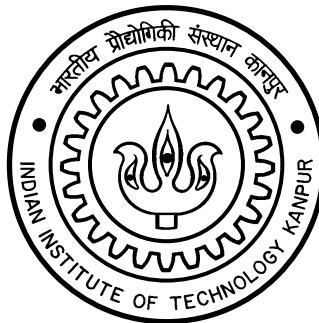


# SEARCH FOR A CHARGED HIGGS BOSON AT 13 TeV IN THE CMS EXPERIMENT AT THE LHC, CERN

*A thesis submitted*  
in partial fulfilment of the requirement  
for the Degree of  
DOCTOR OF PHILOSOPHY

by  
**RAVINDRA KUMAR VERMA (13109883)**



TO  
THE DEPARTMENT OF PHYSICS  
INDIAN INSTITUTE OF TECHNOLOGY, KANPUR, INDIA

AUGUST 2019



# Certificate

It is certified that the work contained in the thesis titled “SEARCH FOR A CHARGED HIGGS BOSON AT 13 TEV IN THE CMS EXPERIMENT AT THE LHC, CERN” by Ravindra Kumar Verma has been carried out under our supervision and this work has not been submitted elsewhere for a degree.

PROF PANKAJ JAIN

SUPERVISOR

DEPARTMENT OF PHYSICS

IIT KANPUR

PROF SHASHI R DUGAD

CO-SUPERVISOR

DEPARTMENT OF HIGH ENERGY PHYSICS

TIFR MUMBAI

AUGUST 2019



# Synopsis

Name:	<b>Ravindra Kumar Verma</b>
Roll No.:	<b>13109883</b>
Degree for which submitted:	<b>Doctor of Philosophy</b>
Department:	<b>Physics</b>
Thesis Title:	<b>Search for a charged Higgs Boson at 13 TeV in the CMS experiment at the LHC, CERN</b>
Thesis Supervisors:	<b>Prof. Pankaj Jain (IIT Kanpur), Prof. Shashi R. Dugad (TIFR Mumbai)</b>
Month and year of submission:	<b>August 2019</b>

---

The discovery of the Higgs boson in 2012 by the ATLAS and CMS experiments at the CERN LHC has ushered a new beginning in the field of particle physics. The Higgs boson could be the first of many elementary scalars present in nature to be observed in the laboratory. Various extensions to the standard model (SM), such as supersymmetry and the two Higgs doublet model (2HDM), predict multiple scalars as the remnants of an additional  $SU(2)_L$  complex doublet introduced to address some known drawbacks of the SM. After spontaneous symmetry breaking, out of the eight degrees of freedom of the two Higgs doublets, three are used to make the W and Z bosons massive, leaving five physical scalar particles. Of these, two ( $h$ ,  $H$ ) are neutral Higgs bosons which are CP-even (scalar), one ( $A$ ) is neutral and CP-odd (pseudoscalar), and the remaining two are charged Higgs  $H^\pm$  bosons.

The 2HDM can be classified into different categories depending on the type of interaction of the two doublets with quarks and charged leptons. For example, in the type I 2HDM, fermions have Yukawa couplings only to the second doublet. The nature of the Yukawa coupling determines the branching fraction of the charged Higgs boson decays into different final states. In this thesis, a search for the charged Higgs boson has been performed in the decay channel  $H^+ \rightarrow c\bar{s}$  (and its charge conjugate), whose branching fraction can range anywhere up to 100% depending on the type of Yukawa couplings. The latter is expressed in terms of

the parameter  $\tan\beta = v_2/v_1$  where  $v_1$  and  $v_2$  are the vacuum expectation values of the two Higgs doublets. In the minimal supersymmetric standard model (MSSM), this is the dominant decay channel for low values of  $\tan\beta$ . In this thesis, we perform a model-independent search for the charged Higgs boson assuming that  $\mathcal{B}(H^+ \rightarrow c\bar{s}) = 100\%$ .

Limits on charged Higgs boson production at hadron colliders have been set by the Tevatron and LHC, assuming the production mode  $t \rightarrow H^+ b$ . The CDF collaboration set a 95% CL upper limit on the branching fraction  $\mathcal{B}(t \rightarrow H^+ b) < 10\text{--}30\%$  for a charged Higgs lying in the mass range 60–150 GeV, assuming that  $H^+$  decays dominantly to  $c\bar{s}$ . Similar limits have been obtained by the D0 experiment. Using the 7 TeV data, the ATLAS collaboration set an upper limit at 95% CL on the product  $\mathcal{B}(t \rightarrow H^+ b)\mathcal{B}(H^+ \rightarrow \tau^+ \nu) < 0.23\text{--}1.3\%$  for a charged Higgs mass in the range 80–160 GeV. A search for a charged Higgs boson decaying into  $c\bar{s}$  was performed at 8 TeV by the CMS collaboration, which set an upper limit at 95% CL on  $\mathcal{B}(t \rightarrow H^+ b)$  in the range 1.2–6.5%. In this thesis, the analysis of collision data recorded in 2016 by the CMS experiment at 13 TeV is presented.

As illustrated in Figure 2.4, the principal SM background to this search consists of  $t\bar{t}$  pair production where one of the top quarks decays by  $t \rightarrow W^+ b$  and the other top quark decays by  $\bar{t} \rightarrow W^- \bar{b}$ ; this is referred to as the “semileptonic” decay channel. In contrast, in the signal process, one of the top quarks decays to  $H^+ b$  and the other to  $W^- \bar{b}$ . The  $W^+/H^+$  decays hadronically into “light” jets (not from a  $b$  quark), whereas the  $W^-$  decays leptonically. We define two channels depending on whether the lepton produced in the  $W^-$  decay is an electron or a muon (events with tau leptons are not considered).

## Summary of research work

In **chapter 1**, we present a brief theoretical overview of the standard model of particle physics where we discuss its successes and failures. In **chapter 2**, the two Higgs doublet model is described where various properties of the charged Higgs such as its interaction with the particles of SM are discussed. In the same chapter, we present the current status of the charged Higgs searches and the search strategy followed in this thesis.

In **chapter 3**, we present a brief description of the LHC and the detectors installed there. A few important parameters of the LHC have also been described along with a few physics parameters. In **chapter 4**, a brief overview of the CMS experiment is presented. “The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the solenoid volume

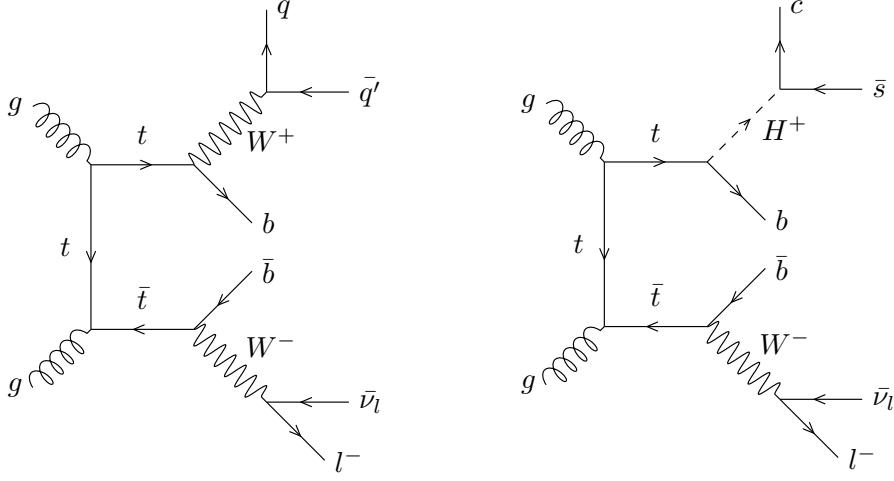


Figure 1.: Production of  $t\bar{t}$  from gluon-gluon scattering. The left plot shows the SM decay of the  $t\bar{t}$  pair in the semileptonic decay channel. The right plot shows the signal process in which the  $t\bar{t}$  pair decay products include a charged Higgs boson.

are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter (HCAL), each composed of a barrel and two endcap sections. Silicon pixel and tracker detector identifies the trajectory of charged particles and accurately measures their transverse momentum up to  $|\eta| \leq 2.5$ . Forward calorimeters extend the pseudorapidity coverage provided by the barrel and endcap calorimeter. Segmented calorimeters provide sampling of electromagnetic and hadronic showers up to  $|\eta| \leq 5$ . Muons are detected in gas-ionization chambers embedded in the steel flux-return yoke outside the solenoid in the range of  $|\eta| \leq 2.4''$ .

In **chapter 5**, we list the collision and simulated data samples. The data used for the analysis was collected by the CMS detector in 2016 in proton-proton (pp) collisions at  $\sqrt{s} = 13\text{TeV}$ , with an integrated luminosity of  $35.9\text{fb}^{-1}$ . The simulated signal and background samples are generated using the MADGRAPH5\_aMC@NLO and POWHEG v2 generators at parton level. In all cases, these parton level events are hadronized using PYTHIA 8 with the CUETP8M1 tune, and then passed to GEANT4 for simulation of the CMS detector response. In the same chapter, we describe the reconstruction and identification of various physics objects.

In **chapter 6**, we describe various corrections applied on simulated samples. In **chapter 7**, the event selection has been described. In the event topology of interest, there are four jets (two b jets and two light jets), one charged lepton, and missing transverse energy. Various selection requirements are applied to ensure the resulting events have this topology.

In **chapter 8**, we perform kinematic fitting to select events coming from true  $t\bar{t}$  decay. In this analysis, the charged Higgs boson is assumed to decay to  $c\bar{s}$ . The invariant mass of the system of the two light jets ( $m_{jj}$ ) is thus used as the final observable. If the two observed light jets come from a semileptonic  $t\bar{t}$  decay, then the  $m_{jj}$  distribution should have a peak at the W boson mass. However, the observed mean of the  $m_{jj}$  distribution is much higher (around 128 GeV), reflecting the fact that the two light jets in each event may not necessarily come from the decay of a W boson. To select true semileptonic  $t\bar{t}$  events, a kinematic fit is performed on the reconstructed objects using the top quark kinematic fitter package. In the output, the top quark kinematic fitter gives exactly four jets (two b jets, one from each of the leptonic and hadronic t decays, and two light jets from the hadronic t decay), a lepton, and a neutrino. The two light jets coming from the hadronic t decay are further used for charm tagging.

In **chapter 9**, we describe the procedure to estimate QCD multijet background from data. The simulation of QCD multijet events is computationally intensive, resulting in a limited number of such events being available. A data-driven approach is used to make a more precise estimation of the QCD multijet background.

In **chapter 10**, the  $m_{jj}$  distribution without and with charm quark tagging is presented. Further, events are divided exclusively into loose, medium, and tight categories, based on whether at least one of the light jets passes the loose but neither passes the medium, at least one passes the medium but neither passes the tight, or at least one passes the tight working points of the charm tagging selection requirements, respectively. The expected signal to background ratio is different in the various charm categories, so partitioning the events into categories results in an improvement in the expected upper limits on  $\mathcal{B}(t \rightarrow H^+ b)$ .

In **chapter 11**, a detailed description of the statistical and systematic uncertainties is given. There are various sources of systematic uncertainty which may arise due to detector calibration effects, uncertainty in the measured reconstructed efficiency, the theoretical modeling of signal events, and other effects.

In **chapter 12**, the final results are presented. The total expected background number of events agrees with the data within uncertainties. The absence of a charged Higgs signal in the data is characterized by setting exclusion limits on the branching ratio  $\mathcal{B}(t \rightarrow H^+ b)$ , assuming that  $\mathcal{B}(H^+ \rightarrow c\bar{s}) = 100\%$ . In the absence of any excess, an asymptotic 95% confidence level (CL) limit using the likelihood ratios on  $\mathcal{B}(t \rightarrow H^+ b)$  is calculated.

In **chapter 13**, we conclude the analysis. We also discuss the future aspects of the analysis and how the current experimental results can be interpreted in different types of the two Higgs doublet model.

# Publications

During the first two years of my Ph.D, I worked in the area of theoretical high energy physics. The work during this period has resulted in three publications (first three entries in the list below). Thereafter, I switched to experiments and became a member of CMS collaboration, LHC. Being a member of this collaboration, I am one of the co-authors of more than 160 papers by now (August 2019). The list of all the papers can be found at [1]. However, the papers where I have contributed directly (enlisted below) and the work due to which I got the authorship in the rest of the papers published by the CMS collaboration is described in the next points.

- **Direct contribution:**

1. **ELKO fermions as dark matter candidates**

B. Agarwal, P. Jain, S. Mitra, A. C. Nayak and R. K. Verma.

arXiv:1407.0797 [hep-ph]

DOI:10.1103/PhysRevD.92.075027

Phys. Rev. D **92**, 075027 (2015)

2. **Effect of VSR invariant Chern-Simon Lagrangian on photon polarization**

A. C. Nayak, R. K. Verma and P. Jain.

arXiv:1504.04921 [hep-ph]

DOI:10.1088/1475-7516/2015/07/031

JCAP **1507**, no. 07, 031 (2015)

3. **The top threshold effect in the  $\gamma\gamma$  production at the LHC**

S. R. Dugad, P. Jain, S. Mitra, P. Sanyal and R. K. Verma.

arXiv:1605.07360 [hep-ph]

DOI:10.1140/epjc/s10052-018-6188-z

Eur. Phys. J. C **78**, no. 9, 715 (2018)

4. CMS Collaboration, "Search for a light charged Higgs boson in the  $H^\pm \rightarrow cs$  channel at 13 TeV

CMS Physics Analysis Summary [CMS-PAS-HIG-18-021](#), 2019

In collaboration wide review, [AN-18-061](#), [HIG-18-021](#)

5. **Search for an excited lepton in the lepton + fatjet final state at 13 TeV, in the CMS experiment at the LHC**

S.B. Beri, K. Hoepfner, S. Dutt, S. Thakur, R.K. Verma

CMS AN-18-126

- **Indirect contribution:** As a policy of the CMS experiment, a person automatically qualifies to become an author of a paper if he/she participates in the data taking process. The data taking is a very extensive process which requires a huge man power to operate and monitor the CMS detector on a daily basis. Therefore, each published paper from the CMS experiment has around 2500 authors because each author has contributed in some aspect of the data taking. Every year each person has to do a *service task* for 4 months to become an author of the papers published in the next year.

On the management front, people are assigned a convener post for handling the responsibility of the given task. The hierarchy is like this - around ten level-3 (L3) convenors report to their level-2 (L2) convener and around ten level-2 (L2) convenors report to the level-1 (L1) convener assigned to them.

As part of my service task, I served as an L3 convener in the alignment calibration and database (AlCaDB) sub-group of the detector performance group (DPG). The detector alignment constants (a set of detector conditions) are derived on a daily or weekly basis to calibrate the simulated data so that it matches with the observed data. To assure this, a dedicated validation is performed for every new set of constants. As part of the validation process, many plots have to be scrutinized to make sure that these constants do not affect the performance of the detector. Once these constants are validated, they are stored in the database for future usage. Below, I have provided a brief summary of my works as an L3 convener.

- **As an L3 convener (2016-18) for the *workflow submission and management*:** I served as an L3 convener from December 2016 to December 2018 for the workflow submission and management with my co-convener Bajrang (TIFR, Mumbai) (from December 2016 to December 2017) and Pritam (IIT, Madras) (from December 2017 to 2018). We performed “tag” validations by submitting workflows and monitoring the DQM (data quality monitoring) plots. A “tag” is a C++ string which contains condition data of the CMS experiment. Whenever there is any change in any sub-detector of the CMS, a “new tag” is created to accommodate the change. Our task was to validate whether the “new tag” is correct or not by looking various DQM plots. The DQM plots are created when we submit workflows corresponding

to the “new tag”. We worked on the submission and validation of 687 (672) workflows in 2017 (2018). The alignment conditions in these workflows were derived from various sub-groups such as tracker, ECAL, HCAL, and muon chambers of the CMS experiment. We interacted with around 100 people from different sub-groups during the validation process. In each year, I got the service task credit for four months.

From June 2017 to June 2018, I also worked on another service task in the AlCaDB group with Amey. We worked on a software package called the LHCInfo. It fetches LHC “Fill” information from an online database and stores them in the offline database. It was originally developed by Salvatore. Various new variables such as lumi per bunch crossings, `lhccState`, `lhccComments`, `ctppsStatus`, `lumiSection` etc, were not present in the old package developed by Salvatore. Therefore, we upgraded this package so that it can fetch many new variables related to the LHC Fill. The LHCInfo O2O was developed, commissioned and deployed in early May 2018. Since its deployment, approximately 0.44 M payloads have been populated in the condition database as of December 2018. The payloads for each Fill (with stable beam) are being populated with lumi-section granularity (that is after every 23 seconds) in the database. With this, one can access the relevant information of any LHC Fill within the CMSSW jobs which helps in correlating the beam information with sub-detector needs. In 2018, I got 2 months of service task credit for this.

- **An an L3 convener (2019-21) for the *AlCa-TSG contact*:** After my tenure as an L3 convener for *workflow submission and management* ended, I am recommended for new position in the AlCaDB group as AlCa-TSG (trigger study group) contact. This term is from Jan 2019 to Jan 2021. Our (mine and Ashish) job is to present AlCaDB interest in the TSG group and vice versa. Currently, we have been assigned three tasks. The first task is to include a lot of AlCa conditions directly in the global tag for the offline data. The second task is to list Alca conditions consumed in the latest CMSSW which are needed at different steps such as GEN, SIM, DIGI, HLT, and RECO. The third task is to automate update or switch off the trigger bits.

During the last three years of the service task in the AlCaDB group, I worked under the guidance and leadership of five outstanding L2 convenors - Giovanni Franzoni, Giacomo Govi, Arun Kumar, Luca Pernie, and Tongguang Cheng.



# Dedication

To my late Fufa (*Santsaran Verma*), Dadi (*Hansha Devi*), Guru (*Mahendra Pratap Singh*), and Mitra (*Siddhant Singh*) who laid the foundation this thesis is built upon.



# Acknowledgments

**A**dream came true for me. A few years back in 2008, when the LHC was about to collide the two oppositely moving proton for the first time in history, there was an avalanche of news around every corner of the world about it. There were misconceptions that during the collision, a microscopic black hole will be created which will start sucking the matter around it and eventually the whole planet earth after some time. As a student of physical science in 12th standard, I was very curious to know more about the LHC in those days. There was one edition of *Vigyan Pragati*, a science magazine in Hindi language, about the physics activities happening at the LHC. After reading that, I became even more curious about the LHC. However, at that time, I had never imagined that in the future, I would end up doing a Ph.D. in the LHC! The journey from one of the villages in the north India to one of the world's most sophisticated experimental setups in Switzerland has been filled with a lot of ups and downs. There have been many people who helped me directly or indirectly during this journey. Here I want to acknowledge a few person for their support during this long journey.

- **12th:** The first step of this journey started in 12th standard in my school (*Pioneer Montessori Inter College*) where the foundation was laid. My ideal teacher (*Mahendra Pratap Singh*) taught us physics and mathematics. He is not only a great teacher but also a great motivator. I am fortunate enough to have many school friends *Siddhant, Ashutosh, Mayank, Shahbaz, Sandeep, Prashant, Shailesh*, etc, who made my school life memorable.
- **B.Sc.:** My best friend *Siddhant* guided me and my late Fufa (*Santsaran*) financially supported me to get admission in *Kamla Nehru Institute of Physical and Social Sciences* into *B.Sc.* program. I got the opportunity to learn college level physics and mathematics there. I am indebted to my professors - *Yogendra Bahadur Singh, Pankaj Singh, Prashant Singh, Lalit Kumar Divedi, Jaysnath Mishra* who cultivated my understanding of physics and mathematics. During these years, I made a lot of friends - *Ravi, Prem, Anil, Zeeshan, Chand, Sagar, Deepoo, Syamoo, Arunesh, Mayank, and Satyajit* and they made the college life super enjoyable. I sincerely thank them for being with me when I needed them the most.

- **M.Sc.:** I was fortunate enough to get admission in the M.Sc.-Ph.D. dual degree program for physics in *Indian Institute of Technology, Kanpur* (IITK) which is one of the best institutes for higher education in India. During the M.Sc. program, I got the opportunity to learn from the advanced physics courses taught by the leading physicist at IITK. I enjoyed the courses taught by *Deshdeep Sahdev, Manoj Kumar Harbola, Pankaj Jain, Debashish Chowdhury, and Joydeep Chakraborty*. During the M.Sc. program, I got the opportunity to attend classes with the best minds of India. I am thankful to my friends *Vimalesh, Sangha Mitra, Chitrasen, Dhananjay, and Santosh* for making IITK life memorable. I enjoyed the discussions/debates with my best friend Vimalesh on a wide range of issues ranging from science, society, politics, philosophy, sports, etc. We spent many hours on these issues and ended most of the time without reaching to any logical conclusion.
- **Ph.D.:** In the first two years of my Ph.D. program, I pursued my research in the area of theoretical high energy physics working on the dark matter, very special relativity (VSR), and top-quark threshold effect. I am very grateful to my senior *Alekhya* for helping and correcting me at many points. I also thank *Prasenjit* and *Subhadip* for many theoretical discussions. In my third year of Ph.D., I started my research in the experimental high energy physics and joined Prof. Shashi Dugad at *Tata Institute of Fundamental Research* as a junior research fellow in the fourth year. During the transition from theoretical to the experimental field, I met one of the most beautiful girl and my one of the best friends, *Swyamsree Patra*. On academic leave from IITK, I deputed to TIFR in May 2016 to explore the LHC world. I will always be indebted to my supervisors *Pankaj Jain* and *Shashi Dugad* for giving me such an excellent opportunity.

It is beyond words to describe the experience that I got while working at TIFR and CERN. I am very thankful to *Gouranga* and *Arun Nayak* for guiding me at every technical detail. Having a physics discussion with *Gagan* and *Shashi* was very useful. I am also grateful to *Giovanni Franzoni, Giacomo Govi, Arun Kumar, Luca, and Tongguang* for giving me the opportunity to work as a level-3 convener in the AlCaDB (alignment calibration and database) group. I also benefited from the discussion with Higgs conveners (*Anne-Marie, Abdollah, Giacomo Ortona, Andra, Giovanni Petrucciani, and Maria Cepeda*). I will always be grateful to *Bajarang* who was senior to me but always treated me like a younger brother. He helped me a lot on personal and professional front in the initial days at TIFR. A very special thank to him. During this period, I got a few amazing friends - *Muzamil, Bibhu, Raghu, Irfan, Sushil, Amey, Manish, Alibordi, Shalini, Akshansh, Pritam, Ashish, etc.* I also would like to thank *Brij* and *Puneet* for being available to fix technical issues related to computing infrastructures, specifically providing condor facility for running parallel jobs.

My special thanks to Prof. Pankaj Jain who is one my ideals and a living legend. Working under him in the past 6 years was quite rewarding on personal as well as professional fronts. I would also thank Prof. Shashi Dugad with my folded hand for always with me during the last 3 years. For him, I would use the phrase *A silent sea never made a skilled sailor*. After working under him, I feel trained enough to sail in any sea.

I am also grateful to the administration members at Department of Physics (IITK) and Department of High Energy Physics (TIFR) for assisting me with various paperworks. I am also thankful to MHRD and DAE for financial support during my PhD at IITK and TIFR respectively.

Finally, I would like to thank my parents for giving me the freedom to choose the academic carrier as per my wish. I also thank the other members of my family and village - *Ramesh, Sachin, Manju, Parvind, Dipak, Akash, Vikash, Ankit, Surendra* for being with me since my childhood. I lost my Fufa (*Santsaran*) in a car accident on February 10, 2010. He was a father-figure for me. He was one of the people who motivated and helped me when I needed the most. Had he been alive, he would have been very happy to see the fruits of the tree, he had planted a decade ago. At the end, I would like to thank my fiance (*Swayamsree Patra*) for taking my care and loving me.

RAVINDRA KUMAR VERMA  
IIT KANPUR  
UTTAR PRADESH, 208016



असतो मा सद्गमय । तमसो मा ज्योर्तिगमय । मृत्योर्मा अमृतं गमय ।

*asato mā sadgamaya, tamaso mā jyotirgamaya, mr̥tyormā amṛtam̄ gamaya*

**Translation:** Take me from falsity to truth, from darkness to light, and from death to immortality. BRĀHAD ĀRAṄYAKA UPANIṢAD 1.3.28, ŠUKLA YAJUR VEDA



# Abbreviations

**ALICE** A Large Ion Collider Experiment

**ATLAS** A Toroidal LHC ApparatuS

**BTV** b tagging and Vertexing

**CERN** European Organisation for Nuclear Research

**CL** Confidence Level

**CMS** Compact Muon Solenoid

**CR** Control Region

**CSC** Cathode Strip Chamber

**CSV** Combined Secondary Vertex

**DT** Drift Tube

**ECAL** Electromagnetic Calorimeter

**HB** Hadron Barrel

**HCAL** Hadron Calorimeter

**HE** Hadron Endcap

**HF** Hadron Forward

**HLT** High Level Trigger

**HO** Hadron Outer

**HPD** Hybrid Photo Diode

**IITK** Indian Institute of Technology, Kanpur

**JER** Jet Energy Resolution

**JES** Jet Energy Scale

**LHC** Large Hadron Collider

**LHCb** Large Hadron Collider beauty

**LHCf** Large Hadron Collider forward

**MC** Monte Carlo

**MET** Missing Transverse Energy

**MoEDAL** Monopole and Exotics Detector at the LHC

**MSSM** Minimal Supersymmetric Standard Model

**NLO** Next to Leading Order

**NP** Nuisance Parameter

**PDF** Probability Distribution Function

**PF** Particle Flow

**POG** Physics Object Group

**PU** Pileup

**PV** Primary Vertex

**QCD** Quantum Chromodynamics

**RPC** Resistive Plate Chamber

**SF** Scale Factor

**SiPM** Silicon Photo Multiplier

**SM** Standard Model

**SR** Signal Region

**SUSY** Supersymmetry

**SU** Special Unitary

**TEC** Tracker End Cap

**TIB** Tracker Inner Barrel

**TID** Tracker Inner Disk

**TIFR** Tata Institute of Fundamental Research

**2HDM** Two Higgs Doublet Model

**TOB** Tracker Outer Barrel

**TOTEM** TOTal Elastic and diffractive cross section Measurement

**WP** Working Point



# Contents

<b>Certificate</b>	i
<b>Synopsis</b>	iii
<b>Publications</b>	vii
<b>Dedication</b>	xi
<b>Acknowledgments</b>	xiii
<b>Abbreviations</b>	xix
<b>I. The Standard Model and Beyond</b>	1
<b>1. The Standard Model</b>	3
1.1. Introduction . . . . .	3
1.2. Group and symmetry . . . . .	4
1.3. The Lagrangian density . . . . .	8
1.4. Mass of particles . . . . .	11
1.5. Success and failure of the SM . . . . .	12
<b>2. The Two Higgs Doublet Model</b>	15
2.1. Introduction . . . . .	15
2.2. The Higgs sector . . . . .	15
2.3. The Yukawa sector . . . . .	16
2.4. Branching ratio of top quark decay . . . . .	17
2.5. Branching ratio of charged Higgs decay . . . . .	18
2.6. Current status of searches for charged Higgs . . . . .	20

2.7. Search strategy for the charged Higgs at 13 TeV . . . . .	22
<b>II. The CMS Experiment at The LHC, CERN</b>	<b>31</b>
<b>3. The LHC at CERN</b>	<b>33</b>
3.1. Introduction . . . . .	33
3.2. Detectors at the LHC . . . . .	35
3.3. The LHC parameters . . . . .	37
3.4. The LHC coordinate system . . . . .	43
<b>4. The CMS Experiment at The LHC</b>	<b>47</b>
4.1. Introduction . . . . .	47
4.2. Silicon tracker . . . . .	50
4.3. Electromagnetic calorimeter . . . . .	53
4.4. Hadron calorimeter . . . . .	58
4.5. Magnet . . . . .	62
4.6. Muon chambers . . . . .	64
4.7. The trigger . . . . .	68
4.8. Data distribution system . . . . .	70
4.9. Event data format and the CMS software . . . . .	72
<b>III. Search for Charged Higgs Boson</b>	<b>75</b>
<b>5. Reconstruction and Identification</b>	<b>77</b>
5.1. Data samples . . . . .	77
5.2. Simulated samples . . . . .	77
5.3. Primary vertex and diffuse offset energy density . . . . .	79
5.3.1. Primary vertex . . . . .	79
5.3.2. Diffuse offset energy density . . . . .	81
5.4. Muon . . . . .	82
5.4.1. Reconstruction . . . . .	82
5.4.2. Identification and selection . . . . .	84
5.4.3. Rochester correction . . . . .	84
5.4.4. Isolation . . . . .	86

5.5.	Electron	86
5.5.1.	Reconstruction	86
5.5.2.	Identification and selection	87
5.5.3.	Isolation	88
5.5.4.	Conversion rejection	88
5.6.	Jet	89
5.6.1.	Reconstruction	89
5.6.2.	Identification and selection	89
5.6.3.	Cleaning	90
5.7.	Missing transverse energy	90
5.8.	b Jet Tagging	91
5.8.1.	Track selection	91
5.8.2.	Reconstruction of secondary vertex	92
5.8.3.	Multi-layer perceptron training	92
<b>6.</b>	<b>Corrections on Simulated Events</b>	<b>95</b>
6.1.	Luminosity scale factors	95
6.2.	Pileup reweighting	95
6.3.	Lepton scale factors	95
6.4.	Jet and $E_T^{\text{miss}}$ correction	96
6.5.	b tag scale factor	96
<b>7.</b>	<b>Event Filter and Selection</b>	<b>99</b>
7.1.	Filters	99
7.2.	Selection	100
7.3.	Comparison of data and background after b jet selection	102
<b>8.</b>	<b>Kinematic Fitting and c Jet Tagging</b>	<b>109</b>
8.1.	Input to the TopKinFitter	109
8.2.	Separation of jets	111
8.3.	Kinematic constraints	112
8.4.	$\chi^2$ of the fit	112
8.5.	Output of the TopKinFitter	113
8.6.	Selection and performance	113

8.7. Comparison of data and background after kinematic fitting	114
8.8. $c$ Jet Tagging	114
8.9. $c$ tag scale factor and selection	118
<b>9. QCD Multijet Estimation</b>	<b>121</b>
9.1. Determination of transfer scale factor	122
9.2. Estimation of QCD multijet in the signal region	122
9.3. Comparison of data-driven QCD multijet shapes	126
<b>10. Mass Reconstruction</b>	<b>129</b>
10.1. Inclusive event category without charm tagging	131
10.2. Inclusive event category with loose charm tagging	132
10.3. Exclusive event categories based on charm tagging	132
<b>11. Systematic Uncertainties</b>	<b>139</b>
11.1. Experimental uncertainties	139
11.2. Theoretical uncertainties	143
11.3. Effect of systematics on $m_{jj}$ distribution	146
<b>12. Exclusion Limits</b>	<b>153</b>
12.1. 95% CL upper limit on $\mathcal{B}(t \rightarrow H^+ b)$	153
12.2. Profile of nuisance parameter: shape vs $\ln N$	154
12.3. Limits from inclusive event category without $c$ tagging	156
12.4. Limits from inclusive event category with loose $c$ tagging	156
12.5. Limits from exclusive event categories based on $c$ tagging	159
12.6. Impact of nuisance parameters	159
12.7. Goodness of fit	164
<b>13. Summary and Future Outlook</b>	<b>173</b>
<b>Bibliography</b>	<b>176</b>

# List of Figures

1.	Production of $t\bar{t}$ from gluon-gluon scattering. The left plot shows the SM decay of the $t\bar{t}$ pair in the semileptonic decay channel. The right plot shows the signal process in which the $t\bar{t}$ pair decay products include a charged Higgs boson. . . . .	v
1.1.	A schematic picture showing the hierarchy of building blocks of the matter when one looks deeper and deeper, from right to left. This figure is adopted from [2]. . . . .	4
1.2.	Physical properties of the fundamental particles [3]. There are three generations of quarks and leptons grouped in the increasing order of mass. The gluon, photon, Z/W, and Higgs bosons are force carriers. . . . .	5
1.3.	A timeline of the fundamental particles from their theoretical prediction to the experimental observation. . . . .	6
1.4.	A schematic diagram showing the interactions of fundamental particles among themselves [4].	6
2.1.	Variation of $\mathcal{B}(t \rightarrow H^+ b)$ as a function of the mass of the charged Higgs (left side plots) and $\tan\beta$ (right side plots). . . . .	19
2.2.	The branching ratio of charged Higgs decay in different channels as a function of $\tan\beta$ for $m_{H^+} = 150$ GeV in different types of the 2HDM. These plots are taken from [5]. . . . .	21
2.3.	Current allowed range at 95% CL from ATLAS and CMS. . . . .	24
2.4.	Production of $t\bar{t}$ from gluon-gluon and quark-quark scattering. The quark-scattering production process has a dominant contribution at Tevatron energies whereas gluon-gluon scattering diagrams are dominant at LHC energies [6, 7]. The SM production of $t\bar{t}$ is shown in (a), (b) and (c). The charged Higgs boson production and its decay are shown in (d), (e), and (f). . . . .	26
2.5.	Representative Feynman diagrams for single t quark production processes. . . . .	27
2.6.	Representative Feynman diagrams for $W + n$ jets channel. The $W$ boson is produced by quark-quark and quark-gluon scattering along with $n$ jets ( $n = 0, 1, 2, 3, 4$ ). . . . .	28

2.7. Representative Feynman diagrams for $Z/\gamma + n$ jets channel. The $Z/\gamma$ is produced by quark-quark and quark-gluon scattering along with $n$ jets ( $n = 0, 1, 2, 3, 4$ ). . . . .	29
3.1. The CERN accelerator complex [8]. The particles are accelerated at various stages, starting from LINAC to PSB to PS to SPS to LHC. Various detectors such as ATLAS, LHCb, CMS, and ALICE are placed at the collision points. . . . .	34
3.2. The layout of insertion regions (IRs) at the LHC ring [9]. The separation between the two colliding beams is not to scale. The separation between them is exaggerated for illustration. The total circumference of the LHC ring is 27 km. Oppositely moving beams (Beam 1 and 2) are made to collide at four points of the ring. At the collision points, detectors such as CMS, ALICE, ATLAS, and LHCb are installed to record the collision. . . . .	35
3.3. Schematic diagram showing quadrupole and dipole magnets for focussing and bending proton beams at the interaction point (IP). The D1 dipole separate beams from both sides of the IP. This figure is adopted from [10]. . . . .	36
3.4. The bunch configuration and crossing angle. The bunches are separated by 7.5 m. The crossing angle shown is exaggerated for illustration. Figure 3.4(a) is adopted from [11] and Figure 3.4(b) is taken from [12]. . . . .	38
3.5. Two oppositely moving bunches with the total number of particles $N_1$ and $N_2$ , and beam density distribution functions $\rho_1(x, y, s, -s_0)$ and $\rho_2(x, y, s, s_0)$ [13]. . . . .	39
3.6. Luminosity measurement from different years of the data taking [14]. In this thesis, the analysis presented uses the one for the 2016 data taking (golden yellow line in the right plot). The total delivered luminosity at 13 TeV from all years is about $160\text{fb}^{-1}$ . . . . .	41
3.7. Variation of few LHC parameters during the data taking on 22-23rd October 2018 (Fill = 7333). All these parameters change over a different point of time. Other parameters such as the center-of-mass energy, number of particles in each beam, etc do not change on a given day. . . . .	42
3.8. Variation of few LHC parameters during the data taking from random days in the period of June-November 2018. Only a few days (Fills) have been chosen for illustration. The last two Fills (7433 and 7477) correspond to Pb-Pb, rest are from proton-proton collisions. The average intensity of the beam is the number of particles in each beam. As can be seen, all these parameters change over a different period of time. . . . .	42
3.9. The LHC coordinate system [15] with the $z$ -axis lying along the beam direction. At CMS, the $z$ -axis points towards the mountain Jura in France. . . . .	43

3.10. The relationship between $\theta$ and $\eta$ [16]. The value of $\eta = 0$ implies $\theta = \pi/2$ . That is, lower the value of $\eta$ the farther away we are from the beam axis in the perpendicular direction and vice versa. . . . .	44
4.1. The CMS detector installed in IR5 at the LHC [17]. . . . .	48
4.2. Diagram of the CMS detector [18]. Various parts of the CMS detector are shown starting from tracker to ECAL to HCAL to magnet to the muon chambers. . . . .	48
4.3. Trajectories of various particles inside the CMS detector [19]. The charged particles such as electrons, muons, and charged pions ( $\pi^\pm$ ) are bent inside and outside the solenoid whereas the neutral particles such as photons and neutral pions ( $\pi^0$ ) traverse without being bent. The silicon tracker is mainly for measuring the momentum of particles whereas the ECAL and HCAL measure the energy deposits. The electrons and photons deposit their energy inside the ECAL whereas the hadrons deposit in both ECAL and HCAL. The muons are the only particles that traverses all the way to the muon chambers. . . . .	49
4.4. The silicon tracker of the CMS. The picture of one of trackers inner barrel is shown in (a). The radial and $\eta$ position of the various parts of the pixel and strip tracker is shown in (b). . . . .	51
4.5. The principle of operation of a silicon detector [20]. The green strips are $p^+$ -type of semiconductor where the electrons are attracted after being produced in the depletion region when a particle passes through. The $\text{SiO}_2$ layer along with metallic surface, on top of the green strips, provides conducting circuits for the electron. The amplifiers/shapers are used to amplify and shape the electron current. The holes are attracted towards the +ve bias voltage. . . . .	52
4.6. The pixel detector of the CMS tracker. The location of pixel layers in the barrel and endcap region is shown in (b). . . . .	52
4.7. Total thickness of the silicon tracker in units of radiation and interaction lengths [21]. The thickness of supporting tube and beam pipe is also shown. . . . .	53
4.8. The absolute resolution of $\phi$ and relative resolution of $p_T$ measurement of muons, electrons, and charged pions from the silicon tracker as function of $\eta$ [21]. Other resolution such as $d_0, z_0$ , and $\cot\theta$ are given in Reference [21]. The open (solid) markers correspond to the half-width for 90% (68%) intervals [21]. . . . .	54
4.9. (a) A $\text{PbWO}_4$ scintillating crystal. (b) The various components of the electromagnetic calorimeter in the barrel and endcap regions. . . . .	55

4.10. The energy band structure in pure (impure) inorganic scintillator shown on left (right) side. The impurity creates intermediate excited states between the valence and conduction band. The photons are emitted when activator de-excites. . . . .	56
4.11. The schematic diagram of a circular and cylindrical PMT (a). The principle of operation of an APD is shown in (b). . . . .	56
4.12. ECAL preshower [22] placed in the endcap region at $z = 298.5$ cm. The preshower is made of alternate layers starting from the moderator, to heating film to foam to cooling block to the absorber (Pb) to silicon detectors to the digital electronics. The width of the lead (Pb) absorber is not the same on both sides. . . . .	57
4.13. The relative energy resolution of electron and photon as a function of supercluster $\eta$ in the barrel and endcap region of the ECAL [23]. . . . .	59
4.14. The CMS hadron calorimeter [17]. A prototype of HCAL barrel is shown in silver color in (a). The actual endcap part of the HCAL is shown in (b). . . . .	60
4.15. The energy levels of a molecule of an organic scintillator [24]. The ground state is $S_0$ followed by vibration sublevels ( $S_{01}, S_{02}$ etc). The $S_{1,2,3}$ are the excited singlet energy levels. In between the singlet, there are triplet energy levels shown as $T_{1,2,3}$ . The fluorescence (phosphorescence) occurs when there is a transition from $S_1(T_1) \rightarrow S_{03,02,01,00}$ . . . . .	60
4.16. The location of various components (HB, HO, HE, and HF) of the HCAL [4]. . . . .	61
4.17. The HCAL energy resolution as a function of pion energy measured using hybrid photodi- odes, and silicon photomultipliers with different pixel density [25]. . . . .	62
4.18. An artistic view of the CMS solenoid [26]. An aluminum alloy shell, with a thickness of 50 mm, is placed on the inner and outer surface of the solenoid and used as a cooling wall. Outside the solenoid are the return iron yokes which hold the magnetic flux. . . . .	63
4.19. Component of the magnetic field inside CMS [16]. . . . .	64
4.20. Location of the various sub-detectors of the CMS [27]. The DTs and CSCs are placed in barrel and endcap regions. The RPCs are placed in both regions. The muon chamber in the barrel region is formed by placing three “wheels” side-by-side along the $z$ -axis whereas in the endcap regions different segments of RPC and CSC are placed along the radial direction. . . . .	65
4.21. Schematic view of one of the DTs of the muon chamber [17]. The red dots are the aluminum wires pointing into the page, the red arrow is the direction of an incoming muon, and the horizontal blue lines are the distance traveled by secondary electrons. . . . .	66

4.22. Schematic view of one of the cathode strip chambers [17]. The wires shown in red and cathode strips shown in cyan are placed perpendicular to each other. . . . .	67
4.23. Schematic diagram of one of the RPCs. The cathode (bottom red plate) and anode (top red plate) is separated by a separator (shown in gray). The region between the two plates is filled with gas. The anode is transparent to the secondary electrons. The orange strips (above anode) picks the secondary electrons. The high voltage aluminum foil is placed at the bottom. This figure is adopted from [17]. . . . .	67
4.24. Workflow of decisions for the Level-1 trigger [28]. . . . .	69
4.25. Muon and electron trigger efficiency as a function of $p_T$ . The trigger turn-on is at about $p_T = 24$ GeV for muon, and 33 GeV for electron trigger. . . . .	71
4.26. Location of different data acquisition systems [29]. The T0 is located at CERN, T1s are located in 7 different places, and 55 T2s (although only 22 are shown in this figure) are located at various countries. There is only one T2 in India which is located at TIFR, Mumbai. . . . .	72
4.27. The flow of collision data among different tiers. [29] . . . . .	73
4.28. A schematic diagram of the event data model [30]. The <i>pool source</i> is a package that provides the input data for every event. The <i>digitizer</i> and <i>tracker</i> are the EDProducer which takes input from the event, do some operation and write the output to the same event. The <i>nTrack filter</i> is the EDFilter which takes input from the event and decides if that event is to be kept or rejected. If the event is to be rejected then the subsequent processing is halted and the CMSSW start looking at the next event through <i>pool source</i> . The <i>pool output</i> package stores the output after all the operations are performed. . . . .	74
5.1. Distribution of number of primary vertices and diffuse offset energy density for $\mu + \text{jets}$ and $e + \text{jets}$ channels. The agreement between data and simulation is better for $\rho$ as compared to $N^{\text{vertex}}$ distribution. However, for lower values of $\rho$ , there is a significant mismatch between data and simulation. Similar trend is seen in other analysis as well [31, 32]. . . . .	83
5.2. The distribution of efficiencies for medium muon ID and tight relative isolation (with medium ID) as a function of $p_T$ and $\eta$ of muon from simulation and data from different eras of 2016 data taking. . . . .	85
5.3. The distribution of efficiency for medium and veto electron ID. . . . .	88
5.4. Primary and secondary vertex with tracks and impact parameter. $\text{IP} > 0$ or $< 0$ if tracks are upstream or downstream along the jet direction. This figure is taken from [33]. . . . .	91

5.5. The b discriminator distributions of all jets, after $N_{jets} \geq 4$ selection as described in Section 7.2, obtained using CSVv2 method for the $\mu +$ jets and $e +$ jets channel. . . . .	93
7.1. Event yields after various selection cuts for the $\mu +$ jets and $e +$ jets channel. . . . .	102
7.2. Distribution of $p_T, \eta$ of reconstructed lepton and $p_T$ of jets after b jet selection as described in Section 7.2, for $\mu +$ jets and $e +$ jets channel. . . . .	105
7.3. Distribution of reconstructed $\eta$ of jets, jet multiplicity, and b jet multiplicity after b jet selection as described in Section 7.2, for $\mu +$ jets and $e +$ jets channel. . . . .	106
7.4. Distribution of reconstructed $E_T^{\text{miss}}$ and $m_{W^+}^T$ (transverse mass of W boson decaying leptonically) after b jet selection as described in Section 7.2, for $\mu +$ jets and $e +$ jets channel. . . .	107
8.1. $m_{jj}$ distributions of two non b, highest $p_T$ jets for the $\mu +$ jets and $e +$ jets channel. The distributions of Figures 8.1(a) and 8.1(b) are obtained using reconstructed jets after applying b tag scale factor as described in Section 7.2. On the other hand, the distributions of Figures 8.1(c) and 8.1(d) are calculated using kinematic fitted jets after kinematic fit selection. The mean of the invariant mass distribution from kinematic fitted jets is closer to the W mass as compared to that of reconstructed jets. . . . .	110
8.2. Distribution of $p_T, \eta$ of kinematic fitted lepton and $p_T$ of jets after kinematic fit selection for $\mu +$ jets and $e +$ jets channel. . . . .	115
8.3. Distribution of kinematic fitted $\eta$ of jets, jet multiplicity, and b jet multiplicity after kinematic fit selection for $\mu +$ jets and $e +$ jets channel. . . . .	116
8.4. Distribution of kinematic fitted $E_T^{\text{miss}}$ and $m_{W^+}^T$ after kinematic fit selection for $\mu +$ jets and $e +$ jets channel. . . . .	117
8.5. The 2D plot between the two c taggers. The region right to the vertical and above to the horizontal line correspond to different charm working points (WPs). Unlike the b tag WPs, there is an overlap between the loose, medium, and tight c tag WPs. This figure is taken from [34]. . . . .	118
8.6. The c discriminator distribution of two non b jets after kinematic fitting selection for $\mu +$ jets and $e +$ jets channel. . . . .	119
9.1. Distribution of relative isolation and missing transverse energy after lepton trigger selection for $\mu +$ jets and $e +$ jets channel. The QCD multijet process is dominated in the high $I_{\text{rel}}$ region. . . . .	123

9.2. Distribution of mean of $I_{\text{rel}}$ in different $E_{\text{T}}^{\text{miss}}$ bins after lepton trigger selection for $\mu + \text{jets}$ and $e + \text{jets}$ channel. There seems to be some correlation between $E_{\text{T}}^{\text{miss}}$ and $I_{\text{rel}}$ in the low $E_{\text{T}}^{\text{miss}} < 50$ GeV region for data as shown in Figures 9.2(a), 9.2(b). However, from simulated QCD multijet events, the $E_{\text{T}}^{\text{miss}}$ and $I_{\text{rel}}$ are uncorrelated, as shown in Figures 9.2(c), 9.2(d). . . . .	124
9.3. Different regions between $I_{\text{rel}}$ and MET (missing transverse energy also known as $E_{\text{T}}^{\text{miss}}$ ) to estimate QCD multijet events from data. For $\mu + \text{jets}$ channel, the isolation region corresponds to $I_{\text{rel}}^{\mu} < 0.15$ whereas the non-isolation corresponds to $0.15 < I_{\text{rel}}^{\mu} < 0.40$ . For $e + \text{jets}$ channel, the isolation region corresponds to $I_{\text{rel}}^e < 0.0821(0.0695)$ whereas the non-isolation region corresponds to $0.0821 < I_{\text{rel}}^e < 0.30 (0.0695 < I_{\text{rel}}^e < 0.30)$ for electron found in barrel (endcap) regions. . . . .	125
9.4. Comparison of data-driven QCD multijet shapes in low $E_{\text{T}}^{\text{miss}}$ region ( $< 20$ GeV), from the isolated and anti-isolated region with reconstructed jets after b jet selection for $\mu + \text{jets}$ and $e + \text{jets}$ channel. . . . .	127
9.5. Comparison of data-driven QCD multijet shapes in low $E_{\text{T}}^{\text{miss}}$ region ( $< 20$ GeV), from the isolated and anti-isolated region with kinematic fitted objects after kinematic fit selection for $\mu + \text{jets}$ and $e + \text{jets}$ channel. . . . .	128
10.1. $m_{\text{jj}}$ distributions of two non b, highest $p_{\text{T}}$ jets from charged Higgs signal samples ( $m_{\text{H}^+} = 80, 90, 100, 120, 140, 150, 155, 160$ GeV) for $\mu + \text{jets}$ and $e + \text{jets}$ channel. The $m_{\text{jj}}$ distributions of Figures 10.1(a), 10.1(b) are evaluated using reconstructed jets after b jet selection. Whereas $m_{\text{jj}}$ distributions of Figures 10.1(c), 10.1(d) are evaluated using kinematic fitted jets after kinematic fit selection. . . . .	130
10.2. Distribution of $m_{\text{jj}}$ for inclusive loose charm tagging as described in Section 10.2 for $\mu + \text{jets}$ and $e + \text{jets}$ channel. . . . .	132
10.3. Signal to background ratio for $m_{\text{H}^+} = 80, 90, 100, 120$ GeV. . . . .	134
10.4. Signal to background ratio for $m_{\text{H}^+} = 140, 150, 155, 160$ GeV. . . . .	135
10.5. Exclusive event categories based on different charm tagging. The events falling in the exclusive “other” category are not considered in this analysis. However, fraction of events falling in this category is very small. . . . .	136
10.6. Distribution of $m_{\text{jj}}$ from exclusive charm categories as described in Section 10.3 for $\mu + \text{jets}$ and $e + \text{jets}$ channel. The signal significance is different across different exclusive categories. . . . .	137

11.1. Correlation of b and c tag scale factors. In the region (a), c quark is tagged as b jet; in region (b), c quark is tagged as c jet; in region (c), b quark is tagged as b jet; in region (d), b quark is tagged as c jet; in region (e), light quark is tagged as b jet; and in region (f), light quark is tagged as c jet. . . . .	140
11.2. The c tag event weight for different cases from $t\bar{t} + \text{jets}$ sample for both channels. For exclusive loose charm category, the value of $CMS_{eff\_ExL} = 1.026/0.9961$ ( $1.025/0.9960$ ) for $\mu + \text{jets}$ ( $e + \text{jets}$ ) channel. For exclusive medium charm category, the value of $CMS_{eff\_ExM} = 1.096/1.046$ ( $1.095/1.046$ ) for $\mu + \text{jets}$ ( $e + \text{jets}$ ) channel. . . . .	142
11.3. Effect of pileup, jet energy scale, and jet energy resolution systematics on the $m_{jj}$ distribution for the charged Higgs signal ( $m_{H^+} = 120$ GeV) for $\mu + \text{jets}$ and $e + \text{jets}$ channel , after $\geq 1$ c jet selection. . . . .	147
11.4. Effect of b/c tagging and t quark $p_T$ systematics on the $m_{jj}$ distribution from charged Higgs signal ( $m_{H^+} = 120$ GeV) for $\mu + \text{jets}$ and $e + \text{jets}$ channel, after $\geq 1$ c jet selection. . . . .	148
11.5. Effect of jet energy scale, jet energy resolution, and b/c tagging systematics on the $m_{jj}$ distribution from $t\bar{t} + \text{jets}$ background for $\mu + \text{jets}$ and $e + \text{jets}$ channel, after $\geq 1$ c jet selection. . . . .	149
11.6. Effect of c tagging, t quark mass, renormalization and factorisation scale, and parton shower matching systematics on the $m_{jj}$ distribution from $t\bar{t} + \text{jets}$ background for $\mu + \text{jets}$ and $e + \text{jets}$ channel, after $\geq 1$ c jet selection. . . . .	150
11.7. Effect of t quark $p_T$ and pileup reweighting systematics on the $m_{jj}$ distribution from $t\bar{t} + \text{jets}$ background for $\mu + \text{jets}$ and $e + \text{jets}$ channel, after $\geq 1$ c jet selection. . . . .	151
12.1. Ratio of up and down with base template for exclusive loose charm category for $\mu + \text{jets}$ channel. . . . .	156
12.2. Ratio of up and down with base template for exclusive medium charm category for $\mu + \text{jets}$ channel. . . . .	157
12.3. Ratio of up and down with base template from exclusive tight charm category for $\mu + \text{jets}$ channel. . . . .	157
12.4. Ratio of up and down with base template from exclusive loose charm category for $e + \text{jets}$ channel. . . . .	158
12.5. Ratio of up and down with base template from exclusive medium charm category for $e + \text{jets}$ channel. . . . .	158
12.6. Ratio of up and down with base template from exclusive tight charm category for $e + \text{jets}$ channel. . . . .	159

12.7. The upper limit in % on $\mathcal{B}(t \rightarrow H^+ b)$ as a function of $m_{H^+}$ using $m_{jj}$ after kinematic fit selection without c tagging, and with loose c tagging, as discussed in Section 10.1 and 10.2, for $\mu +$ jets, $e +$ jets, and $\ell +$ jets channel. . . . .	160
12.8. The upper limit in % on $\mathcal{B}(t \rightarrow H^+ b)$ as a function of $m_{H^+}$ by combining $m_{jj}$ distributions from different exclusive charm categories as discussed in Section 10.3 for $\mu +$ jets $e +$ jets, and $\ell +$ jets channel. . . . .	161
12.9. The upper limit in % on $\mathcal{B}(t \rightarrow H^+ b)$ as a function of $m_{H^+}$ using $m_{jj}$ from all event categories for $\mu +$ jets, $e +$ jets, and $\ell +$ jets channel. . . . .	162
12.10 Distribution of post-fit pulls and impacts of nuisance parameters from exclusive event categories based on charm-tagging for $m_{H^+} = 120$ GeV from $\ell +$ jets channel. Contd ... . . . . .	165
12.11 Distribution of post-fit pulls and impacts of nuisance parameters from exclusive event categories based on charm-tagging for $m_{H^+} = 120$ GeV from $\ell +$ jets channel. Contd ... . . . . .	166
12.12 Distribution of post-fit pulls and impacts of nuisance parameters from exclusive event categories based on charm-tagging for $m_{H^+} = 120$ GeV from $\ell +$ jets channel. Contd ... . . . . .	167
12.13 Distribution of post-fit pulls and impacts of nuisance parameters from exclusive event categories based on charm-tagging for $m_{H^+} = 120$ GeV from $\ell +$ jets channel. . . . .	168
12.14 Goodness of fit for $\ell +$ jets from exclusive charm categories. . . . .	170



# List of Tables

1.1.	The third component of the weak-isospin, hypercharge, and charge quantum numbers for leptons, quarks, and bosons. . . . .	8
1.2.	The mass of all fundamental particles in terms of the free parameters of the standard model with the corresponding observed values. . . . .	13
2.1.	The scalar potential and physical parameters of the Higgs sector of the 2HDM. . . . .	16
2.2.	The different types of 2HDM. The quarks and leptons interact either with the first or the second Higgs doublet in different types. . . . .	17
2.3.	The coupling of quarks and leptons with Higgs $A$ for all types of 2HDM. . . . .	17
2.4.	Allowed range at 95% CL for various parameters from the search for charged Higgs involving $c\bar{s}$ channel at different collider experiments. . . . .	22
2.5.	Allowed range at 95% CL for various parameters from the search for charged Higgs involving other channels such as $\tau^+\nu$ , $c\bar{b}$ , and $t\bar{b}$ at different collider experiments. . . . .	23
3.1.	The energy of protons after passing through various accelerators. . . . .	34
4.1.	Various properties and parameters of the CMS muon subsystems [27]. . . . .	65
4.2.	Event rate after applying level-1 and high-level triggers. . . . .	68
5.1.	Muon and electron data samples recorded in 2016 with an integrated luminosity $35.9 \text{ fb}^{-1}$ . .	77
5.2.	Background and signal simulated samples with corresponding cross section and the number of events. . . . .	80
5.3.	The SM $t\bar{t}$ + jets samples used for systematics study. . . . .	81
5.4.	Selection and veto cuts applied on muon. . . . .	84
5.5.	Selection and veto cuts applied on electron found in barrel and endcap regions. . . . .	87
5.6.	Effective area for different $\eta_{\text{sc}}$ range. . . . .	89
5.7.	Selection criteria applied on jets. . . . .	90

5.8. Track selection criteria. PCA stands for the point of closest approach.	92
5.9. Secondary vertex selection criteria. $d_{\text{PV}, \text{SV}}^{\text{trans}}$ is the 2D flight distance or the distance between a primary and secondary vertex in the transverse plane.	92
5.10. The efficiency of loose (L), medium (M), and tight (T) b tag working points for different quark-flavor of jets [35]. These efficiencies are calculated from $t\bar{t}$ events with jet $p_T > 20$ GeV.	93
 6.1. Jet energy resolution scale factors for different $\eta$ range	97
 7.1. List of event filters applied to data and simulated samples. Most of the filters are available in the MINIAOD dataset through the trigger collection. The unavailable filters are applied on the fly using EDFilters.	100
7.2. Event yields after various selection cuts for the $\mu + \text{jets}$ channel. The $W + \text{jets}$ process is the dominant background upto 1-lepton selection. When the events are required to have $N_{\text{jets}} \geq 4$ , the $t\bar{t} + \text{jets}$ becomes the dominant background. The QCD multijet events are from simulated samples as listed in Table 5.2. The simulated MC signal process corresponds to $m_{H^+} = 120$ GeV.	103
7.3. Event yields after various selection cuts for the $e + \text{jets}$ channel. A similar trend as that of Table 7.2 is seen. The simulated MC signal process corresponds to $m_{H^+} = 120$ GeV.	103
7.4. Signal event yields for the different mass of charged Higgs after various selection cuts for the $\mu + \text{jets}$ channel. Event yields are almost same upto 1-lepton selection for all mass points. However, after the $N_{\text{jets}} \geq 4$ selection (when $t\bar{t} + \text{jets}$ becomes the dominant background as shown in Table 7.2) the event yield for higher masses of the charged Higgs are reduced more compared to lower masses. The event yield reduces because of the less phase space available between top-quark and charged Higgs for higher masses.	104
7.5. Signal event yields for different mass of charged Higgs after various selection cuts for the $e + \text{jets}$ channel. Similar trend as that of Table 7.4 is seen.	104
 8.1. The efficiency of loose (L), medium (M), and tight (T) c tag working points for different quark-flavor of jets [35]. These efficiencies are calculated from $t\bar{t}$ events with jet $p_T > 20$ GeV.	114
9.1. Event yields from different regions, for $\mu + \text{jets}$ channel after kinematic fit selection.	122
9.2. Event yields from different regions, for $e + \text{jets}$ channel after kinematic fit selection.	124

9.3. The QCD multijet transfer scale factors ( $SF_{\text{qcd}}$ ) for $\mu + \text{jets}$ and $e + \text{jets}$ channel after various selection steps. The categorization of events based on exclusive loose, medium, and tight c quark tagging is described in Section 10.3. . . . .	125
10.1. Event yield for inclusive category after kinematic fit selection. The statistical uncertainty in the total background corresponds to the quadratically added uncertainties from individual backgrounds. However, a systematic uncertainty correlated among each background is linearly added for the total background. And each uncorrelated systematic uncertainty for the total background is quadratically added. . . . .	131
10.2. Event yield for inclusive loose charm category after kinematic fit selection. The statistical uncertainty in the total background corresponds to the quadratically added uncertainties from individual backgrounds. However, a systematic uncertainty correlated among each background is linearly added for the total background. And each uncorrelated systematic uncertainty for the total background is quadratically added. . . . .	133
10.3. Event yield from exclusive charm categories after kinematic fit selection. The statistical uncertainty in the total background corresponds to the quadratically added uncertainties from individual backgrounds. However, a systematic uncertainty correlated among each background is linearly added for the total background. And each uncorrelated systematic uncertainty for the total background is quadratically added. . . . .	138
11.1. Inclusive loose c tag types. The acronym yLyMnT refers to the case where a jet satisfies loose and medium, but not tight working point. . . . .	141
11.2. Inclusive medium c tag types. . . . .	141
11.3. Systematic and statistical uncertainties in %, from inclusive category after kinematic fit selection for $\mu + \text{jets}$ ( $e + \text{jets}$ ) channel. The “—” indicates that the corresponding uncertainties are not considered for the given process. . . . .	144
11.4. Systematic and statistical uncertainties in %, from inclusive loose charm tagging category after kinematic fit selection for $\mu + \text{jets}$ ( $e + \text{jets}$ ) channel. The “—” indicates that the corresponding uncertainties are not considered for the given process. . . . .	144
11.5. Systematic and statistical uncertainties in %, from exclusive charm tagging categories after kinematic fit selection for $\mu + \text{jets}$ ( $e + \text{jets}$ ) channel. The “—” indicates that the corresponding uncertainties are not considered for the given process. . . . .	145

12.1. Description of nuisance parameters corresponding to the systematic and statistical uncertainties. . . . .	155
12.2. 95% CL exclusion limit in % for $\mu + \text{jets}$ channel from different event categories. . . . .	160
12.3. 95% CL exclusion limit in % for $e + \text{jets}$ channel from different event categories. . . . .	163
12.4. 95% CL exclusion limit in % for $\ell + \text{jets}$ channel from different event categories. . . . .	163
12.5. Acronym used in the impact plots for naming statistical nuisance parameters. . . . .	164
12.6. Goodness of fit for $\mu + \text{jets}$ channel, from different event categories. . . . .	171
12.7. Goodness of fit for $e + \text{jets}$ channel, from different event categories. . . . .	171
12.8. Goodness of fit for $\ell + \text{jets}$ channel, from different event categories. . . . .	171
 13.1. 95% CL exclusion limits in % for the $\mu + \text{jets}$ from combined c tagging categories. . . . .	175
13.2. 95% CL exclusion limits in % for the $e + \text{jets}$ from combined c tagging categories. . . . .	175
13.3. 95% CL exclusion limits in % for the $\ell + \text{jets}$ from combined c tagging categories. . . . .	175

## **Part I.**

# **The Standard Model and Beyond**



# 1 | The Standard Model

## 1.1. Introduction

Since ancient times, the human mind has been curious to know more and more about nature. There have been fundamental questions about the origin of the universe and the smallest entity the universe is made of. There have been many theoretical and experimental attempts to answer these questions in the past centuries. With the advancement in science and technology, we have theoretically explained and experimentally verified some of the fundamental building blocks of the universe. However, there are still many more questions, for example about the origin of dark matter [36, 37], which we are still investigating. None of the experiments have directly observed it, although there are many theoretical postulates about it. There are many pieces of evidence which suggest that the known part of the universe is only about 4%, rest 96% are unknown to us. The unknown part of the universe is supposed to consist of dark matter and dark energy.

The standard model of particle physics explains the smallest constituents of the known part of the matter. As shown in Figure 1.1, a house is made of sand stones, a stone is made of crystals, a crystal is made of atoms, an atom is made of nucleons (proton, neutron) and electrons, a proton (or neutron) is made of quarks. The quark and electron are the fundamental particles, that is, they are not made of any other sub-particles. Therefore, the whole known universe is made of these fundamental particles by combining them in appropriate proportions. There are other fundamental particles which mediate the force which binds the quarks to form a proton or neutron.

The quarks are of different types such as *up* quark (u), *down* quark (d), *charm* quark (c), *strange* quark (s), *top* quark (t), and *bottom* quark (b). The mass of each quark is different. The electric charge is 2/3 for u, c, t quarks and -1/3 for s, b, and d quarks. The electron has different partners such as *muon* ( $\mu$ ) and *tau* ( $\tau$ ). They differ in mass only, the electric charge is the same. The electron, muon, and tau are collectively called leptons. There are also electric-neutral partners of leptons called neutrinos such as *electron-neutrino* ( $\nu_e$ ), *muon-neutrino* ( $\nu_\mu$ ), and *tau-neutrino* ( $\nu_\tau$ ). The quarks are bound by a strong force mediated by a *gluon*.

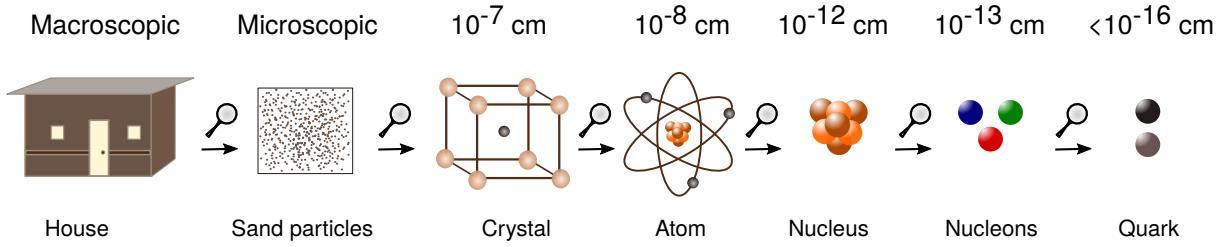


Figure 1.1.: A schematic picture showing the hierarchy of building blocks of the matter when one looks deeper and deeper, from right to left. This figure is adopted from [2].

The leptons interact themselves by an electromagnetic force mediated by a *photon*. The lepton and neutrino interact by electro-weak force mediated by the W and Z boson. The quarks, leptons, photon, gluon, W/Z boson are the fundamental particles. The mass to these particles is given by another fundamental particle called the *Higgs* boson. The mass, electric charge, and spin of all the fundamental particles are shown in Figure 1.2. All the composite particles are made of fundamental particles. All the physics process in nature involves interactions between the fundamental particles. The standard model of particle physics describe all the interactions of the fundamental particles, as shown in Figure 1.4. Here the photon can interact with a lepton, quarks, and W boson. The gluon can interact with only quarks and itself. The Higgs boson interacts with itself, electron, muon, tau, W, Z, and all the quarks.

The discovery of all the fundamental particles took centuries. Some of them were theoretically predicted many years before being experimentally observed. The timeline of the fundamental particles are shown in Figure 1.3. The electron was the first elementary particle to be theorized and discovered in the late 19<sup>th</sup> century. The photon, electron-neutrino, and muon were discovered in the first 50 yrs of the 20<sup>th</sup> century. All the quarks, except the t quark, were discovered in the 60's and 70's of the 20<sup>th</sup> century. The Higgs boson is the only particle discovered in the 21<sup>st</sup> century which took 48 yrs, it was theoretically predicted in 1964. A brief history of elementary particles can be found in [38].

## 1.2. Group and symmetry

In mathematics, a set  $(G, \times)$  is called a group with respect to the multiplicative ( $\times$ ) operator if the following conditions are satisfied:

1. there exist an identity ( $I$ ) element in  $G$  so that  $I \times g = g \times I = g$  where  $g$  is an element of  $G$ ,

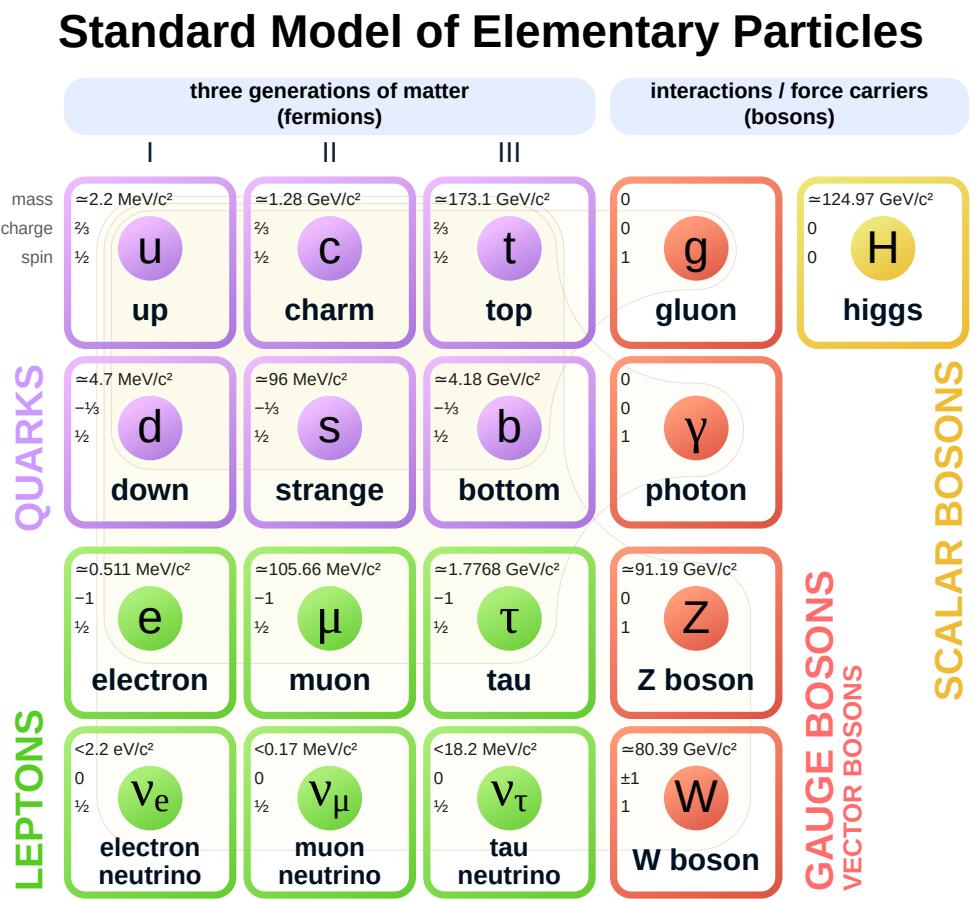


Figure 1.2.: Physical properties of the fundamental particles [3]. There are three generations of quarks and leptons grouped in the increasing order of mass. The gluon, photon, Z/W, and Higgs bosons are force carriers.

