University of Cyprus
Department of Physics

Searches for Beyond the Standard Model Signatures with the CMS Detector at LHC

Leonidas Paizanos

Research Advisors: Dr. Fotios Ptochos and Dr. Halil Saka

May 2022

Contents

List of Figures **iii**

1 Standard Model and Beyond **1**
1.1 Standard Model . . . 1
1.2 Open Questions in Particle Physics . . . 3
1.2.1 Neutrino Oscillations . . . 3
1.2.2 Gravity . . . 4
1.2.3 Dark Matter and Dark Energy . . . 4
1.2.4 Matter and Anti-Matter Asymmetry . . . 5
1.2.5 Hierarchy Problem . . . 5
1.3 Physics Beyond the Standard Model . . . 6
1.3.1 Charged Higgs Boson . . . 6
1.3.2 Heavy Fermions . . . 8

2 The LHC Accelerator and CMS Detector **9**
2.1 Large Hadron Collider (LHC) . . . 9
2.1.1 Luminosity . . . 10
2.2 Compact Muon Solenoid (CMS) . . . 11
2.2.1 CMS Coordinates System . . . 12
2.2.2 Superconducting Magnet . . . 13
2.2.3 Tracker . . . 14
2.2.4 Electromagnetic Calorimeter . . . 15
2.2.5 Hadronic Calorimeter . . . 17
2.2.6 Muon Detector . . . 18
2.2.7 Triggering System . . . 19

3 Event Reconstruction and Simulation **20**
3.1 Event Reconstruction . . . 20
3.1.1 Primary Vertices . . . 20
3.1.2 Electrons and Photons . . . 20
3.1.2.1 Cut-Based Electron Identification . . . 21
3.1.3 Muons . . . 23
3.1.4 Hadrons And Non Isolated Photons . . . 24
3.1.5 Jets . . . 25
3.1.6 B-Tagged Jets . . . 25
3.1.7 Hadronically decaying τ leptons . . . 26
3.1.8 Hadronically decaying top quarks . . . 27
3.1.9 Missing Transverse Energy . . . 28
3.2 Event Simulation . . . 28

4 Search for Charged Higgs boson decaying to a top and a bottom quark in lepton plus jets final state 30
4.1 Trigger and Characteristics of Final State Objects . . . 31
4.1.1 Trigger . . . 32
4.1.2 MET Filters . . . 32
4.1.3 Primary Vertex . . . 33
4.1.4 Electrons . . . 33
4.1.5 Muons . . . 33
4.1.6 Tau Veto . . . 33
4.1.7 Jets and BJets . . . 34
4.2 Initial Selections . . . 34
4.3 Object Reconstruction . . . 36
4.3.1 Hadronicaly Decaying Objects . . . 36
4.3.2 Missing Energy . . . 37
4.3.3 Leptonicaly Decaying Objects . . . 38
4.3.4 BJet From Higgs . . . 39

5 Signal and Background Discrimination 40
5.1 Background Process . . . 40
5.2 Discrimination Between Signal and Background Events . . . 46
5.3 Multivariate BDTG Classifier . . . 47
5.3.1 Expected Limits . . . 61

6 Search for type-III Seesaw Heavy Fermions in muons plus AK8-jets final state 63
6.1 Selections: . . . 64
6.2 Heavy fermion mass reconstruction . . . 67
6.3 Increasing Signal Sensitivity . . . 68
6.4 Extraction of the Expected Limits . . . 74

7 Conclusions 75

Bibliography 77

List of Figures

1.1.1 The SM of particle physiscs. . . . 2
1.2.1 The relation between flavour and mass eigenstates [4] . . . 3
1.2.2 The main loop diagrams that contribute to Higgs mass calculation . . . 5
1.3.1 LO diagram for the production of a light charged Higgs with mass $mH^{\pm} < m_t - m_b$ [8] . . . 7
1.3.2 LO diagram for the production of a heavy charged Higgs with mass $mH^{\pm} > m_t - m_b$ [9] . . . 7
1.3.3 LO diagram for the production of the intermediate Charged Higgs with mass $mH^{\pm} \approx m_t - m_b$ [8] . . . 8
1.3.4 Production process of heavy fermion pairs . . . 8

2.1.1 Cern accelerating system. [12] . . . 10
2.2.1 CMS detector [14]. . . . 11
2.2.2 Transverse cut of the CMS detector. . . . 12
2.2.3 Coordinate convention for CMS detector . . . 12
2.2.4 The tracker detector [16]. . . . 14
2.2.5 CMS ECAL . . . 15
2.2.6 Electromagnetic Shower . . . 16
2.2.7 Hadronic Calorimeter of CMS. . . . 17
2.2.8 Sampling calorimeter . . . 17
2.2.9 CMS muon detector . . . 18
2.2.10 The path for the different particles in CMS detector. . . . 19

3.1.1 Schematic representation of a heavy-flavor jet being produced at the secondary vertex. . . . 26
3.1.2 Tau decay modes . . . 27
3.2.1 Monte Carlo Simulation . . . 29
3.2.2 The amount of energy that each one of the partons carry while the proton's energy increase . . . 29

4.0.1 Leading order Feynman diagram of a charged Higgs boson produced in association with a top quark and decaying into a top and bottom quark-antiquark pair. . . . 31
4.0.2 W boson decay modes [19] . . . 31
4.1.1 Efficiency of the HLT-Ele35-WPTight-Gsf (left) and HLT-IsoMu27 (right) triggers as a function of the electron's pT, using the 2017 datasets . . . 32
4.2.1 Jet multiplicity for Muon final state for three different masses of Charged Higgs . . . 35
4.2.2 Jet multiplicity for Electron final state for three different masses of Charged Higgs . . . 35

4.2.3 bjet multiplicity for Muon final state for three different masses of Charged Higgs . . . 35
4.2.4 bjet multiplicity for Electron final state for three different masses of Charged Higgs . . . 35

5.1.1 Diboson (WW) . . . 40
5.1.2 Diboson (WZ) . . . 40
5.1.3 Diboson (ZZ) . . . 40
5.1.4 $t\bar{t}$ . . . 40
5.1.5 $t\bar{t} + X$ . . . 40
5.1.6 $t\bar{t}t\bar{t}$ . . . 40
5.1.7 Drell Yan (leptons) . . . 41
5.1.8 Diboson (leptons + Jets) . . . 41
5.1.9 Single Top (4FS) . . . 41
5.1.10 Single Top (5FS) . . . 41
5.1.11 W + Jets . . . 41
5.1.12 Jet multiplicity. . . . 43
5.1.13 BJet multiplicity. . . . 44
5.1.14 Lepton's P and eta. . . . 44
5.1.15 MET and HT. . . . 45
5.1.16 P_T of the Leading BJet. . . . 45
5.3.1 Schematic view of a decision tree . . . 47
5.3.2 HT (left), Min ΔR between two bjets (center) and Min ΔR between lepton and bjets (right) for signal hypothesis of charged Higgs with mass 200 GeV. 50
5.3.3 The minimum mass between a b-tagged jet and the system of a di-jet or a jet-bjet (left), Max $\Delta\eta$ between two bjets (center) and ΔR between the third b-tagged jet and the system of a di-jet or a jet-bjet with minimum P_T (right) for signal hypothesis of charged Higgs with mass 200 GeV. . . . 50
5.3.4 ΔR between the first and third bjets (left), ΔR of the two bjets with minimum $\Delta\eta$ (center) and ΔR of lepton and third bjet (right) for signal hypothesis of charged Higgs with mass 200 GeV. . . . 50
5.3.5 Circularity (left), Min ΔR between higgs bjet and a di-bjet (center), $\Delta\phi$ between MET and third jet (right) and $\Delta\phi$ between MET and leading bjet (bottom) for signal hypothesis of charged Higgs with mass 200 GeV. . . . 51
5.3.6 Centrality (left), ΔR between the leading non b-tagged jet and the system of a di-jet or a jet-bjet with minimum P_T (center) and ΔR between first jet and second bjet (right) for signal hypothesis of charged Higgs with mass 200 GeV. . . . 51
5.3.7 Min $\Delta\Phi$ between the leptonic bjet and the di-bjet with the minimum mass (left), Min mass of a tri-bjet (center) and Min mass from lepton plus bjet system (right) for signal hypothesis of charged Higgs with mass 200 GeV. . 52
5.3.8 Correlation matrices for SR1 (left), SR2 (center) and SR3 (right). . . . 52
5.3.9 At least four jets including exactly three b-tagged jets. . . . 53
5.3.10 At least five jets including exactly two b-tagged jets. . . . 53
5.3.11 At least five jets including at least four b-tagged jets. . . . 53
5.3.12 HT (left), Min ΔR between two bjets (center) and Min ΔR between lepton and bjets (right) for signal hypothesis of charged Higgs with mass 500 GeV. 54

5.3.13 The minimum mass between a b-tagged jet and the system of a di-jet or a jet-bjet (left), Max $\Delta\eta$ between two bjets (center) and ΔR between the third b-tagged jet and the system of a di-jet or a jet-bjet with minimum P_T (right) for signal hypothesis of charged Higgs with mass 500 GeV. . . . 54
5.3.14 ΔR between the first and third bjets (left), ΔR of the two bjets with minimum $\Delta\eta$ (center) and ΔR of lepton and third bjet (right) for signal hypothesis of charged Higgs with mass 500 GeV. 54
5.3.15 Circularity (left), Min ΔR between higgs bjet and a di-bjet (center), $\Delta\phi$ between MET and third jet (right) and $\Delta\phi$ between MET and leading bjet (bottom) for signal hypothesis of charged Higgs with mass 500 GeV. 55
5.3.16 Centrality (left), ΔR between the leading non b-tagged jet and the system of a di-jet or a jet-bjet with minimum P_T (center) and ΔR between first jet and second bjet (right) for signal hypothesis of charged Higgs with mass 500 GeV. 55
5.3.17 Min $\Delta\Phi$ between the leptonic bjet and the di-bjet with the minimum mass (left), Min mass of a tri-bjet (center) and Min mass from lepton plus bjet system (right) for signal hypothesis of charged Higgs with mass 500 GeV. . 56
5.3.18 Correlation matrices for SR1 (left), SR2 (center) and SR3 (right). 56
5.3.19 At least four jets including exactly three b-tagged jets. 56
5.3.20 At least five jets including exactly two b-tagged jets. 57
5.3.21 At least five jets including at least four b-tagged jets. 57
5.3.22 HT (left), Min ΔR between two bjets (center) and Min ΔR between lepton and bjets (right) for signal hypothesis of charged Higgs with mass 1000 GeV. 57
5.3.23 The minimum mass between a b-tagged jet and the system of a di-jet or a jet-bjet (left), Max $\Delta\eta$ between two bjets (center) and ΔR between the third b-tagged jet and the system of a di-jet or a jet-bjet with minimum P_T (right) for signal hypothesis of charged Higgs with mass 1000 GeV. . . . 58
5.3.24 ΔR between the first and third bjets (left), ΔR of the two bjets with minimum $\Delta\eta$ (center) and ΔR of lepton and third bjet (right) for signal hypothesis of charged Higgs with mass 1000 GeV. 58
5.3.25 Circularity (left), Min ΔR between higgs bjet and a di-bjet (center), $\Delta\phi$ between MET and third jet (right) and $\Delta\phi$ between MET and leading bjet (bottom) for signal hypothesis of charged Higgs with mass 1000 GeV. . . . 59
5.3.26 Centrality (left), ΔR between the leading non b-tagged jet and the system of a di-jet or a jet-bjet with minimum P_T (center) and ΔR between first jet and second bjet (right) for signal hypothesis of charged Higgs with mass 1000 GeV. 59
5.3.27 Min $\Delta\Phi$ between the leptonic bjet and the di-bjet with the minimum mass (left), Min mass of a tri-bjet (center) and Min mass from lepton plus bjet system (right) for signal hypothesis of charged Higgs with mass 1000 GeV. 60
5.3.28 Correlation matrices for SR1 (left), SR2 (center) and SR3 (right). 60
5.3.29 At least four jets including exactly three b-tagged jets. 60
5.3.30 At least five jets including exactly two b-tagged jets. 61
5.3.31 At least five jets including at least four b-tagged jets. 61
5.3.32 The observed and expected limits 95% CL upper limits for the $H^{\pm}$ mass range of 200-3000 GeV 62

6.0.1 Example processes illustrating production and decay of type-III seesaw heavy fermion pairs at the LHC that result in lepton plus jets final states . 63
6.1.1 Multiplicity of AK8 Jets for events with one or two muons 65

6.1.2 Pruned Mass of Leading (left) and sub-leading (right) in P_T AK8 Jets in single lepton final state . . . 66
6.1.3 Pruned Mass of Leading (left) and sub-leading (right) in P_T AK8 Jets in single lepton final state . . . 66
6.2.1 Minimum and Maximum mass of a muon and a fat jet over all possible combinations for single-muon final state . . . 67
6.2.2 Minimum and Maximum mass of a muon and a fat jet over all possible combinations for di-muon final state . . . 67
6.3.1 P_T of the leading muon for single muon and di-muon final states . . . 68
6.3.2 Reconstructed mass of the heavy fermion for each cut (single-muon final state) . . . 69
6.3.3 Reconstructed mass of the heavy fermion for each cut (di-muon final state) 69
6.3.4 P_T of the leading AK8 jet for single muon and di-muon final states . . . 69
6.3.5 Reconstructed mass of the heavy fermion for each cut (single-muon final state) . . . 69
6.3.6 Reconstructed mass of the heavy fermion for each cut (di-muon final state) 70
6.3.7 MET for single muon and di-muon final states . . . 70
6.3.8 Reconstructed mass of the heavy fermion for each cut (single-muon final state) . . . 70
6.3.9 Reconstructed mass of the heavy fermion for each cut (di-muon final state) 71
6.3.10 MET for single muon final state . . . 71
6.3.11 Reconstructed mass of the heavy fermion for each cut (di-muon final state) 71
6.3.12 P_T of the sub-leading in P_T muon for di-muon final state . . . 72
6.3.13 Reconstructed mass of the heavy fermion for each cut (di-muon final state) 72
6.3.14 BJet multiplicity for single muon and di-muon final states . . . 73
6.3.15 Reconstructed mass of the heavy fermion after all the selections for single-muon and di-muon final states . . . 74

7.0.1 Expected limits for current analyses (left) 2016 CMS analyses (right) . . . 75
7.0.2 Expected limits of heavy fermion production in multi-lepton final state (2021 CMS analyses)[20] . . . 76

Chapter 1

Standard Model and Beyond

1.1 Standard Model

With the discovery of the Higgs boson at the LHC in 2012, the standard model (SM) [1] of particle physics was completed. The SM describes the elementary particles that form our universe and the interactions between them, providing at the same time a successful explanation of the current experimental data. The SM contains twelve fermions with spin -1/2 (six leptons and six quarks), four gauge bosons with spin -1 that are responsible for the mediation of the interactions between the particles, and the scalar boson, Higgs.

Fermions are separated into three generations. The first is composed of the electron, the electron neutrino, and the up and down quarks. The particles of the second generation have the same properties as the associate fermion of the first, but there are heavier and more unstable. The same thing happens between the fermions of the third generation. Muon, muon neutrino, and charm and strange quarks are the fermions of the second generation, while tau, tau neutrino, and top and bottom quarks form the third generation. There is an anti-particle with the same properties but the opposite charge for each of these twelve particles.

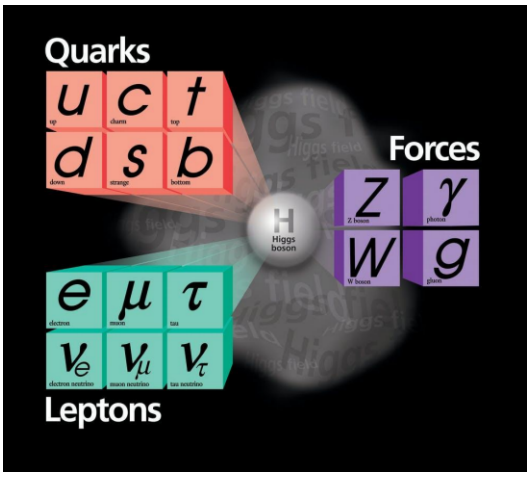

FIGURE 1.1.1: The SM of particle physiscs.

The particles interact via gravitational, electromagnetic, strong and weak forces. The SM includes all the interactions except the gravitational force. All the fundamental particles interact through the weak force because they carry the charge of the weak force (weak isospin). Additionally, apart from neutrinos that are described from SM as massless particles with zero charge, all fermions feel the gravity and the electromagnetic force. Now for the strong interactions, the presence of the QCD charge carrier is required, known as color. Only quarks have the charge carrier of QCD, making them the only particles that can participate in strong interactions. Quarks come in three colors, red, green, and blue. Due to the nature of strong interactions, the properties of quarks are different from those of fermions. They don't exist as free particles. Instead, quarks form bound states known as hadrons.

Hadrons are split into two different categories, baryons, and mesons. Mesons are the bound state of a quark and anti-quark pair that carries the opposite color for the mesons to be color-neutral. Baryons are composed of three quarks with different colors ensuring that baryons will be also color neutral.

Photon, gluon, and W and Z bosons are the carriers of electromagnetic, strong, and weak forces. The first two are massless, and all of them, apart from W, have zero electric charge. As the carrier of the strong interactions, the gluon also has a color charge. Specifically, it carries a combination of two colors, ensuring that the color charge will be conserved during quark-quark with an exchange of a gluon or gluon-gluon interactions. Eight types of gluons represent the eight combinations between the three colors.

Another interesting point is that each interaction has some specific properties regarding the flavor conservation of the particle that participates in that decay. The fermions

that interact via electromagnetic force never change flavor. The same happens with the quarks participating in the strong interactions. The only case where flavor can change is during the weak interactions in which the W boson is present.

The last piece of SM is the scalar Higgs boson. It has a mass of 125 GeV and is the only boson with zero spin. Higgs field's has an essential purpose in the SM. The Higgs mechanism is the process that gives mass to all the other particles.

1.2 Open Questions in Particle Physics

The SM is undoubtedly one of the most outstanding achievements of modern particle physics. Yet it fails to answer many crucial questions about our universe. We will study some of the most critical problems in the following sub-chapters.

1.2.1 Neutrino Oscillations

As we have discussed, SM predicts that the neutrinos are massless particles. However, in an experiment conducted to detect solar neutrinos (ν_e), a discrepancy between the expected and the detected number of the electron neutrinos was observed, something known as the solar neutrino problem. The idea of neutrino oscillations [2][3] that led to their flavor changing was first expressed in 1967 by Bruno Pontecorvo. The definitive answer came at the end of the 1990s from the Super–Kamiokande experiment, which confirmed that neutrinos can change flavor. This phenomenon can't occur if their mass is zero, which contradicts the SM about neutrinos. To understand the effect of this process on the mass of the neutrino, we can study their oscillation between two flavors. Considering that the flavor eigenstates ν_e, ν_μ and ν_τ can be formed from the superimposition of the mass eigenstates ν_1, ν_2 and ν_3 the relation between them is shown in figure (1.2.1) and if we chose only the muon and tau neutrinos just to simplify the situation, the expressions that connect them with the eigenstates ν_2 and ν_3 are equations:

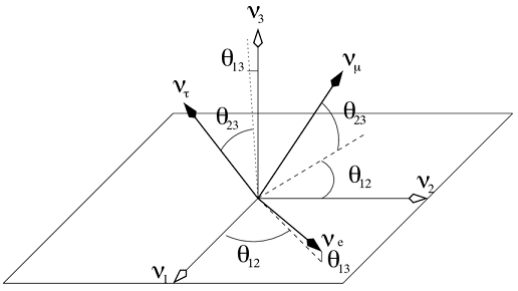

FIGURE 1.2.1: The relation between flavour and mass eigenstates [4]

$$\nu_\mu = \cos\theta_{23} * \nu_2 * e^{-iE_2t} + \sin\theta_{23} * \nu_3 * e^{-iE_3t} \tag{1.2.1}$$

$$\nu_\tau = -\sin\theta_{23} * \nu_2 * e^{-iE_2t} + \cos\theta_{23} * \nu_3 * e^{-iE_3t} \tag{1.2.2}$$

The time dependent probabilities of a muon neutrino to change to a tau neutrino or to remain a muon neutrino are given as:

$$P(\nu_\mu \to \nu_\tau, t) = \sin^2(2\theta_{23}) * \sin^2 \frac{1.27 * (\Delta m)^2 * L}{E} \tag{1.2.3}$$

$$P(\nu_\mu \to \nu_\mu, t) = 1 - \sin^2(2\theta_{23}) * \sin^2 \frac{1.27 * (\Delta m)^2 * L}{E} \tag{1.2.4}$$

- L = Distance from the neutrino's production point (Km).
- E = Neutrino's energy (GeV).
- $\Delta m^2 = m2^2 - m3^2$ (eV^2)

From equations 1.2.3 and 1.2.4, we can obtain that if the neutrino mass is zero, then the Δm quantity will be zero, hence the probability of changing flavour should be also zero, which will be in contradiction with experimental observations.

1.2.2 Gravity

It's the force we observe in our everyday lives, yet it is so weak that it is impossible to measure on the subatomic scale. As a result, SM doesn't offer an explanation for the gravitational interactions [5] between the particles. Graviton is considered to be the carrier of the gravitational force, and in contrast with the carriers of the other forces, it has spin -2 and still remains undetected. Additionally, if we approximately include it in the SM, the experimental data disagree with the SM predictions. Today gravity is being described by the general relativity and there are attempts for the extraction of a quantum field theory.

1.2.3 Dark Matter and Dark Energy

The rotational speed of galaxies in our universe leads to the conclusion that it is impossible for the gravity generated from the observed matter to hold them together. This

can only be explained if we assume that an unknown type of matter that we haven't yet detected provides them with extra mass and subsequently generates extra gravity. This different type of matter is known as dark matter and makes up approximately 27% of the universe. Dark matter [3] remains undetected because it only interacts gravitationally. It can't absorb, emit or reflect light. The only way to be found is from its gravitational interactions with visible matter. Another paradox is the expansion of the universe. If it was formed only by matter, its growth should have slowed down due to the effect of gravity. Dark energy [6] makes up the 68% of the universe and is considered the main suspect for accelerating the universe's expansion. Connected with the vacuum space has no topological effect but rather a global impact on the universe. Considering all these, we can conclude that SM can describe only the 5% of the universe.

1.2.4 Matter and Anti-Matter Asymmetry

In 1933 Paul Dirac won the Nobel prize for his equation that combined quantum theory with special relativity to describe the movement of an electron that travels with relativistic speed. The two solutions of this equation conclude that the electron can have either positive or negative energy. The negative energy case was explained by the introduction of the anti-particles [7]. In this case, the positron, that has the same properties as the electron but the opposite charge. This applies to all particles. Thus, the matter is formed by particles and anti-matter from anti-particles. An interesting effect is that when a particle comes in contact with its anti-particle, they annihilate. The question now is, why our world is made out of particles and not anti-particles, and if during the generation of the universe, the amount of matter and anti-matter were equal, why they did not annihilate.

1.2.5 Hierarchy Problem

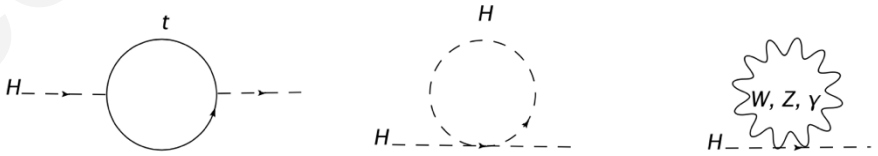

FIGURE 1.2.2: The main loop diagrams that contribute to Higgs mass calculation

The hierarchy problem has to do with the mass of the Higgs boson. To compute Higgs mass, we have to sum the contribution of all the processes that are involved. The first term represents the tree-level contribution. The SM is a stable and well-behaved quantum field theory at tree-level. Thus, this term does not cause any problems. The

"dangerous" terms are the ones that come from the loop diagrams. Their contribution is calculated by integrating over the momentum of the fermions/bosons that participate in each diagram. Hence, it depends on the square of the particle's momentum.

$$m_{H^2} = m_{tree}^2 - \frac{3}{8\pi^2}\lambda_t^2\Lambda^2 + \frac{1}{16\pi^2}g^2\Lambda^2 + \frac{3}{8\pi^2}\lambda^2\Lambda^2 \tag{1.2.5}$$

For high energy scales, some loop diagrams, such as the terms containing the top quark, the W and Z bosons, or even the Higgs boson itself, diverge. To avoid this effect, we introduce a cut-off energy scale Λ at approximately 10 TeV, assuming that below this energy scale, SM is a correct and effective theory. Additionally, we must carefully adjust the values of the parameters λ_t, g, and λ to ensure the cancellation of diverging terms (fine-tuning). Fine-tuning can fix the problem but is an unnatural solution, and as the energy scale increases, it gets worst. This leads to the necessity of new physics for energy scales above Λ that provide a more natural solution.

