jet clustering algorithm". In: *Journal of High Energy Physics* 2008.04 (Apr. 2008), pp. 063–063. DOI: 10.1088/1126-6708/2008/04/063. URL: https://doi.org/10.1088%5C%2F1126-6708%5C%2F2008%5C%2F04%5C%2F063.

[62] 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". In: *JINST* 13.10 (2018), P10005. DOI: 10.1088/1748-0221/13/10/P10005. arXiv: 1809.02816 [hep-ex].

[63] Francesco Guerra, Matteo Leone, and Nadia Robotti. "When Energy Conservation Seems to Fail: The Prediction of the Neutrino". en. In: *Science & Education* 23.6 (June 2014), pp. 1339–1359. ISSN: 0926-7220, 1573-1901. DOI: 10.1007/s11191-012-9567-0. URL: http://link.springer.com/10.1007/s11191-012-9567-0 (visited on 01/04/2023).

[64] C. D. Ellis and W. A. Wooster. "The β-ray Type of Disintegration". In: *Mathematical Proceedings of the Cambridge Philosophical Society* 22.6 (1925), pp. 849–860. DOI: 10.1017/S030500410001447X.

[65] "Atomic Stability and Conservation Laws". In: *Nuclear Physics (1929–1952)*. Ed. by Rudolf Peierls. Vol. 9. Niels Bohr Collected Works. Elsevier, 1986, pp. 99–114. DOI: https://doi.org/10.1016/S1876-0503(08)70456-0. URL: https://www.sciencedirect.com/science/article/pii/S1876050308704560.

[66] Wolfgang Pauli. "Pauli letter collection: letter to Lise Meitner". Typed copy. URL: https://cds.cern.ch/record/83282.

[67] C. L. Cowan et al. "Detection of the free neutrino: A Confirmation". In: *Science* 124 (1956), pp. 103–104. DOI: 10.1126/science.124.3212.103.

[68] Raymond Davis Jr. "Attempt to detect the antineutrinos from a nuclear reactor by the Cl37(anti-nu,e-) A37 reaction". In: *Phys. Rev.* 97 (1955), pp. 766–769. DOI: 10.1103/PhysRev.97.766.

[69] Raymond Davis. "Solar Neutrinos. II. Experimental". In: *Phys. Rev. Lett.* 12 (11 Mar. 1964), pp. 303–305. DOI: 10.1103/PhysRevLett.12.303. URL: https://link.aps.org/doi/10.1103/PhysRevLett.12.303.

[70] B. T. Cleveland et al. "Measurement of the solar electron neutrino flux with the Homestake chlorine detector". In: *Astrophys. J.* 496 (1998), pp. 505–526. DOI: 10.1086/305343.

[71] Salvador Centelles Chuliá, Rahul Srivastava, and José W.F. Valle. "Seesaw roadmap to neutrino mass and dark matter". In: *Physics Letters B* 781 (June 2018), pp. 122–128. DOI: 10.1016/j.physletb.2018.03.046. URL: https://doi.org/10.1016%5C%2Fj.physletb.2018.03.046.

[72] Yi Cai et al. "Lepton Number Violation: Seesaw Models and Their Collider Tests". In: *Frontiers in Physics* 6 (May 2018), p. 40. ISSN: 2296-424X. DOI: 10.3389/fphy.2018.00040. URL: http://journal.frontiersin.org/article/10.3389/fphy.2018.00040/full (visited on 01/04/2023).

[73] Céline Degrande et al. "Fully automated precision predictions for heavy neutrino production mechanisms at hadron colliders". In: *Phys. Rev. D* 94 (2016), p. 053002. DOI: 10.1103/PhysRevD.94.053002. arXiv: 1602.06957 [hep-ph].

[74] Daniel Alva, Tao Han, and Richard Ruiz. "Heavy Majorana neutrinos from Wγ fusion at hadron colliders". In: *JHEP* 02 (2015), p. 072. ISSN: 1029-8479. DOI: 10.1007/JHEP02(2015)072. arXiv: 1411.7305 [hep-ph].