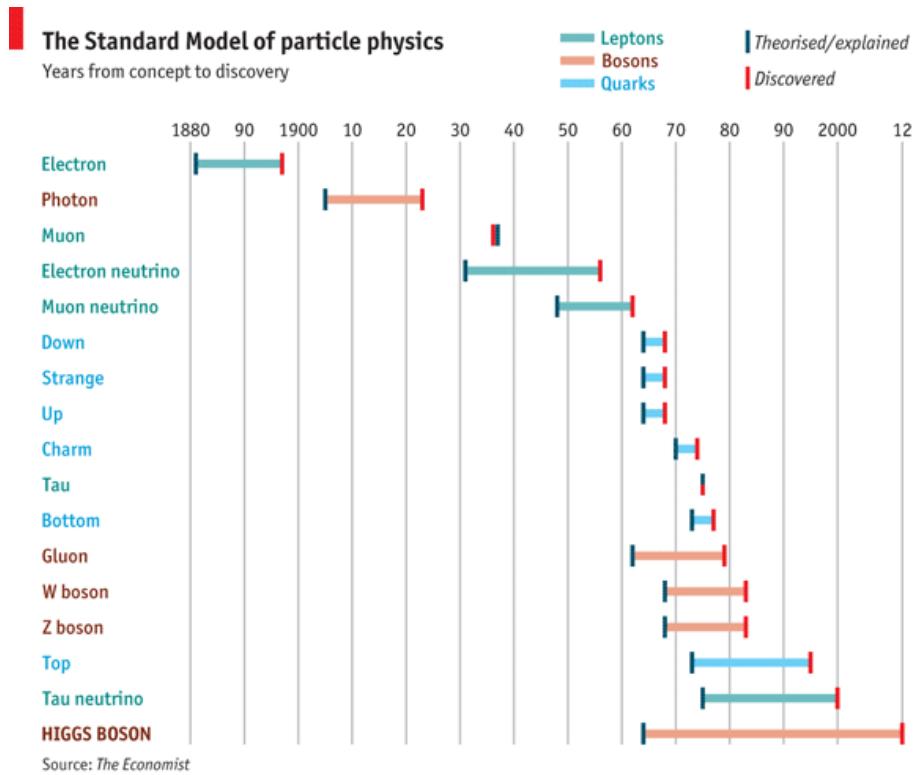


Figure 1.3.: A timeline of the fundamental particles from their theoretical prediction to the experimental observation.

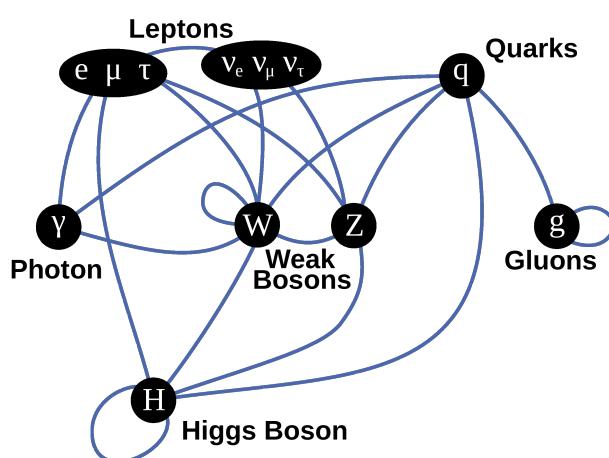


Figure 1.4.: A schematic diagram showing the interactions of fundamental particles among themselves [4].

2. the elements of the group follow associative law that is  $(g_1 \times g_2) \times g_3 = g_1 \times (g_2 \times g_3)$ , and
3. there exist an inverse of every element of  $G$  so that  $g \times g^{-1} = g^{-1} \times g = I$ .

The group  $(G, \times)$  is called an abelian group if the elements also follow commutative law that is  $g_1 \times g_2 = g_2 \times g_1$ . An example of a group in physics is the continuous rotation group. In general, all the elements of a group can be generated by the parameters and generators of the group. For Lie groups, an element is given by  $e^{i\alpha_a \tau_a}$  where  $\alpha_a$  ( $a = 1, 2, \dots, n$ ) are the parameters and  $\tau_a$  are the generators. For the theory of the standard model, we need the following three groups.

- **U(1):** The unitary  $1 \times 1$  matrices form the U(1) group. An element of the U(1) is given by  $e^{i\alpha}$  where  $\alpha$  is the parameter. The identity element is the generator of U(1).
- **SU(2):** The elements of special unitary (SU) group are  $2 \times 2$  unitary matrices. The “special” implies that the determinant of each matrix is 1. There are in total 3 generators (the three Pauli matrices). All the elements of the SU(2) can be constructed using these generators ( $\tau_a$ ) and a set of parameters ( $\alpha_a$ ). An element of the SU(2) is given by  $e^{i\alpha_a \tau_a}$ , where  $a = 1, 2, 3$ . Out of three generators, one is diagonal. In general, the elements of the SU(2) group do not commute. Therefore, SU(2) is a non-Abelian group.
- **SU(3):** A set of  $3 \times 3$  unitary matrices form the SU(3) group. There are 8 generators (Gell-Mann matrices) of the SU(3), two out of eight are diagonal. An element of the SU(3) is given by  $e^{i\alpha_a \lambda_a}$  where  $a = 1, 2, \dots, 8$ , and  $\lambda_a$  are the Gell-Mann matrices. In general, for a SU(N), there are  $N^2 - 1$  generators, out of which  $N - 1$  are diagonal. The SU(3) group is also non-Abelian.

A physical system is said to be invariant if it does not change under the transformation of a group. For example, a function  $f(\psi)$  is invariant under SU(2) if  $f(\psi') = f(\psi)$  where  $\psi' = e^{i\alpha_a \tau_a} \psi$ . There could be a function of more than one variable such as  $f(\psi, \phi)$  where one of the variables (say  $\psi$ ) might transform under SU(2) however, the other variable may not transform at all under this group. In this case also, the function  $f(\psi, \phi)$  will be invariant under SU(2). Therefore it is important to know about the transformation law of every variable of a function. One can also construct a function where one variable transforms under one group (say SU(2)) and the other variable transforms under a different group (say U(1)). In this case, the function is said to be invariant under the combined group  $SU(2) \times U(1)$ . However, the generators of the individual group must commute for the invariance under combined group.

Noether theorem [39] demands that there is a conserved charge with every global (the parameters of the group are space-time independent) symmetry. For example, invariance under rotations leads to conservation of angular momentum. The quantum numbers for U(1) and SU(2) are denoted by  $Q$  (called charge) and

Table 1.1.: The third component of the weak-isospin, hypercharge, and charge quantum numbers for leptons, quarks, and bosons.

Fermion type	Generation			$T_3$	Y	Q
	1st	2nd	3rd			
Leptons	$\begin{pmatrix} v_e \\ e \end{pmatrix}_L$	$\begin{pmatrix} v_\mu \\ \mu \end{pmatrix}_L$	$\begin{pmatrix} v_\tau \\ \tau \end{pmatrix}_L$	$\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$	-1	$\begin{pmatrix} 0 \\ -1 \end{pmatrix}$
	$e_R$	$\mu_R$	$\tau_R$	0	-2	-1
Quarks	$\begin{pmatrix} u \\ d \end{pmatrix}_L$	$\begin{pmatrix} c \\ s \end{pmatrix}_L$	$\begin{pmatrix} t \\ b \end{pmatrix}_L$	$\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$	$\frac{1}{3}$	$\begin{pmatrix} \frac{2}{3} \\ -\frac{1}{3} \end{pmatrix}$
	$u_R$	$c_R$	$t_R$	0	$\frac{4}{3}$	$\frac{2}{3}$
	$d_R$	$s_R$	$b_R$	0	$-\frac{2}{3}$	$-\frac{1}{3}$
	$W^+$			+ 1	0	+ 1
	$W^-$			- 1	0	- 1
	Z			0	0	0
	$\gamma$			0	0	0
	H			$-\frac{1}{2}$	1	0

$T$  (called the weak-isospin). The third component of weak-isospin is denoted by  $T_3$ . In the context of the standard model, where electro-weak lagrangian (refer Section 1.3) is invariant under the  $SU(2)_L \times U(1)$  group, an other quantum number called the hypercharge ( $Y = 2(Q - T_3)$ ) is assigned to every particle. The  $L$  in the  $SU(2)_L$  implies that only left-handed fermions transform under the  $SU(2)$  group. The right-handed fermions are singlets, that is, they don't transform under this group. The various quantum numbers associated with all the fundamental particles are shown in Table 1.1. From this table, it can be seen that vector bosons ( $W^+$ ,  $W^-$ , and Z) form weak-isospin triplet, left-handed fermions form weak-isospin doublet, the Higgs boson is also part of the weak-isospin doublet, and the rest of the other particles (right-handed quarks and photon) are weak-isospin singlet.

### 1.3. The Lagrangian density

The physics of all fundamental particles is described by the Standard Model. The starting point of all physical theories is the construction of a Lagrangian density. The equations of motion for different fields,

such as Klein-Gordon equation for scalar field, Dirac equation for spinor field, Maxwell's equations of electromagnetism, etc, can be derived from a Lagrangian density using the Euler-Lagrange equation. For the sake of convenience, we will call "Lagrangian density" as "Lagrangian" only. The Lagrangian has to satisfy some basic principles. First, it has to follow some symmetry principle, that is, it should be invariant under the transformation of the prescribed symmetry group. Second, the Lagrangian has to be "renormalizable". One of the criteria of renormalization demands that each interaction term in the Lagrangian should have a coupling with mass dimension greater than or equals to zero (the mass dimension of fermion field, boson field, space-time derivative, and mass is 3/2, 1, 1, 1, respectively). The advantage of Lagrangian formulation is that the Feynman rules can be obtained just by looking at the mathematical structure of individual terms, for example, the propagator is determined by the quadratic terms, and the vertex factor is determined by the interaction terms. The standard model (SM) is a unified theory incorporating electromagnetic, weak, and strong interactions. The SM Lagrangian is invariant under the  $SU(3)_C \times SU(2)_L \times U(1)_Y$  group and written as

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{EW}} + \mathcal{L}_{\text{QCD}} + \mathcal{L}_{\text{H}} + \mathcal{L}_{\text{Yukawa}} \quad (1.1)$$

where individual terms are from electroweak, quantum chromodynamics, Higgs, and Yukawa sectors. A detailed description of each individual term is given below.

- $\mathcal{L}_{\text{EW}}$ : The electro-weak (EW) Lagrangian follows  $SU(2)_L \times U(1)_Y$  symmetry and incorporates the interaction between leptons (and quarks) and electro-weak gauge bosons along with the respective kinetic energy terms. It is written as

$$\mathcal{L}_{\text{EW}} = \bar{L} \gamma^\mu (i\partial_\mu - g' \frac{1}{2} Y B_\mu - g \frac{1}{2} \tau_a W_\mu^a) L + \bar{R} \gamma^\mu (i\partial_\mu - g' \frac{1}{2} Y B_\mu) R - \frac{1}{4} W_a^{\mu\nu} W_{\mu\nu}^a - \frac{1}{4} B^{\mu\nu} B_{\mu\nu}, \quad (1.2)$$

where  $L$  ( $\bar{L} = L^\dagger \gamma^0$ ) is the doublet of left-handed leptons ( $e_L, \nu_{e_L}; \mu_L, \nu_{\mu_L}; \tau_L, \nu_{\tau_L}$ ) as shown in the first row of Table 1.1,  $R$  is the singlet of right-handed lepton (only  $e_R, \mu_R, \tau_R$ ) as shown in the second row of Table 1.1, the  $\gamma^\mu$  are the Dirac matrices,  $\tau_a$  are Pauli matrices,  $g'$  and  $g$  are coupling strength of the  $U(1)$  and  $SU(2)$  group, respectively. In Equation (1.2), we have not explicitly shown the terms corresponding to quarks. The  $B_\mu$  is the  $U(1)$  gauge field,  $W^a$  are the  $SU(2)$  gauge fields,  $B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$  and  $W_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a + g \epsilon^{abc} W_\mu^b W_\nu^c$  are the corresponding field strength tensors. Here  $\epsilon^{abc}$  is totally antisymmetric Levi-Civita symbol. The  $Y$  in Equation (1.2) is the hypercharge of the  $U(1)$  group. The summation in Equation (1.2) runs over three spaces; three generation of lepton, three generators of the  $SU(2)$  ( $a = 1, 2, 3$ ), and the 4 Lorentz indices ( $\mu = 0, 1, 2, 3$ ). In Equation (1.2), there is no coupling of  $W^a$  with the right-handed lepton, as they interact only with the left-handed lepton. Under the  $SU(2)_L \times U(1)_Y$  group, the field  $L$ ,  $R$ ,  $B_\mu$ , and  $W_\mu^a$  transform as  $L' = e^{i\alpha_a(x)T^a + i\beta(x)Y} L$ ,  $R' = e^{i\beta(x)Y} R$ ,  $B'_\mu = B_\mu +$

$\frac{1}{e}\partial_\mu\alpha(x)$ ,  $W_\mu'^a = W_\mu^a - \frac{1}{g}\partial_\mu\alpha_a(x) - f_{abc}\alpha_b(x)W_\mu^c$ , where  $\alpha(x)$  and  $\beta(x)$  are the local gauge parameters of the SU(2) and U(1) group, respectively. The Lagrangian  $\mathcal{L}_{EW}$  is invariant under these transformations.

- $\mathcal{L}_{QCD}$ : The Lagrangian for the quantum chromodynamics involves the interactions of quarks and gluons and respects SU(3) symmetry. Unlike the lepton, the quarks have an additional color quantum number called “color”. The Lagrangian is given as

$$\mathcal{L}_{QCD} = \bar{Q}_i (i\gamma^\mu (\partial_\mu \delta_{ij} - ig_s G_\mu^a T_{ij}^a)) Q_j - \frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu}, \quad (1.3)$$

where the  $i$  in  $Q_i$  is the color index for  $R$ ,  $G$ , and  $B$ . The  $Q$  is the Dirac spinor for each quark,  $g_s$  is the strength of strong coupling,  $T^a$  are 8 ( $a = 1, 2, \dots, 8$ ) generators of the SU(3) also called Gell-Mann matrices,  $G^a$  are the gauge fields of the strong interaction,  $G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a + g f^{abc} G_\mu^b G_\nu^c$  are the gluon field strength tensors, and  $f^{abc}$  are the structure constants of SU(3). There are five summations involved in Equation (1.3): quark family ( $Q = u, d, c, s, t, b$ ), chirality of quarks, color ( $i = 1, 2, 3$ ),  $a = 1, 2, \dots, 8$ , and Lorentz ( $\mu = 0, 1, 2, 3, 4$ ). Under the SU(3), the two fields transform as  $Q' = e^{i\alpha_a(x)T^a} Q$ ,  $G'_\mu = G_\mu^a - \frac{1}{g}\partial_\mu\alpha_a(x) - f_{abc}\alpha_b(x)G_\mu^c$ . The Lagrangian  $\mathcal{L}_{QCD}$  is invariant under these transformations.

- $\mathcal{L}_H$ : The Higgs sector of the SM incorporates interaction of Higgs fields with gauge bosons of SU(2) and U(1) groups as well as a potential energy term for Higgs. The Higgs field is part of a complex scalar doublet of the SU(2) group. The Lagrangian density is given as

$$\mathcal{L}_H = \left| \left( \partial_\mu + \frac{i}{2} (g' Y B_\mu + g \tau_a W_\mu^a) \right) \varphi \right|^2 - V(\varphi), \quad (1.4)$$

where the Higgs doublet is defined as

$$\varphi = \frac{1}{\sqrt{2}} \begin{pmatrix} \varphi^+ \\ \varphi^0 \end{pmatrix}, \quad (1.5)$$

with  $\varphi^+$  and  $\varphi^0$  having electrical charge +1 and 0, respectively. The hypercharge of both fields is 1. The potential term for the Higgs field  $V(\varphi)$  is given in Section 1.4. Under the SU(2) group, the Higgs doublet transforms as  $\varphi' = e^{i\alpha_a(x)\tau^a} \varphi$ . The transformation law for the other two fields of Equation (1.4) is described in  $\mathcal{L}_{EW}$ . The mass to the vector bosons is generated in  $\mathcal{L}_H$  as discussed in Section 1.4.

- $\mathcal{L}_{Yukawa}$ : The interactions of fermions (quarks and leptons) with Higgs boson are described by the Yukawa Lagrangian. It is written as

$$\mathcal{L}_{Yukawa} = G_l \bar{L} \varphi R + G_d \bar{q}_L \varphi d_R + G_u \bar{q}_L \varphi_c u_R + h.c., \quad (1.6)$$

where  $G_l$  is the coupling of Higgs to lepton,  $G_d$  ( $G_u$ ) is the coupling of Higgs to up (down) quark, and  $\varphi_c = -i\tau_2 \varphi^*$ .

## 1.4. Mass of particles

The Lagrangian, given in Equation (1.1), is complete in the sense that it is invariant under the  $SU(3)_C \times SU(2)_L \times U(1)_Y$  group and it is renormalizable. However, except for the Higgs boson, the mass terms for other particles, such as the fermion mass term  $m\bar{\psi}\psi$ , are not included in Equation (1.1). The mass term for the Higgs boson is included in the potential

$$V(\varphi) = (\partial_\mu \varphi)^\dagger (\partial^\mu \varphi) - \mu^2 \varphi^\dagger \varphi - \lambda (\varphi^\dagger \varphi)^2, \quad (1.7)$$

where,

$$\varphi = \frac{1}{\sqrt{2}} \begin{pmatrix} \varphi_1(x) + i\varphi_2(x) \\ \varphi_3(x) + i\varphi_4(x) \end{pmatrix}, \quad (1.8)$$

After substituting  $\varphi$  from Equation (1.8) to the potential of Equation (1.7), it can be easily seen that all the four Higgs fields have a mass term of type  $\mu^2 \varphi_i^2$  where  $i = 1, 2, 3, 4$ , where  $\mu^2 < 0$ . This lead to spontaneous symmetry breaking as discussed in the next section. Note that  $\mu$  is not the physical mass of the Higgs particle. The physical masses of all particles are generated by the spontaneously breaking of symmetry [40] and the Higgs mechanism [4, 41–46] as discussed below.

- **Spontaneous symmetry breaking (SSB) and Higgs mechanism:** The minima of the potential  $V(\varphi)$  correspond to  $\varphi_1^2 + \varphi_2^2 + \varphi_3^2 + \varphi_4^2 = v^2$ , where  $v = \sqrt{-\mu^2/\lambda}$  is the vacuum expectation value (vev),  $\mu^2 < 0$ , and  $\lambda > 0$ . Next we expand the potential about the minima. The minima can be chosen at  $\varphi_1 = \varphi_2 = \varphi_4 = 0, \varphi_3 = v$ , and the potential in the Lagrangian  $\mathcal{L}_H$  is expanded about this minimum. That is, the field given by Equation (1.8) is re-written as

$$\varphi = \frac{1}{\sqrt{2}} \begin{pmatrix} \varphi_1(x) + i\varphi_2(x) \\ v + h(x) + i\varphi_4(x) \end{pmatrix}, \quad (1.9)$$

It turns out that the fields  $\varphi_1, \varphi_2$  and  $\varphi_4$  are not physical and get eliminated from the physical particle spectrum. We see this by reparametrizing this equation as,

$$\varphi = \frac{e^{i\tau_a \theta^a(x)/v}}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \quad (1.10)$$

where  $\theta^a$  are three new fields. We can now eliminate these fields by a gauge transformation. The Higgs multiplet can now be expressed as

$$\varphi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \quad (1.11)$$

The resulting gauge choice is called the unitary gauge. In this gauge, only the physical Higgs field appears

in the Lagrangian and all the other scalar fields get eliminated. After substituting  $\varphi$  from Equation (1.11) in  $\mathcal{L}_H$ , mass terms appear for gauge bosons  $B_\mu$ ,  $W_1$ ,  $W_2$ , and  $W_3$ . Furthermore the Higgs field acquires physical mass. The other three scalar degrees of freedom now act as the longitudinal modes of the massive gauge bosons.

The  $\mathcal{L}_H$  now contains only gauge fields  $B_\mu, W_\mu^a$  and the  $h(x)$  with respective mass terms. We next identify the physical gauge boson particle spectrum. We note that the vacuum preserves the electromagnetic  $U(1)_{em}$  symmetry since  $\varphi_3$  has zero electric charge ( $T^3 = -1/2$ ,  $Y = 2$ , hence  $Q = 0$ ) hence does not transform under the  $U(1)_{em}$  group. The gauge bosons are expressed in a new basis as  $W^\pm = (W^1 \mp iW^2)/\sqrt{2}$ ,  $A_\mu = \cos\theta_W B_\mu + \sin\theta_W W_\mu^3$ ,  $Z_\mu = -\sin\theta_W B_\mu + \cos\theta_W W_\mu^3$ . After writing  $\mathcal{L}_H$  in terms of  $A_\mu$ ,  $Z_\mu$ ,  $W^\pm$ , one can see that the  $A_\mu$  becomes massless and  $W^\pm$ ,  $Z_\mu$  remain massive with masses  $\frac{1}{2}vg$ ,  $\frac{1}{2}v\sqrt{g^2 + g'}$ , respectively. The mass of leptons (electron, muon, and tau) and quarks are generated using Equation (1.10) in  $\mathcal{L}_{\text{Yukawa}}$ . The neutrino in the SM remains massless.

Notice that before the symmetry breaking, there were 4 massless (gauge bosons) and 4 scalar particles. The polarization degree of freedom ( $P_{\text{dof}}$ ) for each Higgs is 1 and for each massless gauge boson is 2. Therefore, total  $P_{\text{dof}}$  before the SSB is  $4 \times 1 + 4 \times 2 = 12$ . After the SSB, the massive gauge bosons get an additional  $P_{\text{dof}}$  from the three scalar boson. Therefore, the total number of  $P_{\text{dof}}$  after the SSB is also 12 ( $1 \times 1 + 3 \times 0 + 1 \times 2 + 3 \times 3$ ).

The masses of all the fundamental particles are shown in Table 1.2 in terms of free parameters. However, the value of these parameters is determined from experiment. Until now, the Lagrangian  $\mathcal{L}_{\text{SM}}$  contains 14 free parameters which are 3 couplings of gauge groups ( $g$ ,  $g'$ ,  $g_s$ ), 2 from Higgs field ( $v$ ,  $\lambda$ ), 3 couplings ( $G_l$ ) of Higgs-lepton interactions, and 6 couplings ( $G_q$ ) of Higgs-quark interactions. There are 5 additional free parameters in the SM which are 3 CKM mixing angles ( $\theta_{12}$ ,  $\theta_{23}$ ,  $\theta_{13}$ ) [47], 1 CKM CP violation phase [48], and 1 QCD vacuum angle ( $\theta_{\text{QCD}}$ ). Therefore, there are 19 free parameters in the SM.

## 1.5. Success and failure of the SM

The standard model is a mathematically complete and experimentally tested theory. However, it does not provide explanations to all of the observed phenomena in nature. A brief description of successes and failures of the SM is given below.

- **Success:** The SM describes the interaction among the fundamental particles with excellent accuracy. The electro-weak gauge bosons ( $W^\pm$ ,  $Z$ ), gluon, top, and charm quarks were predicted by the SM before

Table 1.2.: The mass of all fundamental particles in terms of the free parameters of the standard model with the corresponding observed values.

Particles	Theoretical mass	Experimental mass
H	$v\sqrt{2\lambda}$	125.1 GeV
$W^\pm$	$\frac{1}{2}vg$	80.4 GeV
Z	$\frac{1}{2}v\sqrt{g^2 + g'}$	91.2 GeV
photon	0	0
gluon	0	0
$e, \mu, \tau$	$\frac{G_F v}{\sqrt{2}}$	shown in Figure 1.2
neutrino	0	shown in Figure 1.2
Quarks	$\frac{G_q v}{\sqrt{2}}$	shown in Figure 1.2

being observed experimentally. The Higgs boson was predicted in 1964 and experimentally observed in 2012 [49]. Value to all the free parameters of the SM has been experimentally assigned. The predictions from the SM agree very well with the experimental observations.

- **Failure:** There are many shortcomings of the SM. A few of them are mentioned below:

- The  $\mathcal{L}_{SM}$  includes three generations of leptons. However, only the first generation ( $e, \nu_e$ ) is enough to form the nucleus and hence the visible content of the universe. The SM does not provide any explanation for the existence of the second and the third generations or why there should not be more than three?
- The neutrinos in the SM are massless. However, it is observed by the SNO experiment that neutrinos do have a small mass.
- Third, why the Lagrangian  $\mathcal{L}_{SM}$  does not include gravity?
- The SM does not explain the origin of dark matter [36, 37] and dark energy which account for about 96% of the universe.
- Why there are so many (19) free parameters in the SM?
- Why there is more matter than antimatter, etc?

To explain the shortcomings of the SM, many beyond the standard model (BSM) theories have been proposed such as grand unified theories, supersymmetry [50], etc. One of the extensions of the SM that predicts the existence of charged Higgs boson is described in the next chapter.



# 2 | The Two Higgs Doublet Model

## 2.1. Introduction

As described in Section 1.5, although the predictions of the SM match well with the experimental observation, the SM has no explanation to many other observed phenomena. To be able to explain such phenomena, many extensions of the SM have been proposed. One of the simplest extensions is the two Higgs doublet model (2HDM) where an extra Higgs doublet is added to Equation (1.1). In Lagrangian of the SM, there is only one Higgs SU(2) doublet comprising four Higgs fields. After the spontaneous symmetry breaking, only one Higgs ( $h$ ) remains in the model as the other three Higgses are eaten away by the three gauge bosons ( $W^\pm$ ,  $Z$ ). In the 2HDM, there are two SU(2) Higgs doublets comprising eight Higgs fields. Three of them are again eaten away by the gauge bosons and the five Higgses ( $h, H, A, H^\pm$ ) remains in the model. Further, the 2HDM is divided into four types (I, II, X, Y) based on the coupling of quarks and leptons with the two Higgs doublets. The type II of the 2HDM corresponds to the MSSM (minimal supersymmetric extension of the standard model).

## 2.2. The Higgs sector

The Higgs potential term for the 2HDM is constructed in such a way so that it is invariant under  $Z_2$  symmetry, where the Higgs doublets transform as  $\Phi_1 \rightarrow \Phi_1$ ,  $\Phi_2 \rightarrow -\Phi_2$ . The  $Z_2$  symmetry is imposed to avoid the flavor changing neutral currents (FCNC). The CP-conserving potential is given by [5]

$$V^{2\text{HDM}} = m_1^2 \Phi_1^\dagger \Phi_1 + m_2^2 \Phi_2^\dagger \Phi_2 - m_3^2 (\Phi_1^\dagger \Phi_2 + \Phi_2^\dagger \Phi_1) + \frac{\lambda_1}{2} (\Phi_1^\dagger \Phi_1)^2 + \frac{\lambda_2}{2} (\Phi_2^\dagger \Phi_2)^2 + \lambda_3 (\Phi_1^\dagger \Phi_1)(\Phi_2^\dagger \Phi_2) + \lambda_4 (\Phi_1^\dagger \Phi_2)(\Phi_2^\dagger \Phi_1) + \frac{\lambda_5}{2} \left[ (\Phi_1^\dagger \Phi_2)^2 + (\Phi_2^\dagger \Phi_1)^2 \right], \quad (2.1)$$

where  $m_i$  and  $\lambda_i$  are scalar potential parameters taken to be real. The parameters are chosen such that the potential admits spontaneous symmetry breaking. Let  $v_i$  be the vacuum expectation value (vev) of the Higgs field  $\Phi_i$ . We expand the Higgs field around its vev [5], such that,

Table 2.1.: The scalar potential and physical parameters of the Higgs sector of the 2HDM.

Scalar potential parameters	$m_1, m_2, m_3, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5$
Physical parameters	$m_h^2, m_H^2, m_{H^\pm}^2, m_A^2, m_3^2, v, \tan\beta, \cos(\beta - \alpha)$

$$\Phi_i = \begin{pmatrix} \omega_i^+ \\ \frac{1}{\sqrt{2}}(v_i + h_i - iz_i) \end{pmatrix}. \quad (2.2)$$

After substituting new fields from Equation (2.2) into (2.1), one can see that there are off-diagonal terms for  $\omega_i^+$ ,  $h_i$ ,  $z_i$  fields. To remove these off-diagonal terms from the potential, these fields are reparametrized in terms of new fields  $\omega^\pm$ ,  $H^+$ ,  $z$ ,  $A$ ,  $h$ , and  $H$  by using the following transformations [5]

$$\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = R(\alpha) \begin{pmatrix} H \\ h \end{pmatrix}, \quad \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = R(\beta) \begin{pmatrix} z \\ A \end{pmatrix}, \quad \begin{pmatrix} \omega_1^+ \\ \omega_2^+ \end{pmatrix} = R(\beta) \begin{pmatrix} \omega^+ \\ H^+ \end{pmatrix}, \quad (2.3)$$

where  $R(\alpha)$ ,  $R(\beta)$  are the rotation matrices given by [5]

$$R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \quad (2.4)$$

After the Higgs mechanism, the three Higgses are eliminated from the potential terms. They are eaten away by the three gauge bosons ( $W^\pm$ ,  $Z$ ). The free parameters in the potential are masses of the remaining five Higgses,  $m_3^2$ ,  $v$ ,  $\tan\beta$ , and  $\cos(\beta - \alpha)$ , where  $v = \sqrt{v_1^2 + v_2^2}$  and  $\tan\beta = v_1/v_2$ . The scalar parameters in the potential and the physical parameters are listed in Table 2.1.

## 2.3. The Yukawa sector

The Yukawa sector of the 2HDM is considerably more complicated in comparison to that of a single Higgs doublet. The general Lagrangian free from FCNC for the Yukawa sector is given as

$$\mathcal{L}_{\text{Yukawa}}^{\text{2HDM}} = G^l \bar{L} \varphi_i R + G^d \bar{q}_L \varphi_i d_R + G^u \bar{q}_L \varphi_i^c u_R + h.c., \quad (2.5)$$

where  $i$  can be 1 or 2. The interactions of the two Higgs doublets are different for different flavors of the fermions. Based on the different types of interaction, the 2HDM is divided into four types as shown in Table 2.2. For example, in Type II, the up-type quarks interact with the second Higgs doublet ( $\varphi_2$ ), down type-quarks and leptons interact with the first Higgs doublet ( $\varphi_1$ ).

The general form of the Lagrangian, as given by Equation (2.5), for all types of the 2HDM can be written

Table 2.2.: The different types of 2HDM. The quarks and leptons interact either with the first or the second Higgs doublet in different types.

Coupling of	Type-I	Type-II	Type-X	Type-Y
up-type quarks	$\Phi_2$	$\Phi_2$	$\Phi_2$	$\Phi_1$
down-type quarks	$\Phi_2$	$\Phi_1$	$\Phi_2$	$\Phi_2$
leptons	$\Phi_2$	$\Phi_1$	$\Phi_1$	$\Phi_2$

Table 2.3.: The coupling of quarks and leptons with Higgs  $A$  for all types of 2HDM.

Coupling	Type-I	Type-II	Type-X	Type-Y
$G_A^u$	$\cot\beta$	$\cot\beta$	$\cot\beta$	$\cot\beta$
$G_A^d$	$-\cot\beta$	$\tan\beta$	$-\cot\beta$	$\tan\beta$
$G_A^l$	$-\cot\beta$	$\tan\beta$	$\tan\beta$	$-\cot\beta$

as [5]

$$\mathcal{L}_{\text{yukawa}}^{2\text{HDM}} = - \sum_{f=u,d,\ell} \left( \frac{m_f}{v} G_h^f \bar{f} f h + \frac{m_f}{v} G_H^f \bar{f} f H - i \frac{m_f}{v} G_A^f \bar{f} \gamma_5 f A \right) - \left\{ \frac{\sqrt{2} V_{ud}}{v} \bar{u} \left( m_u G_A^u P_L + m_d G_A^d P_R \right) d H^+ + \frac{\sqrt{2} m_\ell G_A^\ell}{v} \bar{v}_L \ell_R H^+ + \text{H.c.} \right\}, \quad (2.6)$$

where  $f$  is the fermion,  $G_h^f$  is the coupling of fermion  $f$  with the Higgs boson  $h$ , and  $P_{L/R}$  are projection operators. The value of  $G^f$  is different for different types of the 2HDM as shown in Table 2.3. Only couplings with the Higgs  $A$  is shown in this table as only that is needed for the study of branching ratios of the charged Higgs boson. Couplings of fermions with all the Higgses are given in [5]. The branching ratio of t quark decays and the charged Higgs decays have been given in the next sections.

## 2.4. Branching ratio of top quark decay

Being a very short-lived particle with a lifetime of  $5 \times 10^{-25}$  s, the t quark decays immediately. In the SM, it decays only to W boson and b quark as its decay involving other quarks (s, d) is suppressed by the corresponding element of the CKM matrix. The ratio  $R = \mathcal{B}(t \rightarrow W^+ b)/\mathcal{B}(t \rightarrow W^+ q)$  where  $q = b, s$ , and  $d$  is  $0.90 \pm 0.04$  as was earlier determined from D0 experiment using  $5.4 \text{ fb}^{-1}$  data [51]. The CDF experiment using  $162 \text{ pb}^{-1}$  luminosity has put a lower bound at 95% CL on it,  $R > 0.61$  [52]. The latest lower bound,  $R > 0.95$ , is obtained from the CMS experiment at 8 TeV using  $19.7 \text{ fb}^{-1}$  luminosity [53] which suggests that the t quark can have other decay channels beyond the SM. In the 2HDM, it is allowed to decay into  $H^+ q$

channel. The branching ratio of this decay process is given by

$$BR(t \rightarrow H^+ q) = \frac{\Gamma(t \rightarrow H^+ q)}{\Gamma(t \rightarrow W^+ q) + \Gamma(t \rightarrow H^+ q)} \quad (2.7)$$

where  $\Gamma(t \rightarrow H^+ q)$  and  $\Gamma(t \rightarrow W^+ q)$  are the partial decay widths. For the b quark, these decay widths are given as [5]

$$\begin{aligned} \Gamma(t \rightarrow H^+ b) &= \frac{G_F |V_{tb}|^2}{8\sqrt{2}\pi m_t} \lambda \left( \frac{m_b^2}{m_t^2}, \frac{m_{H^\pm}^2}{m_t^2} \right)^{1/2} \\ &\times \left\{ m_t^2 \left[ m_t^2 G_A^u {}^2 \left( 1 + \frac{m_b^2}{m_t^2} - \frac{m_{H^\pm}^2}{m_t^2} \right) + m_b^2 G_A^d {}^2 \right] + 4m_t^2 m_b^2 G_A^u G_A^d \right\}, \end{aligned} \quad (2.8)$$

and

$$\Gamma(t \rightarrow W^+ b) = \frac{G_F |V_{tb}|^2}{8\sqrt{2}\pi} m_t^3 \quad (2.9)$$

where

$$\lambda(x, y) = 1 + x^2 + y^2 - 2x - 2y - 2xy, \quad (2.10)$$

and  $m_t = 166$  GeV (running mass of the t quark),  $m_b$  (running mass of b quark) = 3.0 GeV,  $G_F = 1.1663 \times 10^{-5}$ , and  $V_{tb} = 0.99$ . The variation of  $\mathcal{B}(t \rightarrow H^+ b)$  as a function of  $\tan\beta$  and the mass of the charged Higgs is shown in Figure 2.1. From the plots shown on the left side, it can be seen that for low  $\tan\beta < 1$ , the  $\mathcal{B}(t \rightarrow H^+ b)$  is the same for all the types of 2HDM. For higher  $\tan\beta$ , the  $\mathcal{B}(t \rightarrow H^+ b)$  increases in the Type II model as compared to Type X. The similar  $\tan\beta$  dependence can be explicitly seen from the right hand plots of this figure. Note that  $\mathcal{B}(t \rightarrow H^+ b)$  is the same for Type I and X, and for Type II and Y as the coupling of quarks with charged Higgs is same as shown in Table 2.3. We also see that  $\mathcal{B}(t \rightarrow H^+ b)$  does not change much with respect to the mass of charged Higgs for  $m_{H^\pm} < 150$  GeV.

## 2.5. Branching ratio of charged Higgs decay

For the search of charged Higgs boson, it is very important to study all of its possible decay channels to have an estimate of the dominance of respective channels for a given value of  $\tan\beta$ ,  $m_{H^\pm}$ , etc. In the 2HDM, the charged Higgs is allowed to decay into fermions and gauge bosons, for example,  $H^+ \rightarrow c\bar{s}$ ,  $H^+ \rightarrow c\bar{b}$ ,  $H^+ \rightarrow t\bar{b}$ ,  $H^+ \rightarrow \tau^+ \nu_\tau$ ,  $H^+ \rightarrow W^+ \gamma$ ,  $H^+ \rightarrow W^+ Z$ ,  $H^+ \rightarrow hW^+$ ,  $H^+ \rightarrow HW^+$ , and  $H^+ \rightarrow AW^+$ . The partial decay width of charged Higgs decaying to quarks and leptons is given by [5]

$$\Gamma(H^+ \rightarrow ud) = N_C \frac{G_F m_{H^\pm} |V_{ud}|^2}{4\sqrt{2}\pi} \beta_{ud} \left\{ \left( m_u^2 G_A^u {}^2 + m_d^2 G_A^d {}^2 \right) \left( 1 - \frac{m_u^2 + m_d^2}{m_{H^\pm}^2} \right) - \frac{4m_u^2 m_d^2 G_A^u G_A^d}{m_{H^\pm}^2} \right\}, \quad (2.11)$$

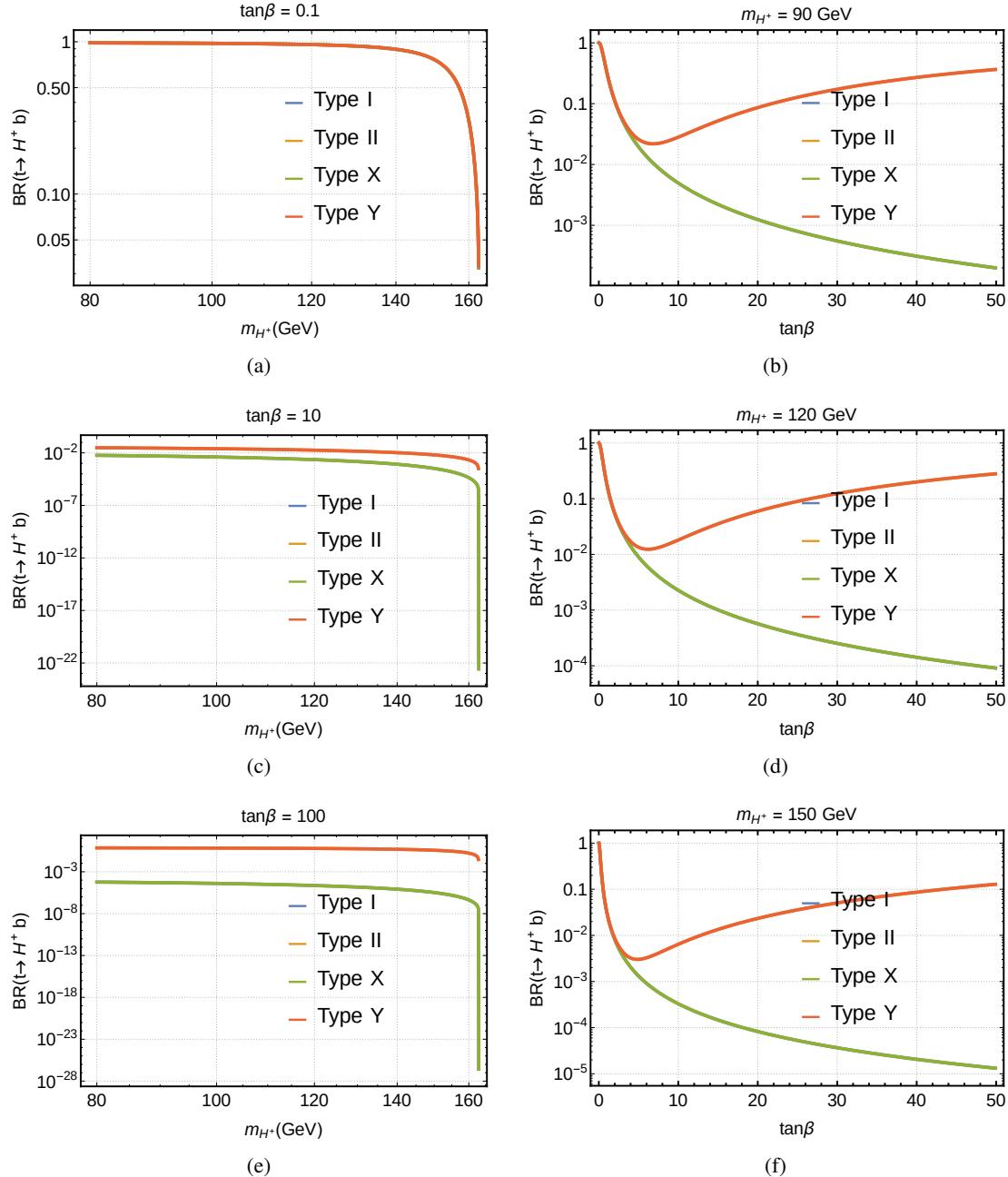


Figure 2.1.: Variation of  $\mathcal{B}(t \rightarrow H^+ b)$  as a function of the mass of the charged Higgs (left side plots) and  $\tan\beta$  (right side plots).

and

$$\Gamma(H^+ \rightarrow \ell^+ v) = \frac{G_F m_{H^\pm} m_\ell^2}{4\sqrt{2}\pi} G_A^{\ell^2} \left(1 - \frac{m_\ell^2}{m_{H^\pm}^2}\right)^2 \quad (2.12)$$

where

$$\beta_{XY} = \lambda^{1/2} \left( \frac{m_X^2}{m_\phi^2}, \frac{m_Y^2}{m_\phi^2} \right), \quad (2.13)$$

and  $N_C = 3$ ,  $V_{cs} = 0.97$ ,  $m_c = 0.81$  GeV,  $m_s = 0.046$  GeV,  $V_{cb} = 0.0412$ ,  $m_\tau = 1.77$  GeV,  $m_\mu = 0.105$  GeV.

The total decay width of charged Higgs decaying into fermions is given by

$$\Gamma(H^+ \rightarrow \text{All}) = \Gamma(H^+ \rightarrow \tau^+ v) + \Gamma(H^+ \rightarrow c\bar{s}) + \Gamma(H^+ \rightarrow c\bar{b}) + \Gamma(H^+ \rightarrow \mu^+ v) \quad (2.14)$$

The branching ratio of charged Higgs decay into different channels is shown in Figure 2.2. As shown in this figure, the  $\tau^+ v$  channel is dominant for all  $\tan\beta$  in the Type I model. For low  $\tan\beta$ , the  $c\bar{s}$  channel is dominant in Type II and Type X.

## 2.6. Current status of searches for charged Higgs

There have been several searches for the charged Higgs boson in last two decades, particularly in the context of Type II of the 2HDM. However, no signature of the existence of the charged Higgs has been found as of now. In the absence of any signal for the charged Higgs, various constraints have been imposed from a wide range of theoretical studies as well as low and high energy collider experiments. The theoretical constraints come from the vacuum stability, perturbative unitarity, and electro-weak precision observables [54]. The low energy constraints are from the indirect searches [55] where a charged Higgs appears in the loop such as the decay of B-mesons ( $B \rightarrow X_s \gamma$ ). The high energy constraints are from the direct production of charged Higgs at the collider experiments [56–63]. The constraints from direct searches are more reliable as compared to those from indirect searches or theoretical predictions.

The direct searches of the charged Higgs have been performed at the different experiments such as ALEPH, L3, DELPHI, and OPAL at the large electron-positron collider (LEP), D0 and CDF at the Tevatron, and finally CMS and ATLAS at the large hadron collider (LHC) in different decay channels of the charged Higgs. These searches have been performed at different center of mass energies, different luminosities, and for a wide range of the mass of the charged Higgs boson. Constraints have been imposed on various parameters such as the mass of charged Higgs,  $\tan\beta$ , branching ratio of t quark decay into the charged Higgs and b quark, etc. A summary of all constraints from the collider experiment is shown in Tables 2.4 and 2.5. The current best constraints are:  $m_{H^+} > 79.3$  GeV from the ALEPH at 0.2 TeV ( $0.63 \text{ fb}^{-1}$ ),  $\mathcal{B}(t \rightarrow H^+ b) < 0.8-$

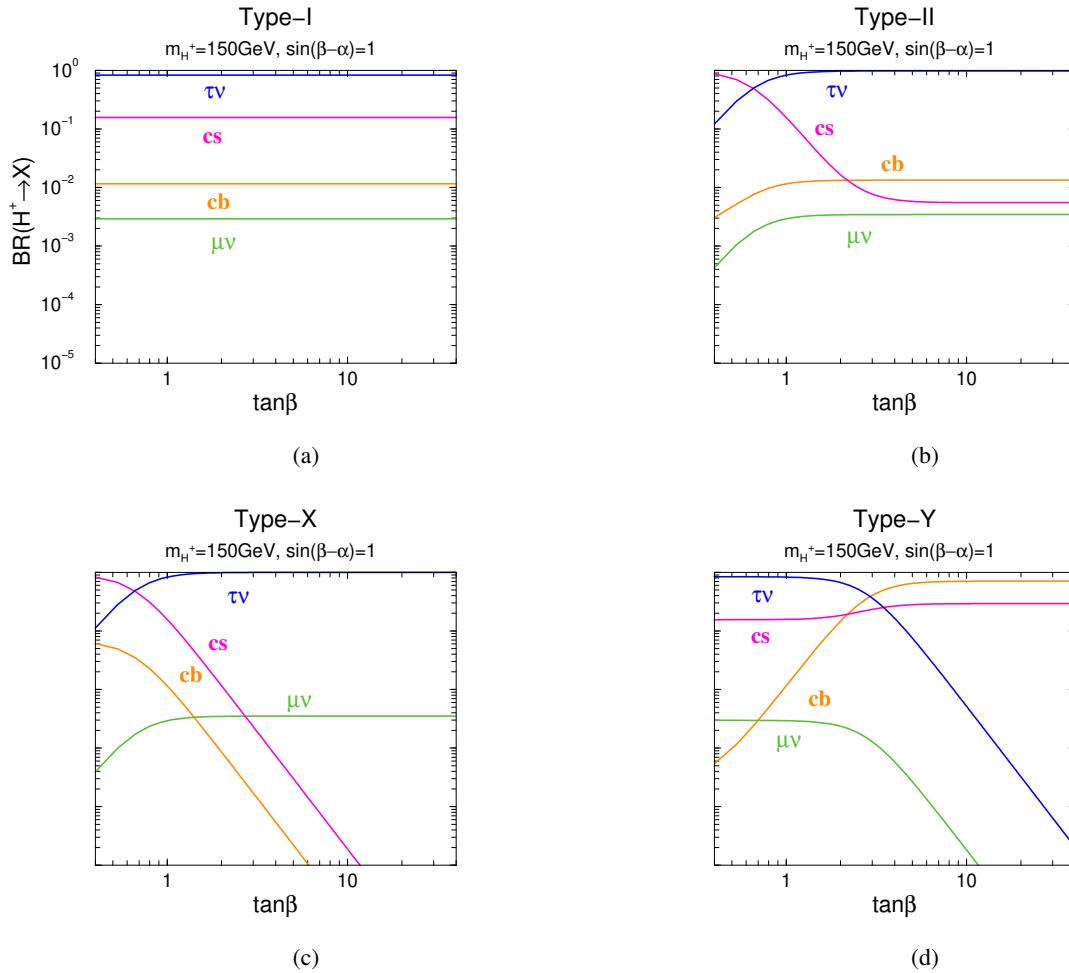


Figure 2.2.: The branching ratio of charged Higgs decay in different channels as a function of  $\tan\beta$  for  $m_{H^+} = 150 \text{ GeV}$  in different types of the 2HDM. These plots are taken from [5].

Table 2.4.: Allowed range at 95% CL for various parameters from the search for charged Higgs involving  $c\bar{s}$  channel at different collider experiments.

Experiment	Year	$\sqrt{s} (\mathcal{L}_{int})$ TeV ( $\text{fb}^{-1}$ )	Allowed range at 95% CL	Assumption
ALEPH	2002 [57]	0.189-0.209 (0.629)	$m_{H^+} > 79.3 \text{ GeV}$	$\mathcal{B}(H^+ \rightarrow \tau^+ \nu_\tau) + \mathcal{B}(H^+ \rightarrow c\bar{s}) = 1$
L3	2003 [56]	0.189-0.209 (0.629)	$m_{H^+} > 76.5 \text{ GeV}$	$\mathcal{B}(H^+ \rightarrow \tau^+ \nu_\tau) + \mathcal{B}(H^+ \rightarrow c\bar{s}) = 1$
DELPHI	2004 [58]	0.189-0.209 (0.220)	$m_{H^+} > 74.5 \text{ GeV}$	Pair production of $H^+ H^-$ in channels $\tau^+ \nu_\tau \tau^- \bar{\nu}_\tau, c\bar{s}c\bar{s}, c\bar{s}\tau^- \bar{\nu}_\tau$
OPAL	2012 [59]	0.189-0.209 (0.6)	$m_{H^+} > 76.3 \text{ GeV}$	$\mathcal{B}(H^+ \rightarrow \tau^+ \nu_\tau) + \mathcal{B}(H^+ \rightarrow q\bar{q}') = 1$
CDF	2009 [60]	1.96 (2.2)	$\mathcal{B}(t \rightarrow H^+ b) < 10\text{-}30\%, \text{ for } m_{H^+} \text{ in range } 60 \text{ to } 150 \text{ GeV}$	$\mathcal{B}(H^+ \rightarrow c\bar{s}) = 1$
D0	2009 [61]	1.96 (1.0)	$\mathcal{B}(t \rightarrow H^+ b) < 22\%, \text{ for } m_{H^+} \text{ in range } 80 \text{ to } 155 \text{ GeV}$	$\mathcal{B}(H^+ \rightarrow c\bar{s}) = 1$
ATLAS	2013 [62]	7 (4.7)	$\mathcal{B}(t \rightarrow H^+ b) < 5.1\%, \text{ for } m_{H^+} \text{ in range } 90 \text{ to } 150 \text{ GeV}$	$\mathcal{B}(H^+ \rightarrow c\bar{s}) = 1$
CMS	2016 [63]	8 (19.7)	$\mathcal{B}(t \rightarrow H^+ b) < 6.5\text{-}1.2\%, \text{ for } m_{H^+} \text{ in range } 90 \text{ to } 150 \text{ GeV}$	$\mathcal{B}(H^+ \rightarrow c\bar{s}) = 1$
CMS	2019	13 (35.9)	Presented in this thesis	$\mathcal{B}(H^+ \rightarrow c\bar{s}) = 1$

0.5% for  $m_{H^+}$  in range 90 to 150 GeV at 13 TeV ( $35.9 \text{ fb}^{-1}$ ) from the CMS experiment,  $\sigma(pp \rightarrow tH^+ + X) \times \mathcal{B}(H^+ \rightarrow \tau^+ \nu_\tau) < 4.2\text{-}0.0025 \text{ pb}$  for  $m_{H^+}$  in range 90 to 2000 GeV at 13 TeV ( $36.1 \text{ fb}^{-1}$ ) from the ATLAS,  $\sigma(pp \rightarrow tH^+ + X) \times \mathcal{B}(H^+ \rightarrow t\bar{b}) < 9.6\text{-}0.01 \text{ pb}$  for  $m_{H^+}$  in range 200 to 3000 GeV at 13 TeV ( $35.9 \text{ fb}^{-1}$ ) from the CMS, and  $\tan\beta < 1$  values are excluded for all  $m_{H^+}$  up to 160 GeV from both CMS as well as ATLAS. A few of the current best constraints are shown in Figure 2.3.

## 2.7. Search strategy for the charged Higgs at 13 TeV

The search for the charged Higgs boson in the  $c\bar{s}$  channel at 13 TeV in the CMS experiment adopted a similar strategy as that of the previous analysis at 8 TeV [63]. An additional charm quark tagging have been further exploited to improve sensitivity. The invariant mass of the jets originating from charm and strange

Table 2.5.: Allowed range at 95% CL for various parameters from the search for charged Higgs involving other channels such as  $\tau^+ \nu$ ,  $c\bar{b}$ , and  $t\bar{b}$  at different collider experiments.

Experiment	Year	$\sqrt{s} (\mathcal{L}_{int})$ TeV ( $\text{fb}^{-1}$ )	Allowed range at 95% CL	Assumption
D0	2009 [61]	1.96 (1.0)	$\mathcal{B}(t \rightarrow H^+ b) < 15\text{-}19\%$ , for $m_{H^+}$ in range 80 to 155 GeV	$\mathcal{B}(H^+ \rightarrow \tau^+ \nu_\tau) = 1$
ATLAS	2013 [64]	7 (4.6)	$\mathcal{B}(t \rightarrow H^+ b) < 0.8\text{-}3.4\%$ , for $m_{H^+}$ in range 90 to 160 GeV	$\mathcal{B}(H^+ \rightarrow \tau^+ \nu_\tau) = 1$
ATLAS	2015 [65]	8 (19.5)	$\sigma(pp \rightarrow tH^+ + X) \times \mathcal{B}(H^+ \rightarrow \tau^+ \nu_\tau) < 4.5\text{-}0.76 \text{ pb}$ for $m_{H^+}$ in range 180 to 1000 GeV, $\mathcal{B}(t \rightarrow H^+ b) \times \mathcal{B}(H^+ \rightarrow \tau^+ \nu_\tau) < 0.23\text{-}1.3\%$ for $m_{H^+}$ in range 80 to 160 GeV, and $\tan\beta < 1$ for $m_{H^+}$ in range 80 to 160 GeV	$\mathcal{B}(H^+ \rightarrow \tau^+ \nu_\tau) = 1$
ATLAS	2016 [66]	13 (3.2)	$\sigma(pp \rightarrow tH^+ + X) \times \mathcal{B}(H^+ \rightarrow \tau^+ \nu_\tau) < 1.9\text{-}0.015 \text{ pb}$ for $m_{H^+}$ in range 200 to 2000 GeV	$\mathcal{B}(H^+ \rightarrow \tau^+ \nu_\tau) = 1$
ATLAS	2016 [67]	13 (14.2)	$\sigma(pp \rightarrow tH^+ + X) \times \mathcal{B}(H^+ \rightarrow \tau^+ \nu_\tau) < 2.0\text{-}0.008 \text{ pb}$ for $m_{H^+}$ in range 200 to 2000 GeV	$\mathcal{B}(H^+ \rightarrow \tau^+ \nu_\tau) = 1$
ATLAS	2018 [68]	13 (36.1)	$\sigma(pp \rightarrow tH^+ + X) \times \mathcal{B}(H^+ \rightarrow \tau^+ \nu_\tau) < 4.2\text{-}0.0025 \text{ pb}$ for $m_{H^+}$ in range 90 to 2000 GeV, $\mathcal{B}(t \rightarrow H^+ b) \times \mathcal{B}(H^+ \rightarrow \tau^+ \nu_\tau) < 0.25\text{-}0.031\%$ for $m_{H^+}$ in range 90 to 160 GeV, and $\tan\beta < 1$ is excluded for all $m_{H^+}$ up to 160 GeV	$\mathcal{B}(H^+ \rightarrow \tau^+ \nu_\tau) = 1$
CMS	2018 [69]	8 (19.7)	$\mathcal{B}(t \rightarrow H^+ b) < 0.8\text{-}0.5\%$ , for $m_{H^+}$ in range 90 to 150 GeV	$\mathcal{B}(H^+ \rightarrow c\bar{b}) = 1$
CMS	2019 [70]	13 (35.9)	$\sigma(pp \rightarrow tH^+ + X) \times \mathcal{B}(H^+ \rightarrow \tau^+ \nu_\tau) < 6.0\text{-}0.005 \text{ pb}$ for $m_{H^+}$ in range 80 to 3000 GeV, and $\tan\beta < 1$ is excluded for all $m_{H^+}$ up to 160 GeV	$\mathcal{B}(H^+ \rightarrow \tau^+ \nu_\tau) = 1$
CMS	2019 [71]	13 (35.9)	$\sigma(pp \rightarrow tH^+ + X) \times \mathcal{B}(H^+ \rightarrow t\bar{b}) < 9.6\text{-}0.01 \text{ pb}$ , for $m_{H^+}$ in range 200 to 3000 GeV	$\mathcal{B}(H^+ \rightarrow t\bar{b}) = 1$

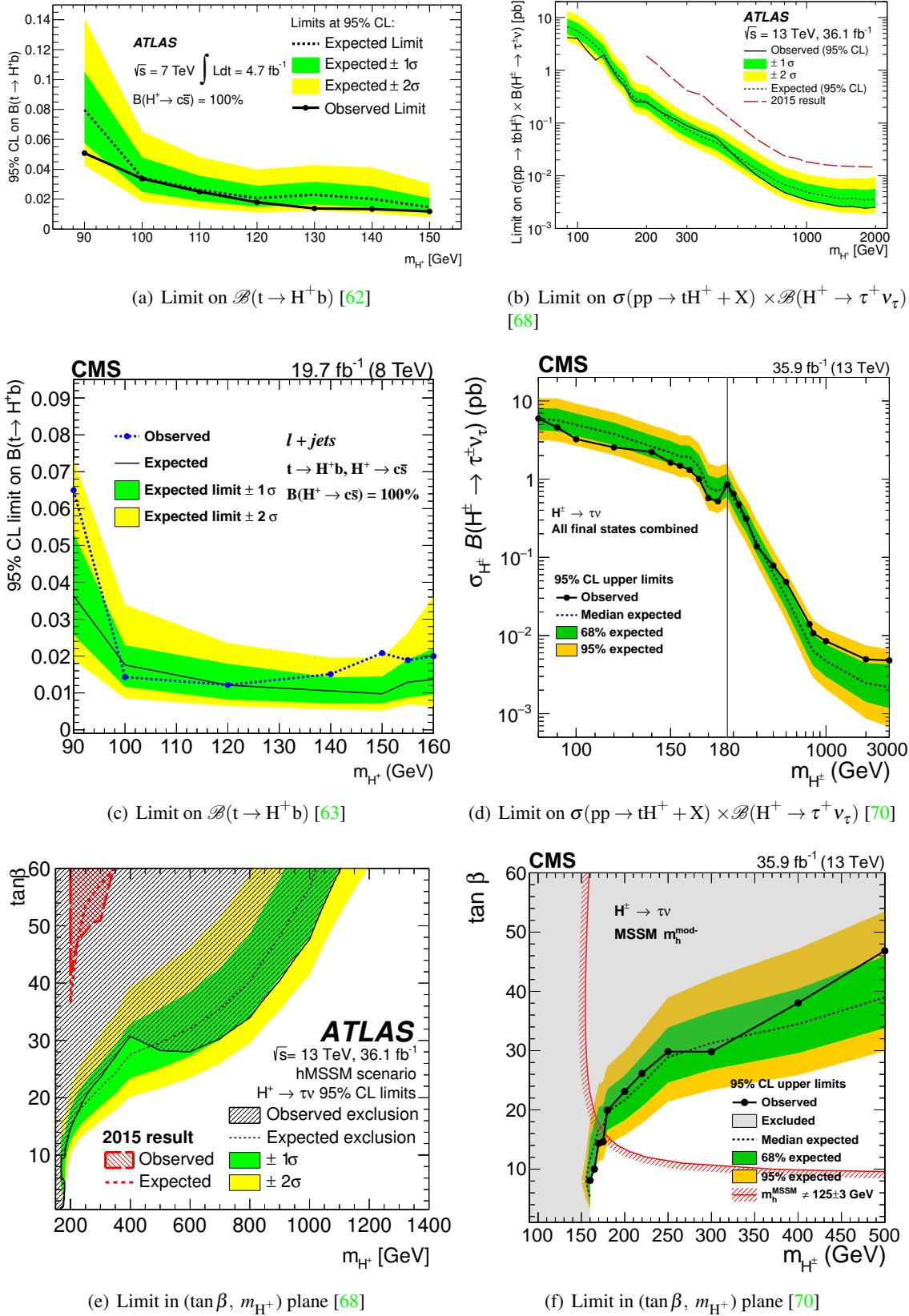


Figure 2.3.: Current allowed range at 95% CL from ATLAS and CMS.

antiquark is taken as the observable for the search of charged Higgs, in the low mass region from 80 to 160 GeV. In the absence of an excess in the observed data, a 95% CL limit is put on the  $\mathcal{B}(t \rightarrow H^+ b)$ . The charm tagging is extensively used to improve this limit. As shown on the right side of the Figure 2.4, for the signal process, one t quark decays to  $H^+ b$  and the other one to  $W^- \bar{b}$ . The  $W^+/H^+$  decays hadronically, whereas the  $W^-$  decays leptonically. As a result, in the final states, there are four jets (2 b jets, 1 c jet, 1 s jet), one lepton (electron or muon,  $\tau$  is not considered) and missing transverse energy attributed to neutrino. In this thesis, we assume that the  $\mathcal{B}(H^+ \rightarrow c\bar{s}) = 100\%$ .

The standard Model processes that give same final states (4 jets + 1 lepton + missing energy) are considered as backgrounds for this analysis. The standard model  $t\bar{t}$  production is the most dominant, irreducible background process. As shown in the left side of the Figure 2.4, for SM  $t\bar{t}$  process, one t quark decays to the  $W^+$  and b quark ( $t \rightarrow W^+ b$ ) and the other decays to  $W^- \bar{b}$  ( $\bar{t} \rightarrow W^- \bar{b}$ ). The SM  $t\bar{t}$  contributes around 94% of the total backgrounds. Other sub-dominant backgrounds that give rise to similar final states are single t quark production, QCD multijet,  $W + \text{jets}$ ,  $Z/\gamma + \text{jets}$ , and vector boson pair production processes. The following background processes are considered for the search for charged Higgs. They are ordered in their significance of contribution.

1. **SM  $t\bar{t} + \text{jets}$ :** Feynman diagrams for  $t\bar{t} + \text{jets}$  production are shown on the left hand side of Figure 2.4. This is the most dominant background channel in the signal search region (SR).
2. **Single t:** The single t quark production process can also mimic the signal topology. Three different ways, as shown in Figure 2.5, of production of single top quark considered in this analysis. It is produced through s-channel, t-channel, and tW-channel. In the s-channel and t-channel the initial quarks can be u, d, c and s (4-flavour scheme). However, in the tW-channel, the initial quark is only b quark (5-flavor scheme).
3. **QCD multijet:** The QCD multijet events contain only jets at parton level. However, after event reconstruction, they can still have leptons from misidentifications, and  $E_T^{\text{miss}}$  due to poor measurement of energy in the detector. Thus these events also mimic the signal topology.
4.  **$W + \text{jets}$ :** In this process, a W boson is produced in the proton-proton collisions which subsequently decays leptonically. Following  $W + \text{jets}$  background process are considered in this analysis:
  - 4.1  $W^\pm \rightarrow l^\pm \nu(l^- \bar{\nu})$
  - 4.2  $W + 1 \text{ jet } (W^\pm \rightarrow l^\pm \nu(l^- \bar{\nu}))$
  - 4.3  $W + 2 \text{ jets } (W^\pm \rightarrow l^\pm \nu(l^- \bar{\nu}))$

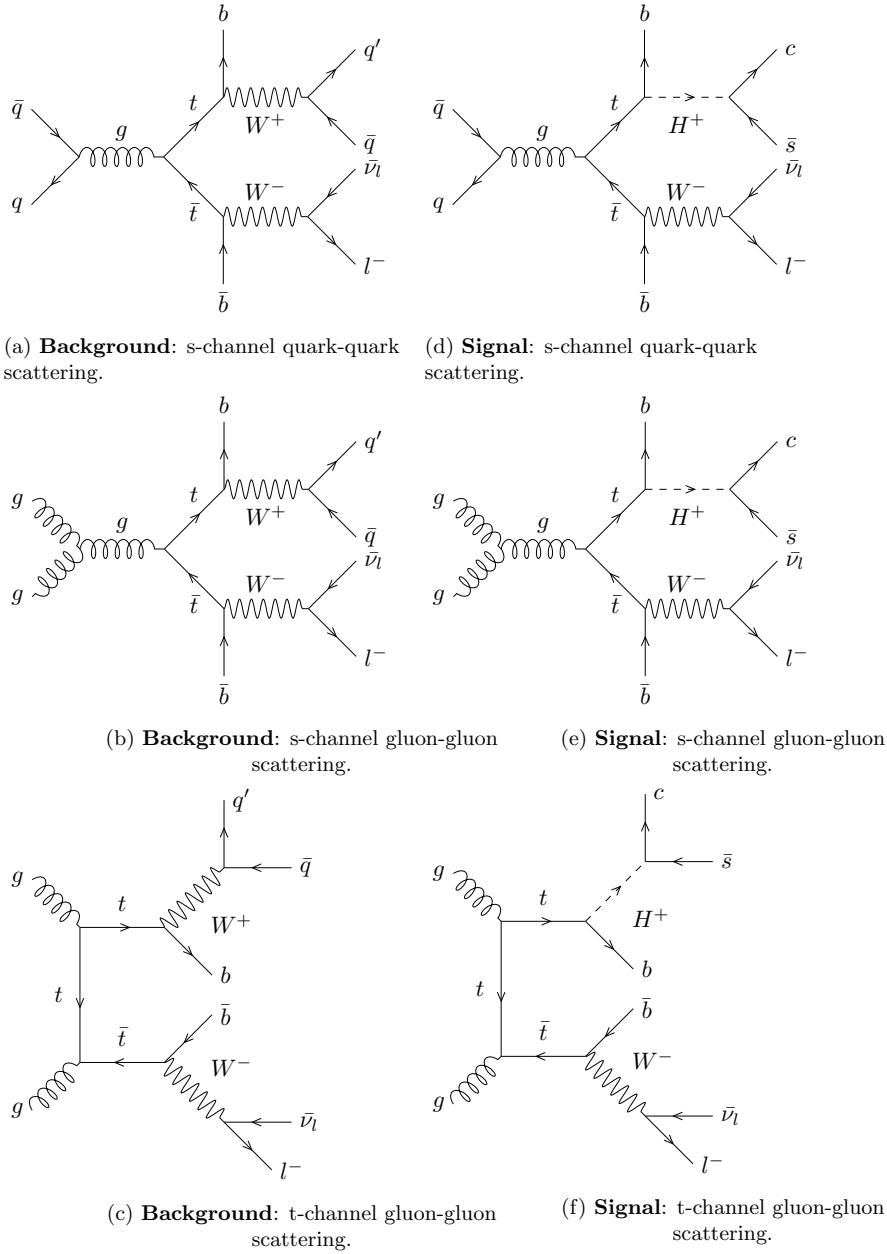


Figure 2.4.: Production of  $t\bar{t}$  from gluon-gluon and quark-quark scattering. The quark-scattering production process has a dominant contribution at Tevatron energies whereas gluon-gluon scattering diagrams are dominant at LHC energies [6, 7]. The SM production of  $t\bar{t}$  is shown in (a), (b) and (c). The charged Higgs boson production and its decay are shown in (d), (e), and (f).

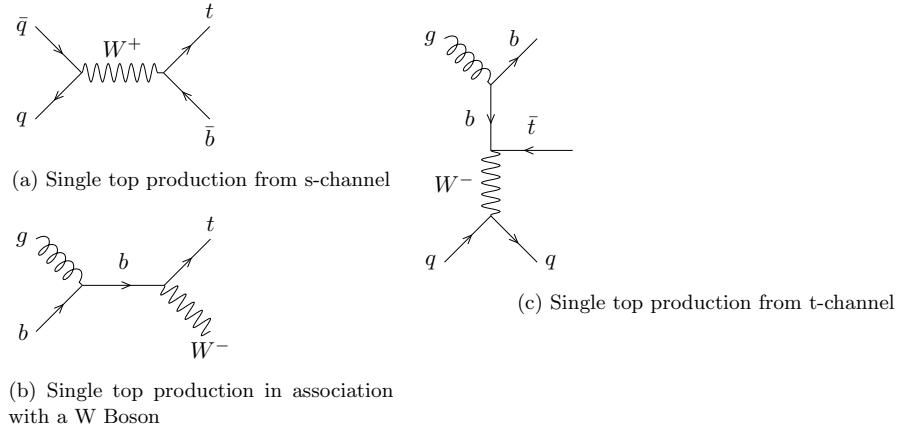


Figure 2.5.: Representative Feynman diagrams for single t quark production processes.