1.3 Physics Beyond the Standard Model

Some of the open questions that we have discussed in the previous section can be explained by other models that contain new physics. In our analysis we are going to search for the production of charged higgs bosons and heavy fermions.

1.3.1 Charged Higgs Boson

Adding a Higgs field doublet, it's one of the most straightforward possible extensions of SM and is included in several BSM models such as the Two-Higgs-Doublet models (2HDM). The introduction of an additional Higgs doublet leads to the production of 5 Higgs bosons (3 neutral and 2 charged). Furthermore, based on the coupling between the fermions and the Higgs doublet fields we have 4 types of 2HDM:

- Type-I: All quarks and leptons couple only to the second doublet.
- Type-II: Up-type quarks couple to the second doublet while down-type and leptons are coupled to the first.
- Type-III: Quarks couple to the second doublet and leptons to the first.
- Type-IV: Up-type quarks and leptons couple to the second doublet and down-type are couple to the first.

SUSY is a popular extension of SM in which each SM particle has a super-partner (sparticle). It solves the hierarchy problem because there is also an additional term originating from the partner of the particle that participates in the first loop diagram for each Higgs mass correction. Thus, the two terms will cancel each other. If the particles and the sparticles have the same mass, the cancellation will be exact. Furthermore, SUSY predicts the existence of a lightest supersymmetric particle which is stable and therefore is a possible candidate for explaining dark matter.

Minimal Supersymmetric Standard Model (MSSM) is a Type-II 2HDM thus, it contains 2 charged Higgs bosons. Thus, the detection of a charged Higgs particle will prove the existence of BSM physics.

Production Process

This section discusses the different production processes of the charged Higgs boson that is predict in the MSSM. The production process can split into three distinct regions connected with the charged Higgs mass.

- The light charged Higgs with mass smaller than the difference between the top quark and bottom quark mass.

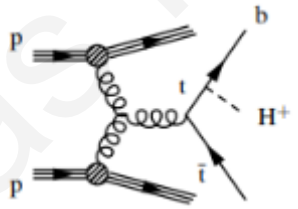

FIGURE 1.3.1: LO diagram for the production of a light charged Higgs with mass $mH^{\pm} < m_t - m_b$ [8]

- The heavy charged Higgs with mass bigger from the difference between the top and bottom quarks mass.

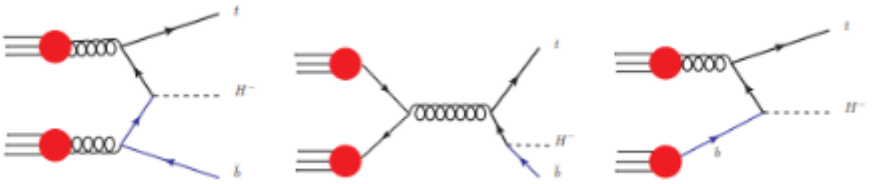

FIGURE 1.3.2: LO diagram for the production of a heavy charged Higgs with mass $mH^{\pm} > m_t - m_b$ [9]

- The intermediate region where the mass of the charged Higgs is similar to the mass of the top quark.

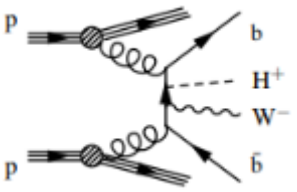

FIGURE 1.3.3: LO diagram for the production of the intermediate Charged Higgs with mass $mH^{\pm} \approx m_t - m_b$ [8]

1.3.2 Heavy Fermions

The seesaw mechanism [10] is a model that introduces neutrino mass to the SM. The first part of this mechanism includes a new Dirac mass term added to the Lagrangian of SM. This term represents the interaction with the Higgs field and is similar to the analogous term that gives electrons their mass. However, this solution explains the non-zero mass of the neutrinos, but it provides them with a much bigger mass than the one measured during experiments. Here comes the second part of this model, introducing an additional Majorana mass term. The combination of the two terms gives two possible physical states. The first refers to a light neutrino with a mass of approximately 0.01 eV that can be the SM neutrino. The second is a heavy neutrino with mass at around 10^{11} GeV. Finally, a Majorana term allows the transition of a particle to its anti-particle. This would be a problem for charged particles because it violates charge conservation. Since neutrinos are particles with zero charge, this is not a problem in our case. Thus, neutrinos can be Majorana particles and be their own anti-particles.

The Type-III seesaw mechanism predicts the existence of a fermionic triplet consist of two charged leptons ($\Sigma^{\pm}$) and a Majorana neutral lepton (Σ^0). The three particles are mainly produced in pairs, either charged-charged or charged-neutral, as it shown in figure.

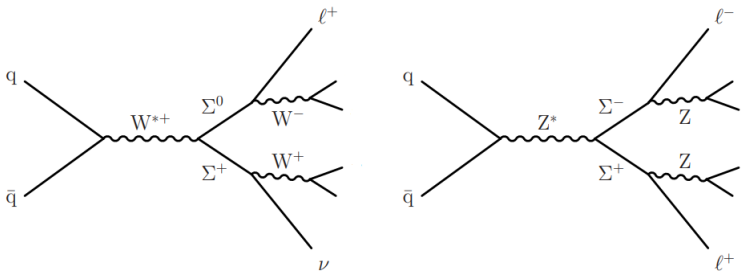

FIGURE 1.3.4: Production process of heavy fermion pairs

The discovery of heavy fermions with these properties will confirm the seesaw mechanism and solve the problem of neutrino masses in the SM.

Chapter 2

The LHC Accelerator and CMS Detector

2.1 Large Hadron Collider (LHC)

With a perimeter of 27 km and center of mass energy of 13 TeV, LHC [11] is the world's largest and most powerful accelerator. Located at CERN, it first started in 2008, replacing the electron-positron accelerator LEP. Its purpose is to explore the SM physics and search for physics beyond the SM during proton-proton collisions. The fact that the collisions occur between two particles and not a particle with its anti-particle leads to the necessity of two separate rings. The beams accelerate in the two rings to reach their final energy of 6.5 TeV and collide. The collisions occur at four points of LHC where the two rings are connected. At the collision places, there are the four detectors of CERN: Compact Muon Solenoid (CMS), A Toroidal LHC Apparatus (ATLAS), A Large Ion Collider Experiment (ALICE), and Large Hadron Collider beauty (LHCb). The journey of the two proton beams doesn't start at the LHC. There is an accelerating system that gradually increases their energy. Linac 2 is their first station. A linear accelerator that uses RF cavities to accelerate the beam to 50 MeV. Afterwards, the beam enters the Proton Synchrotron Booster, consisting of four synchrotrons that accelerate the beam at 1.4 GeV. Then the Proton Synchrotron arrives, where the beam's energy increases to 25 GeV. Finally, the last stop before the LHC is the Super Proton Synchrotron, where the beam accelerates until its energy reaches 450 GeV. Each beam has 2808 packets, and each one contains 10^{11} protons.

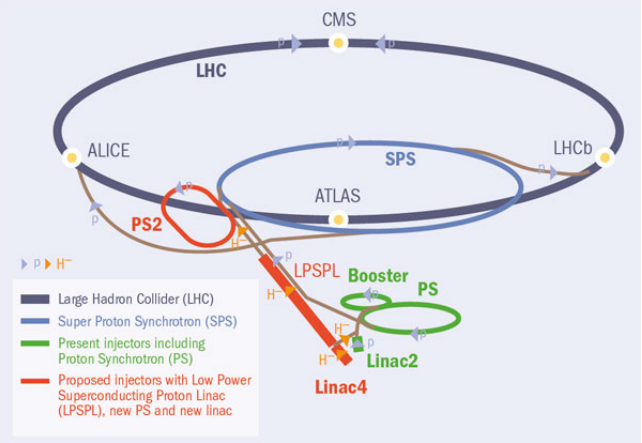

FIGURE 2.1.1: Cern accelerating system. [12]

2.1.1 Luminosity

One of the most important variables for accelerators is the luminosity of the beam. It is the number of particles that can fit in a given space for a given period of time. It is a quantity that tells us the number of interactions that can occur if all the particles of the two beams collide with each other. Interaction rate can be defined as the luminosity multiplied by the cross-section. That means that the rate depends on the possibility of the interaction to happen and on the number of particles that will be located at the collision point. As a result, if we want to increase the interaction rate, we can achieve this by increasing the luminosity of the accelerator. The current luminosity of LHC is $10^{34} cm^{-2} s^{-1}$.
Luminosity is defined as follows:

$$L = fn\frac{N_1 N_2}{A} \tag{2.1.1}$$

- N_1 and N_2 : The number of particles in each packet.
- A: Cross-sectional area of the beam.
- f: Frequency of the beam.
- n: The number of the packets.

Another important quantity is the integrated luminosity (over time), which specifies the number of the collected dataset.

$$L_{int} = \int L(t)dt \tag{2.1.2}$$

Furthermore, the integrated luminosity is a more useful quantity than instantaneous luminosity because if we multiply it with the cross section of the desire interaction we can compute the expected number of events.

$$N = \sigma L_{int} \tag{2.1.3}$$

2.2 Compact Muon Solenoid (CMS)

CMS detector [13] is built for the effective detection of all the stable particles from which all the unstable particles produced in proton-proton collisions can be reconstructed. It consists of different sub-detectors with dedicated purposes, such as identifying the particle tracks or measuring their energy and momentum.

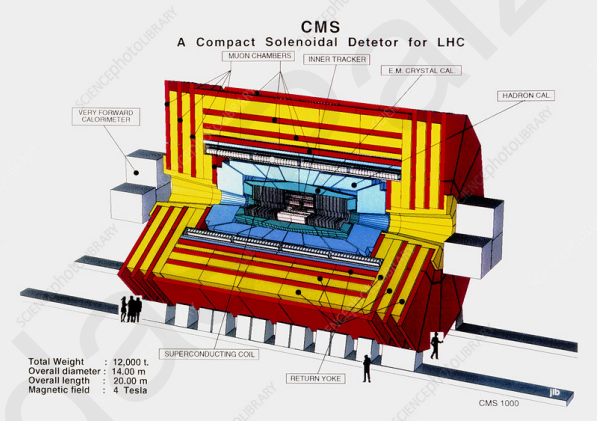

FIGURE 2.2.1: CMS detector [14].

Figures 2.2.1 and 2.2.2 illustrate the different parts of the CMS detector. The tracker, the electromagnetic and hadronic calorimeters, the superconductive magnet and the muon detector.

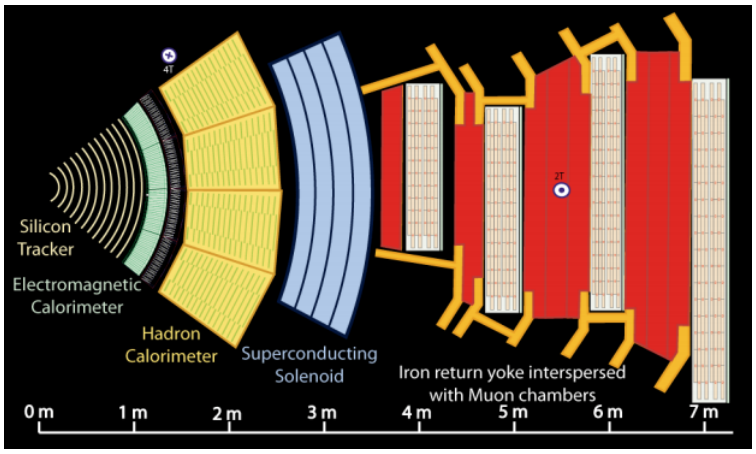

FIGURE 2.2.2: Transverse cut of the CMS detector.

2.2.1 CMS Coordinates System

The CMS detector has a cylindrical shape. Thus, it will be easier to replace the Cartesian coordinates with a different coordinate system. In the new system, the center of the axis is the collision point. The z-axis is parallel to the beam direction, and the x-y plane is replaced from the azimuthal angle Φ that takes values between $-\pi$ and π. To describe all the possible points of the three-dimensional space, we introduce the polar angle θ, which is defined between 0 and π, and it is measured clockwise from the z-axis.

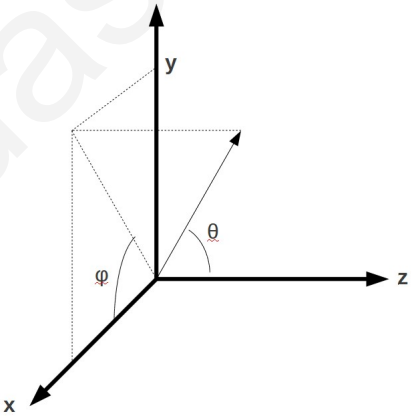

FIGURE 2.2.3: Coordinate convention for CMS detector

In our calculations, instead of the polar angle ϑ, we are use a different quantity known as rapidity (y):

$$y = \frac{1}{2} ln(\frac{E + Pz}{E - Pz}) \tag{2.2.1}$$

For relativistic particles, rapidity is transformed to a more simplified expression called pseudorapididy (η):

$$\eta = -ln(\tan\frac{\theta}{2}) \tag{2.2.2}$$

Another important point is the definition of transverse momentum (P_T) and the angular distance between two particles (ΔR):

$$P_T = \sqrt{(P_x^2 + P_x^2)} \tag{2.2.3}$$

$$\Delta R = \sqrt{(\eta 1 - \eta 2)^2 + (\phi 1 - \phi 2)^2} \tag{2.2.4}$$

2.2.2 Superconducting Magnet

The CMS Superconducting Magnet [15] has a length of 12.5m, a diameter of 6m, and is located around the tracker, the ECAL, and the HCAL. It has a solenoid shape that ensures the production of a strong and uniformly distributed magnetic field. The superconductivity is achieved with the use of liquid helium that cools the magnet down to approximately 4K. Generating a uniform magnetic field of 3.8T, its purpose is to bend the tracks of charged particles, allowing their momentum measurement. From equation 2.2.5, we obtain that as the particle's momentum increases, the less its track is bent. Additionally, the existence of the magnetic field can lead to the identification of the charge of the different particles that travel through the detector, because the tracks originated from particles and anti-particles bend in opposite directions. Finally, it is covered by an iron return yoke to ensure that the effect of the magnetic field will be confined only to the sub-detectors inside the magnet.

$$p(\frac{GeV}{c}) = 0.3B(T)R(m) \tag{2.2.5}$$

Where:

- p = Momentum of the particle.
- B = Magnetic field.
- R = Radius of the particle's track.

2.2.3 Tracker

The tracker [16] is the first detector that the particles will meet after the proton-proton collision. It is the part of CMS in which the tracks of charged particles are reconstructed, their momenta are computed, and their charges are defined. It is a cylindrical detector with a length of 5.2m and a diameter of 2.5m. It covers a region of pseudorapidity less than 2.5.

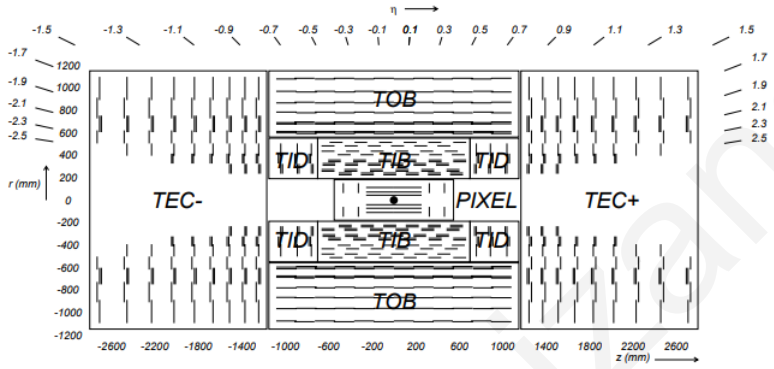

FIGURE 2.2.4: The tracker detector [16].

The space between two proton bunches is 25 ns, and the number of proton pairs that interact is usually more than one. Thus, the amount of particles that are expected to reach the detector is enormous. To deal with this situation and provide sufficient information about the particle path, it must have the needed accuracy and response time. The tracker is made up of two sub-detectors: the silicon pixel and the silicon strips. The first one is located at the center of the detector and consists of three barrel layers and two forward discs. Its pixels can achieve a resolution of 150 $\times$ 100 μm^2. Silicon strips sub-detector is composed of Tracker Inner Barrel (TIB), Tracker Inner Disks (TID), Tracker Outer Barrel (TOB), and Tracker Endcaps (TEC$\pm$). The first two can provide position measurement with a resolution from 13-38 μm and the last two from 18-47 μm. The working principle of the two detectors is based on the electricity production when a charged particle strikes the depletion region of a p-n junction diode. If the particle has kinetic energy at least equal to the energy gap of the semiconductor, it can knock a bound electron out of its bound state in the valence band and excite it to a state in the conduction band. This leads to the generation of electron-hole pairs. The electron that has been produced will be accelerated due to the electric field of the depletion region. The result of the above process is the production of an electric current that can be translated to an electric signal and reveal that a particle has passed through a given pixel.

For the reconstruction of charged particle trajectories, the Combinatorial Track Finder (CTF) algorithm is used. It scans the event repeatedly in order to reconstruct as many

tracks as possible. The idea behind the repeated process is to initially detect the clearest trajectories that possibly belong to particles with high momentum that originated close to the collision point. Afterward, the first identified tracks are removed, making the detection and reconstruction of lesser quality trajectories easier. Each track building includes four steps. The first is seed generation, which is a first estimation using two or three pixels. The next step is the track fitting which includes the propagation of the seed in the detector and the Kalman filter algorithm to find tracks that may be compatible with the trajectory. Afterwards, we have the final track fit that uses all the previously selected hits and obtains a more precise result using the Kalman filter and a smoother algorithm. Finally, the reconstructed trajectories are categorized based on selection criteria that define their quality.

2.2.4 Electromagnetic Calorimeter

Electromagnetic calorimeter [15] is the second detector that the particles meet during their journey. Its main task is to identify electrons and photons, measure their energy, and stop them, preventing their entrance to the next detector.

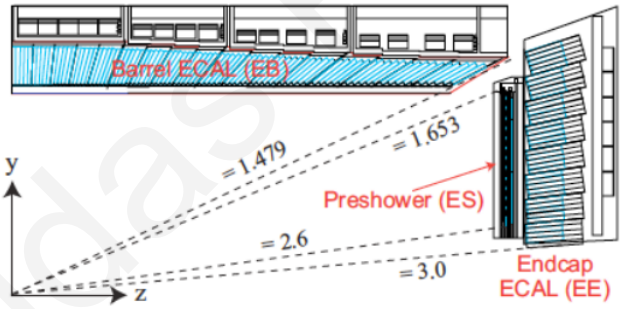

FIGURE 2.2.5: CMS ECAL

As we can see from figure 2.2.5 it is composed by two parts the Barrel ECAL (EB) that covers the region that has $|\eta| < 1.479$ and the Endcap ECAL (EE) that is responsible for larger values of pseudorapididy, specifically for $1.479 < |\eta| < 3.0$.

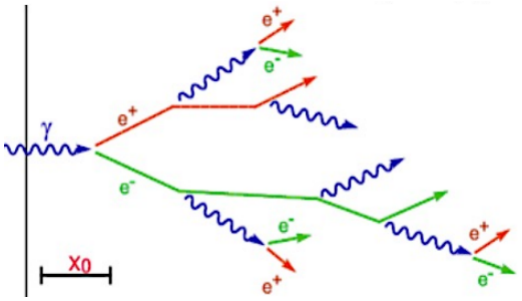

FIGURE 2.2.6: Electromagnetic Shower

To understand how the ECAL is works, we have to study the behavior of electrons and photons as they traverse the material. When an electron enters a medium, it loses energy due to ionization and bremsstrahlung radiation (emitting photons). In the first stage, we can take into account only the energy lost due to bremsstrahlung radiation since the energy lost depends on the particle's energy. In contrast, in the ionization case, the logarithm of electrons energy is the quantity that defines the amount of the deposited energy. A photon that reaches the material of the detector converts into an electron and a positron pair, what we have discussed above is repeated. Finally, another vital quantity is the length of interaction x_0, defined as the distance an electron typically travels until its energy decreases by a factor of 1/e. For photons, this distance is 9/7 that of the electrons. This quantity is unique for each material since it depends on its density.

$$E = E_0 e^{-\frac{x}{x_0}} \tag{2.2.6}$$

The above process is known as an electromagnetic shower. This procedure stops when the particle's energy decreases enough to reach the value of the critical energy. Where the rate of losing energy due to ionization and bremsstrahlung radiation are equal.

ECAL's working principle is based on the scintillation mechanism. There is a material that causes electromagnetic shower, and at the end, there are avalanche photo-diodes. During this process, the produced photons will hit the depletion region of the photo-diodes and generate an electron-hole pair that will lead to an avalanche breakdown and to photo-current production that will be translated to an electric signal. Depending on the produced signal, we can estimate the initial energy of the particle that entered the medium.

The probability density function of the energy lost by electron during its travel in a material, in contrast with other particles, is not a single Gaussian distribution. It is a mixture of various Gaussian distributions. This means that the reconstruction of the electrons can't be performed using the Kalman filter, which is typically used for particles

with Gaussian probability density functions. In our case, the optimal algorithm is the Gaussian-sum filter (GSF) that considers multiple Gaussian distributions.

2.2.5 Hadronic Calorimeter

The next stop of the particles is the HCAL, which is located around ECAL. Its purpose is to measure the energy of hadrons and absorb them completely. It consists by four parts, hadron barrel (HB) for $|\eta| < 1.3$, endcap (HE) for $1.3 < |\eta| < 3.0$, outer (HO) and forward (HF) for $3.0 < |\eta| < 5.0$. The last two are located behind the superconducting magnet.

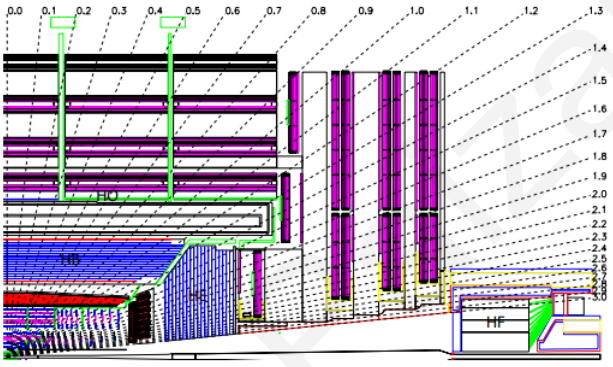

FIGURE 2.2.7: Hadronic Calorimeter of CMS.

The working principle of HCAL is similar to that of the ECAL. The difference between them is that HCAL is a sampling calorimeter. This means that it is not composed of only one material, but instead it is built from two different materials placed alternately. A plastic scintillator causes ionization, and the absorber (brass or steel) generates hadronic shower. The photons produced at each calorimeter layer are transferred to the photo-detectors and are used to measure the particle energy. HF sub-detector, instead of plastic scintillator, has quartz fibers that cause Cerenkov radiation when a particle with sufficient energy enters the material.

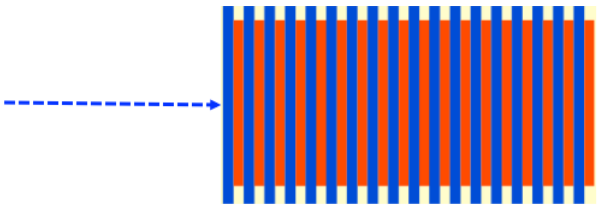

FIGURE 2.2.8: Sampling calorimeter

2.2.6 Muon Detector

Finally, all the particles that didn't get absorbed in the previous detectors face the iron yoke and the muon detector[15][16]. Muon detection is crucial for high energy physics analyses. The fact that muons belong to the same family with the electrons but are much heavier minimizes their interactions with the different materials. It makes them very important for the studies of many decay processes. For example, the best reconstruction of the Higgs boson was achieved through its decay into four muons. The interactions of muons with materials are so weak that they can travel through all the previous detectors and deposit only a small amount of energy. A muon can also pass the iron yoke without losing much energy. That's why their detection occurs far from the detecting point of the other particles. The Muon detector includes three gas sub-detectors separated with iron layers.

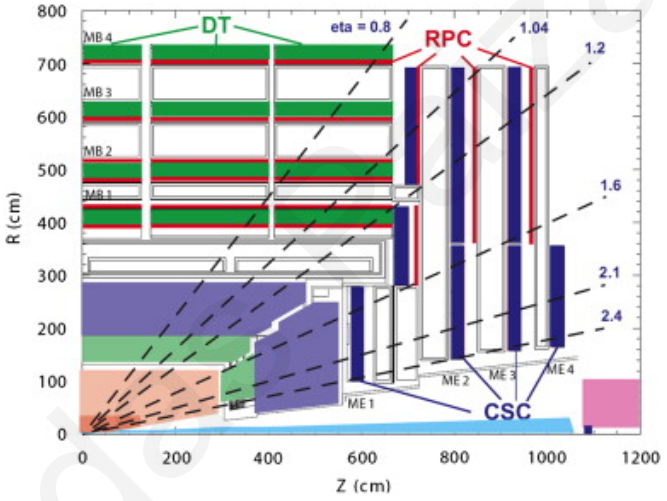

FIGURE 2.2.9: CMS muon detector

The first sub-detector is composed of the Drift Tubes (DT) which cover regions with pseudorapidity with $|\eta| < 1.2$. Next are the Cathode Strip Chambers (CSC) for the region of $0.9 < |\eta| < 2.4$. Finally, we have the Resistive Plate Chambers (RPC) responsible for $|\eta| < 1.6$. The physics behind the three detectors is based on gas ionization, where the electrons that are excited from the gas atoms excite other electrons in a cascade. All these electrons connect the cathodes with the anodes of the electrodes and produce electric signals that mark the trajectory of the muons. RPC is also used as a trigger since it is fast.