[75] Benjamin Fuks, Miha Nemevšek, and Richard Ruiz. "Doubly Charged Higgs Boson Production at Hadron Colliders". In: *Phys. Rev. D* 101.7 (2020), p. 075022. DOI: 10.1103/PhysRevD.101.075022. arXiv: 1912.08975 [hep-ph].

[76] Olivier Mattelaer, Manimala Mitra, and Richard Ruiz. "Automated Neutrino Jet and Top Jet Predictions at Next-to-Leading-Order with Parton Shower Matching in Effective Left-Right Symmetric Models". In: (Oct. 2016). arXiv: 1610.08985 [hep-ph].

[77] Vardan Khachatryan et al. "Search for heavy Majorana neutrinos in $\mu^\pm\mu^\pm$ + jets events in proton–proton collisions at $\sqrt{s}$ = 8 TeV". In: *Phys. Lett. B* 748 (2015), pp. 144–166. DOI: 10.1016/j.physletb.2015.06.070. arXiv: 1501.05566 [hep-ex].

[78] Vardan Khachatryan et al. "Search for heavy Majorana neutrinos in $e^\pm e^\pm$ + jets and $e^\pm\mu^\pm$ + jets events in proton-proton collisions at $\sqrt{s}$ = 8 TeV". In: *JHEP* 04 (2016), p. 169. DOI: 10.1007/JHEP04(2016)169. arXiv: 1603.02248 [hep-ex].

[79] Vardan Khachatryan et al. "Search for heavy neutrinos and W bosons with right-handed coupling in proton-proton collisions at $\sqrt{s}$ = 8 TeV". In: *Eur. Phys. J. C* 74.11 (2014), p. 3149. DOI: 10.1140/epjc/s10052-014-3149-z. arXiv: 1407.3683 [hep-ex].

[80] Georges Aad et al. "Search for heavy Majorana neutrinos with the ATLAS detector in pp collisions at $\sqrt{s}$ = 8 TeV". In: *JHEP* 07 (2015), p. 162. DOI: 10.1007/JHEP07(2015)162. arXiv: 1506.06020 [hep-ex].

[81] A. M. Sirunyan et al. "Search for a heavy composite Majorana neutrino in the final state with two leptons and two quarks at $\sqrt{s}$ = 13 TeV". In: *Phys. Lett. B* 775 (2017), pp. 315–337. DOI: 10.1016/j.physletb.2017.11.001. arXiv: 1706.08578 [hep-ex].

[82] Albert M Sirunyan et al. "Search for heavy Majorana neutrinos in same-sign dilepton channels in proton-proton collisions at $\sqrt{s}$ = 13 TeV". In: *JHEP* 01 (2019), p. 122. DOI: 10.1007/JHEP01(2019)122. arXiv: 1806.10905 [hep-ex].

[83] Albert M Sirunyan et al. "Search for a heavy right-handed W boson and a heavy neutrino in events with two same-flavor leptons and two jets at $\sqrt{s} = 13$ TeV". In: *JHEP* 05 (2018), p. 148. DOI: 10.1007/JHEP05(2018)148. arXiv: 1803.11116 [hep-ex].

[84] Armen Tumasyan et al. "Search for a right-handed W boson and a heavy neutrino in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *JHEP* 04 (2022), p. 047. DOI: 10.1007/JHEP04(2022)047. arXiv: 2112.03949 [hep-ex].

[85] Christopher Brust et al. "Identifying boosted new physics with non-isolated leptons". In: *JHEP* 04 (2015), p. 079. DOI: 10.1007/JHEP04(2015)079. arXiv: 1410.0362 [hep-ph].

[86] Armen Tumasyan et al. "Search for long-lived heavy neutral leptons with displaced vertices in proton-proton collisions at $\sqrt{s} = 8$ TeV". In: *JHEP* 07 (2022), p. 081. DOI: 10.1007/JHEP07(2022)081. arXiv: 2201.05578 [hep-ex].

[87] Huilin Qu and Loukas Gouskos. "ParticleNet: Jet Tagging via Particle Clouds". In: *Phys. Rev. D* 101.5 (2020), p. 056019. DOI: 10.1103/PhysRevD.101.056019. arXiv: 1902.08570 [hep-ph].