4.4  $W + 3 \text{ jets}$  ( $W^\pm \rightarrow l^+ \nu(l^- \bar{\nu})$ )4.5  $W + 4 \text{ jets}$  ( $W^\pm \rightarrow l^+ \nu(l^- \bar{\nu})$ )

The Feynman diagram for these processes are shown in Figure 2.6.

5.  **$Z/\gamma + \text{jets}$ :** The Drell-Yan processes in which  $Z/\gamma$  are produced along with jets, have lepton and jets at parton level as shown in Figure 2.7. However, after the reconstruction, the  $E_T^{\text{miss}}$  is also found in the events due to the poor measurement of energy in the detector.

5.1  $Z/\gamma + \text{jets}$  ( $Z/\gamma \rightarrow l^+ l^-$ )5.2  $Z/\gamma + 1 \text{ jet}$  ( $Z/\gamma \rightarrow l^+ l^-$ )5.3  $Z/\gamma + 2 \text{ jets}$  ( $Z/\gamma \rightarrow l^+ l^-$ )5.4  $Z/\gamma + 3 \text{ jets}$  ( $Z/\gamma \rightarrow l^+ l^-$ )5.5  $Z/\gamma + 4 \text{ jets}$  ( $Z/\gamma \rightarrow l^+ l^-$ )

6. **VV:** Vector boson fusion processes are the smallest background in the signal search region. The fusion happens via tri-linear coupling between the  $W^\pm$  and Z. The Z boson further decays to  $l^+ l^-$ . The vector boson pair production process (VV) has three sub-categories: WW, WZ, and ZZ.

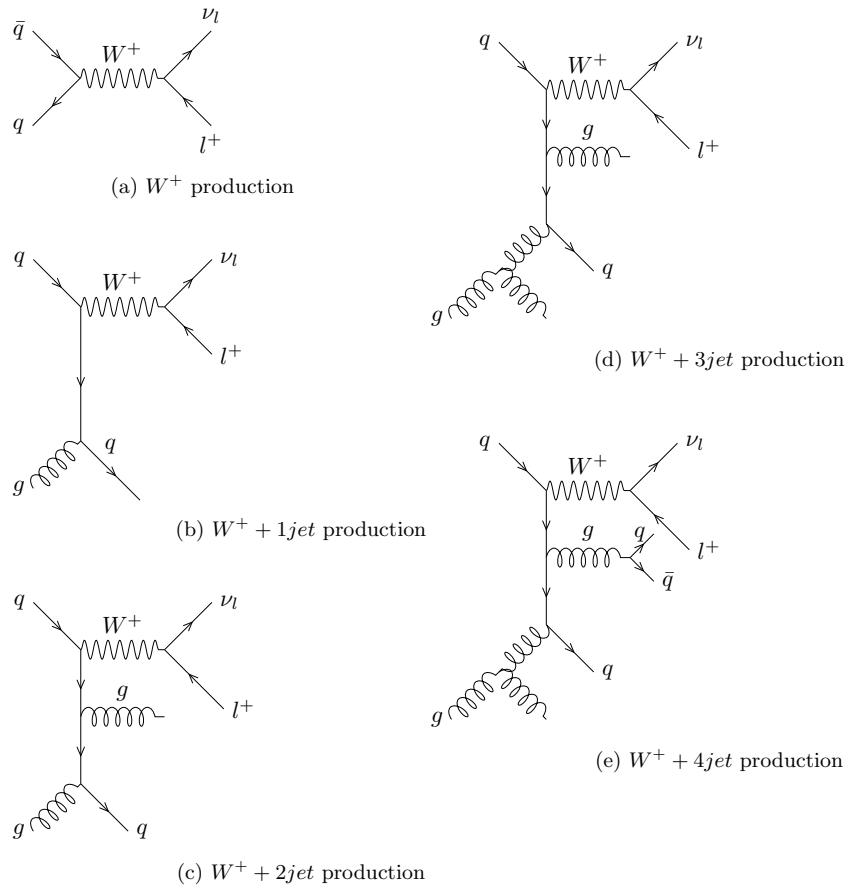


Figure 2.6.: Representative Feynman diagrams for  $W + n$  jets channel. The  $W$  boson is produced by quark-quark and quark-gluon scattering along with  $n$  jets ( $n = 0, 1, 2, 3, 4$ ).

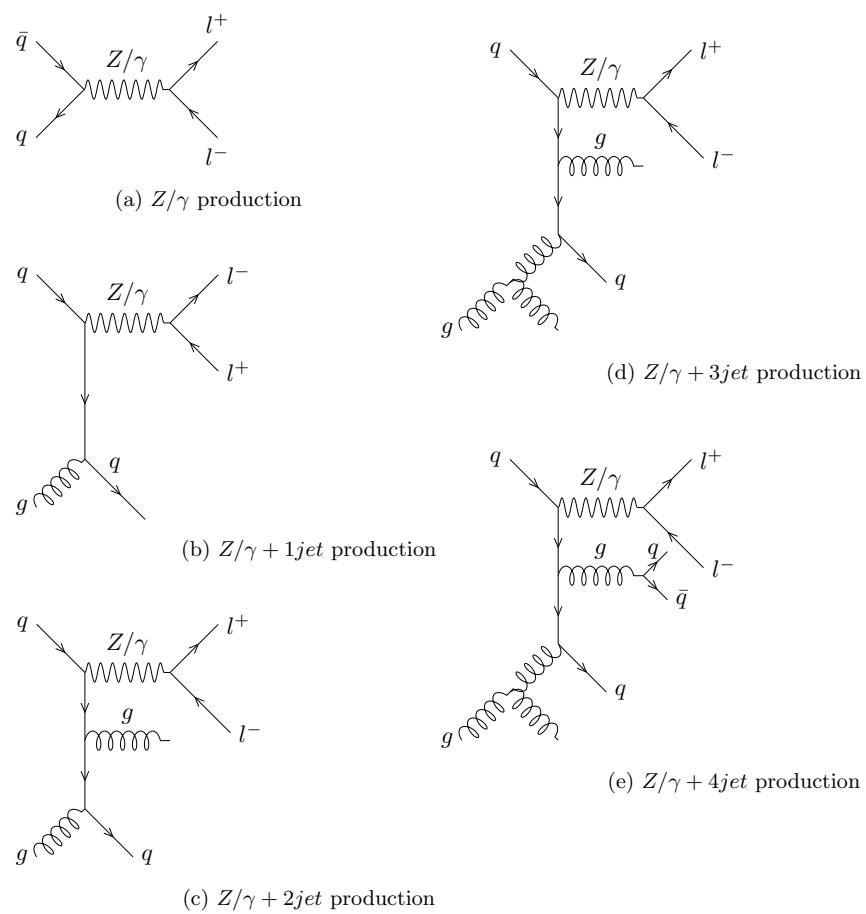


Figure 2.7.: Representative Feynman diagrams for  $Z/\gamma + n$  jets channel. The  $Z/\gamma$  is produced by quark-quark and quark-gluon scattering along with  $n$  jets ( $n = 0, 1, 2, 3, 4$ ).



## **Part II.**

# **The CMS Experiment at The LHC, CERN**



# 3 | The LHC at CERN

## 3.1. Introduction

The European Organisation for Nuclear Research (CERN- *Conseil Europén pour la Recherche Nucléaire*), located at Geneva, Switzerland, is one of the leading laboratories in the world in the field of experimental high energy physics. The experimental collaborations at CERN involve more than 1700 people from 41 countries. The research activity at CERN ranges from the understanding of fundamental constituents of matter such as the Higgs Boson to the hunt of dark matter. Particles such as protons (p), lead nuclei (Pb) are accelerated to a very high speed (nearly the speed of light) and collided head-on. The acceleration of these particles proceeds in several stages, such as shown in Figure 3.1.

Bunches of protons ( $H^+$ , after the electron is stripped-off from a hydrogen atom) are first passed through the Linear Accelerator (LINAC2), which accelerates the protons to an energy of 50 MeV. Subsequently, the proton bunches are passed through the Proton Synchrotron Booster (PSB) which accelerates the protons to an energy of 1.4 GeV. Further, the bunches are circulated in a bigger Proton Synchrotron (PS) accelerating protons up to an energy of 26 GeV. After that, the proton bunches are circulated inside the Super Proton Synchrotron (SPS) which accelerates them to an energy of 450 GeV. Finally, the bunches are injected into the Large Hadron Collider (LHC) ring. Where protons are accelerated to a speed of 99.99% of the speed of light. At such a large speed the proton bunches come so close that they look like a beam. The increment in the energy of protons at various stages are listed in Table 3.1.

A typical synchrotron consists of many components such as dipole magnets to bend the beams, quadrupole magnets to focus the beams, radio frequency (RF) cavities to accelerate the beams, cryogenics for superconducting and cooling the LHC ring, and beam diagnostics to monitor the beam movement. Two proton beams, moving in the opposite direction, are collided at four points of the LHC ring, as shown in Figure 3.2.

There are eight insertion regions (IRs) at the LHC ring as shown in Figure 3.2. At four of them, where

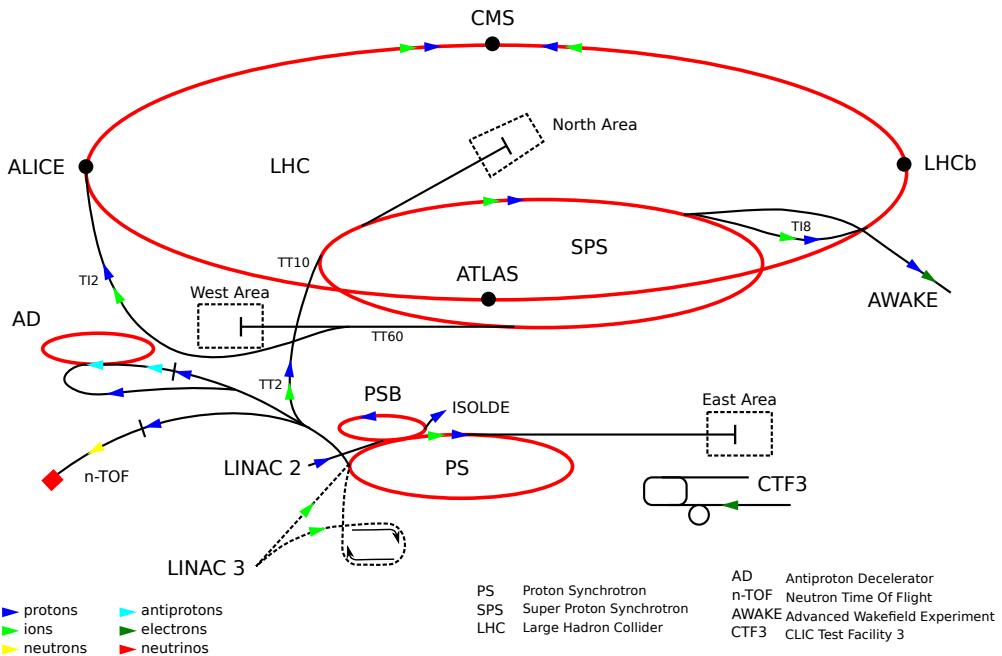


Figure 3.1.: The CERN accelerator complex [8]. The particles are accelerated at various stages, starting from LINAC to PSB to PS to SPS to LHC. Various detectors such as ATLAS, LHCb, CMS, and ALICE are placed at the collision points.

Table 3.1.: The energy of protons after passing through various accelerators.

LINAC2    50 MeV	
PSB	1.4 GeV
PS	26 GeV
SPS	450 GeV
LHC	7 TeV

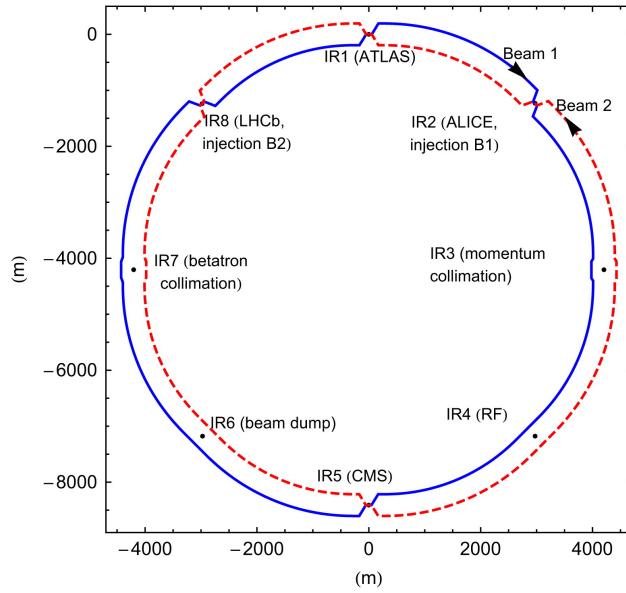


Figure 3.2.: The layout of insertion regions (IRs) at the LHC ring [9]. The separation between the two colliding beams is not to scale. The separation between them is exaggerated for illustration. The total circumference of the LHC ring is 27 km. Oppositely moving beams (Beam 1 and 2) are made to collide at four points of the ring. At the collision points, detectors such as CMS, ALICE, ATLAS, and LHCb are installed to record the collision.

collision happens, four main detectors are installed, ATLAS at IR1, ALICE at IR2, CMS at IR5, and LHCb at IR8. Three other small detectors namely LHCf, TOTEM, and MoEDAL are placed in the same cavern as ATLAS, CMS, and LHCb, respectively. At other IRs, the collimators are installed such as the momentum collimator at IR3, betatron at IR7, and RF at IR4. The beam dumping system is placed at IR6. Near the collision points, various systems such as dipole and quadrupole magnets are installed to bend and focus the proton beams as shown in Figure 3.3. Using these magnets the two counter rotating beams are brought closer for a head-on collision. A brief description of various detectors is given in the next section.

## 3.2. Detectors at the LHC

- **ATLAS (A Toroidal LHC ApparatuS):** The ATLAS [72] is a general purpose detector built to study the physics of the standard model and beyond it such as the origin of dark matter. It was one of the two detectors that discovered the Higgs Boson in 2012 [73]. It is 25 m in diameter, contains around 3000 km of cable, 46 m long, and weighs nearly 7000 tonnes.
- **LHCf (Large Hadron Collider forward):** The LHCf [74] is placed at IR1, on both sides of the ATLAS detector. It is designed to study collisions in the forward region that appreciates very high radiation.

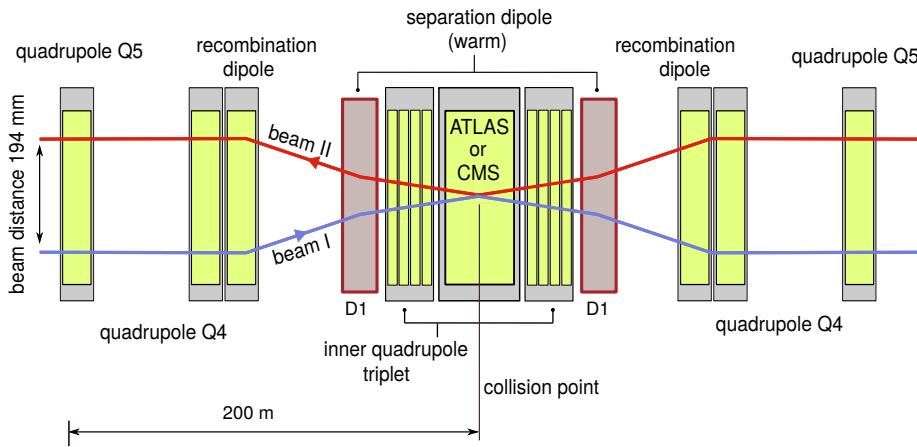


Figure 3.3.: Schematic diagram showing quadrupole and dipole magnets for focussing and bending proton beams at the interaction point (IP). The D1 dipole separate beams from both sides of the IP. This figure is adopted from [10].

One of the physics goals of LHCf is to study the neutral pions ( $\pi^0$ ) produced in collisions. A Proper measurement of the energy of  $\pi^0$  will help to understand and double check the origin of ultra-high-energy cosmic rays which have already been measured by other experiments such as the Telescope Array Project in Utah, and the Pierre Auger Observatory in Argentina.

- **ALICE** (A Large Ion Collider Experiment): The ALICE [75] detector is mainly designed for the study of heavy ion collisions such as proton-lead (p-Pb) and lead-lead (Pb-Pb). At the collision point, due to an extremely high temperature, the quark-gluon plasma is produced. It is believed that similar conditions existed just after the Big Bang where quarks and gluons were in a free state before combining to form hadrons. In 2011, the ALICE experiment measured the size of the fireball in Pb-Pb collision [76]. The ALICE with a weight of 10000 tonnes, weighs more than the Eiffel tower.
- **CMS** (Compact Muon Solenoid): The CMS [17] is a general purpose detector like the ATLAS. A detailed information about this experiment is given in Section 4.
- **TOTEM** (TOTal Elastic and diffractive cross section Measurement): The TOTEM [77] is a small detector placed at IR5 in the same cavern where the CMS is. As the name suggests, its aim is to measure the total cross-section, and to study diffractive processes and elastic scattering.
- **LHCb** (Large Hadron Collider beauty): The LHCb [78] experiment specially built for the study of physics processes involving b and c quarks such as the parameters of CP violation from the decays of b hadrons. During 2010-2012, LHCb has published many physics results including the measurement of branching fraction of the  $B_s \rightarrow \mu^+ \mu^-$  decay [79], the forward-backward symmetry of muon pair from the  $B_d \rightarrow$

$K^*\mu^+\mu^-$  decay, the properties of radiative B decays, the determination of the unitarity triangle parameters, and two-body charmless decay of B mesons.

- **MoEDAL** (Monopole and Exotics Detector At the LHC): The MoEDAL [80] detector is placed in IR8 adjacent to the LHCb. Its physics goal is to search for the existence of magnetic monopoles. So far it has not found any evidence for magnetic monopoles and has accordingly set an exclusion limit on their production cross section.

### 3.3. The LHC parameters

The proton beams have a number of associated parameters such as the energy of each beam, collision frequency, number of particles in each beam, luminosity etc. An accurate and updated knowledge of these parameters is very important during the data taking. Because they are used in the reconstruction of various physics objects, as well as to apply correction on simulated Monte Carlo samples for a better comparison of the observed data with simulated results. The information about these parameters is referred to as “non-event” data since they only contain information about the hardware configuration. These parameters are regularly updated on the online master database system (OMDS). They are later retrieved from the OMDS for offline usage using dedicated online-to-offline tools. There are thousands of such parameters. A few important ones are described below:

- **The number of particles in each beam:** One of the basic parameters about the beam is the number of particles in each beam. A proton beam consists of proton bunches. There are nearly 2556 proton bunches in each beam. A typical configuration of the bunch is shown in Figure 3.4(a). There are  $1.15 \times 10^{11}$  protons in one bunch. Each bunch is separated by 7.5 m from each other across the LHC ring. The collision frequency of the bunches is 25 ns, that is after each 25 ns counter rotating proton bunches are made to collide at the IP. There is another parameter called “bunches filled for each beam”. It tells out of 2556 bunches which are filled and which are empty during the data taking. Assuming that all the bunches of a beam are filled, the number of protons in each beam will be  $2556 \times 1.15 \times 10^{11} \approx 10^{14}$ .
- **Crossing angle ( $\alpha$ ):** The  $\alpha$  is the angle at which the two beams cross at the IP as shown in Figure 3.4(b). Its typical value lies in the  $\mu\text{rad}$  range. Even though  $\alpha$  is very small, a slight change in it affects the number of interactions happening at the collision point. A relatively smaller  $\alpha$  implies that the two colliding beams are closer to each other away from the IP. This results in a long-range collision away from the IP. However, a larger value of  $\alpha$  reduces the number of collisions as the overlapping area between the

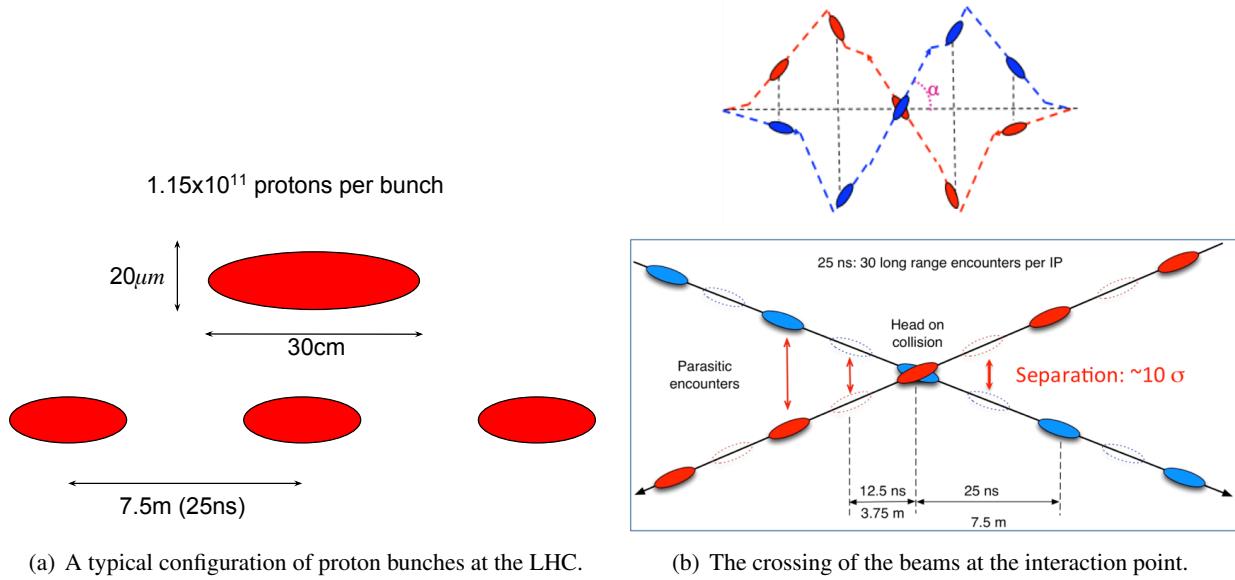


Figure 3.4.: The bunch configuration and crossing angle. The bunches are separated by 7.5 m. The crossing angle shown is exaggerated for illustration. Figure 3.4(a) is adopted from [11] and Figure 3.4(b) is taken from [12].

beams is smaller.

- **The  $\beta^*$ :** The beam sizes do not remain constant throughout the ring. Near the IP, the beams are bent using dipole magnets as shown in Figure 3.3. Therefore, the beams are a bit squeezed near the IP, the extent of which is quantified with a parameter, called  $\beta$ . The  $\beta^*$  is  $\beta$  value at IP. It is the distance from the IP up to a point where the beam size is doubled. The value of  $\beta^*$  lies in the mm range. Smaller value of  $\beta^*$  implies the beam is squeezed near the IP and vice-versa. The formula of  $\beta^*$  is given by [81]

$$\beta^* = \frac{\pi \sigma_b^2}{\varepsilon} \quad (3.1)$$

where  $\sigma_b^2$  is the cross-sectional size of the bunch and  $\varepsilon$  is the emittance, the smallest opening a beam can be squeezed through [81].

- **Center-of-mass energy ( $\sqrt{s}$ ):** The counter rotating protons collide head-on along the z-axis. In the center-of-mass frame,  $p_{1z} + p_{2z} = 0$ . In addition, the momentum in the transverse plane must be conserved. If there is an imbalance in  $p_T$ , it would imply the production of some (unknown) particles which has passed the detector without being detected; neutrino is an example. For the two colliding protons with energy  $E_1$  and  $E_2$ , the center-of-mass energy is given by

$$\sqrt{s} = E_1 + E_2 \quad (3.2)$$

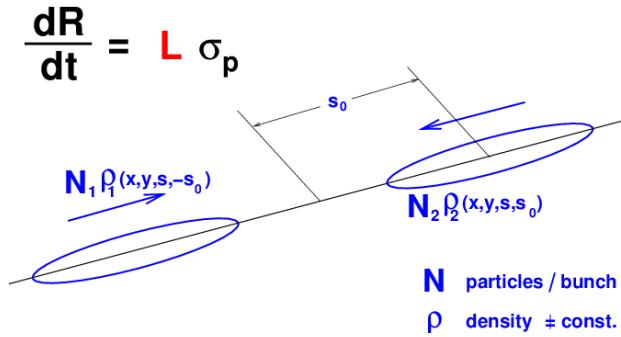


Figure 3.5.: Two oppositely moving bunches with the total number of particles  $N_1$  and  $N_2$ , and beam density distribution functions  $\rho_1(x, y, s, -s_0)$  and  $\rho_2(x, y, s, s_0)$  [13].

For the 2016-18 data taking, the energy of each beam was 6.5 TeV, leading to  $\sqrt{s} = 13$  TeV. The production of new particles in the physics process largely depends on  $\sqrt{s}$ . Higher the value of  $\sqrt{s}$ , more is the number of produced events for a given process. Many beyond the standard model theories predict the existence of a new particle at higher  $\sqrt{s}$ . Therefore, the  $\sqrt{s}$  is an important parameter of the LHC.

- **Luminosity:** Like the center-of-mass energy, luminosity is another important parameter which is a measure of how many collisions occur. Higher luminosity means a higher chance of producing rare physics process in the collisions. With higher luminosity, the standard model predictions can be tested with more accuracy thanks to more statistics, hence less uncertainty. It is defined as:

$$L = \frac{1}{\sigma_p} \frac{dR}{dt} \quad (3.3)$$

where  $\frac{dR}{dt}$  is the number of events per second and  $\sigma_p$  is the cross section. In the CGS unit, the dimension of  $L$  is  $\text{cm}^{-2} \text{s}^{-1}$ . However,  $\text{fb}^{-1}$  (inverse femtobarn) is often used in the collider experiments. The luminosity is recorded by experiments over a certain period of time. The integrated luminosity is the integration of instantaneous luminosity over a period of time

$$L_{\text{int}} = \int L dt \quad (3.4)$$

At the LHC, the luminosity is calculated in terms of beam parameters such as number of particles ( $N$ ), collision frequency ( $f$ ), the number of bunches in the beam ( $N_b$ ), etc [13]. For two counter rotating bunches with total number of particles  $N_1$  and  $N_2$ , and beam density distribution functions  $\rho_1(x, y, s, -s_0)$  and  $\rho_2(x, y, s, s_0)$  as shown in Figure 3.5, the luminosity is given by [13]

$$L = 2N_1 N_2 f N_b \int \int \int \int \rho_1(x, y, s, -s_0) \rho_2(x, y, s, s_0) dx dy ds ds_0. \quad (3.5)$$

where  $s$  corresponds to the  $z$ -coordinate which is along the beam direction and  $s_0$  is the 4<sup>th</sup> component of

space-time coordinate, that is,  $s_0 = ct$ . As shown in Figure 3.5, the two bunches meet at  $s_0 = 0$ . Assuming that the two bunches collide head-on and the beam density is uncorrelated along all directions, Equation (3.5) can be written as

$$L = 2N_1 N_2 f N_b \int \int \int \rho_{1x}(x) \rho_{1y}(y) \rho_{1s}(s - s_0) \rho_{2x}(x) \rho_{2y}(y) \rho_{2s}(s + s_0) dx dy ds ds_0. \quad (3.6)$$

The beam density distribution at the LHC follows a Gaussian distribution, that is, the proton density is more in the middle of the beam and lesser as one goes away from the middle point. Therefore, for the two beams, the beam density functions in Equation (3.6) can be written as

$$\rho_{ix}(x) = \frac{1}{\sigma_{ix} \sqrt{2\pi}} \exp\left(-\frac{(x)^2}{2\sigma_{ix}^2}\right), i = 1, 2 \quad (3.7)$$

$$\rho_{iy}(x) = \frac{1}{\sigma_{iy} \sqrt{2\pi}} \exp\left(-\frac{(y)^2}{2\sigma_{iy}^2}\right), i = 1, 2 \quad (3.8)$$

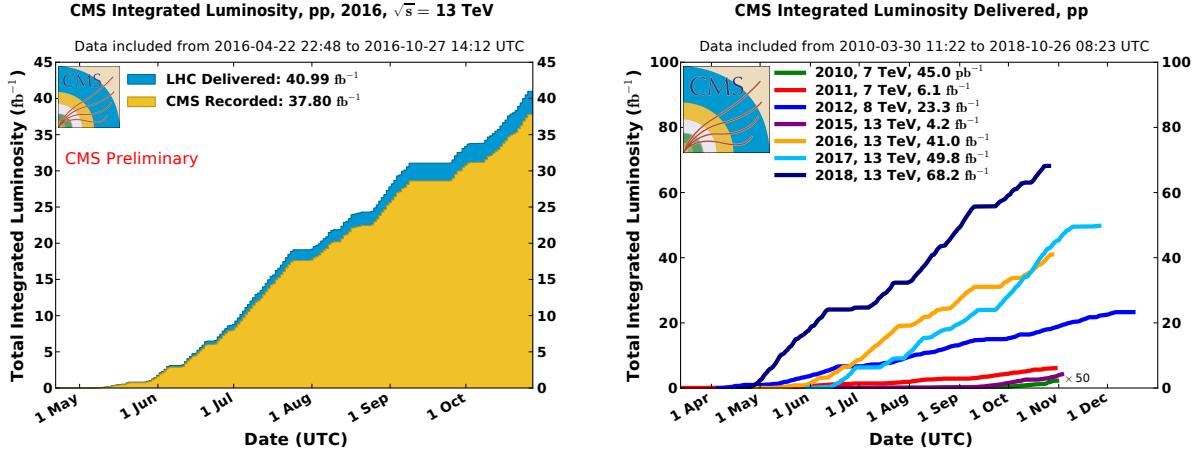
$$\rho_s(s \pm s_0) = \frac{1}{\sigma_s \sqrt{2\pi}} \exp\left(-\frac{(s \pm s_0)^2}{2\sigma_s^2}\right) \quad (3.9)$$

where  $\sigma_{ix}$  and  $\sigma_{iy}$  are the standard deviation of Gaussian distribution of the two beams along the  $x$  and  $y$ -direction. Using these, Equation (3.6) becomes [13]

$$L = \frac{N_1 N_2 f N_b}{2\pi \sqrt{\sigma_{1x}^2 + \sigma_{2x}^2} \sqrt{\sigma_{1y}^2 + \sigma_{2y}^2}} \quad (3.10)$$

Equation (3.10) holds in the ideal situation where the beam profile is uncorrelated in all directions and the machine operates in an ideal condition. However in real life, there are various additional machine effects that need to be incorporated such as finite crossing angle, collision offset, hourglass effect, non-Gaussian beam profiles, nonzero dispersion at the collision point, and so on [13]. The luminosity given by Equation (3.10) is what delivered by the LHC. Each detector records the luminosity individually during the collision. The recorded luminosity is always smaller than the delivered one due to the detector effects. Five sub-detectors of the CMS experiment are used to monitor and measure the luminosity. These are the silicon pixel detector, the Hadron Forward Calorimeter (HF), the Drift Tubes in the barrel (DT), Pixel Luminosity Telescope (PLT), and the Fast Beam Conditions Monitor (BCM1f) [82].

The luminosity delivered by the LHC and recorded by CMS for 2016 data taking is shown in Figure 3.6(a) over a period of time. The integrated recorded luminosity is  $37.8 \text{ fb}^{-1}$  for the 2016 data taking. However, a luminosity mask is applied to reject some amount where some parts of the CMS detector were not operating in the normal mode. Accordingly, the recorded luminosity reduced to  $35.9 \text{ fb}^{-1}$ . The analysis of 2016 data is what presented in this thesis. Delivered luminosity for different center-of-mass energies



(a) Delivered luminosity by the LHC and recorded by CMS in 2016 data taking.  
(b) Delivered luminosity by the LHC for different years at the different center-of-mass energies, starting from 2010 to 2018.

Figure 3.6.: Luminosity measurement from different years of the data taking [14]. In this thesis, the analysis presented uses the one for the 2016 data taking (golden yellow line in the right plot). The total delivered luminosity at 13 TeV from all years is about  $160 \text{ fb}^{-1}$ .

and during different years are shown in Figure 3.6(b). The total delivered luminosity at 13 TeV during 2015-18 is about  $160 \text{ fb}^{-1}$ .

At the LHC, there are various beam modes indicating the beam status such as “Squeeze” (the beams are squeezed by betatron towards the target collision, for most of the 2016 data taking the  $\beta^* = 30 \text{ cm}$ ), “Adjust” (the separation between the beams is made to collapse, bringing them for collision), “Stable beams” (it signals that the physics data taking can be started), and “Dump” (the beams are dumped after the data taking ends). The stable beam mode lasts for a few hours during which the collisions are recorded.

A “Fill” is created as soon as the stable beams are achieved and all the detector information is updated in the online database. A Fill lasts for few hours and is divided in “Runs” for better management. Further, a Run is divided in “lumi sections”. The duration of each lumi section (LS) is 23 s. Therefore, the data taking by a detector has a granularity of 23 s. That is, the detector information is updated/stored after each 23 s. An LS is marked as good or bad depending on whether all the parts of the detector were on or some of them were off. The integrated luminosity is obtained by summing over all LS.

All the LHC parameters are measured regularly during the data taking. Most of them do not change on a regular day-to-day basis. However, few do change on every day during the data taking such as the ones shown in Figure 3.7. Other parameters change over a period of time, for example on a month-to-month basis, and some even on a yearly basis such as the ones shown in Figure 3.8.

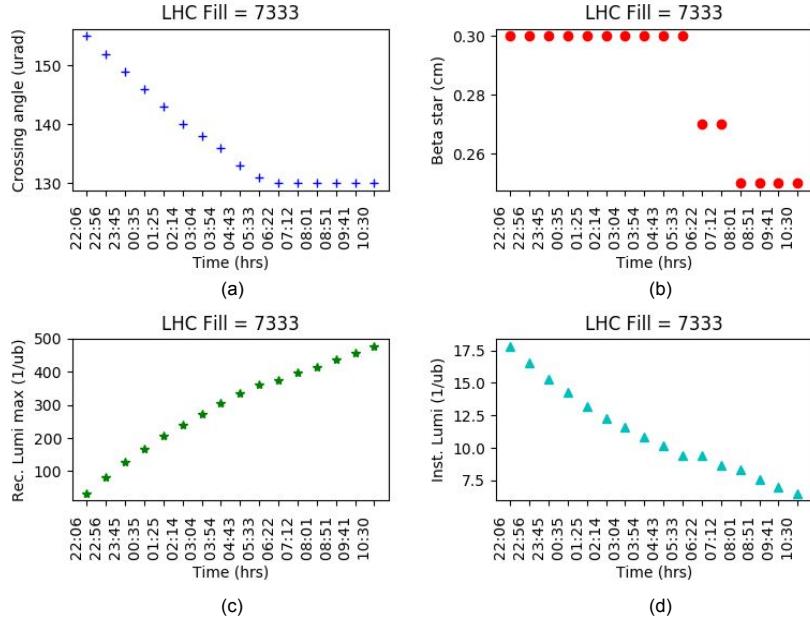


Figure 3.7.: Variation of few LHC parameters during the data taking on 22-23rd October 2018 (Fill = 7333). All these parameters change over a different point of time. Other parameters such as the center-of-mass energy, number of particles in each beam, etc do not change on a given day.

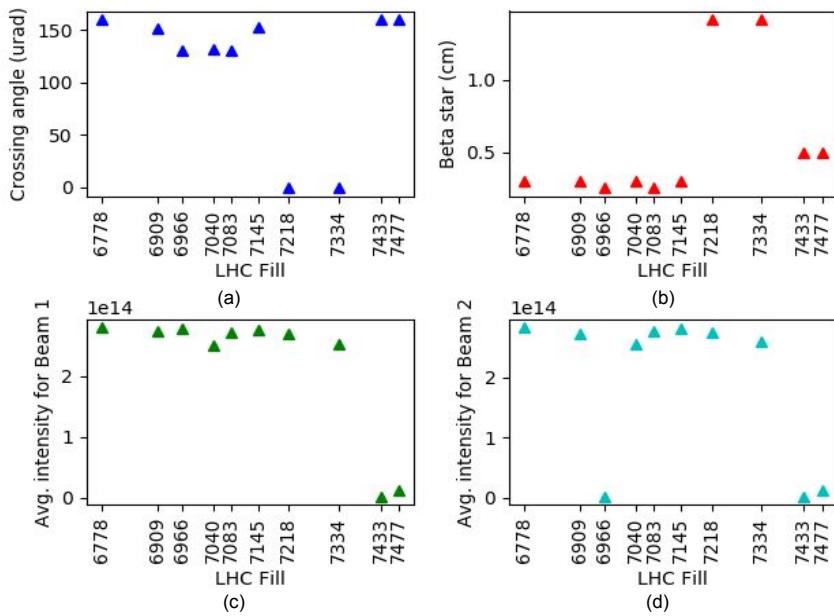


Figure 3.8.: Variation of few LHC parameters during the data taking from random days in the period of June-November 2018. Only a few days (Fills) have been chosen for illustration. The last two Fills (7433 and 7477) correspond to Pb-Pb, rest are from proton-proton collisions. The average intensity of the beam is the number of particles in each beam. As can be seen, all these parameters change over a different period of time.

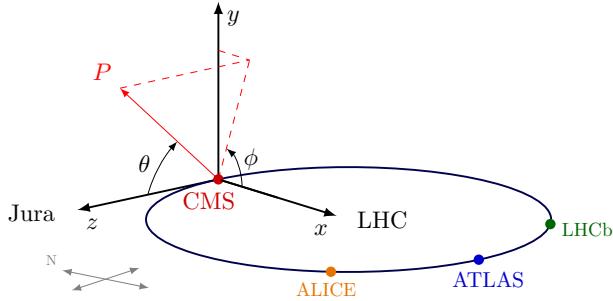


Figure 3.9.: The LHC coordinate system [15] with the  $z$ -axis lying along the beam direction. At CMS, the  $z$ -axis points towards the mountain Jura in France.

## 3.4. The LHC coordinate system

Most of the detectors placed at the IRs have a cylindrical shape with the axis of the cylinder lying along the beam direction. Within the detectors, various sub-detectors are placed either parallel (barrel region) to the beam direction or perpendicular (endcap region) to it. At the LHC, the  $z$ -axis lies along the beam direction, while the  $x$ - $y$  plane is perpendicular to it as shown in Figure 3.9. The following derived quantities are often used by the LHC experiments:

- **Transverse momentum ( $p_T$ ):** Out of  $10^{11}$  protons in a bunch, only about 20 hard collisions occur in one bunch crossing. Rest of the protons move in the forward region, some of them with soft collision and some even without being collided. The hard collisions produce actual physics process of interest. Due to the head-on collision, the particles produced in hard collisions mostly go in the transverse plane. The momentum in the  $x$ - $y$  plane,  $p_T$ , is defined as

$$p_T = \sqrt{p_x^2 + p_y^2} \quad (3.11)$$

where  $p_x$  and  $p_y$  are the  $x$  and  $y$  components of the momentum vector, respectively. Detecting a higher  $p_T$  in an event indicates the occurrence of actual physics process.

- **Pseudorapidity ( $\eta$ ):** As shown in Figure 3.9, the azimuthal angle ( $\phi$ ) is the angle between  $\vec{p}_x$  and the  $x$ -axis. It covers the full barrel region and ranges from 0 to  $2\pi$ . On the other hand  $\theta$  is the angle between the momentum vector  $\vec{p}$  and  $z$ -axis which varies from 0 to  $\pi$  to cover both sides of the endcap. In the collider experiments, one measures 4-vector  $(E, \vec{p})$  of a particle. Just looking at the 4-vector, its not easy to predict the direction of the particle with respect to the beam axis. In view of this, there is a quantity called “rapidity”, given by [83]

$$y = \frac{1}{2} \ln \left( \frac{E + p_z}{E - p_z} \right) \quad (3.12)$$

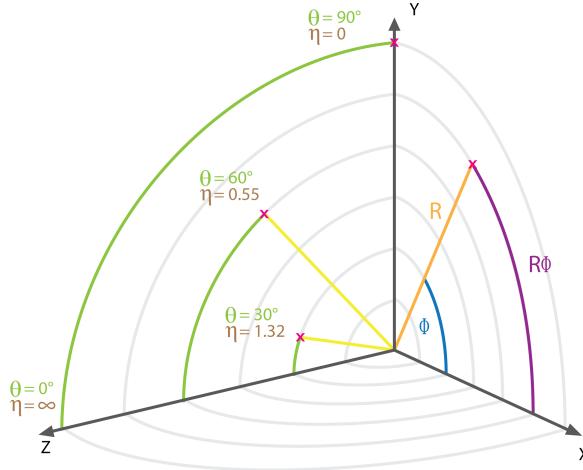


Figure 3.10.: The relationship between  $\theta$  and  $\eta$  [16]. The value of  $\eta = 0$  implies  $\theta = \pi/2$ . That is, lower the value of  $\eta$  the farther away we are from the beam axis in the perpendicular direction and vice versa.

which is very useful to know the value of  $\theta$ . For example,  $y = 0$  implies that  $E > p_z$ . That is, the particle is produced in the transverse plane which implies  $\theta = \pi/2$ . For  $y = \pm 1$ ,  $E < p_z$ . That is the particle is moving mostly in the beam direction implying  $\theta = 0$  or  $\pi$ . The another advantage of rapidity is that the difference in rapidities is Lorentz invariant. For example, under the Lorentz boost, the 4-momentum transforms as  $E' = \gamma(E - \beta p_z)$ ,  $p'_x = p_x$ ,  $p'_y = p_y$ ,  $p'_z = \gamma(p_z - \beta E)$ , so the rapidity is transformed as

$$y' = y - \tanh^{-1} \beta \quad (3.13)$$

where  $\beta = v$  ( $v$  is the velocity of the boosted frame), and  $\gamma = \sqrt{1 - \beta^2}$ . If an observer measures rapidities of two particles in the boosted frame then difference  $y'_1 - y'_2 = y_1 - y_2$ . That is, the difference in rapidities is Lorentz invariant. This is very useful when one is interested in the angular separation between two physics objects such as a muon and a hadronic jet.

The relativistic limit of rapidity is called pseudorapidity ( $\eta$ ). In such limit, the momentum of a particle is much larger than its rest mass , that is,  $E = \sqrt{p^2 + m_0^2} \approx p$ . Therefore, Equation (3.12) becomes

$$y = -\ln \tan\left(\frac{\theta}{2}\right) \rightarrow \eta \quad (3.14)$$

The  $\eta$  variable is mostly used everywhere at the LHC as it is directly related to the angle  $\theta$ . The relationship between  $\eta$  and  $\theta$  is shown in Figure 3.10. One can see that for a large value of  $\eta$ , e.g.,  $\eta = 4(-4)$ , the particle is close to the beam axis [ $\theta = 0(\pi)$ ]. Similarly, the lower value of  $\eta$  indicates that the particle is produced in the transverse plane. The lower and higher values of  $\eta$  correspond to barrel and endcap

region of the detector, respectively.



# 4 | The CMS Experiment at The LHC

## 4.1. Introduction

The Compact Muon Solenoid (CMS) detector is installed in IR5 at the LHC ring. CMS is one of the biggest international collaborations involving 43 countries, 199 institutes, with around 4000 people. The weight of the CMS detector is about 14000 tonnes (around the weight of 2500 African elephants). Such a huge weight is accommodated in a small volume (the shape of CMS is cylindrical with length of 21.5 m, and diameter of 15.6 m). That is why the “compact” word has been attached to it. One of the main physics goals of the CMS experiment is to detect “muons” as they are produced in most of the Standard Model processes. Of course, other particles such as an electron, photon, neutral and charged hadrons are also detected. A huge “solenoid” magnet is placed inside the CMS to bend the tracks of charged particles so that their momentum can be measured precisely. An image of the CMS experiment is shown in Figure 4.1, while various parts of the detectors are shown in Figure 4.2. The beam axis is along the center of the cylinder. The tracker (silicon pixel and strip) is the first sub-detector followed by the electromagnetic calorimeter (ECAL), and the hadron calorimeter (HCAL). The magnetic solenoid is placed outside the HCAL followed by the muon chambers. The iron yoke provides stand for the muon chambers and contains the magnetic flux outside the solenoid. A detailed description of each sub-detector is given in the next sections.

The trajectories of various particles inside the CMS detector are shown in Figure 4.3. The particles are produced at the IP and move towards various layers of the detector. The charged particles are bent thanks to the presence of a strong magnetic field inside the solenoid. Outside of it, the charged particles are bent in the opposite direction. On the other hand, the neutral particles do not bend inside the detector. The particles detected by the CMS experiment are muon, electron, charged hadron (e.g. pion), neutral hadron (e.g. neutron), and photon.

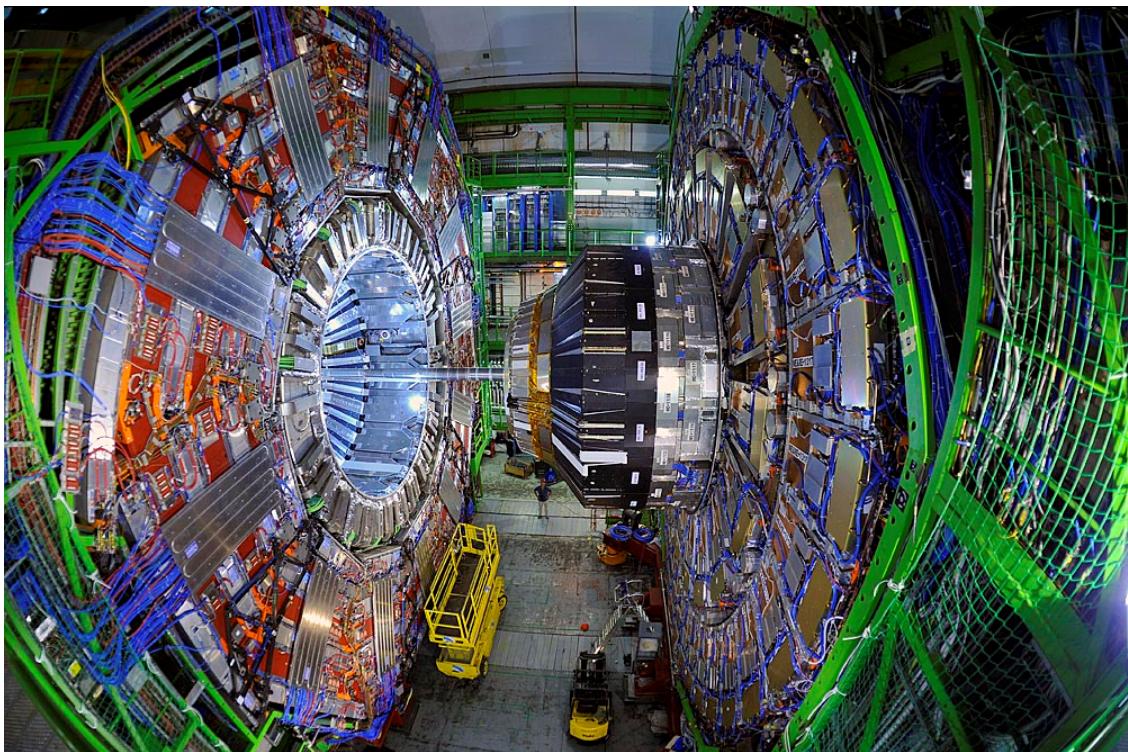


Figure 4.1.: The CMS detector installed in IR5 at the LHC [17].

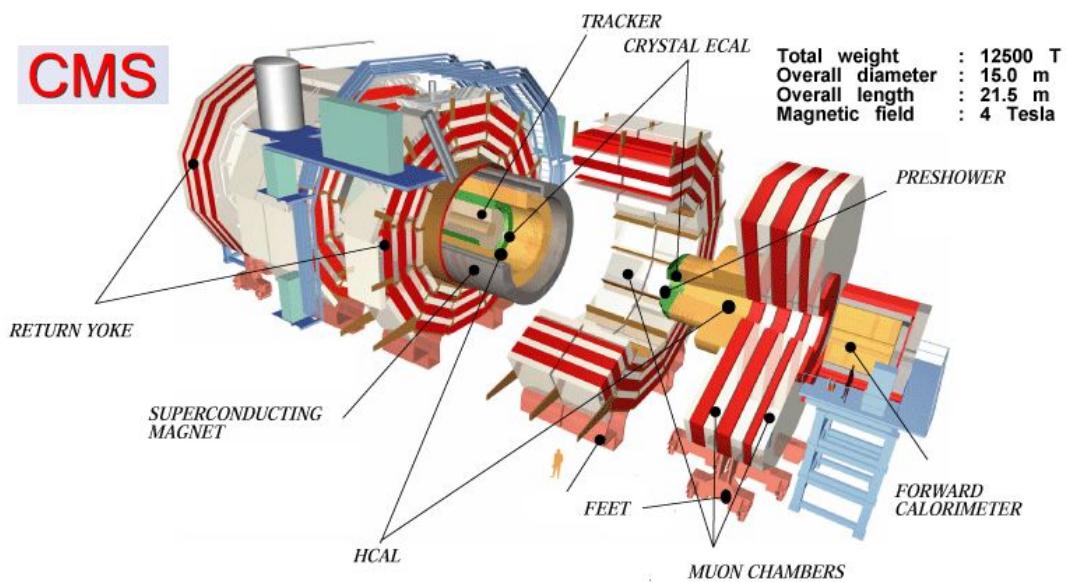


Figure 4.2.: Diagram of the CMS detector [18]. Various parts of the CMS detector are shown starting from tracker to ECAL to HCAL to magnet to the muon chambers.

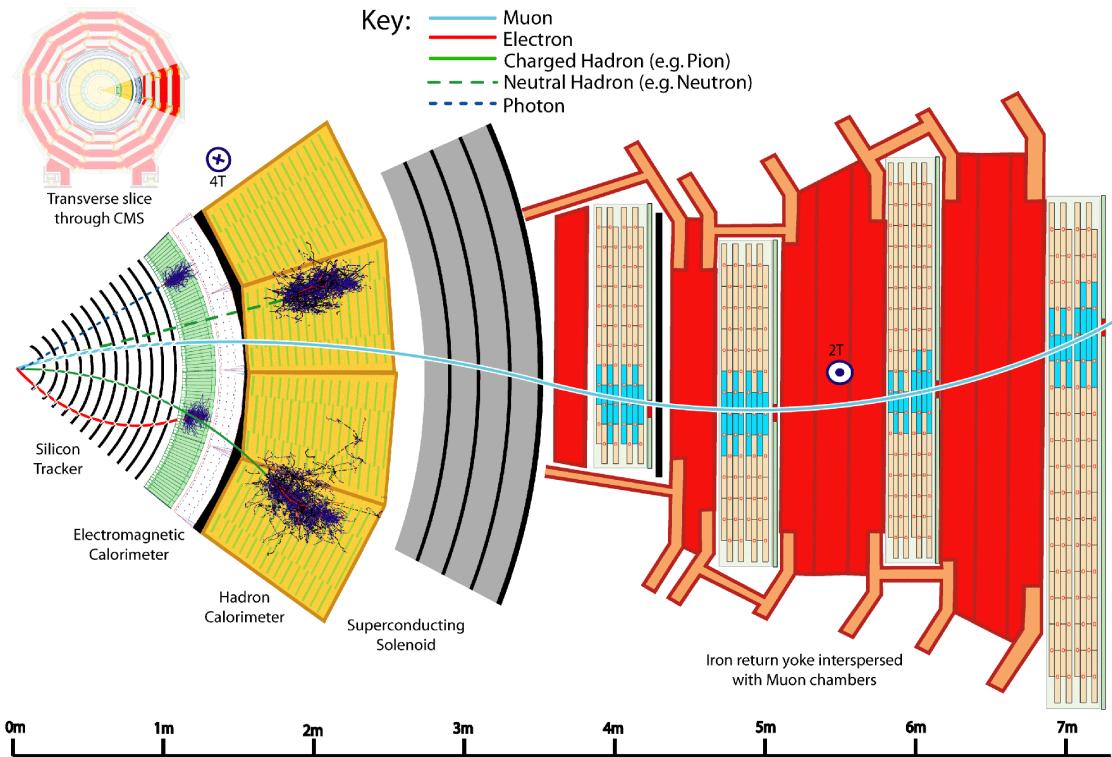


Figure 4.3.: Trajectories of various particles inside the CMS detector [19]. The charged particles such as electrons, muons, and charged pions ( $\pi^\pm$ ) are bent inside and outside the solenoid whereas the neutral particles such as photons and neutral pions ( $\pi^0$ ) traverse without being bent. The silicon tracker is mainly for measuring the momentum of particles whereas the ECAL and HCAL measure the energy deposits. The electrons and photons deposit their energy inside the ECAL whereas the hadrons deposit in both ECAL and HCAL. The muons are the only particles that traverses all the way to the muon chambers.

## 4.2. Silicon tracker

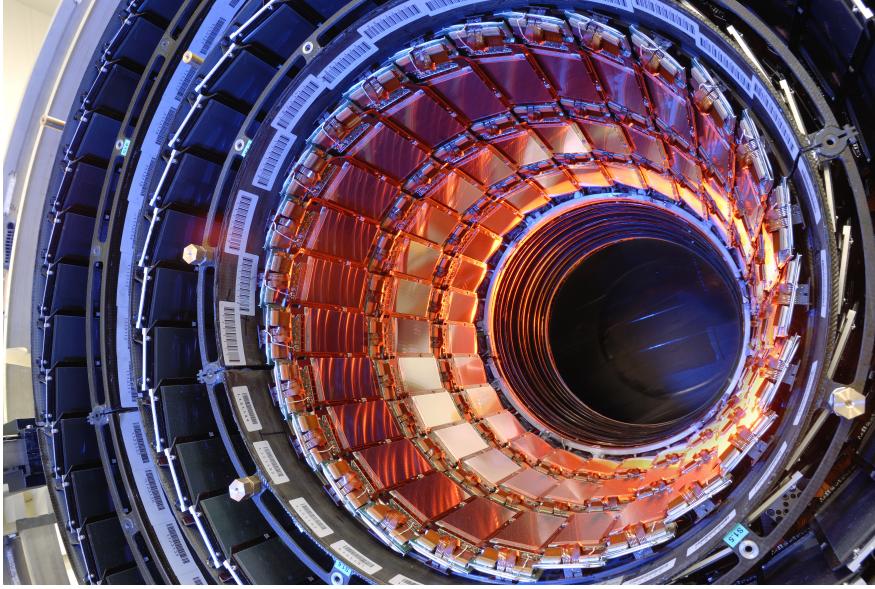
Due to the small band gap ( $1.12\text{ eV}$ ), high carrier mobility ( $1450\text{ cm}^2/\text{Vs}$ ), and high specific density ( $2.33\text{ g/cm}^3$ ) the silicon detectors are used in the CMS experiment for the measurements of trajectory and momentum of the particles produced in the collision. The silicon tracker of the CMS is closest to the beam axis. Its location is shown in Figure 4.4. It has 3 (2) layers of pixel tracker in the barrel (endcap) region, and 10 (9) layers of the strip tracker in the barrel (endcap) region. There are about 48 (18) million pixels in the barrel (endcap) region covering an area of  $200\text{ m}^2$  which makes it the biggest silicon detector ever built in the world.

The silicon detector is a reverse biased pn junction, whose principle of operation is shown in Figure 4.5. Under the reverse bias voltage ( $80\text{ V}$ ), when a particle (for example, a minimum ionising particle) traverses through the bulk n-type region, several electron-hole pairs (about 24,000) are created in the depletion region. The electrons, thus created, are attracted towards the  $\text{p}^+$  implant due to the applied electric field within very short time of about  $10\text{ ns}$ . The electrons collected at  $\text{p}^+$  implant are amplified and readout by the backend electronics.

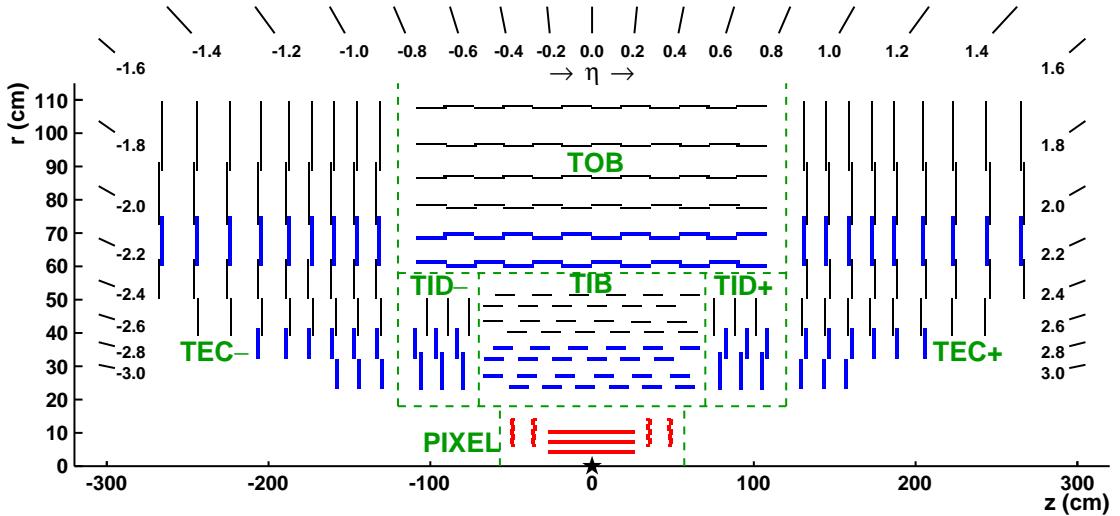
As alluded earlier, the CMS silicon tracker is divided into two types:

- **Pixel tracker:** The innermost component of the tracker, the pixel detector occupies the central region close to which the collision happens. It operates in a high-radiation environment. The length of the pixel detector from interaction point along the beam axis is  $46.5\text{ cm}$ , while two endcaps are placed at  $z = 34.5$  and  $46.5\text{ cm}$  from both sides of the IP. The 3-layers of the detector are placed at  $r = 4.4, 7.3,$  and  $10.2\text{ cm}$  in the barrel region as shown in Figure 4.6.
- **Strip tracker:** The major part of the tracker is covered by the silicon strip detectors as shown in Figure 4.4. The length of strip detectors is  $5.8\text{ m}$ , and diameter is  $2.5\text{ m}$ . There are over 15000 modules covering  $0 \geq \eta \geq 2.5$ . The charge generated, when a particle passes through the strip detectors, is read and amplified from each micro-strip as shown in Figure 4.5.

The total thickness of tracker material in units of radiation length (it is a characteristic of a material defined as the distance traveled by an electron when it loses  $1/e$  of its energy due to the Bremsstrahlung radiation) and nuclear interaction length (it is the mean distance a hadronic particle travels before undergoing an inelastic nuclear interaction) is shown in Figure 4.7. In the direction perpendicular to the  $z$ -axis ( $\eta = 0$ ), the thickness is small ( $t/X_0 \approx 0.4$ ). That is, an electron produced at collision point can traverse in the  $\eta = 0$  direction by losing only 14.7% ( $0.4/2.718$ ) of its initial energy. The thickness is maximum in the endcap



(a) The inner part of one of the tracker inner barrels (TIBs) of the CMS experiment [84].



(b) The position of silicon tracker on both sides of the IP ( $z = 0$  cm) [21]. The pixel trackers are very close to the IP. For the strip tracker, there are tracker inner barrel (TIB), tracker outer barrel (TOB) in the barrel region, tracker endcap (TEC), and tracker inner disk (TID) in the transition region. The black (blue) lines correspond to single (double) sided silicon strips.

Figure 4.4.: The silicon tracker of the CMS. The picture of one of trackers inner barrel is shown in (a). The radial and  $\eta$  position of the various parts of the pixel and strip tracker is shown in (b).

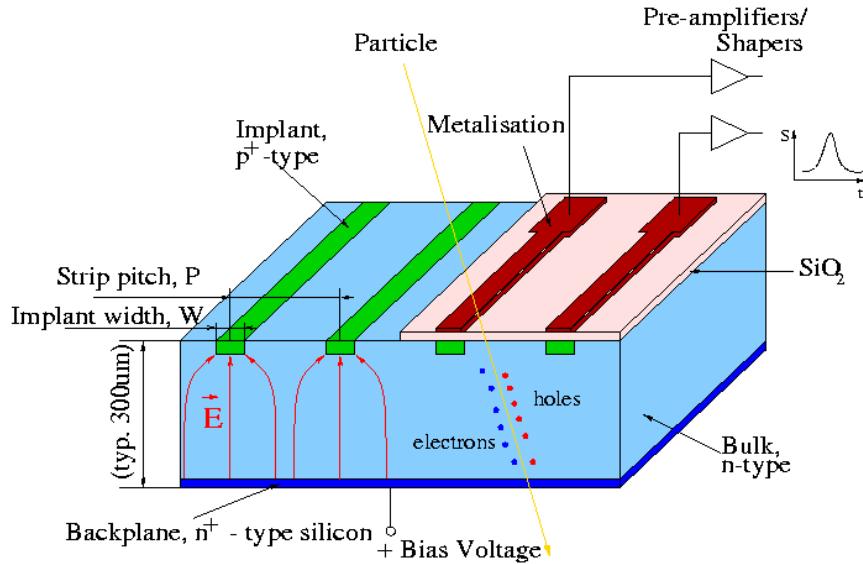


Figure 4.5.: The principle of operation of a silicon detector [20]. The green strips are  $p^+$ -type of semiconductor where the electrons are attracted after being produced in the depletion region when a particle passes through. The  $\text{SiO}_2$  layer along with metallic surface, on top of the green strips, provides conducting circuits for the electron. The amplifiers/shapers are used to amplify and shape the electron current. The holes are attracted towards the +ve bias voltage.

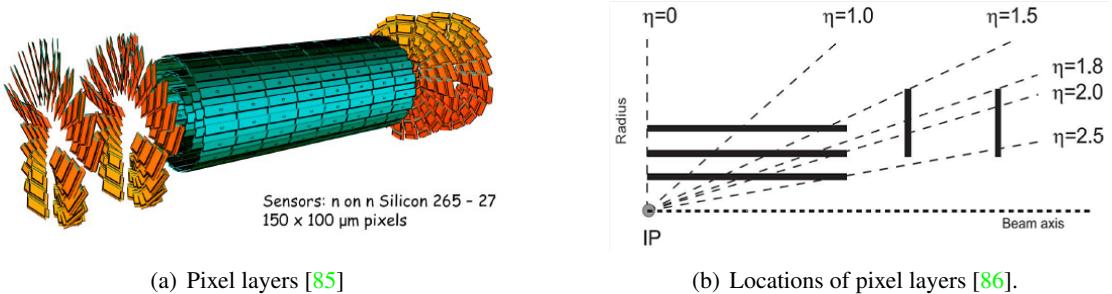


Figure 4.6.: The pixel detector of the CMS tracker. The location of pixel layers in the barrel and endcap region is shown in (b).

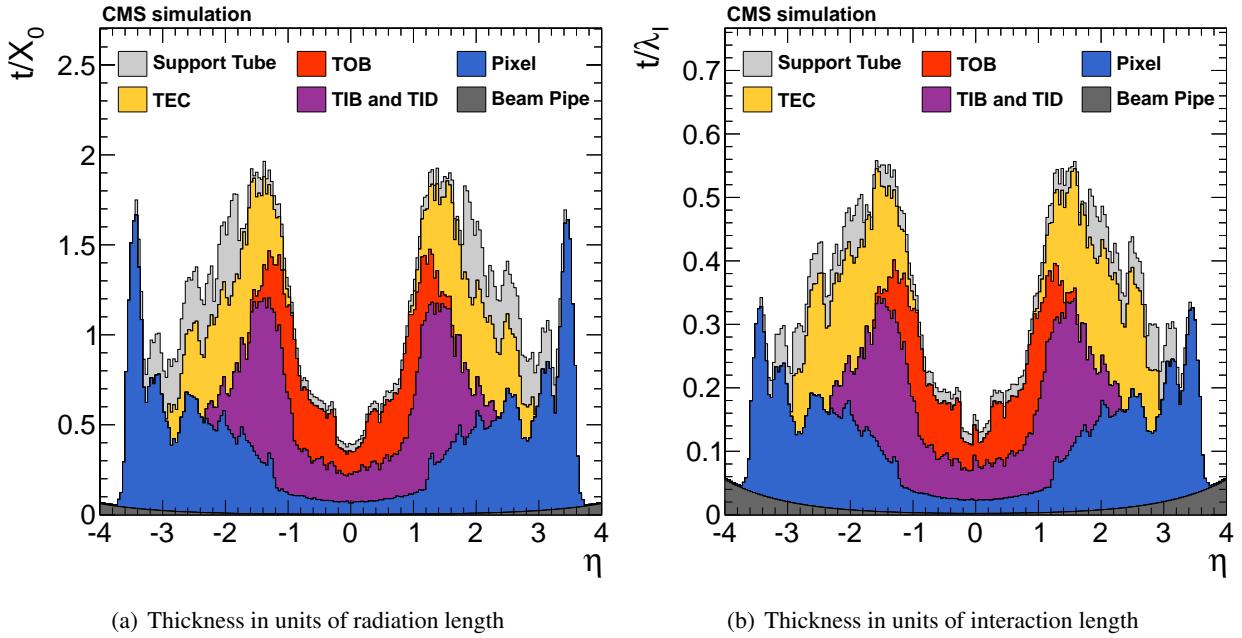


Figure 4.7.: Total thickness of the silicon tracker in units of radiation and interaction lengths [21]. The thickness of supporting tube and beam pipe is also shown.

region as shown in Figure 4.7.

The resolution is the parameter that characterizes how accurately a measurement is performed. For example, if a detector “measures” the momentum of a particle to be 100 GeV and the momentum “resolution” is 5 GeV then it implies that the “actual” value of the momentum would be in the range of 95 to 105 GeV. Hence lower the value of detector resolution, the more precise is the measurement. The absolute (relative) resolution of  $\phi$  ( $p_T$ ) of the silicon tracker as a function of  $\eta$  is shown in Figure 4.8 of muons, electrons, and charged pions [21]. The  $\eta$  and  $p_T$  resolution of all particles are better in the barrel region as compared to that in endcap regions. The resolution for muons is better as compared to electrons and charged pions in all regions. Also, the higher the  $p_T$ , better is the resolution for all charged particles. For  $p_T = 100$  GeV, the momentum resolution at  $\eta = 0$  is 2%, 30%, and 3% for muons, electrons, and pions respectively. The silicon tracker is essential for the reconstruction of trajectories of charged particles. Being very close to the collision point, these are also useful for high level triggering as discussed in Section 4.7.

### 4.3. Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) is used to measure the energy deposited by electrons, photons, and jets. Due to the high collision rate (collisions happen after every 25 ns), the ECAL has to be very fast

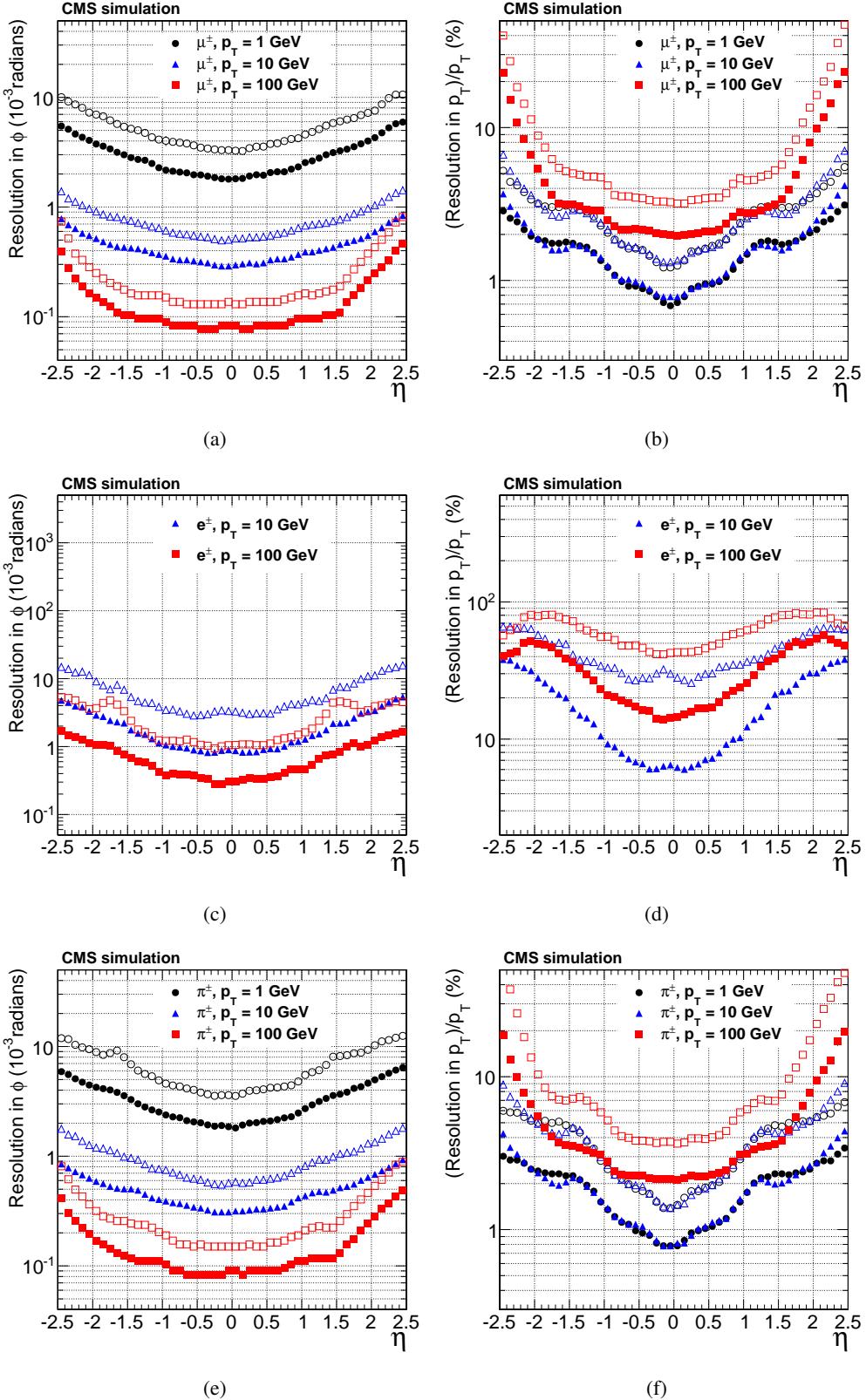
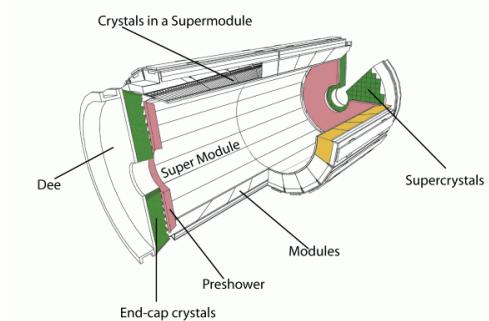


Figure 4.8.: The absolute resolution of  $\phi$  and relative resolution of  $p_T$  measurement of muons, electrons, and charged pions from the silicon tracker as function of  $\eta$  [21]. Other resolution such as  $d_0, z_0$ , and  $\cot\theta$  are given in Reference [21]. The open (solid) markers correspond to the half-width for 90% (68%) intervals [21].