The following figure illustrates the trajectory of the different particles during their journey in the CMS detector.

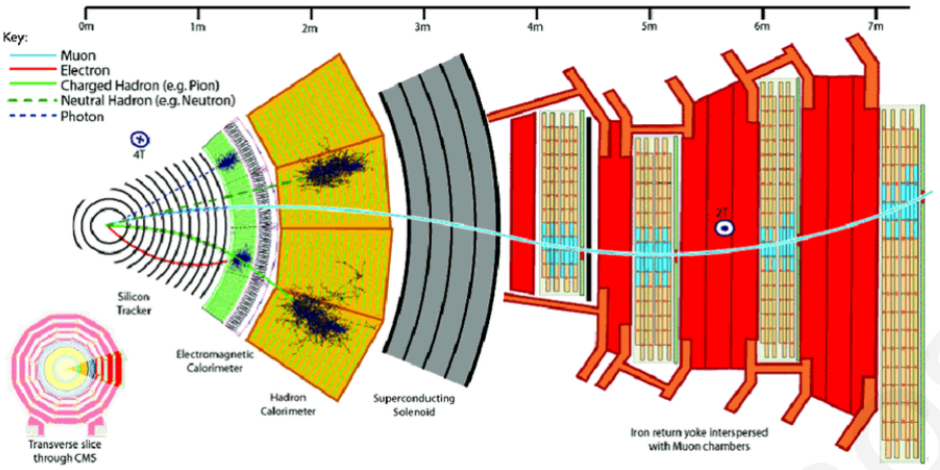

FIGURE 2.2.10: The path for the different particles in CMS detector.

2.2.7 Triggering System

During the collisions, the amount of information that reaches the detectors is enormous and the CMS detector can't collect all of this data. For this reason, we have the CMS real-time trigger system. It includes two steps, the Level-1 trigger (L1) and the High-Level Trigger (HLT). Its purpose is to reduce the amount of collected data by discarding events that may not contain interesting physics. This decision is taken based on some quality criteria of the event and the participating particles.

L1 trigger decreases the amount of collected data from 40 MHz to 100 kHz in a few μs. It consists of different sub-triggers responsible for each reconstructed object (muons, jets, τ_h, electrons and photons, MET and HT). The procedure needs to be fast therefor the L1 trigger can't use the information from the tracker. It only uses what is provided by the calorimeters and the muon detector.

The HLT trigger has access to all the detector data and the L1 results. Using this information and dedicated reconstruction algorithms, it does a more careful scan of the event to cluster deposit energies and form the particle candidates. After the triggering process of HLT, the incoming data rate is reduced to 1 kHz.

Chapter 3

Event Reconstruction and Simulation

In this chapter, we will investigate how the information provided by the detector is deployed for the reconstruction and identification of the physics objects. For this purpose, responsible is the Particle-Flow algorithm, which combines the data collected from all the detectors of CMS.

3.1 Event Reconstruction

3.1.1 Primary Vertices

With approximately 2400 bunches per beam and 10^{11} protons per bunch, the number of collisions in each beam crossing it much larger than one. Thus, the association of all the reconstructed objects with a collision point is important. The primary vertex is the point in the accelerator where the proton-proton collision takes place. The vertex candidate with the largest transverse momentum square ($\sum_i P_{T,i}^2$) of all the participating tracks in its reconstruction is considered as the primary vertex.

3.1.2 Electrons and Photons

The reconstruction process [17] begins with clustering ECAL crystals with energy ($>$ 80 MeV EB and $>$ 300 MeV EE), typically 2-3 times bigger than the electronic noise expected for these types of crystals. Then, after scanning a specific region, the group of crystals with the higher transverse momentum is defined as the seed cluster. For this part of the process, the seed must pass a threshold of 1GeV of transverse momentum.

The reconstruction continues with the "superclustering" procedure (SC). During the SC step, crystals that exist in a specific area around the seed cluster are combined. Afterwards, there is matching procedure to find if there is a trajectory from the pixel detector that can be associated with the cluster. The promising trajectories are then used as input to a dedicated algorithm based on Gaussian Sum Filtering (GSF) for track reconstruction. Parallel to the above procedure, there is a "reverse" check where all the reconstructed trajectories of the event are tested to determine whether they can be matched with an electron trajectory hypothesis. Another algorithm is being used for this reconstruction, known as the Kalman Filter (KF). Selected trajectories with transverse momentum greater than 2 GeV are then used to seed the GSF algorithm. The extracted information is imported to the PF algorithm that creates the particle candidates. At this stage, there is no difference between electron and photon reconstruction. What separates them is that photons, as zero-charged particles, can't interact with the tracker material. Thus, they are electrons with no track.
After electron and photon candidates are reconstructed we can compute some important variables for electron identification.

3.1.2.1 Cut-Based Electron Identification

In order to define the quality of the reconstructed electrons some identification variables are computed [17]. One the most important is the isolation quantity.
Isolation:
The isolation tells us the ratio of the amount of transverse energy stored in a cone of fixed radius around the electron to that of the electron. Specifically, this quantity is defined as:

$$Isolation = \frac{I_{Combine}}{ET_{Electron}} \tag{3.1.1}$$

$$I_{Combine} = I_{CH} + max(0, I_n + I_\gamma - I_{PU}) \tag{3.1.2}$$

Where I_{CH}, I_n and I_γ is the sum of the transverse energy of charged hadrons (I_{CH}) photons (I_n) and neutral hadrons (I_γ) inside an isolation cone of radius $\Delta R = 0.3$.
I_{PU} is the correction related to the event pileup.

The additional identification variables are:
Hadronic over Electromagnetic energy ratio (H/E):
This quantity is the energy deposited in the HCAL in a cone of radius $\Delta R = 0.15$ around the electron candidate over the energy of the electron or photon. Energy in HCAL is

expected due to detector noise, pileup contributions, and electrons or photons that failed to stop in the ECAL.

The second moment of the log-weighted distribution of crystal energies in η ($\sigma_{i\eta i\eta}$):

This variable is defined as:

$$\sigma_{i\eta i\eta} = \sqrt{\frac{\sum_i^{5\times5} w_i(\eta i - \overline{\eta}_{5\times5})}{\sum_i^{5\times5} w_i}} \tag{3.1.3}$$

where,

$$w_i = max(0, 4.7 + ln(\frac{E_i}{E_{5\times5}})). \tag{3.1.4}$$

Looking at equation 3.1.4 we can see that if E_i is less than 0.9% of $E_{5\times5}$, the value of w_i is set to zero. This quantity ensures that the electron or photon candidate isn't affected by ECAL noise.

Additional variables:

- $|\frac{1}{E} - \frac{1}{p}|$, where E is the energy of SC candidate and p is the track momentum at the point closest to the vertex.
- $|\Delta\eta_{in}^{seed} = |\eta_{seed} - \eta_{track}|$
- $|\Delta\Phi_{in}^{seed} = |\Phi_{seed} - \Phi_{track}|$
- Number of missing hits at the pixel detector.
- Pass conversion veto : Association between the SC cluster and a reconstructed track.

Depending on different cuts on these variables, various workings points are defined.
The four working points that are currently being used in CMS are the following:

- Veto: Corresponds to 95% signal efficiency and it's used in cases where the final state doesn't include electrons.
- Loose: Correspond to 90% signal efficiency and it's used in analyses where the background contributions due to misidentified electrons are low.
- Medium: Corresponds to 80% signal efficiency and it's suitable for W and Z bosons measurements.
- Tight: Corresponds to 70% signal efficiency and it is used to analyses where the background contributions due to misidentified electrons are high.

The following table illustrates the selections for the Cut-based electron identification, for the tight working point in the barrel and in the endcaps.

Variable	Barrel (tight WP)	Endcaps (tight WP)
$\sigma_{i\eta i\eta}$	<0.010	<0.035
$\lvert\Delta\eta_{\text{in}}^{\text{seed}}\rvert$	<0.0025	<0.005
$\lvert\Delta\phi_{\text{in}}\rvert$	<0.022 rad	<0.024 rad
H/E	$<0.026 + 1.15\,\text{GeV}/E_{\text{SC}}$ $+0.032\rho/E_{\text{SC}}$	$<0.019 + 2.06\,\text{GeV}/E_{\text{SC}}$ $+0.183\rho/E_{\text{SC}}$
$I_{\text{combined}}/E_{\text{T}}$	$<0.029 + 0.51\,\text{GeV}/E_{\text{T}}$	$<0.0445 + 0.963\,\text{GeV}/E_{\text{T}}$
$\lvert 1/E - 1/p\rvert$	$<0.16\,\text{GeV}^{-1}$	$<0.0197\,\text{GeV}^{-1}$
Number of missing hits	≤ 1	≤ 1
Pass conversion veto	Yes	Yes

3.1.3 Muons

Muons [18] are the only charged particles that traverse all the detectors, leaving signs in the tracker and the muon chamber. The process starts with the independent track reconstruction in the tracker (tracker track) and the muon system (standalone muon track). The final purpose is the full muon track reconstruction, and three different procedures are used.

Standalone Muons:

These muon candidates are reconstructed from the muon system information. An algorithm based on the Kalman-filter technique gathers hits from CSC, DT, and RPC to finally built muon trajectories.

Tracker Muons:

This method starts from the tracker, where all the reconstructed tracks with transverse momentum larger than 0.5 GeV and total momentum larger than 2.5 GeV are extrapolated up to the muon system. If the tracker track is matched with at least one segment in the muon system, then it is defined as a tracker muon track.

Global Muons:

This procedure combines the previous two. Global muon collections consist of trajectories built from matching the standalone tracks and the tracker tracks. Kalman-Filter algorithm is used for the combined fit.

In cases where two or all methods reconstruct the same track, the trajectories are merged into one candidate.

As for electrons, there are also some identification variables such as the track fit χ^2, the number of hits per track and the quality of matching between trackers, and compatibility with the primary vertex and standalone tracks. The last variable is being used only for global muons.

The different selections that we can obtain depending on the values of the identification variables and their combinations are the following:

- Loose: This collection targets muons originating from the primary vertex and from light or heavy flavor decays. Additionally, the selections applied for the loose muons collection decrease the possibility of the candidate being a misidentified charged hadron.
- Medium: A loose muon with a tracker track and at least 80% hits in the inner tracker. If the reconstructed muon has only a tracker track, the required muon segment compatibility must be bigger than 0.451, and if it has been also reconstructed as a global muon, the previous value decreases to 0.303.
- Tight: A loose muon that has fired at least six tracker layers, including a pixel hit, and its track belong to both tracker and global tracks. Additional selections about χ^2 and compatibility with the primary vertex are also included. This collection aims to reject muons from hadronic punch through and from decaying in flight.
- Soft: Tracker muons with low transverse momentum that as the tight muons have fired at least six tracker layers including a pixel hit.
- High Momentum: Muons that have been reconstructed as tracker and global muons with transverse momentum greater than 200 GeV.

In contrast with electrons, the isolation quantity for muons is not included in the identification variables. Thus, if the analysis requires isolated muons, we must add this additional selection. Muon isolation is computed in the same way as electron isolation (equations 3.2.1 and 3.2.2), but the cone radius is 0.4.
There are two working points, tight (isolation < 0.15) and loose(isolation < 0.25) with 95% and 98% signal efficiency respectively.

3.1.4 Hadrons And Non Isolated Photons

The last task for object identification is the detection of neutral hadrons and non-isolated photons. Once the electrons, isolated photons, and muons are reconstructed and removed from the PF blocks, the remaining deposit energy at the ECAL and HCAL is expected to originate from the non-identified particles. ECAL clusters are identified as photons and HCAL as hadrons. The presence or absence of a track associated with the particle candidate determines whether the particle is charged or neutral. ECAL clusters linked to a given HCAL cluster are assumed to arise from the same hadron shower.

3.1.5 Jets

Jets are a group of hadrons produced from a high-energy quark or gluon decay. The short lifetime of quarks and gluons makes the possibility of reaching the detectors impossible. Thus, what remains from quarks and gluons is the group of the produced hadrons that travel in eliminated bunches of particles "Jets". The challenging part is the definition of a clustering process that will transform a set of hadrons into individual jets. The two methods that can be applied for this task are the recombination algorithms and the cone algorithms. The idea behind the first procedure is the repeated combination of the two particles that are closest to each other. This process continues until a specific criterion is reached. On the other hand, the second algorithm has a cone of fixed radius, and the jet is the clustering of the particles in that cone. The algorithm currently being used in CMS is the anti-K_T algorithm. A sequential recombination algorithm that defines as follows:

$$d_{ij} = min(\frac{1}{P_{ti}^2}, \frac{1}{P_{tj}^2}) * \frac{\Delta R_{ij}^2}{R^2} \tag{3.1.5}$$

$$\Delta R_{ij}^2 = (y_i - y_j)^2 + (\Phi_i - \Phi_j)^2 \tag{3.1.6}$$

$$d_{iB} = \frac{1}{P_{ti}^2} \tag{3.1.7}$$

After the calculation of the d_{ij} and d_{iB} variables for two particles i and j, if the smallest one is d_{ij} then the two particles are removed from the list with the available hadrons and merge to one particle that is the sum of their four-momenta. This process is repeated until no particles are left. Finally, R is the jet radius parameter. Two of the most used values of this parameter are R = 0.4, which leads to the formation of AK4 jets, and R = 0.8, which has to do with another jet collection, AK8 jets.

3.1.6 B-Tagged Jets

Identifying the quark that produced a jet is very important because it gives us additional information about a process. Without this categorization, a process with jets originated from light-flavored quarks, and a process that has jets from b quarks would be the same. Heavy flavor jets contain a bottom or a charm hadron. In these cases, the quark transfers a large part of its momentum to the hadron that has a lifetime of around 1ps. As a result, it travels a distance of some mm before its decay, creating a secondary vertex displaced from the PV.

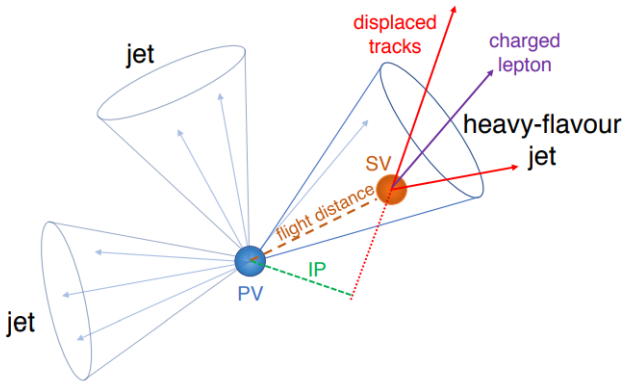

FIGURE 3.1.1: Schematic representation of a heavy-flavor jet being produced at the secondary vertex.

This displacement is the main difference between light-flavored and heavy-flavored jets. There are dedicated b-tagging algorithms for this kind of identification for better results. The three algorithms currently being used at CMS are CSV_{v2}, DeepCSV, and DeepJet. DeepJet is the newest b-tagger, and it was chosen for this analysis. It's a DNN-based algorithm with 650 input variables. There are three different working points for the DeepJet algorithm. The following table illustrates the different B-tagging working points of the DNN score, based on the misidentification rate:

Working Point	Mistag Efficiency	DeepJet		
		2016	2017	2018
loose	10%	0.0614	0.0521	0.0494
medium	1%	0.3093	0.3033	0.2770
tight	0.1%	0.7221	0.7489	0.7264

3.1.7 Hadronically decaying τ leptons

Tau is a third generation lepton with mass 1.78GeV and life time 290.3*10^{-15} s. Due to it's mass, it is the only lepton that can decay hadronically.

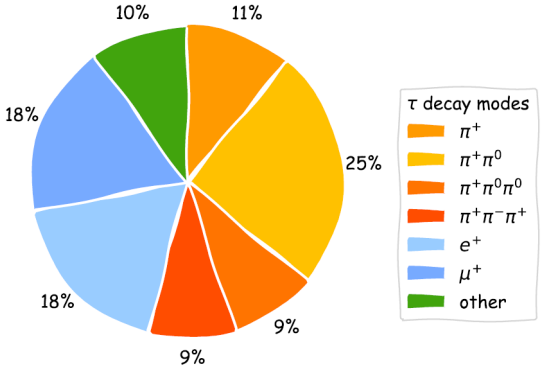

FIGURE 3.1.2: Tau decay modes

As we can see from the pie-chart, tau lepton decays hadronically around 64% of the time, and the rest 36% produce an electron or muon. Unfortunately, we can't distinguish the leptons made from a tau from other leptons. Thus, the only way to learn about its existence is from its hadronic products. Usually, tau decays into one or three charged pions and up to two neutral pions. These are vital for its detection because hadronically decaying taus form jets with specific properties different from those of other jets. The jets originating from tau leptons are reconstructed by the Hadrons-Plus-Strips (HPS) algorithm that uses the anti-K_T clustering method combined with PF information for hadron identification since, as we have discussed, the majority of tau decays contain a fixed number of charged and neutral hadrons. A dedicated tau identification algorithm is used to increase the discrimination efficiency between tau jets and jets originating from quarks or gluons. The algorithm is known as DeepTauV2.1. It's based on a convolutional deep neural network that uses information from different subdetectors.

3.1.8 Hadronically decaying top quarks

Moderately boosted top quarks participate in many decays of BSM particles. This makes their identification very important because it can significantly assist in discriminating signal versus background contributions. For this task, we use a resolved top tagger, a classifier based on a multivariate deep neural network. During the top quark decay, a W boson and a b quark are produced. This algorithm selects hadronically decaying top quarks. Thus, the W is expected to provide us with two additional quarks. The final objects are one b-tagged jet and two jets. Resolved top tagger is trained to distinguish combinations of three jets that belong to the top quark decay from tri-jets originating from the background process. For efficient identification, variables associated with top and jets properties are used as an input to the DNN. Distributions with high discrimination between signal and background process are dangerous for machine learning analysis because they bias the algorithm's output. Top quark mass belongs to this type

of variable. It's a characteristic distribution and leads to modification of the background mass distribution that imitates the signal mass distribution. To prevent this, there are mass decorrelated methods that leave the algorithm's output unaffected by the top quark mass. distribution. Depending on signal and background efficiencies and the b-tagged discriminator that is used, there are three different working points (for each b-tagged discriminator). The different working points are illustrated in the following table:

	WP	Signal eff.	Bkg eff.
Loose	0.4055	0.90	0.10
Medium	0.6035	0.78	0.05
Tight	0.8755	0.43	0.01

	WP	Signal eff.	Bkg eff.
Loose	0.3865	0.95	0.10
Medium	0.5985	0.87	0.05
Tight	0.8725	0.54	0.01

Top tagging working points and corresponding signal and background efficiency using the DeepCSV (left) and DeepFlavour (right) b jet discriminators

3.1.9 Missing Transverse Energy

During the acceleration process, the proton beams move along the z-axis. Thus, we expect that the vectorial sum of transverse momenta of all the detected particles' will be zero due to the conservation of momentum. This is violated by the presence of neutrinos since, if they are produced, they escape the detector undetected. As a result, this negative vectorial sum gives an estimation of the total transverse momentum of neutrinos in the event and is known as the Missing Transverse Energy (MET)

$$MET = -\sum_{i} ET(i) \tag{3.1.8}$$

3.2 Event Simulation

We use Monte Carlo simulations to study events produced in proton-proton collisions. Depending on the provided information, one can do detailed studies. Starting from a simple branching ratio of a particle's decay, we can add other piece of information such as the interaction between them, the parton shower, the hadronization process, etc. The simulation of a process in high-energy physics is critical because it provides us with details about the process that we can't obtain during an actual proton-proton collision. Finally, it is also possible for us to study and compare different processes and the behavior of various undetected particles.

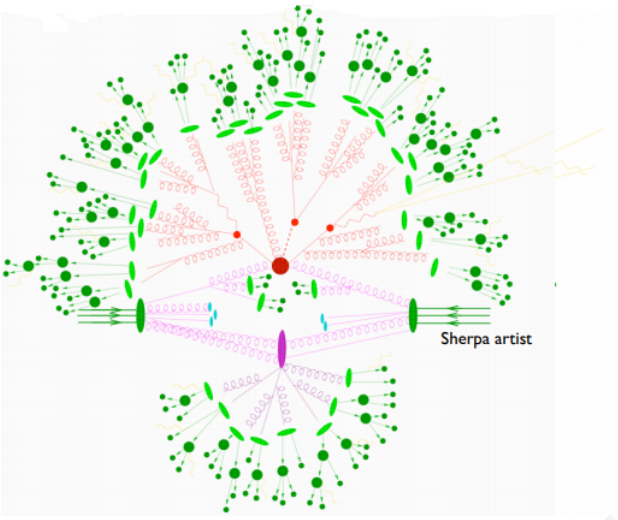

FIGURE 3.2.1: Monte Carlo Simulation

When a proton has zero kinetic energy, its energy is equally distributed to its three quarks. But when it is accelerated in such energies, the distribution of its energy changes. The gluons start to share also an amount of proton's momentum and decay to a quark and anti-quark pair. The amount of energy that each one of the partons carry must be taken into account in Monte Carlo simulations.

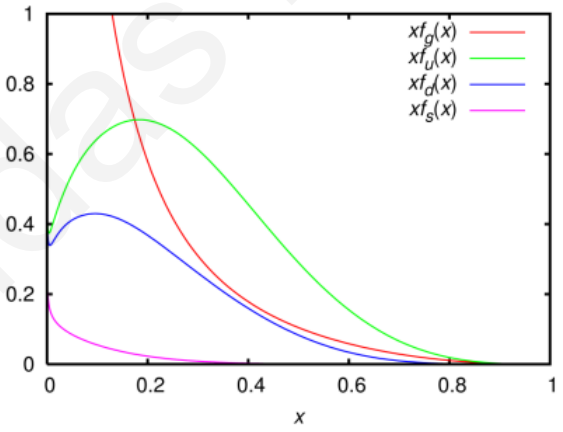

FIGURE 3.2.2: The amount of energy that each one of the partons carry while the proton's energy increase

Finally, the PYTHIA generator is responsible for other processes such as the pile-up interactions, the parton showers, and the hadronization of baryons and mesons.

Chapter 4

Search for Charged Higgs boson decaying to a top and a bottom quark in lepton plus jets final state

In the previous chapter, we have discussed the detection of final state objects. Using this information, we will search for the production of a charged Higgs boson with mass heavier than the top quark through its decays to a top and a bottom quark in the lepton plus jets final states.

This search is based on data collected by the CMS detector during the 2017 LHC collider run and correspond to 41.5 fb^{-1} of integrated luminosity. For signal we have used eight different masses of charged Higgs (200, 300, 400, 500, 800, 1000, 2000 and 3000 GeV). But all the analysis strategy is based on the 200 GeV charged Higgs, a challenging case due to the small difference between the mass of the charged Higgs and the top quark (173 GeV), and the fact that for the current final state $t\bar{t}$ is the dominant background process. This makes the discrimination between signal and background processes even more difficult (Chapter 5).

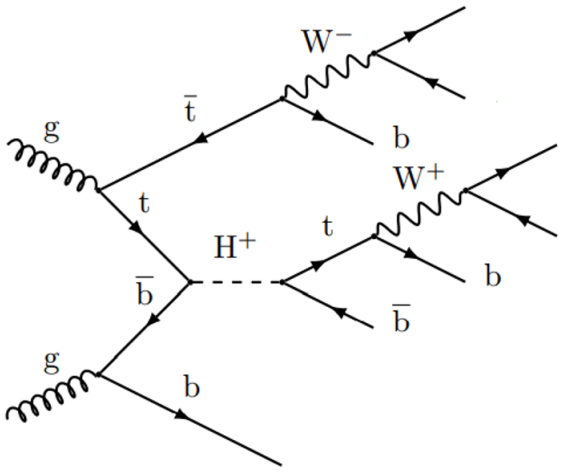

FIGURE 4.0.1: Leading order Feynman diagram of a charged Higgs boson produced in association with a top quark and decaying into a top and bottom quark-antiquark pair.