[88] "Probing heavy Majorana neutrinos and the Weinberg operator through vector boson fusion processes in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: (June 2022). arXiv: 2206.08956 [hep-ex].

[89] Albert M Sirunyan et al. "Search for physics beyond the standard model in multilepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *JHEP* 03 (2020), p. 051. DOI: 10.1007/JHEP03(2020)051. arXiv: 1911.04968 [hep-ex].

[90] P. Abreu et al. "Searches for heavy neutrinos from Z decays". In: *Phys. Lett. B* 274 (1992), pp. 230–238. DOI: 10.1016/0370-2693(92)90528-C.

[91] P. Achard et al. "Search for heavy isosinglet neutrino in e^+e^- annihilation at LEP". In: *Phys. Lett. B* 517 (2001), pp. 67–74. DOI: 10.1016/S0370-2693(01)00993-5. arXiv: hep-ex/0107014.

[92] J. Alwall et al. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations". In: *JHEP* 07 (2014), p. 079. DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph].

[93] R. Frederix et al. "The automation of next-to-leading order electroweak calculations". In: *JHEP* 07 (2018). [Erratum: JHEP 11, 085 (2021)], p. 185. DOI: 10.1007/JHEP11(2021)085. arXiv: 1804.10017 [hep-ph].

[94] Enrique Fernandez-Martinez, Josu Hernandez-Garcia, and Jacobo Lopez-Pavon. "Global constraints on heavy neutrino mixing". In: *Journal of High Energy Physics* 2016.8 (Aug. 2016). arXiv:1605.08774 [hep-ph], p. 33. ISSN: 1029-8479. DOI: 10.1007/JHEP08(2016)033. URL: http://arxiv.org/abs/1605.08774 (visited on 01/04/2023).

[95] Pierre Artoisenet et al. "Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations". In: *JHEP* 03 (2013), p. 15. ISSN: 1029-8479. DOI: 10.1007/JHEP03(2013)015. arXiv: 1212.3460 [hep-ph].

[96] Pier Francesco Monni et al. "MiNNLO$_{PS}$: a new method to match NNLO QCD to parton showers". In: *JHEP* 05 (2020). [Erratum: JHEP 02, 031 (2022)], p. 143. DOI: 10.1007/JHEP05(2020)143. arXiv: 1908.06987 [hep-ph].

[97] Carlo Oleari. "The POWHEG-BOX". In: *Nucl. Phys. B Proc. Suppl.* 205-206 (2010). Ed. by Johannes Blümlein, Sven-Olaf Moch, and Tord Riemann, pp. 36–41. DOI: 10.1016/j.nuclphysbps.2010.08.016. arXiv: 1007.3893 [hep-ph].

[98] Elisabetta Barberio and Zbigniew Was. "PHOTOS: A Universal Monte Carlo for QED radiative corrections. Version 2.0". In: *Comput. Phys. Commun.* 79 (1994), pp. 291–308. DOI: 10.1016/0010-4655(94)90074-4.

[99] Enrico Bothmann et al. "Event Generation with Sherpa 2.2". In: *SciPost Phys.* 7.3 (2019), p. 034. DOI: 10.21468/SciPostPhys.7.3.034. arXiv: 1905.09127 [hep-ph].

[100] Keith Hamilton, Paolo Nason, and Giulia Zanderighi. "MINLO: Multi-Scale Improved NLO". In: *JHEP* 10 (2012), p. 155. DOI: 10.1007/JHEP10(2012)155. arXiv: 1206.3572 [hep-ph].

[101] Rikkert Frederix and Stefano Frixione. "Merging meets matching in MC@NLO". In: *JHEP* 12 (2012), p. 061. DOI: 10.1007/JHEP12(2012)061. arXiv: 1209.6215 [hep-ph].

[102] Albert M Sirunyan et al. "Pileup mitigation at CMS in 13 TeV data". In: *JINST* 15 (2020), P09018. DOI: 10.1088/1748-0221/15/09/p09018. arXiv: 2003.00503 [hep-ex].