(a) PbWO<sub>4</sub> crystal [17]

(b) Components of the ECAL [17]

Figure 4.9.: (a) A PbWO<sub>4</sub> scintillating crystal. (b) The various components of the electromagnetic calorimeter in the barrel and endcap regions.

and responsive. In view of this, most part of ECAL is made of lead tungstate (PbWO<sub>4</sub>) crystals which are birefringent, tetragonal, and radiation-hard [87]. A small part is made of the Pb absorber (placed in the endcap region, primarily to differentiate events coming from  $H \rightarrow \gamma\gamma$  and  $\pi^0 \rightarrow \gamma\gamma$ ). A typical PbWO<sub>4</sub> crystal is shown in Figure 4.9(a). The volume of a PbWO<sub>4</sub> crystal is roughly equal to the volume of a coffee cup, and the weight of each crystal is 1.5 kg (the density is 8.3 g/cm<sup>3</sup>). The PbWO<sub>4</sub> crystals are grouped in to form modules, and the modules are grouped to form super-modules. These super-modules are placed in barrel and endcap region of the ECAL as shown in Figure 4.9(b). The ECAL is made up of 76,000 lead tungstate crystals (61,200 in the barrel and 14,800 in the endcap region).

When a particle passes through the PbWO<sub>4</sub> crystals, photons are produced by the scintillating process. The PbWO<sub>4</sub> are inorganic scintillator doped with impurities. Other inorganic scintillators are alkali halide (NaI+TI, CsI+Na), Cerium-activated fast inorganics (LaBr<sub>3</sub>), etc. Pure crystals such as NaI (octahedral) don't produce visible photons so impurities such as TI (activators) are added which create special sites in the lattice as shown in Figure 4.10. When a particle passes through the crystal it creates electron-hole pairs. The electrons are excited to the conduction band whereas the holes ionise the impurity atoms. The excited electron recombines the ionised impurity atoms creating its own excited configurations. The impurity excited configuration de-excites to the ground state by emitting photons.

The photons thus produced are converted into an electrical signal by the photomultiplier tubes (PMTs) or avalanche photodiodes (APDs). The APDs in the barrel and vacuum phototriodes (VPTs) are placed in the endcap region. The electrical signals are amplified and stored. The working principle of PMT is shown in Figure 4.11(a). The photons are absorbed on photocathodes. Due to the photoelectric effect, electrons are emitted, if the energy of incoming photons is more than  $\approx 3$  eV. The emitted electrons are directed towards

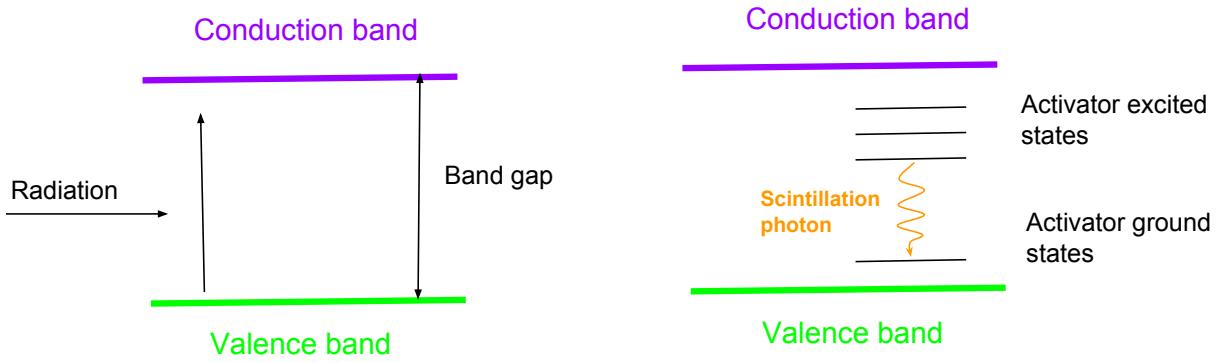
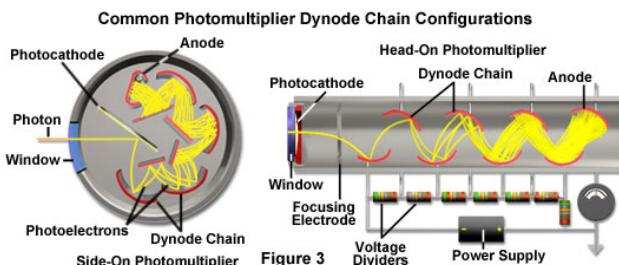
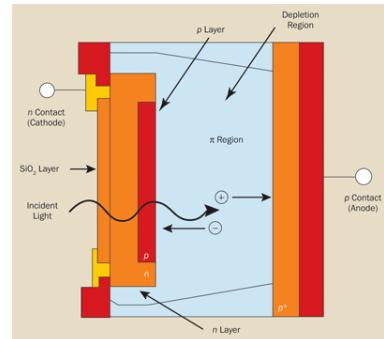


Figure 4.10.: The energy band structure in pure (impure) inorganic scintillator shown on left (right) side. The impurity creates intermediate excited states between the valence and conduction band. The photons are emitted when activator de-excites.



(a) Schematic diagram of photomultiplier tube [88]



(b) Principle of operation of the avalanche photo-diode [89]

Figure 4.11.: The schematic diagram of a circular and cylindrical PMT (a). The principle of operation of an APD is shown in (b).

the dynodes by an electric field. From the dynodes, the secondary electrons are emitted with an efficiency (the ratio of number of secondary and primary electrons) of  $\approx 6\text{-}8$ . These electrons are then directed towards other dynodes and the process continues. In the end, there are many electrons. A schematic diagram of APD is shown in Figure 4.11(b). When a photon enters the depletion region, electron-hole pairs are created. The created electron-holes start moving towards the respective electrode due to large reverse bias voltage. While moving, they create additional electron-hole pairs. The process continues producing an avalanche multiplication.

An electromagnetic pre-shower, as shown in Figure 4.12(a) is placed in the endcap region at  $z = 298.5$  cm to distinguish photons coming from Higgs and pion. The photons produced by the pion ( $\pi^0$ ) are likely to be narrowly separated and will be traveling in the forward direction. However, those coming from the Higgs

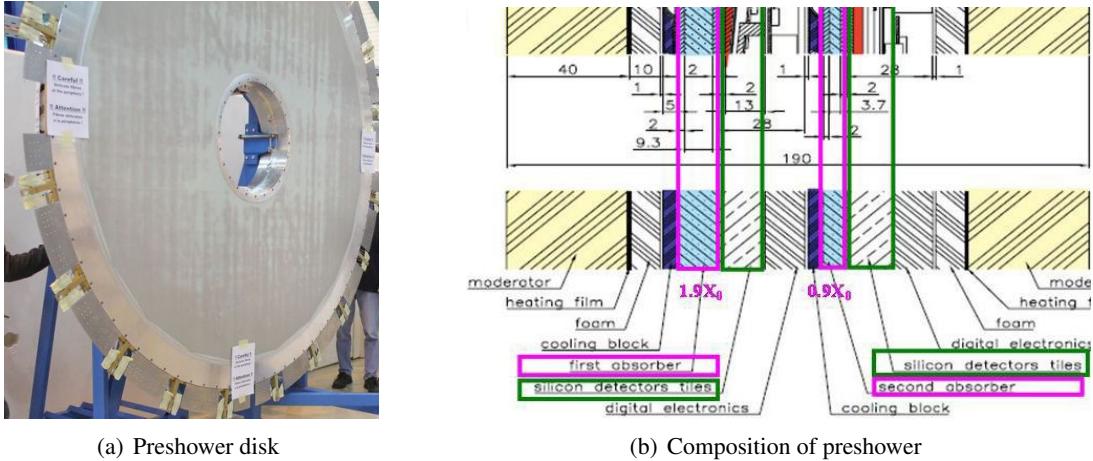


Figure 4.12.: ECAL preshower [22] placed in the endcap region at  $z = 298.5$  cm. The preshower is made of alternate layers starting from the moderator, to heating film to foam to cooling block to the absorber (Pb) to silicon detectors to the digital electronics. The width of the lead (Pb) absorber is not the same on both sides.

are well separated and mostly in the transverse direction. Therefore, if there is a photon in an event which was reconstructed using the preshower then it most likely came from the  $\pi$  decay and can be rejected for the study of  $H \rightarrow \gamma\gamma$  process. The preshower has a diameter of 240 cm with a thickness of 20 cm. It has alternate layers of absorber and detectors on both sides of its center. Electromagnetic showers are produced from lead layers when a photon passes through it, which silicon sensors detect. The two layers are used to determine the position of the particles.

For a given energy  $E$  (in GeV), the energy resolution of ECAL is given by [23]

$$\frac{\sigma_E}{E} = \frac{2.8\%}{\sqrt{E}} \oplus \frac{12.8\%}{E} \oplus 0.3\% \quad (4.1)$$

where the terms in the quadrature sum correspond to the stochastic correction which accounts for the energy deposited in the preshower and event-by-event fluctuations, the noise term due to noise in backend electronics, and a constant term which takes into account the non-uniformity of the detector and temperature gradient. From Equation (4.1), it is clear that the stochastic and noise term dominate at lower energies whereas the constant term becomes dominant at higher energies.

The relative energy resolution as a function of  $\eta$  of supercluster ( $3 \times 3$  crystal) for electron and photon is shown in Figure 4.13. The relative resolution is shown in both barrel and endcap region for  $R_9 < 0.94$ ,  $R_9 > 0.94$ , where  $R_9$  is a cluster shape parameter defined as  $\max(E_{3 \times 3})/E_{sc}$ . The  $\max(E_{3 \times 3})$  is the maximum energy deposited in a particular crystal out of 9 crystal, and  $E_{sc}$  is the energy deposited in the supercluster, that is, the sum of energy deposited in all 9 crystals [23]. The relative energy resolution from observed data,

simulated MC, and smeared MC were determined at  $\sqrt{s} = 7$  TeV. The relative resolutions are better in the barrel as compared to the endcap region for  $R_9 < 0.94$  as well as  $R_9 > 0.94$  for electron and photon. From data, the minimum (maximum) resolution is about 1% (5%) for both electron and photon.

## 4.4. Hadron calorimeter

The hadron calorimeter (HCAL) of the CMS experiment is used to measure the energy deposited by hadrons (protons, pions, neutrons, kaons, etc). It is also useful in estimating the missing transverse energy (as described in Section 5.7) which is attributed to the neutrinos and unknown particles such as the dark matter. The HCAL is made of alternating layers of absorbers and plastic scintillators. The absorber is made of brasses and steels whereas the scintillators are made of Bicron-BC408 and Kuraray-SCSN81. When a particle passes through the absorber, it produces hadronic showers (a cascade of secondary particles). The secondary particles of the shower interact with the scintillating material producing photons. The photons are converted into an electronic signal by hybrid photodiodes (HPDs). The HCAL is made up of 36 barrel, 36 endcap wedges (weight of 1 wedge = 26 tonnes), 70000 tiles, and contains over 400 optical decoder units which take light to the HPDs.

Plastic (organic) scintillators are used in HCAL, unlike the inorganic scintillators of the ECAL. The organic scintillators have a wide range of varieties such as a pure organic crystal (anthracene  $C_{14}H_{10}$ ), organic liquids (p-terphenyl  $C_{18}H_{14}$ ), and a plastic scintillator (polyethylene naphthalate). The typical gap between the electronic (vibrational) energy levels is 3-4 eV (0.15 eV). The scintillation mechanism in an organic scintillator is shown in Figure 4.15. At room temperature, all molecules having an energy about 0.025 eV are in the singlet ground state ( $S_{00}$ ). When a particle passes through the scintillator, it transfers part of its kinetic energy to the molecules exciting them to higher states. The de-excitations from  $S_2 \rightarrow S_1$  and  $S_3 \rightarrow S_1$  are radiationless transitions. That is, a photon is not produced in such a transition. However, as a result of these transitions, the population at  $S_1$  increases considerably. The transitions from  $S_1$  to lower energy states result in the emission of photons. There are two types of such transitions. The first is from  $S_1 \rightarrow S_0$  called fluorescence, and the second is  $S_1 \rightarrow T_1 \rightarrow S_0$  called phosphorescence. The light produced from such transitions is converted into electric signals by HPDs. The HPD uses photomultiplier-cum-semiconductor photodiode principle.

The HCAL calorimeters are placed in the barrel and endcap regions as shown in Figure 4.16 along with its various components. The hadron calorimeters are in the barrel (HB), endcap (HE), outer (HO), and forward (HF) region. The HO is placed outside of the magnetic solenoid acting as a tail catcher. The HF is used to

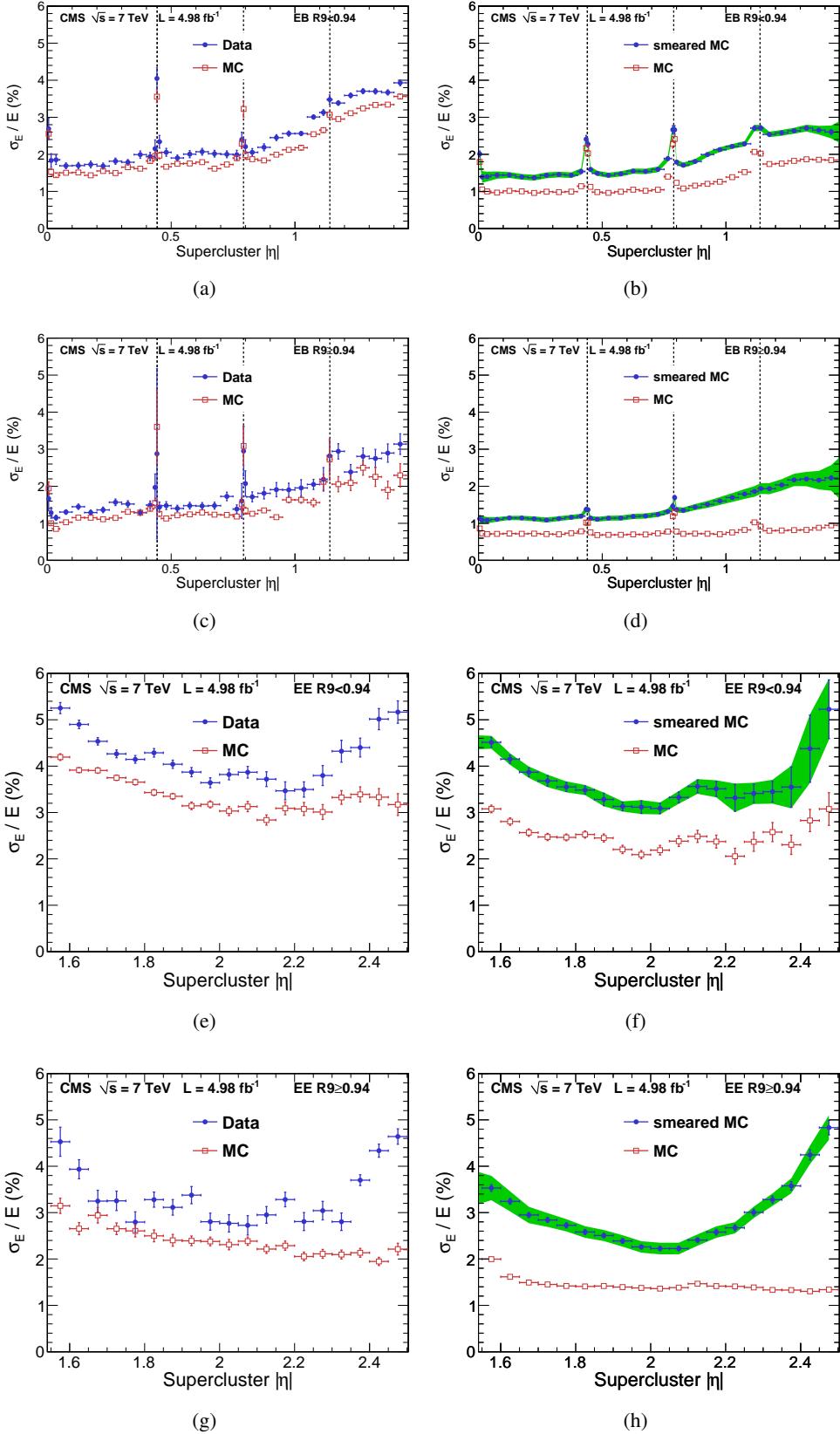


Figure 4.13.: The relative energy resolution of electron and photon as a function of supercluster  $\eta$  in the barrel and endcap region of the ECAL [23].

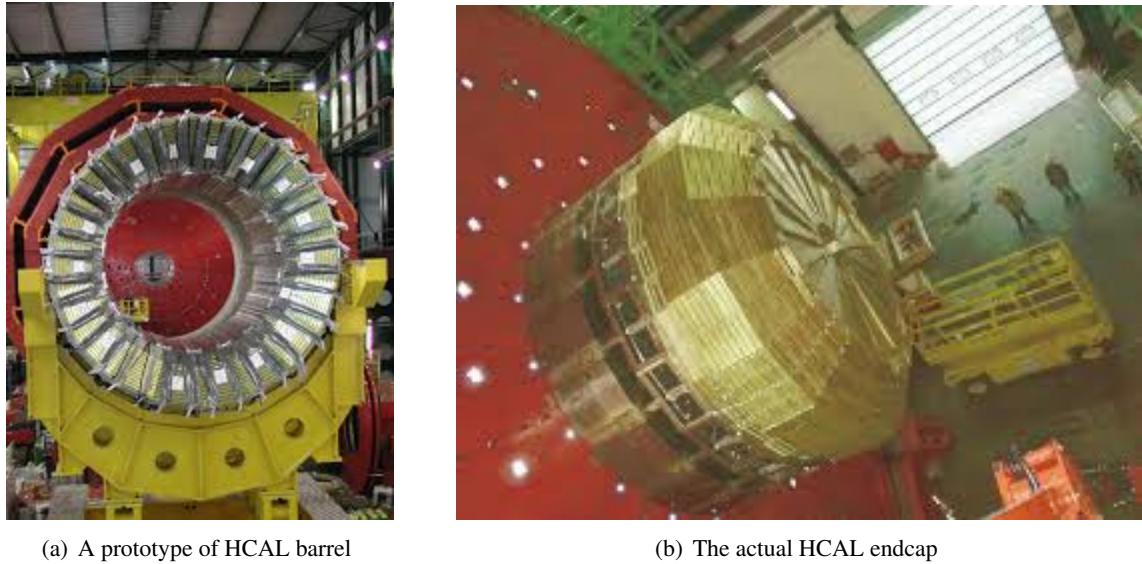


Figure 4.14.: The CMS hadron calorimeter [17]. A prototype of HCAL barrel is shown in silver color in (a). The actual endcap part of the HCAL is shown in (b).

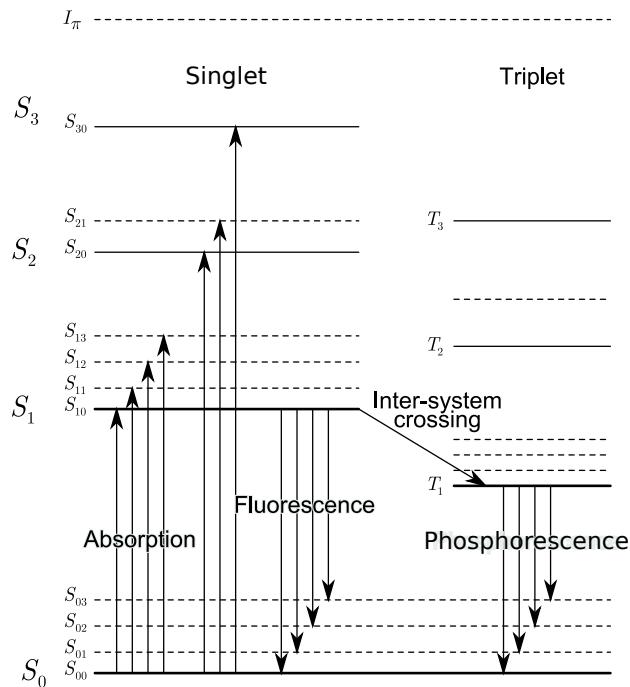


Figure 4.15.: The energy levels of a molecule of an organic scintillator [24]. The ground state is  $S_0$  followed by vibration sublevels ( $S_{01}, S_{02}$  etc.). The  $S_{1,2,3}$  are the excited singlet energy levels. In between the singlet, there are triplet energy levels shown as  $T_{1,2,3}$ . The fluorescence (phosphorescence) occurs when there is a transition from  $S_1(T_1) \rightarrow S_{03,02,01,00}$ .

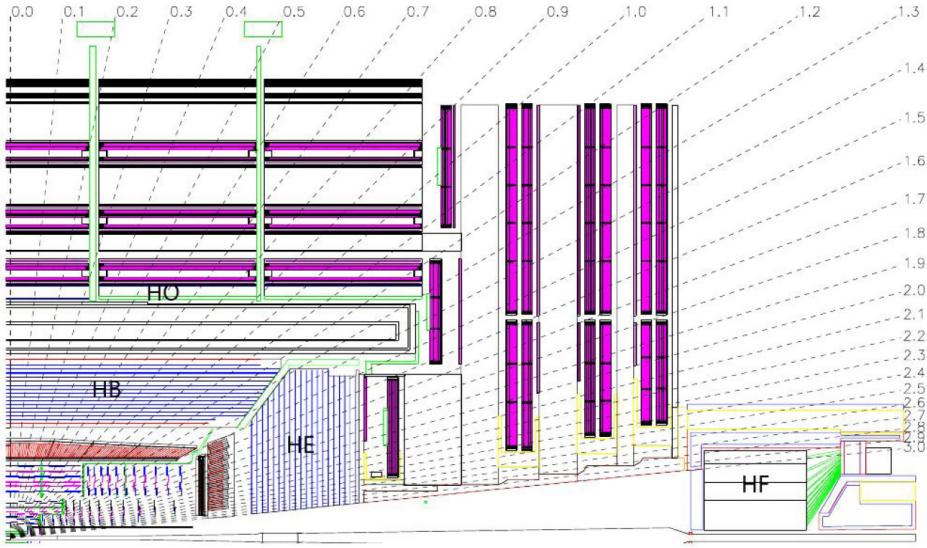


Figure 4.16.: The location of various components (HB, HO, HE, and HF) of the HCAL [4].

study jets in the forward region coming from high radiations. More description of various components of the HCAL is given below:

- **Hadron Barrel:** The HB has a fine and uniform granularity in the  $\eta - \phi$  plane, for example,  $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ . Due to high thickness, around 7-11 times of the interaction length, the HB stops almost all the hadrons passing through it.
- **Hadron Outer:** Even though all the hadrons are supposed to be stopped in the HB, there is a chance for high energetic hadrons to go outside the HB and magnetic solenoid. To measure the energy of hadronic showers from such hadrons, the HO is placed outside the magnet. The HO covers  $|\eta| = 1.26$  range.
- **Hadron Endcap:** The HE covers the  $1.3 \geq |\eta| \geq 3.0$  region, with 14 additional calorimeter towers. Being close to the beam pipe, it is radiation hard. The response of the HE is regularly calibrated during the data taking. The HB and HE overlap in the  $1.305 \geq |\eta| \geq 1.392$  region.
- **Hadron Forward:** The HF is located in the forward region in the range  $1.3 \geq |\eta| \geq 3.0$ , and  $z = \pm 11.2$  m from the collision point. The length of the HF is 1.65 m, divided into 900 towers. The hadronization product of the beam remnants and jets with very high  $\eta$  value are detected by the HF.

The energy resolution of the HCAL is given by

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{s}{\sqrt{E}}\right)^2 + c^2, \quad (4.2)$$

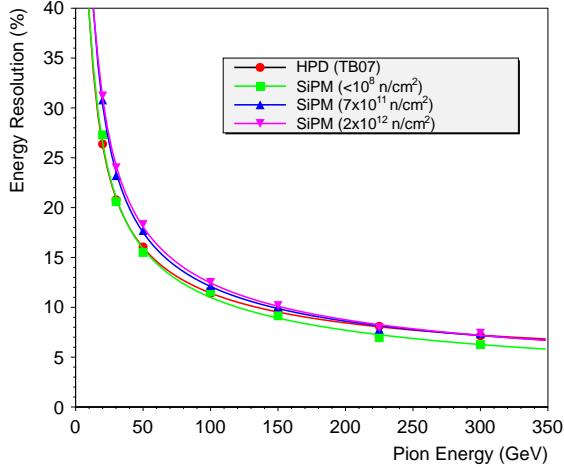


Figure 4.17.: The HCAL energy resolution as a function of pion energy measured using hybrid photodiodes, and silicon photomultipliers with different pixel density [25].

where the energy of hadron ( $E$ ) is measured in GeV, the constant  $s = (0.847) \text{ GeV}^{1/2}$  and  $c = 0.074$  for all part of the HCAL except the HF. For the HF,  $s = 1.98 \text{ GeV}^{1/2}$  and  $c = 0.09$ . The pion energy resolution from the HCAL is shown in Figure 4.17. For low  $E$ , the resolution is poor, while for  $E > 100 \text{ GeV}$ , it lies in 5-10% range.

## 4.5. Magnet

The magnet bends charged particles moving inside and outside of it. Lower the momentum of a particle, the more it is bent and vice versa. The weight of the magnet is 12,000 tonnes, being the largest superconducting magnet ever built. There are three main parts of the magnet: a solenoid, a vacuum tank, and the return iron yoke. A brief description of each component is given below:

- **Solenoid:** The solenoid is made of NbTi (Niobium Titanium) coils which is an aluminum-stabilized and mechanically reinforced superconducting material. The solenoid is formed by placing 5 modules side-by-side along the  $z$ -axis. With the length of 2.5 m, each module has 4 internal layers of winding, each with 109 turns. Therefore, the total length of the solenoid is  $5 \times 2.5 \text{ m} = 12.5 \text{ m}$ , with  $5 \times 4 \times 109 = 2180$  turns. The diameter of the solenoid is 6 m. The magnetic field inside the solenoid is given by

$$B = k\mu_0 \times N \times I/L \quad (4.3)$$

Where  $k$  is the relative permeability. Therefore, with  $k = 0.90$ ,  $\mu_0 = 4\pi \times 10^{-7} \text{ mT/A}$ ,  $N = 2180$ ,  $I =$

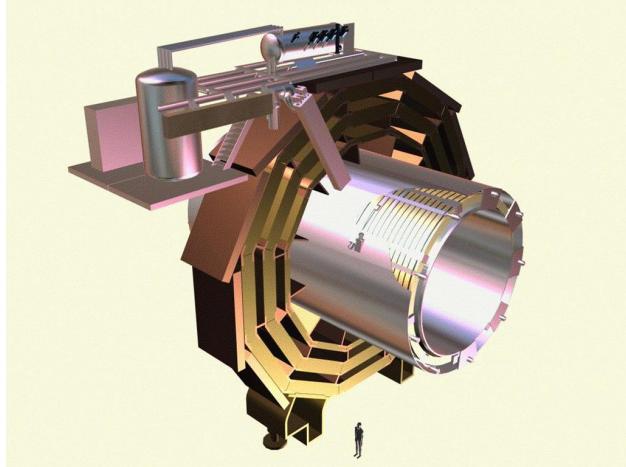


Figure 4.18.: An artistic view of the CMS solenoid [26]. An aluminum alloy shell, with a thickness of 50 mm, is placed on the inner and outer surface of the solenoid and used as a cooling wall. Outside the solenoid are the return iron yokes which hold the magnetic flux.

19.14 kA, and  $L = 12.5$  m, the magnetic field inside the solenoid,  $B = 3.8$  T, which is 100,000 times stronger than the Earth's magnetic field. Being a very high energy experiment, such a large magnetic field is necessary for better tracking resolution.

- **Vacuum tank:** A huge amount of heat is produced due to the flow of such a large current. The energy contained inside the solenoid is 2.7 GJ which is more than enough to melt 18 tonnes of gold. A dedicated cooling system is placed to cool the solenoid down to  $-268^{\circ}\text{C}$ .
- **Return yoke:** With four layers in the barrel region, and six disks in the endcap region, a giant iron return yoke is placed outside the solenoid to contain the magnetic flux. The iron yokes in the barrel region are shown in Figure 4.18. Total weight of the iron yoke is 10000 tonnes.

The longitudinal and transverse component of the magnetic field produced from the solenoid is shown in Figure 4.19. The longitudinal component is the strongest compared to the transverse component. The transverse one is everywhere zero except at the endcap in the region  $6 \text{ m} < z < 10 \text{ m}$ . The magnetic field of the longitudinal component is 3.8 T inside the solenoid, and about 2 T inside the return iron yoke. Further, the longitudinal component inside and outside the solenoid is in the opposite direction. By virtue of this property, the charged particles are bent in the opposite direction.

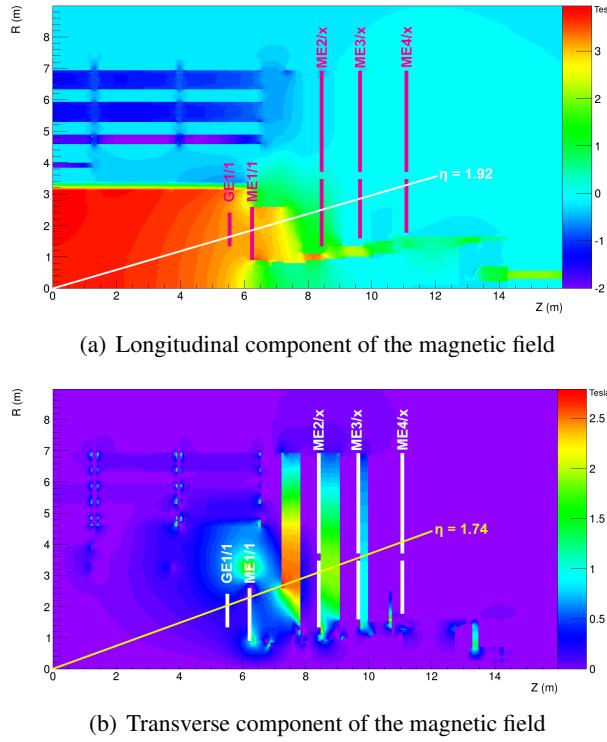


Figure 4.19.: Component of the magnetic field inside CMS [16].

## 4.6. Muon chambers

Detection of muons was the primary physics motive for the CMS experiment. Being a minimum ionizing particle, the muon traverses most part of CMS without losing much energy in the calorimeters. Except for the muons, other particles produced in the collision are stopped either in ECAL or HCAL. In view of this, muon chambers are placed at the outermost side of CMS where only muons are supposed to be detected. The muon chambers have a different configuration in the barrel and endcap regions. When a muon passes through the muon chambers, filled with gases, atomic electrons are knocked out. Subsequently, the current corresponding to these electrons is amplified and stored. As shown in Figure 4.20, the CMS experiment has the following three types of the muon chambers: drift tubes (DTs), cathode strip chambers (CSCs), and resistive plate chambers (RPCs). The momentum resolution of the muon detector is 10% (20%) in the barrel (endcap) region for a muon with  $p_T = 40$  GeV. Various properties and parameters of the muon subsystem are listed in Table 4.1. A brief description of these subsystems is given below.

- **Drift tube:** The DTs are placed in the barrel region. A schematic diagram of a DT is shown in Figure 4.21, where the red dots are the aluminum wires pointing into the page. The region between the wires are filled 85% with Argon (Ar) and 15% with CO<sub>2</sub> gases. When a muon passes through the DTs, atomic electrons

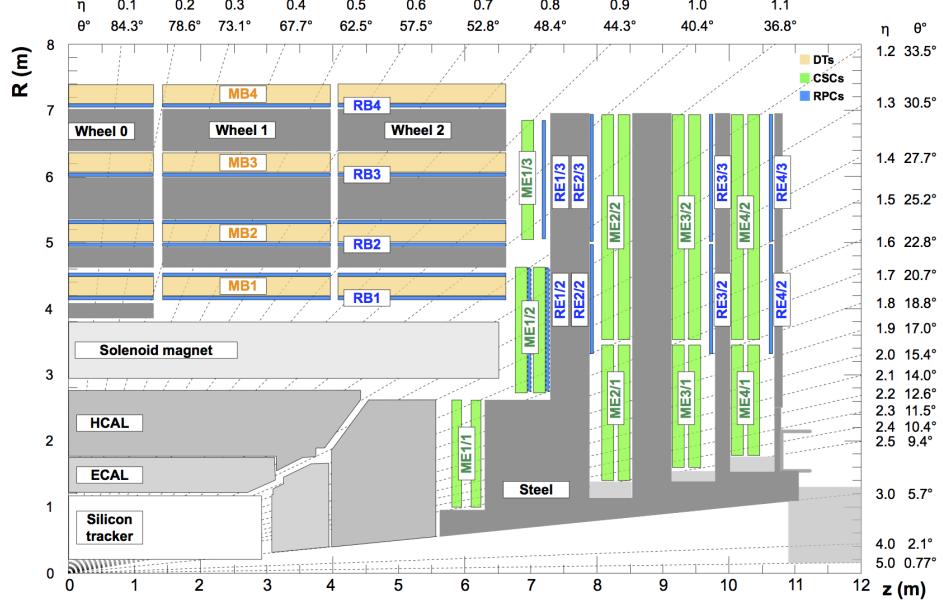


Figure 4.20.: Location of the various sub-detectors of the CMS [27]. The DTs and CSCs are placed in barrel and endcap regions. The RPCs are placed in both regions. The muon chamber in the barrel region is formed by placing three “wheels” side-by-side along the  $z$ -axis whereas in the endcap regions different segments of RPC and CSC are placed along the radial direction.

Table 4.1.: Various properties and parameters of the CMS muon subsystems [27].

Muon subsystem	DT	CSC	RPC
$ \eta $ coverage	0.0–1.2	0.9–2.4	0.0–1.9
Number of stations	4	4	4
Number of chambers	250	540	Barrel: 480 Endcap: 576
Number of layers per chamber	$R\text{-}\phi$ : 8; $z$ : 4	6	2 in RB1 and RB2 1 elsewhere
Number of readout channels	172000	Strips: 266112 Anode channels: 210 816	Barrel: 68136 Endcap: 55296
Percentage of active channels	98.4%	99.0%	98.3%

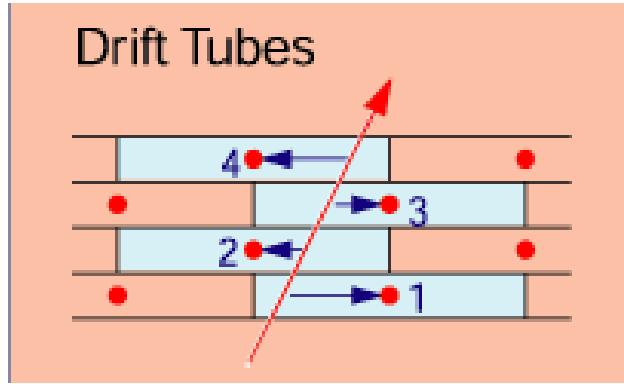


Figure 4.21.: Schematic view of one of the DTs of the muon chamber [17]. The red dots are the aluminum wires pointing into the page, the red arrow is the direction of an incoming muon, and the horizontal blue lines are the distance traveled by secondary electrons.

are knocked-out of the gases. The anode wires attract electrons towards themselves. The distance traveled by electrons up to the anode wire is shown by horizontal arrows. The information about the length of such arrows gives direction in which the muon could have traversed inside the DTs, which give the  $z$  and  $\phi$  coordinates of the muon.

- **Cathode strip chamber:** A schematic diagram of the cathode strip chamber is shown in Figure 4.22. The CSC is placed in the endcap region where particle multiplicity is very high. It consists of anode wires of gold-plated tungsten and cathode strips made of copper. The wires and strips are placed perpendicular to each other. A mixture of different gases consisting of 50% CO<sub>2</sub>, 40% Ar, and 10% CF<sub>4</sub> is filled between anode and cathode. When muon passes through the CSC, atomic electrons are knocked out of the gasses and attracted towards the anode wires. Whereas the ions are attracted towards the cathode strips. Therefore, two coordinates is determined from the CSC, unlike the DT where only one coordinate is determined.
- **Resistive plate chamber:** The response from the RPC is very fast, hence, it helps in quick triggering (see Section 4.7). The time resolution of the RPC is 3 ns. A schematic diagram of one of the RPCs is shown in Figure 4.23. There are two resistive plates: one acts as the anode and the other as the cathode. Only the anode plate is transparent to the secondary electrons. The region between the plates is filled with a mixture of gases (95.2% C<sub>2</sub>H<sub>2</sub>F<sub>4</sub>, 4.5% i-C<sub>4</sub>H<sub>10</sub>, and 0.3% SF<sub>6</sub>). When a muon passes through the gases, atomic electrons are knocked out. These electrons are picked up by anode strips placed above the anode plate. The momentum of a muon is measured from the hit pattern in the strips. The muon information is conveyed to the global muon trigger (see Section 4.7).

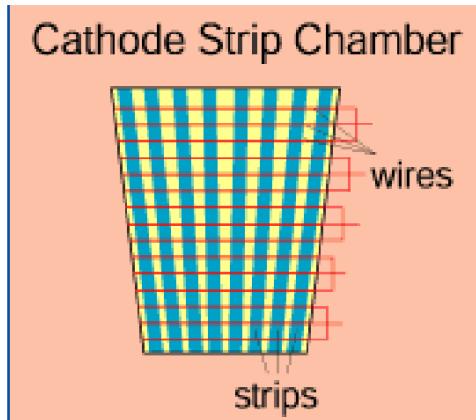


Figure 4.22.: Schematic view of one of the cathode strip chambers [17]. The wires shown in red and cathode strips shown in cyan are placed perpendicular to each other.

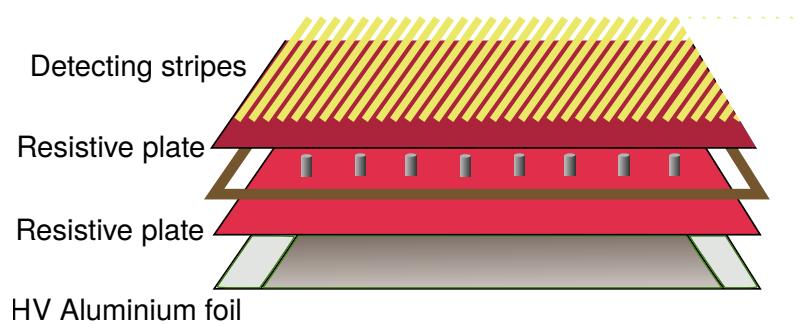


Figure 4.23.: Schematic diagram of one of the RPCs. The cathode (bottom red plate) and anode (top red plate) is separated by a separator (shown in gray). The region between the two plates is filled with gas. The anode is transparent to the secondary electrons. The orange strips (above anode) picks the secondary electrons. The high voltage aluminum foil is placed at the bottom. This figure is adopted from [17].

Table 4.2.: Event rate after applying level-1 and high-level triggers.

Input event rate	Trigger	Type	Output event rate
40 MHz	L1	Hardware	100 KHz
100 KHz	HLT	Software	200 Hz

## 4.7. The trigger

At the LHC, the proton beams are made to collide after 25 ns. Due to such a high collision rate a huge number of events, roughly 40 million, are produced at every second. Storing all these events would require many petta bytes (1 PB = 1000000 GB) of computer memory, which is practically not possible as the collision happens over a large period of time (many months). Even if somehow one can get around this challenge, most of the events are produced from soft collision and contain no interesting physics process. To filter events which are of physics interest, a two-tier trigger system is deployed at the CMS experiment. The first is called Level-1 (L1) trigger which is based on hardware (by collecting information from various sub-detectors) and the second is called High Level Trigger (HLT) which is based on software algorithms. The event rate after passing through these triggers are shown in Table 4.2. After applying the HLT, the event rate is reduced to 200 Hz which requires about 200 MB of computer memory. A brief description of the L1 and HLT triggers are given in this thesis. A detailed description of these can be found in Reference [28].

- **Level-1 Trigger:** L1 is also called online trigger as events are triggered during the data taking. The L1 trigger is required to be very fast with a latency of  $4\ \mu\text{s}$ . The decision to either keep or reject events by this trigger is based on a set of information provided by various sub-detectors as shown in Figure 4.24. Over 8000 towers of the calorimeters (ECAL and HCAL) send transverse energies and quality flags to the regional calorimeter trigger (RCT). These pieces of information are processed in parallel by the RCT. The  $e/\gamma$  candidates along with  $E_T$  sums based on  $4 \times 4$  towers are sent by RCT to the global calorimeter trigger (GCT). The GCT finds central-jets, forward-jets, tau-jets, and missing transverse energy. The decision based on the findings of GCT is conveyed to the global trigger (GT) whether to keep an event or not. The other decision conveyed to the GT comes from the global muon trigger (GMT). The GMT takes input from different muon chambers. The hits from RPC is sent to the pattern comparator (PC) which finds muon candidates. The information about the number of muon candidates from the PC is shared with the GMT. The front-end trigger electronics identify segments from the hits in CSC and DT. These segments are sent to the regional track finder (TF). In the TF, the PC comparator identifies muon candidates and measures momentum from the bending of tracks due to the magnetic field. The muon candidates with the

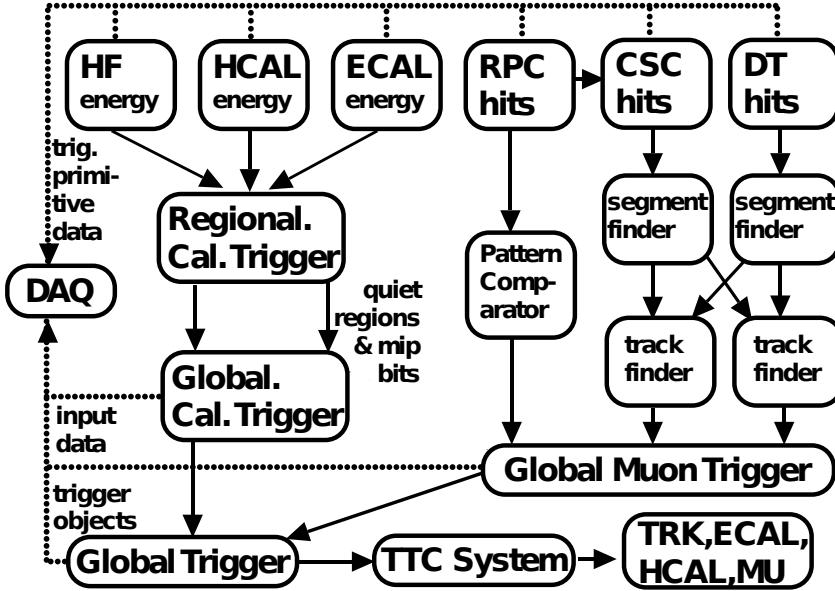


Figure 4.24.: Workflow of decisions for the Level-1 trigger [28].

corresponding momentum from the TF are shared with GMT. The GMT decides whether an event is worth keeping from muon's point of view. Finally, the collective decision from the GCT and GMT is passed to the global trigger which takes the final call. The decision of the GT is conveyed to the data acquisition system (DAQ) which stores data from sub-detectors such as tracker (TRK), ECAL, HCAL, and muon chambers. If the event is going to be rejected by the GT then the DAQ does not store corresponding data from the sub-detectors.

- **High Level Trigger:** After applying L1 trigger, the filtered events are passed as input to the HLT. The L1 bits available in the event serve as the seed for HLT. The decision to reject or keep an event from the HLT is taken based on whether an event has a reasonable number of physics objects such as electrons, muons, and jets with transverse momentum above some threshold. In order to reconstruct these physics objects, a set of dedicated software algorithms is run for every event. During the run, the input from one step is passed to the subsequent steps. For every event, the HLT decision is stored in the form of programming strings also called trigger paths.

The HLTs are applied to select events enriched with desired physics objects. Depending on the topology of analysis, an event is selected if the corresponding trigger path is available in that event. In this thesis, we have 1 lepton, at least 4 jets, and neutrino in the final state. Therefore, an event is selected if it has

a trigger path of `HLT_IsoMu24*` and `HLT_IsoTkMu24*` with logical OR for the muon channel, and `HLT_Ele27_WPTight_Gsf` for the electron channel. Having these muon trigger paths in the event ensures that there is at least one isolated muon with  $p_T$  more than 24 GeV (see Section 5.4) whereas the electron trigger path implies the presence of at least one electron with  $p_T$  more than 27 GeV with tight working point (WP) reconstructed with Gaussian sum filter algorithm (see Section 5.5). The efficiency of these triggers depends on  $p_T$  and  $\eta$  of the lepton. For lower  $p_T$ , the efficiency is very poor and has a sudden increase at a particular value. That sudden increase is called trigger turn-on and the leptons are required to have a  $p_T$  greater than that at the turn-on.

The muon and electron trigger efficiency as a function of  $p_T$  is shown in Figure 4.25. The trigger turn-on for muon trigger is at about  $p_T = 24$  GeV, hence in the subsequent selection (Section 5.4), the muon is required to have  $p_T > 26$  GeV. For the electron trigger (`HLT_Ele27_WPTight_Gsf`), the trigger turn-on is at around  $p_T = 33$  GeV, therefore a cut of  $p_T > 35$  GeV is applied on electrons in the subsequent selection (Section 5.5).

## 4.8. Data distribution system

The collision and simulated data are stored at different tiers located at different places in the world. Being a very large collaboration, where thousands of people from across the world work on the same data, the CMS distributes its data in such a way that any user can access these through the worldwide LHC grid. The data distribution system is divided into various tiers such as tier-0 (T0), tier-1 (T1), and tier-2 (T2) depending on the type of operations performed on the data. These are located at different places, as shown in Figure 4.26. A brief description of all the data tiers is given below.

- **Tier-0:** The collision happening at IR5 is recorded by the CMS experiment in RAW (response in the form of electrical signals from sub-detectors plus triggers plus high-level objects reconstructed while running HLT) format. These RAW data are repacked in roughly 10 different datasets based on trigger information. The repacked data is archived at T0, one copy is saved at T0, and other copy is transferred to every T1 site. Besides the storage and transfer of the data, few operations are also performed at T0. The automated prompt calibration is performed during the data taking to get calibration constants needed for reconstruction. With the calibration constants, a *prompt* reconstruction is also performed on the RAW data. After the prompt reconstruction, AOD (analysis object data) is sent to all T1 centers. All the operations performed at T0 are automated.

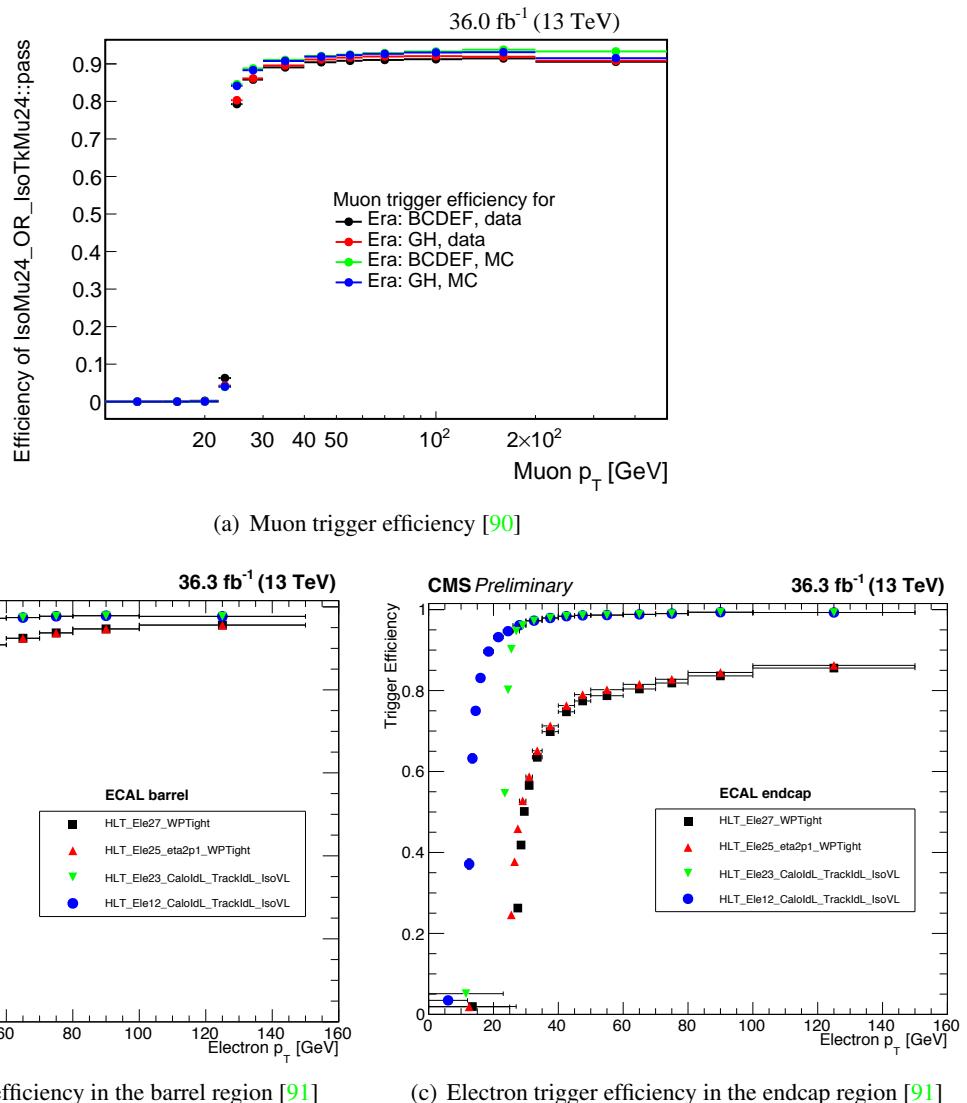


Figure 4.25.: Muon and electron trigger efficiency as a function of  $p_T$ . The trigger turn-on is at about  $p_T = 24 \text{ GeV}$  for muon, and  $33 \text{ GeV}$  for electron trigger.

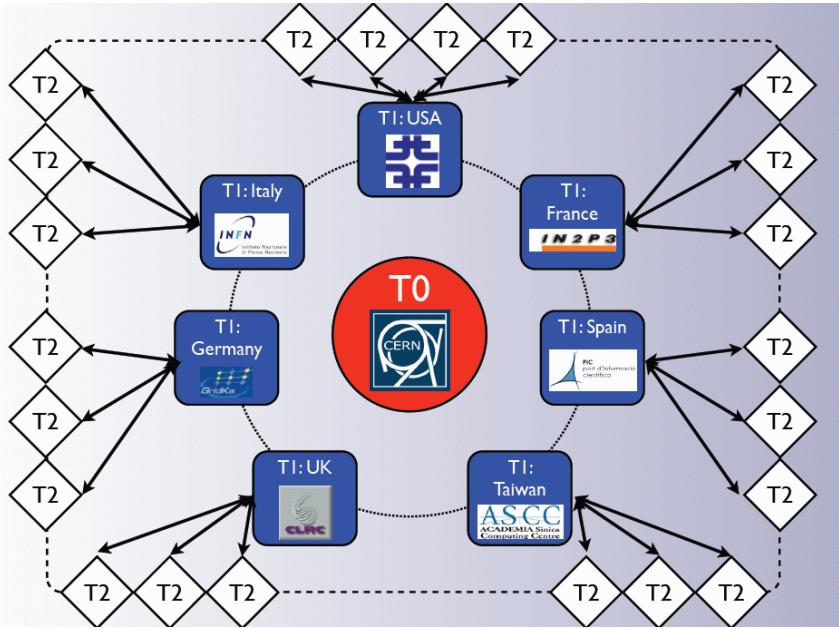


Figure 4.26.: Location of different data acquisition systems [29]. The T0 is located at CERN, T1s are located in 7 different places, and 55 T2s (although only 22 are shown in this figure) are located at various countries. There is only one T2 in India which is located at TIFR, Mumbai.

- **Tier-1:** One copy of RAW data received from T0 is archived at each T1 site. User-based and long operations are performed at T1 sites. The full reconstruction chain (RAW-RECO-AOD-MiniAOD) is performed at T1. Each copy of AOD/MiniAOD is saved at every T1 site and transferred at the associated T2 sites.
- **Tier-2:** The simulated data are produced at T2 sites. The whole MC production chain (GEN-SIM-Raw-RECO-AOD-MiniAOD) is performed at T2. The simulated samples are also transferred to each T1 site. The T2 sites are located at different universities and most of the resources are used by the corresponding members of that university.

There is also tier-3 sites located at different universities. However, all the user-based operations are performed only by the local members on these T3s. The flow of collision data between the different tiers are shown in Figure 4.27.

## 4.9. Event data format and the CMS software

The particles produced during collision interact with various sub-detectors of the CMS experiment. The response of each sub-detector is stored in the form of electronic signals. The physics objects from these

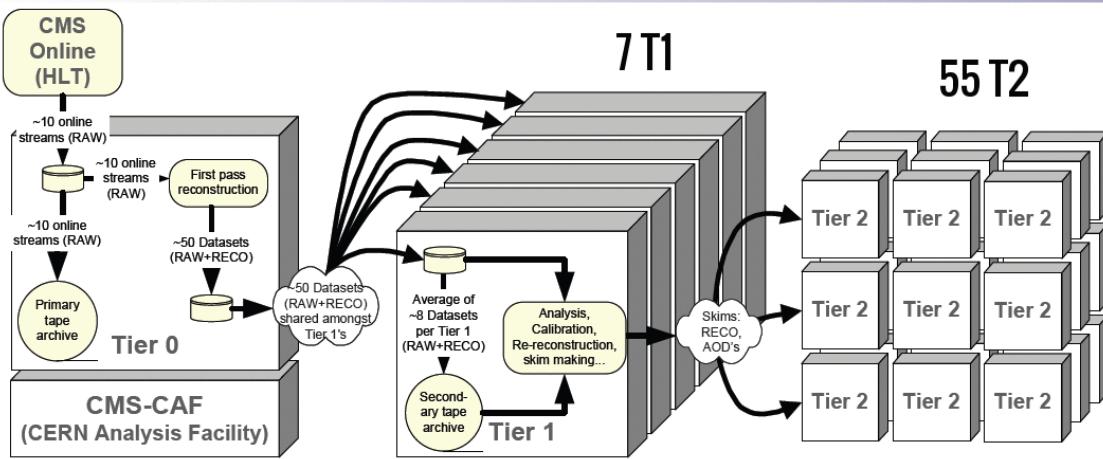


Figure 4.27.: The flow of collision data among different tiers. [29]

electronic signals are reconstructed following a chain of processes comprising many steps. The output data at every step are stored in a different format and serve as input for the subsequent steps. The detectors response is stored in the form of RAW data, which are further digitized (electronic response is converted into the energy deposits) and the output is stored in the form of DIGI data. The DIGI step itself involves various substeps such as  $\text{RAW} \rightarrow \text{DIGI} \rightarrow \text{L1} \rightarrow \text{DIGI2RAW} \rightarrow \text{HLT} \rightarrow \text{RAW2DIGI}$ . The reconstruction algorithms are run on the DIGI data to reconstruct physics objects (see Chapter 5). The data for the physics objects are stored in the form of RECO. The RECO data are further processed to get the analysis-specific objects and the output is stored in AOD (analysis object data) format. To reduce the size of AOD dataset as well as to retain only the most relevant physics objects, the AOD dataset is slimmed further and the output is stored in the form of a mini-AOD dataset. In this thesis, we have used the mini-AOD datasets. All the datasets are written on the disk with `.root` extension. The ROOT framework is used to browse the content of the root file.

The CMS software (CMSSW) is one of the most advanced packages which provides thousands of sub-packages needed to analyze data at different steps. The CMSSW is mostly written in C++ and python. The core design of the CMSSW relies on the event data model (EDM). All the sub-packages (modules) of the CMSSW run on each event separately, that is none of them run on two events at the same time. Different types of modules are available in the CMSSW to perform different types of operation. A schematic diagram of the event data model is shown in Figure 4.28. A brief description of these modules is given below.

- **Source:** The source opens the root files and provides each event to the subsequent modules. For the simulation where events are generated from scratch, an empty source is used.

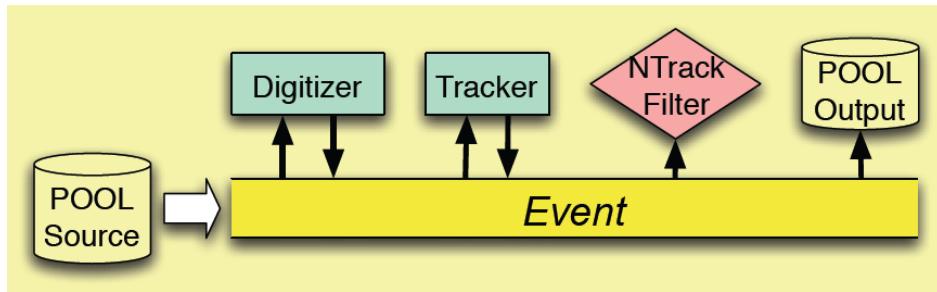


Figure 4.28.: A schematic diagram of the event data model [30]. The *pool source* is a package that provides the input data for every event. The *digitizer* and *tracker* are the EDProducer which takes input from the event, do some operation and write the output to the same event. The *nTrack filter* is the EDFilter which takes input from the event and decides if that event is to be kept or rejected. If the event is to be rejected then the subsequent processing is halted and the CMSSW start looking at the next event through *pool source*. The *pool output* package stores the output after all the operations are performed.

- **EDProducer:** These modules take input from the input root file, perform some operations such as modifying physics objects, and put back the modified objects to the event. These modified objects are further used by subsequent modules.
- **EDFilter:** All the events of the input root file may not be needed for every analysis. In view of this, a set of filter modules are available in the CMSSW. These filters read the event and take the decision about whether to keep that event or not. If the event is to be rejected then the operation of the subsequent modules is halted. And the loop goes to the next event.
- **EDAnalyzer:** The EDAnalyzers are the most frequently used modules for the physics analysis. The mini-AOD datasets contain all the relevant physics objects and the same dataset is used in every physics analysis. However, different physics analysis has a different final state, therefore, the full content of mini-AOD dataset is not needed for every analysis. The EDAnalyzers are used to slim the miniAOD dataset and keep only the analysis-specific physics objects. The EDAnalyzer reads the event, selects desired objects and relevant kinematic properties and writes out the necessary information into private data format known as ntuple.
- **OutputModule:** The output module stores the output event after all the other modules are executed.

## **Part III.**

# **Search for Charged Higgs Boson**



# 5 | Reconstruction and Identification

The physics objects of our interest are primary and secondary vertices, lepton, jets, and missing transverse energy attributed to neutrino. The particle flow (PF) algorithm [92, 93] is used to reconstruct these objects. In this chapter, we first list the collision data and simulated Monte Carlo (MC) samples used for this analysis and then describe the reconstruction process for the primary and secondary vertices, muon, electron, jets, and b tagging.

## 5.1. Data samples

The reconstructed data is broadly divided into different event topologies. Only those datasets matching with event topology under consideration are used for this analysis. Such approach help in optimising the computing resources. The single muon and single electron data sample used for this analysis are shown in Table 5.1. The data was collected in 2016. A luminosity mask is applied to select good quality data, where most of the sub-detectors of the CMS operates in a normal condition. The integrated luminosity after applying the mask reduces marginally from  $37.0\text{ fb}^{-1}$  to  $35.9\text{ fb}^{-1}$ .

## 5.2. Simulated samples

The centrally produced official simulated samples are used in this analysis. All the simulated samples are shown in Table 5.2 with corresponding cross section and number of events. The k-factor is the ratio of cross

Table 5.1.: Muon and electron data samples recorded in 2016 with an integrated luminosity  $35.9\text{ fb}^{-1}$ .

Dataset	$L_{\text{int}}(\text{fb}^{-1})$	Events
SingleMuon	35.86	786767595
SingleElectron	35.86	937954379

sections from next-to-leading order and leading order,  $k_f = \sigma_{\text{NLO}}/\sigma_{\text{LO}}$ .

The signal and background simulated samples are generated using **MADGRAPH5\_aMC@NLO** [94, 95] as well as **POWHEG** [96–98] generators at parton level. These parton level events are hadronised using **PYTHIA** [99, 100]. The hadronised events were tuned using **CUETP8M1** [101] which are then passed through **GEANT4** [102] for detector simulation. Finally, the events are reconstructed after complete detector simulation.

The  $t\bar{t} + \text{jets}$  channel is the most significant irreducible background for this analysis. It contributes around 94% of the total backgrounds in the signal region. The parton level events for  $t\bar{t} + \text{jets}$  samples of Table 5.2 are generated using **POWHEG** [96–98] at next-to-leading (NLO) order. The **NNPDF30\_nlo\_as\_0118** [103] parton distribution function (PDF) was used for this purpose. The partonic events were then hadronised using **PYTHIA** [99, 100]. To simulate this channel at higher orders and to take care of non-perturbative effects,  $t\bar{t} + \text{jets}$  samples were tuned with **CUETP8M1** [101]. The NNLO cross section for  $t\bar{t} + \text{jets}$  process is estimated to be  $831.76 \pm^{20}_{29}$  (scale)  $\pm 35(\text{PDF} + \alpha_s)$  pb [104].

Single t quark samples are shown in Table 5.2, where a t quark is produced with jets in t-channel (**ST\_t**), s-channel (**ST\_s**), and t W-channel (**ST\_tw**). The **ST\_tw** and **ST\_t** samples are produced using **POWHEG + PYTHIA** and **CUETP8M1**, where **ST\_tw** has 5-flavor scheme whereas **ST\_t**, has 4-flavor scheme and further, the t quark decays inclusively. The **ST\_s** are generated using **MADGRAPH5\_aMC@NLO** [95] in the 4-flavor scheme and was hadronised using **PYTHIA**. The NLO cross section for single t process [105, 106] are shown in Table 5.2.

The inclusive  $W + \text{jets}$  and  $Z/\gamma + \text{jets}$  samples were generated using **MADGRAPH5\_aMC@NLO** and are hadronised using **PYTHIA**. The MLM [107] technique was used to take care of double counting of partons. There are exclusive  $W + n$  jets and  $Z/\gamma + n$  jets samples available for  $n = 1, 2, 3, 4$ . The inclusive ( $W + \text{jets}$  and  $Z/\gamma + \text{jets}$ ) and exclusive ( $W/Z/\gamma + n$  jets) samples have our event topology i.e. 4 jets + 1 lepton +  $E_T^{\text{miss}}$ . Therefore, the inclusive and exclusive samples are added appropriately to have more statistics. To avoid double counting, the exclusive samples are normalized w.r.t the inclusive one. For example, the  $W + 1$  jet events are multiplied by  $1/(L_{W+\text{jets}} + L_{W+1\text{jet}})$  and then are added linearly with that of  $W + \text{jets}$ . Where  $L_{W+\text{jets}}$  is the luminosity (number of events divided by the cross section) of  $W + \text{jets}$  sample. A similar procedure is followed to add  $Z/\gamma + n$  jets with  $Z/\gamma + \text{jets}$ . The NLO cross section for these processes is shown in Table 5.2.

The vector boson pair production process (WW, WZ, and ZZ) samples are generated and hadronised using **PYTHIA** and are tuned using **CUETP8M1**. The NLO cross section for these samples are shown in Table 5.2.

The muon (electron) enriched QCD multijet samples are used for  $\mu + \text{jets}$  ( $e + \text{jets}$ ) channel. These samples are generated using PYTHIA and CUETP8M1. After multiplying with filter efficiency, the NLO cross sections are shown in Table 5.2. From this table, it can be seen that the luminosity for these samples are significantly smaller as compared to the observed luminosity and therefore, a data-driven approach is used to make more precise estimation of QCD multijet background.

The charged Higgs signal samples were generated using MADGRAPH5\_aMC@NLO, and hadronized using PYTHIA. The signal sample for several mass points in the range of 80 to 160 GeV (80, 90, 100, 120, 140, 150, 155, 160) are generated for the search for charged Higgs. The cross section for the signal is  $831.76 \times 0.2132 \text{ pb}$ , where 831.76 pb is the inclusive (fully-hadronic, fully-leptonic, semi-leptonic)  $t\bar{t} + \text{jets}$  production cross section and factor 0.2132 is the branching fraction of  $W^- \rightarrow l^-\bar{\nu}_l$  (where  $l = \mu, e, \tau$  is not considered in this analysis) [108]. The factor 0.2132 is multiplied because  $t\bar{t}$  decays semi-leptonically ( $(t \rightarrow H^+ b, H^+ \rightarrow c\bar{s}), (\bar{t} \rightarrow W^- \bar{b}, W^- \rightarrow l^-\bar{\nu}_l)$ ), where  $l = \mu, e$  in the charged Higgs signal samples. Furthermore, the signal events are scaled by the maximum observed upper limit obtained at 8 TeV. The upper observed limit on  $\mathcal{B}(t \rightarrow H^+ b)$  is 6.5% for the  $m_{H^+} = 90$  GeV [63]. Therefore, the signal samples are scaled by a factor  $2 \times 0.065 \times (1 - 0.065) = 0.12155$  in every plot and table except in the data cards used for limit computation.

In the  $t\bar{t} + \text{jets}$  samples, the NLO matrix element parton shower matching is varied by the damping parameter ( $h_{\text{damp}}$ ). Additional  $t\bar{t} + \text{jets}$  samples are generated by varying  $h_{\text{damp}}$  up and down and are used to observe the effect of  $h_{\text{damp}}$ . Similarly,  $t\bar{t} + \text{jets}$  samples where the renormalization and factorisation scales have been varied up and down are used to evaluate the uncertainties due to these scales. Also, alternate  $t\bar{t} + \text{jets}$  samples with  $m_t = 171.5$  and 173.5 GeV are considered to observe the effect of mass of the t quark. All the additional  $t\bar{t} + \text{jets}$  samples used for systematics study are shown in Table 5.3.

## 5.3. Primary vertex and diffuse offset energy density

### 5.3.1. Primary vertex

To reconstruct a primary vertex (PV) following three steps are performed [21]:

- **Track selection:** Primary vertices are associated with a large number of tracks emerging from the same point. Tracks satisfying following criteria are used to reconstruct the collision point (vertex).
  - distance of closest approach from the center of beamspot  $< 2.4 \text{ cm}$ ,

Table 5.2.: Background and signal simulated samples with corresponding cross section and the number of events.