The two W bosons that are produced in the process define the possible final states and the associated branching ratios. Lepton (e/μ) plus jets is the second most common final state with branching ratio of approximately 30%.

	e	μ	τ	Hadrons
e	1/81	1/81	1/81	6/81
μ	1/81	1/81	1/81	6/81
τ	1/81	1/81	1/81	6/81
Hadrons	6/81	6/81	6/81	36/81

W^+ DECAY MODES		Fraction (Γ_i/Γ)
$\ell^+\nu$	[b]	(10.86± 0.09) %
$e^+\nu$		(10.71± 0.16) %
$\mu^+\nu$		(10.63± 0.15) %
$\tau^+\nu$		(11.38± 0.21) %
hadrons		(67.41± 0.27) %

FIGURE 4.0.2: W boson decay modes [19]

4.1 Trigger and Characteristics of Final State Objects

At this stage of our analysis, we need to apply some filtering based on the quality of the event and the characteristics of the identified objects, increasing the possibility of keeping events that "hide" the final state of the process that we want to study.

4.1.1 Trigger

The trigger does the first cleanup. It is necessary to choose events that have fired a trigger to discard the ones with "bad quality" based on some criteria. Since this analysis includes a lepton, we required the event to pass electron or muon High-Level Trigger (HLT-Ele35-WPTight-Gsf or HLT-IsoMu27). This is achieved if the event contains at least one tight isolated electron with P_T greater than 35 GeV and Eta less than 2.1, or at least one tight isolated muon with P_T greater than 27 GeV and Eta less than 2.4.

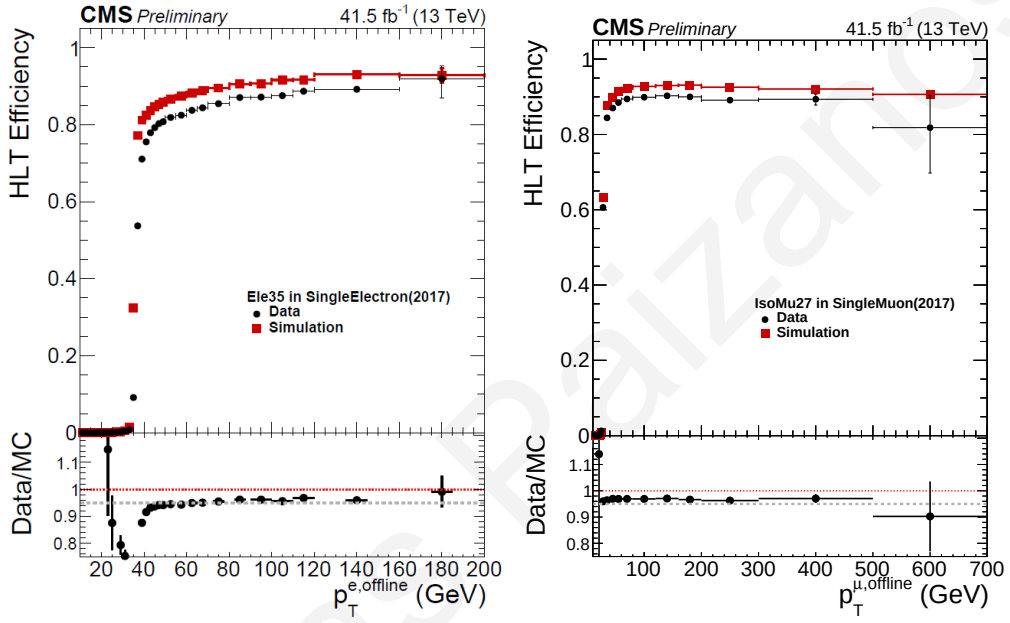

FIGURE 4.1.1: Efficiency of the HLT-Ele35-WPTight-Gsf (left) and HLT-IsoMu27 (right) triggers as a function of the electron's pT, using the 2017 datasets

From figures 4.1.1, we obtain that to reach the plateau of the trigger efficiency the P_T (transverse momenta) thresholds should be above 38 GeV and 28 GeV for electrons and muons, respectively.

4.1.2 MET Filters

The final state that we want to study is characterized by the production of a lepton and the associate neutrino. Thus, expecting the production of a neutrino, we can set the initial requirement of MET to be greater than 20 GeV.

4.1.3 Primary Vertex

As we have seen in Chapter 3, using the tracks of all the detected charged particles, we can reconstruct the Primary Vertex, the point where the proton-proton collision occurred. If there isn't a primary vertex or it's displaced from the beams axis, we have to skip that event.

4.1.4 Electrons

Selections	Tight Electron Base	Electron Veto
HLT-Ele32-WPTight-Gsf	√	-
$p_T >$	38 GeV	10 GeV
$\|\eta\| <$	2.1	2.4
Isolation:	Tight	Loose
ID :	tight	veto

ID: cutBasedElectronID-Fall17-94X-V2

4.1.5 Muons

Selections	Tight Muon Base	Muon Veto
HLT-IsoMu27	√	-
$p_T >$	28 GeV	10 GeV
$\|\eta\| <$	2.4	2.4
Isolation:	Tight	Loose
ID:	isCutBasedIDTight	isCutBasedIDLoose

4.1.6 Tau Veto

Selections	Taus
Trigger Matching	-
$p_T >$	20 GeV
$\|\eta\| <$	2.3
Isolation Discrimination:	byMediumDeepTau2017v2p1VSjet
Isolation Raw Discrimination:	byDeepTau2017v2p1VSjetraw
Against Electrons:	byMediumDeepTau2017v2p1VSe
Against Muons:	byTightDeepTau2017v2p1VSmu

4.1.7 Jets and BJets

Selections	Jets	BJets
$p_T >$	30 GeV	30 GeV
$\|\eta\| <$	2.4	2.4
Jet ID:	Tight	Tight
BJet Discrimination	-	pfDeepFlavourBJetTags
BJet Discrimination WP	-	Medium

4.2 Initial Selections

Ideally, the result of lepton plus jets final state for this process would be one lepton, two jets, four bjets, and MET from the neutrino coming from the W that decays leptonically. Unfortunately, things are slightly different since the detection of the final state products is not always successful.

We select events that have passed the single-lepton HLT trigger. Thus, it's reasonable to start the selection by requiring exactly one tight isolated lepton that is HLT matched (electron or muon) depending on which one of the two HLT triggers was fired.
The final state that interests us has only one lepton. For this reason, events with loose leptons or tight opposite flavor leptons (from the selected one) should be skipped. The last check that has to be done to ensure that we selected a "good quality" lepton is about the distance between the lepton and the jets of the event. Sometimes during the reconstruction process, the lepton can be found inside a jet cone, which will cause problems in the analysis because it means that the lepton and the jet are the same objects. To prevent this, we require leptons and each jet ΔR to be greater than 0.4. Furthermore the final state does not include any taus so in our selections is added that the numbers of taus in the event should be zero. The last selection to be made is the number of jets and bjets. We expect six jets, that four of them will be bjets. The following plots show the jet multiplicity after the selection that we applied up to this point.

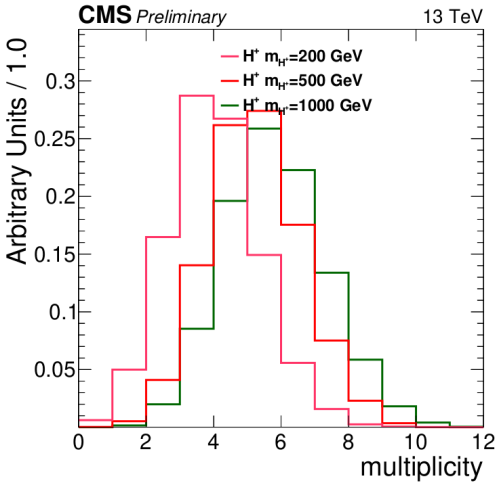

FIGURE 4.2.1: Jet multiplicity for Muon final state for three different masses of Charged Higgs

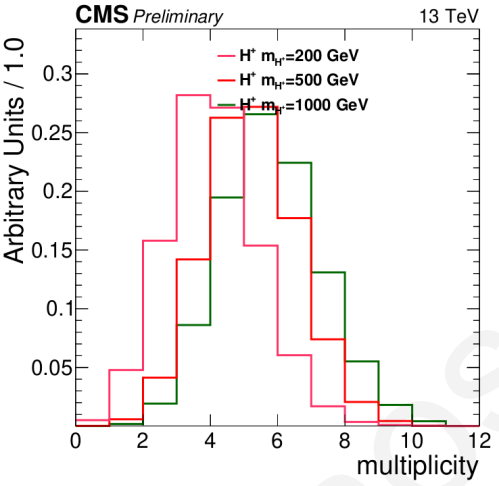

FIGURE 4.2.2: Jet multiplicity for Electron final state for three different masses of Charged Higgs

From figures 4.2.1 and 4.2.2, we can observe three things. First, as expected, the jet multiplicity behaves similarly for both electron and muon final states. Secondly, the number of jets in the event increases as the charged Higgs mass increases. This happens because when a particle decays, depending on its mass, it gives its products more momentum making them more easily detected. Finally, despite expecting exactly six jets, the result is a distribution from 0 to 10 jets due to initial and final state radiation or unsuccessful reconstruction. Looking at the jet multiplicity, the optimal cut is at four jets, trying to balance the low mass and high mass charged Higgs cases.

The following plots show the bjet multiplicity after the selection that we applied up to this point (including at least 4 jet selection).

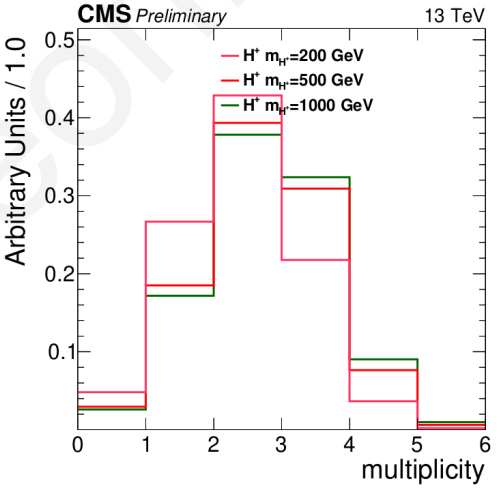

FIGURE 4.2.3: bjet multiplicity for Muon final state for three different masses of Charged Higgs

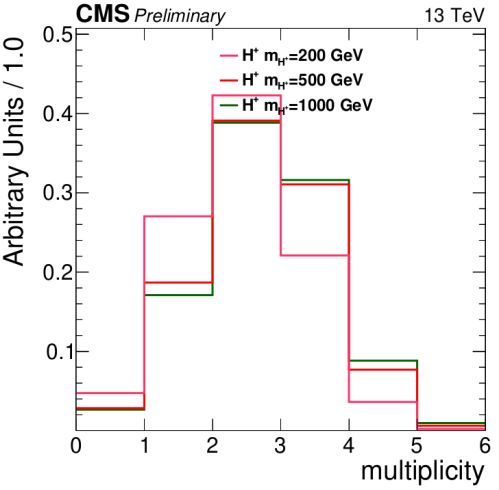

FIGURE 4.2.4: bjet multiplicity for Electron final state for three different masses of Charged Higgs

BJet multiplicity is also distributed between 0 and 5 bjets while the expected number was precisely four. Again the reason for this fluctuation is what we have discussed in the previous paragraph. An example is that the bjet that is produced in association with the charged Higgs usually has very low P_T. As a result, it fails the P_T cut of 30 GeV and remains undetected. Having in mind that our final state has four bjets and looking at their multiplicity distribution, we decided to require at least two bjets.

The next table shows all the event selection criteria and the associate efficiency for muon and electron plus jets final states (MuJets/EJets).

Selections	MuJets	EJets
All Events	100%	100%
Pass HLT Trigger	100%	100%
MET Filter	100%	100%
Primary Vertex	100%	100%
Exactly One Muon/Electron	45%	30%
Veto Loose Muons/Electrons	43%	27%
Lepton-Jets $\Delta R > 0.4$	41%	22%
Veto Electrons/Muons	36%	19%
Veto Taus	34%	18%
At Least Four Jets	17%	10%
At Least Two BJets	13%	7%

4.3 Object Reconstruction

In this section, we will take our analysis a step further by tagging more particles that participate in our process using the final state products that we have selected. This is essential because a successful detection of these particles can be used to discriminate signal from background processes.

4.3.1 Hadronicaly Decaying Objects

Top:

- For reconstructing the hadronically decaying top quark, we use the collection of the candidates provided from the resolve top tagger classifier. Afterward, from the reconstructed trijets that pass the tagger's medium working point, we chose the one with a mass closer to the top quark mass(172.8 GeV).

W:

- Hadronic W is the Di-Jet of the hadronically decaying top.
- For events with no hadronically decaying top the W is reconstructed from the Di-Jet with the mass closer to the W mass(80.4 GeV)

BJet:

- Hadronic bjet is the bjet of the hadronic top.
- For events with no hadronically decaying top, the bjet is not reconstructed.

A cleanup procedure is enforced so objects used in the reconstruction of top quark decays or W or Bjets are not used. For example, the bjet from hadronically decaying objects can't be Higgs bjet.

4.3.2 Missing Energy

The MET vector is a 2D vector. Thus, we can compute its x and y components Px and Py. So we need to determine the z component to reconstruct the missing energy of the event entirely. To do this, we are going to solve the equation of W bosons mass:

$$m_W^2 = (E_{lepton} + E_\nu)^2 - (\vec{P}_{lepton} + \vec{P}_\nu)^2 \tag{4.3.1}$$

- $m_W = 80.4 GeV$
- $E_\nu = \sqrt{P_x^2 + P_y^2 + P_z^2}$
- P_x, P_y are the two components of Met

With this information, we can solve equation 4.3.1 and compute the z component of MET, Pz. We have a quadratic equation, so depending on the value of the discriminant, we can have either one, two, or no real solutions. In the first case, the process is straightforward, the discriminant is zero, and the result it's just the double root. We have to decide what strategy we will follow for the two other cases.

- **Two Solutions:** Since we have more than one solutions, we need to choose one of them. After some studies we figure out that approximately in 80% of the cases the **smallest** P_z in absolute value is closer to the generator level P_z.

- **No Solutions:** In this situation, the discriminant is negative. The first and easiest thing that someone can do is skip this type of events. But before, one must wonder how many events we will lose and whether this will improve our results. After investigation, it turns out that 30% of the events the discriminant was negative, and losing one-third of the total events for this reconstruction wasn't the optimal solution. So we tried to find another method to solve this problem. The decision was to force the discriminant to be zero and obtain one solution. Instead of stopping here to improve this method, we assumed that if the MET (P_x and P_y) that we have used in the equation 4.3.1 gave us a negative discriminant, it might need a correction. So for this type of event, we calculate MET again for the discriminant to be zero. In this case, we will also get two solutions, and we will keep the one closer to the MET before our correction.

4.3.3 Leptonicaly Decaying Objects

After tagging the final state products of the leptonic decay (lepton and missing energy) we can now use them for the reconstruction of other particles of the event.
W: The W boson is decaying to a lepton plus a neutrino in this final state. For this reason, it is reconstructed as the sum of the two Lorentz vectors (lepton and missing energy).
BJet: From the available jets we choose the jet that if it's summed up with the leptonic W, their mass it's closer to the top's mass. This jet it is also removed from the available jets of the event.
Top: The sum of leptonic W and leptonic bjet.

Score for the Leptonic Top:
As we have seen, for the reconstruction of the hadronically decaying top, there is a DNN based top-tagger. Thus, we can estimate the possibility of the reconstructed object being a top. For the leptonic top, we don't have this information. We use the χ^2 method for the top's mass to check how well we reconstructed the top.

- After fitting a Gaussian function to the top mass distribution knowing that all the products of top and top itself are matched to generator particles, we can compute the standard deviation (σ).

- $$x^2 = (\frac{M_{Reconstructed} - M_{Top}}{\sigma})^2 \tag{4.3.2}$$

- x^2 is the score that tell us how well we have reconstruct the leptonic top based on it's mass.

4.3.4 BJet From Higgs

From the remaining available jets that are not participating in the top quark reconstruction, we choose the one with the highest MVA score. This jet will be the bjet coming from Higgs decay.

Chapter 5

Signal and Background Discrimination

5.1 Background Process

After a proton-proton collision, our process won't be the only one produced and detected. There are a lot processes that share the same final state. We present here the different process that we expect to find from lepton plus jets detection.

- Diboson production

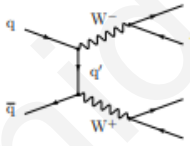

FIGURE 5.1.1: Diboson (WW)

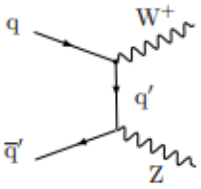

FIGURE 5.1.2: Diboson (WZ)

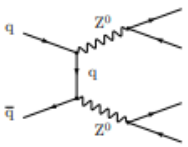

FIGURE 5.1.3: Diboson (ZZ)

- More than one top quark production (TT)

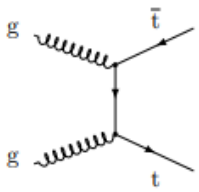

FIGURE 5.1.4: $t\bar{t}$

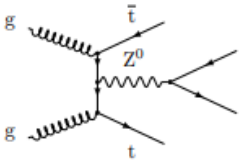

FIGURE 5.1.5: $t\bar{t} + X$

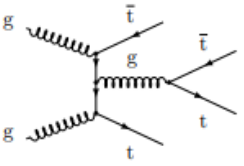

FIGURE 5.1.6: $t\bar{t}t\bar{t}$

- Drell Yan

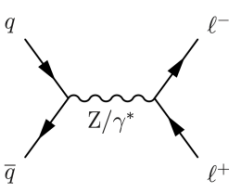

FIGURE 5.1.7: Drell Yan (leptons)

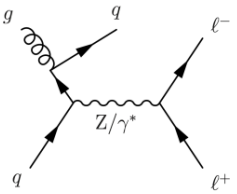

FIGURE 5.1.8: Diboson (leptons + Jets)

- Single Top

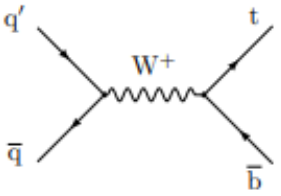

FIGURE 5.1.9: Single Top (4FS)

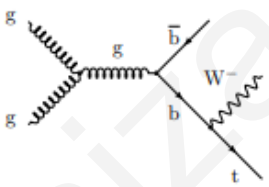

FIGURE 5.1.10: Single Top (5FS)

- W + Jets production

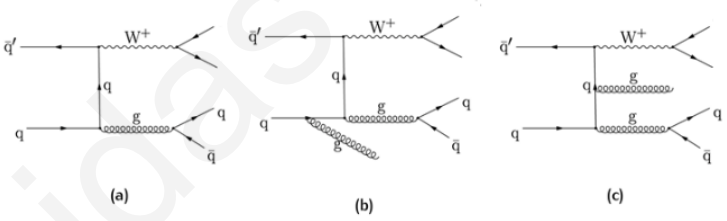

FIGURE 5.1.11: W + Jets

The following table contains a summary of the signal and background samples used for this analysis:

Category	Sample
Signal:	ChargedHiggs-HplusTB-HplusToTBLep-M-200
	ChargedHiggs-HplusTB-HplusToTBLep-M-300
	ChargedHiggs-HplusTB-HplusToTBLep-M-400
	ChargedHiggs-HplusTB-HplusToTBLep-M-500
	ChargedHiggs-HplusTB-HplusToTBLep-M-800
	ChargedHiggs-HplusTB-HplusToTBLep-M-1000
	ChargedHiggs-HplusTB-HplusToTBLep-M-2000
	ChargedHiggs-HplusTB-HplusToTBLep-M-3000

Category	Sample
TT	TTToSemiLeptonic
	TTToHadronic
	TTTo2L2Nu
Single Top (ST)	ST-s-channel-4f-hadronicDecays
	ST-t-channel-antitop-4f-InclusiveDecays
	ST-t-Channel-top-4f-InclusiveDecays
	ST-tW-antitop-5f-inclusiveDecays
	ST-tW-top-5f-inclusiveDecays
Electroweak (EW)	DYJetsToLL-M-50-HT-70to100
	DYJetsToLL-M-50-HT-100to200
	DYJetsToLL-M-50-HT-100to200-ext1
	DYJetsToLL-M-50-HT-200to400
	DYJetsToLL-M-50-HT-200to400-ext1
	DYJetsToLL-M-50-HT-400to600
	DYJetsToLL-M-50-HT-400to600-ext1
	DYJetsToLL-M-50-HT-600to800
	DYJetsToLL-M-50-HT-800to1200
	DYJetsToLL-M-50-HT-1200to2500
	DYJetsToLL-M-50-HT-2500toInf
	WJetsToLNu-HT-70To100
	WJetsToLNu-HT-100To200
	WJetsToLNu-HT-200To400
	WJetsToLNu-HT-400To600
	WJetsToLNu-HT-600To800
	WJetsToLNu-HT-800To1200
	WJetsToLNu-HT-1200To2500
	WJetsToLNu-HT-2500ToInf
tt + X	TTZToLLNuNu-M-10
	ttHJetTobb-M125
	ttHJetToGG-M125-ext1
	ttHJetToNonbb$_M$125
	TTTT
	TTWJetsToLNu
	TTZToQQ

We look at some distributions using both MC process (signal and background) and 2017 Data. The following plots include the combined final state leptons plus jets (electron or

muon plus jets).
The selections that we applied until this point are:

- MET greater than 20 GeV.
- Exactly one electron or muon with $|\eta| > 1.56$ or $|\eta| < 1.44$
- Zero Taus
- At least 4 Jets
- At least 2 of the Jets must be B-tagged Jets

The characteristics of all these objects are described in chapter 4.
The additional selection about lepton's η was applied due to Data and MC disagreement in that region (ECAL transition region).

Jet Multiplicity

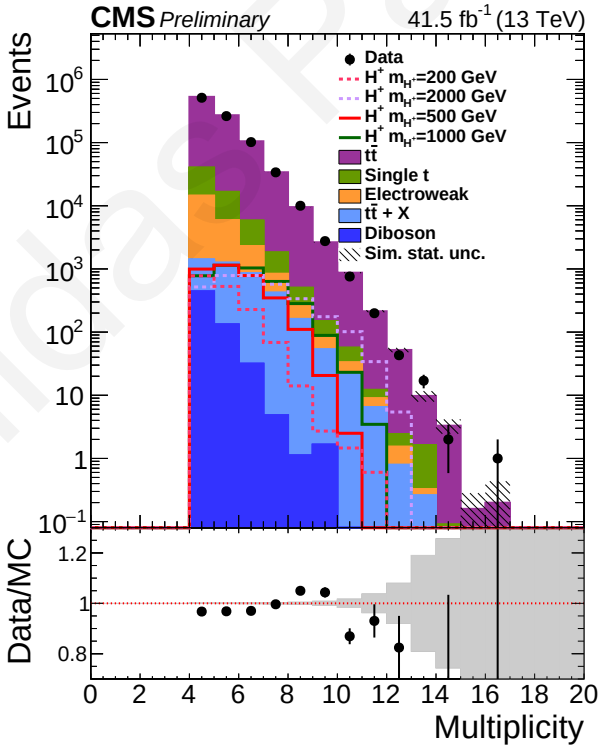

FIGURE 5.1.12: Jet multiplicity.

BJet Multiplicity

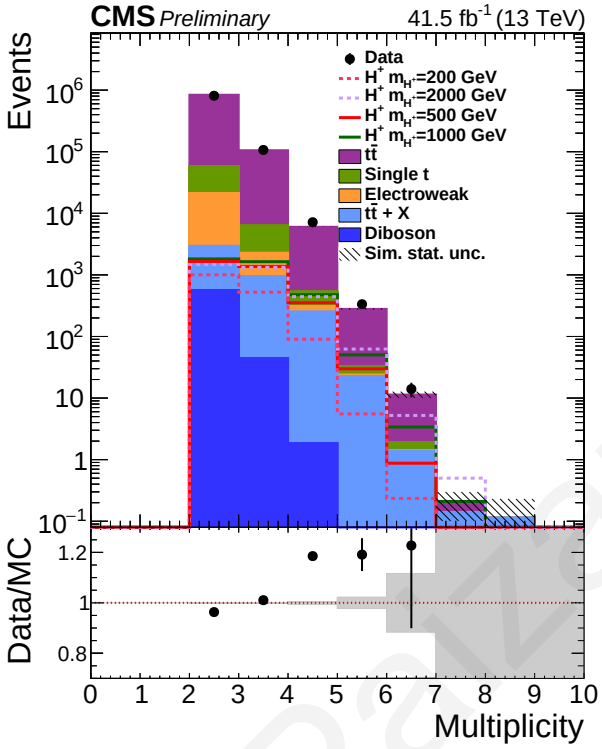

FIGURE 5.1.13: BJet multiplicity.