[103] Andrew J. Larkoski et al. "Soft drop". In: *Journal of High Energy Physics* 2014.5 (May 2014). DOI: 10.1007/jhep05(2014)146. URL: https://doi.org/10.1007%5C%2Fjhep05%5C%282014%5C%29146.

[104] Emil Bols et al. "Jet Flavour Classification Using DeepJet". In: *JINST* 15.12 (2020), P12012. DOI: 10.1088/1748-0221/15/12/P12012. arXiv: 2008.10519 [hep-ex].

[105] Vardan Khachatryan et al. "Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8$ TeV". In: *JINST* 10.02 (2015), P02006. DOI: 10.1088/1748-0221/10/02/P02006. arXiv: 1411.0511 [physics.ins-det].

[106] Daniele Bertolini et al. "Pileup Per Particle Identification". In: *JHEP* 10 (2014), p. 059. DOI: 10.1007/JHEP10(2014)059. arXiv: 1407.6013 [hep-ph].

[107] A. Bodek et al. "Extracting Muon Momentum Scale Corrections for Hadron Collider Experiments". In: *Eur. Phys. J. C* 72 (2012), p. 2194. DOI: 10.1140/epjc/s10052-012-2194-8. arXiv: 1208.3710 [hep-ex].

[108] Jon Butterworth et al. "PDF4LHC recommendations for LHC Run II". In: *J. Phys. G* 43 (2016), p. 023001. DOI: 10.1088/0954-3899/43/2/023001. arXiv: 1510.03865 [hep-ph].

[109] A L Read. "Presentation of search results: the $\mathrm{CL_s}$ technique". In: *J. Phys. G* 28 (2002), p. 2693. DOI: 10.1088/0954-3899/28/10/313. URL: https://doi.org/10.1088%5C%2F0954-3899%5C%2F28%5C%2F10%5C%2F313.

[110] Thomas Junk. "Confidence level computation for combining searches with small statistics". In: *Nucl. Instrum. Meth. A* 434 (1999), p. 435. DOI: 10.1016/S0168-9002(99)00498-2. arXiv: hep-ex/9902006 [hep-ex].

[111] Waleed Abdallah et al. "Reinterpretation of LHC Results for New Physics: Status and Recommendations after Run 2". In: *SciPost Phys.* 9.2 (2020), p. 022. DOI: 10.21468/SciPostPhys.9.2.022. arXiv: 2003.07868 [hep-ph].

[112] Manuel Drees et al. *CheckMATE: Confronting your Favourite New Physics Model with LHC Data.* 2013. DOI: 10.48550/ARXIV.1312.2591. URL: https://arxiv.org/abs/1312.2591.

[113] Eric Conte, Benjamin Fuks, and Guillaume Serret. "MadAnalysis 5, a user-friendly framework for collider phenomenology". In: *Computer Physics Communications* 184.1 (Jan. 2013), pp. 222–256. DOI: 10.1016/j.cpc.2012.09.009. URL: https://doi.org/10.1016%5C%2Fj.cpc.2012.09.009.

[114] Sabine Kraml et al. "SModelS: a tool for interpreting simplified-model results from the LHC and its application to supersymmetry". In: *The European Physical Journal C* 74.5 (May 2014). DOI: 10.1140/epjc/s10052-014-2868-5. URL: https://doi.org/10.1140%5C%2Fepjc%5C%2Fs10052-014-2868-5.

[115] Eamonn Maguire, Lukas Heinrich, and Graeme Watt. "HEPData: a repository for high energy physics data". In: *Journal of Physics: Conference Series* 898 (Oct. 2017), p. 102006. DOI: 10.1088/1742-6596/898/10/102006. URL: https://doi.org/10.1088%5C%2F1742-6596%5C%2F898%5C%2F10%5C%2F102006.

[116] Christian Bierlich et al. "Robust Independent Validation of Experiment and Theory: Rivet version 3". In: *SciPost Phys.* 8 (2020), p. 026. DOI: 10.21468/SciPostPhys.8.2.026. arXiv: 1912.05451 [hep-ph].