<b>Process</b>	$\sigma$ (pb) (Order)	$k_f$	<b>Events</b>
t̄ + jets	$831.76^{+20}_{-29} \pm 35$ (NNLO)	-	77081156
Single t (tW-channel)	$71.7 \pm 1.80 \pm 3.40$ (NLO)	-	6933094
Single t (t-channel)	$80.95^{+5.8}_{-5.2} \pm 0.16$ (NLO)	-	38811017
Single t (s-channel)	$10.32^{+0.6}_{-0.5} \pm 0.01$ (NLO)	-	2989199
W + jets	$50690 \pm 389.1$ (LO)	1.21	29181900
W + 1 jet	$9493 \pm 25.52$ (LO)	1.21	44813600
W + 2 jets	$3120 \pm 78.5$ (LO)	1.21	29878415
W + 3 jets	$942.3 \pm 36.8$ (LO)	1.21	19798117
W + 4 jets	$524.2 \pm 23.6$ (LO)	1.21	9170576
Drell-Yan + jets	$4895 \pm 41$ (LO)	1.17	48103700
Drell-Yan + 1 jet	$1016 \pm 16.8$ (LO)	1.17	62079400
Drell-Yan + 2 jets	$331.3 \pm 8.5$ (LO)	1.17	19970551
Drell-Yan + 3 jets	$96.6 \pm 3.9$ (LO)	1.17	5856110
Drell-Yan + 4 jets	$51.4 \pm 2.5$ (LO)	1.17	4197868
WW	$118.7^{+2.5}_{-2.2}$ (NNLO)	-	994012
WZ	$46.74^{+1.9}_{-1.5}$ (NLO)	-	1000000
ZZ	$17.72^{+0.6}_{-0.4}$ (NLO)	-	990064
QCD_Pt-15to20_MuEnriched	3819570 (NLO)	-	4141251
QCD_Pt-20to30_MuEnriched	2960198 (NLO)	-	31475157
QCD_Pt-30to50_MuEnriched	1652471 (NLO)	-	29954815
QCD_Pt-50to80_MuEnriched	437504 (NLO)	-	19806915
QCD_Pt-80to120_MuEnriched	106033 (NLO)	-	13786971
QCD_Pt-120to170_MuEnriched	25190 (NLO)	-	8042721
QCD_Pt-170to300_MuEnriched	8654 (NLO)	-	7947159
QCD_Pt-300to470_MuEnriched	797 (NLO)	-	7937590
QCD_Pt-15to20_EMEEnriched	254600 (NLO)	-	5652601
QCD_Pt-20to30_EMEEnriched	5352960 (NLO)	-	9218954
QCD_Pt-30to50_EMEEnriched	9928000 (NLO)	-	4730195
QCD_Pt-50to80_EMEEnriched	2890800 (NLO)	-	22337070
QCD_Pt-80to120_EMEEnriched	350000 (NLO)	-	35841783
QCD_Pt-120to170_EMEEnriched	62964 (NLO)	-	35817281
QCD_Pt-170to300_EMEEnriched	18810 (NLO)	-	11540163
QCD_Pt-300toInf_EMEEnriched	1350 (NLO)	-	7373633
ChargedHiggsToCS_M080_To160	$0.2132 \times 831.76$ (NNLO)	-	1000000

Table 5.3.: The SM  $t\bar{t}$  + jets samples used for systematics study.

Process	$\sigma$ (pb)	Events
$t\bar{t}$ + jets, scale up	730	29310620
$t\bar{t}$ + jets, scale down	730	28354188
$t\bar{t}$ + jets, $m_t = 173.5$ GeV	711	19419050
$t\bar{t}$ + jets, $m_t = 171.5$ GeV	750	19578812
$t\bar{t}$ + jets, $h_{\text{damp}}$ up	750	29689380
$t\bar{t}$ + jets, $h_{\text{damp}}$ down	750	29117820

- number of the associated pixel hits  $\geq 2$ , and
- number of the associated pixel+strips hits  $\geq 5$ .
- **Track clustering:** After track selection, the level of association of tracks from a vertex is quantified. This procedure is called track clustering where tracks are clustered according to their  $z$ -position from the center of beamspot. The Deterministic Annealing algorithm [109, 110] is used for such clustering. After clustering, the candidate vertices are found along with the tracks associated with them.
- **Vertex-position fitting:** The  $z$ -position of candidate vertices with at least two tracks are fitted using the *adaptive vertex fitter* [111]. The fit returns  $x$ ,  $y$ ,  $z$ -position, and covariance matrix of the PV along with other parameters such as the number of degrees of freedom ( $n_{\text{dof}}$ ) and a weight associated with each track. We require  $n_{\text{dof}} > 4$ .

The simulated samples listed in Table 5.2 were generated before the actual data taking in 2016. While simulating these samples, the pileup distribution as initially seen in the data was used. During the actual data taking, however, the pileup distributions were changed. To have a similar pileup distribution as in data, the simulated events are re-weighted. After applying the pileup weights, as discussed in Section 6.2, the distribution of number of primary vertices per event ( $N^{\text{vertex}}$ ) is shown in Figures 5.1(a) and 5.1(b) for the  $\mu + \text{jets}$  and  $e + \text{jets}$  channel, respectively.

### 5.3.2. Diffuse offset energy density

From Figures 5.1(a) and 5.1(b), it can be seen that there is a poor agreement between data and simulation even after applying the pileup reweighting. A similar trend is seen in many analyses at 13 TeV [31, 32]. There is another variable similar to  $N^{\text{vertex}}$  called diffuse offset energy density ( $\rho$ ) [112, 113]. For  $n$  number

of jets in an event,  $\rho$  is given by [114]

$$\rho = \text{median} \left( \frac{p_{T,n}}{A_n} \right) \quad (5.1)$$

where  $p_{T,n}$  and  $A_n$  are the transverse momentum and area of  $n^{\text{th}}$  jet. The  $\rho$  distribution is shown in Figures 5.1(c) and 5.1(d). It can be seen that  $\rho$  has a better agreement between data and simulation as compared to the  $N^{\text{vertex}}$  distribution. The  $\rho$  is used in the calculation of isolation variable for muon and electron as shown in Equation (5.4). The jet energy corrections are also derived using  $\rho$  as discussed in Section 5.6.2.

## 5.4. Muon

### 5.4.1. Reconstruction

Muons are often referred to as minimum ionising particles. Their energy loss is much smaller than electrons due to its larger mass. Muon in CMS detector with sufficiently higher  $p_T$  can traverse through most of the sub-detectors such as tracker, calorimeter, and the muon chambers. Most of the high  $p_T$  muons go out of the CMS detector. Due to the presence of a magnetic field, trajectories of muon are bent in the CMS detector. The number of hits in the tracker, energy deposit in the calorimeter and the hits in the muon chambers (DT, CSC, and RPC) are used in the Kalman filtering (KF) technique [115] to reconstruct muon candidates. There are several categories of muons [116] described below:

- **Standalone muons:** These are reconstructed from the muon chambers only. Here, the KF technique uses the hits from DT, CSC, and RPC to reconstruct standalone muons.
- **Global muons:** The trajectory of a standalone muon is extrapolated up to the tracker. The best pair consisting of the standalone muon and tracker track is selected using the KF technique. Finally, these two tracks are combined to reconstruct global muon. As the global muons are reconstructed starting from the outer part (muon chambers) of the CMS detector back to the inner part (strip tracker), this approach of reconstruction is called *outside-in*.
- **Tracker muons:** There is an *inside-out* approach to reconstruct muons, which starts from the tracks reconstructed in the strip tracker to the energy deposits in the calorimeters up to the hits in the muon chamber. Muons reconstructed with this approach are called tracker muons.
- **RPC muons:** These muons are also reconstructed using the *inside-out* approach where only RPC hits of the muon chamber are used [117].

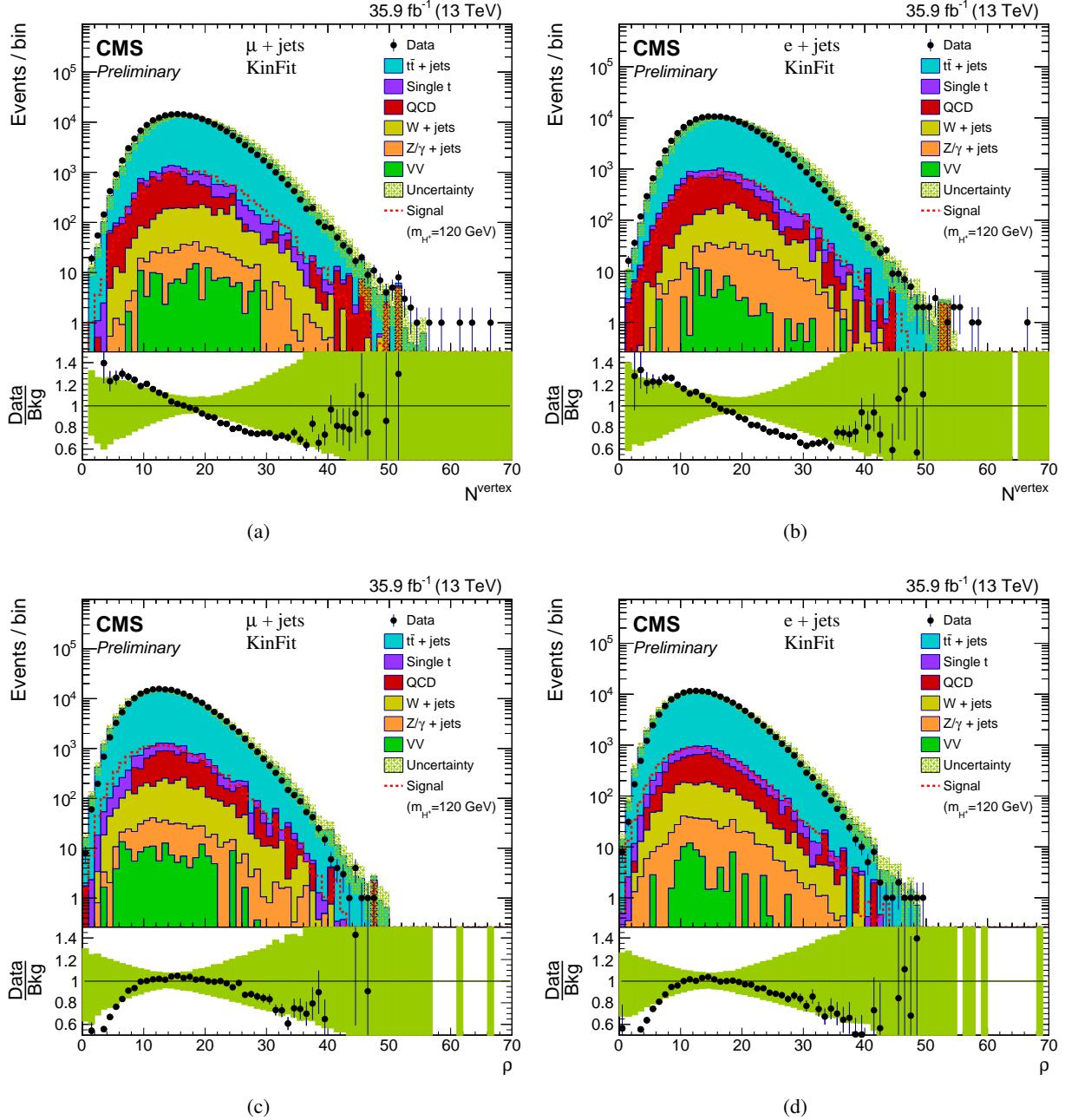


Figure 5.1.: Distribution of number of primary vertices and diffuse offset energy density for  $\mu + \text{jets}$  and  $e + \text{jets}$  channels. The agreement between data and simulation is better for  $\rho$  as compared to  $N^{\text{vertex}}$  distribution. However, for lower values of  $\rho$ , there is a significant mismatch between data and simulation. Similar trend is seen in other analysis as well [31, 32].

Table 5.4.: Selection and veto cuts applied on muon.

Variable	Selection	Veto
$p_T$ (GeV)	$> 26$	$> 15$
$ \eta $	$< 2.4$	$< 2.4$
Global or tracker muon	Yes	Global
Normalised $\chi^2$	$< 3$	
$\chi^2$ local position	$< 12$	
Tracker kink	$< 20$	
Segment compatibility	$> 0.303$	
$ d_{xy}(\text{vertex}) $ (cm)	$< 0.05$	
$ d_z(\text{vertex}) $ (cm)	$< 0.2$	
Inner track valid fraction	$> 0.8$	
$I_{\text{rel}}^\mu$	$< 0.15$	$< 0.25$

- **Calorimeter muons:** Some of the reconstructed muon tracks consist of energy deposits from the calorimeters. Such muons are called calorimeter muons.

#### 5.4.2. Identification and selection

In the event topology of our interest, there only one lepton (electron or muon) as shown in Figure 2.4. These are the 'prompt' lepton coming directly from the electroweak decay of W bosons. However, there could be 'extra' leptons that come from other sources such as misidentification of light flavored jets or charged pion decay in case of muons. Medium identification (ID) criteria, as listed Table 5.4, are applied to select prompt muons. The efficiency of medium muon ID as a function of  $p_T$  and  $\eta$  is shown in Figure 5.2 from data and simulation. From this figure, it can be seen that the efficiency at lower  $p_T$  is relatively low compared to that from high  $p_T$ . However there is a sudden drop in the efficiency for data from era BCDEF (The data taking periods are divided in different eras. For example, the first few months after the data taking starts are assigned era-A and so on. Note that some of the eras may have number of days less than 30) in the  $p_T > 120$  GeV region.

#### 5.4.3. Rochester correction

Due to misalignment of the magnetic field and azimuth dependent in-efficiency of the muon detector, the efficiency of the muon needs to be corrected. The efficiency depends on the charge,  $p_T$ ,  $\eta$  and  $\phi$  of muon. These corrections referred as Rochester corrections are applied on data as well as simulated samples. The

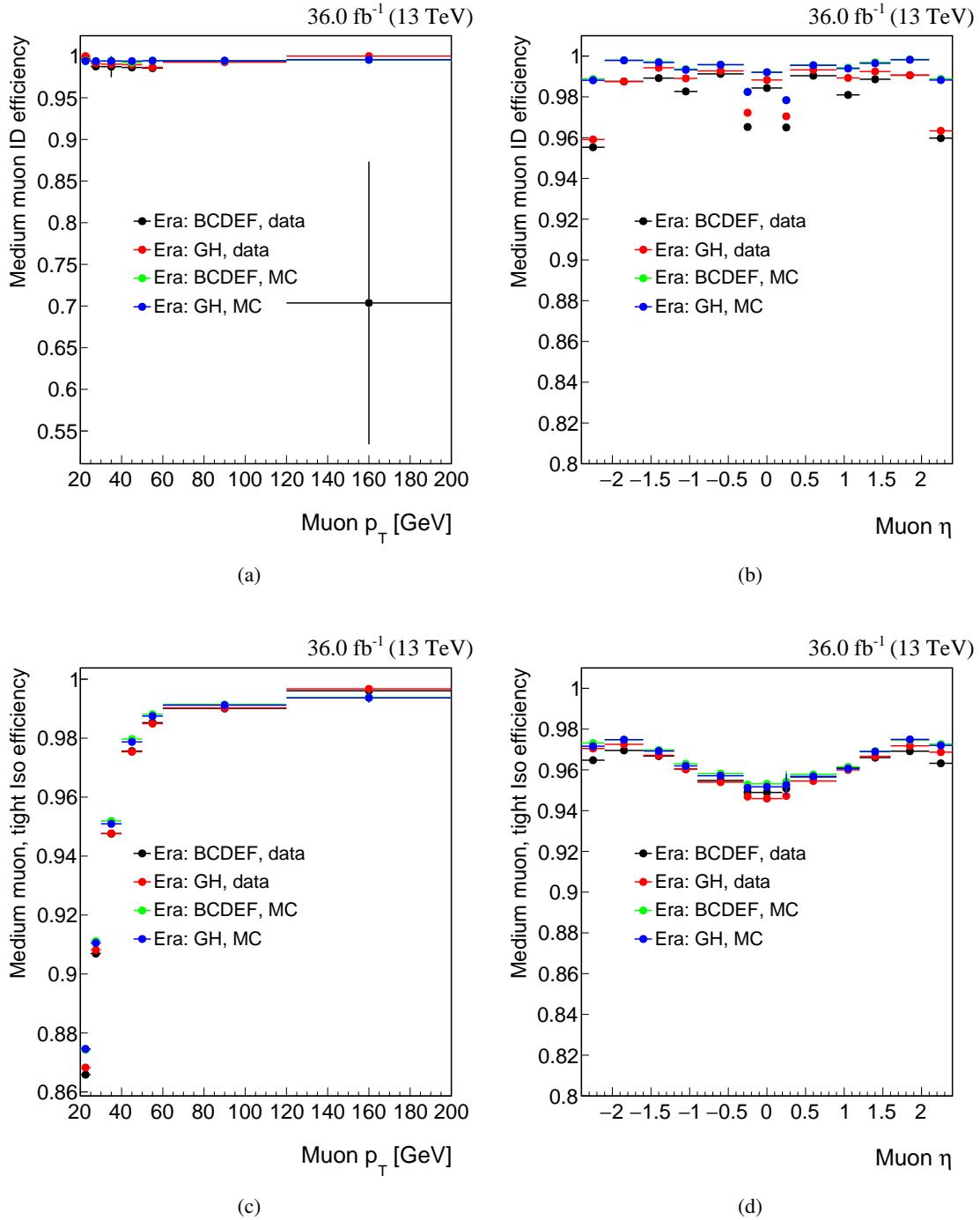


Figure 5.2.: The distribution of efficiencies for medium muon ID and tight relative isolation (with medium ID) as a function of  $p_T$  and  $\eta$  of muon from simulation and data from different eras of 2016 data taking.

muon  $p_T$  is corrected using Rochester correction tools at 13 TeV [118]. The Rochester scale factors depend on the  $p_T$ ,  $\eta$ ,  $\phi$  and charge of the muon. Selection criteria (Table 5.4) on muon are applied after Rochester correction.

#### 5.4.4. Isolation

A muon track in the proton-proton collision is not necessarily isolated i.e., it can be surrounded by other particles. On the other hand, muons from the W decay are expected to be isolated. To achieve this, we apply a criterion on the relative isolation, defined as the  $p_T$  sum of all particles excluding the muon within a cone of radius 0.4 around the muon direction divided by the muon  $p_T$  i.e.  $I_{\text{rel}}^\mu = \sum_{i \neq \mu} p_T^i / p_T^\mu$ . We use a more specific definition given by

$$I_{\text{rel}}^\mu = \frac{\sum p_T^{\text{ch}} + \max[(\sum E_T^\gamma + \sum E_T^{\text{neut}} - 0.5 \times \sum p_T^{\text{chPU}}), 0]}{p_T^\mu}, \quad (5.2)$$

where  $p_T^{\text{ch}}$  is the transverse momentum of charged hadrons;  $E_T^\gamma$  and  $E_T^{\text{neut}}$  are the transverse energy of photons and neutral hadrons; and  $p_T^{\text{chPU}}$  is the transverse momentum of charged hadrons associated with the pileup vertices. The factor 0.5 takes into account the neutral tracks from the leading vertex and charged tracks from the pileup vertices [27]. A tight muon isolation cut of  $I_{\text{rel}}^\mu < 0.15$  is applied. The efficiency as a function of  $p_T$  and  $\eta$  of tight muon isolation with medium ID is shown in Figure 5.2 [119].

### 5.5. Electron

#### 5.5.1. Reconstruction

Electrons are also reconstructed using the PF algorithm [92,93] based on the tracks in the tracker and energy deposits in the electromagnetic calorimeter [120]. Being a charged particle electrons radiate as they travel inside the huge magnetic field of the CMS. Thus the energy deposited by an electron in the ECAL is spread in  $\eta$  and  $\phi$  direction across several crystals. The energy deposited in various crystals are clustered using “hybrid” algorithm in the barrel and “multi-5x5” in endcap part of the ECAL. For given hits in the innermost layer of the tracker, the KF technique is used to find hits in the next layer of the tracker. The hits collected from all layers of the tracker is mapped to the energy deposit in the ECAL to form an electron candidate. To take care of the effect of bremsstrahlung, the Gaussian sum filter (GSF) method is used to fit the tracks. The electron charge is estimated from the curvature of the electron track and by matching associated KF track with GSF tracks. The vector joining (supercluster, beamspot) and (supercluster, first hit of the GSF track) in  $\phi$  is also used for electron charge estimation. The measurements from tracker and ECAL are combined to

Table 5.5.: Selection and veto cuts applied on electron found in barrel and endcap regions.

Variable	Selection		Veto	
	$ \eta_{\text{sc}}  \leq 1.479$	$ \eta_{\text{sc}}  > 1.479$	$ \eta_{\text{sc}}  \leq 1.479$	$ \eta_{\text{sc}}  > 1.479$
$p_{\text{T}}$ (GeV)	> 30	> 30	> 15	> 15
$ \eta $	< 2.5	< 2.5	—	—
full $5 \times 5 \sigma_{i\eta i\eta}$	< 0.00998	< 0.0298	< 0.0115	< 0.037
$ \Delta\eta(\text{sc}, \text{track}) $	< 0.00311	< 0.00609	< 0.00749	< 0.00895
$ \Delta\phi(\text{sc}, \text{track}) $	< 0.103	< 0.045	< 0.228	< 0.213
$E_{\text{had}}/E_{\text{em}}$	< 0.253	< 0.0878	< 0.356	< 0.211
$I_{\text{rel}}^e$	< 0.0821	< 0.0695	< 0.175	< 0.159
$ \frac{1}{E} - \frac{1}{p}  (\text{GeV})^{-1}$	< 0.134	< 0.13	< 0.299	< 0.15
Missing inner hits	$\leq 1$	$\leq 1$	$\leq 2$	$\leq 3$
Conversion veto	Yes	Yes	Yes	Yes
$ d_{xy}(\text{vertex})  (\text{cm})$	< 0.05	< 0.10	< 0.05	< 0.05
$ d_z(\text{vertex})  (\text{cm})$	< 0.10	< 0.20	< 0.1	< 0.1

estimate the electron momentum [120].

### 5.5.2. Identification and selection

To select a prompt electron, medium cut-based ID as listed in Table 5.5 is used. Events having an electron with  $p_{\text{T}} > 30$  GeV and  $|\eta| < 2.5$  are excluded.

A brief description of the variables listed in Table 5.5 is given below:

- full  $5 \times 5 \sigma_{i\eta i\eta}$  is defined as [120]

$$\sigma_{i\eta i\eta} = \sqrt{\sum (0.0175 \times \eta_i + \eta_{\text{cryst}}^{\text{seed}} - \bar{\eta}_{5 \times 5})^2 \times w_i} / \sum w_i, \quad w_i = 4.2 + \log(E_i/E_{5 \times 5}), \quad (5.3)$$

the summation is carried over the  $5 \times 5$  matrix around the highest  $E_T$  crystal of the supercluster (sc).

- $|\Delta\eta(\text{sc}, \text{track})|$ : Difference in  $\eta$  of the supercluster and the electron track.
- $|\Delta\phi(\text{sc}, \text{track})|$ : Difference in  $\phi$  of the supercluster and the electron track.
- $E_{\text{had}}/E_{\text{em}}$ : Ratio of the electron energy deposited in the HCAL and ECAL.
- $|\frac{1}{E} - \frac{1}{p}|$ :  $E$  is the energy of the supercluster and  $p$  is the track momentum at the point of closest approach from the PV.
- Missing inner hits: Number of missing hits in the strip tracker.

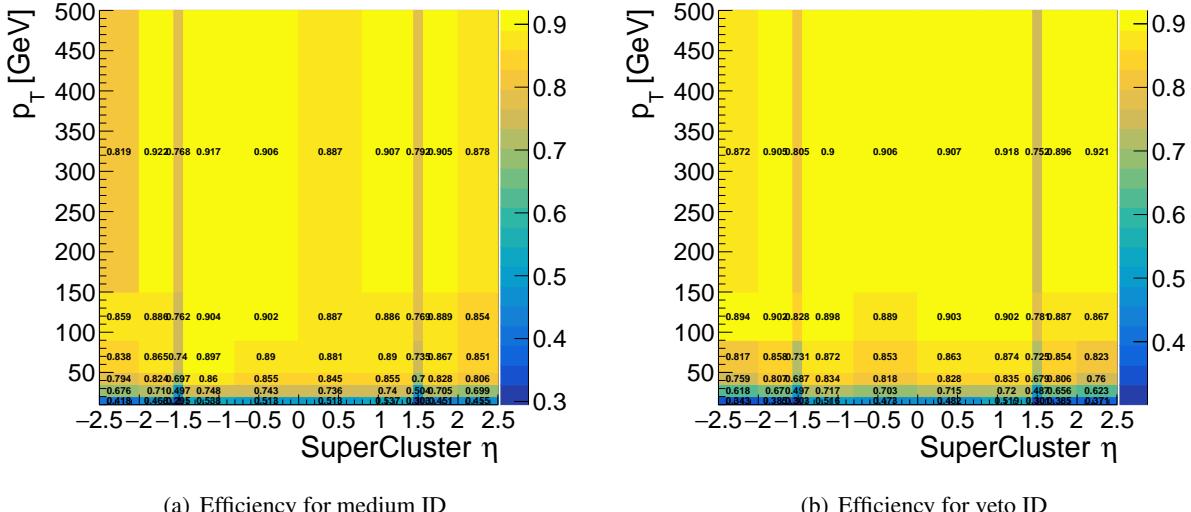


Figure 5.3.: The distribution of efficiency for medium and veto electron ID.

- $|d_{xy}(\text{vertex})|$ : Distance of closest approach of the electron track from the PV in the  $xy$  plane.
- $|d_z(\text{vertex})|$ : Difference in the  $z$ -coordinate of the PV and the electron track.

The relative isolation  $I_{\text{rel}}^e$  and conversion veto of Table 5.5 are discussed in Sections 5.5.3 and 5.5.4. The efficiency of medium and veto electron ID as a function of  $p_T$  and  $\eta$  is shown in Figure 5.3 [121].

### 5.5.3. Isolation

The relative isolation of an electron is given by:

$$I_{\text{rel}}^e = \frac{\sum p_T^{\text{ch}} + \max[(\sum E_T^\gamma + \sum E_T^{\text{neut}} - \rho \times A_{\text{eff}}), 0]}{p_T^e}, \quad (5.4)$$

where  $\rho$  is as defined in Equation (5.1) and  $A_{\text{eff}}$  is the effective area. These two quantities are used to correct electron isolation from pileup. The effective area for different  $\eta_{\text{sc}}$  are shown in Table 5.6. The tight isolation cut of  $I_{\text{rel}}^e < 0.0821$  (0.0695) is applied on electrons from the barrel (endcap) region.

### 5.5.4. Conversion rejection

Some of the reconstructed electrons are end products of photon conversion inside the tracker. While selecting prompt electrons these secondary electrons need to be rejected. As the secondary electrons don't have

Table 5.6.: Effective area for different  $\eta_{\text{sc}}$  range.

Super-cluster eta ( $\eta_{\text{sc}}$ )	$A_{\text{eff}}$
$0.0 \leq \eta_{\text{sc}} < 1.0$	0.1703
$1.0 \leq \eta_{\text{sc}} < 1.5$	0.1715
$1.5 \leq \eta_{\text{sc}} < 2.0$	0.1213
$2.0 \leq \eta_{\text{sc}} < 2.2$	0.1230
$2.2 \leq \eta_{\text{sc}} < 2.3$	0.1635
$2.3 \leq \eta_{\text{sc}} < 2.4$	0.1703
$2.4 \leq \eta_{\text{sc}}$	0.2393

hits in the innermost layer of the tracker [120], one can use this feature to reject them. By fitting the electron tracks, these secondary electrons are also rejected [120].

## 5.6. Jet

### 5.6.1. Reconstruction

Due to color confinement [122], the quarks and gluons produced in proton-proton collisions cannot exist in the free states. Instead, they hadronise into a cluster of colorless particles such as hadrons, leptons, and photons. We form a jet out of these particles, emanating in a cone around the initial direction of quarks and gluons. In CMS, the PF algorithm is used to reconstruct all possible particles that can be inside a jet. Subsequently, the clustering of, or the reconstruction of jets from, these particles are done with the anti- $k_T$  algorithm [123].

### 5.6.2. Identification and selection

The selection criteria applied on jets are listed in Table 5.7. The identification of b and c jet are discussed in Sections 5.8 and 8.8. The charged hadron subtraction technique [124] is used to take care of the pileup effects. It ignores the charged particles coming from the pileup vertices. The jet energy in data and simulated samples are corrected to account for the non-linear response of the ECAL and HCAL. A set of corrections are applied in a factored manner, that is, the output of previous serves as the input to the next step. The raw  $p_T$  of a jet is corrected at every step and the corrected  $p_T$  is used in the subsequent analysis. These set of corrections are known as L1FastJet (pileup correction through offset energy density) L2RelativeL3Absolute (detector response corrections), and L2L3Residual (residual corrections

Table 5.7.: Selection criteria applied on jets.

Variable	selection
$p_T$ (GeV)	> 25
$ \eta $	< 2.4
Neutral hadron energy fraction	< 0.99
Neutral electromagnetic energy fraction	< 0.99
Number of constituent	> 1
Charged hadron energy fraction	> 0
Charged hadron multiplicity	> 0
Charged hadron electromagnetic energy fraction	< 0.99

on data). The first two are applied on simulated samples whereas all the three are applied on data. Furthermore, the jet  $p_T$  in the simulation are smeared to have same resolution as that of data using the jet energy resolution (JER) scale factors. Before applying the selection cuts listed in Table 5.7, the jet  $p_T$  is corrected as discussed in Section 6.4.

### 5.6.3. Cleaning

The prompt leptons (electrons and muons) are required to be outside the jet cone. This procedure is called jet cleaning, where we require the angular separation between the jet and lepton,  $\Delta R > 0.5$ .

## 5.7. Missing transverse energy

The missing transverse energy ( $E_T^{\text{miss}}$ ) is the negative vector sum of the momenta of all objects reconstructed with the PF algorithm. Being a weakly interacting particle, neutrinos are not directly detected through the CMS detector and thus contribute to  $E_T^{\text{miss}}$ . Note that this is not true only for the SM; there are some beyond-the-SM particles e.g., neutralinos in supersymmetry that could also contribute to  $E_T^{\text{miss}}$ . Nevertheless, as there is a neutrino in our signal events (see Figure 2.4) we expect large  $E_T^{\text{miss}}$  due to its escape from the detector. The  $E_T^{\text{miss}}$  in simulation and data can have different resolutions. Therefore,  $E_T^{\text{miss}}$  from simulated events are smeared using the Type-1 correction in which jet energy correction is propagated to the  $E_T^{\text{miss}}$ . The events having  $E_T^{\text{miss}} > 20$  GeV are selected as described in Section 7.2.

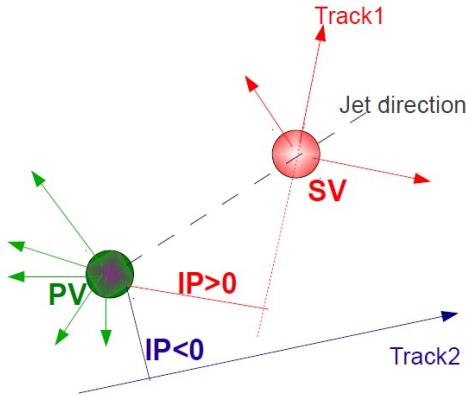


Figure 5.4.: Primary and secondary vertex with tracks and impact parameter. IP  $> 0$  or  $< 0$  if tracks are upstream or downstream along the jet direction. This figure is taken from [33].

## 5.8. b Jet Tagging

There are two b jets in the final state, as shown in Figure 2.4 for charged Higgs signal as well as SM  $t\bar{t}$  + jets background. Accurate identification of these b jets will substantially reduce the SM backgrounds such as  $W + \text{jets}$ ,  $Z/\gamma + \text{jets}$ ,  $VV$ , etc. The *combined secondary vertex* (CSVv2) method [125] is used for b tagging. The main idea behind the method is that b hadrons produced from the b quark (originating from the PV) have a relatively larger lifetime due to which they can travel measurable distance w.r.t. primary vertex in the transverse plane before decaying further (Figure 5.4). The position resolution of vertex in the transverse plane is about few  $\mu\text{m}$ . Therefore, a secondary vertex displaced by few hundred microns w.r.t. primary vertex can be identified and reconstructed. The transverse distance of tracks from the PV is called impact parameter (IP). To calculate the b jet discriminator values, first tracks of reasonable quality are selected, following which secondary vertex is reconstructed. Finally, a multi-layer perceptrons [126] is used to accept or reject the secondary vertex, discriminator value of the jet.

### 5.8.1. Track selection

The tracks are selected with the criteria listed in Table 5.8. These requirements ensure the tracks are closer to the PV and not coming from pileup vertices. After track selection, the SV is reconstructed as described in the next section.

Table 5.8.: Track selection criteria. PCA stands for the point of closest approach.

Variable	Selection
$p_T$ associated with tracks	$> 0.8$ GeV
number of hits associated with tracks in the tracker	$> 8$
longitudinal component of IP	$< 0.3$ cm
distance between PCA of tracks from the PV and jet-axis	$< 0.07$ cm
distance between PCA of tracks from the jet-axis and PV	$< 5$ cm
the displaced track should have	$IP > 50 \mu\text{m}$ and $\frac{IP}{\sigma_{IP}} > 1.2$

Table 5.9.: Secondary vertex selection criteria.  $d_{\text{PV}, \text{SV}}^{\text{trans}}$  is the 2D flight distance or the distance between a primary and secondary vertex in the transverse plane.

Variable	selection
number of tracks associated with the SV	$> 2$
$d_{\text{PV}, \text{SV}}^{\text{trans}}$	$> 0.1$ mm
$d_{\text{PV}, \text{SV}}^{\text{trans}}$	$< 2.5$ cm
mass associated with the SV	$< 6.5$ GeV
$\Delta R$ (jet axis, secondary flight direction)	$< 0.4$

### 5.8.2. Reconstruction of secondary vertex

For Run-I (proton-proton collisions at 8 TeV), the *combined secondary vertex* (CSV) method used the adaptive vertex reconstruction algorithm [127], which is based on adaptive vertex fitter [111] to reconstruct the SV. On the other hand, in case of Run-II (proton-proton collisions at 13 TeV), the CSVv2 uses the inclusive vertex finder (IVF) algorithm [128] for the same purpose. The collection of selected tracks are used as inputs to the IVF. The displaced tracks are used as seed. The cluster of tracks is formed from seed track depending on the angles and distance between them. The adaptive vertex fitter is used to fit the clusters. The selection listed in Table 5.9 are applied for the SV.

### 5.8.3. Multi-layer perceptron training

Three types of discriminating variable, as shown below, are used in multi-layer perceptron training to identify the secondary vertex.

- category-I: jet with  $N_{\text{SV}} > 1$ ,

Table 5.10.: The efficiency of loose (L), medium (M), and tight (T) b tag working points for different quark-flavor of jets [35]. These efficiencies are calculated from  $t\bar{t}$  events with jet  $p_T > 20$  GeV.

Working point	$\varepsilon^b$ (%)	$\varepsilon^c$ (%)	$\varepsilon^{uds}$ (%)	CSVv2
b tagger L	81	37	8.9	$> 0.5426$
b tagger M	63	12	0.9	$> 0.8484$
b tagger T	41	2.2	0.1	$> 0.9535$

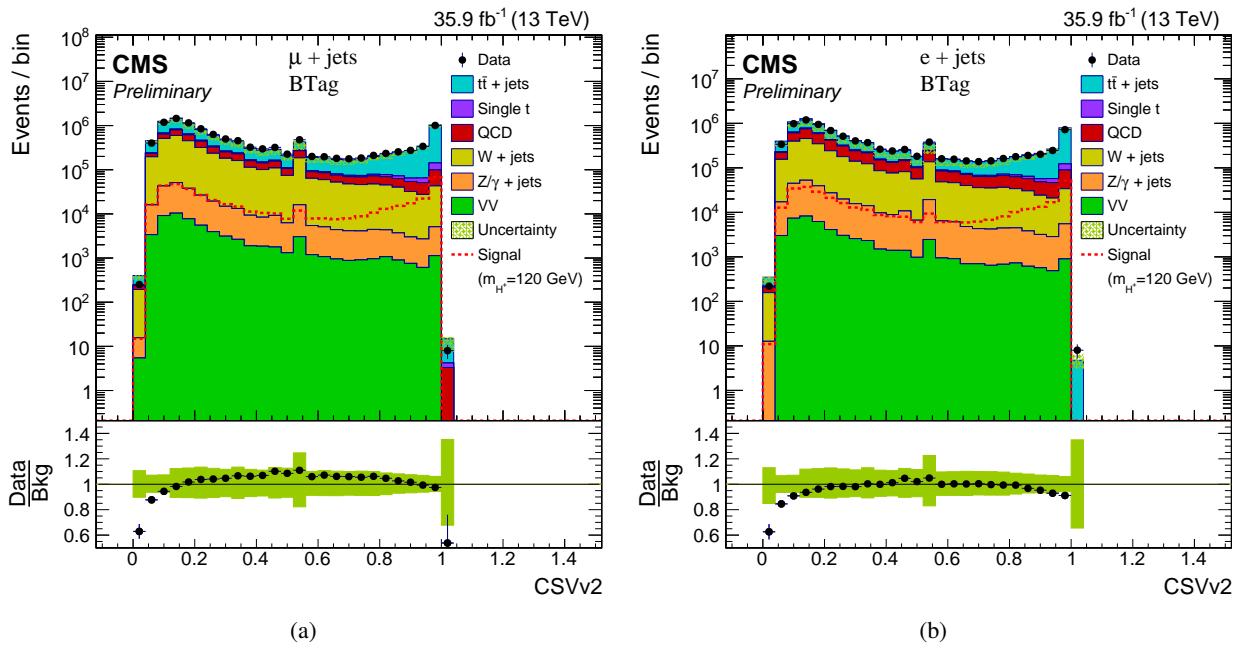


Figure 5.5.: The b discriminator distributions of all jets, after  $N_{\text{jets}} \geq 4$  selection as described in Section 7.2, obtained using CSVv2 method for the  $\mu + \text{jets}$  and  $e + \text{jets}$  channel.

- category-II: jet with pseudo-vertex [126], and
- category-III: jet with no SV or no pseudo-vertex.

The discriminating variables from the vertex-categories are combined using a likelihood ratio, which gives the b discriminator value. The b discriminator value for the  $\mu + \text{jets}$  and  $e + \text{jets}$  channel is shown in Figure 5.5. There are three official b tagging working points: *loose*, *medium*, and *tight* with b tag efficiency [defined in Eq. (6.9)] 81%, 63%, and 41%, respectively. The corresponding probability of a light jet being misidentified as b jet is 10%, 1%, and 0.1%. In this analysis, we use the medium working point i.e., b discriminator  $> 0.8484$  for b tagging. The efficiency of all b taggers for b, c, and light quarks are shown in Table 5.10.



# 6 | Corrections on Simulated Events

## 6.1. Luminosity scale factors

The simulated events listed in Table 5.2 are arbitrarily generated. While comparing distributions obtained using simulated events with that of data, the former has to be normalised with the integrated luminosity of data. To achieve this, each simulated sample is multiplied by the following luminosity scale factor ( $SF_L$ ):

$$SF_L = \frac{L_{\text{data}}}{L_{\text{simulation}}} = \frac{L_{\text{data}} \times \sigma_{\text{simulation}}}{N_{\text{simulation}}}, \quad (6.1)$$

where  $L_{\text{data}} = 35.9 \text{ fb}^{-1}$  and  $N_{\text{simulation}}$  is the number of simulated events and  $\sigma_{\text{simulation}}$  is the corresponding cross section for the given sample. As can be seen from Table 5.2, the  $SF_L$  value is less than 1 for most of the background samples except for QCD multijet. For QCD multijet samples, it is very large compared to other backgrounds owing to less number of generated events compared to its cross section. Simulation of large number of QCD multijet events is not feasible because of computing limitations.

## 6.2. Pileup reweighting

The pileup distributions in the data and simulation are different. To have the same pileup distribution, the simulated events are reweighted. Depending on the number of primary vertices, the simulated event is multiplied by corresponding pileup weight. The pileup weights are the ratio of pileup distributions from data and simulation.

## 6.3. Lepton scale factors

Trigger, tracking, isolation and identification efficiencies are different between simulation and data. A  $p_T$  and  $\eta$  dependent scale factors are applied to simulated events to take care of this difference. The muon scale factors are era dependent. There are different muon scale factors for era BCDEF and GH. However, electron

scale factors are the same for full 2016 data taking period. Separate scale factors are combined into one as

$$SF^\mu = SF_{ID}^\mu \times SF_{iso}^\mu \times SF_{track}^\mu \times SF_{trig}^\mu, \quad SF^{ele} = SF_{ID}^{ele} \times SF_{reco}^{ele} \times SF_{trig}^{ele}, \quad (6.2)$$

where  $SF_{ID}^\mu$ ,  $SF_{iso}^\mu$ ,  $SF_{track}^\mu$  and  $SF_{trig}^\mu$  are weighted average with luminosity ( $L$ ) for different eras (BCDEF and GH) e.g.

$$SF_{ID}^\mu = \frac{SF_{ID}^\mu(\text{BCDEF}) \times L(\text{BCDEF}) + SF_{ID}^\mu(\text{GH}) \times L(\text{GH})}{L(\text{BCDEF}) + L(\text{GH})}, \quad (6.3)$$

where, for example,  $SF_{ID}^\mu(\text{BCDEF})$  is the scale factor for era-B to era-F and  $L(\text{BCDEF})$  is the luminosity during this period. A similar formula for  $SF_{iso}^\mu$ ,  $SF_{track}^\mu$  and  $SF_{trig}^\mu$  is used. The scale factors of Equation (6.2) are applied on simulated events.

## 6.4. Jet and $E_T^{\text{miss}}$ correction

Jets from simulations are smeared using jet energy scale (JES) and jet energy resolution (JER) to have the same resolution as that in data [129]. For smearing,  $p_T$  and  $\eta$  dependent scale factors are used as listed in Table 6.1. The  $p_T$  of jet in simulations is scaled by the following factor:

$$p_{T\text{scale}} = \max[0.0, 1.0 + (SF - 1) \times (p_{T\text{jet}} - p_{T\text{jet}}^{\text{gen}})/p_{T\text{jet}}^{\text{gen}}] \quad (6.4)$$

with the following constraint:

$$p_{T\text{jet}}^{\text{gen}} > 0, \quad \Delta R < 0.2, \quad |p_{T\text{jet}} - p_{T\text{jet}}^{\text{gen}}| < 3 \times \sigma_{\text{JER}} \times p_{T\text{jet}}, \quad (6.5)$$

where  $\Delta R$  is the angular separation between the reconstructed and generated jet and  $\sigma_{\text{JER}}$  is resolution of  $p_T$  of reconstructed jet.

## 6.5. b tag scale factor

The b tagging efficiency is different between simulation and data. An event weight as given by Equation (6.8) is applied on the simulated events to take care of this difference [130].

$$P(\text{Simulation}) = \prod_{i=\text{tagged}} \varepsilon_i \prod_{j=\text{not tagged}} (1 - \varepsilon_j) \quad (6.6)$$

$$P(\text{Data}) = \prod_{i=\text{tagged}} SF_i \varepsilon_i \prod_{j=\text{not tagged}} (1 - SF_j \varepsilon_j) \quad (6.7)$$

$$w = \frac{P(\text{Data})}{P(\text{Simulation})} \quad (6.8)$$

Table 6.1.: Jet energy resolution scale factors for different  $\eta$  range

$\eta$ range	base	down	up
$0.0 \leq  \eta  < 0.5$	1.109	1.044	1.174
$0.5 \leq  \eta  < 0.8$	1.138	1.072	1.204
$0.8 \leq  \eta  < 1.1$	1.114	1.050	1.178
$1.1 \leq  \eta  < 1.3$	1.123	1.022	1.224
$1.3 \leq  \eta  < 1.7$	1.084	0.985	1.183
$1.7 \leq  \eta  < 1.9$	1.082	0.973	1.191
$1.9 \leq  \eta  < 2.1$	1.140	1.020	1.260
$2.1 \leq  \eta  < 2.3$	1.067	0.953	1.181
$2.3 \leq  \eta  < 2.5$	1.177	0.967	1.387
$2.5 \leq  \eta  < 2.8$	1.364	1.203	1.525
$2.8 \leq  \eta  < 3.0$	1.857	1.654	2.060
$3.0 \leq  \eta  < 3.2$	1.328	1.203	1.453
$3.2 \leq  \eta  < 5.0$	1.160	1.013	1.307

The b tagging efficiency  $\varepsilon_f(m, n)$  in the  $(m, n)$  bin of  $p_T$  and  $\eta$  is calculated using the following formula

$$\varepsilon_f(m, n) = \frac{N_f^{\text{b-tagged}}(m, n)}{N_f^{\text{total}}(m, n)}, \quad (6.9)$$

where  $N_f^{\text{b-tagged}}(m, n)$  is the number of b-tagged jets with flavor f (b quark, c quark, light quark, and gluon) and  $N_f^{\text{total}}(m, n)$  is the total number of events. The scale factor ( $SF_i$ ) depends on  $p_T$ ,  $\eta$ , parton flavor of jet, and b jet discriminator value.



# 7 | Event Filter and Selection

## 7.1. Filters

A series of filters [131] are applied to filter “bad” events, such as those where there is a huge difference in the  $p_T$  of muon measured from the tracker and the muon chambers. Although the occurrence of such “bad” events is rare, these filters are applied as sanity measures. These filters mostly affect the actual data, most of them have no effect on the simulated events. The efficiency of all combined filters is 99.4% for data, and 99.9% for simulated  $t\bar{t} + \text{jets}$  sample. These filters are particularly useful in those analyses where there is a requirement of very high  $E_T^{\text{miss}}$  and the final event yield is small. The list of filters with their availability and applicability is shown in Table 7.1. A detailed description of each filter is given below:

- **HBHE noise filter:** It is known that the barrel and endcap part of the HCAL records noise (sporadic anomalous signals) at a fixed rate even if there is no stable beam for physics data taking [131]. Such an event is filtered.
- **HBHE isolated noise filter:** This filter first finds the candidates for rechits and rejects the event if the sum of iso-tagged energy is  $> 50$  GeV, the sum of iso-tagged transverse energy  $> 25$  GeV, and the number of iso-tagged rechits is  $> 9$ .
- **CSC beam halo filter:** The actual physics objects of interest are those which are produced in the proton-proton collisions. However, there are machine induced particles, produced through beam-gas or beam-pipe interactions, which fly with the beam at a large radius (up to 5 m). These machine induced particles are reconstructed mainly in the cathode strip chambers (CSC) as muon candidate. An event is filtered if such particles are present.
- **Primary vertex filter:** The events are required to have at least one good primary vertex. Therefore, those events are filtered where there is no good vertex. A good vertex is required to be real, the z-component of the position is  $< 24$  cm, the  $\rho$  component of the position is  $< 2$  cm, and the minimum number of degrees

Table 7.1.: List of event filters applied to data and simulated samples. Most of the filters are available in the MINIAOD dataset through the trigger collection. The unavailable filters are applied on the fly using EDFilters.

Name of the filter	Available in MINIAOD	Applied on
HBHE noise filter	Yes	Data & Simulation
HBHE isolated noise filter	Yes	Data & Simulation
CSC beam halo filter	Yes	Data & Simulation
Good primary vertex filter	Yes	Data & Simulation
Bad EE supercrystal filter	Yes	Data
ECAL dead cell trigger primitive filter	Yes	Data & Simulation
Bad PF Muon Filter	No	Data & Simulation
Bad Charged Hadron Filter	No	Data & Simulation

of freedom is  $> 4$ .

- **Bad EE supercrystal filter:** The events are filtered if there is an unusually large energy deposit in the four superclusters of ECAL endcaps [131]. This filter is applied only on data.
- **ECAL dead cell trigger primitive filter:** If the energy of the ECAL crystals are estimated from the trigger primitive and lies near the saturation energy then that event is filtered [131].
- **Bad PF muon filter:** An event is filtered if the muon is although qualified as a PF muon but the quality is very low, and the  $p_T$  is large. The events are filtered if a muon is found which has  $p_T > 100$  GeV, segment compatibility  $< 0.3$  or relative percentage error in the  $p_T$  of the best (inner) track is more than 200% (100%), and  $\Delta R$  (muon, PF muon)  $< 0.001$ .
- **Bad charged hadron filter:** Those events are filtered where the quality of the muon is low and it is not declared as PF muon. However, this non-PF muon is used in the calculation of  $\text{PF-}E_T^{\text{miss}}$  as a candidate for charged hadron. To reject such events, the muon is required to have same  $p_T$ , segment compatibility, and error in track cuts as that of the above filter. An additional requirement of  $\Delta R$  (muon, charge hadron)  $< 0.00001$ , and  $\text{PtDiffRe1} < 0.00001$  is applied where  $\text{PtDiffRe1} = (p_T \text{ of PF candidate} - p_T \text{ of muon inner track}) / (0.5 \times (p_T \text{ of PF candidate} + p_T \text{ of muon inner track}))$ .

## 7.2. Selection

In the final state, there are four jets, one charged lepton and  $E_T^{\text{miss}}$ . Various selection cuts are applied to ensure the resulting events to have this topology. Cutflow plots after various selection cuts are shown in

Figure 7.1. Below we list out various selection requirements applied.

1. **Event filter and trigger:** The events are passed to the filters and triggers (Section 4.7). The lepton trigger is used to enrich events with lepton and significantly reduces QCD multijet compared to other SM background. The number of events surviving filters and trigger cuts for simulated signal and background and data are shown in the 2nd column of Tables 7.2, 7.3, 7.4 and 7.5.
2. **Lepton selection:** The event topology has only one lepton thus events having a second loose lepton are not selected. The lepton-veto selection for the  $\mu + \text{jets}$  and  $e + \text{jets}$  channel are listed in Tables 5.4 and 5.5. For the  $\mu + \text{jets}$  ( $e + \text{jets}$ ) channel there cannot be any electron (muon) in an event. Those events where there is any electron (muon) in the  $\mu + \text{jets}$  ( $e + \text{jets}$ ) channel are rejected. Lepton scale factors as described in Section 6.3 are applied to the simulated events. The relative isolation cut for muon is  $I_{\text{rel}}^\mu < 0.15$  and for electron it is  $I_{\text{rel}}^e < 0.0821(0.0695)$  in the barrel (endcap) region. Event yields after one lepton selection are shown in the 3rd column of Tables 7.2, 7.3, 7.4 and 7.5.
3. **Jet selection:** There should be at least four jets in an event. Jet selection cuts are listed in Table 5.7. The jet energy is corrected using JES and JER scale factors. After correcting jet energy, the selections on jet as shown in Table 5.7 are applied. Number of events after jet selection are shown in the 4th column of Tables 7.2, 7.3, 7.4 and 7.5.
4.  **$E_T^{\text{miss}}$  selection:** The missing transverse energy should be greater than 20 GeV. Event yields after  $E_T^{\text{miss}}$  selection are shown in the 5th column of Tables 7.2, 7.3, 7.4 and 7.5. After applying  $E_T^{\text{miss}}$  selection QCD multijet,  $Z/\gamma + \text{jets}$  event yields reduce more compared to  $t\bar{t} + \text{jets}$ , single t,  $W + \text{jets}$ , and  $VV$  processes as there is no neutrino at the parton level.
5. **b jet selection:** For b jet selection, the medium working point is used i.e,  $b_{\text{discr}} > 0.8484$ . The b tag event weights, as described in Section 6.5, are applied on simulated events. The events are required to have at least two b jets. Event yields for simulated signal, background and data are shown in the 6th column of Tables 7.2, 7.3, 7.4 and 7.5.

From Table 7.2 (7.3) it can be seen that the  $W + \text{jets}$  process is the dominant background upto 1-lepton selection. However, when the events are required to have  $N_{\text{jets}} \geq 4$ , the  $t\bar{t} + \text{jets}$  becomes the dominant background. The  $t\bar{t} + \text{jets}$  remains dominant background after subsequent selection cuts. The QCD multijet events are from simulated samples as shown in Table 5.2. The signal event yields for a different mass of charged Higgs after various selection cuts are shown in Table 7.4 (7.5) for the  $\mu + \text{jets}$  ( $e + \text{jets}$ ) channel. It can be seen from these tables that the event yields are almost the same upto 1-lepton selections for all masses

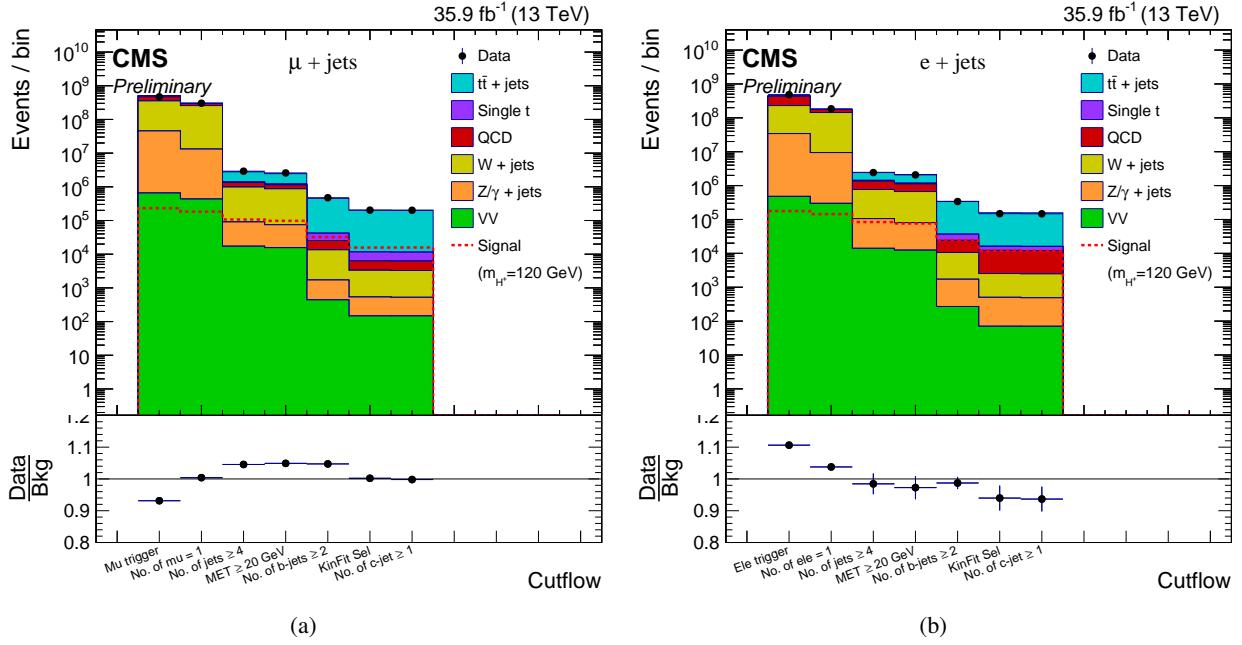


Figure 7.1.: Event yields after various selection cuts for the  $\mu + \text{jets}$  and  $e + \text{jets}$  channel.

of charged Higgs. However, after the  $N_{\text{jets}} \geq 4$  selection, when  $t\bar{t} + \text{jets}$  becomes the dominant background as shown in Table 7.2 (7.3), the event yield for higher masses of the charged Higgs are reduced more compared to lower masses. The event yield reduces because of the less phase space available between  $t$  quark and charged Higgs for higher masses.

### 7.3. Comparison of data and background after b jet selection

Data to background comparison of variables from reconstructed objects after applying b jet selection as described in Section 7.2 are shown in Figures 7.2, 7.3, and 7.4 for  $\mu + \text{jets}$  and  $e + \text{jets}$  channel. There is a good agreement between data and simulated background within statistical and systematics uncertainties for all variables except for the  $p_T$  (of lepton and jets) and  $E_T^{\text{miss}}$ . There is a poor agreement between data and background for higher values of  $p_T$  and  $E_T^{\text{miss}}$ . The disagreement in these distributions can be fixed by applying the  $t$  quark  $p_T$  weights. However, the  $t$  quark  $p_T$  weights do not affect the  $m_{jj}$  distribution which is the final observable for this analysis. Therefore, we do not apply  $t$  quark  $p_T$  weights. The  $m_{jj}$  distribution after b jet selection is shown in Figures 8.1(a) and 8.1(b) for both channels where we do see a good agreement between data and SM backgrounds.

Table 7.2.: Event yields after various selection cuts for the  $\mu + \text{jets}$  channel. The  $W + \text{jets}$  process is the dominant background upto 1-lepton selection. When the events are required to have  $N_{\text{jets}} \geq 4$ , the  $t\bar{t} + \text{jets}$  becomes the dominant background. The QCD multijet events are from simulated samples as listed in Table 5.2. The simulated MC signal process corresponds to  $m_{H^+} = 120$  GeV.

Process	Trigger	$N_\mu = 1$	$N_{\text{jets}} \geq 4$	$E_T^{\text{miss}} \geq 20 \text{ GeV}$	$\geq 2 \text{ b jets}$	KF selection	$\geq 1 \text{ c jet}$
MC signal	232088	182966	107237	97792	32309	15719.1	15784.7
SM $t\bar{t} + \text{jets}$	4.02837e+06	2.6683e+06	1.36181e+06	1.24994e+06	411521	191136	190228
Single $t\bar{t}$	589196	428074	81804	74924.6	16599.5	5311.76	5272.39
$W + \text{jets}$	3.09765e+08	2.44546e+08	894318	797849	11767.7	2795.35	2760.97
$Z/\gamma + \text{jets}$	4.48481e+07	1.27487e+07	73347.7	59133.7	1272.68	389.494	377.772
QCD multijet	1.4453e+08	4.07546e+07	355254	265303	11962.6	3007.9	2960.23
VV	653661	433207	17232.4	15614.6	440.143	146.569	146.862
Bkg	5.04414e+08	3.01579e+08	2.78377e+06	2.46276e+06	453564	202787	201746
Data	4.69828e+08	3.02791e+08	2.91052e+06	2.58346e+06	475011	203209	201361
Data/Bkg	0.931433	1.00402	1.04553	1.04901	1.04729	1.00208	0.998092

Table 7.3.: Event yields after various selection cuts for the  $e + \text{jets}$  channel. A similar trend as that of Table 7.2 is seen. The simulated MC signal process corresponds to  $m_{H^+} = 120$  GeV.

Process	Trigger	$N_e = 1$	$N_{\text{jets}} \geq 4$	$E_T^{\text{miss}} \geq 20 \text{ GeV}$	$\geq 2 \text{ b jets}$	KF selection	$\geq 1 \text{ c jet}$
MC signal	176709	143701	83955.8	76103.8	24488.2	11805.2	11835.4
SM $t\bar{t} + \text{jets}$	3.00511e+06	2.04055e+06	1.03705e+06	949027	306375	141807	141111
Single $t$	436844	314868	63532.5	57857.3	12741.7	3988.54	3953.51
$W + \text{jets}$	1.95739e+08	1.34112e+08	660657	586662	8878.8	2031.55	1991.47
$Z/\gamma + \text{jets}$	3.38256e+07	9.13592e+06	91830.9	68347.4	1449.81	433.03	418.256
QCD multijet	2.05104e+08	3.2298e+07	595856	453793	13497.6	9678.07	9613.93
VV	479881	298702	14058.6	12404.1	270.078	70.432	70.2293
Bkg	4.3859e+08	1.782e+08	2.46299e+06	2.12809e+06	343213	158009	157158
Data	4.85205e+08	1.84925e+08	2.42496e+06	2.07011e+06	338837	148500	147210
Data/Bkg	1.10628	1.03774	0.984561	0.972755	0.98725	0.939821	0.936699

Table 7.4.: Signal event yields for the different mass of charged Higgs after various selection cuts for the  $\mu + \text{jets}$  channel. Event yields are almost same upto 1-lepton selection for all mass points. However, after the  $N_{\text{jets}} \geq 4$  selection (when  $t\bar{t} + \text{jets}$  becomes the dominant background as shown in Table 7.2) the event yield for higher masses of the charged Higgs are reduced more compared to lower masses. The event yield reduces because of the less phase space available between top-quark and charged Higgs for higher masses.

Process	Trigger	$N_\mu = 1$	$N_{\text{jets}} \geq 4$	$E_T^{\text{miss}} \geq 20 \text{ GeV}$	$\geq 2 \text{ b jets}$	KF selection	$\geq 1 \text{ c jet}$
$m_{H^+} = 80 \text{ GeV}$	233610	182859	103424	94638.9	33949.4	15577.6	15613.5
$m_{H^+} = 90 \text{ GeV}$	232966	182661	105445	96234.4	33860.8	15721	15732.8
$m_{H^+} = 100 \text{ GeV}$	234002	183693	107477	97946.7	34203	16320.8	16371.9
$m_{H^+} = 120 \text{ GeV}$	232088	182966	107237	97792	32309	15719.1	15784.7
$m_{H^+} = 140 \text{ GeV}$	233830	185420	101550	92786.9	26039	12460.9	12521.1
$m_{H^+} = 150 \text{ GeV}$	232622	185015	94656.9	86379.9	19889	8909.13	8955.59
$m_{H^+} = 155 \text{ GeV}$	233266	185615	90834.8	83035.3	16384.8	7014.31	7056.58
$m_{H^+} = 160 \text{ GeV}$	234577	186955	87634.3	80181.9	13517.2	5383.27	5407.94

Table 7.5.: Signal event yields for different mass of charged Higgs after various selection cuts for the  $e + \text{jets}$  channel. Similar trend as that of Table 7.4 is seen.

Process	Trigger	$N_e = 1$	$N_{\text{jets}} \geq 4$	$E_T^{\text{miss}} \geq 20 \text{ GeV}$	$\geq 2 \text{ b jets}$	KF selection	$\geq 1 \text{ c jet}$
$m_{H^+} = 80 \text{ GeV}$	174717	141754	79743.7	72559.8	25357.2	11736.4	11757.5
$m_{H^+} = 90 \text{ GeV}$	175758	142497	82160.4	74725.2	25885.4	12075.8	12075.9
$m_{H^+} = 100 \text{ GeV}$	175150	141911	82368.1	74892.7	25663.2	12206.6	12231.6
$m_{H^+} = 120 \text{ GeV}$	176709	143701	83955.8	76103.8	24488.2	11805.2	11835.4
$m_{H^+} = 140 \text{ GeV}$	175445	142881	78422.5	71394.9	19835.5	9494.45	9546.35
$m_{H^+} = 150 \text{ GeV}$	177098	144326	74020.9	67223	15325.4	6902.44	6937.26
$m_{H^+} = 155 \text{ GeV}$	176518	143676	70351.4	64047.7	12594.7	5397.84	5435.16
$m_{H^+} = 160 \text{ GeV}$	176109	143527	68012.6	62001.6	10333.9	4184.27	4204.59

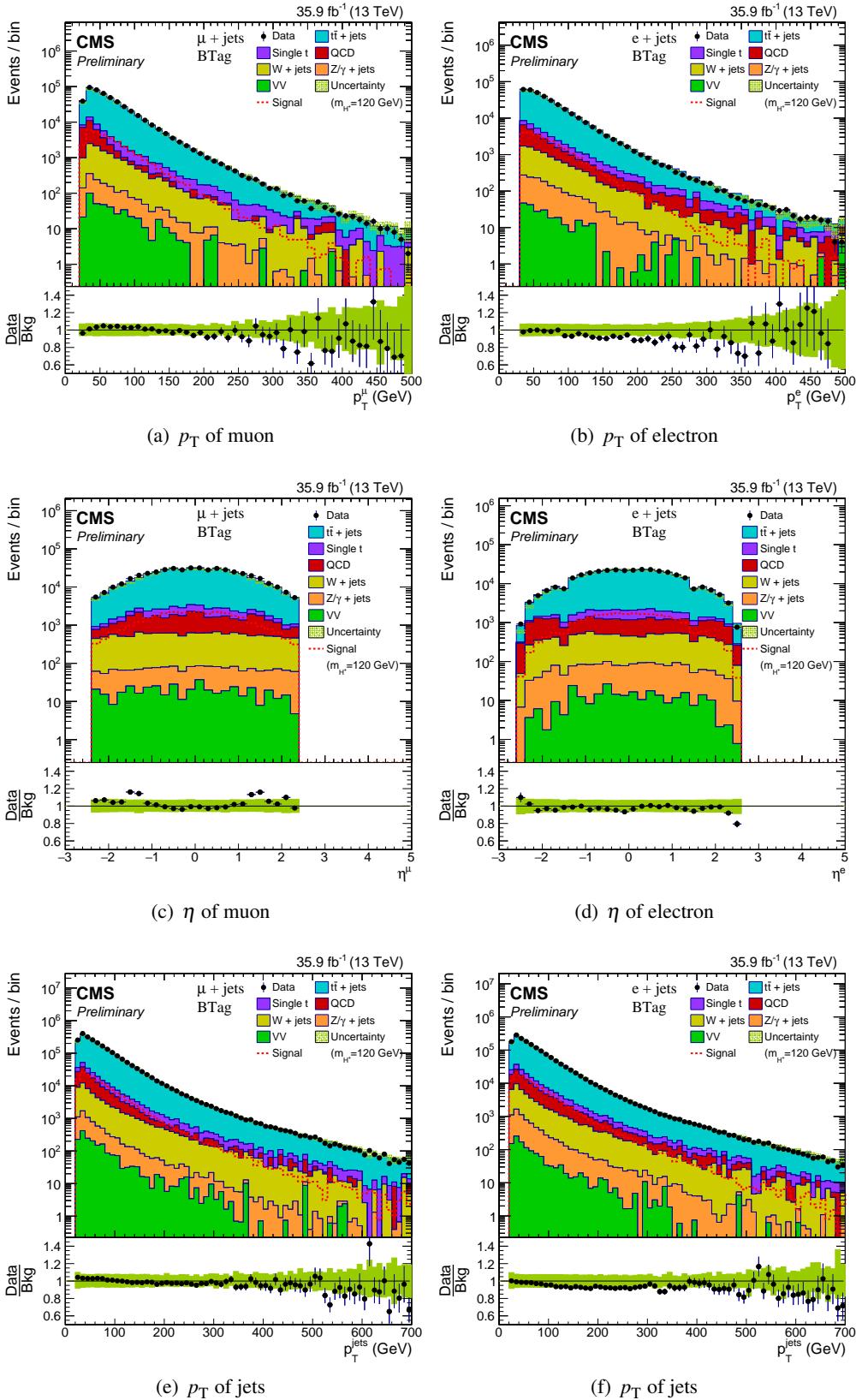


Figure 7.2.: Distribution of  $p_T$ ,  $\eta$  of reconstructed lepton and  $p_T$  of jets after b jet selection as described in Section 7.2, for  $\mu + \text{jets}$  and  $e + \text{jets}$  channel.

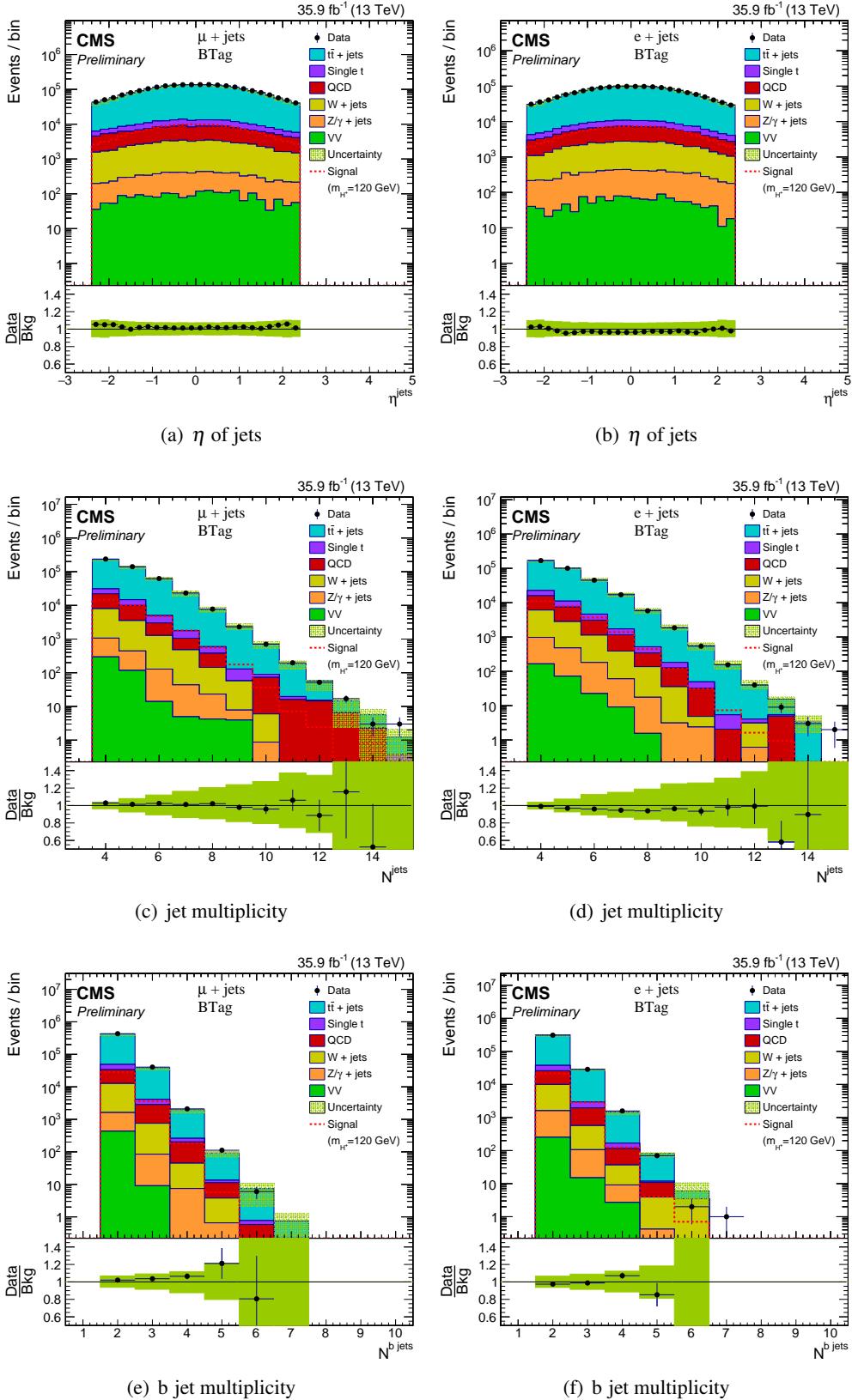


Figure 7.3.: Distribution of reconstructed  $\eta$  of jets, jet multiplicity, and b jet multiplicity after b jet selection as described in Section 7.2, for  $\mu + \text{jets}$  and  $e + \text{jets}$  channel.