Lepton's P_T and Eta

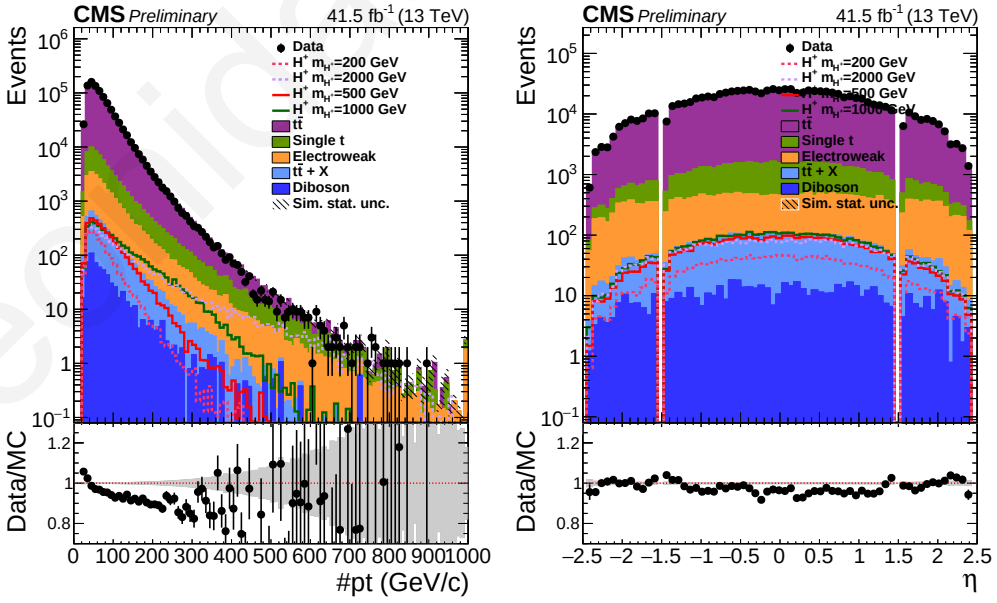

FIGURE 5.1.14: Lepton's P and eta.

MET and HT

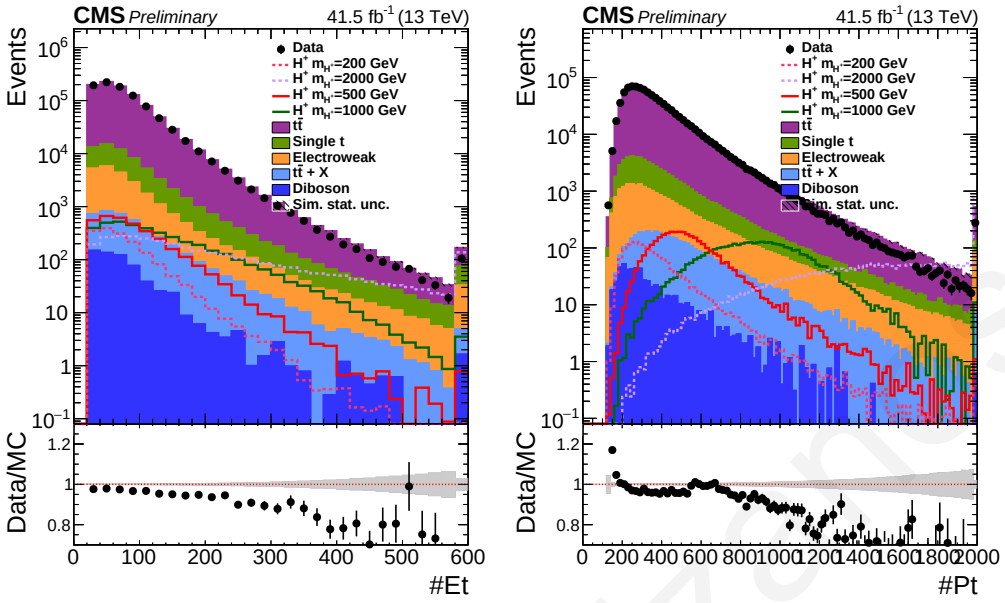

FIGURE 5.1.15: MET and HT.

P_T of the Leading BJet

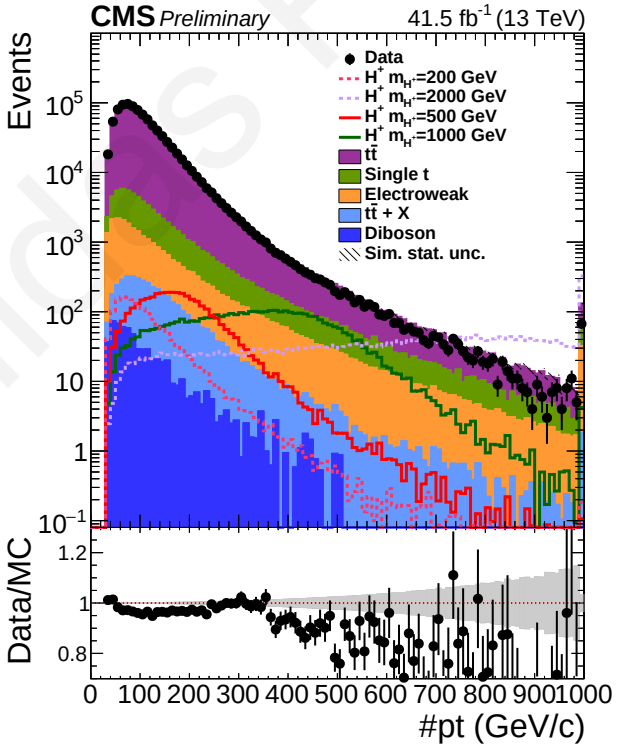

FIGURE 5.1.16: P_T of the Leading BJet.

From these plots we can observe that for the majority of the variables there is agreement between data and MC in regions with high statistic. Additionally, apart from lepton's

η as the charged Higgs mass increases, signal distributions behave different from the background. As we have discussed before, this happens because when a particle decays, depending on its mass,it gives its products more momentum, something that affect kinematic variables and object multiplicities.

5.2 Discrimination Between Signal and Background Events

Looking at the plots of section 5.1, we observe that signal and background distributions behave similarly for the low mass charged Higgs. For larger charged Higgs masses, despite the difference, the contribution of the background process is enormous compared to the signal. To detect the signal process, we need to find a way to discriminate between charged Higgs process and the background process. There are different ways of doing that, and there isn't a way that is always correct. It depends on many things, such as the final state we are interested in, the variables we are using, the statistics, etc. The way to decide is always based on which method gives the best discrimination between signal and background events.

But before this, let's start by looking at how many signal and background events exist in different categories based on the number of jets and bjets. A helpful process to choose the regions dominated by the background process (control region) and the regions with higher signal sensitivity (signal regions).

We decided to split the phase space in three orthogonal signal regions and one control region based on the jet and b-tagged jet multiplicity:

- CR: Exactly four jets including two b-tagged jets
- SR1: At least four jets including exactly three b-tagged jets.
- SR2: At least five jets including exactly two b-tagged jets.
- SR3: At least four jets including at least four b-tagged jets.

In our case, the background is dominated by $t\bar{t}$ process. Thus, this is the background that we have to discriminate our signal from. For the hypothesis of charged Higgs mass close to top quark mass and lepton plus jets final state, this is challenging because most distributions behave similarly for the two processes. But this is the purpose of our analysis. The goal is to search for variables that may provide even small discrimination between signal and background events for charged Higgs with mass 200GeV and use them for all

the other Higgs mass hypotheses.

The fact that most variables behave similarly for the two processes prevents us from doing a cut-based analysis. We couldn't find enough variables in which we could have applied some cuts to decrease $t\bar{t}$ contribution. This kind of analysis is easier for a hypothesis of charge Higgs with higher masses, where their products are boosted, something that is not happening in the $t\bar{t}$ case. For our analysis where the contribution of each variable is minimal, we choose to use the Boosted Decision Tree method.

5.3 Multivariate BDTG Classifier

Decision tree logic is based on event categorization as signal or background after repeatedly yes/no decisions taken on one single variable at a time. The selected variable in each split node is the one that, after its cut, gives the best discrimination between signal and background. Thus, some variables can be used more than once, and others are not going to be used. This finally defines the importance of each distribution. Figure 5.3.1 illustrates the structure of a decision tree classifier.

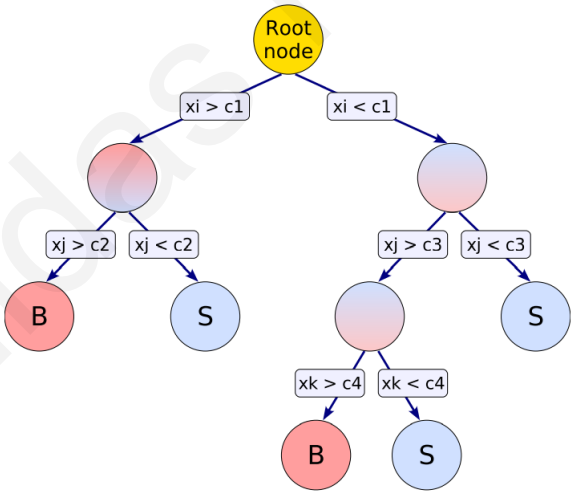

FIGURE 5.3.1: Schematic view of a decision tree

Starting from the root node, each variable's distribution is scanned to select the optimal cut. The scanning step is a fixed number that we can decide depending on the situation. After this process, there is an initial splitting using the variable that it's cut that maximizes node quantity $p * (1 - p)$, where purity (p) is the ratio of signal events over all the events in that node.
Unlike a cut-based method from which we can select only one region of phase space, a

decision tree splits the phase space into many regions and labels them as signal or background depending on what type of events have reached that node. This node splitting stops when a maximum depth value is reached.

The boosting of a decision tree uses the above logic, but instead of using one tree, it combines several trees, forming a forest. This is useful because more than one variable might give a similar separation at a specific node. At this point, the question is which one of these variables is the best to use. The result can be affected by the statistical fluctuations of the training sample. Using several trees can prevent this type of problem by minimizing the effect of the statistical changes and producing more reliable selections between the variables.

The selected variables are chosen based on their discrimination between signal and background at the signal hypothesis with a mass of 200 GeV using 2017 signal samples. As we have already seen, the background is dominated by $t\bar{t}$ and a single top process. For this reason, the background that was used for the BDTG method includes the following 2017 samples:

- TTToSemiLeptonic
- TTTo2L2Nu
- ST-s-channel-4f-hadronicDecays
- ST-t-channel-antitop-4f-InclusiveDecays
- ST-t-channel-top-4f-InclusiveDecays
- ST-tW-antitop-5f-inclusiveDecays
- ST-tW-top-5f-inclusiveDecays

The task of finding discriminant variables in our case is really difficult so we need to plan our strategy carefully.

A good starting point is the combination of generator-level information and the physics behind the signal and background process. It's a way to examine if some kinematic or topological variables have discrimination between the signal and the background process and search for more complicated distributions that are expected to behave differently. After the first investigation, it is necessary to also look at the same variables at the reconstruction level to examine whether the discrimination exists or if they look different due to incorrect particle detection and reconstruction. Unfortunately, apart from some

variables, most of the time, the latter happens. As a result, we decided to build an algorithm that will automatically compute some specific quantities using all the possible combinations between the reconstructed final and not-final state objects. We applied this for five variables (Invariant Mass, P_T, $\Delta\Phi$, $\Delta\eta$ and ΔR) using final state products such as the leading in Pt jet/bjet and tagged objects, like the leptonically decaying top and bjets from Higgs decay.

Selected variables:

A/A	Variable	Description	SR1	SR2	SR3
1	HT	$\sum_{i=0}^{N_{Jets}} P_{T\,Jet_i}$	√	√	√
2	Circularity	-	√	√	√
3	Centrality	-	√	√	√
4	Min[$\Delta R(b,b)$]	Min ΔR between two bjets	√	√	√
5	Min[$\Delta R(l,b)$]	Min ΔR between lepton and bjets	√	√	√
6	Min[Mass(b,J + J/b)]	-	√	√	√
7	Max[$\Delta\eta(b,b)$]	Max $\Delta\eta$ between two bjets	√	√	√
8	$\Delta R(J_1, Min[P_T(J,J/b)])$	-	√	√	√
9	$\Delta R(J1, b2)$]	ΔR between first jet and second bjet	√	√	√
10	Min[$\Delta\phi(b_l, MinMass(b,b)$]	-	√	-	√
11	$\Delta R(b3, Min[P_T(J,J/b)])$]	-	√	-	√
12	$\Delta R(b1, b3)$]	ΔR between the first and third bjets	√	-	√
13	Min[Mass(bbb)]	Min mass of a tri-bjet	√	-	√
14	$\Delta R(b,b)$ with Min[$\Delta\eta$(b,b)]	ΔR of the two bjets with minimum $\Delta\eta$	√	-	√
15	$\Delta R(l,b3)$]	ΔR of lepton and third bjet	√	-	√
16	Min[Mass(l,b)]	Min mass from lepton plus bjet system	-	√	√
17	Min[$\Delta R(b_h, bb)$]	Min ΔR between higgs bjet and a di-bjet	-	√	-
18	$\Delta\phi(Met, J3)$	$\Delta\phi$ between MET and third jet	-	√	-
19	$\Delta\phi(Met, b1)$	$\Delta\phi$ between MET and leading bjet	-	√	-

- Circularity: Is a variable constructed from the Jets of the event. It is define as the complement of planarity, a variable that measure the alignment of the Jets in the plane perpendicular to the beam.
- Centrality: The ratio of the sum of the transverse momentum of all jets and their total energy. It determines the degree of the collision peripherality.
- Min[Mass(b,J + J/b)]: The minimum mass between a b-tagged jet and the system of a di-jet or a jet-bjet.
- $\Delta R(J_1, Min[P_T(J,J/b)])$: ΔR between the leading non b-tagged jet and the system of a di-jet or a jet-bjet with minimum P_T
- Min[$\Delta\phi(b_l, MinMass(b,b)$]: Min $\Delta\Phi$ between the leptonic bjet and the di-bjet with the minimum mass
- $\Delta R(b3, Min[P_T(J,J/b)])$: ΔR between the third b-tagged jet and the system of a di-jet or a jet-bjet with minimum P_T

The following figures illustrates some of the selected variables for the first signal region (SR1), for the muon plus jets final state (signal hypothesis with 200 GeV charged Higgs):

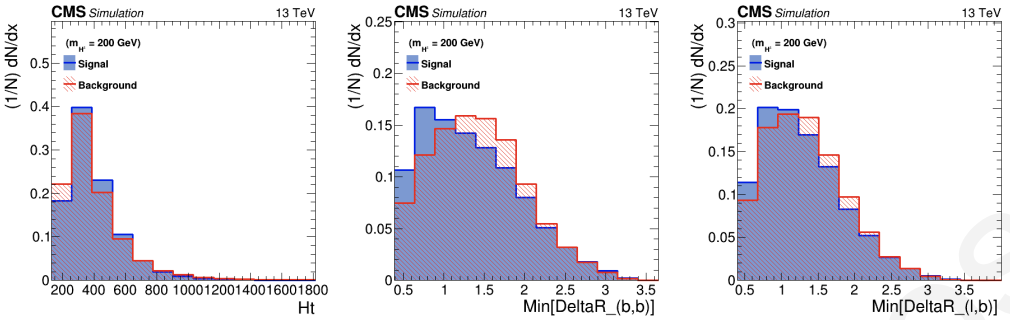

FIGURE 5.3.2: HT (left), Min ΔR between two bjets (center) and Min ΔR between lepton and bjets (right) for signal hypothesis of charged Higgs with mass 200 GeV.

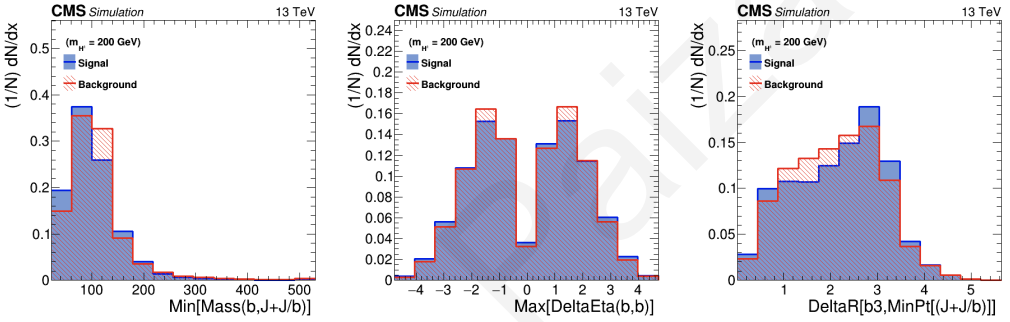

FIGURE 5.3.3: The minimum mass between a b-tagged jet and the system of a di-jet or a jet-bjet (left), Max $\Delta\eta$ between two bjets (center) and ΔR between the third b-tagged jet and the system of a di-jet or a jet-bjet with minimum P_T (right) for signal hypothesis of charged Higgs with mass 200 GeV.

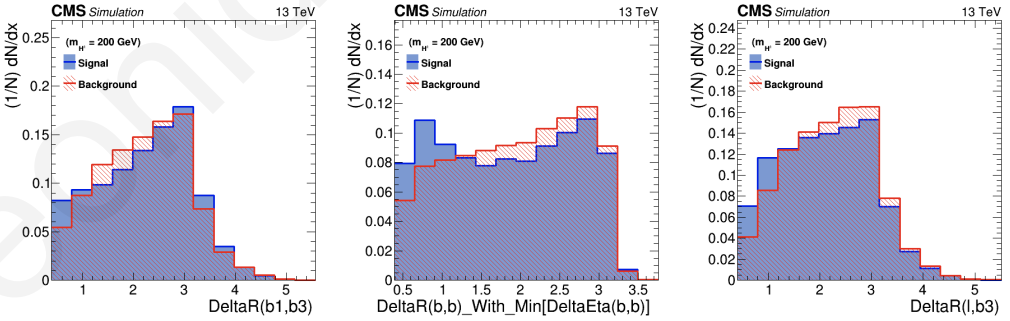

FIGURE 5.3.4: ΔR between the first and third bjets (left), ΔR of the two bjets with minimum $\Delta\eta$ (center) and ΔR of lepton and third bjet (right) for signal hypothesis of charged Higgs with mass 200 GeV.

The following figures illustrates some of the selected variables for the first signal region (SR2), for the muon plus jets final state:

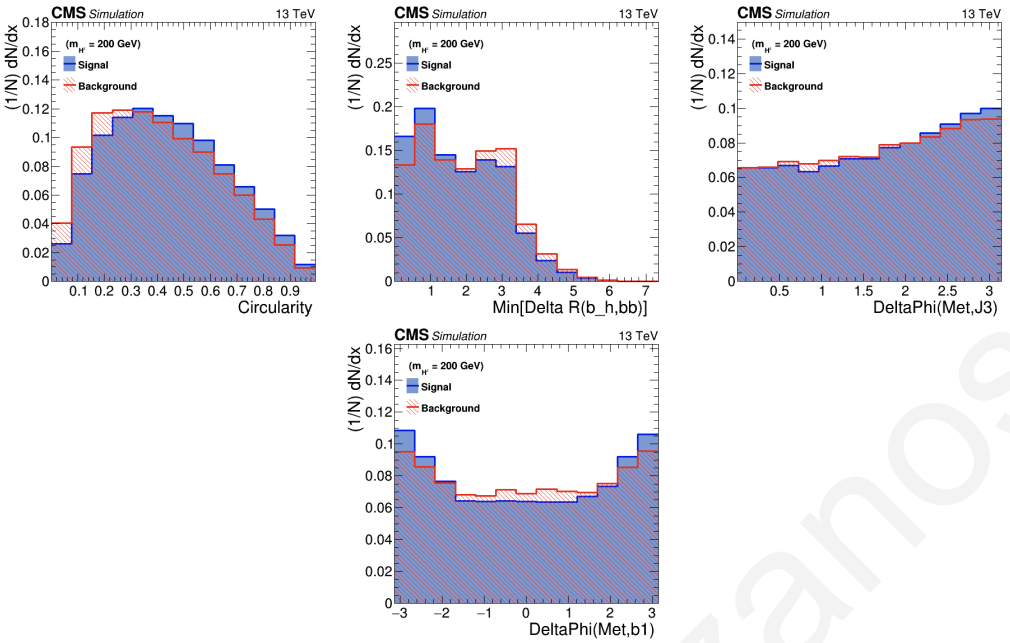

FIGURE 5.3.5: Circularity (left), Min ΔR between higgs bjet and a di-bjet (center), $\Delta\phi$ between MET and third jet (right) and $\Delta\phi$ between MET and leading bjet (bottom) for signal hypothesis of charged Higgs with mass 200 GeV.

Finally, the last variables are plot for the third signal region (SR3), for muon plus jets final state.

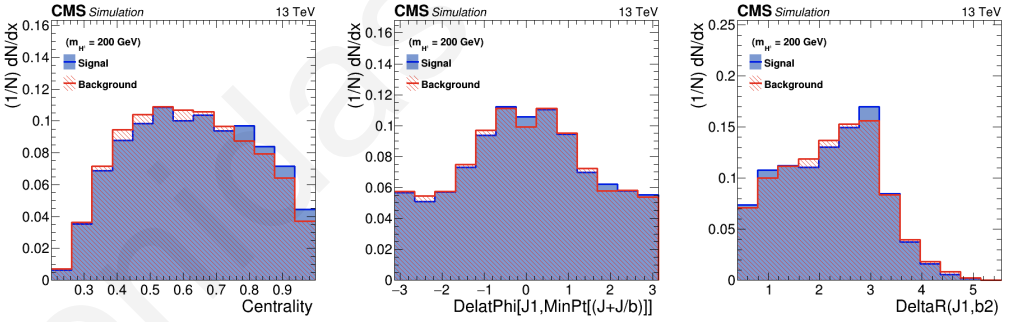

FIGURE 5.3.6: Centrality (left), ΔR between the leading non b-tagged jet and the system of a di-jet or a jet-bjet with minimum P_T (center) and ΔR between first jet and second bjet (right) for signal hypothesis of charged Higgs with mass 200 GeV.

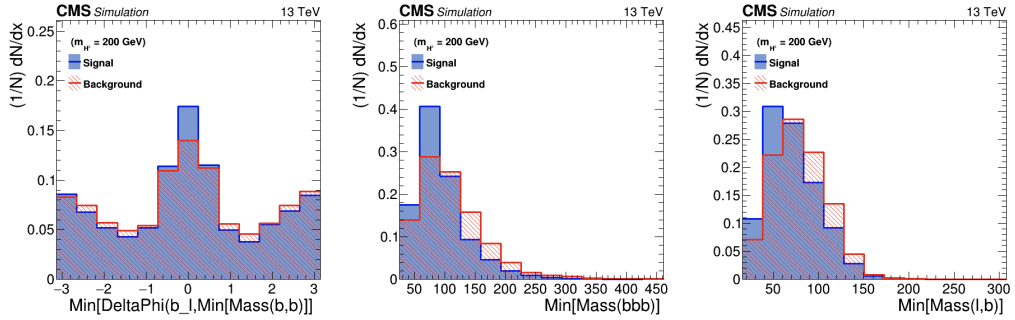

FIGURE 5.3.7: Min $\Delta\Phi$ between the leptonic bjet and the di-bjet with the minimum mass (left), Min mass of a tri-bjet (center) and Min mass from lepton plus bjet system (right) for signal hypothesis of charged Higgs with mass 200 GeV.

The following figures shows the correlation between the selected variable for the three signal regions:

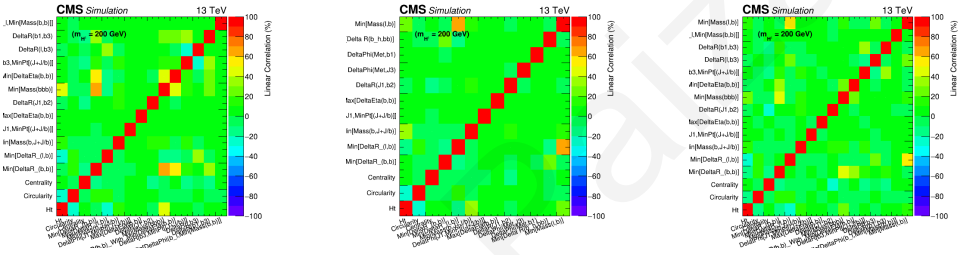

FIGURE 5.3.8: Correlation matrices for SR1 (left), SR2 (center) and SR3 (right).

As we can observe from the matrices their is no significant correlation between the selected variables.

The BDTG classifier uses these variables as input and provides us with two final distributions. The MVA output and the efficiencies of signal and background for training and testing events. From the first, we can study the achieved discrimination between the signal and background events based on the selected variables. From the second distribution, we can get the corresponding signal and background efficiencies for the different scores of MVA. Additionally, we can observe the behavior of training and testing events. This is essential because it's the only way to validate the classifier's performance. If the two curves of the training and testing event are not close to each other, it means that the classifier learns something wrong. Maybe it is caused by under-training or over-training effects.