[117] Vardan Khachatryan et al. "Phenomenological MSSM interpretation of CMS searches in pp collisions at sqrt(s) = 7 and 8 TeV". In: *JHEP* 10 (2016), p. 129. DOI: 10.1007/JHEP10(2016)129. arXiv: 1606.03577 [hep-ex].

[118] Georges Aad et al. "Summary of the ATLAS experiment's sensitivity to supersymmetry after LHC Run 1 — interpreted in the phenomenological MSSM". In: *JHEP* 10 (2015), p. 134. DOI: 10.1007/JHEP10(2015)134. arXiv: 1508.06608 [hep-ex].

[119] Sezen Sekmen. "Recent Developments in CMS Fast Simulation". In: *PoS* ICHEP2016 (2016), p. 181. DOI: 10.22323/1.282.0181. arXiv: 1701.03850 [physics.ins-det].

[120] "RECAST framework reinterpretation of an ATLAS Dark Matter Search constraining a model of a dark Higgs boson decaying to two b-quarks". In: (2019).

[121] A. Buckley et al. "Testing new physics models with global comparisons to collider measurements: the Contur toolkit". In: *SciPost Phys. Core* 4 (2021), p. 013. DOI: 10.21468/SciPostPhysCore.4.2.013. arXiv: 2102.04377 [hep-ph].

[122] Adam Alloul et al. "FeynRules 2.0 - A complete toolbox for tree-level phenomenology". In: *Comput. Phys. Commun.* 185 (2014), pp. 2250–2300. DOI: 10.1016/j.cpc.2014.04.012. arXiv: 1310.1921 [hep-ph].

[123] Richard D. Ball et al. "Parton distributions for the LHC Run II". In: *JHEP* 04 (2015), p. 040. DOI: 10.1007/JHEP04(2015)040. arXiv: 1410.8849 [hep-ph].

[124] Ashutosh V. Kotwal. "The W Boson Mass Measurement". In: *The Standard Theory of Particle Physics*. Chap. Chapter 10, pp. 185–200. DOI: 10.1142/9789814733519_0010. eprint: https://www.worldscientific.com/doi/pdf/10.1142/9789814733519_0010. URL: https://www.worldscientific.com/doi/abs/10.1142/9789814733519_0010.

[125] G. Arnison et al. "Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV". In: *Phys. Lett. B* 122 (1983), pp. 103–116. DOI: 10.1016/0370-2693(83)91177-2.

[126] M. Banner et al. "Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider". In: *Phys. Lett. B* 122 (1983), pp. 476–485. DOI: 10.1016/0370-2693(83)91605-2.

[127] V. M. Abazov et al. "Combination of CDF and D0 Results on W Boson Mass and Width". In: *Phys. Rev. D* 70 (2004), p. 092008. DOI: 10.1103/PhysRevD.70.092008. arXiv: hep-ex/0311039.

[128] "Precision Electroweak Measurements and Constraints on the Standard Model". In: (Dec. 2010). arXiv: 1012.2367 [hep-ex].

[129] Timo Antero Aaltonen et al. "Combination of CDF and D0 W-Boson Mass Measurements". In: *Phys. Rev. D* 88.5 (2013), p. 052018. DOI: 10.1103/PhysRevD.88.052018. arXiv: 1307.7627 [hep-ex].

[130] Morad Aaboud et al. "Measurement of the W-boson mass in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". In: *Eur. Phys. J. C* 78.2 (2018). [Erratum: Eur.Phys.J.C 78, 898 (2018)], p. 110. DOI: 10.1140/epjc/s10052-017-5475-4. arXiv: 1701.07240 [hep-ex].

[131] Roel Aaij et al. "Measurement of the W boson mass". In: *JHEP* 01 (2022), p. 036. DOI: 10.1007/JHEP01(2022)036. arXiv: 2109.01113 [hep-ex].

[132] P. A. Zyla et al. "Review of Particle Physics". In: *PTEP* 2020.8 (2020), p. 083C01. DOI: 10.1093/ptep/ptaa104.