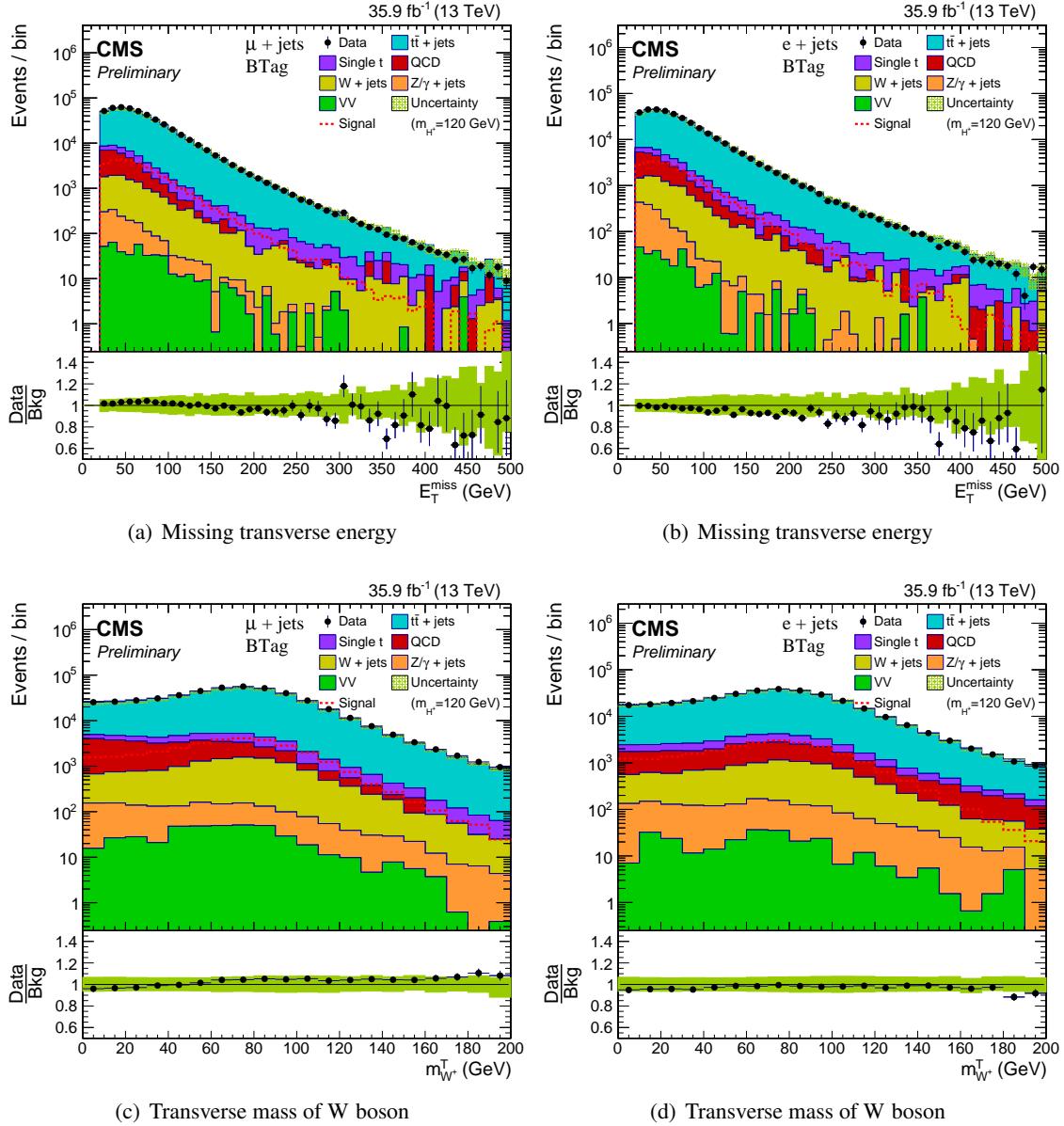


Figure 7.4.: Distribution of reconstructed  $E_T^{\text{miss}}$  and  $m_{W^+}^T$  (transverse mass of W boson decaying leptonically) after b jet selection as described in Section 7.2, for  $\mu + \text{jets}$  and  $e + \text{jets}$  channel.



# 8 | Kinematic Fitting and c Jet Tagging

In this analysis, the charged Higgs boson is decaying to the c and  $\bar{s}$  quark. The invariant mass of the c  $\bar{s}$  system ( $m_{jj}$ ) is thus used as the final observable. The  $m_{jj}$  distribution of two highest  $p_T$ , non b jets is shown in Figures 8.1(a) and 8.1(b) for both channels. For the true semileptonic  $t\bar{t}$  events, the mean of the  $m_{jj}$  distribution should be close to 80 GeV, which is the mass of the W boson. However, the mean of  $m_{jj}$ , as shown in Figures 8.1(a) and 8.1(b), is about 128 GeV because the two light jets in every event may not necessarily come from the decay of a W boson. Secondly, the  $m_{jj}$  distribution has a long tail which might constrain the search for new resonances in the dijet mode.

To select true semileptonic  $t\bar{t}$  events, a kinematic fit is performed on the reconstructed objects using the top kinematic fitter package [132]. The `TopKinFitter` takes physics objects such as lepton, jets,  $E_T^{\text{miss}}$ , and their resolutions as input, and gives improved four-vectors of lepton, jets, and neutrino with associated  $\chi^2$  and probability of the fit as output. It constrains the reconstructed t quark mass to its nominal value ( $m_t = 172.5$  GeV). In the output, the `TopKinFitter` gives only four jets (2 b jets from leptonic and hadronic modes, and 2 light jets from hadronic mode), 1 lepton, and neutrino. It also separates jets coming from leptonic and hadronic decay modes of  $t\bar{t}$ . The 2 light jets coming from hadronic decay mode are further used for c tagging as described in Section 8.8. A more detailed description of the kinematic fitting is given below.

## 8.1. Input to the TopKinFitter

After applying selection criteria on physics objects as described in Section 7.2, the event which is passed to the `TopKinFitter` contains only one lepton,  $E_T^{\text{miss}}$  and at least 4 jets. The b discriminator value with the medium working point is also given as the input to the `TopKinFitter` to separate b jet from light jets (more details in Section 8.2). The constraints on the  $t\bar{t}$  system along with the theoretical mass of the t quark are also specified (more details in Section 8.3). All the constraints and invariant mass are parameterized in

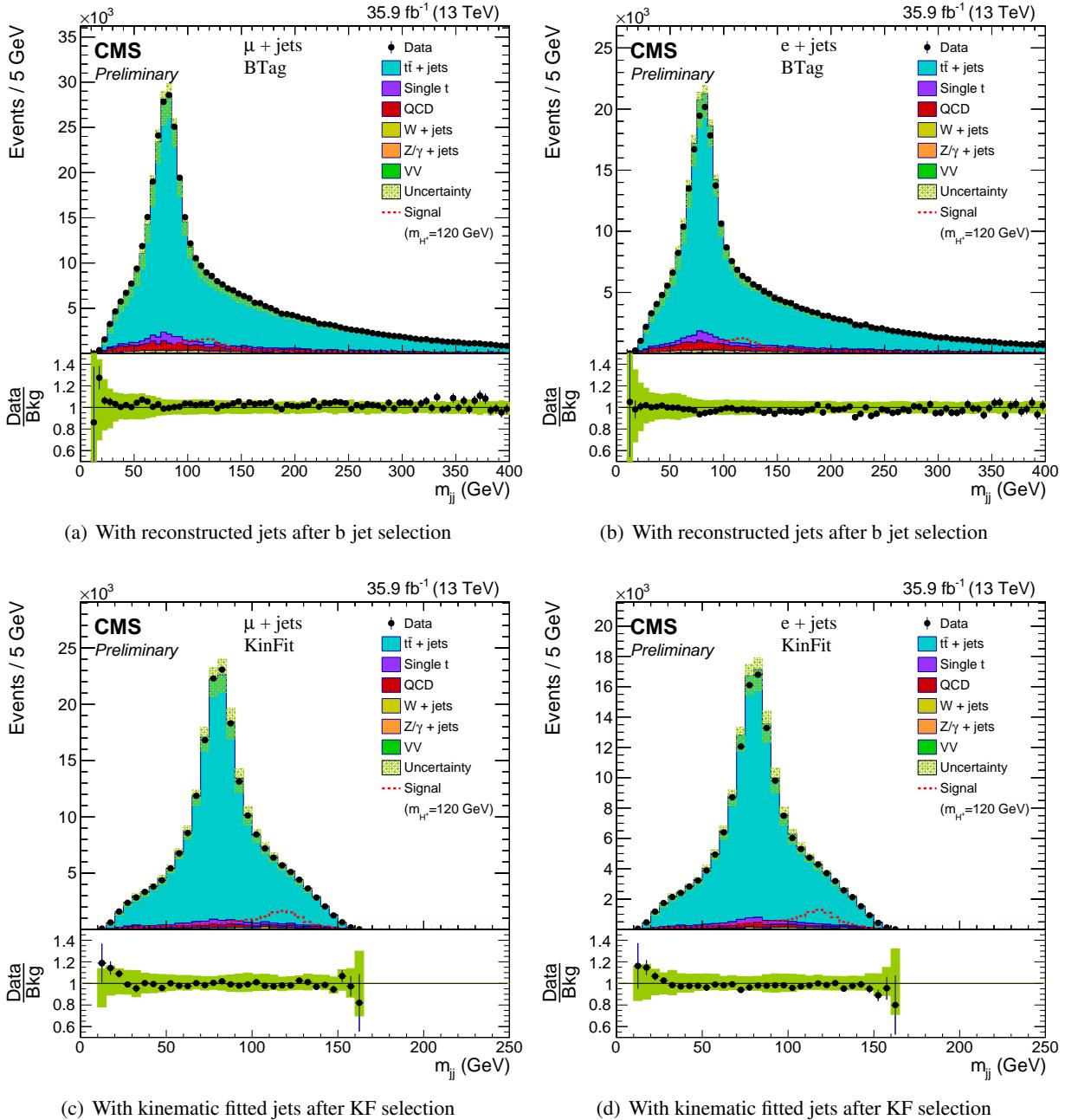


Figure 8.1.:  $m_{jj}$  distributions of two non b, highest  $p_T$  jets for the  $\mu + \text{jets}$  and  $e + \text{jets}$  channel. The distributions of Figures 8.1(a) and 8.1(b) are obtained using reconstructed jets after applying b tag scale factor as described in Section 7.2. On the other hand, the distributions of Figures 8.1(c) and 8.1(d) are calculated using kinematic fitted jets after kinematic fit selection. The mean of the invariant mass distribution from kinematic fitted jets is closer to the W mass as compared to that of reconstructed jets.

terms of  $E_T$ ,  $\eta$ , and  $\phi$  variables of the physics objects. The components of 4-momentum vector in terms of these variables are given as

$$E = E_T \sin \theta, p_x = E_T \cos \phi, p_y = E_T \sin \phi, p_z = E_T \cot \theta, \quad (8.1)$$

and the invariant mass of two particles, in the relativistic limit, is given by

$$m_{\text{inv}} = \sqrt{2E_{T_1}E_{T_2}(\cosh(\eta_1 - \eta_2) - \cos(\phi_1 - \phi_2))} \quad (8.2)$$

where  $\eta = \ln \cot(\theta/2)$ . The resolution of each physics object as a function of  $p_T$  and  $\eta$ , and the JER scale factors from different  $\eta$  binning are also given in the input. Apart from the physics objects, various inputs related to the minimization of  $\chi^2$  such as the maximum number of iterations, criteria on the convergence of the  $\chi^2$  are also specified (more details in Section 8.4).

## 8.2. Separation of jets

For the semileptonic decay mode of the  $t\bar{t}$ , the first step is to separate 2 b jets from light jets and the second step is to identify each b jet from hadronic and leptonic modes. All the jets are sorted in their  $p_T$  order and first 4 jets are selected. To find out the two b jets, a b tag probability is calculated using following formula [132]

$$L_b(x) = \frac{\text{PDF}_b(x)}{\sum_{i=1}^5 \text{PDF}_i(x)} \quad (8.3)$$

Where  $\text{PDF}_i(x)$  is the probability distribution function of b discriminator value (x) for flavour  $i$ . An event is selected if two jets have b tag probability of more than 60%. The jet-parton matching is also performed for each jet. For  $n$  number of jets, there are exactly  $n$  number of partons. The number of permutations in the jet-parton matching is given by

$$N_{\text{permutations}} = \frac{n!}{(n-4)!} \quad (8.4)$$

Therefore for 4 jets, there are 24 permutations. All the 24 permutations are shown in [133] through diagrams. However, after identifying two b jets, the number of permutations reduces to 12 as the two b jets are interchangeable. For each jet-parton permutation, a  $\chi^2$  is constructed, as described in Section 8.4, and the permutation with the lowest value of  $\chi^2$  is treated as the correct matching. Afterward, three jets (1 b jet and 2 light jets) have to be selected to form hadronic decaying t quark. For this, several sensitive variables are used [132] such as the angle between jets and lepton, angular separation between the generated and reconstructed jets, etc.

### 8.3. Kinematic constraints

The semileptonic decay mode of the  $t\bar{t}$  has 4 jets, 1 lepton and the neutrino in the final state. The x and y-component of the neutrino are taken from the  $E_T^{\text{miss}}$ , as the missing transverse energy is attributed to the neutrino. And the z-component of the neutrino,  $p_z^\nu$ , is determined from the fit. The following kinematic constraints are imposed on the semileptonic  $t\bar{t}$  system

$$m_{\text{inv}}(b^{\text{had}} q\bar{q}) = m_t = 172.5 \text{ GeV} \quad (8.5a)$$

$$m_{\text{inv}}(b^{\text{lep}} l\nu_l) = m_{\bar{t}} = 172.5 \text{ GeV} \quad (8.5b)$$

After the fit, the  $p_z^\nu$  is determined from the Equation (8.5b). For every event, a  $\chi^2$  is constructed as discussed in Section 8.4. The  $\chi^2$  is minimized by varying  $p_T$ ,  $\eta$ , and  $\phi$  of each object within their resolution. Those values of  $p_T$ ,  $\eta$ , and  $\phi$  variables are finally selected which minimises the  $\chi^2$  and at the same time satisfies Equation (8.5).

### 8.4. $\chi^2$ of the fit

For every selected event, a  $\chi^2$  is constructed which is given by

$$\chi^2 = \left( \frac{m_t - 172.5}{\sigma_{m_t}} \right)^2 + \left( \frac{m_{\bar{t}} - 172.5}{\sigma_{m_{\bar{t}}}} \right)^2 + \sum_i \left( \frac{p_i^{\text{fit, lep}} - p_i^{\text{lep}}}{\sigma_{p_i^{\text{lep}}}} \right)^2 + \sum_j \sum_i \left( \frac{p_i^{\text{fit, jet}_j} - p_i^{\text{jet}_j}}{\sigma_{p_i^{\text{jet}_j}}} \right)^2 \quad (8.6)$$

where  $\sigma_{m_t}$  is the resolution of the mass of the t quark and  $\sigma_{p_i}$  is the momentum resolution of three component ( $i = 1, 2, 3$ ) of the corresponding lepton and jets. The  $p^{\text{fit}}$  is the fitted momentum of the given lepton and jets. The  $\chi^2$ , given in Equation (8.6), is minimized using the Lagrangian multipliers under the constraints given in Equation (8.5). A detailed mathematical description for the minimization of the  $\chi^2$  is given in [132]. The  $\chi^2$  is minimized iteratively where in each step the 4 components of the momentum vector of each object are varied within their resolution and a  $\chi^2$  and the constraint is calculated. The fit is declared to be converged if the following constraints are satisfied

$$\frac{\chi^2(n-1) - \chi^2(n)}{\text{ndf}} < \varepsilon_\chi \quad (8.7a)$$

$$m_{\text{inv}}(b^{\text{had}} q\bar{q}) - 172.5 < \varepsilon_c \quad (8.7b)$$

$$m_{\text{inv}}(b^{\text{lep}} l\nu_l) - 172.5 < \varepsilon_c \quad (8.7c)$$

where  $\varepsilon_\chi = 5 \times 10^{-5}$ ,  $\varepsilon_c = 0.0001$ , and ndf is the number of degrees of freedom which is equal to 1 as we have two constraints and one free parameter ( $p_z^\nu$ ). The total number of iterations ( $n$ ) is 500.

## 8.5. Output of the TopKinFitter

For every event, the `TopKinFitter` gives the status of fit which is 0 if the fit converged and 1 if it did not,  $\chi^2$  of the fit, probability ( $P_\chi$ ) of the goodness-of-fit of  $\chi^2$ , and the 4 momentum vector of physics objects such as jets, lepton and a neutrino. The  $\chi^2$  and  $P_\chi$  are related by the following equation

$$P_\chi = \exp\left(-\frac{\chi^2}{2}\right) \quad (8.8)$$

The fit also separates jets from hadronic and leptonic decay mode of  $t\bar{t}$ , that is, it gives one hadronic b jet, one leptonic b jet, and 2 light jets from hadronic decay modes. The c tagging criteria are applied further on the 2 light jets as discussed in Section 8.8. The  $m_{jj}$  distribution of the light jets is used in further analysis.

## 8.6. Selection and performance

Due to wrong jet combinatorics, the fit does not converge for every event. Only those events are selected for which the fit converges. The efficiency of fit convergence is 73% for simulated  $t\bar{t}$  sample and 71% for data for both channels. Further, the angular separation between reconstructed and kinematic fitted lepton and jets are required to be less than 0.2 to make sure that they are almost in the same direction. The same  $\Delta R$  cut is applied on reconstructed and kinematic fitted jets. Also, the  $p_T$  cut on kinematic fitted lepton and jets are required to be the same as that on the reconstructed jets and lepton. After applying these additional cuts on  $\Delta R$  and  $p_T$ , the kinematic fit efficiency reduces to 47% for  $t\bar{t}$  and 44% for data for both channels. In summary, the following kinematic fit selections are applied:

- fit is converged,
- $\Delta R(\text{fitted lepton}, \text{reconstructed lepton}) < 0.2$ ,
- $\Delta R(\text{fitted jet}, \text{reconstructed jet}) < 0.2$

Event yields after kinematic fit selections are shown in the 7th column of Tables 7.2, 7.3, 7.4 and 7.5. From these tables, we see that almost half the number of events is reduced after these selections. The  $m_{jj}$  distribution from the two light jets after kinematic fit selection is shown in Figures 8.1(c) and 8.1(d) for both channels. The mean of these distributions is 84 GeV which is close to the mass of W boson.

The kinematic fit is also performed after applying jet energy corrections such as JES and JER. For up and down systematics of JES and JER, the fit is performed separately on every event after correcting jet  $p_T$  using the corresponding scale factors. The output of the fit for each systematic is stored in a different collection

Table 8.1.: The efficiency of loose (L), medium (M), and tight (T) c tag working points for different quark-flavor of jets [35]. These efficiencies are calculated from  $t\bar{t}$  events with jet  $p_T > 20$  GeV.

Working point	$\epsilon^c$ (%)	$\epsilon^b$ (%)	$\epsilon^{uds}$ (%)	pfCCvsL	pfCBvsB
c tagger L	88	36	91	> -0.48	> -0.17
c tagger M	40	17	19	> -0.1	> 0.08
c tagger T	19	20	1.2	> 0.69	> -0.45

using the `EDProducer`. The kinematic fitting is a time-consuming process which takes, on average 0.002 seconds per event for  $t\bar{t}$  process. Therefore, performing all (nominal, JESup, JESdown, JERup, and JERdown) the kinematic fitting takes a reasonable amount of time.

## 8.7. Comparison of data and background after kinematic fitting

Data to background comparison of variables from the kinematic fitted objects after kinematic fit selection are shown in Figures 8.2, 8.3, and 8.4. There is also a good agreement between data and simulation within the statistical and systematic uncertainties. Here also we see a similar disagreement for the higher value of  $p_T$  and  $E_T^{\text{miss}}$  as we have after b jet selection described in Section 7.2.

## 8.8. c Jet Tagging

As the charged Higgs boson decays to a charm and a strange antiquark, the identification of charm jet is expected to increase the signal significance. The charm tagging or c tagging is recently developed in the CMS collaboration [34] based on the CSVv2 method for the analyses at 13 TeV data. This procedure is similar to that of b tagging described in Section 5.8. At the end, we have two charm discriminators: c vs. b jet (`CvsB`) and c vs. light jet (`CvsL`). These are collectively used to tag a jet as c jet. There are three working points for the c tagging as shown in Table 8.1 and Figure 8.5. Distribution of these two discriminators after kinematic fit selection is shown in Figure 8.6 for the  $\mu + \text{jets}$  and  $e + \text{jets}$  channel.

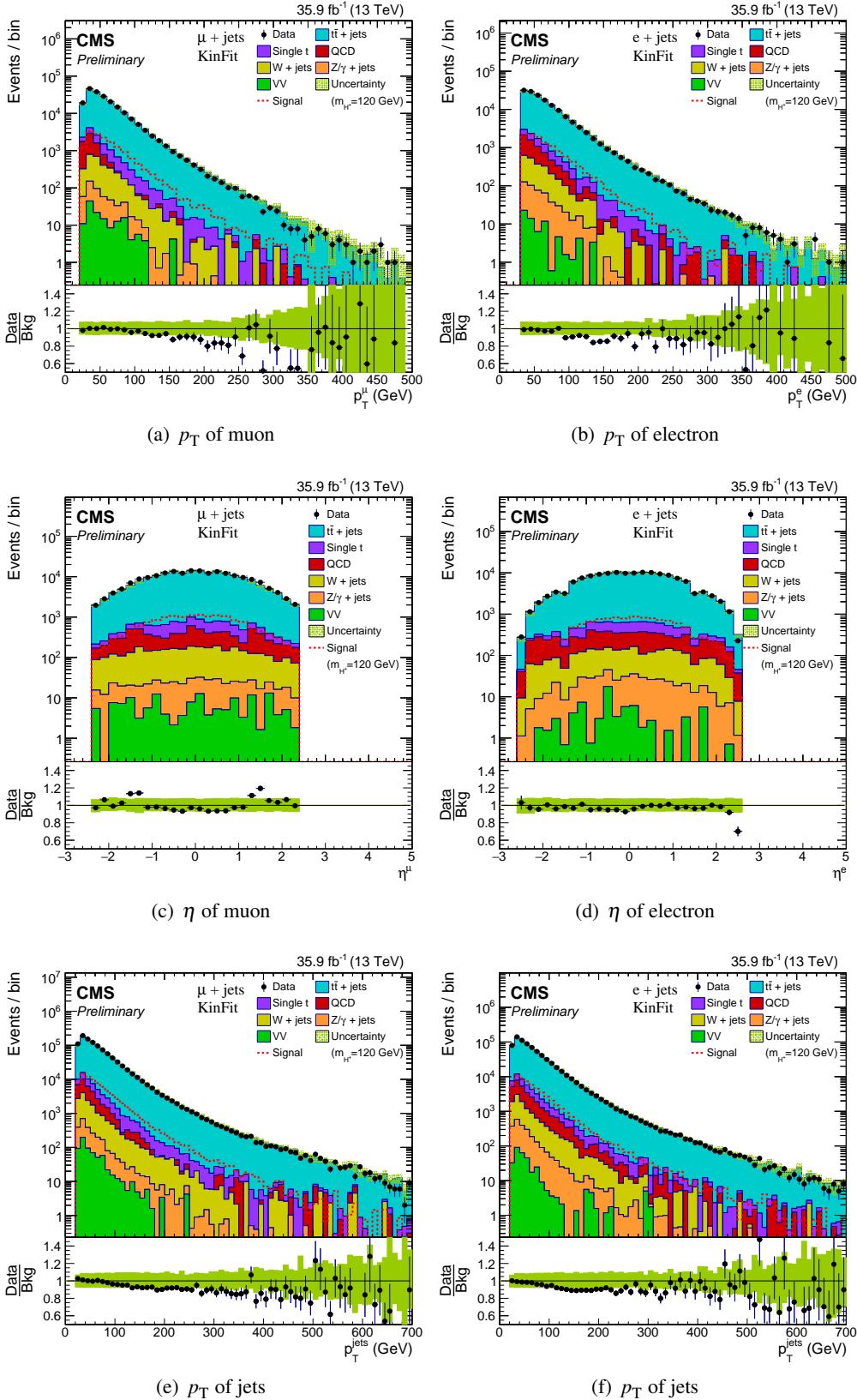


Figure 8.2.: Distribution of  $p_T$ ,  $\eta$  of kinematic fitted lepton and  $p_T$  of jets after kinematic fit selection for  $\mu + \text{jets}$  and  $e + \text{jets}$  channel.

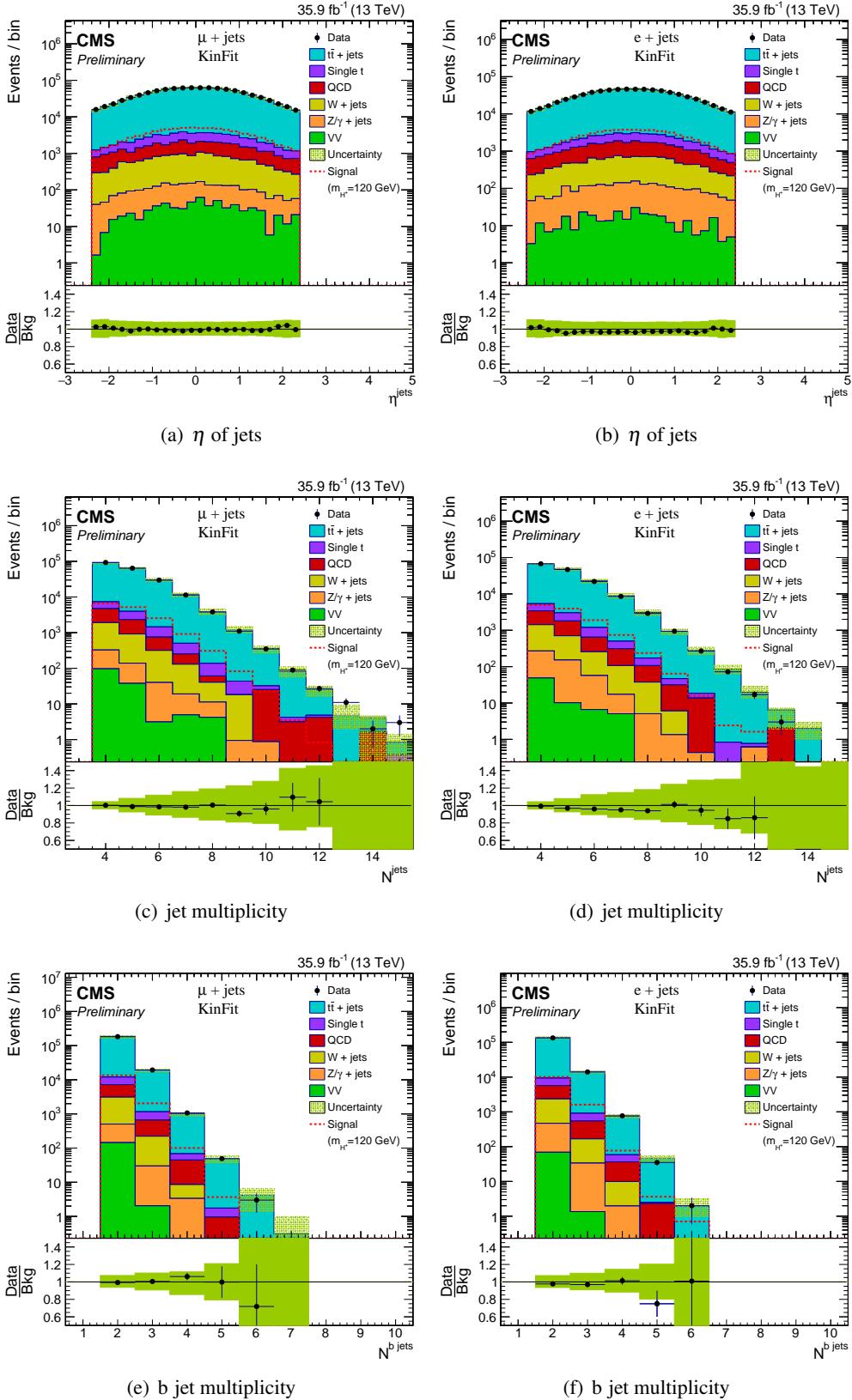


Figure 8.3.: Distribution of kinematic fitted  $\eta$  of jets, jet multiplicity, and b jet multiplicity after kinematic fit selection for  $\mu + \text{jets}$  and  $e + \text{jets}$  channel.

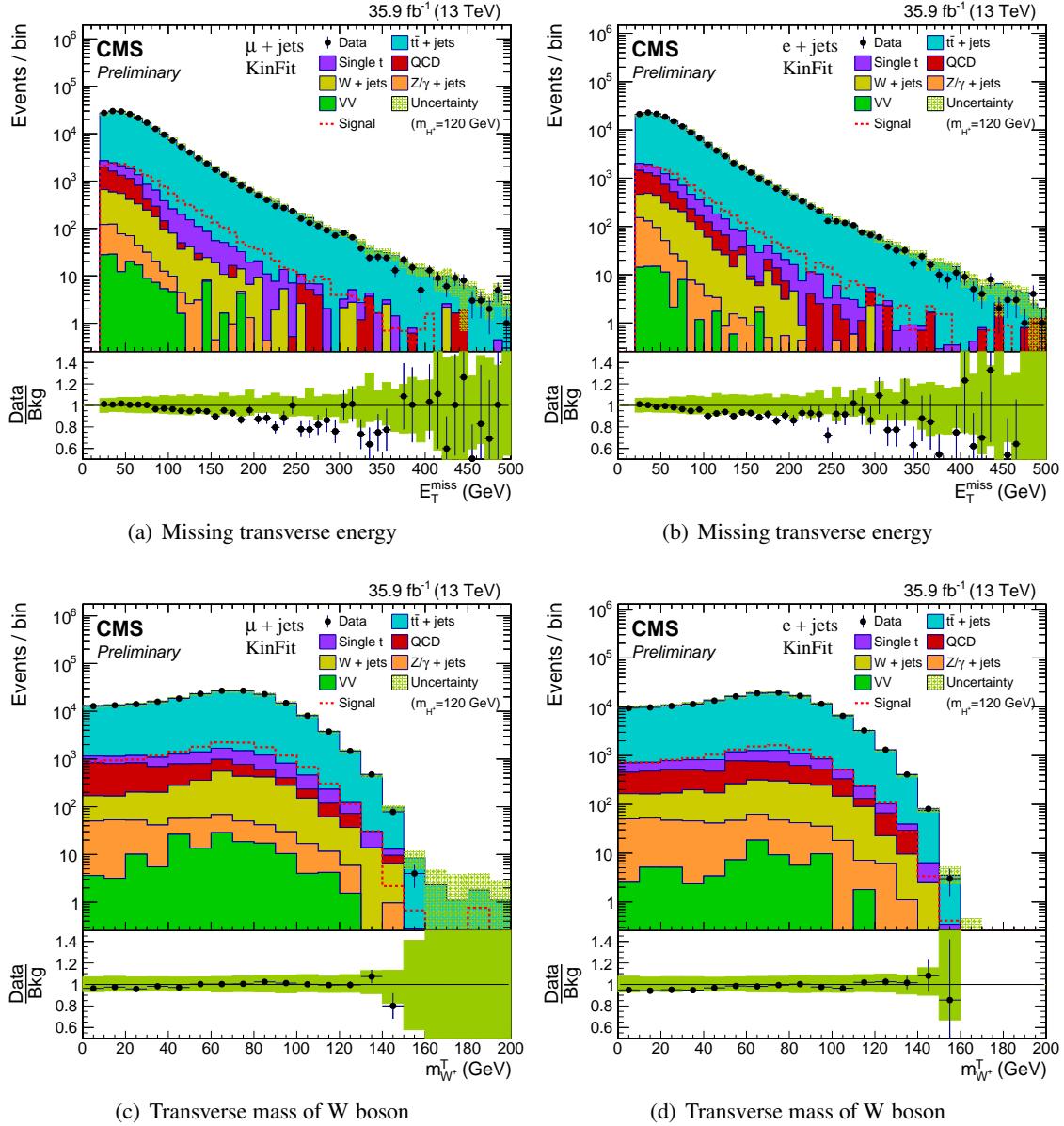


Figure 8.4.: Distribution of kinematic fitted  $E_T^{\text{miss}}$  and  $m_{W^+}^T$  after kinematic fit selection for  $\mu + \text{jets}$  and  $e + \text{jets}$  channel.

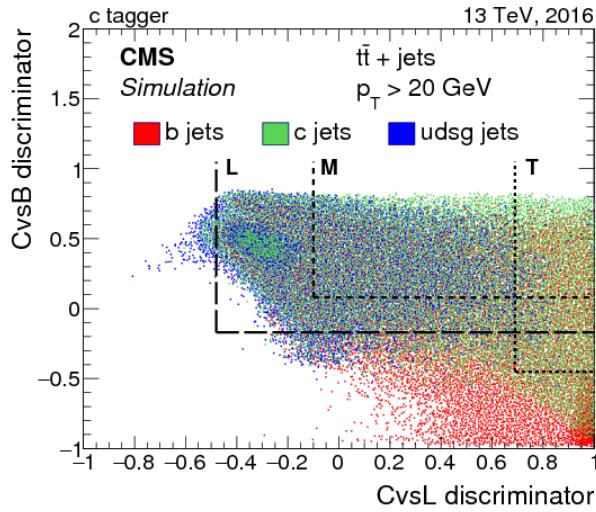


Figure 8.5.: The 2D plot between the two c taggers. The region right to the vertical and above to the horizontal line correspond to different charm working points (WPs). Unlike the b tag WPs, there is an overlap between the loose, medium, and tight c tag WPs. This figure is taken from [34].

## 8.9. c tag scale factor and selection

Similar to the event weights for b tagging, the event weight for c jet tagging is also applied due to the mismatch in the efficiencies between simulation and data. The mistagging of c-tagged jet or tagging a non c tagged jet as c jet is done using the same procedure as that of Section 6.5. The c tagging requirement is applied on the two light jets from the kinematic fit. Event yield after demanding at least one c-tagged jet with the loose working point for signal, background, and data are shown in 8th column of Tables 7.2, 7.3, 7.4 and 7.5. The  $m_{jj}$  distribution for the loose c tagging is shown in Figure 10.2.

Initially, each working point was separately used for c tagging. However, the limits were not getting improved for medium and tight working points. Therefore the loose working is finally used. Although the signal to background ratio increases as one go from loose to tight working points, the limits are not improved because the event yield also goes down. However, it is found that, as described in Section 10.3, that if the events after kinematic fit selection are *exclusively* divided into the categories based on loose, medium, and the tight charm working points and the limit is computed by combining the  $m_{jj}$  distributions from these three categories, as described in Section 12.5, then the limit gets improved.

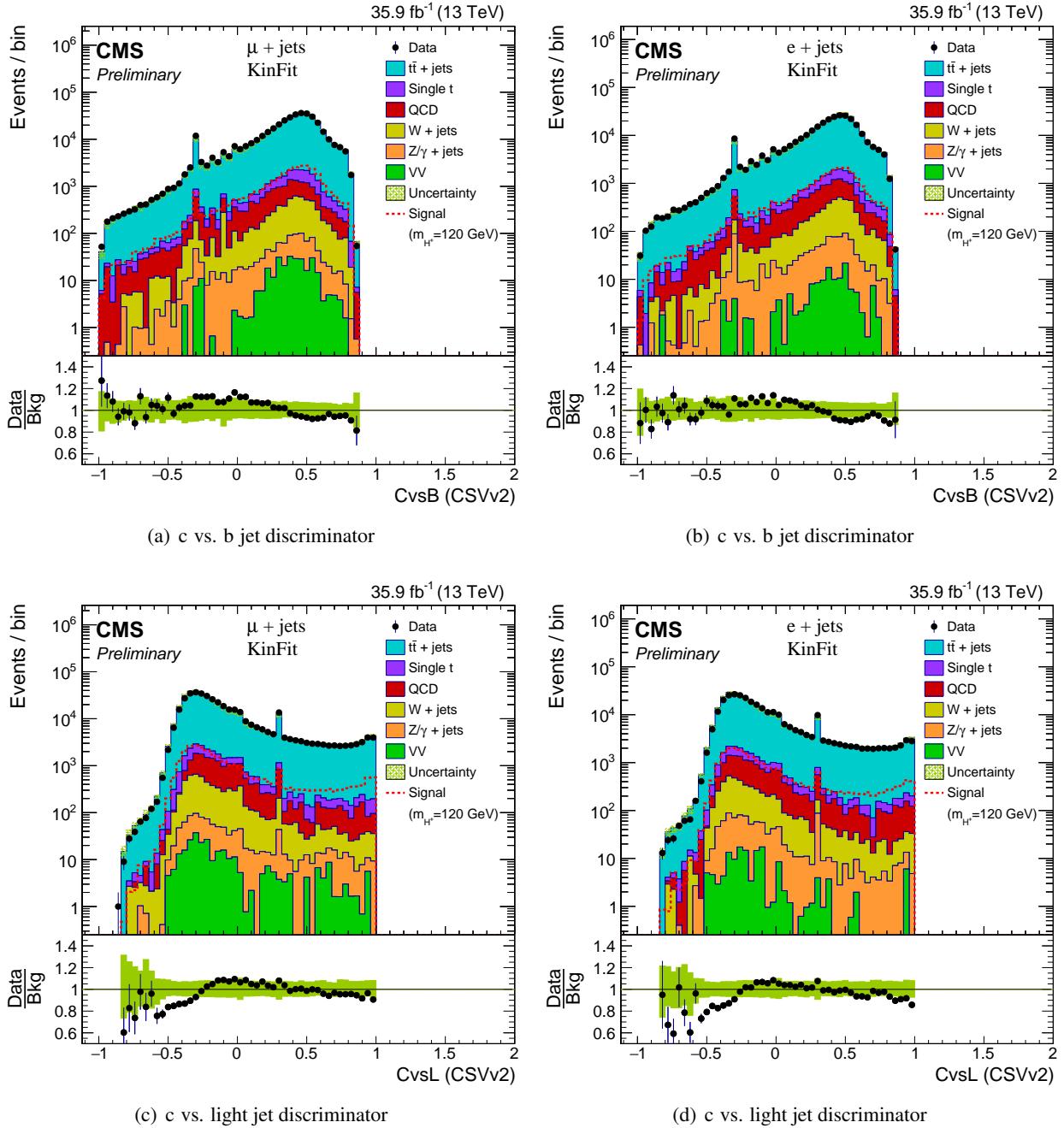


Figure 8.6.: The c discriminator distribution of two non b jets after kinematic fitting selection for  $\mu + \text{jets}$  and  $e + \text{jets}$  channel.



## 9 | QCD Multijet Estimation

In the signal region, events consist of 1 lepton, at least 4 jets, and  $E_T^{\text{miss}}$ . The QCD multijet events contain only jets at parton level. However, after event reconstruction, these events can still have leptons from the misidentifications of light jets, and  $E_T^{\text{miss}}$  due to the poor measurement of energy in the detector. The leptons can also be produced from the decay of bottom and charm hadrons. The  $E_T^{\text{miss}}$  can also arise due to the mismeasurement of hadronic activity inside the CMS detector. Having the same event topology as that of signal, the QCD multijet process is also considered as a background for this analysis. However, it contributes only around 2% of the total backgrounds.

The simulation of QCD multijet events are CPU time consuming due to which we can not generate them in large numbers due to computing limitations. Due to the lesser number of events, compared to the cross section, the statistical uncertainty is high. Also, the simulated events are not properly modeled for high jet multiplicities. After lepton trigger selection, the relative isolation plots from data and simulated events are shown in Figures 9.1(a), 9.1(b) for  $\mu + \text{jets}$  and  $e + \text{jets}$  channel, respectively. From these figures, it is obvious that the events with  $I_{\text{rel}}^\mu \geq 0.20$  and  $I_{\text{rel}}^e \geq 0.10$  are dominated by QCD multijet processes. However, there is a poor agreement between data and simulation for  $I_{\text{rel}}^\mu > 0.20$  and  $I_{\text{rel}}^e > 0.10$  for both  $\mu + \text{jets}$  and  $e + \text{jets}$  channel. In view of the above, the QCD multijet background is estimated from data to have a precise estimation.

To estimate QCD multijet events from data, a method similar to the ABCD method is followed. The ABCD region, as shown in Figure 9.3, is formed from two uncorrelated variables,  $E_T^{\text{miss}}$  and  $I_{\text{rel}}$  (relative isolation as defined in Equations (5.2) and (5.4)). The correlation between  $E_T^{\text{miss}}$  and  $I_{\text{rel}}$  are shown in Figures 9.2(a) (9.2(b)), 9.2(c) (9.2(d)) from data and simulated QCD multijet process for  $\mu + \text{jets}$  ( $e + \text{jets}$ ) channel. The comparison between data and simulation for  $E_T^{\text{miss}}$  and  $I_{\text{rel}}$  are shown in Figure 9.1 after 1-lepton selection, of Section 7.2, for  $\mu + \text{jets}$  and  $e + \text{jets}$  channel. From this figure, it can be seen that the QCD multijet process is dominated in high  $I_{\text{rel}}$  region. The signal region lies in high  $E_T^{\text{miss}}$  and low  $I_{\text{rel}}$  (region-A of Figure 9.3). We estimate QCD multijet in the region-A from other regions as described in the

Table 9.1.: Event yields from different regions, for  $\mu + \text{jets}$  channel after kinematic fit selection.

Process	Region-A (Iso, high $E_T^{\text{miss}}$ )	Region-B (Non iso, high $E_T^{\text{miss}}$ )	Region-C (Non iso, low $E_T^{\text{miss}}$ )	Region-D (Iso, low $E_T^{\text{miss}}$ )
Simulated QCD	$3008 \pm 1013$	$2323 \pm 645$	$1180 \pm 794$	$3260 \pm 2138$
$t\bar{t} + \text{jets}$	$191135 \pm 268$	$9405 \pm 60$	$955 \pm 19$	$23520 \pm 94$
Single t	$5312 \pm 40$	$248 \pm 9$	$20 \pm 2$	$701 \pm 15$
W + jets	$2795 \pm 123$	$141 \pm 18$	$7 \pm 4$	$442 \pm 34$
Z/ $\gamma$ + jets	$389 \pm 15$	$13 \pm 3$	$1 \pm 1$	$115 \pm 8$
VV	$147 \pm 20$	$6 \pm 5$	$0 \pm 27$	$35 \pm 10$
nonQCDBkg	$204409 \pm 383$	$9813 \pm 63$	$983 \pm 20$	$24812 \pm 102$
Data	$203207 \pm 451$	$12767 \pm 113$	$1949 \pm 44$	$26312 \pm 162$

following sections.

## 9.1. Determination of transfer scale factor

The phase space of  $E_T^{\text{miss}}$  and  $I_{\text{rel}}$  are divided into 4 regions as shown in Figure 9.3. The transfer scale factors are determined using the following formula

$$\text{SF}_{\text{qcd}} = \frac{N_{(\text{Data} - \text{other backgrounds})}^{\text{D}}}{N_{(\text{Data} - \text{other backgrounds})}^{\text{C}}} \quad (9.1)$$

The number of events in different regions after kinematic fit selection are shown in Table 9.1 for  $\mu + \text{jets}$  and in Table 9.2 for  $e + \text{jets}$  channel. The  $\text{SF}_{\text{qcd}}$  is calculated after each selection step, of Section 7.2, using Equation (9.1). The transfer scale factors are shown in Table 9.3 for both channels after various selection steps. From this table, it can be seen that the transfer scale factors are of the same order for all selection steps.

## 9.2. Estimation of QCD multijet in the signal region

All the other backgrounds are subtracted from data in region-B and the resulting distribution is multiplied with the transfer scale factor to estimate the contribution of data-driven QCD multijet events in region-A (signal region) i.e.

$$(\text{QCD})_{\text{A}} = \text{SF}_{\text{qcd}} \times (\text{Data} - \text{other backgrounds})_{\text{B}} \quad (9.2)$$

In the control plots of Figures 7.2, 7.3, 7.4 and 8.2, 8.3, 8.4, the data-driven QCD multijet events are shown.

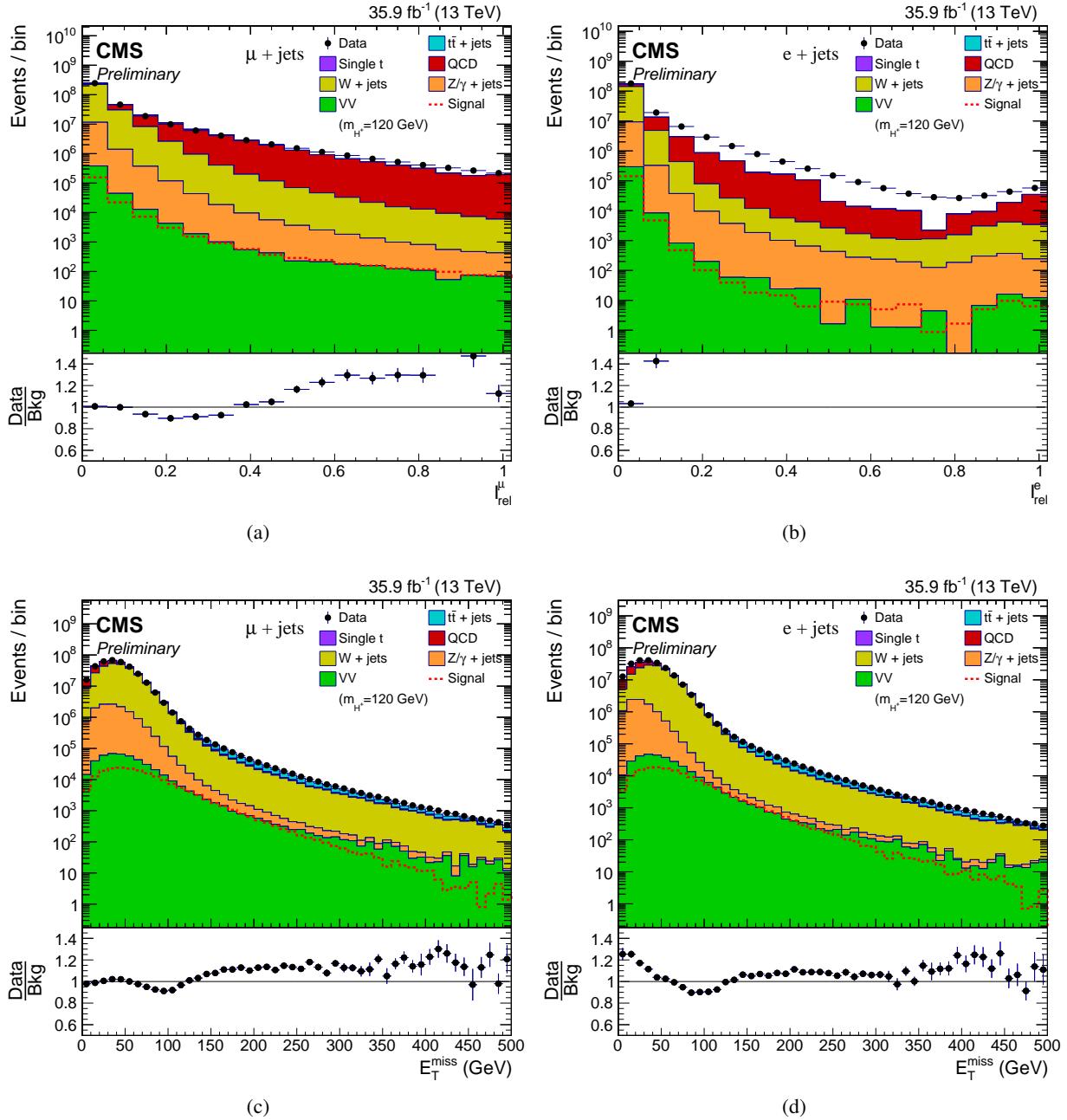


Figure 9.1.: Distribution of relative isolation and missing transverse energy after lepton trigger selection for  $\mu + \text{jets}$  and  $e + \text{jets}$  channel. The QCD multijet process is dominated in the high  $I_{\text{rel}}$  region.

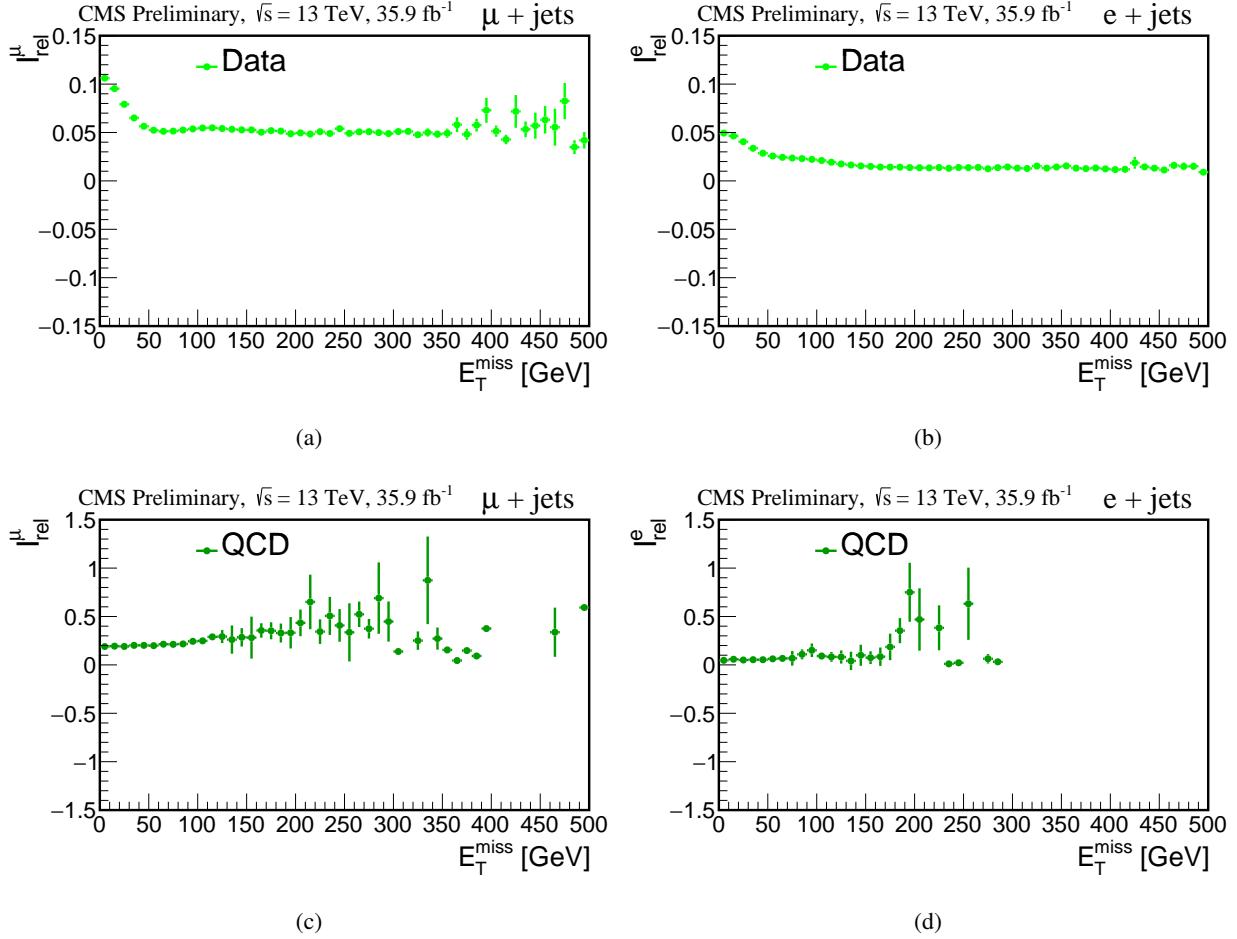


Figure 9.2.: Distribution of mean of  $I_{\text{rel}}$  in different  $E_{\text{T}}^{\text{miss}}$  bins after lepton trigger selection for  $\mu + \text{jets}$  and  $e + \text{jets}$  channel. There seems to be some correlation between  $E_{\text{T}}^{\text{miss}}$  and  $I_{\text{rel}}$  in the low  $E_{\text{T}}^{\text{miss}} < 50$  GeV region for data as shown in Figures 9.2(a), 9.2(b). However, from simulated QCD multijet events, the  $E_{\text{T}}^{\text{miss}}$  and  $I_{\text{rel}}$  are uncorrelated, as shown in Figures 9.2(c), 9.2(d).

Table 9.2.: Event yields from different regions, for  $e + \text{jets}$  channel after kinematic fit selection.

Process	Region-A (Iso, high $E_{\text{T}}^{\text{miss}}$ )	Region-B (Non iso, high $E_{\text{T}}^{\text{miss}}$ )	Region-C (Non iso, low $E_{\text{T}}^{\text{miss}}$ )	Region-D (Iso, low $E_{\text{T}}^{\text{miss}}$ )
Simulated QCD	$9678 \pm 6556$	$0 \pm 53$	$48 \pm 48$	$751 \pm 358$
$t\bar{t} + \text{jets}$	$141807 \pm 228$	$3899 \pm 38$	$421 \pm 12$	$18575 \pm 83$
Single $t$	$3989 \pm 35$	$97 \pm 5$	$9 \pm 2$	$542 \pm 13$
$W + \text{jets}$	$2032 \pm 71$	$42 \pm 10$	$0 \pm 18$	$336 \pm 28$
$Z/\gamma + \text{jets}$	$433 \pm 15$	$9 \pm 2$	$4 \pm 1$	$184 \pm 9$
VV	$70 \pm 13$	$0 \pm 53$	$0 \pm 18$	$28 \pm 7$
nonQCDBkg	$153308 \pm 315$	$4051 \pm 39$	$428 \pm 13$	$19394 \pm 89$
Data	$148499 \pm 385$	$6995 \pm 84$	$1271 \pm 36$	$20740 \pm 144$

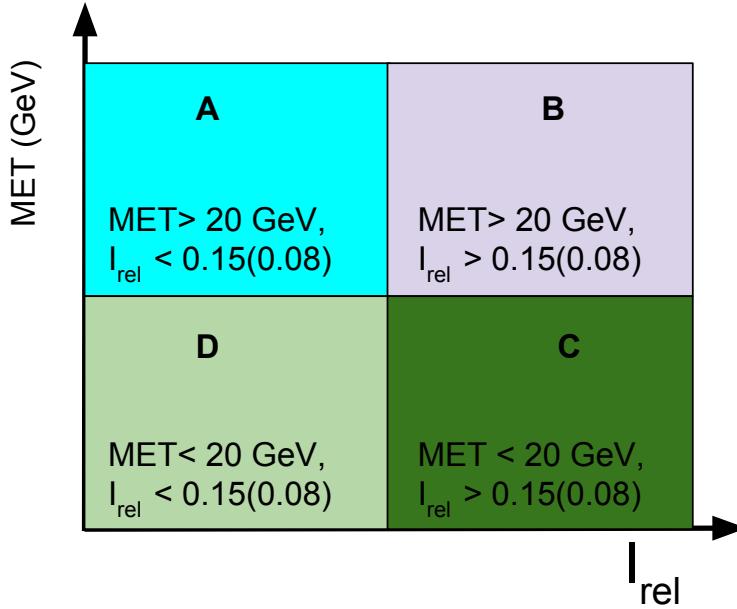


Figure 9.3.: Different regions between  $I_{\text{rel}}$  and MET (missing transverse energy also known as  $E_{\text{T}}^{\text{miss}}$ ) to estimate QCD multijet events from data. For  $\mu + \text{jets}$  channel, the isolation region corresponds to  $I_{\text{rel}}^{\mu} < 0.15$  whereas the non-isolation corresponds to  $0.15 < I_{\text{rel}}^{\mu} < 0.40$ . For  $e + \text{jets}$  channel, the isolation region corresponds to  $I_{\text{rel}}^e < 0.0821(0.0695)$  whereas the non-isolation region corresponds to  $0.0821 < I_{\text{rel}}^e < 0.30$  ( $0.0695 < I_{\text{rel}}^e < 0.30$ ) for electron found in barrel (endcap) regions.

Cuts	$SF_{\text{qcd}}$ ( $\mu + \text{jets}$ )	$SF_{\text{qcd}}$ ( $e + \text{jets}$ )
$N_{\text{lepton}} = 1$	$1.863 \pm 0.005$	$1.911 \pm 0.004$
$N_{\text{jets}} \geq 4$	$1.871 \pm 0.011$	$1.831 \pm 0.008$
$N_{\text{b jets}} \geq 2$	$2.567 \pm 0.141$	$1.854 \pm 0.114$
Kinematic fit selection	$1.554 \pm 0.213$	$1.282 \pm 0.210$
$N_{\text{c jets}} \geq 1$	$1.436 \pm 0.214$	$1.181 \pm 0.212$
Exclusive loose	$1.141 \pm 0.278$	$0.907 \pm 0.265$
Exclusive medium	$1.238 \pm 0.396$	$1.098 \pm 0.404$
Exclusive tight	$1.177 \pm 0.728$	$1.126 \pm 0.730$

Table 9.3.: The QCD multijet transfer scale factors ( $SF_{\text{qcd}}$ ) for  $\mu + \text{jets}$  and  $e + \text{jets}$  channel after various selection steps. The categorization of events based on exclusive loose, medium, and tight c quark tagging is described in Section 10.3.

### 9.3. Comparison of data-driven QCD multijet shapes

The validity of ABCD method relies on the fact that the data-driven QCD multijet shape should be similar in all the four regions. For sake of comparison, we compare shapes of few variables after b jet and kinematic fit selection steps. In the low  $E_T^{\text{miss}}$  region, the shape of  $p_T$ ,  $\eta$  of jets and lepton from the isolated and anti-isolated region are shown in Figures 9.4, 9.5. From these figures, it can be seen that the data-driven shape in isolated and anti-isolated matches quite well for these distributions. In few bins, the number of events in other background is greater than that in data. Therefore the number of data-driven QCD multijet events ( $n_{\text{Data}} - \text{other background}$ ) becomes negative. We set bin content of such bin to zero keeping the statistical uncertainty as it is.

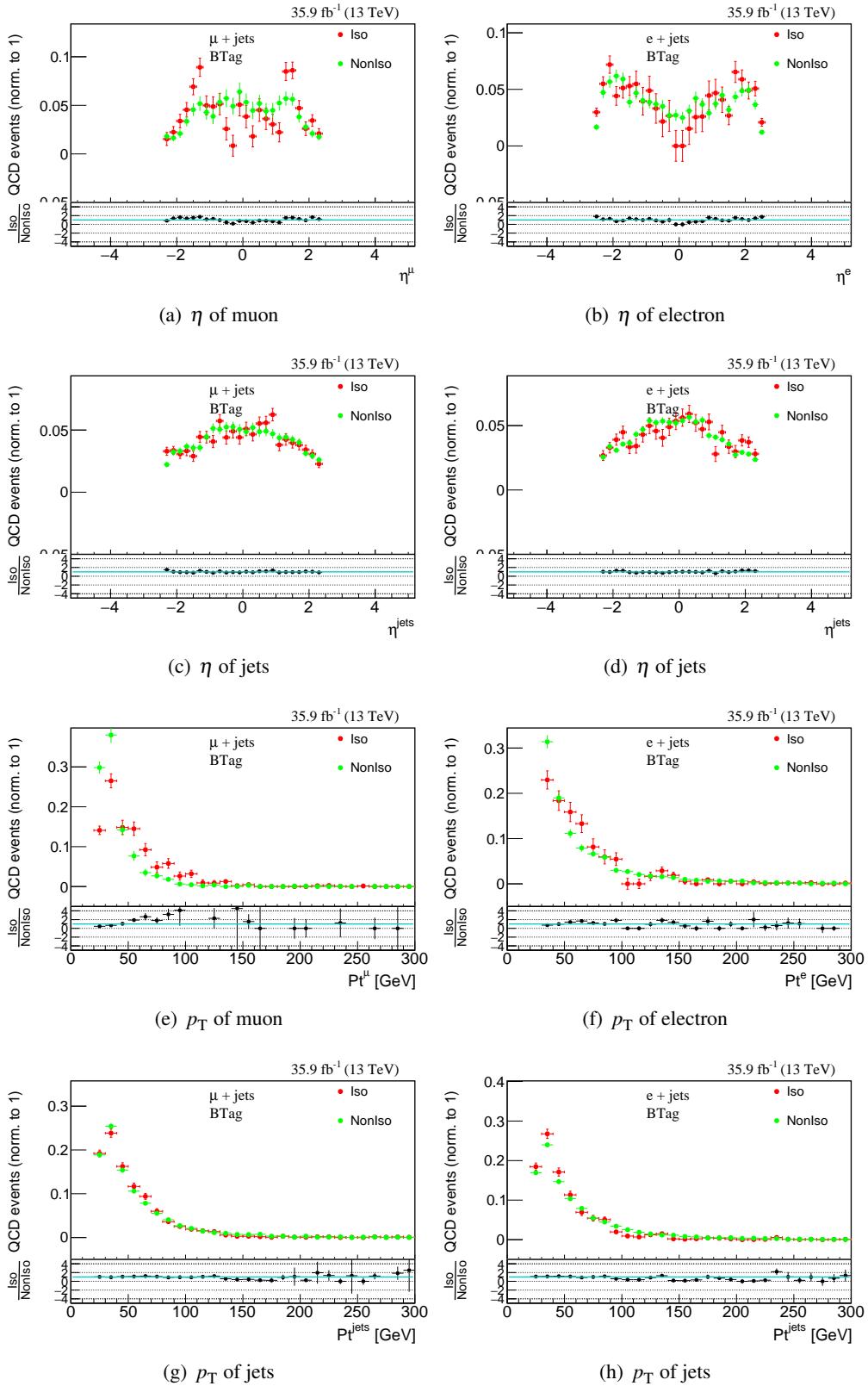


Figure 9.4.: Comparison of data-driven QCD multijet shapes in low  $E_T^{\text{miss}}$  region ( $< 20$  GeV), from the isolated and anti-isolated region with reconstructed jets after b jet selection for  $\mu + \text{jets}$  and  $e + \text{jets}$  channel.

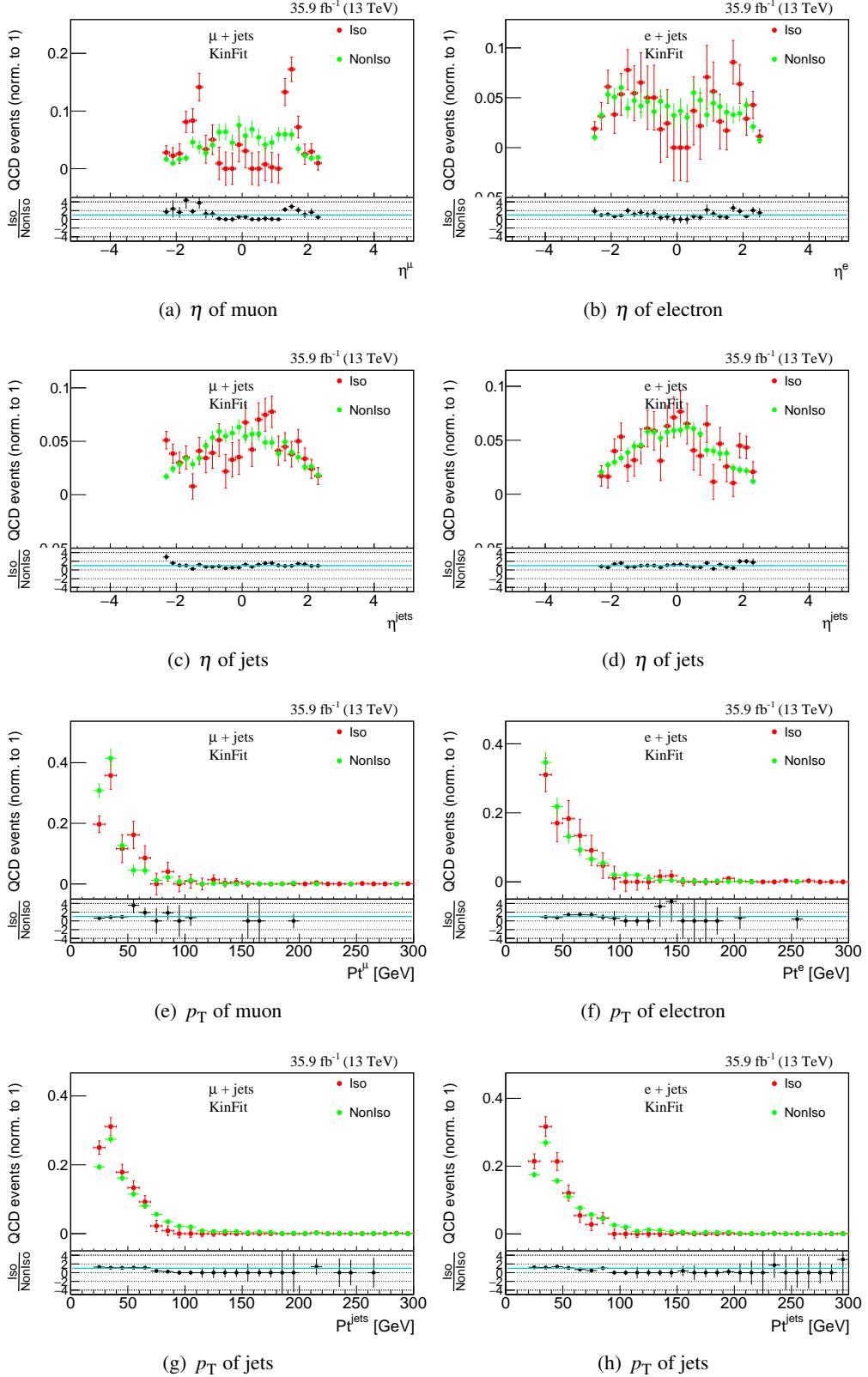


Figure 9.5.: Comparison of data-driven QCD multijet shapes in low  $E_T^{\text{miss}}$  region ( $< 20$  GeV), from the isolated and anti-isolated region with kinematic fitted objects after kinematic fit selection for  $\mu + \text{jets}$  and  $e + \text{jets}$  channel.

# 10 | Mass Reconstruction

The charged Higgs decays to c quark and s antiquark. Reconstruction and identification of jets coming from these quarks is very important to reconstruct the mass of charged Higgs ( $m_{H^+}$ ). We reconstruct ( $m_{H^+}$ ) before and after kinematic fit as follows:

- **$m_{H^+}$  reconstruction before kinematic fitting:** The jets are reconstructed and selected as described in Section 5.6. Events are required to have at least 4 jets. We use b tagging as described in Section 5.8 to tag two jets as b-tagged jet. The jets are sorted in descending order of b-discriminator values and the first two jets are considered as b jets. If an event contains exactly 4 jets then the other 2 non b jets are used to evaluate  $m_{jj}$ . However, if an event contains more than 4 jets, then the non b jets are sorted in descending order of  $p_T$  and two highest  $p_T$  non b jets are used to evaluate  $m_{jj}$ . The invariant mass ( $m_{jj}$ ) distribution after b jet selection as described in Section 7.2 is shown in Figures 8.1(a), 8.1(b) for data and background, and in Figures 10.1(a), 10.1(b) for all charged Higgs signal samples for  $\mu +$  jets and  $e +$  jets channel respectively.
- **$m_{H^+}$  reconstruction after kinematic fitting:** The kinematic fit is performed on the reconstructed jets of Section 5.6, as described in Chapter 8, in the semi-leptonic decay mode of  $t\bar{t}$  where W boson from one of the t quarks decays leptonically and W boson from the other t quark decays hadronically. The  $m_{jj}$  distribution from two light jets after kinematic fit selection is shown in Figures 8.1(c), 8.1(d) for data and background, and in Figures 10.1(c), 10.1(d) for all charged Higgs signal samples for  $\mu +$  jets and  $e +$  jets channel respectively.