The MVA output with associate efficiencies for the signal mode hypothesis with mass 200 GeV for the three signal regions and muon plus jets final states is:

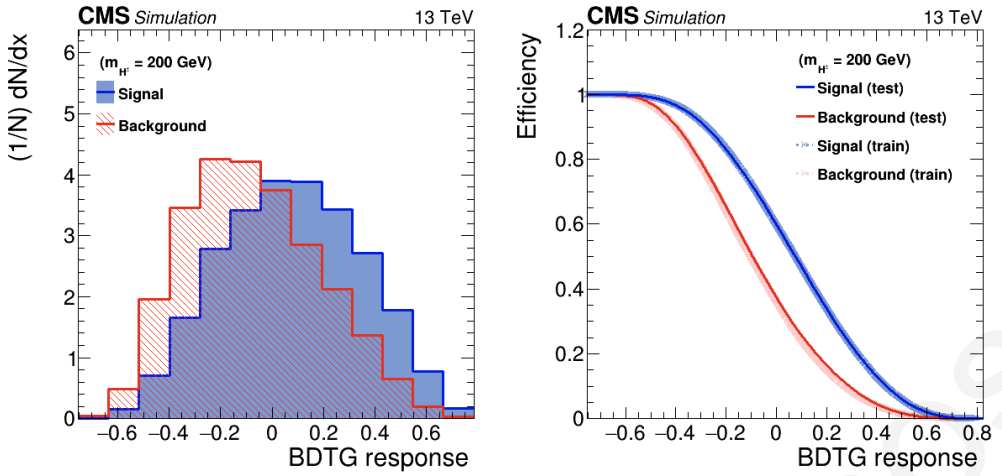

FIGURE 5.3.9: At least four jets including exactly three b-tagged jets.

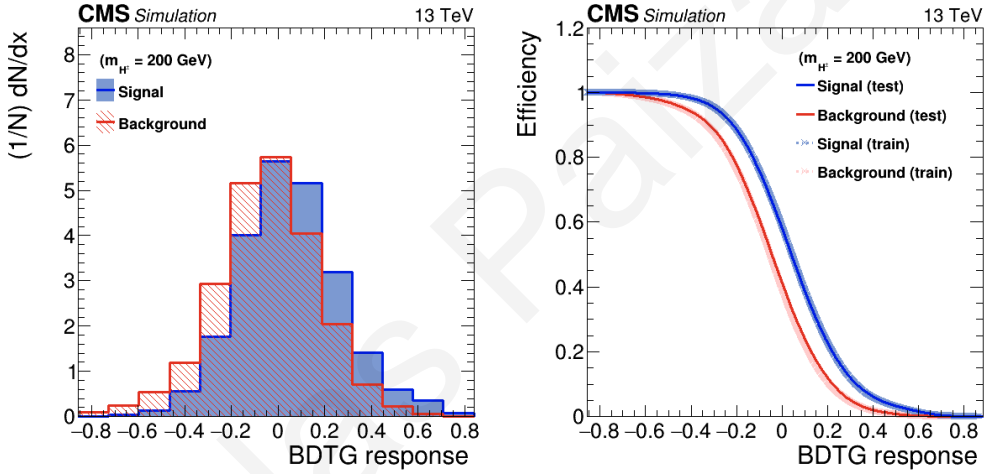

FIGURE 5.3.10: At least five jets including exactly two b-tagged jets.

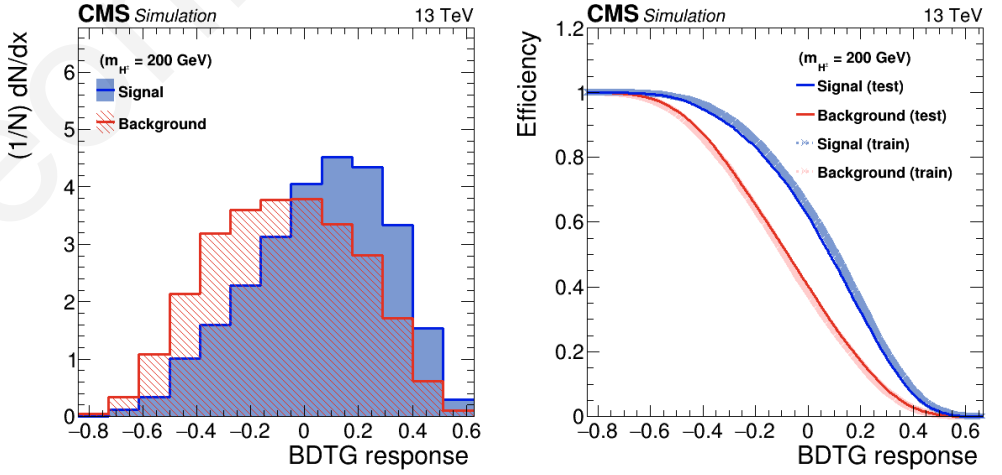

FIGURE 5.3.11: At least five jets including at least four b-tagged jets.

Let's now look at the same plots for a higher mass of charged Higgs to investigate their behaviour.

Selected variables for charged Higgs with mass 500 GeV in muon plus jets final state:

The following figures illustrates some of the selected variables for the first signal region (SR1), for the muon plus jets final state:

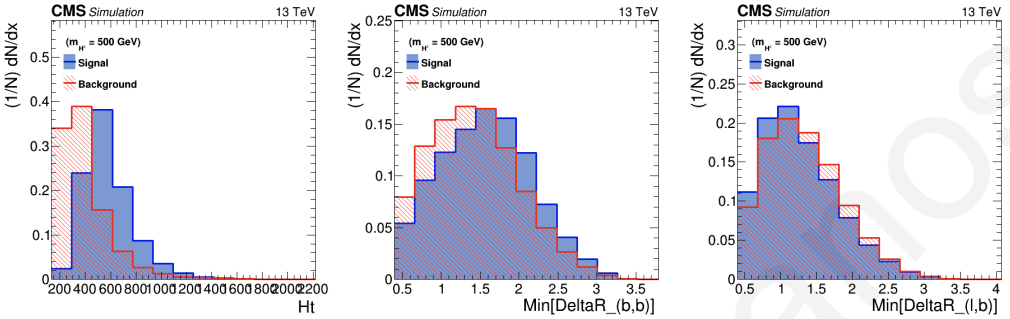

FIGURE 5.3.12: HT (left), Min ΔR between two bjets (center) and Min ΔR between lepton and bjets (right) for signal hypothesis of charged Higgs with mass 500 GeV.

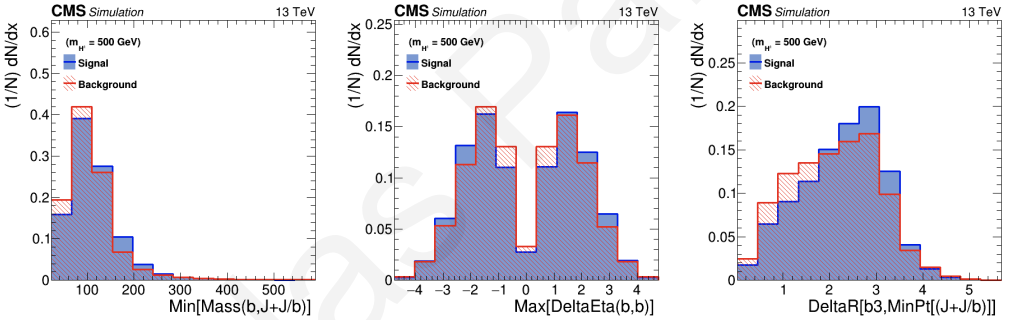

FIGURE 5.3.13: The minimum mass between a b-tagged jet and the system of a di-jet or a jet-bjet (left), Max $\Delta\eta$ between two bjets (center) and ΔR between the third b-tagged jet and the system of a di-jet or a jet-bjet with minimum P_T (right) for signal hypothesis of charged Higgs with mass 500 GeV.

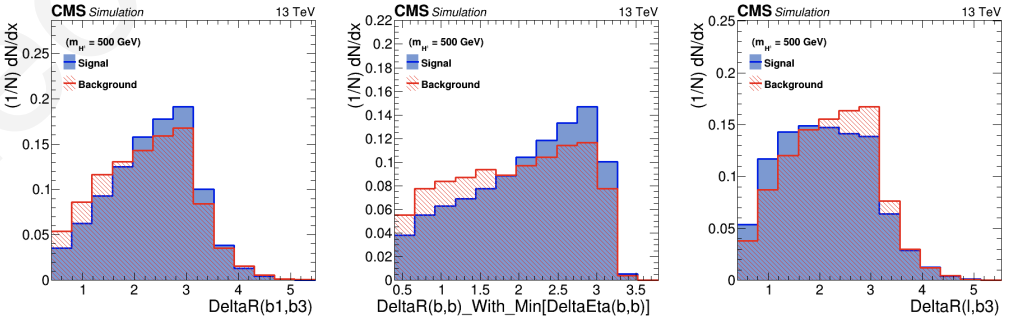

FIGURE 5.3.14: ΔR between the first and third bjets (left), ΔR of the two bjets with minimum $\Delta\eta$ (center) and ΔR of lepton and third bjet (right) for signal hypothesis of charged Higgs with mass 500 GeV.

The following figures illustrates some of the selected variables for the first signal region (SR2), for the muon plus jets final state:

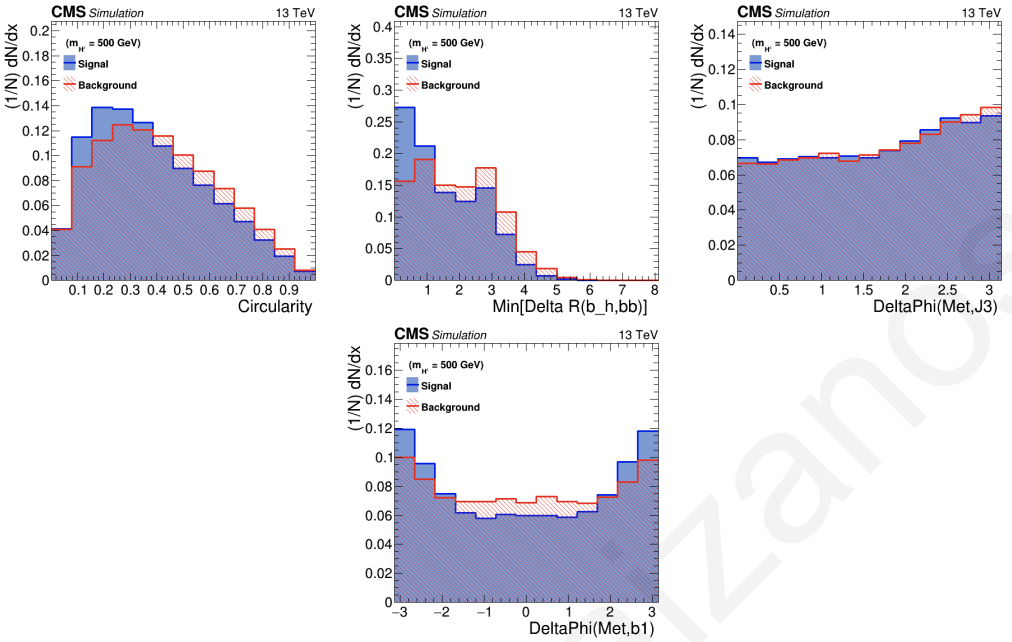

FIGURE 5.3.15: Circularity (left), Min ΔR between higgs bjet and a di-bjet (center), $\Delta\phi$ between MET and third jet (right) and $\Delta\phi$ between MET and leading bjet (bottom) for signal hypothesis of charged Higgs with mass 500 GeV.

Finally, the last variables are plot for the third signal region (SR3), for muon plus jets final state.

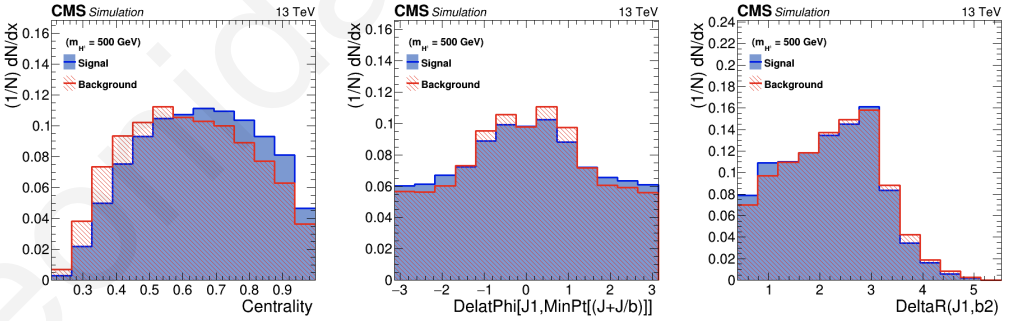

FIGURE 5.3.16: Centrality (left), ΔR between the leading non b-tagged jet and the system of a di-jet or a jet-bjet with minimum P_T (center) and ΔR between first jet and second bjet (right) for signal hypothesis of charged Higgs with mass 500 GeV.

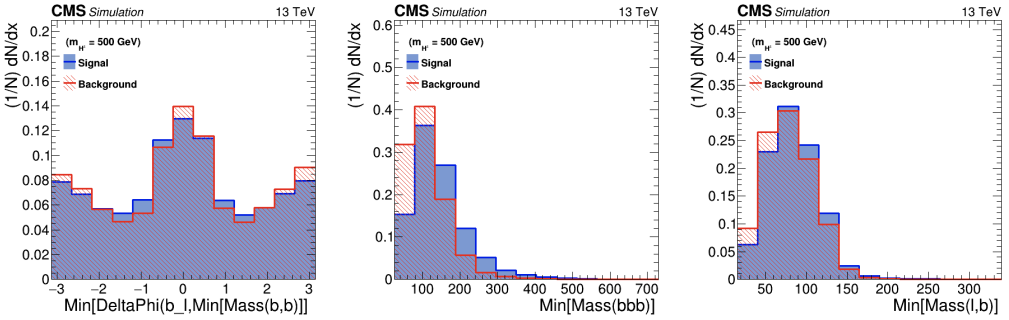

FIGURE 5.3.17: Min $\Delta\Phi$ between the leptonic bjet and the di-bjet with the minimum mass (left), Min mass of a tri-bjet (center) and Min mass from lepton plus bjet system (right) for signal hypothesis of charged Higgs with mass 500 GeV.

The following figures shows the correlation between the selected variable for the three signal regions:

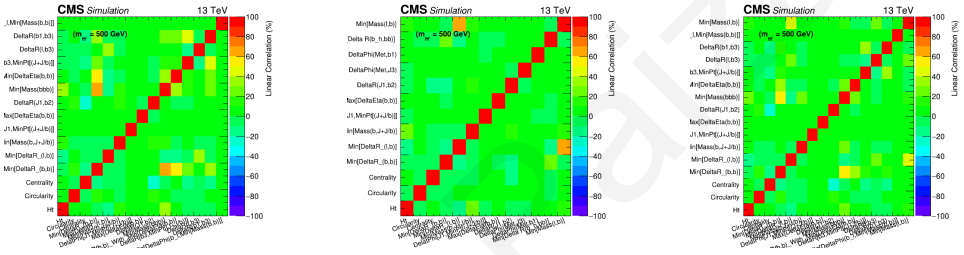

FIGURE 5.3.18: Correlation matrices for SR1 (left), SR2 (center) and SR3 (right).

The MVA output with associate efficiencies for the signal mode hypothesis with mass 500 GeV for the three signal regions and muon plus jets final states is:

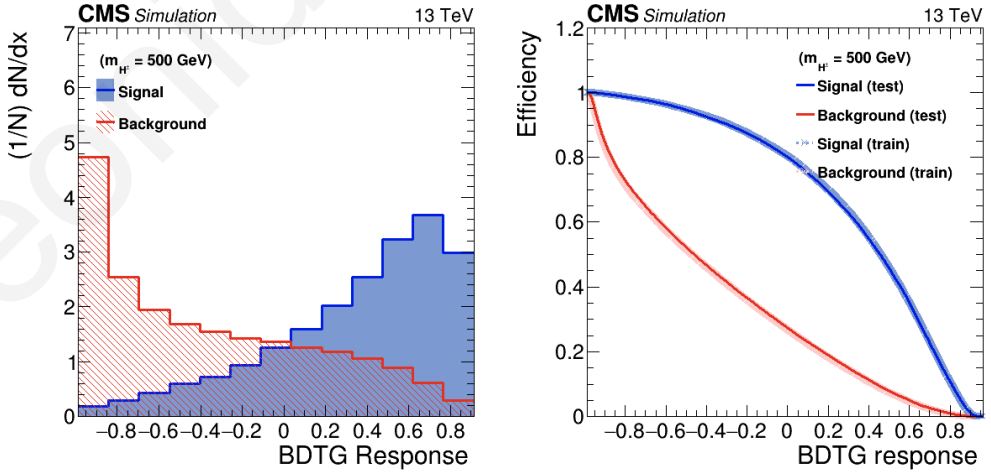

FIGURE 5.3.19: At least four jets including exactly three b-tagged jets.

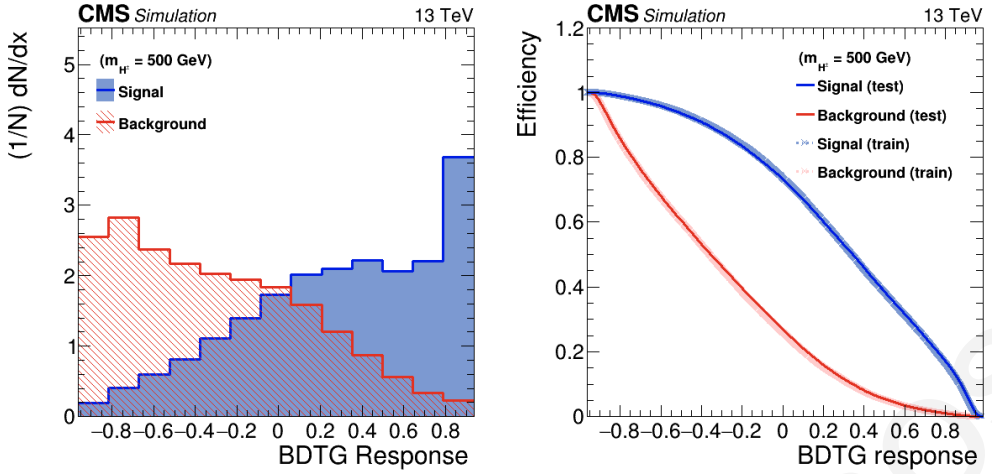

FIGURE 5.3.20: At least five jets including exactly two b-tagged jets.

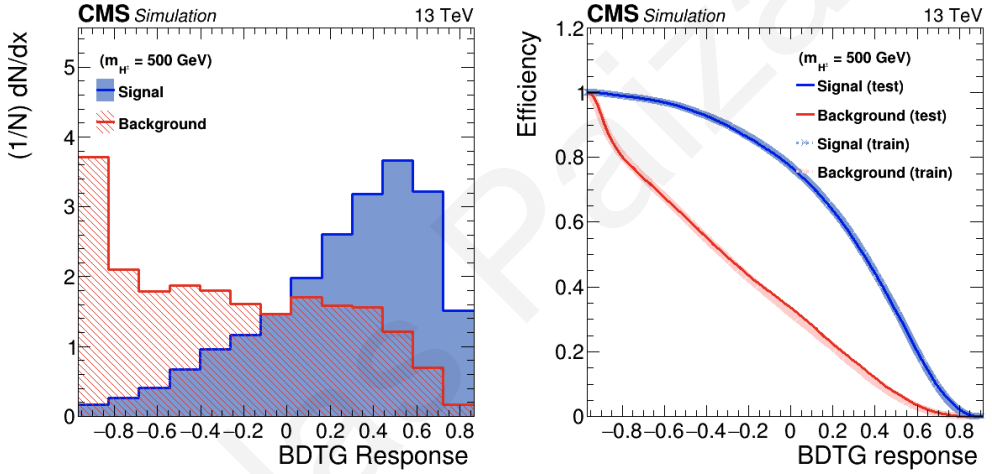

FIGURE 5.3.21: At least five jets including at least four b-tagged jets.

Selected variables for charged Higgs with mass 1000 GeV in muon plus jets final state:

The following figures illustrates some of the selected variables for the first signal region (SR1), for the muon plus jets final state:

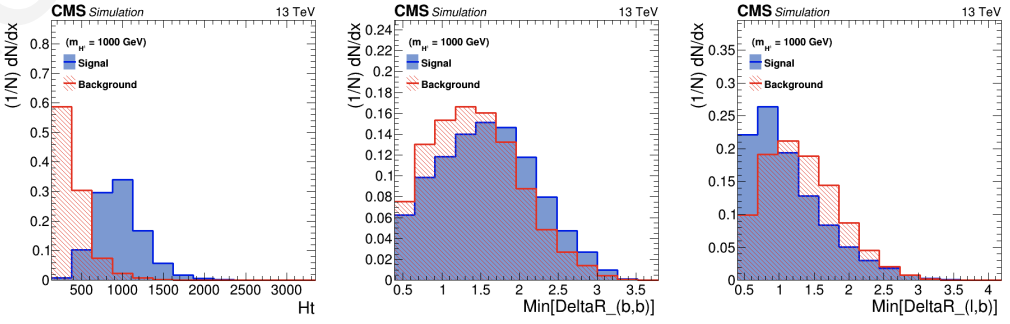

FIGURE 5.3.22: HT (left), Min ΔR between two bjets (center) and Min ΔR between lepton and bjets (right) for signal hypothesis of charged Higgs with mass 1000 GeV.

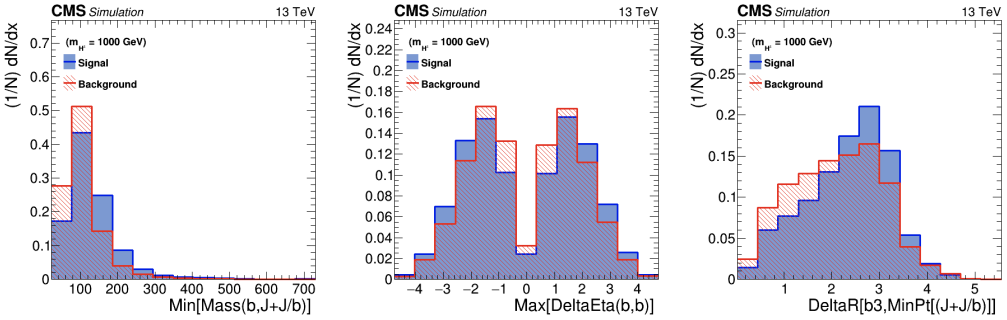

FIGURE 5.3.23: The minimum mass between a b-tagged jet and the system of a di-jet or a jet-bjet (left), Max $\Delta\eta$ between two bjets (center) and ΔR between the third b-tagged jet and the system of a di-jet or a jet-bjet with minimum P_T (right) for signal hypothesis of charged Higgs with mass 1000 GeV.

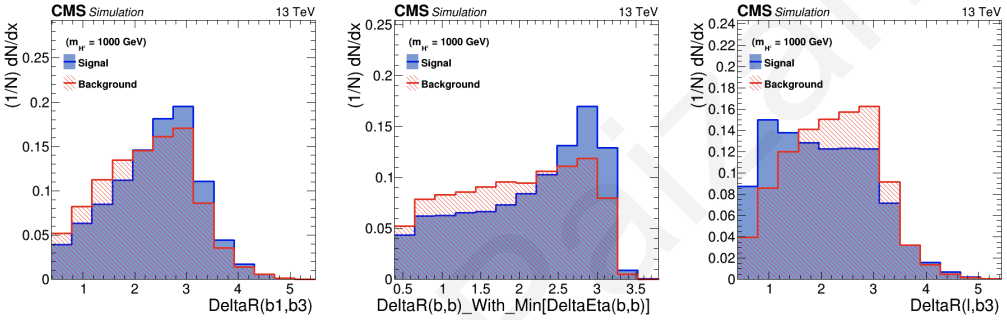

FIGURE 5.3.24: ΔR between the first and third bjets (left), ΔR of the two bjets with minimum $\Delta\eta$ (center) and ΔR of lepton and third bjet (right) for signal hypothesis of charged Higgs with mass 1000 GeV.