[133] T. Aaltonen et al. "High-precision measurement of the W boson mass with the CDF II detector". In: *Science* 376.6589 (2022), pp. 170–176. DOI: 10.1126/science.abk1781.

[134] Jun Gao, DianYu Liu, and Keping Xie. "Understanding PDF uncertainty in W boson mass measurements*". In: *Chin. Phys. C* 46.12 (2022), p. 123110. DOI: 10.1088/1674-1137/ac930b. arXiv: 2205.03942 [hep-ph].

[135] Stephen P. Martin. "Three-loop QCD corrections to the electroweak boson masses". In: *Phys. Rev. D* 106.1 (2022), p. 013007. DOI: 10.1103/PhysRevD.106.013007. arXiv: 2203.05042 [hep-ph].

[136] Joshua Isaacson, Yao Fu, and C. -P. Yuan. "ResBos2 and the CDF W Mass Measurement". In: (May 2022). arXiv: 2205.02788 [hep-ph].

[137] S. Schael et al. "Precision electroweak measurements on the Z resonance". In: *Phys. Rept.* 427 (2006), pp. 257–454. DOI: 10.1016/j.physrep.2005.12.006. arXiv: hep-ex/0509008.

[138] P. A. Zyla et al. "Review of Particle Physics". In: *PTEP* 2020.8 (2020), p. 083C01. DOI: 10.1093/ptep/ptaa104.

[139] Michael E. Peskin and Tatsu Takeuchi. "Estimation of oblique electroweak corrections". In: *Phys. Rev. D* 46 (1992), pp. 381–409. DOI: 10.1103/PhysRevD.46.381.

[140] Alejandra Melfo et al. "Type II Seesaw at LHC: The Roadmap". In: *Phys. Rev. D* 85 (2012), p. 055018. DOI: 10.1103/PhysRevD.85.055018. arXiv: 1108.4416 [hep-ph].

[141] "Search for doubly charged Higgs boson production in multi-lepton final states using 139 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: (Nov. 2022). arXiv: 2211.07505 [hep-ex].

[142] Morad Aaboud et al. "Search for doubly charged scalar bosons decaying into same-sign W boson pairs with the ATLAS detector". In: *Eur. Phys. J. C* 79.1 (2019), p. 58. DOI: 10.1140/epjc/s10052-018-6500-y. arXiv: 1808.01899 [hep-ex].

[143] Georges Aad et al. "Search for doubly and singly charged Higgs bosons decaying into vector bosons in multi-lepton final states with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *JHEP* 06 (2021), p. 146. DOI: 10.1007/JHEP06(2021)146. arXiv: 2101.11961 [hep-ex].

[144] Shinya Kanemura et al. "LHC Run-I constraint on the mass of doubly charged Higgs bosons in the same-sign diboson decay scenario". In: *PTEP* 2015 (2015), 051B02. DOI: 10.1093/ptep/ptv071. arXiv: 1412.7603 [hep-ph].

[145] Shinya Kanemura et al. "Bounds on the mass of doubly-charged Higgs bosons in the same-sign diboson decay scenario". In: *Phys. Rev. D* 90.11 (2014), p. 115018. DOI: 10.1103/PhysRevD.90.115018. arXiv: 1407.6547 [hep-ph].

[146] Shinya Kanemura, Kei Yagyu, and Hiroshi Yokoya. "First constraint on the mass of doubly-charged Higgs bosons in the same-sign diboson decay scenario at the LHC". In: *Phys. Lett. B* 726 (2013), pp. 316–319. DOI: 10.1016/j.physletb.2013.08.054. arXiv: 1305.2383 [hep-ph].

[147] Julian Heeck. "W-boson mass in the triplet seesaw model". In: *Phys. Rev. D* 106.1 (2022), p. 015004. DOI: 10.1103/PhysRevD.106.015004. arXiv: 2204.10274 [hep-ph].

[148] Jon Butterworth et al. "Testing the Scalar Triplet Solution to CDF's Fat W Problem at the LHC". In: (Oct. 2022). arXiv: 2210.13496 [hep-ph].