To improve the exclusion limit on the  $\mathcal{B}(t \rightarrow H^+ b)$ , the  $m_{jj}$  from different event categories have been used for limit computation.

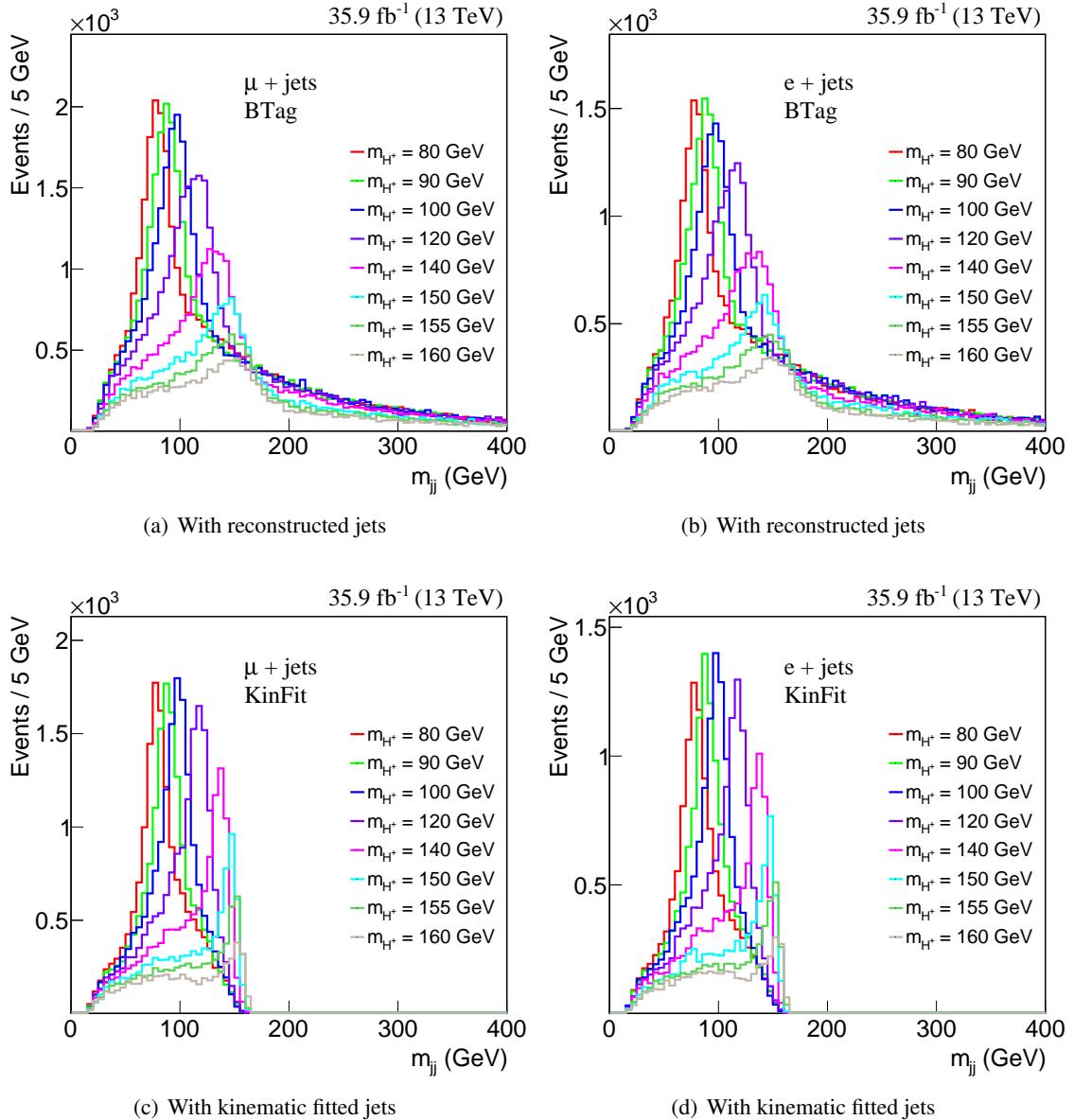


Figure 10.1.:  $m_{jj}$  distributions of two non b, highest  $p_T$  jets from charged Higgs signal samples ( $m_{H^+} = 80, 90, 100, 120, 140, 150, 155, 160$  GeV) for  $\mu +$  jets and  $e +$  jets channel. The  $m_{jj}$  distributions of Figures 10.1(a), 10.1(b) are evaluated using reconstructed jets after b jet selection. Whereas  $m_{jj}$  distributions of Figures 10.1(c), 10.1(d) are evaluated using kinematic fitted jets after kinematic fit selection.

Table 10.1.: Event yield for inclusive category after kinematic fit selection. The statistical uncertainty in the total background corresponds to the quadratically added uncertainties from individual backgrounds. However, a systematic uncertainty correlated among each background is linearly added for the total background. And each uncorrelated systematic uncertainty for the total background is quadratically added.

<b>Process</b>	$N_{\text{events}} \pm \text{stat} \pm \text{sys}$	$N_{\text{events}} \pm \text{stat} \pm \text{sys}$
	$\mu + \text{jets}$	$e + \text{jets}$
$m_{H^+} = 80 \text{ GeV}$	$15578 \pm 110 \pm 1101$	$11736 \pm 94 \pm 844$
$m_{H^+} = 90 \text{ GeV}$	$15721 \pm 109 \pm 1202$	$12076 \pm 95 \pm 899$
$m_{H^+} = 100 \text{ GeV}$	$16321 \pm 111 \pm 1191$	$12206 \pm 95 \pm 826$
$m_{H^+} = 120 \text{ GeV}$	$15719 \pm 109 \pm 1108$	$11805 \pm 93 \pm 830$
$m_{H^+} = 140 \text{ GeV}$	$12461 \pm 97 \pm 996$	$9494 \pm 84 \pm 738$
$m_{H^+} = 150 \text{ GeV}$	$8909 \pm 82 \pm 776$	$6902 \pm 72 \pm 630$
$m_{H^+} = 155 \text{ GeV}$	$7014 \pm 74 \pm 713$	$5398 \pm 64 \pm 587$
$m_{H^+} = 160 \text{ GeV}$	$5383 \pm 64 \pm 567$	$4184 \pm 56 \pm 451$
SM $t\bar{t} + \text{jets}$	$191135 \pm 268 \pm 14066$	$141807 \pm 228 \pm 10492$
Single t	$5312 \pm 40 \pm 506$	$3989 \pm 35 \pm 405$
QCD multijet	$4631 \pm 240$	$3780 \pm 179$
$W + \text{jets}$	$2795 \pm 123 \pm 575$	$2032 \pm 71 \pm 289$
$Z/\gamma + \text{jets}$	$389 \pm 15 \pm 76$	$433 \pm 15 \pm 76$
VV	$147 \pm 20 \pm 33$	$70 \pm 13 \pm 14$
All background	$204409 \pm 383 \pm 15153$	$152111 \pm 301 \pm 11241$
Data	$203207 \pm 451$	$148499 \pm 385$

## 10.1. Inclusive event category without charm tagging

The  $m_{jj}$  distribution without further categorizing events based on charm tagging was used at 8 TeV [63] for limit computation due to the unavailability of dedicated charm taggers. For sake of comparison, at 13 TeV also the exclusion limit is computed using  $m_{jj}$  after kinematic fit selection, without the charm tagging. The event yields are shown in Table 10.1. Corresponding  $m_{jj}$  distributions are shown in Figures 8.1(c), 8.1(d), 10.1(c), 10.1(d). The exclusion limit for inclusive event category is computed in Section 12.3.

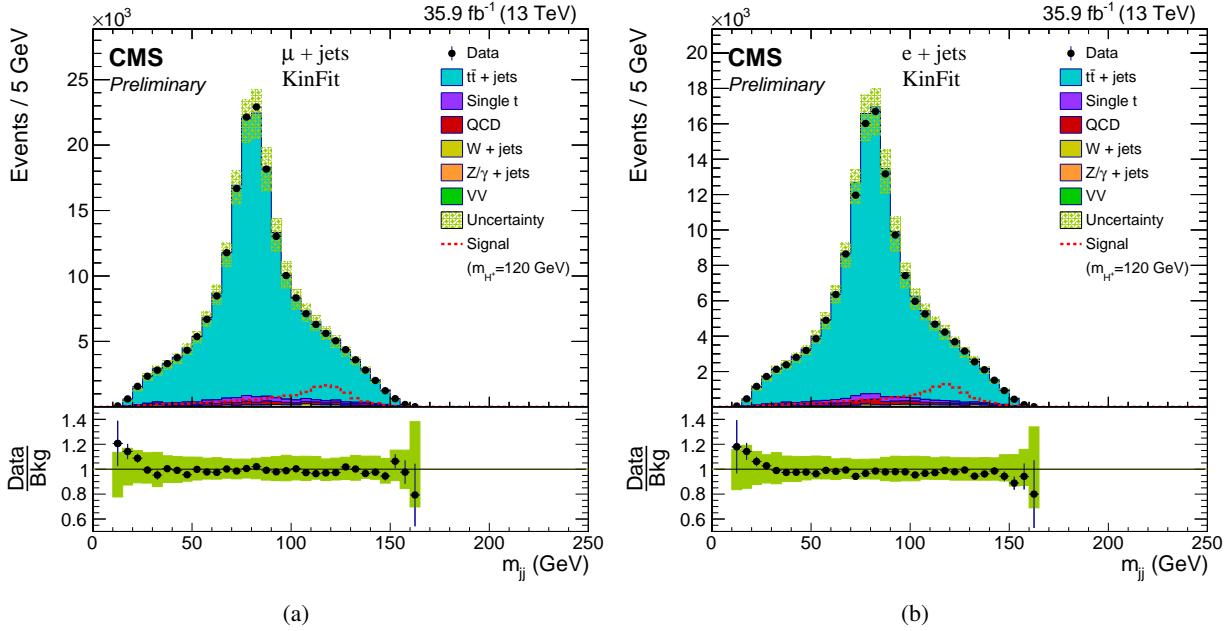


Figure 10.2.: Distribution of  $m_{jj}$  for inclusive loose charm tagging as described in Section 10.2 for  $\mu + \text{jets}$  and  $e + \text{jets}$  channel.

## 10.2. Inclusive event category with loose charm tagging

The charm taggers as discussed in Section 8.8 are used to tag a jet as c jet. After kinematic fit selection, one of the jets forming  $m_{jj}$  is required to pass the c tagging requirement. The event yields for loose charm working points after applying c tag event weights are shown in Table 10.2. The  $m_{jj}$  distributions with inclusive loose charm tagging are shown in Figure 10.2. Exclusion limits using  $m_{jj}$  after applying loose charm tagging is computed in Section 12.4.

### 10.3. Exclusive event categories based on charm tagging

The signal significance is different in different charm working points as can be seen from Figure 10.3 and 10.4. This property is exploited in improving the exclusion limits. Events are exclusively divided into the loose, medium, and tight categories as shown in Figure 10.5, based on whether one of the jets forming  $m_{jj}$  pass loose but not medium, medium but not tight, and tight working points of the c tagging requirements, respectively. The QCD multijet events are estimated in each category following the procedure given in Chapter 9. The resulting  $m_{jj}$  from various categories are combined to compute the limit.

The  $m_{jj}$  distribution from exclusive charm categories is shown in Figure 10.6. The corresponding event

Table 10.2.: Event yield for inclusive loose charm category after kinematic fit selection. The statistical uncertainty in the total background corresponds to the quadratically added uncertainties from individual backgrounds. However, a systematic uncertainty correlated among each background is linearly added for the total background. And each uncorrelated systematic uncertainty for the total background is quadratically added.

<b>Process</b>	$N_{\text{events}} \pm \text{stat} \pm \text{sys}$	$N_{\text{events}} \pm \text{stat} \pm \text{sys}$
	$\mu + \text{jets}$	$e + \text{jets}$
$m_{H^+} = 80 \text{ GeV}$	$15614 \pm 111 \pm 1405$	$11757 \pm 95 \pm 1054$
$m_{H^+} = 90 \text{ GeV}$	$15733 \pm 110 \pm 1484$	$12076 \pm 96 \pm 1115$
$m_{H^+} = 100 \text{ GeV}$	$16372 \pm 112 \pm 1488$	$12231 \pm 96 \pm 1081$
$m_{H^+} = 120 \text{ GeV}$	$15785 \pm 110 \pm 1447$	$11835 \pm 94 \pm 1084$
$m_{H^+} = 140 \text{ GeV}$	$12521 \pm 98 \pm 1281$	$9546 \pm 85 \pm 940$
$m_{H^+} = 150 \text{ GeV}$	$8956 \pm 83 \pm 978$	$6937 \pm 72 \pm 781$
$m_{H^+} = 155 \text{ GeV}$	$7057 \pm 75 \pm 907$	$5435 \pm 65 \pm 698$
$m_{H^+} = 160 \text{ GeV}$	$5408 \pm 65 \pm 686$	$4205 \pm 57 \pm 539$
SM $t\bar{t} + \text{jets}$	$190227 \pm 269 \pm 17385$	$141111 \pm 229 \pm 12842$
Single t	$5272 \pm 40 \pm 589$	$3954 \pm 35 \pm 464$
QCD multijet	$4185 \pm 226$	$3426 \pm 172$
$W + \text{jets}$	$2761 \pm 121 \pm 631$	$1991 \pm 70 \pm 330$
$Z/\gamma + \text{jets}$	$378 \pm 15 \pm 82$	$418 \pm 15 \pm 80$
VV	$147 \pm 20 \pm 35$	$70 \pm 13 \pm 15$
All background	$202970 \pm 375 \pm 18612$	$150970 \pm 297 \pm 13700$
Data	$201360 \pm 449$	$147210 \pm 384$

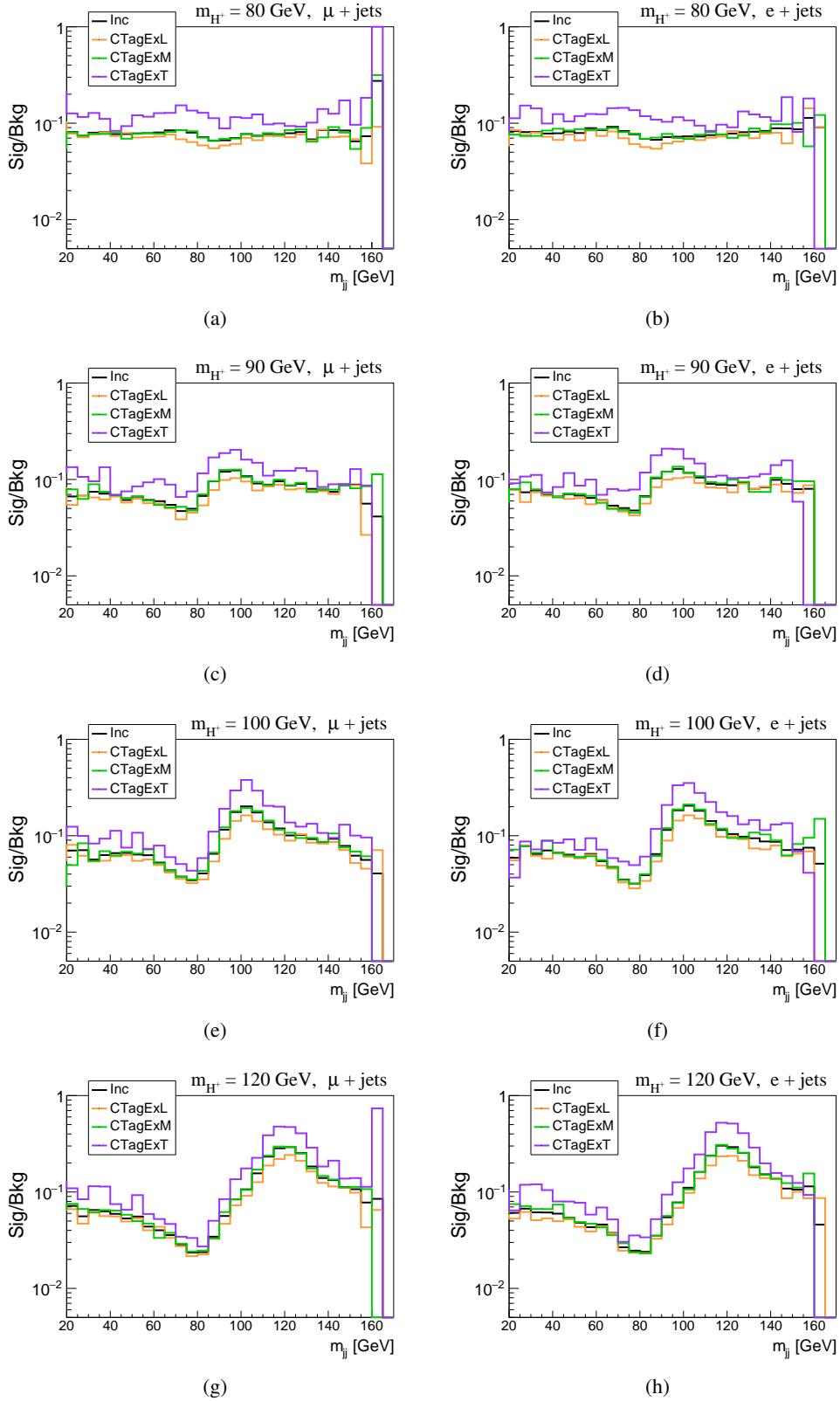
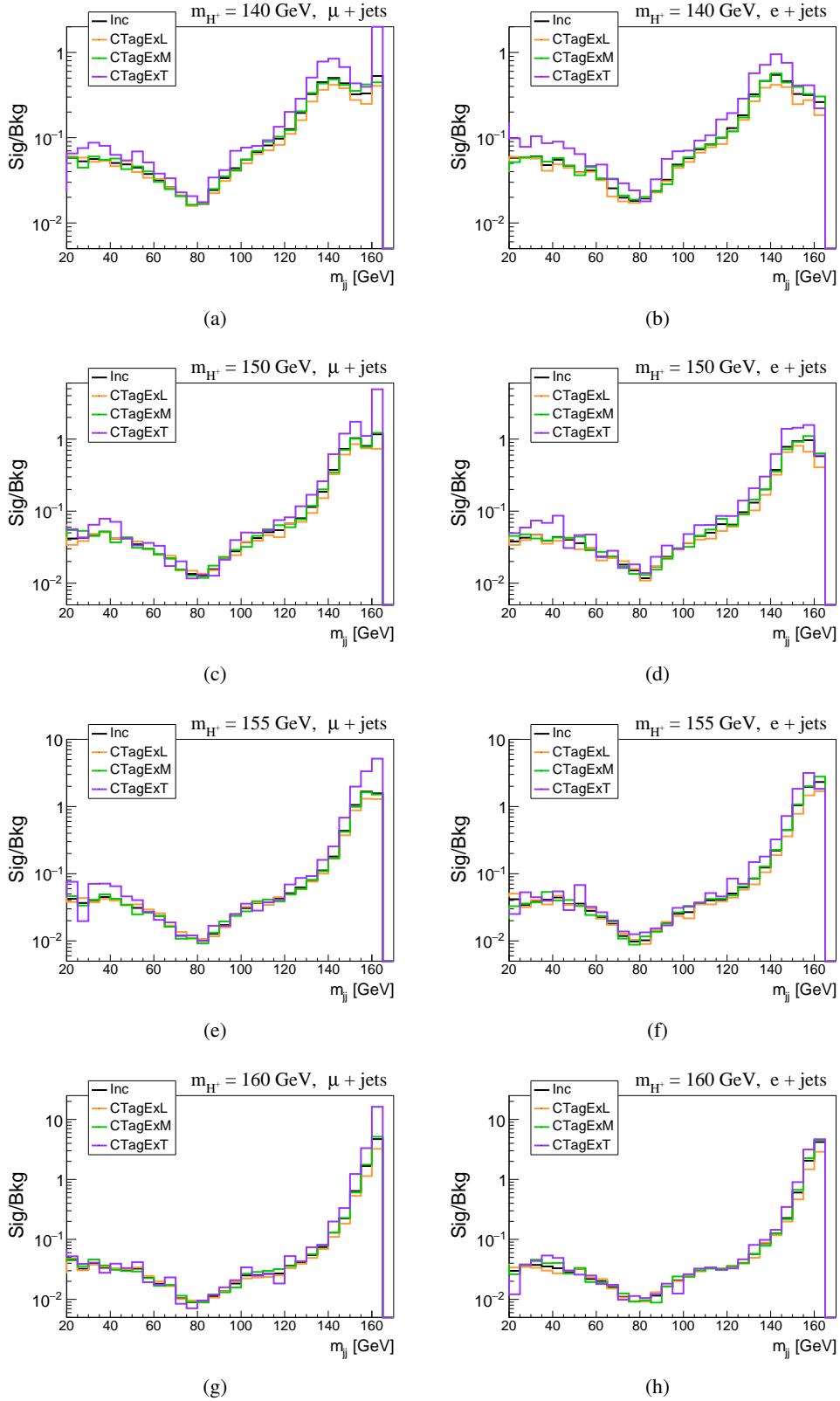


Figure 10.3.: Signal to background ratio for  $m_{H^+} = 80, 90, 100, 120 \text{ GeV}$ .

Figure 10.4.: Signal to background ratio for  $m_{H^+} = 140, 150, 155, 160 \text{ GeV}$ .

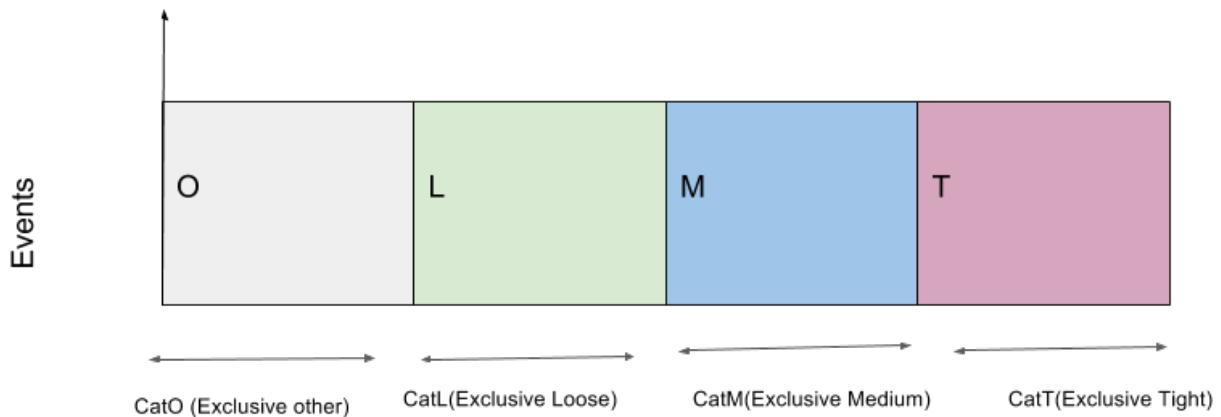


Figure 10.5.: Exclusive event categories based on different charm tagging. The events falling in the exclusive “other” category are not considered in this analysis. However, fraction of events falling in this category is very small.

yields are shown in Table 10.3. The signal to background ratio is shown in Figures 10.3, 10.4 for different masses of charged Higgs for both channels. From these figures, one can see that the signal to background ratio is different in different categories. The  $m_{jj}$  distributions from these three categories are combined to compute limits as described in Section 12.5.

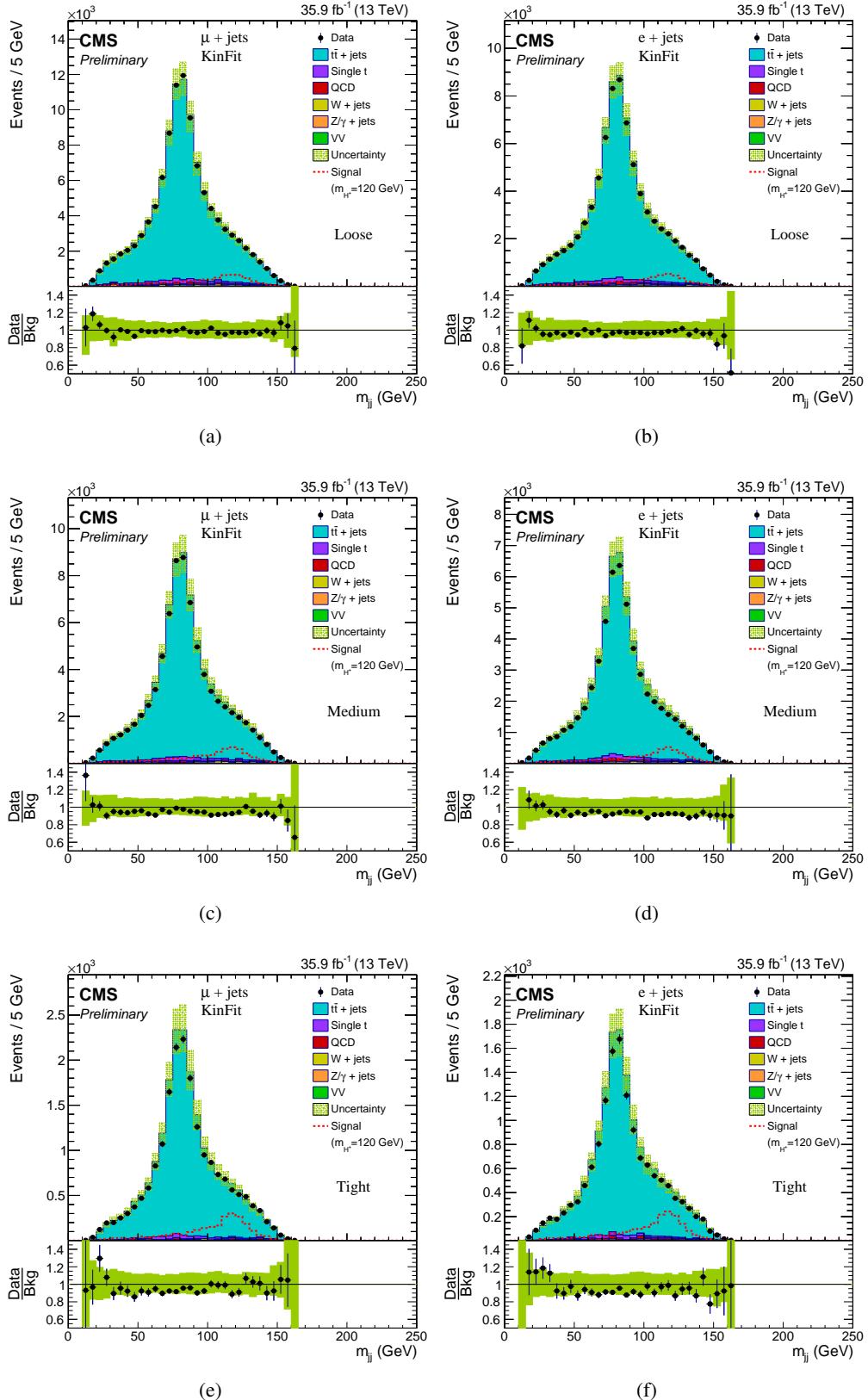


Figure 10.6.: Distribution of  $m_{jj}$  from exclusive charm categories as described in Section 10.3 for  $\mu + \text{jets}$  and  $e + \text{jets}$  channel. The signal significance is different across different exclusive categories.

Table 10.3.: Event yield from exclusive charm categories after kinematic fit selection. The statistical uncertainty in the total background corresponds to the quadratically added uncertainties from individual backgrounds. However, a systematic uncertainty correlated among each background is linearly added for the total background. And each uncorrelated systematic uncertainty for the total background is quadratically added.

Process	Exclusive loose category		Exclusive medium category		Exclusive tight category	
	$N_{\text{events}} \pm \text{stat} \pm \text{sys}$ $\mu + \text{jets}$	$N_{\text{events}} \pm \text{stat} \pm \text{sys}$ $e + \text{jets}$	$N_{\text{events}} \pm \text{stat} \pm \text{sys}$ $\mu + \text{jets}$	$N_{\text{events}} \pm \text{stat} \pm \text{sys}$ $e + \text{jets}$	$N_{\text{events}} \pm \text{stat} \pm \text{sys}$ $\mu + \text{jets}$	$N_{\text{events}} \pm \text{stat} \pm \text{sys}$ $e + \text{jets}$
$m_{H^+} = 80 \text{ GeV}$	$7171 \pm 75 \pm 678$	$5375 \pm 64 \pm 499$	$6137 \pm 71 \pm 551$	$4683 \pm 61 \pm 435$	$2471 \pm 43 \pm 278$	$1826 \pm 37 \pm 203$
$m_{H^+} = 90 \text{ GeV}$	$7171 \pm 74 \pm 718$	$5556 \pm 64 \pm 535$	$6300 \pm 72 \pm 602$	$4818 \pm 62 \pm 468$	$2442 \pm 42 \pm 274$	$1843 \pm 37 \pm 208$
$m_{H^+} = 100 \text{ GeV}$	$7382 \pm 75 \pm 718$	$5480 \pm 64 \pm 519$	$6591 \pm 73 \pm 638$	$4911 \pm 62 \pm 413$	$2570 \pm 43 \pm 290$	$1967 \pm 38 \pm 218$
$m_{H^+} = 120 \text{ GeV}$	$7073 \pm 73 \pm 689$	$5291 \pm 63 \pm 524$	$6378 \pm 72 \pm 602$	$4732 \pm 61 \pm 410$	$2466 \pm 42 \pm 261$	$1923 \pm 37 \pm 213$
$m_{H^+} = 140 \text{ GeV}$	$5716 \pm 66 \pm 625$	$4338 \pm 57 \pm 439$	$5043 \pm 64 \pm 507$	$3816 \pm 55 \pm 401$	$1869 \pm 37 \pm 208$	$1472 \pm 32 \pm 167$
$m_{H^+} = 150 \text{ GeV}$	$4218 \pm 57 \pm 477$	$3215 \pm 49 \pm 363$	$3584 \pm 54 \pm 403$	$2774 \pm 47 \pm 317$	$1233 \pm 30 \pm 154$	$1015 \pm 27 \pm 135$
$m_{H^+} = 155 \text{ GeV}$	$3447 \pm 52 \pm 416$	$2522 \pm 44 \pm 333$	$2758 \pm 48 \pm 357$	$2205 \pm 43 \pm 293$	$935 \pm 27 \pm 150$	$762 \pm 24 \pm 98$
$m_{H^+} = 160 \text{ GeV}$	$2593 \pm 45 \pm 315$	$2053 \pm 39 \pm 249$	$2210 \pm 43 \pm 296$	$1699 \pm 37 \pm 244$	$682 \pm 22 \pm 105$	$506 \pm 19 \pm 65$
SM $t\bar{t} + \text{jets}$	$100295 \pm 195 \pm 9239$	$74479 \pm 166 \pm 6851$	$76226 \pm 178 \pm 7396$	$56556 \pm 152 \pm 5472$	$19728 \pm 86 \pm 2319$	$14528 \pm 73 \pm 1641$
Single t	$2861 \pm 29 \pm 326$	$2159 \pm 25 \pm 261$	$2126 \pm 27 \pm 248$	$1585 \pm 23 \pm 189$	$483 \pm 13 \pm 65$	$355 \pm 11 \pm 48$
QCD multijet	$1772 \pm 140$	$1402 \pm 108$	$1067 \pm 124$	$1077 \pm 109$	$202 \pm 56$	$296 \pm 54$
W + jets	$1609 \pm 109 \pm 369$	$1181 \pm 54 \pm 191$	$1131 \pm 57 \pm 282$	$792 \pm 46 \pm 136$	$144 \pm 19 \pm 43$	$110 \pm 16 \pm 34$
Z/ $\gamma$ + jets	$198 \pm 12 \pm 45$	$250 \pm 11 \pm 52$	$143 \pm 9 \pm 35$	$153 \pm 9 \pm 33$	$61 \pm 6 \pm 14$	$35 \pm 5 \pm 7$
VV	$70 \pm 15 \pm 10$	$52 \pm 11 \pm 12$	$68 \pm 14 \pm 27$	$16 \pm 6 \pm 9$	$14 \pm 6 \pm 7$	$4 \pm 4 \pm 1$
All background	$106805 \pm 266 \pm 9932$	$79523 \pm 207 \pm 7346$	$80761 \pm 226 \pm 7915$	$60179 \pm 194 \pm 5826$	$20632 \pm 106 \pm 2420$	$15329 \pm 93 \pm 1718$
Data	$105485 \pm 325$	$77244 \pm 278$	$76811 \pm 277$	$56051 \pm 237$	$19451 \pm 139$	$14179 \pm 119$

# 11 | Systematic Uncertainties

The event yields, as shown in Table 10.3, have two types of uncertainties: statistical and systematics. For  $N$  number of events, the absolute statistical uncertainty is  $\sqrt{N}$ . The systematic uncertainties arise due to improper calibration of the detector, the uncertainty in the efficiency of the reconstruction algorithm, uncertainty difference in theoretical prediction etc. The systematics uncertainties considered in this analysis are divided into two categories: experimental and theoretical uncertainties.

## 11.1. Experimental uncertainties

- **Luminosity:** The uncertainty in the integrated luminosity is 2.5% for 2016 data taking [134].
- **Pileup reweighting:** The pileup weights as discussed in Section 6.2 are calculated using minimum bias cross-section 69.2 mb. The uncertainty in the minimum bias cross-section affects the pileup distribution. The minimum bias cross-section is varied up and down by 4.7% and corresponding pileup weights are applied to simulated events. The effect of pileup uncertainty on the  $m_{jj}$  shape is shown in Figures 11.3(a), 11.3(b) and Figures 11.7(c), 11.7(d) for charged Higgs and  $t\bar{t} + \text{jets}$  sample, respectively for both channels.
- **Lepton trigger, tracking, identification and isolation:** The lepton scale factors as discussed in Section 6.3 are applied to the simulated samples. A conservative 3% uncertainty is considered on combined lepton scale factors for both channels.
- **Jet energy,  $E_T^{\text{miss}}$  correction:** The  $p_T$  of jet in the simulation are corrected using JES and JER as discussed in Section 6.4. The jet correction is propagated to correct  $E_T^{\text{miss}}$ . The  $m_{jj}$  shape are also affected by the uncertainties in the JES, JER as shown in Figures 11.3(c), 11.3(d), 11.3(e), and 11.3(f) for charged Higgs signal ( $m_{H^+} = 120$  GeV) and in Figures 11.5(a), 11.5(b), 11.5(c), and 11.5(d) for  $t\bar{t} + \text{jets}$  background for both channels.

	b-tagged	c-tagged
c-quark	(a)	(b)
b-quark	(c)	(d)
l-quark	(e)	(f)

Figure 11.1.: Correlation of b and c tag scale factors. In the region (a), c quark is tagged as b jet; in region (b), c quark is tagged as c jet; in region (c), b quark is tagged as b jet; in region (d), b quark is tagged as c jet; in region (e), light quark is tagged as b jet; and in region (f), light quark is tagged as c jet.

- **b and c tagging uncertainty:** The b and c jets are present within the same event. The b and c tag scale factors (as described in Section 6.5 and Section 8.9) are correlated to each other as shown in Figure 11.1. The correlated scale factors are varied up/down simultaneously. Therefore, we have three systematic uncertainties from b and c tagging. Based on their grouping in Figure 11.1, we have the following three nuisance parameters:

- CMS\_eff\_bcInc3 (correlating (a), (c), (d))
- CMS\_eff\_bcInc2 (correlating (e), (f))
- CMS\_eff\_bcInc1 (from (b))

After  $\geq 1$  c jet selection, the effect of b/c tagging systematics on the shape of  $m_{jj}$  distribution are shown in Figures 11.4(a), 11.4(c), 11.4(e), and 11.4(b), 11.4(d), 11.4(f) for charged Higgs signal ( $m_{H^+} = 120$  GeV) and in Figures 11.5(e), 11.5(g), 11.6(a), and 11.5(f), 11.5(h), 11.6(b) for  $t\bar{t} + \text{jets}$  background for both channels.

- **Uncertainties for exclusive charm categories:** The c tag scale factors are officially not available for exclusive event categories. Therefore, conservatively, the inclusive scale factors are used for exclusive event categories. However, one extra systematic uncertainty is considered for exclusive categories. The mean of inclusive c tag weights (Equation (6.8)) is calculated based on whether a jet is satisfying one of the charm working points and not satisfying the other (e.g. a jet satisfies loose and medium but not

Table 11.1.: Inclusive loose c tag types. The acronym yLyMnT refers to the case where a jet satisfies loose and medium, but not tight working point.

Loose	Medium	Tight	Acronym
Y	Y	Y	yLyMyT
Y	Y	N	yLyMnT
Y	N	Y	yLnMyT
Y	N	N	yLnMnT

Table 11.2.: Inclusive medium c tag types.

Medium	Tight	Acronym
Y	Y	yMyT
Y	N	yMnT

tight working, as shown in Tables 11.1, 11.2). The following two nuisance parameters are considered for exclusive categories.

- CMS\_eff\_ExL: The ratio (max/min) of mean of inclusive c-tag weights from four combinations (yLyMyT, yLyMnT, yLnMyT, yLnMnT) is considered.
- CMS\_eff\_ExM: For exclusive medium charm category, the ratio (max/min) of event weights from two combinations (yMyT, yMnT) are considered.

The inclusive and exclusive event categories are same for tight charm working points, hence no extra systematic uncertainty is considered. The loose and medium c tag event weight corresponding to Tables 11.1, 11.2 are shown in Figure 11.2 for both channels, from  $t\bar{t} + \text{jets}$  sample. Similar event weights are found for simulated signal samples.

- **QCD multijet data-driven uncertainty:** The relative isolation of muon (electron) are conservatively shifted to 0.17 (0.11) and the corresponding change in the QCD multijet yield is determined. The percentage change in the yield is considered as systematic uncertainty in the QCD multijet estimation.
- **Bin-by-bin uncertainty:** In some of the bins of  $m_{jj}$  distributions, the statistics are small hence statistical uncertainty is large. In each bin of summed  $m_{jj}$  from all backgrounds, one shape nuisance parameter (as described in Section 12.2) is considered.

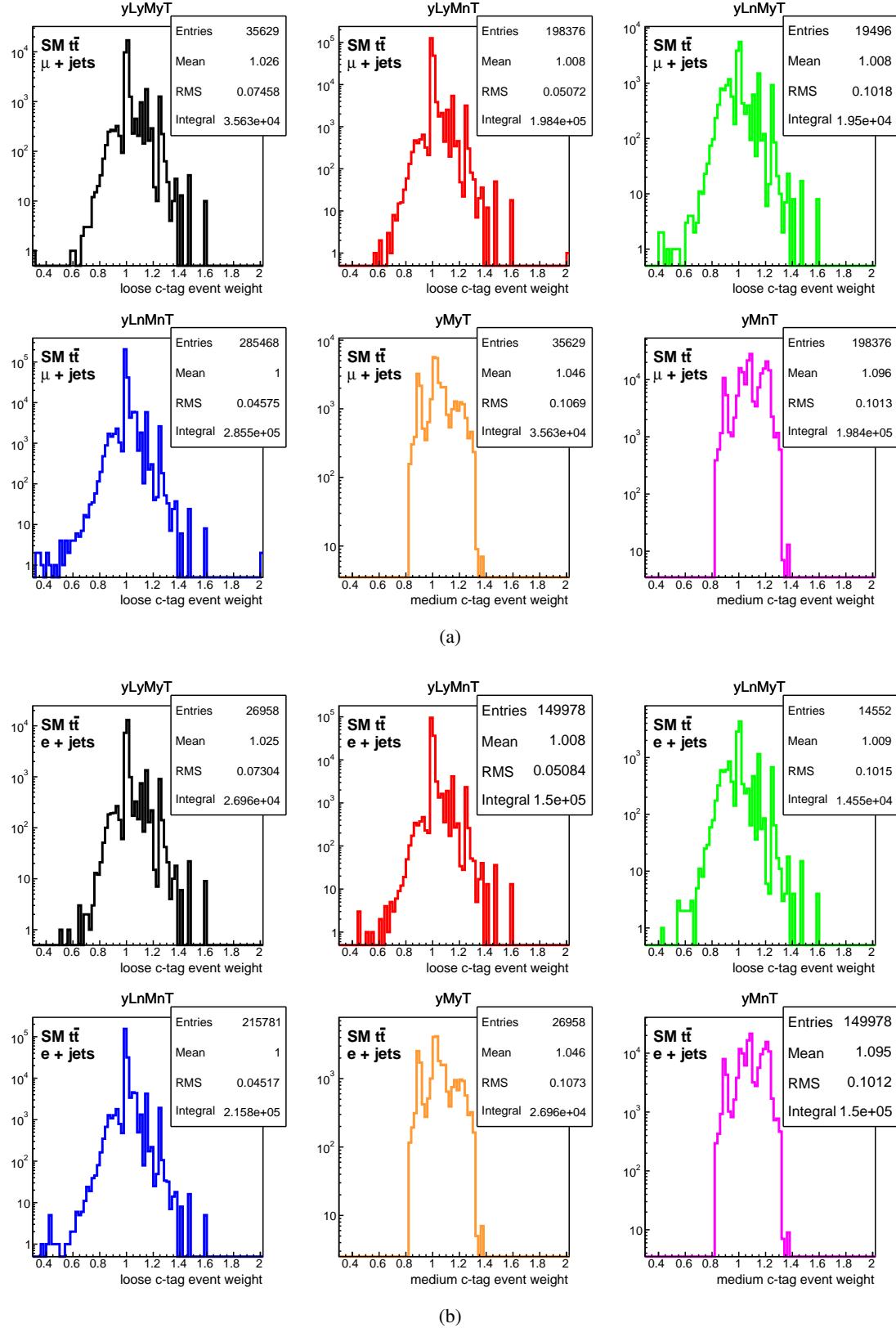


Figure 11.2.: The c tag event weight for different cases from  $t\bar{t} + \text{jets}$  sample for both channels. For exclusive loose charm category, the value of  $CMS_{eff\_ExL} = 1.026/0.9961$  ( $1.025/0.9960$ ) for  $\mu + \text{jets}$  ( $e + \text{jets}$ ) channel. For exclusive medium charm category, the value of  $CMS_{eff\_ExM} = 1.096/1.046$  ( $1.095/1.046$ ) for  $\mu + \text{jets}$  ( $e + \text{jets}$ ) channel.

## 11.2. Theoretical uncertainties

- **Normalization:** The uncertainty in the cross section of each simulated sample is considered.

- **t quark  $p_T$  re-weighting:** The t quark  $p_T$  weights are determined from following equation

$$w_{\text{top}} = \sqrt{\text{SF}(t) \times \text{SF}(\bar{t})}, \quad \text{where } \text{SF} \equiv \exp(0.0615 - 0.0005 \times p_T). \quad (11.1)$$

The values in the exponent are given in [135, 136]. The generator level  $p_T$  of t and  $\bar{t}$  quark is used to calculate the SF. The *particle data group* code “6” is used to identify t quark at parton level. The  $w_{\text{top}}$  is taken to be 1 for “down” and  $w_{\text{top}}^2$  for “up” systematic. The t quark  $p_T$  re-weighting is applied only on  $t\bar{t} + \text{jets}$  and charged Higgs samples shown in Table 5.2. The effect of t quark  $p_T$  re-weighting on the  $m_{jj}$  is shown in Figures 11.4(g), 11.4(h), for charged Higgs signal ( $m_{H^+} = 120$  GeV) and in Figures 11.7(a), 11.7(b) for  $t\bar{t} + \text{jets}$  background for both channels.

- **t quark mass:** The  $t\bar{t} + \text{jets}$  sample shown in Table 5.2 was generated with  $m_t = 172.5$  GeV. We used  $t\bar{t} + \text{jets}$  sample with  $m_t = 171.5$  GeV and  $m_t = 173.5$  GeV, as shown in Table 5.3, to observe the effect of t quark mass on the  $m_{jj}$ . The  $m_{jj}$  for different  $m_t$  are shown in Figures 11.6(c), 11.6(d) for  $t\bar{t} + \text{jets}$  background for both channels.
- **Parton shower matching:** In the POWHEG + PYTHIA generated samples, the next-to-leading order (NLO) matrix element parton shower matching is varied by  $h_{\text{damp}}$  parameter. The  $t\bar{t} + \text{jets}$  samples generated by varying  $h_{\text{damp}}$  “up” and “down”, as shown in Table 5.3 are used to see the effect of  $h_{\text{damp}}$  on the  $m_{jj}$  distribution as shown in Figures 11.6(g), 11.6(h) for  $t\bar{t} + \text{jets}$  background for both channels.
- **Renormalization and factorisation scale:** To see the effect of renormalization and factorisation scale, the “up” and “down”  $t\bar{t} + \text{jets}$  samples as shown in Table 5.3 are used. The effect of these scales on the  $m_{jj}$  are shown in Figures 11.6(e), 11.6(f) for  $t\bar{t} + \text{jets}$  background for both channels.

The QCD multijet events are estimated from data. Hence none of the systematics (except the data-driven uncertainty) are considered on the QCD multijet background. The relative systematic and statistical uncertainties for all the process, from all event categories, and for both channels are shown in Tables 11.3, 11.4, and 11.5.

Table 11.3.: Systematic and statistical uncertainties in %, from inclusive category after kinematic fit selection for  $\mu + \text{jets}$  ( $e + \text{jets}$ ) channel. The “—” indicates that the corresponding uncertainties are not considered for the given process.

Process	Luminosity	Pileup	Lepton	JES + $E_T^{\text{miss}}$	b & c tagging_1	b & c tagging_2	b & c tagging_3	Normalization	Statistical	t quark $p_T$
$m_{H^+} = 80 \text{ GeV}$	2.5 (2.5)	0.86 (0.83)	3.0 (3.0)	6.1 (6.2)	0.12 (0.089)	0.33 (0.3)	3.3 (3.3)	6.1 (6.1)	0.7 (0.8)	1.1 (1.3)
$m_{H^+} = 90 \text{ GeV}$	2.5 (2.5)	1 (0.93)	3.0 (3.0)	6.7 (6.4)	0.012 (0.075)	0.35 (0.43)	3.4 (3.2)	6.1 (6.1)	0.69 (0.79)	0.83 (1.7)
$m_{H^+} = 100 \text{ GeV}$	2.5 (2.5)	0.77 (0.9)	3.0 (3.0)	6.3 (5.7)	0.12 (0.14)	0.3 (0.21)	3.5 (3.2)	6.1 (6.1)	0.68 (0.78)	0.6 (1.3)
$m_{H^+} = 120 \text{ GeV}$	2.5 (2.5)	0.74 (1.3)	3.0 (3.0)	6.1 (5.9)	0.31 (0.058)	0.57 (0.5)	3.4 (3.4)	6.1 (6.1)	0.69 (0.79)	0.64 (1.2)
$m_{H^+} = 140 \text{ GeV}$	2.5 (2.5)	0.83 (1.1)	3.0 (3.0)	6.7 (6.5)	0.45 (0.031)	0.28 (0.56)	4 (3.6)	6.1 (6.1)	0.78 (0.88)	1.2 (1.9)
$m_{H^+} = 150 \text{ GeV}$	2.5 (2.5)	0.49 (0.87)	3.0 (3.0)	7.4 (7.6)	0.24 (0.017)	0.69 (0.75)	4 (3.9)	6.1 (6.1)	0.92 (1)	2 (3)
$m_{H^+} = 155 \text{ GeV}$	2.5 (2.5)	0.97 (0.57)	3.0 (3.0)	8.8 (9.4)	0.44 (0.24)	1.3 (0.72)	4.2 (4.1)	6.1 (6.1)	1.1 (1.2)	2.4 (3.3)
$m_{H^+} = 160 \text{ GeV}$	2.5 (2.5)	0.27 (0.81)	3.0 (3.0)	9 (9.4)	0.59 (0.23)	1.3 (1.1)	4.1 (3.9)	6.1 (6.1)	1.2 (1.3)	3.3 (3.4)
SM $t\bar{t} + \text{jets}$	2.5 (2.5)	0.74 (0.88)	3.0 (3.0)	6.6 (6.5)	0.18 (0.019)	0.43 (0.35)	3.2 (3.1)	6.1 (6.1)	0.14 (0.16)	0.68 (1.4)
Single t	2.5 (2.5)	0.44 (0.59)	3.0 (3.0)	8.8 (9.4)	0.39 (0.14)	0.8 (0.84)	3.5 (3.5)	5 (5)	0.76 (0.87)	—
W + jets	2.5 (2.5)	1 (0.81)	3.0 (3.0)	19 (13)	1.2 (0.079)	4.4 (3.5)	6.7 (4.6)	5 (5)	4.4 (3.5)	—
Z/ $\gamma$ + jets	2.5 (2.5)	1.4 (3)	3.0 (3.0)	18 (16)	1.7 (0.12)	4.3 (3.7)	5.3 (3.8)	4.5 (4.5)	3.9 (3.4)	—
VV	2.5 (2.5)	5 (6.5)	3.0 (3.0)	17 (17)	6.4 (3.3)	8.2 (3.7)	6 (5.8)	4 (4)	14 (19)	—
QCD multijet	—	—	—	—	—	—	—	13 (18)	5.2 (4.7)	—

Table 11.4.: Systematic and statistical uncertainties in %, from inclusive loose charm tagging category after kinematic fit selection for  $\mu + \text{jets}$  ( $e + \text{jets}$ ) channel. The “—” indicates that the corresponding uncertainties are not considered for the given process.

Process	Luminosity	Pileup	Lepton	JES + $E_T^{\text{miss}}$	b & c tagging_1	b & c tagging_2	b & c tagging_3	Normalization	Statistical	t quark $p_T$
$m_{H^+} = 80 \text{ GeV}$	2.5 (2.5)	0.88 (0.82)	3.0 (3.0)	6.1 (6)	2 (2)	4 (4)	4.7 (4.7)	6.1 (6.1)	0.71 (0.81)	1.1 (1.3)
$m_{H^+} = 90 \text{ GeV}$	2.5 (2.5)	1.1 (0.95)	3.0 (3.0)	6.7 (6.5)	1.6 (1.4)	4.1 (4.2)	4.8 (4.5)	6.1 (6.1)	0.7 (0.79)	0.84 (1.7)
$m_{H^+} = 100 \text{ GeV}$	2.5 (2.5)	0.77 (0.89)	3.0 (3.0)	6.3 (5.7)	1.8 (2.5)	3.9 (3.9)	4.9 (4.7)	6.1 (6.1)	0.68 (0.78)	0.59 (1.2)
$m_{H^+} = 120 \text{ GeV}$	2.5 (2.5)	0.72 (1.3)	3.0 (3.0)	5.9 (5.9)	1.6 (1.6)	4.3 (4.2)	5.2 (5.1)	6.1 (6.1)	0.7 (0.8)	0.65 (1.1)
$m_{H^+} = 140 \text{ GeV}$	2.5 (2.5)	0.88 (1.1)	3.0 (3.0)	6.7 (6.5)	2.1 (1.5)	4 (4.2)	6.1 (5.5)	6.1 (6.1)	0.79 (0.89)	1.1 (1.9)
$m_{H^+} = 150 \text{ GeV}$	2.5 (2.5)	0.5 (0.87)	3.0 (3.0)	7.4 (7.6)	2 (1.7)	4.5 (4.5)	6 (5.9)	6.1 (6.1)	0.93 (1)	2 (3)
$m_{H^+} = 155 \text{ GeV}$	2.5 (2.5)	1.1 (0.58)	3.0 (3.0)	8.8 (9.4)	2.6 (1.8)	5.3 (4.7)	6.8 (6.4)	6.1 (6.1)	1.1 (1.2)	2.4 (3.3)
$m_{H^+} = 160 \text{ GeV}$	2.5 (2.5)	0.37 (0.84)	3.0 (3.0)	9 (9.3)	1.3 (1.5)	5.2 (5)	6.3 (6.1)	6.1 (6.1)	1.2 (1.3)	3.3 (3.4)
SM $t\bar{t} + \text{jets}$	2.5 (2.5)	0.74 (0.89)	3.0 (3.0)	6.5 (6.5)	1.1 (0.9)	4.6 (4.5)	4.2 (4.1)	6.1 (6.1)	0.14 (0.16)	0.68 (1.4)
Single t	2.5 (2.5)	0.46 (0.62)	3.0 (3.0)	8.7 (9.3)	1 (0.78)	5.1 (5.3)	4.6 (4.7)	5 (5)	0.76 (0.88)	—
W + jets	2.5 (2.5)	0.94 (0.68)	3.0 (3.0)	19 (13)	1.7 (0.37)	9.3 (8.4)	7.6 (5.7)	5 (5)	4.4 (3.5)	—
Z/ $\gamma$ + jets	2.5 (2.5)	1.2 (2.8)	3.0 (3.0)	19 (16)	1.4 (0.31)	9 (8.2)	6.9 (5.6)	4.5 (4.5)	4 (3.6)	—
VV	2.5 (2.5)	4.9 (6.4)	3.0 (3.0)	17 (16)	5.9 (4.2)	11 (7.8)	7.1 (6.2)	4 (4)	14 (19)	—
QCD multijet	—	—	—	—	—	—	—	14 (21)	5.4 (5)	—

Table 11.5.: Systematic and statistical uncertainties in %, from exclusive charm tagging categories after kinematic fit selection for  $\mu + \text{jets}$  ( $e + \text{jets}$ ) channel. The “—” indicates that the corresponding uncertainties are not considered for the given process.

Category	Process	Luminosity	Pileup	Lepton	JES + JER + $E_T^{\text{miss}}$	b & c tagging-1	b & c tagging-2	b & c tagging-3	Normalization	Statistical	t quark $p_T$
Loose	$m_{H^+} = 80 \text{ GeV}$	2.5 (2.5)	0.96 (0.94)	3.0 (3.0)	6.8 (6.7)	1.6 (1.8)	4.5 (4.1)	4.4 (4.3)	6.1 (6.1)	1 (1.2)	0.86 (1.2)
	$m_{H^+} = 90 \text{ GeV}$	2.5 (2.5)	1.2 (0.88)	3.0 (3.0)	7.5 (7.1)	1.3 (1.2)	4.5 (4.5)	4.4 (4.3)	6.1 (6.1)	1 (1.2)	0.78 (1.4)
	$m_{H^+} = 100 \text{ GeV}$	2.5 (2.5)	0.81 (1.2)	3.0 (3.0)	7.4 (6.8)	1.5 (2.1)	4.2 (4.2)	4.4 (4.3)	6.1 (6.1)	1 (1.2)	0.49 (1)
	$m_{H^+} = 120 \text{ GeV}$	2.5 (2.5)	0.86 (1.4)	3.0 (3.0)	6.9 (7.3)	1.4 (1.1)	4.7 (4.3)	4.7 (4.6)	6.1 (6.1)	1 (1.2)	0.49 (1.3)
	$m_{H^+} = 140 \text{ GeV}$	2.5 (2.5)	0.9 (1.4)	3.0 (3.0)	7.8 (7.1)	1.8 (1.1)	4.2 (4.4)	5.9 (5.2)	6.1 (6.1)	1.2 (1.3)	1.3 (1.7)
	$m_{H^+} = 150 \text{ GeV}$	2.5 (2.5)	0.65 (0.96)	3.0 (3.0)	7.7 (7.8)	2.1 (1.6)	4.9 (4.6)	6.1 (5.7)	6.1 (6.1)	1.3 (1.5)	1.9 (3)
	$m_{H^+} = 155 \text{ GeV}$	2.5 (2.5)	1.3 (0.59)	3.0 (3.0)	8.8 (10)	1.5 (1.6)	4.7 (4.9)	5.7 (6.2)	6.1 (6.1)	1.5 (1.8)	3 (3.1)
	$m_{H^+} = 160 \text{ GeV}$	2.5 (2.5)	0.91 (1.2)	3.0 (3.0)	9.3 (8.9)	0.83 (1)	4.8 (5.1)	5.3 (5.6)	6.1 (6.1)	1.7 (1.9)	2.7 (2.8)
	SM $t\bar{t} + \text{jets}$	2.5 (2.5)	0.97 (1.2)	3.0 (3.0)	6.9 (6.9)	0.74 (0.58)	4.5 (4.5)	3.8 (3.7)	6.1 (6.1)	0.19 (0.22)	0.73 (1.3)
	Single t	2.5 (2.5)	1.1 (0.82)	3.0 (3.0)	9.1 (10)	0.73 (0.63)	5 (5.2)	4.5 (4.2)	5 (5)	1 (1.2)	—
	W + jets	2.5 (2.5)	1.4 (1.1)	3.0 (3.0)	18 (13)	1.2 (0.29)	10 (8.3)	9.8 (5.1)	5 (5)	6.8 (4.6)	—
	Z/ $\gamma$ + jets	2.5 (2.5)	0.4 (2.3)	3.0 (3.0)	20 (18)	1.5 (0.44)	8.7 (8.2)	6.9 (4.4)	4.5 (4.5)	5.9 (4.4)	—
	VV	2.5 (2.5)	4.5 (11)	3.0 (3.0)	8.7 (16)	2.1 (4)	6.5 (8.5)	7.3 (7.2)	4 (4)	21 (22)	—
	QCD multijet	—	—	—	—	—	—	—	13 (23)	7.9 (7.7)	—
Medium	$m_{H^+} = 80 \text{ GeV}$	2.5 (2.5)	1.1 (0.55)	3.0 (3.0)	5.6 (6.1)	3 (2.9)	5 (4.9)	3.6 (3.7)	6.1 (6.1)	1.2 (1.3)	1.3 (1.5)
	$m_{H^+} = 90 \text{ GeV}$	2.5 (2.5)	0.88 (0.69)	3.0 (3.0)	6.5 (6.3)	3 (2.8)	5 (5.4)	3.8 (3.8)	6.1 (6.1)	1.1 (1.3)	0.77 (1.9)
	$m_{H^+} = 100 \text{ GeV}$	2.5 (2.5)	0.65 (0.36)	3.0 (3.0)	6.7 (4.8)	2.8 (2.7)	5 (4.9)	3.9 (3.8)	6.1 (6.1)	1.1 (1.3)	0.65 (1.5)
	$m_{H^+} = 120 \text{ GeV}$	2.5 (2.5)	0.63 (0.83)	3.0 (3.0)	5.9 (5.1)	2.8 (2.8)	5.2 (5)	4.2 (3.8)	6.1 (6.1)	1.1 (1.3)	0.63 (1.2)
	$m_{H^+} = 140 \text{ GeV}$	2.5 (2.5)	0.6 (0.89)	3.0 (3.0)	6.2 (6.8)	3.3 (2.9)	5.2 (5.6)	4.9 (4.4)	6.1 (6.1)	1.3 (1.4)	0.78 (2.3)
	$m_{H^+} = 150 \text{ GeV}$	2.5 (2.5)	0.98 (0.73)	3.0 (3.0)	8 (7.7)	2.3 (2.7)	5.5 (5.7)	4.7 (4.5)	6.1 (6.1)	1.5 (1.7)	2.1 (3.2)
	$m_{H^+} = 155 \text{ GeV}$	2.5 (2.5)	0.45 (0.48)	3.0 (3.0)	8.7 (10)	3.1 (2.6)	7 (5.7)	5.5 (4.9)	6.1 (6.1)	1.8 (1.9)	1.8 (3.1)
	$m_{H^+} = 160 \text{ GeV}$	2.5 (2.5)	0.2 (0.37)	3.0 (3.0)	9.4 (11)	2.1 (2.6)	6.9 (5.9)	5.2 (4.9)	6.1 (6.1)	1.9 (2.2)	3.3 (4)
	SM $t\bar{t} + \text{jets}$	2.5 (2.5)	0.25 (0.45)	3.0 (3.0)	6.2 (6.2)	1.8 (1.6)	6.3 (6.2)	3.4 (3.5)	6.1 (6.1)	0.23 (0.27)	0.62 (1.4)
	Single t	2.5 (2.5)	0.73 (0.29)	3.0 (3.0)	8.5 (8.6)	1.6 (1.3)	7 (7.2)	3.4 (3.9)	5 (5)	1.3 (1.5)	—
	W + jets	2.5 (2.5)	1.5 (2.5)	3.0 (3.0)	21 (12)	2.1 (1.6)	11 (11)	5.1 (5.4)	5 (5)	5 (5.8)	—
	Z/ $\gamma$ + jets	2.5 (2.5)	1.5 (3.5)	3.0 (3.0)	21 (17)	2 (0.51)	11 (11)	4.2 (5.2)	4.5 (4.5)	6.3 (6)	—
	VV	2.5 (2.5)	13 (7.2)	3.0 (3.0)	28 (54)	5.1 (6.4)	20 (8.3)	8.9 (3.2)	4 (4)	21 (35)	—
	QCD multijet	—	—	—	—	—	—	—	15 (16)	12 (10)	—
Tight	$m_{H^+} = 80 \text{ GeV}$	2.5 (2.5)	0.65 (1.4)	3.0 (3.0)	5.6 (4.4)	5.7 (6.2)	4.5 (5.5)	6.3 (5.8)	6.1 (6.1)	1.7 (2)	1.1 (0.78)
	$m_{H^+} = 90 \text{ GeV}$	2.5 (2.5)	1.2 (1.6)	3.0 (3.0)	5.8 (5.4)	5.5 (6.4)	5.8 (5)	5.1 (5.1)	6.1 (6.1)	1.7 (2)	0.92 (1.9)
	$m_{H^+} = 100 \text{ GeV}$	2.5 (2.5)	1.3 (1.4)	3.0 (3.0)	5.2 (5.6)	6.1 (5.9)	5.5 (5)	5.6 (5.3)	6.1 (6.1)	1.7 (1.9)	0.85 (1.1)
	$m_{H^+} = 120 \text{ GeV}$	2.5 (2.5)	1 (1.7)	3.0 (3.0)	4.8 (4.5)	5.6 (5.9)	5.2 (5.4)	5.4 (5.9)	6.1 (6.1)	1.7 (1.9)	0.95 (0.59)
	$m_{H^+} = 140 \text{ GeV}$	2.5 (2.5)	1.1 (1.6)	3.0 (3.0)	5.5 (4.8)	6.3 (6.4)	4.7 (5.3)	5.3 (5.7)	6.1 (6.1)	2 (2.2)	1.6 (1.1)
	$m_{H^+} = 150 \text{ GeV}$	2.5 (2.5)	0.84 (1.8)	3.0 (3.0)	6.3 (7.7)	7.1 (6)	5.2 (5.9)	5.9 (6.3)	6.1 (6.1)	2.4 (2.6)	1.9 (2.2)
	$m_{H^+} = 155 \text{ GeV}$	2.5 (2.5)	2.1 (1.1)	3.0 (3.0)	9.3 (6)	7.3 (6.1)	7.3 (5.5)	7.5 (6.4)	6.1 (6.1)	2.8 (3.1)	1.9 (4.5)
	$m_{H^+} = 160 \text{ GeV}$	2.5 (2.5)	0.63 (1.5)	3.0 (3.0)	8.9 (6.9)	7.1 (5.8)	5.7 (6.3)	7.2 (5.6)	6.1 (6.1)	3.3 (3.8)	4.9 (3.6)
	SM $t\bar{t} + \text{jets}$	2.5 (2.5)	1.3 (0.93)	3.0 (3.0)	5.9 (6)	4.9 (4.9)	6.8 (6.3)	5.5 (5.1)	6.1 (6.1)	0.44 (0.5)	0.47 (1.3)
	Single t	2.5 (2.5)	0.72 (0.38)	3.0 (3.0)	7.8 (8.9)	4.7 (4.2)	7.6 (7)	6.2 (5.8)	5 (5)	2.6 (3)	—
	W + jets	2.5 (2.5)	2.4 (3)	3.0 (3.0)	23 (27)	9.7 (5.4)	9.2 (9.5)	13 (7.9)	5 (5)	13 (14)	—
	Z/ $\gamma$ + jets	2.5 (2.5)	8.8 (4.6)	3.0 (3.0)	8.6 (13)	4.5 (3.6)	16 (9.5)	10 (8.1)	4.5 (4.5)	9.5 (15)	—
	VV	2.5 (2.5)	0.66 (8.9)	3.0 (3.0)	27 (0)	32 (7.4)	9.1 (6.9)	10 (2.3)	4 (4)	40 (1e+02)	—
	QCD multijet	—	—	—	—	—	—	—	11 (13)	28 (18)	—

### 11.3. Effect of systematics on $m_{jj}$ distribution

All the systematic uncertainties described above may affect the  $m_{jj}$  distribution differently. For example, some of the uncertainties may have a different effect in each bin and the others may have the same effect in every bin. Therefore, it is very important to compare the base  $m_{jj}$  distributions with those obtained considering systematic uncertainties. The scale factors for the given systematic are varied up/down and the corresponding  $m_{jj}$  distributions are obtained. We have studied the effect of following systematics: JES, JER, t quark  $p_T$  reweighting (`topPt`), b and c tagging (`bcTag1`, `bcTag2`, and `bcTag3`), t quark mass (`topMass_tt`), factorisation and renormalization scale (`scaleRF_tt`), and parton shower matching (`hDamp_tt`). The  $m_{jj}$  distributions for these systematics are shown in Figures 11.3, 11.4 for charged Higgs signal ( $m_{H^+} = 120$  GeV) and in Figures 11.5, 11.6 for  $t\bar{t} + \text{jets}$  process. From these figures one can see that some of the systematics (JES, b and c tagging) have a significant effect, others have a marginal effect.

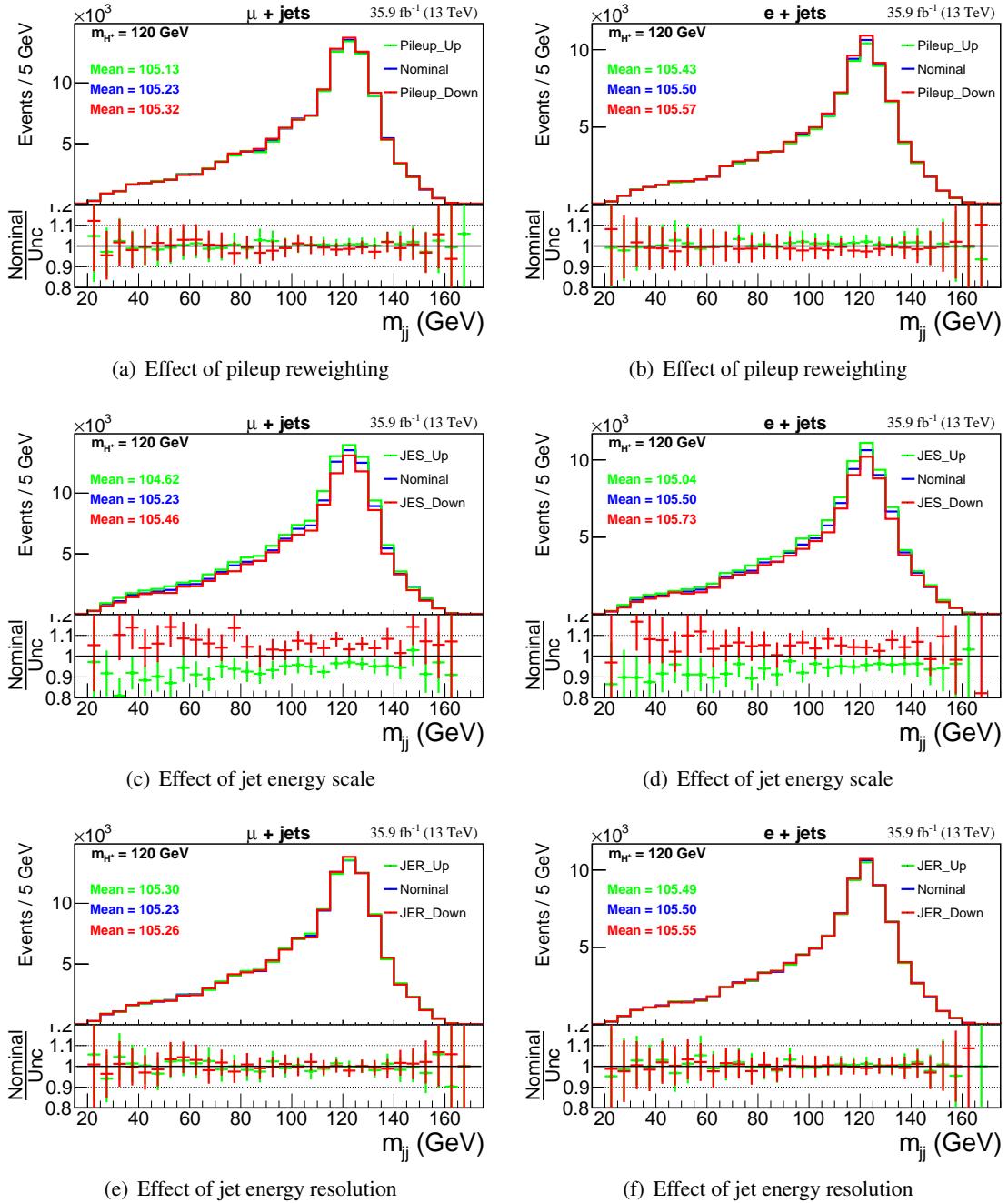


Figure 11.3.: Effect of pileup, jet energy scale, and jet energy resolution systematics on the  $m_{\text{jj}}$  distribution for the charged Higgs signal ( $m_{H^+} = 120 \text{ GeV}$ ) for  $\mu + \text{jets}$  and  $e + \text{jets}$  channel , after  $\geq 1 \text{ c}$  jet selection.