The following figures illustrates some of the selected variables for the first signal region (SR2), for the muon plus jets final state:

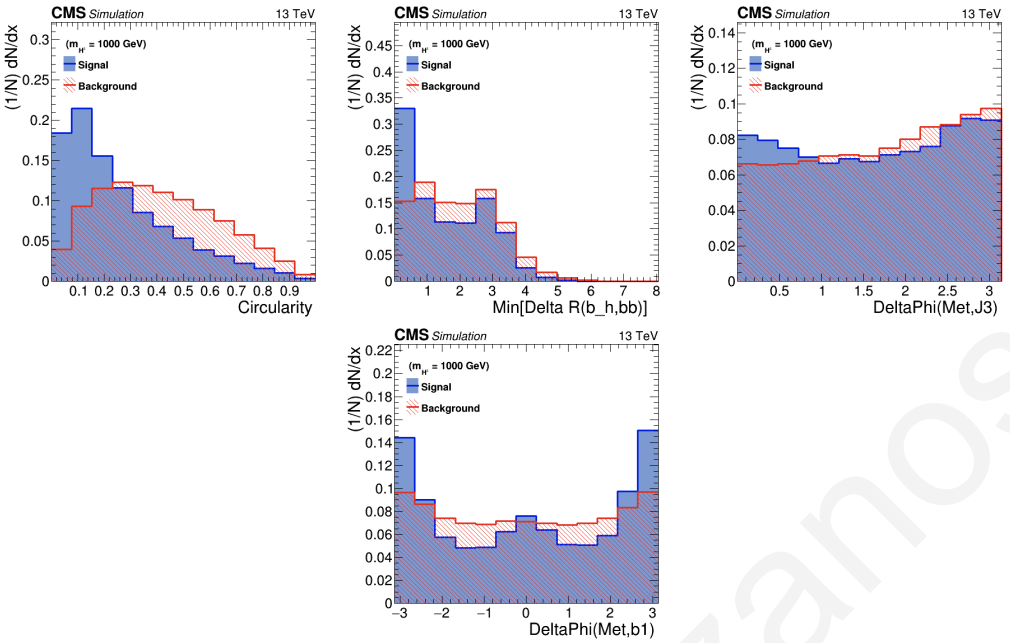

FIGURE 5.3.25: Circularity (left), Min ΔR between higgs bjet and a di-bjet (center), $\Delta\phi$ between MET and third jet (right) and $\Delta\phi$ between MET and leading bjet (bottom) for signal hypothesis of charged Higgs with mass 1000 GeV.

Finally, the last variables are plot for the third signal region (SR3), for muon plus jets final state.

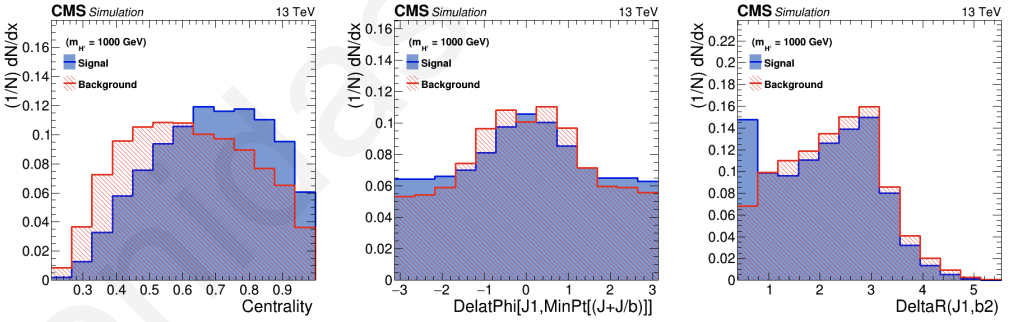

FIGURE 5.3.26: Centrality (left), ΔR between the leading non b-tagged jet and the system of a di-jet or a jet-bjet with minimum P_T (center) and ΔR between first jet and second bjet (right) for signal hypothesis of charged Higgs with mass 1000 GeV.

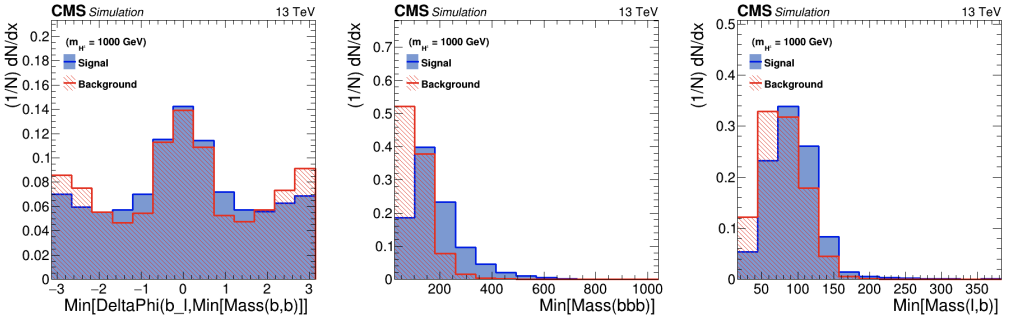

FIGURE 5.3.27: Min $\Delta\Phi$ between the leptonic bjet and the di-bjet with the minimum mass (left), Min mass of a tri-bjet (center) and Min mass from lepton plus bjet system (right) for signal hypothesis of charged Higgs with mass 1000 GeV.

The following figures shows the correlation between the selected variable for the three signal regions:

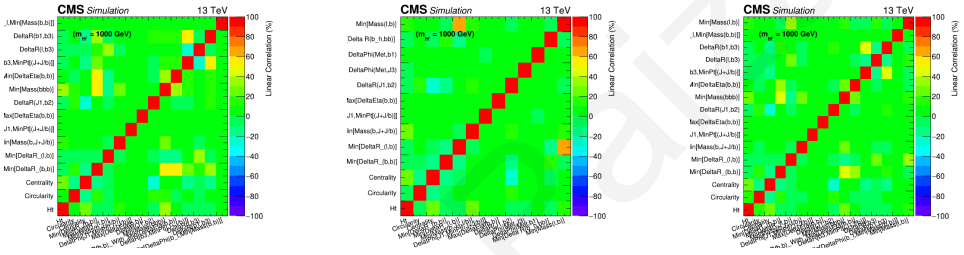

FIGURE 5.3.28: Correlation matrices for SR1 (left), SR2 (center) and SR3 (right).

The MVA output with associate efficiencies for the signal mode hypothesis with mass 1000 GeV for the three signal regions and muon plus jets final states is:

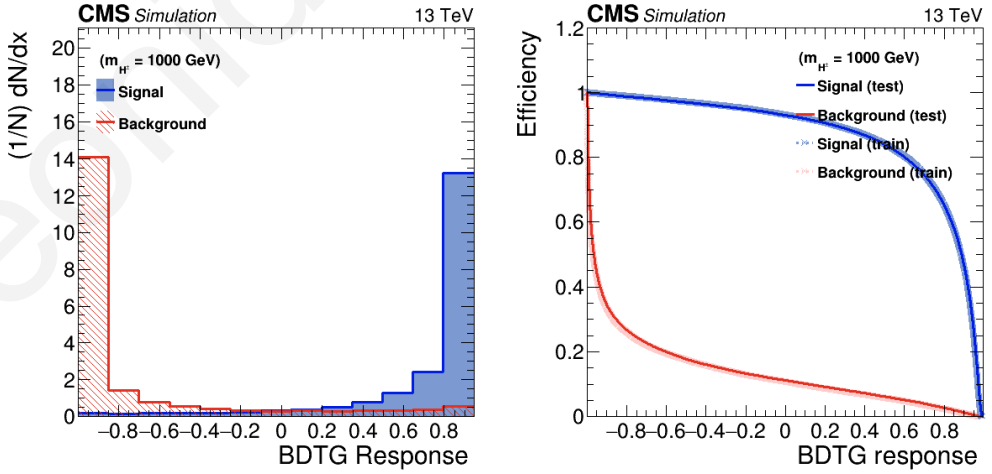

FIGURE 5.3.29: At least four jets including exactly three b-tagged jets.

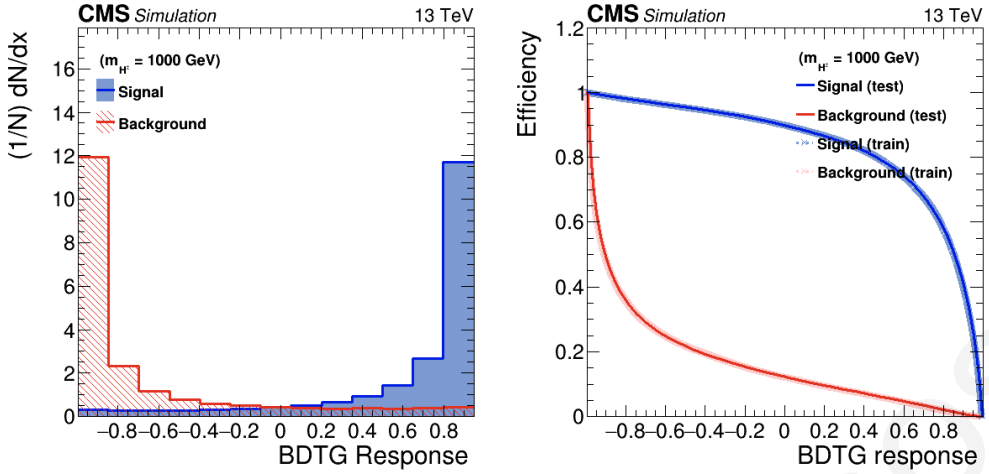

FIGURE 5.3.30: At least five jets including exactly two b-tagged jets.

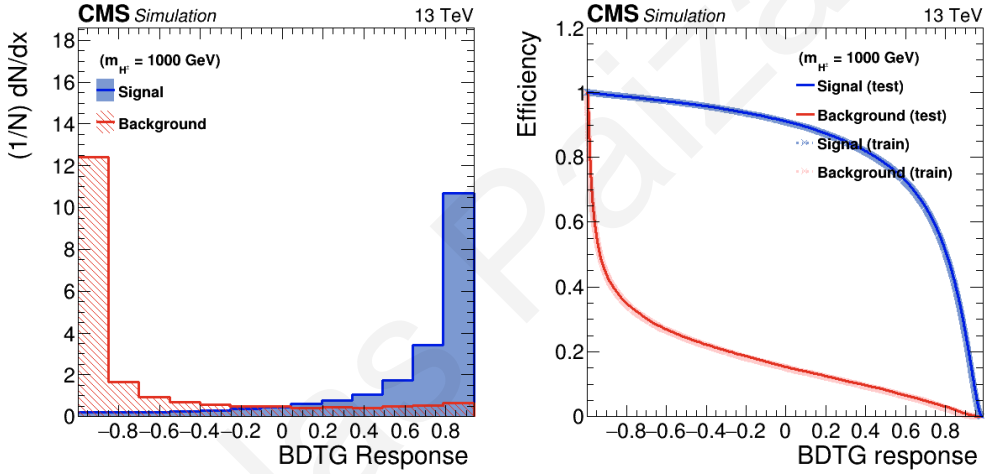

FIGURE 5.3.31: At least five jets including at least four b-tagged jets.

5.3.1 Expected Limits

The output distribution of the BDTG classifier's output is used as a fit discriminant. The training for each charged Higgs mass is assumed to extract the limit for all the other mass hypotheses. The final limit is a complication of the best limits of each mass point, obtained during the training of each mass hypothesis. The limits obtained without considering any systematic uncertainties. The results shown in figure 5.3.11, illustrate the upper expected limits for the combined category lepton plus jets final state in the mass range of 200 - 3000 GeV and are calculated at 95% confidence level (CL) using the asymptotic LHC-type CL_s criterion.

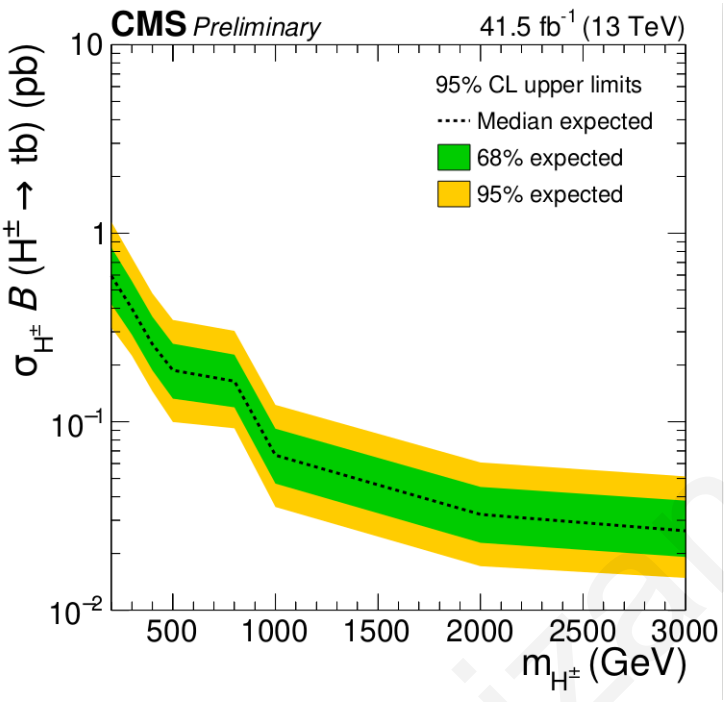

FIGURE 5.3.32: The observed and expected limits 95% CL upper limits for the $H^{\pm}$ mass range of 200-3000 GeV

The values of the expected upper limits and for 68% and 95% expected regions at each of the mass points are listed in the following table:

$M_{H^{\pm}}$	Expected Limit				
(GeV)	-2σ	-1σ	Median	+1σ	+2σ
200	0.31543	0.41980	0.59375	0.83989	1.14919
300	0.22412	0.28949	0.39844	0.55726	0.74129
400	0.14502	0.18732	0.25781	0.36058	0.47966
500	0.09229	0.13257	0.18750	0.25925	0.34669
800	0.09229	0.11920	0.16406	0.22684	0.30336
1000	0.03528	0.04695	0.06641	0.0 9182	0.12279
2000	0.01712	0.02279	0.03223	0.04507	0.06078
3000	0.01483	0.01916	0.02637	0.03814	0.05128

The behaviour of the expected limit distribution for masses of charged Higgs between 500 and 800 GeV is because of the absence of signal samples in that region. This leads to a not accurate prediction of the expected limits in that region.

Chapter 6

Search for type-III Seesaw Heavy Fermions in muons plus AK8-jets final state

This analysis main purpose is to search for heavy fermions decaying to W, Z, or Higgs Bosons in muons plus AK8 Jets final state. The search is based on decays that produce a pair of heavy fermions, either charged-charged or charged-neutral, and it was done for three different signal hypotheses with masses 850,1000, and 1500 GeV. This study includes 2017 proton-proton collision data collected by CMS detector at a center of mass energy of $\sqrt{s} = 13$TeV and integrated luminosity 41.5 fb^{-1}.

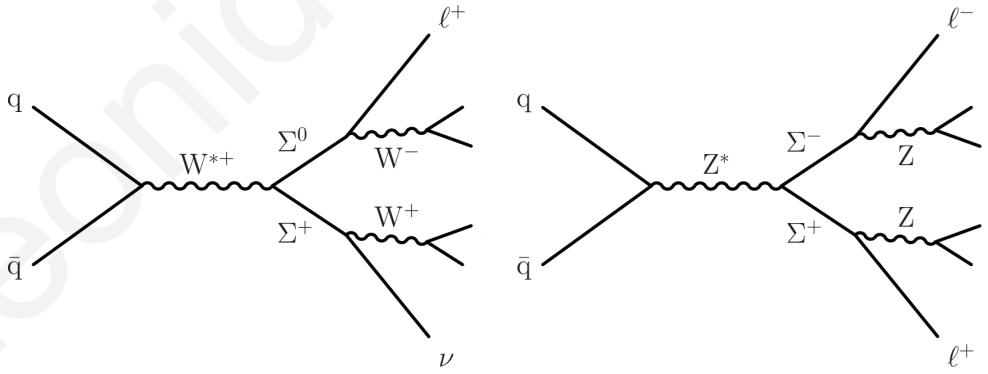

FIGURE 6.0.1: Example processes illustrating production and decay of type-III seesaw heavy fermion pairs at the LHC that result in lepton plus jets final states

The motivation for this analysis is that final states with AK8 jets can possibly increase the signal sensitivity for large heavy fermions masses over CMS analyses that only use leptons final state. The hypothesis is that a massive particle such as a heavy fermion with a mass around 1 TeV can produce bosons (W, Z, H) with high P_T. The products

of a particle with high P_T usually are boosted (close to each other). Thus, we aim for events where the bosons of the process decay hadronically and the produced quarks are close to each other, forming AK8 Jets. This effect is not easy to occur in the SM process. For this reason, we want to investigate whether the presence of AK8 Jets will increase our signal sensitivity.

From figure 6.0.1, we can observe that the final state depends on the bosons decay. The case that we want to investigate includes the following two categories:

- One Muon plus two AK8 Jets (left diagram of figure 6.0.1)
- Two Muons plus two AK8 Jets (right diagram of figure 6.0.1)

After the reconstruction process discussed in chapter 3, we have final state objects with different qualities. Thus, firstly it's necessary to apply some initial selections that clean up the particle collections from objects with bad quality. Once the the particles that pass the identification criteria are selected, we can move to the event selections. Requiring specific numbers of some particles is important because we can reject events originated from other background contributions.

6.1 Selections:

Muons:

Selections	Exactly One or Two Muons (Opposite Sign)
Pt $>$	30 GeV
Eta $<$	2.4
ID:	Tight (Muons-isCutBasedIDTight)
Isolation:	Tight (Muons-is-PFIsoTight)

Tight ID muons are reconstructed with more than 99% efficiency. Additionally, the tight working point of the isolation corresponds to a 98% efficiency. Thus, we can achieve an "accurate" reconstruction of the muons participating in the events. Since we expect events with one or two muons, we can use the muons selection as an initial definition of the two final states splitting the events into two categories. The first includes events with precisely one muon and covers the case of the charged-neutral heavy fermion production. The second selection is exactly two muons with opposite sign, aiming for the charged-charged heavy fermion production.

Electrons:

Selections	Electrons
Pt >	30 GeV
Eta <	2.4
ID:	Medium (cutBasedElectronID-Fall17-94X-V2-medium)

AK8 Jets:

Selections	At least One AK8-Jet	
Pt >	200 GeV	
Eta <	2.4	
ID:	Tight (AK8Jets-IDTight) AK8 Jet - Lepton $\Delta R >$	0.8

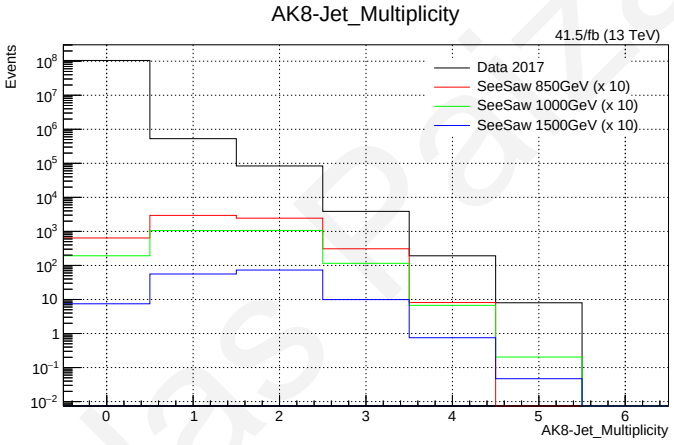

FIGURE 6.1.1: Multiplicity of AK8 Jets for events with one or two muons

Both final states include two AK8 Jets. Unfortunately, the jet reconstruction isn't so successful as the muon reconstruction. Thus, the requirement of exactly two AK8 Jets might increase the purity of the final state but at the same time rejects a significant number of signal events making the analysis difficult due to low statistics. For this reason, our selection includes at least one AK8 Jet and not precisely two. This allows the possibility of losing one AK8 Jet during the reconstruction process.

B-Tagged Jets:

Selections	B-Tagged Jets
Pt >	30 GeV
Eta <	2.4
Discriminator:	Jets-pfDeepFlavourBJetTags (Medium)
BJet - Lepton $\Delta R >$	0.4

After setting our initial selections and decrease the contribution of other final states that may originate from other process, we can now look at the mass distributions of the two leading in P_T AK8 Jets and discuss their shape. For this scope we are going to use the pruned mass distribution.
Pruning is the process of jet reclustering removing soft large angle particles.

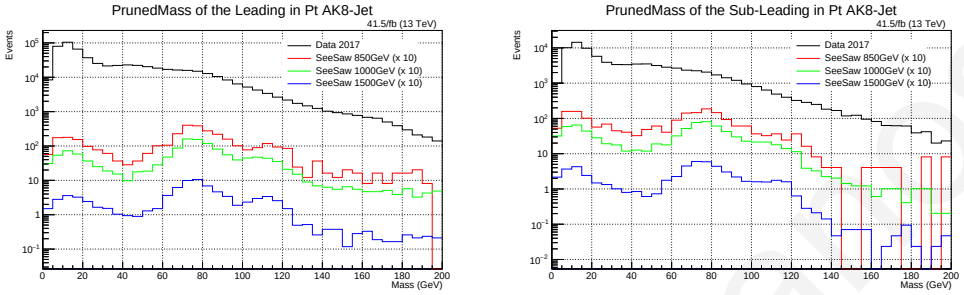

FIGURE 6.1.2: Pruned Mass of Leading (left) and sub-leading (right) in P_T AK8 Jets in single lepton final state

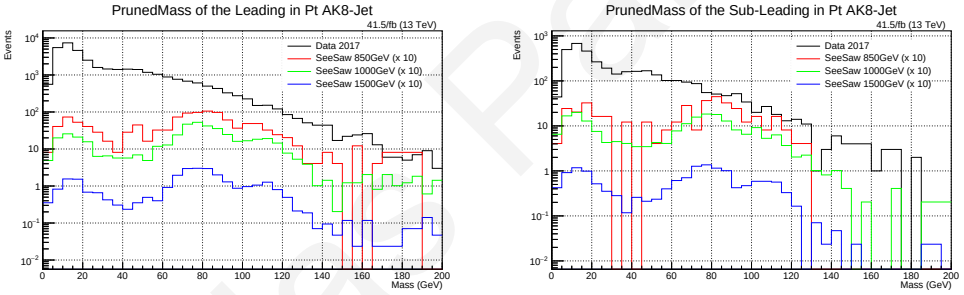

FIGURE 6.1.3: Pruned Mass of Leading (left) and sub-leading (right) in P_T AK8 Jets in single lepton final state

From figure 6.1.3 we can see that there are three peaks in the pruned mass distribution for both categories and both AK8 Jets.

- 0 - 40 GeV : These region possibly contain jets originated from QCD.
- 60 - 100 GeV : In this mass window there are events in which the mass of an AK8 jet is close to the mass of W (80.4 GeV) and Z (91.2 GeV) bosons.
- 100 - 140 GeV : Similar to the previous mass region but in this case the AK8 jets mass is close to higgs boson mass (125 GeV).

The second and third mass regions contain the bosons that participate in the process that interests us. Thus, requiring the mass of the selected AK8 jets to be in one of these two

regions is an efficient selection to reject events that include AK8 Jets originating from QCD and, at the same time, increase the possibility of finally tagging a heavy fermion.

Final selection:

- 60 GeV $<$ AK8 Jet Pruned Mass $<$ 140 GeV

6.2 Heavy fermion mass reconstruction

The next task is the reconstruction of the heavy fermion mass. We checked the minimum and maximum mass of a muon with an AK8 jet over all possible combinations. The results for the single-muon and di-muon final states present in the figures below.

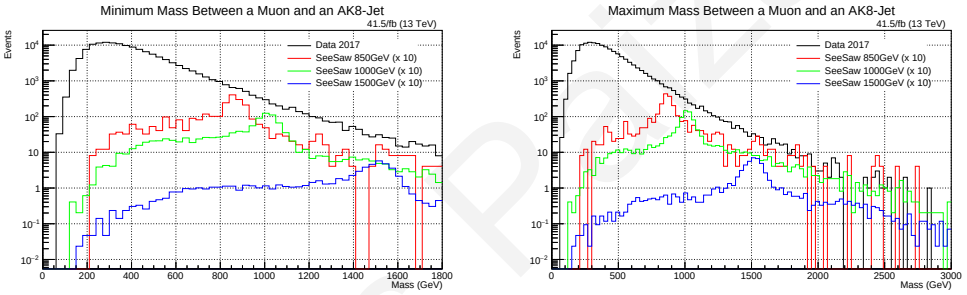

FIGURE 6.2.1: Minimum and Maximum mass of a muon and a fat jet over all possible combinations for single-muon final state

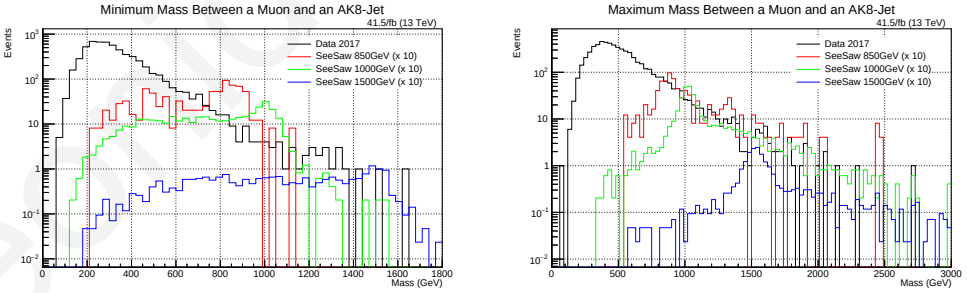

FIGURE 6.2.2: Minimum and Maximum mass of a muon and a fat jet over all possible combinations for di-muon final state

The figures above show that the peaks of the heavy fermion mass distribution (850,1000 and 1500 GeV) are more evident in the maximum mass distributions. In our case, this method seems to achieve better mass reconstruction.