[149] Georges Aad et al. "Measurements of differential cross-sections in four-lepton events in 13 TeV proton-proton collisions with the ATLAS detector". In: *JHEP* 07 (2021), p. 005. DOI: 10.1007/JHEP07(2021)005. arXiv: 2103.01918 [hep-ex].

[150] Morad Aaboud et al. "Search for triboson $W^{\pm}W^{\pm}W^{\mp}$ production in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *Eur. Phys. J. C* 77.3 (2017), p. 141. DOI: 10.1140/epjc/s10052-017-4692-1. arXiv: 1610.05088 [hep-ex].

[151] Georges Aad et al. "Measurements of $W^+W^- + \geq 1$ jet production cross-sections in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *JHEP* 06 (2021), p. 003. DOI: 10.1007/JHEP06(2021)003. arXiv: 2103.10319 [hep-ex].

[152] Shinya Kanemura and Kei Yagyu. "Implication of the W boson mass anomaly at CDF II in the Higgs triplet model with a mass difference". In: *Phys. Lett. B* 831 (2022), p. 137217. DOI: 10.1016/j.physletb.2022.137217. arXiv: 2204.07511 [hep-ph].

[153] Rituparna Ghosh, Biswarup Mukhopadhyaya, and Utpal Sarkar. "The ρ parameter and the CDF W-mass anomaly: observations on the role of scalar triplets". In: (May 2022). arXiv: 2205.05041 [hep-ph].

Abstract

힉스 보존의 발견으로 표준 모형은 오늘날 가장 성공적인 물리 모형으로 여겨질 수 있다. 하지만 이 모형에서 예측하는 바와 위배되는 실험적인 관측 결과들이 있어 여전히 몇 가지 질문들에 답을 할 수가 없기도 하다. 중성미자의 진동 현상은 그 대표적인 예시인데 이는 표준 모형을 넘어서는 새로운 물리가 존재한다는 명백한 증거가 된다. 중성미자가 진동한다는 사실은 중성미자들의 질량 고유 상태와 맛깔 고유 상태가 서로 일치하지 않는다는 점을 드러내며 이는 중성미자들의 질량이 모두 0일 수 없다는 것을 보여준다. 이는 표준 모형에서 중성미자의 질량이 0이라는 가정과 배치되는 실험적인 관측 결과가 된다. 많은 이론들이 중성미자가 질량을 부여받는 방식을 설명하고자 제시되었다. 그 중 시소 모형은 중성미자 질량의 근원과 매우 작은 크기의 중성미자 질량을 동시에 설명할 수 있는 모형이다.

이 학위논문에서는 중성미자 질량에 대한 해답을 찾기 위해 대형 강입자 충돌기를 이용한 두 가지 연구를 진행하였다. 첫번째는 타입-1 시소 모형에서 가정되는 무거운 중성미자를 하나의 렙톤과 반경이 넓은 젯 혹은 반경이 좁은 젯들을 이용해 찾으려는 연구이다. 이 때, 양성자-양성자 충돌 에너지가 13 TeV일 때 CMS 검출기로 수집된 137 fb^{-1}의 런 2 데이터를 사용하였다. 이 연구는 두 가지 방향으로 진행 되었는데, 이는 사건이 어떤 젯으로 구성되는지에 따라 달라지며 이를 통해 더 효과적인 탐색이 가능했다. 추가적으로 이 연구는 무거운 중성미자가 힉스 보존으로 붕괴하는 경우를 처음으로 고려하였다. 이 학위논문의 두번째 연구는 타입-2 시소 모형에서 정의되는 변수들을 이미 출판된 LHC의 결과들을 재해석하는 방식으로 진행되었다. 특히 CDF 검출기에서 최근에 발표한 W 보존의 질량이 표준 모형에서 예측한 값과 큰 차이를 보였는데 이 결과에서 선호되는 변수들에 대한 분석을 하였다.

Keywords: High Energy Physics, LHC, CMS, Thesis, Neutrino, Heavy Neu-

trino, Heavy Neutral Lepton, Seesaw Model

Student Number: 2015-22596