6.3 Increasing Signal Sensitivity

At this stage of the analysis, we need to search for some variables that have different behavior between signal and background events to reduce background contribution and increase signal sensitivity.

As we have discussed before, we are looking for a signal hypothesis with a mass around 1 TeV, which means that we will have the production of particles with high P_T. Thus it is reasonable to search for discriminant variables associated with the P_T of the final state products.

The selected variables are the following:

- P_T of the leading in P_T muon.
- P_T of the leading in P_T AK8 jet.
- Scalar sum of all jets P_T (HT).
- Missing Transverse Energy (MET) (Single-Muon category only).
- P_T of the sub-leading in P_T muon (Di-Muon category only).
- Bosons tagging
- B-Tagged Jet multiplicity

Looking at one variable each time, we applied some cuts and selected the one that gave the best result for the reconstructed mass of the heavy fermion.

Transverse Momentum of the Leading Muon

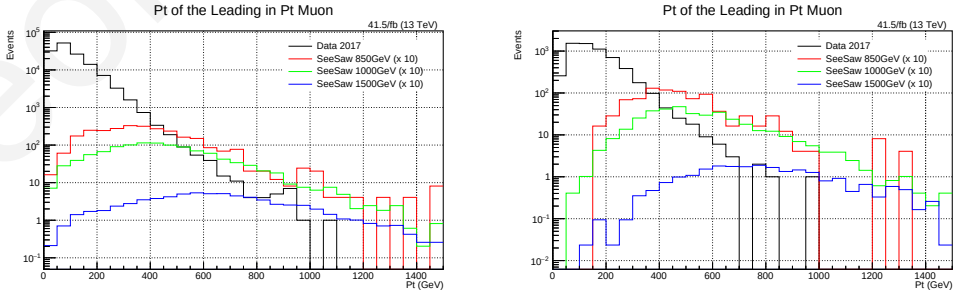

FIGURE 6.3.1: P_T of the leading muon for single muon and di-muon final states

The cuts for the P_T of the leading muon were 50, 100, and 150 GeV. The following figures illustrate the heavy fermion mass distribution for each of these cuts for single and di-muon final states.

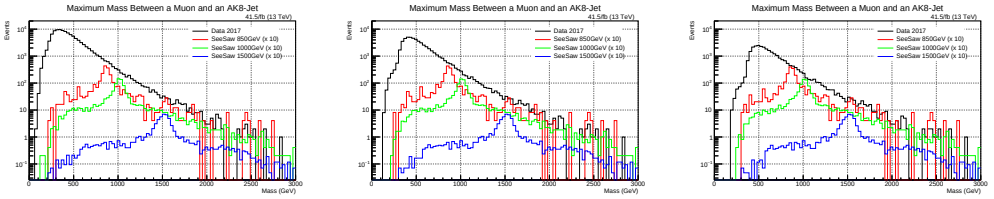

FIGURE 6.3.2: Reconstructed mass of the heavy fermion for each cut (single-muon final state)

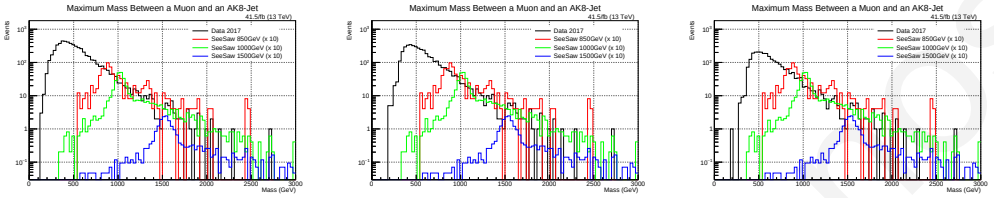

FIGURE 6.3.3: Reconstructed mass of the heavy fermion for each cut (di-muon final state)

For both cases, we decided that the third cut is optimal. Thus, we increase the leading muons P_T threshold at 150 GeV. Now let's look at the P_T distribution of the leading AK8 Jet after this new selection.

Transverse Momentum of the Leading AK8 Jet

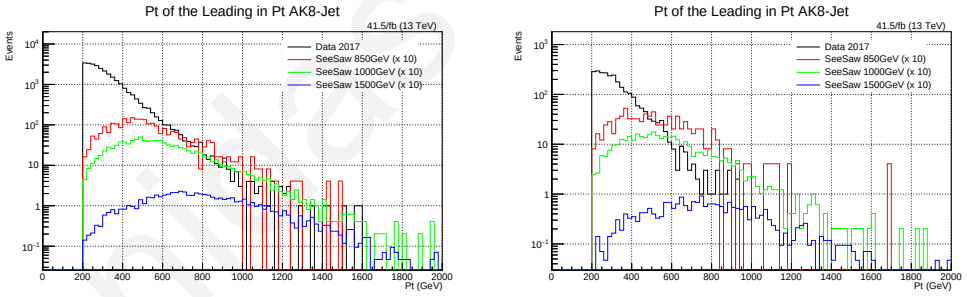

FIGURE 6.3.4: P_T of the leading AK8 jet for single muon and di-muon final states

From these plots, we can observe that most of the background events are stored below 400 GeV, so the cuts we investigated were 200, 300, and 400 GeV. The heavy fermion mass is plotted for each cut in the figures below.

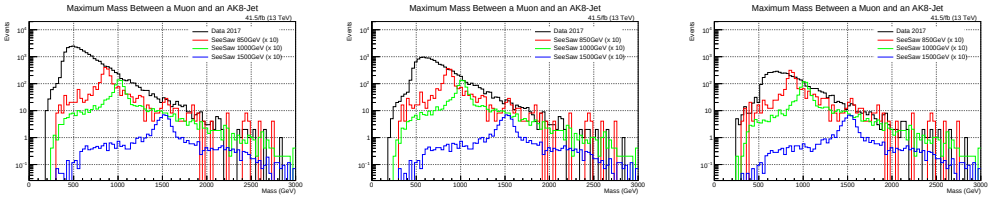

FIGURE 6.3.5: Reconstructed mass of the heavy fermion for each cut (single-muon final state)

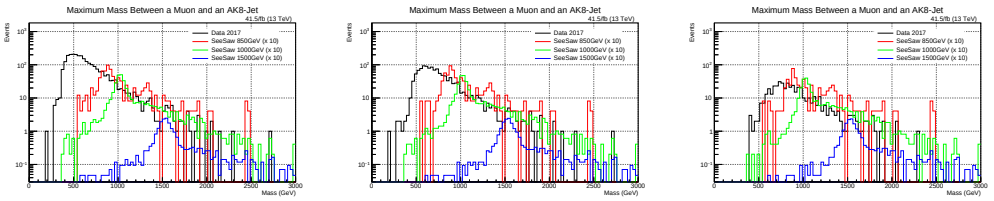

FIGURE 6.3.6: Reconstructed mass of the heavy fermion for each cut (di-muon final state)

Looking at the mass distribution for the three cuts, we can observe that as the P_T threshold increases, there is a drop in the background contribution while signal change is minimal. For this reason, the maximum cut was applied for the first final states and the second for the di-muon category. Thus, the transverse momentum of the leading AK8 Jet has to be greater than 400 GeV and 300GeV for the two categories, respectively.

HT

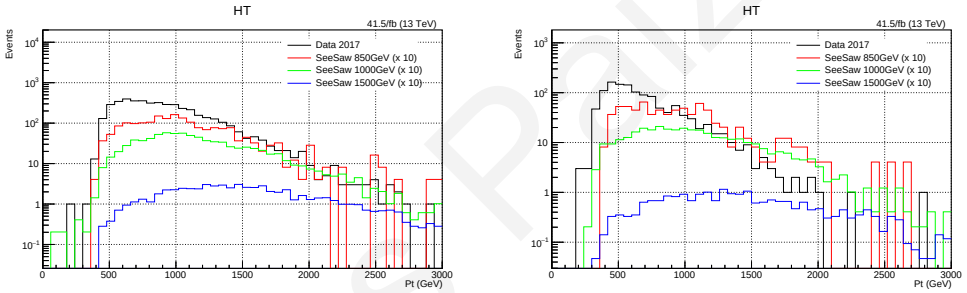

FIGURE 6.3.7: MET for single muon and di-muon final states

For the HT variable the examined cuts are 400, 500, 600 GeV. The following figures shows the mass distribution for these selections.

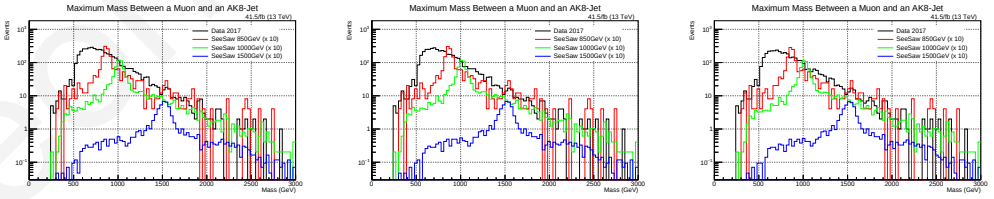

FIGURE 6.3.8: Reconstructed mass of the heavy fermion for each cut (single-muon final state)

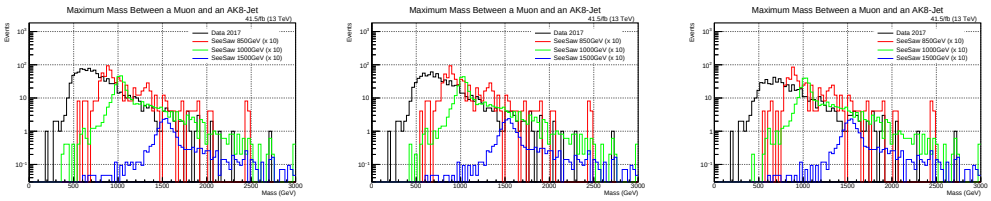

FIGURE 6.3.9: Reconstructed mass of the heavy fermion for each cut (di-muon final state)

From HT and mass distribution at each cut the HT selection has been set to 600 GeV for both final states.

MET

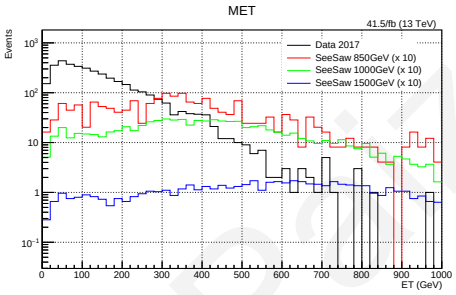

FIGURE 6.3.10: MET for single muon final state

We don't expect neutrino in the charged-charged heavy fermion production. Thus, MET was examined only for the singe-muon final state. Looking at the distribution, we set three possible thresholds. MET above 100, 150, and 200 GeV.

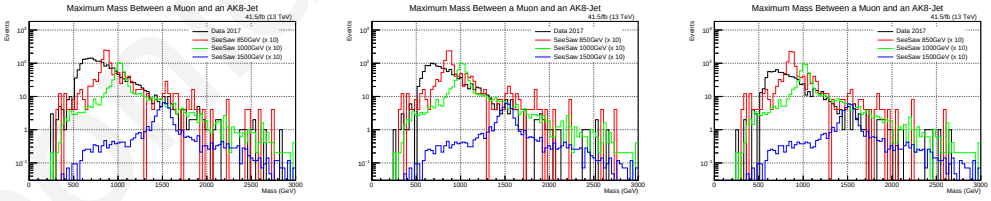

FIGURE 6.3.11: Reconstructed mass of the heavy fermion for each cut (di-muon final state)

We decided that the optimal cut is at 150 GeV.

Transverse Momentum of the Sub-Leading Muon

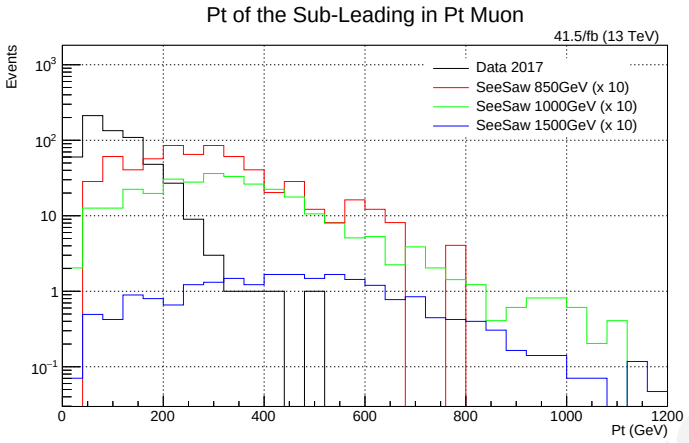

FIGURE 6.3.12: P_T of the sub-leading in P_T muon for di-muon final state

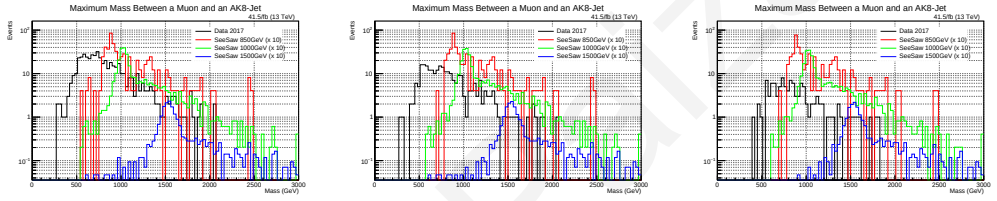

FIGURE 6.3.13: Reconstructed mass of the heavy fermion for each cut (di-muon final state)

Here we can observe a different behavior between signal and background. The second is gathered below 400 GeV, while the first is spread to 1000 GeV. The P_T cuts we examined were 50,100 and 150 GeV, where most of the background events are stored. After the investigation, the P_T threshold for the sub-leading in P_T muon was set to 150 GeV.

Boson Tagging Since we have the participation of W, Z, and Higgs bosons in the signal process, it is reasonable to use tagging discriminators. This will increase the possibility of selecting an AK8 Jet originating from one of these bosons and not from QCD.

The selected disriminators are the follwing:

- AK8Jet-WvsQCD-Discriminator
- AK8Jet-ZvsQCD-Discriminator
- AK8Jet-ZbbvsQCD-Discriminato
- AK8Jet-HbbvsQCD-Discriminator

The first is responsible for W tagging, the next two for Z bosons and the last for Higgs boson.

Our selection is that the AK8 Jet that participates in the heavy fermion's mass reconstruction must achieve a score of at least 0.7 in at least one of the four discriminators.

B-Tagged Jet Multiplicity

The last variable we will examine is the number of bjets in the event after all the applied selections.

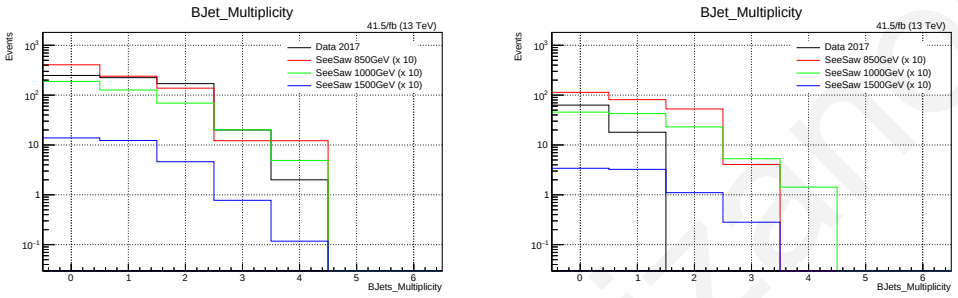

FIGURE 6.3.14: BJet multiplicity for single muon and di-muon final states

In the single-muon category, the behavior between signal and background is similar. But for the di-muon final state, the difference is noticeable. Most of the background events have zero bjets, while the multiplicity for the signal spreads up to three and four bjets. Thus we decided to discard events with zero bjets.

The following table illustrates the selected cuts for both final states.

Variables	Single-Muon	Di-Muon
P_T of the leading Muon $>$	150 GeV	150 GeV
P_T of the leading Muon $>$	400 GeV	300 GeV
HT $>$	600 GeV	600 GeV
MET $>$	150 GeV	-
P_T of the sub-leading Muon $>$	-	150 GeV
Boson Tagging	$\surd$	$\surd$
Number of BJets	≥ 1	≥ 1

After applying the final cuts we can look at the final result of the heavy fermion mass for the two final states.

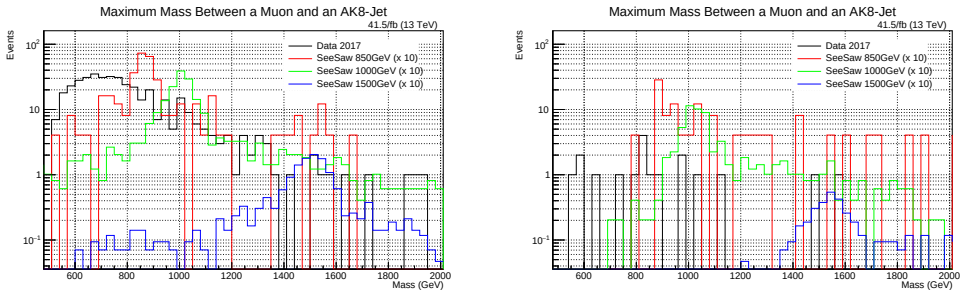

FIGURE 6.3.15: Reconstructed mass of the heavy fermion after all the selections for single-muon and di-muon final states

6.4 Extraction of the Expected Limits

Using the reconstructed mass of the heavy fermion we estimate signal sensitivity by extracting the expected limit of heavy fermion production for the two final states. For this purpose Data, background, and signal are scaled to the integrated luminosity of the complete run2 dataset.

Fermion Mass	Cross section (σ)	Upper limit on signal cross-section	
GeV	pb	Single-Muon	Di-Muon
850	$5*10^{-3}$	$3.13*10^{-3}$	$9.84*10^{-4}$
1000	$1.96*10^{-3}$	$1.81*10^{-3}$	$7.25*10^{-4}$
1500	$1.19*10^{-4}$	$3.64*10^{-4}$	$7.13*10^{-4}$

Chapter 7

Conclusions

A preliminary search for $H^{\pm}$ decaying to lepton plus jets has been performed using 2017 data and a BDTG classifier constructed with event kinematic variables. The limits obtained without considering any systematic uncertainties.

The following figure illustrate the expected limit for the production of a charged Higgs that decays to a top and a bottom quark in lepton plus jets final state. The left plot are the result of the current analyses and the right the limits extracted from the 2016 CMS analyses.

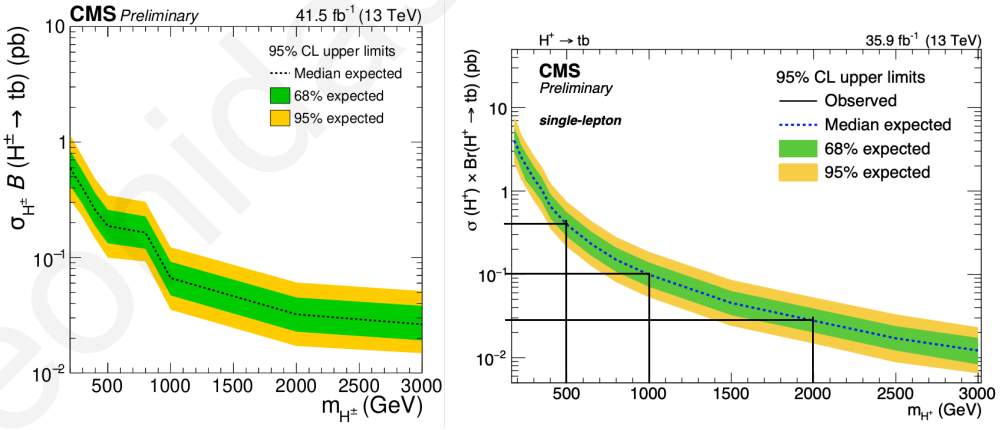

FIGURE 7.0.1: Expected limits for current analyses (left) 2016 CMS analyses (right)

Looking and compare the two plots and the marked points we can observed that expected limits of the results of the current analyses are better than the 2016 expected limits.

Additionally, a preliminary search for heavy fermions decaying to W Z or Higgs Bosons in muon plus AK8 jets is also performed.

Fermion Mass	Cross section (σ)	Upper limit on signal cross-section	
GeV	pb	Single-Muon	Di-Muon
850	$5*10^{-3}$	$3.13*10^{-3}$	$9.84*10^{-4}$
1000	$1.96*10^{-3}$	$1.81*10^{-3}$	$7.25*10^{-4}$
1500	$1.19*10^{-4}$	$3.64*10^{-4}$	$7.13*10^{-4}$

Looking at the table that illustrates the upper limit of the heavy fermion production in lepton plus AK8 jets final state we can firstly observe that a promising sensitivity for heavy fermion masses around 1 TeV is achieved. Furthermore, the results for the di-muon final state are better compared to the single-muon category. One explanation for this is that the dominant background for the single-muon is W plus jets while in di-muon final state the main background process is Z plus jets. The cross section of W plus jets is much bigger than Z plus jets. Thus, we expect much more background in the single-muon case.

The following figure shows the expected limits for the heavy fermion production calculated by the 2021 CMS analyses for multi-lepton final state.

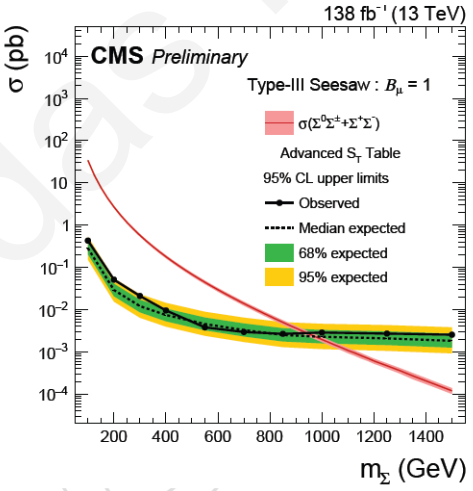

FIGURE 7.0.2: Expected limits of heavy fermion production in multi-lepton final state (2021 CMS analyses)[20]

Looking at the results of the current analyses and the expected limits of 2021 CMS analyses we can observe that the the signal sensitivity of the two final states single-muon plus AK8 jets (current analyses) and multi-lepton (2021 CMS analyses) are similar. While in the di-muon plus AK8 jets compared to multi-lepton (2021 CMS analyses) we have achieved to increase our signal sensitivity.

Bibliography

[1] Mark Thomson. *Modern Particle Physics*. Cambridge University Press, 2013. Chap. 1.1.

[2] Mark Thomson. *Modern Particle Physics*. Cambridge University Press, 2013. Chap. 13.

[3] Alessandro Bettini. *Introduction to Elementary Particle Physics*. Cambridge University Press, 2008. Chap. 10.

[4] URL: https://www.researchgate.net/figure/Neutrino-mixing-angles-assuming-zero-CP-violation-may-be-represented-as-Euler-angles_fig1_282814865.

[5] Francis Halzen and Alan D. Martin. *QUARKS AND LEPTONS: An Introductory Course in Modern Particle Physics*. 1984. Chap. 1.8.

[6] *Dark Matter*. URL: https://home.cern/science/physics/dark-matter.

[7] *Antimatter*. URL: https://home.cern/science/physics/antimatter.

[8] CMS Collaboration. *Search for charged Higgs bosons in the $H^{\pm} \to \tau^{\pm}\nu_{\tau}$ decay channel in proton-proton collisions at $\sqrt{s}$ = 13 TeV (arXiv:1903.04560 [hep-ex])*. 28/07/2019.

[9] Marius Wiesemann Celine Degrande Maria Ubiali and Marco Zaro. *Heavy charged Higgs boson production at the LHC*. 9/07/2015.

[10] Stuart Raby and Richard Slansky. *Neutrino Masses, How to add them to the Standard Model*.

[11] Lyndon Evans and Philip Bryant. *LHC Machine*. 14/08/2008.

[12] Lyn Evans and CERN Lucie Linssen. *The Super-LHC is on the starting blocks*. 8/07/2008.

[13] *CMS*. URL: https://cms.cern/detector.

[14] URL: https://www.sciencephoto.com/media/795/view/diagram-of-cms-detector-for-lhc-at-cern.

[15] Lucas Taylor. *CMS DETECTOR*. 23/11/2011.

[16] CMS Collaboration. *The CMS experiment at the CERN LHC*. 14/08/2008.

[17] CMS Collaboration. *Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC*. 19/05/2021. URL: https://arxiv.org/pdf/2012.06888.pdf.

[18] CMS Collaboration. *Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at* $\sqrt{s} = 13TeV$. 21/06/2018. Chap. 3. URL: https://arxiv.org/pdf/1804.04528.pdf.

[19] Particle Data Group. *W Boson*. URL: https://pdg.lbl.gov/2020/listings/rpp2020-list-w-boson.pdf.

[20] CMS Collaboration. *Inclusive nonresonant multilepton probes of new phenomena at* $\sqrt{s} = 13TeV$. 12/11/2021. URL: https://cms.cern.ch/iCMS/jsp/db_notes/noteInfo.jsp?cmsnoteid=CMS%20AN-2020/052.