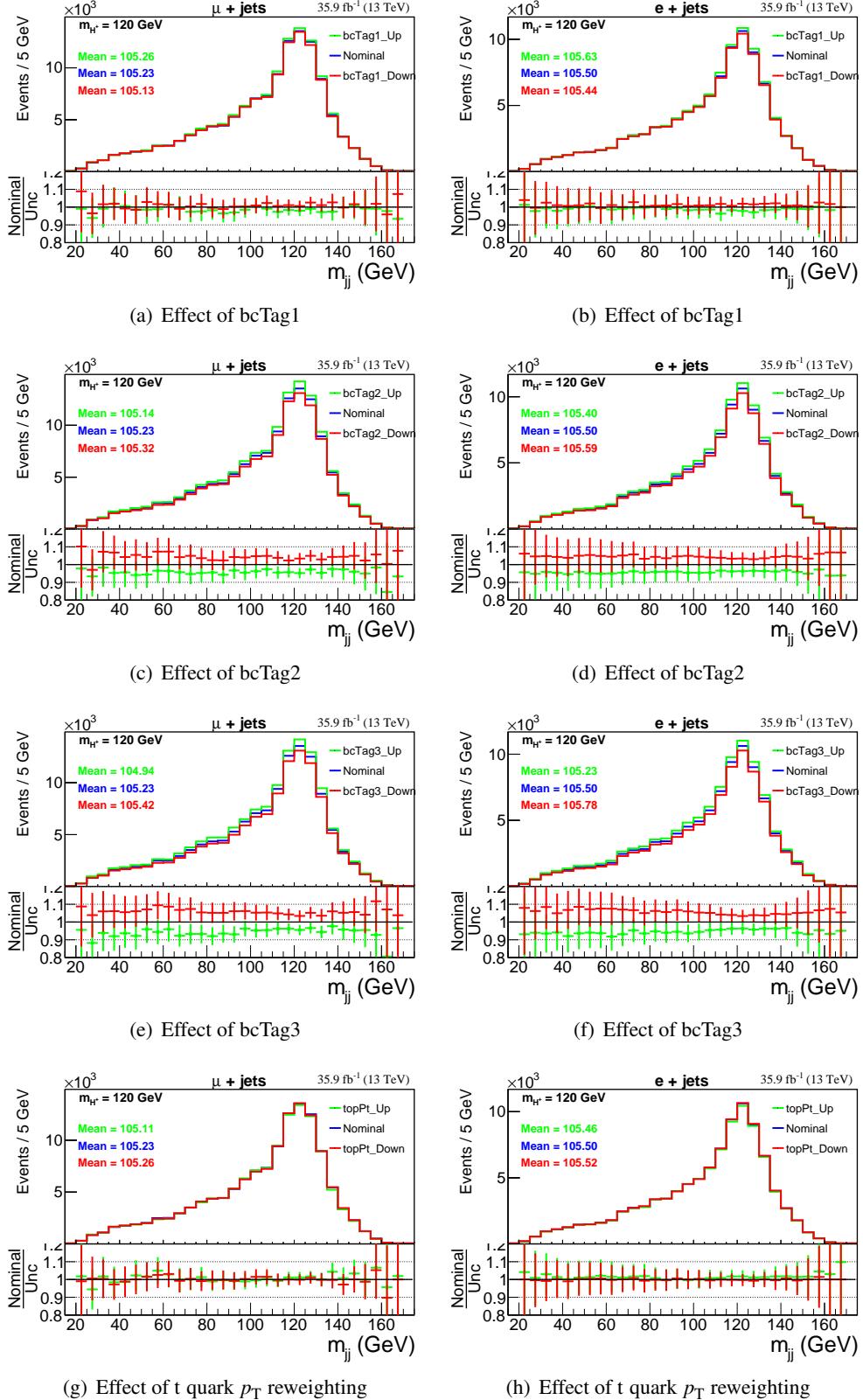


Figure 11.4.: Effect of b/c tagging and t quark  $p_T$  systematics on the  $m_{jj}$  distribution from charged Higgs signal ( $m_{H^+} = 120$  GeV) for  $\mu + \text{jets}$  and  $e + \text{jets}$  channel, after  $\geq 1$  c jet selection.

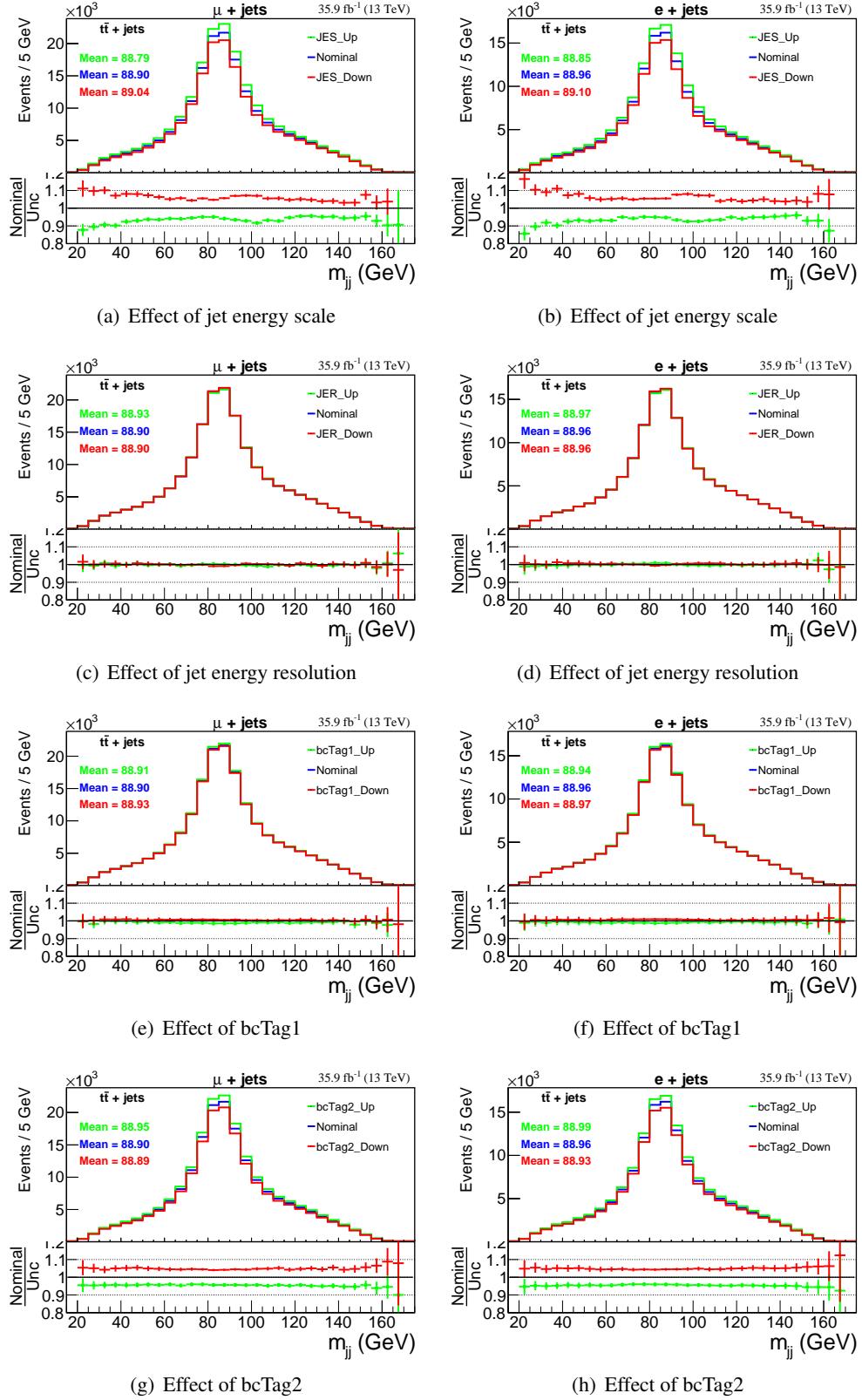


Figure 11.5.: Effect of jet energy scale, jet energy resolution, and b/c tagging systematics on the  $m_{jj}$  distribution from  $t\bar{t} + \text{jets}$  background for  $\mu + \text{jets}$  and  $e + \text{jets}$  channel, after  $\geq 1$  c jet selection.

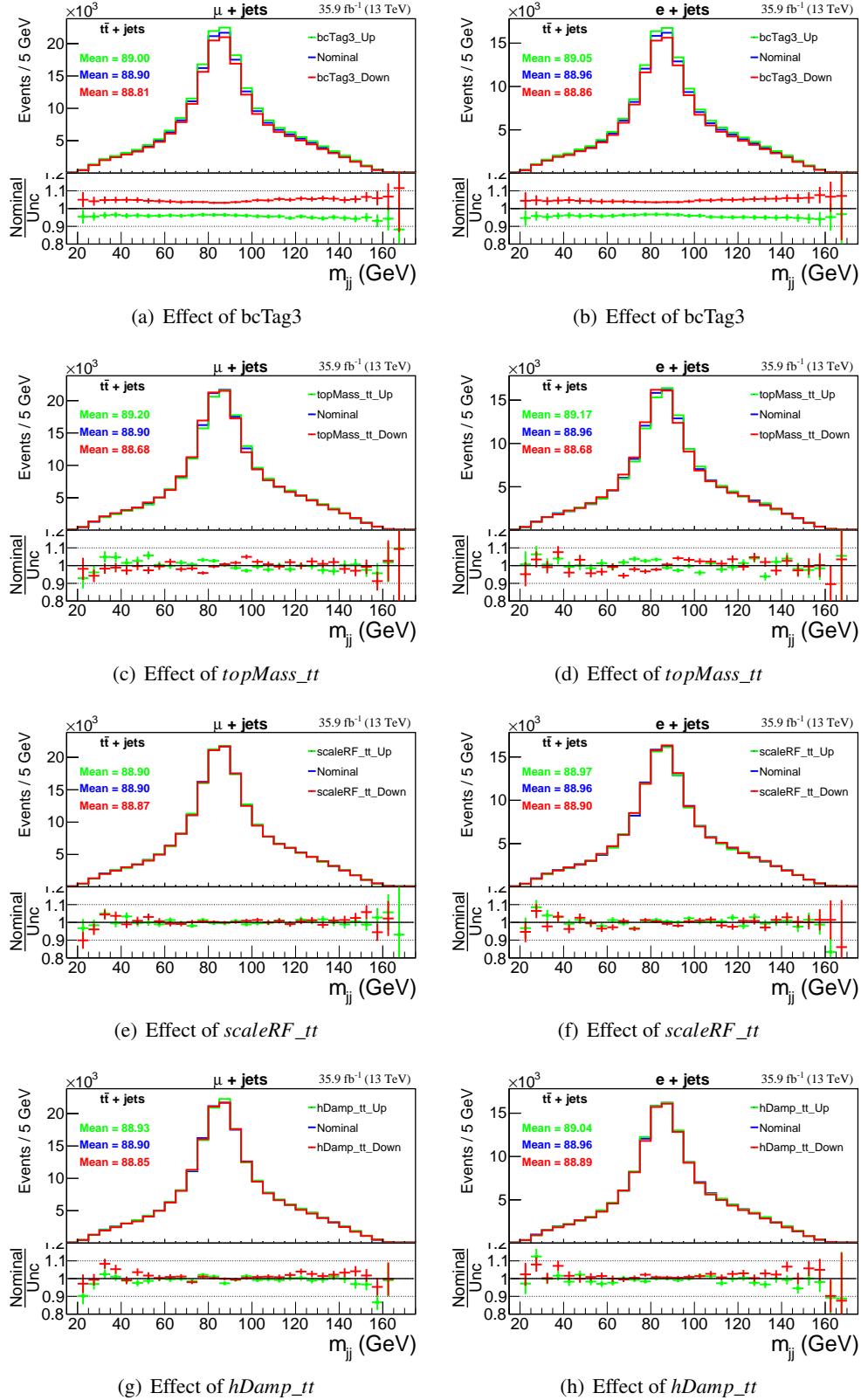


Figure 11.6.: Effect of c tagging, t quark mass, renormalization and factorisation scale, and parton shower matching systematics on the  $m_{jj}$  distribution from  $t\bar{t} + \text{jets}$  background for  $\mu + \text{jets}$  and  $e + \text{jets}$  channel, after  $\geq 1$  c jet selection.

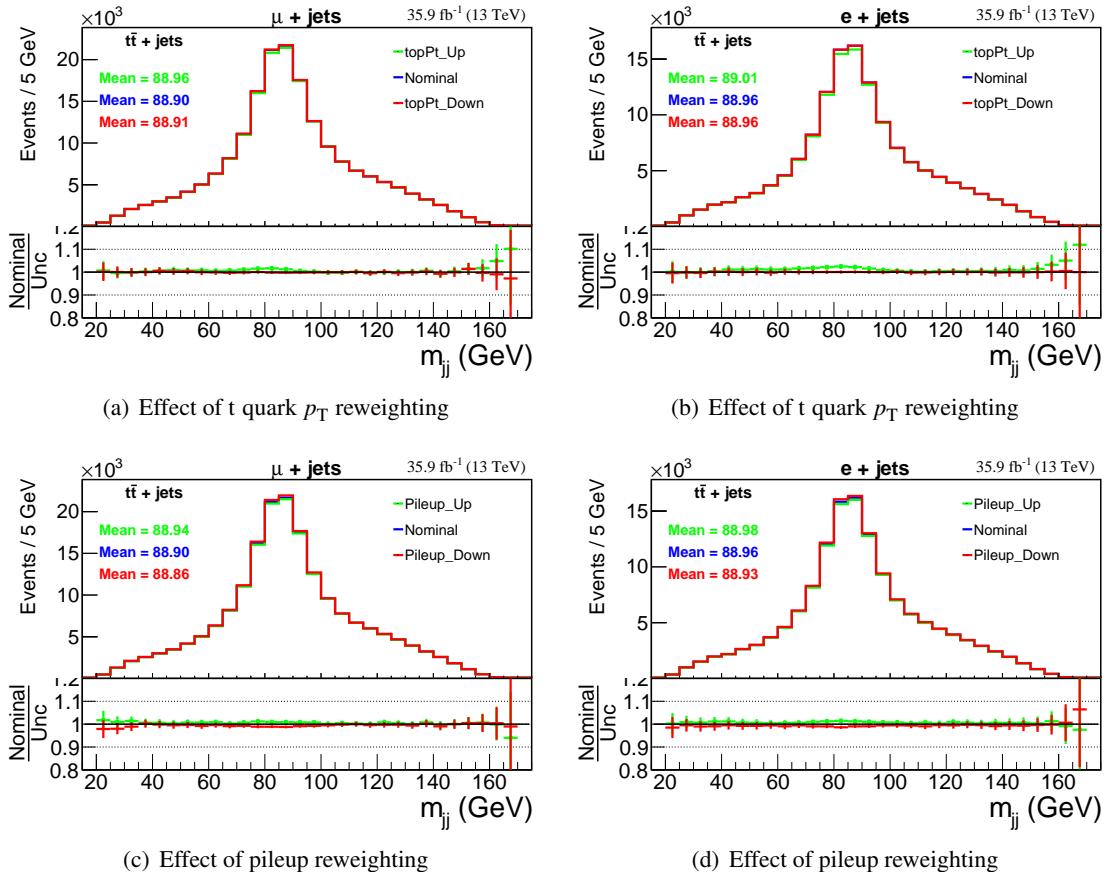


Figure 11.7.: Effect of t quark  $p_T$  and pileup reweighting systematics on the  $m_{\text{JJ}}$  distribution from  $t\bar{t} + \text{jets}$  background for  $\mu + \text{jets}$  and  $e + \text{jets}$  channel, after  $\geq 1$  c jet selection.



# 12 | Exclusion Limits

## 12.1. 95% CL upper limit on $\mathcal{B}(t \rightarrow H^+ b)$

The event yields from different event categories with systematics and statistical uncertainties are shown in Tables 10.1, 10.2, and 10.3 for  $\mu +$  jets and  $e +$  jets channel. From these tables, it can be seen that the total background event matches well with the data within the uncertainties. From Figures 8.1, 10.2, and 10.6 it can be seen that no excess has been observed for the mass of charged Higgs in the range 80 to 160 GeV. In view of this, an exclusion limit at 95% CL [137] using profile likelihood ratio has been put on the  $\mathcal{B}(t \rightarrow H^+ b)$ , assuming  $\mathcal{B}(H^+ \rightarrow c\bar{s}) = 100\%$ . The absence or presence of charged Higgs in the data is characterized by the exclusion limits. Less is the value of limit, the more is the possibility of the absence of charged Higgs signal and vice-versa.

However, the procedure for setting the upper limit is not straight-forward as it involves many mathematical techniques and intensive computing resources as there so many bins of the  $m_{jj}$  distributions, statistical and systematics uncertainties, etc. A detailed mathematical description of the limit setting procedure is given in [138]. A precise description of it can be found in Chapter 2 and 5 of [139]. Since there are good literature how the 95% CL limit is calculated, we do not discuss mathematical details about it in this thesis. The only thing that is specific to this analysis is setting the limit on  $\mathcal{B}(t \rightarrow H^+ b)$  rather than on the signal strength. For this, the signal and background yield that goes in to the likelihood is scaled appropriately. For example, denoting  $x$  as the branching ratio of  $t$  quark decaying to charged Higgs and a bottom quark

$$x = \mathcal{B}(t \rightarrow H^+ b), \quad (12.1)$$

the different yields are scaled as

$$x^2 N^{HH} + 2x(1-x)N^{HW} + (1-x)^2 N_{t\bar{t}+jets}^{WW} + N_{other\ bkg}, \quad (12.2)$$

where we have assumed that  $\mathcal{B}(t \rightarrow H^+ b) + \mathcal{B}(t \rightarrow W^+ b) = 1$ . The  $N^{HW}$ ,  $N_{t\bar{t}+jets}^{WW}$ , and  $N_{other\ bkg}$  is the number of events from simulated signal, SM  $t\bar{t} +$  jets, and other background process as shown in Table 10.3

from exclusive categories for  $\mu + \text{jets}$  and  $e + \text{jets}$  channel. The contribution of  $N^{HH}$  is very small, hence this process is not considered in this analysis. The Higgs combine tool package is used to compute asymptotic limit on  $x$  using the CLs method [137]. Since the signal search region lies in the range 80 to 160 GeV of  $m_{jj}$ , a  $20 \geq m_{jj} \geq 170$  cut is used to avoid low statistical bins for further analysis. The limits are computed for each event category described in Chapter 10.

## 12.2. Profile of nuisance parameter: shape vs lnN

The statistical and systematic uncertainties are treated as nuisance parameters (NP) for limit computation. The statistical uncertainty is propagated through the `autoMCStats` tool. Which adds histograms of each individual background process and assigns one nuisance in each bin of the total background. For each systematic uncertainty, one NP is assigned. There are different types of probability distribution function (PDF) associated with each NP that goes in the likelihood for limit computation. If for the given NP, the PDF is log-normal (Gaussian) distribution then the corresponding NP is said to be profiled as lnN (shape). Since the final observable is  $m_{jj}$ , a NP is profiled as shape (lnN) based on whether its value is different (same) for all the bins of  $m_{jj}$ . From Figures 12.1–12.6, it can be seen that the ratio of up and down template with the base is compatible with a flat line for jet energy scale, jet energy resolution, b/c tagging, and t quark  $p_T$  systematics. Therefore, these NPs are profiled as lnN. The trend for t quark mass (`topMass_tt`), renormalization and factorization scale (`scaleRF_tt`), and parton showering matching (`hDamp_tt`) is mixed. Hence these are also, conservatively, profiled as lnN. A description of the acronyms used to denote the NPs is shown in Table 12.1.

Table 12.1.: Description of nuisance parameters corresponding to the systematic and statistical uncertainties.

<b>NP name</b>	<b>Profile</b>	<b>NP description</b>
<b>Systematics:</b>		
lumi_13TeV	lnN	uncertainty in the luminosity measurement at 13 TeV
CMS_eff_l	lnN	lepton ( $e, \mu$ ) selection uncertainty
CMS_eff_bcInc1	lnN	uncertainty from inclusive b/c tagging-1
CMS_eff_bcInc2	lnN	uncertainty from inclusive b/c tagging-2
CMS_eff_bcInc3	lnN	uncertainty from inclusive b/c tagging-3
CMS_pileup	lnN	uncertainty from pileup reweighting
CMS_scale_j	lnN	uncertainty from jet energy scale (JES)
CMS_res_j	lnN	uncertainty from jet energy resolution (JER)
CMS_norm_tt	lnN	uncertainty on the $t\bar{t}$ + jets cross section
CMS_norm_stop	lnN	uncertainty on the single t cross section
CMS_norm_wjet	lnN	uncertainty on the W + jets cross section
CMS_norm_zjet	lnN	uncertainty on the Z/ $\gamma$ + jets cross section
CMS_norm_qcd	lnN	uncertainty from the data-driven QCD multijet
CMS_norm_vv	lnN	uncertainty on the VV cross section
CMS_topPtReweight	lnN	uncertainty from t quark $p_T$ reweighting
scaleRF_tt	lnN	uncertainty from renormalisation (R) and factorization (F) scale
hDamp_tt	lnN	uncertainty from parton-shower matching
topMass_tt	lnN	uncertainty from t mass $\pm 1$ GeV
<b>Statistical:</b>		
prop_binch*bin*	shape	bin-by-bin statistical uncertainties from autoMCStats

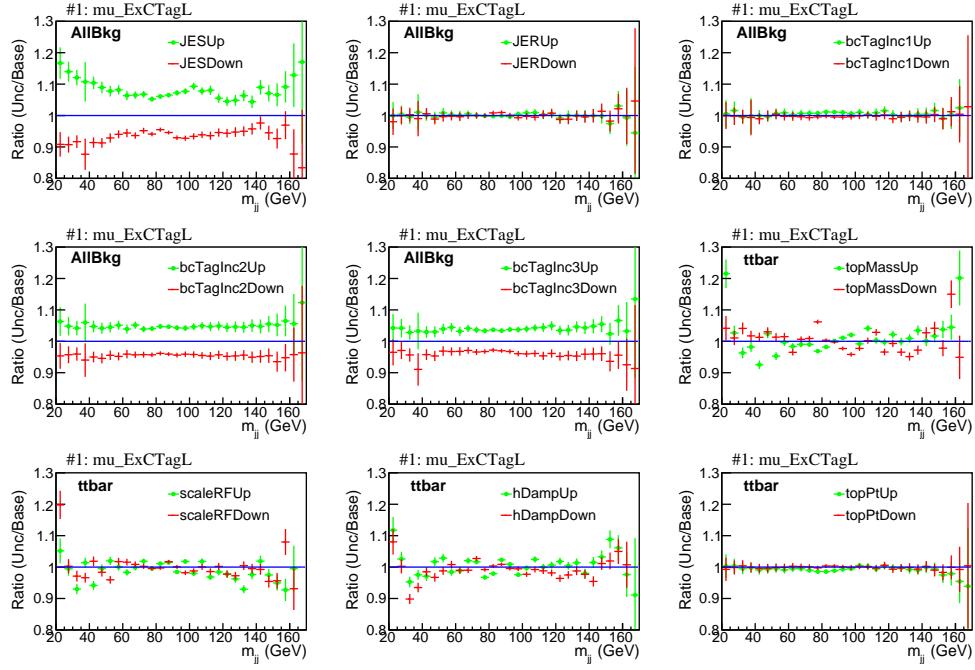


Figure 12.1.: Ratio of up and down with base template for exclusive loose charm category for  $\mu + \text{jets}$  channel.

## 12.3. Limits from inclusive event category without c tagging

To extract possible signal, the  $m_{jj}$  distribution as shown in Figures 8.1(c), 8.1(d), 10.1(c), and 10.1(d) are used in a binned maximum-likelihood fit. The upper limit on the  $\mathcal{B}(t \rightarrow H^+ b)$  as a function of  $m_{H^+}$  is shown in Figures 12.7(a), 12.7(b), and 12.7(c) for the  $\mu + \text{jets}$  and  $e + \text{jets}$  and  $\ell + \text{jets}$  channel, respectively. Corresponding values of the expected (observed) limits, for the charged Higgs mass from 90 to 160 GeV, are in the range 0.40–1.42% (0.32–0.98%), 0.44–1.57% (0.28–2.35%), and 0.32–1.08% (0.18–1.04) respectively. The expected (observed) limits using  $m_{jj}$  from inclusive event category at 8 TeV was in range 1.4–3.6% (1.2–6.5%) for lepton + jets channel [63].

## 12.4. Limits from inclusive event category with loose c tagging

The  $m_{jj}$  distribution from the events with inclusive loose c tagging as discussed in Section 10.2 are used to compute upper limits as shown in Figures 12.7(d), 12.7(e), and 12.7(f). From these figures, it can be seen that the limits are marginally better compared to that of Figures 12.7(a), 12.7(b), and 12.7(c). For different charged Higgs masses, the expected (observed) limits for loose c tagging are in the range 0.40–1.41% (0.26–0.82%), 0.44–1.55% (0.25–1.77%), and 0.31–1.07% (0.16–0.76%) for  $\mu + \text{jets}$ ,  $e + \text{jets}$ , and

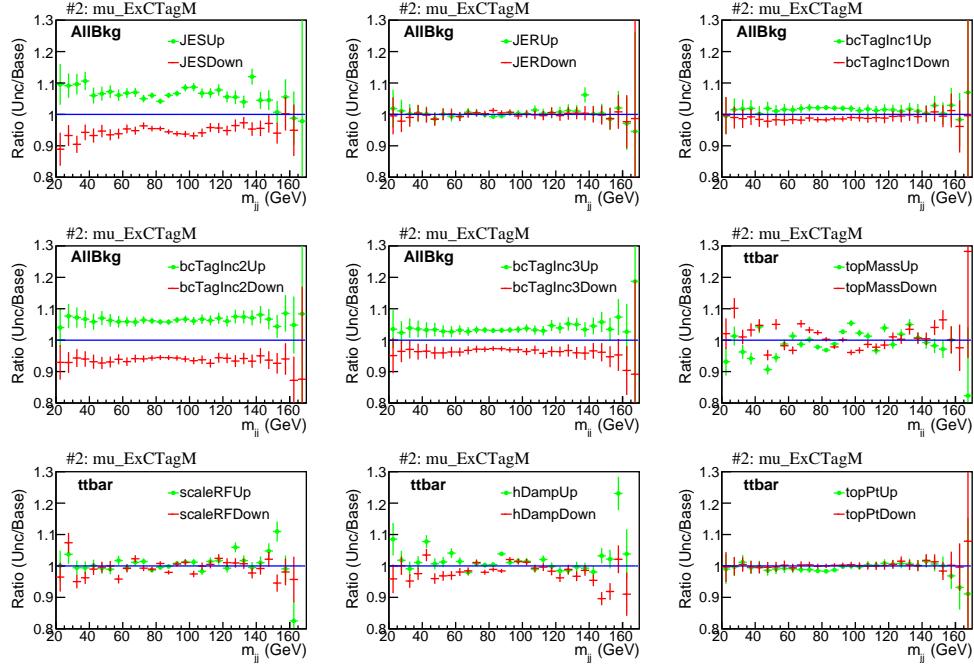


Figure 12.2.: Ratio of up and down with base template for exclusive medium charm category for  $\mu + \text{jets}$  channel.

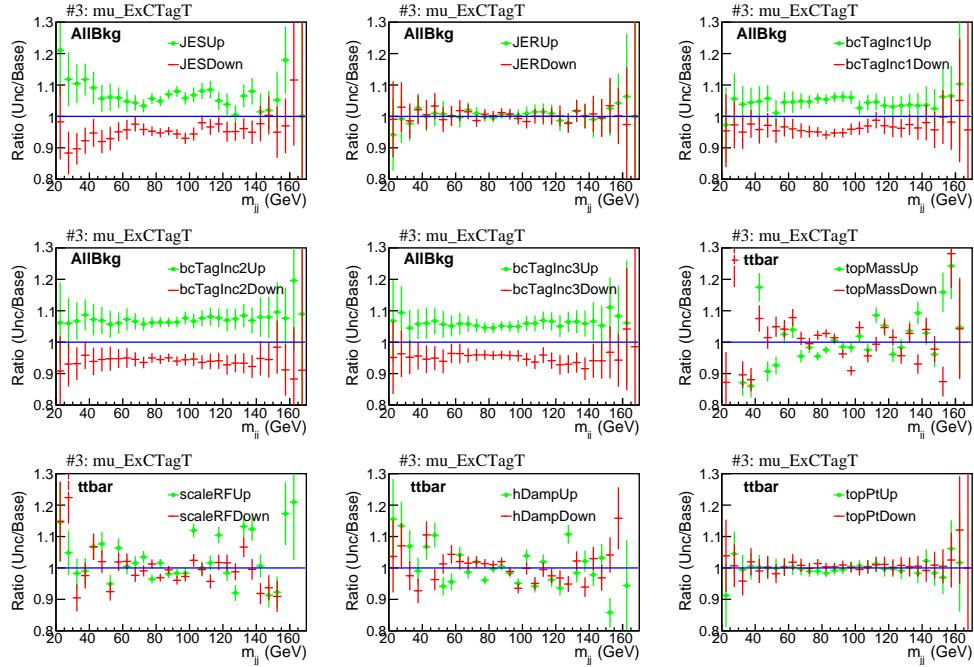


Figure 12.3.: Ratio of up and down with base template from exclusive tight charm category for  $\mu + \text{jets}$  channel.

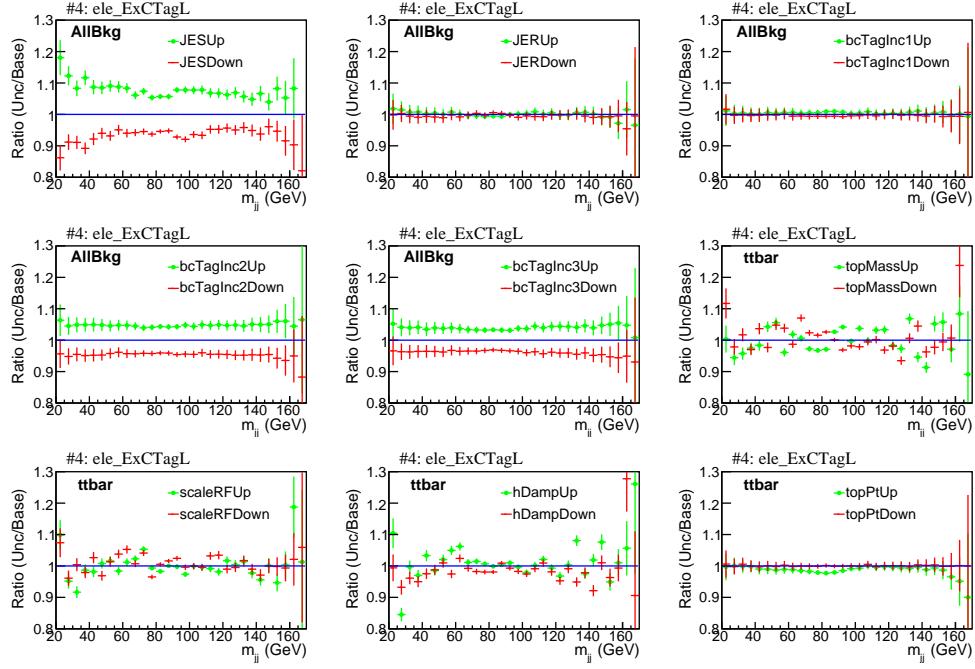


Figure 12.4.: Ratio of up and down with base template from exclusive loose charm category for  $e + \text{jets}$  channel.

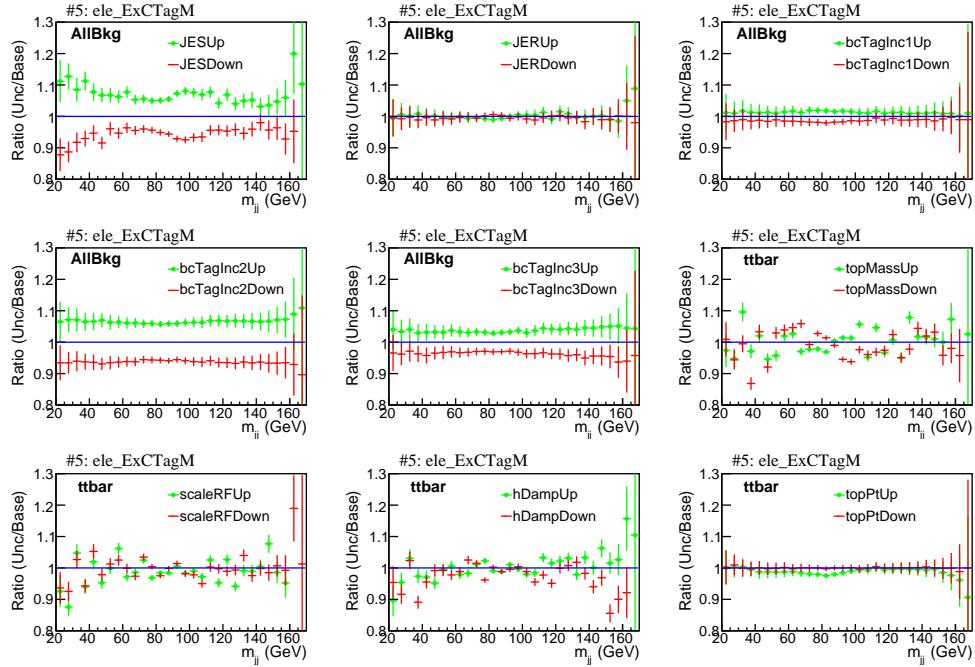


Figure 12.5.: Ratio of up and down with base template from exclusive medium charm category for  $e + \text{jets}$  channel.

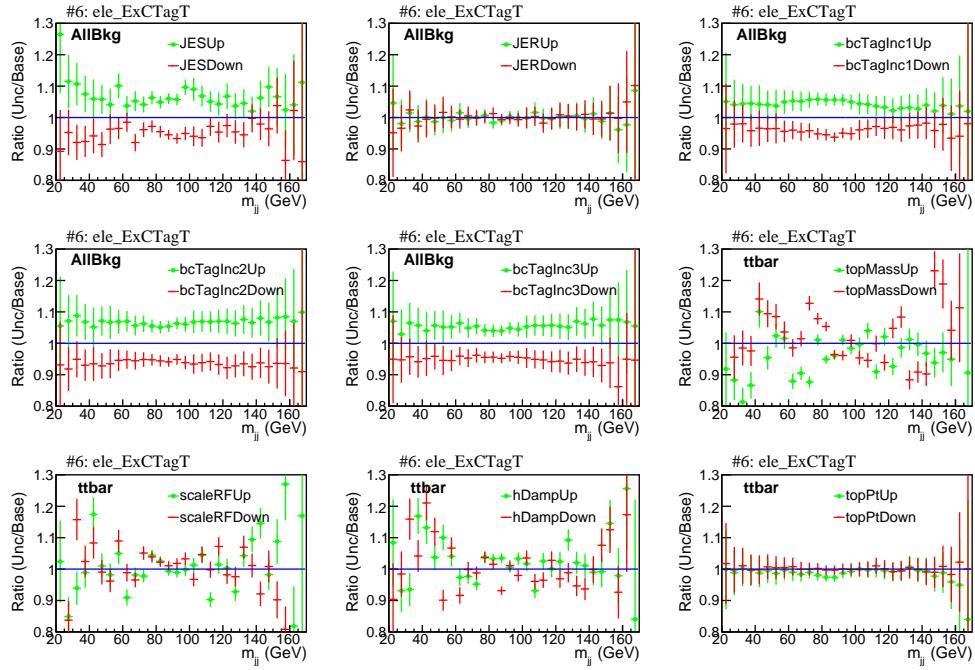


Figure 12.6.: Ratio of up and down with base template from exclusive tight charm category for  $e + \text{jets}$  channel.

$\ell + \text{jets}$  channel, respectively.

## 12.5. Limits from exclusive event categories based on c tagging

The signal significance is different in different charm categories as described in Section 10.3. The  $m_{jj}$  distributions from loose, medium, and tight exclusive charm tagging categories are combined to compute the expected limits as shown in Figure 12.8. From this figure, it can be seen that the upper limits from exclusive categories are better compared to that obtained from inclusive working points. For different charged Higgs masses, the expected (observed) limits from combined categories are in the range 0.39–1.34% (0.29–0.79%), 0.43–1.48% (0.27–1.45%), and 0.31–1.03% (0.19–0.68) for  $\mu + \text{jets}$ ,  $e + \text{jets}$ , and  $\ell + \text{jets}$  channel, respectively.

## 12.6. Impact of nuisance parameters

All the systematics and statistical uncertainties (nuisance parameters) are included in the likelihood function along with the parameter of interest (POI). The POI is the  $\mathcal{B}(t \rightarrow H^+ b)$  of this analysis. For the sake of convenience, we will denote  $\mathcal{B}(t \rightarrow H^+ b)$  as BR in this section. The impact of nuisance parameters (NP)

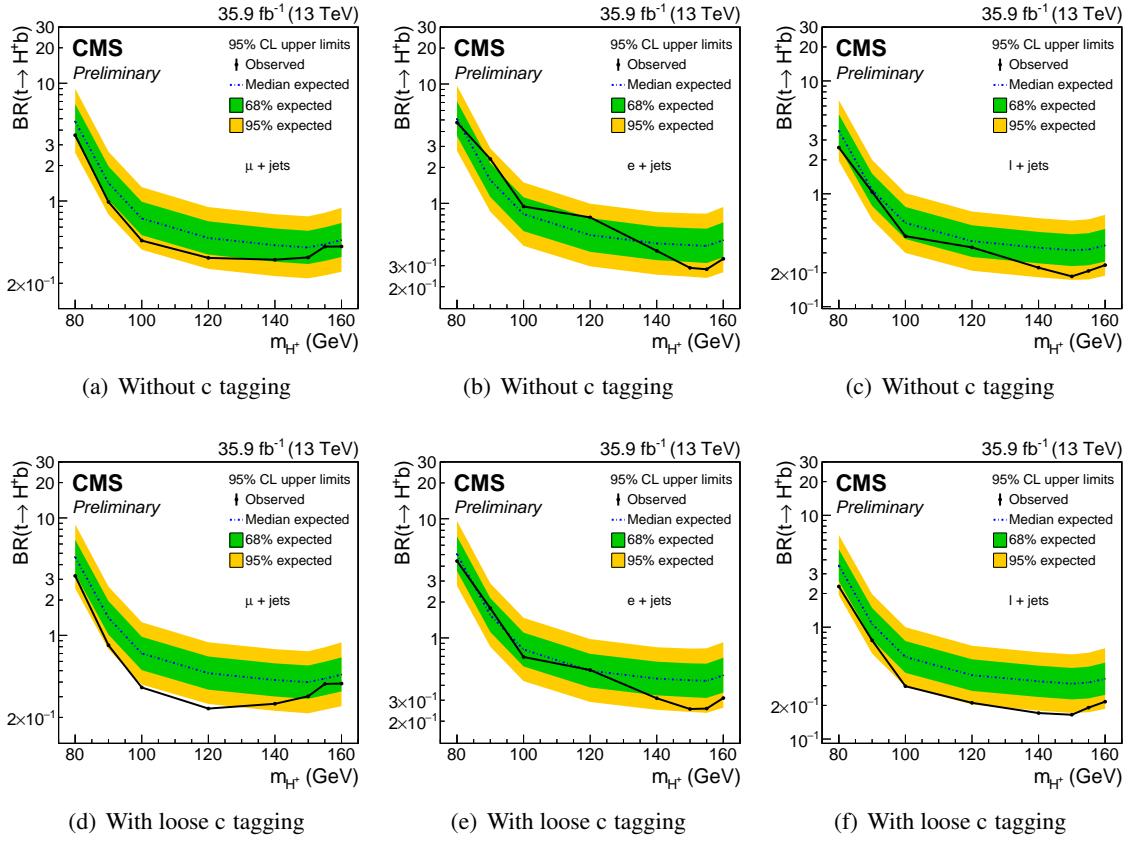


Figure 12.7.: The upper limit in % on  $\mathcal{B}(t \rightarrow H^+ b)$  as a function of  $m_{H^+}$  using  $m_{jj}$  after kinematic fit selection without c tagging, and with loose c tagging, as discussed in Section 10.1 and 10.2, for  $\mu +$  jets,  $e +$  jets, and  $\ell +$  jets channel.

Table 12.2.: 95% CL exclusion limit in % for  $\mu +$  jets channel from different event categories.

$m_{H^+}$ (GeV)	$m_{jj}(\text{Inc})$		$m_{jj}(\text{Inc CTagL})$		$m_{jj}(\text{Ex CTag})$	
	Expected (%)	Observed (%)	Expected (%)	Observed (%)	Expected (%)	Observed (%)
80	$4.75^{+1.91}_{-1.33}$	3.62	$4.65^{+1.87}_{-1.32}$	3.21	$3.06^{+1.25}_{-0.873}$	2.12
90	$1.42^{+0.549}_{-0.399}$	0.982	$1.41^{+0.546}_{-0.396}$	0.823	$1.34^{+0.528}_{-0.376}$	0.793
100	$0.71^{+0.275}_{-0.198}$	0.461	$0.701^{+0.271}_{-0.195}$	0.36	$0.676^{+0.261}_{-0.19}$	0.379
120	$0.485^{+0.187}_{-0.134}$	0.329	$0.476^{+0.184}_{-0.131}$	0.238	$0.464^{+0.179}_{-0.129}$	0.288
140	$0.421^{+0.163}_{-0.116}$	0.317	$0.416^{+0.161}_{-0.115}$	0.261	$0.411^{+0.159}_{-0.115}$	0.335
150	$0.403^{+0.156}_{-0.111}$	0.332	$0.399^{+0.154}_{-0.111}$	0.303	$0.393^{+0.152}_{-0.109}$	0.371
155	$0.43^{+0.168}_{-0.12}$	0.41	$0.428^{+0.167}_{-0.119}$	0.385	$0.414^{+0.163}_{-0.116}$	0.447
160	$0.465^{+0.185}_{-0.131}$	0.411	$0.463^{+0.186}_{-0.13}$	0.388	$0.437^{+0.179}_{-0.125}$	0.439

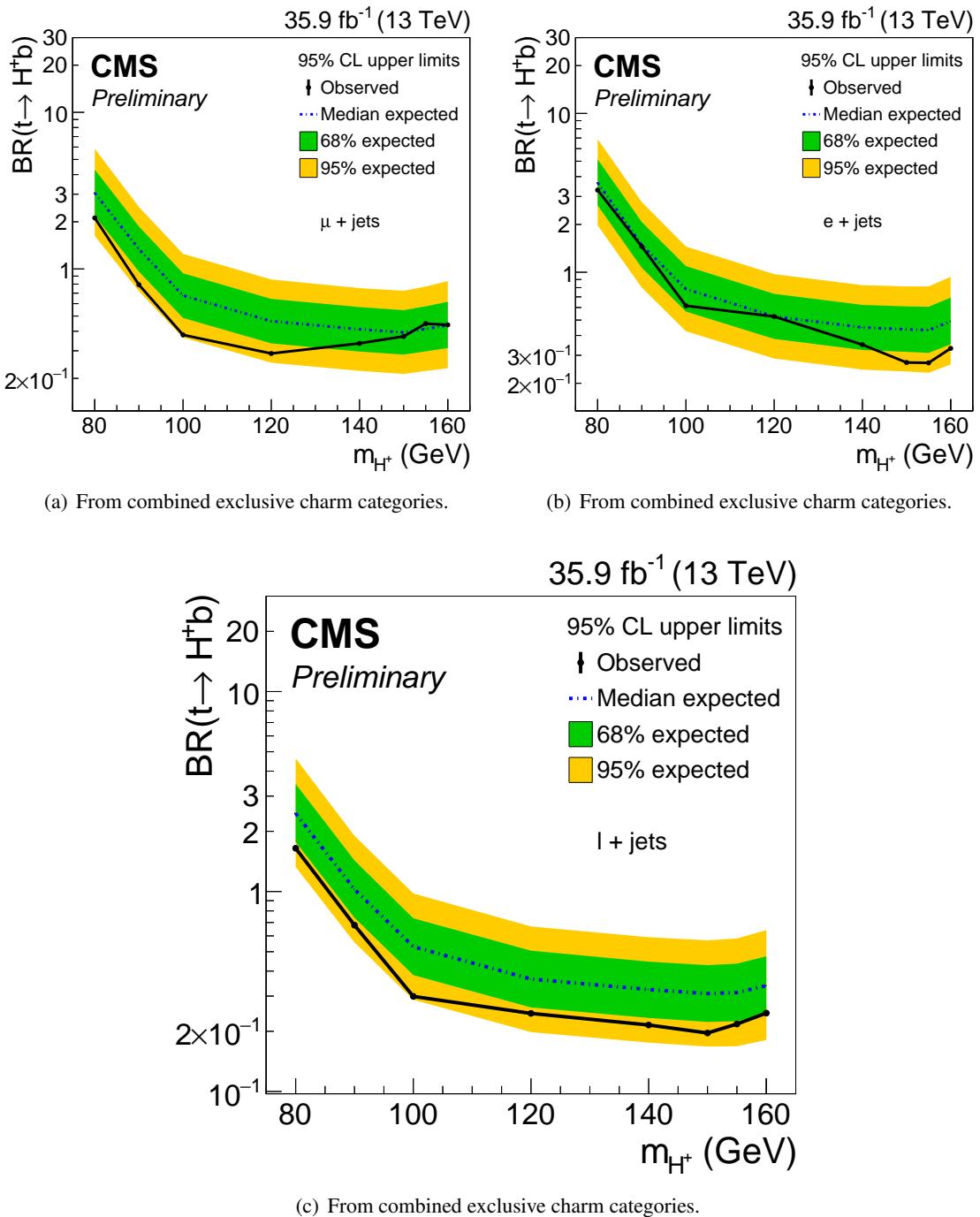
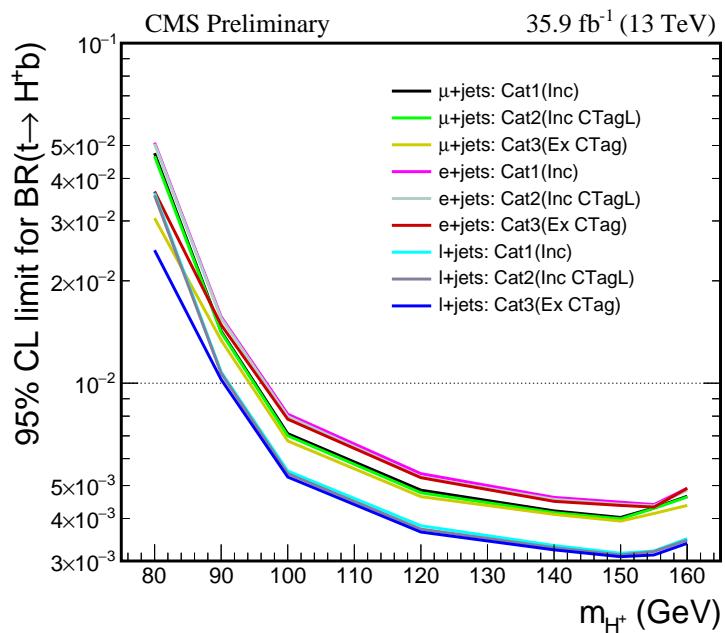
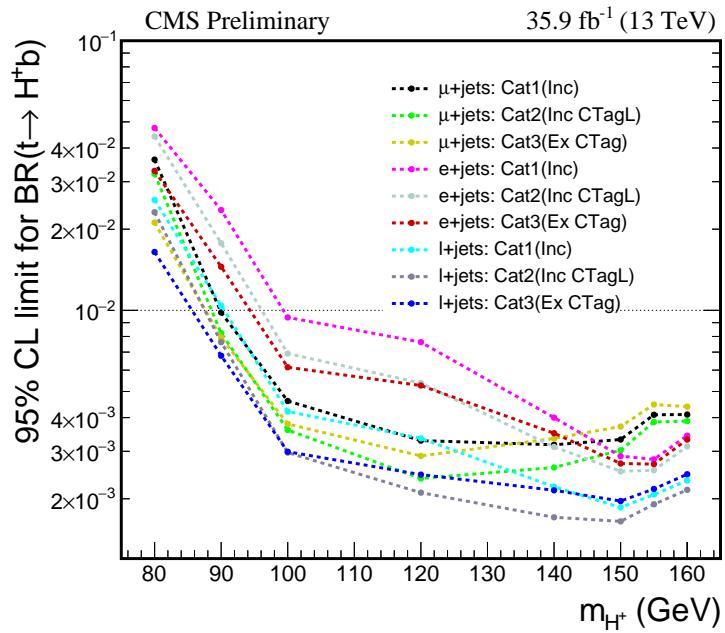


Figure 12.8.: The upper limit in % on  $\mathcal{B}(t \rightarrow H^+ b)$  as a function of  $m_{H^+}$  by combining  $m_{jj}$  distributions from different exclusive charm categories as discussed in Section 10.3 for  $\mu + \text{jets}$   $e + \text{jets}$ , and  $\ell + \text{jets}$  channel.



(a) Expected limits.



(b) Observed limits.

Figure 12.9.: The upper limit in % on  $\mathcal{B}(t \rightarrow H^+ b)$  as a function of  $m_{H^+}$  using  $m_{jj}$  from all event categories for  $\mu + \text{jets}$ ,  $e + \text{jets}$ , and  $\ell + \text{jets}$  channel.

Table 12.3.: 95% CL exclusion limit in % for e + jets channel from different event categories.

$m_{H^+}$ (GeV)	$m_{jj}(Inc)$		$m_{jj}(Inc\ CTagL)$		$m_{jj}(Ex\ CTag)$	
	Expected (%)	Observed (%)	Expected (%)	Observed (%)	Expected (%)	Observed (%)
80	$5.1^{+2.05}_{-1.43}$	4.75	$5.06^{+2.04}_{-1.42}$	4.42	$3.66^{+1.47}_{-1.03}$	3.29
90	$1.57^{+0.606}_{-0.437}$	2.35	$1.55^{+0.607}_{-0.432}$	1.77	$1.48^{+0.586}_{-0.417}$	1.45
100	$0.811^{+0.313}_{-0.226}$	0.941	$0.798^{+0.309}_{-0.222}$	0.691	$0.784^{+0.309}_{-0.218}$	0.615
120	$0.542^{+0.21}_{-0.15}$	0.762	$0.53^{+0.205}_{-0.146}$	0.537	$0.527^{+0.204}_{-0.147}$	0.526
140	$0.461^{+0.177}_{-0.127}$	0.4	$0.455^{+0.176}_{-0.126}$	0.311	$0.449^{+0.174}_{-0.125}$	0.35
150	$0.447^{+0.173}_{-0.124}$	0.288	$0.441^{+0.17}_{-0.123}$	0.253	$0.437^{+0.172}_{-0.122}$	0.27
155	$0.439^{+0.173}_{-0.123}$	0.28	$0.436^{+0.172}_{-0.122}$	0.255	$0.432^{+0.174}_{-0.121}$	0.269
160	$0.491^{+0.201}_{-0.138}$	0.343	$0.485^{+0.197}_{-0.137}$	0.313	$0.491^{+0.201}_{-0.139}$	0.332

Table 12.4.: 95% CL exclusion limit in % for  $\ell + \text{jets}$  channel from different event categories.

$m_{H^+}$ (GeV)	$m_{jj}(Inc)$		$m_{jj}(Inc\ CTagL)$		$m_{jj}(Ex\ CTag)$	
	Expected (%)	Observed (%)	Expected (%)	Observed (%)	Expected (%)	Observed (%)
80	$3.62^{+1.43}_{-1.02}$	2.57	$3.56^{+1.41}_{-0.993}$	2.31	$2.46^{+0.991}_{-0.691}$	1.65
90	$1.08^{+0.426}_{-0.298}$	1.04	$1.07^{+0.413}_{-0.298}$	0.762	$1.03^{+0.405}_{-0.286}$	0.678
100	$0.552^{+0.213}_{-0.154}$	0.421	$0.542^{+0.21}_{-0.151}$	0.297	$0.53^{+0.205}_{-0.148}$	0.299
120	$0.381^{+0.146}_{-0.105}$	0.335	$0.372^{+0.144}_{-0.103}$	0.211	$0.365^{+0.141}_{-0.102}$	0.246
140	$0.333^{+0.126}_{-0.0908}$	0.222	$0.328^{+0.124}_{-0.0907}$	0.171	$0.323^{+0.122}_{-0.0901}$	0.215
150	$0.316^{+0.122}_{-0.0873}$	0.186	$0.312^{+0.12}_{-0.0863}$	0.165	$0.309^{+0.119}_{-0.086}$	0.196
155	$0.321^{+0.127}_{-0.0883}$	0.207	$0.32^{+0.124}_{-0.0891}$	0.191	$0.312^{+0.123}_{-0.0878}$	0.218
160	$0.349^{+0.139}_{-0.0981}$	0.234	$0.345^{+0.136}_{-0.097}$	0.216	$0.338^{+0.136}_{-0.0958}$	0.247

Table 12.5.: Acronym used in the impact plots for naming statistical nuisance parameters.

Acronym	channel	Event category
binch1	$\mu + \text{jets}$	exclusive loose
binch2	$\mu + \text{jets}$	exclusive medium
binch3	$\mu + \text{jets}$	exclusive tight
binch4	$e + \text{jets}$	exclusive loose
binch5	$e + \text{jets}$	exclusive medium
binch6	$e + \text{jets}$	exclusive tight

on the BR indicates which NP plays dominant role in the limit computation. To see the impact, each NP is individually varied by  $\pm 1\sigma$  of its pre-fit value and the corresponding change in the BR is computed. The distribution of post-fit pulls and impacts of nuisance parameters is shown in Figures 12.10, 12.11, 12.12, and 12.13, using  $m_{jj}$  from exclusive event categories based on c tagging for  $m_{H^+} = 120$  GeV from  $\ell + \text{jets}$  channel. In these figures,  $\theta_0$  is the pre-fit value,  $\widehat{\theta}$  is the post-fit value, and  $\Delta\theta = 1$  is the difference in the uncertainties from pre and post-fit values. The  $\widehat{BR}$  is the best-fit value of BR. The  $\Delta\widehat{BR}$  is the change in the value of  $\widehat{BR}$  when the pull of a nuisance parameter is varied  $\pm 1\sigma$  of its pre-fit value.

Larger the value of  $\Delta\widehat{BR}$ , larger is the correlation between BR and the corresponding nuisance parameter. Also, if the change in BR is positive (negative) when an NP is varied  $+1\sigma$  ( $-1\sigma$ ) of its pre-fit value then they are positively correlated. However, if the change in BR is opposite to the variation in NP then they are said to be anti-correlated. One of the most important feature of the plots shown in Figures 12.10, 12.11, 12.12, and 12.13 is the ordering (or indexing) of the NPs. The nuisance parameter having highest impact on the BR is placed at the top (index is 1) of the plot. Those placed at bottom have relatively small impact on BR thereby don't affect the limit. One another important feature in these plots is the  $\pm 1\sigma$  values of the NPs. If an NP has a smaller  $\pm 1\sigma$  value then it is said to be constrained. In such case the constraint on that NP has to be relieved by scrutinizing its pre-fit values. A large constraint on the NPs leads to larger mismatch between the expected and observed limits. The acronym used in these plots for statistical uncertainties are given in Table 12.5 for reference.

## 12.7. Goodness of fit

The goodness of the fit (GOF) indicates how well the model PDF (probability distribution function) agrees with the data. There are various algorithms to compute the GOF such as **saturated** (Baker-Cousins) [140],

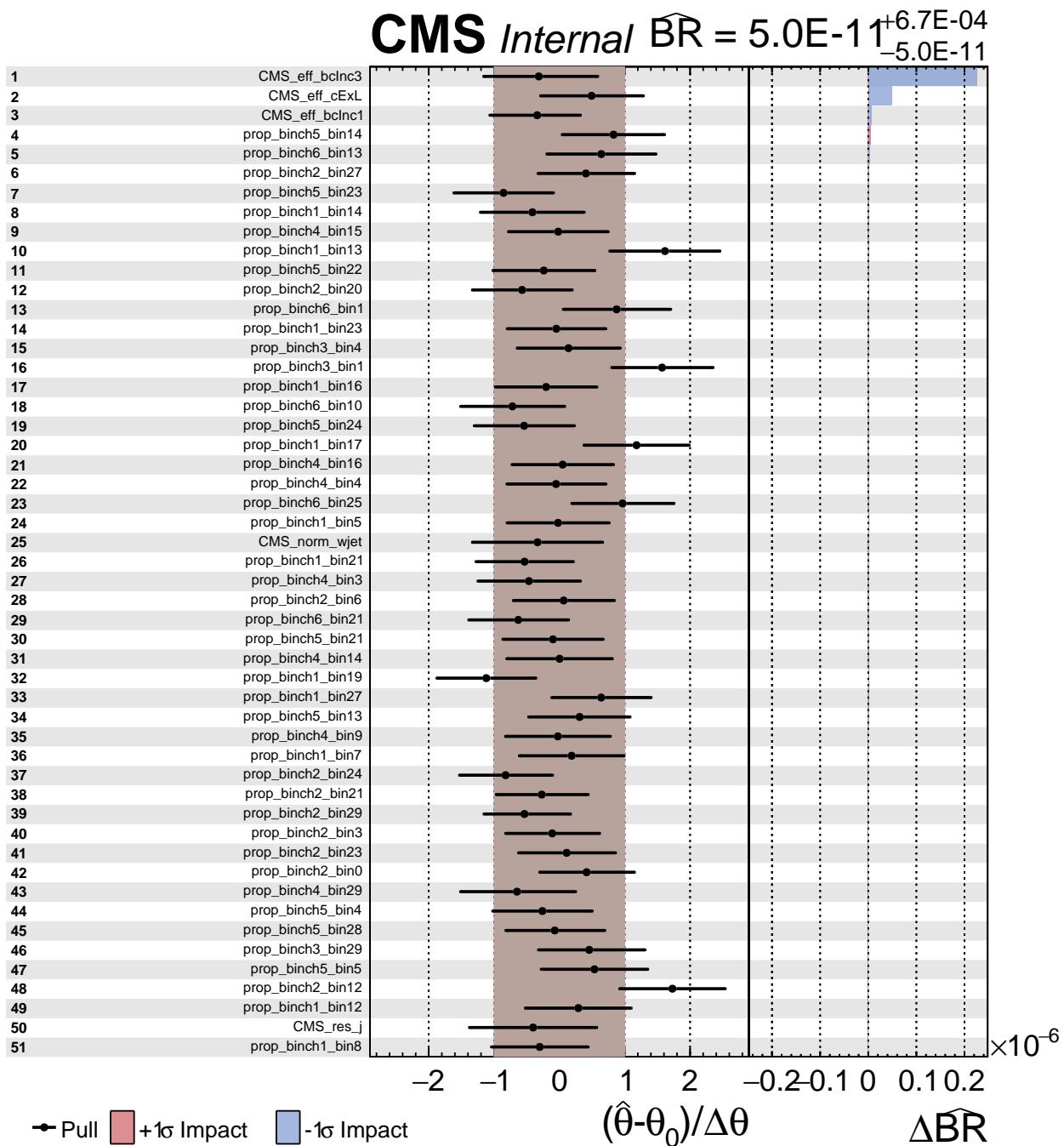


Figure 12.10.: Distribution of post-fit pulls and impacts of nuisance parameters from exclusive event categories based on charm-tagging for  $m_{H^+} = 120$  GeV from  $\ell + \text{jets}$  channel. Contd ...

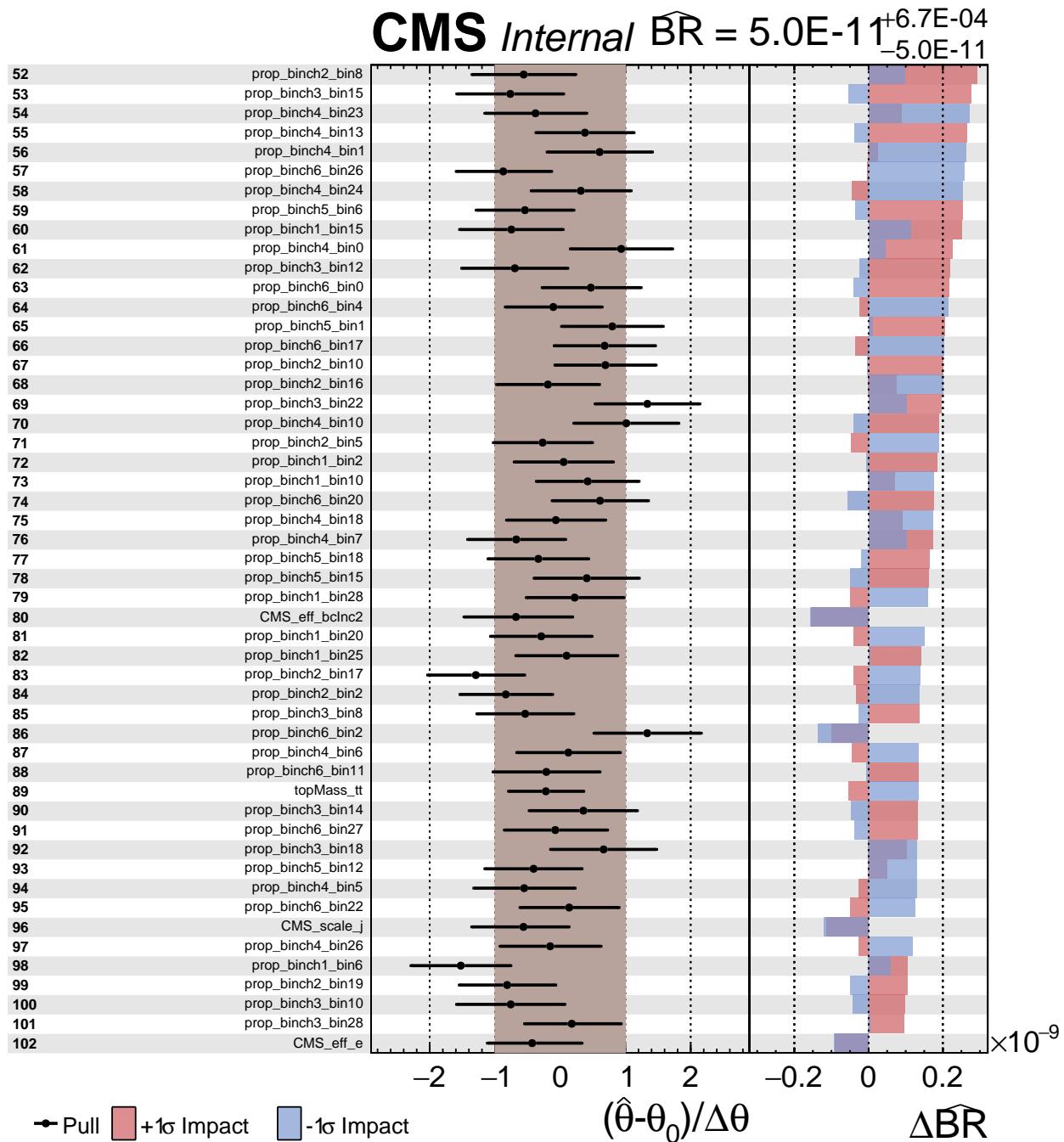


Figure 12.11.: Distribution of post-fit pulls and impacts of nuisance parameters from exclusive event categories based on charm-tagging for  $m_{H^+} = 120$  GeV from  $\ell + \text{jets}$  channel. Contd ...

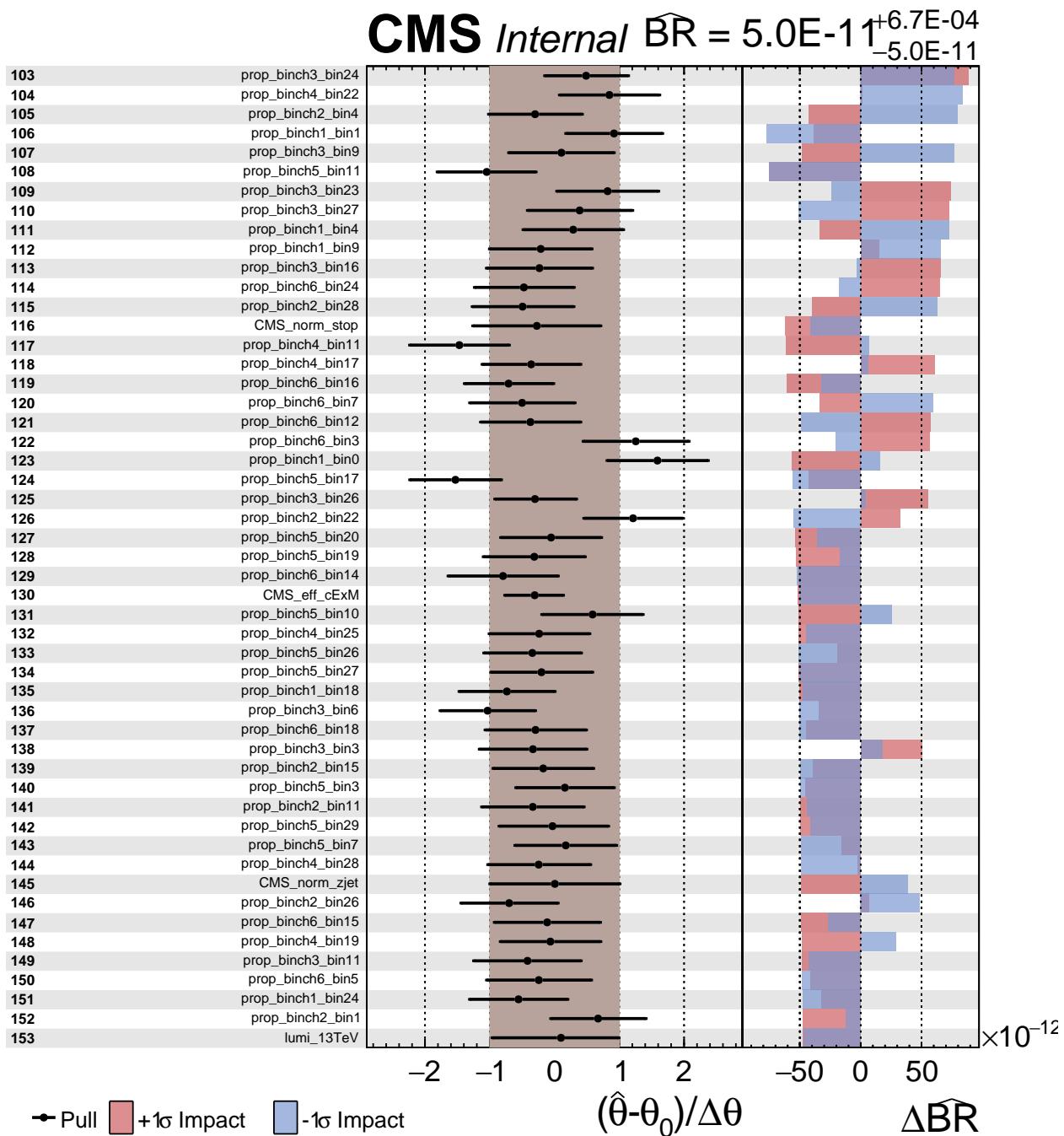


Figure 12.12.: Distribution of post-fit pulls and impacts of nuisance parameters from exclusive event categories based on charm-tagging for  $m_{H^+} = 120$  GeV from  $\ell + \text{jets}$  channel. Contd ...

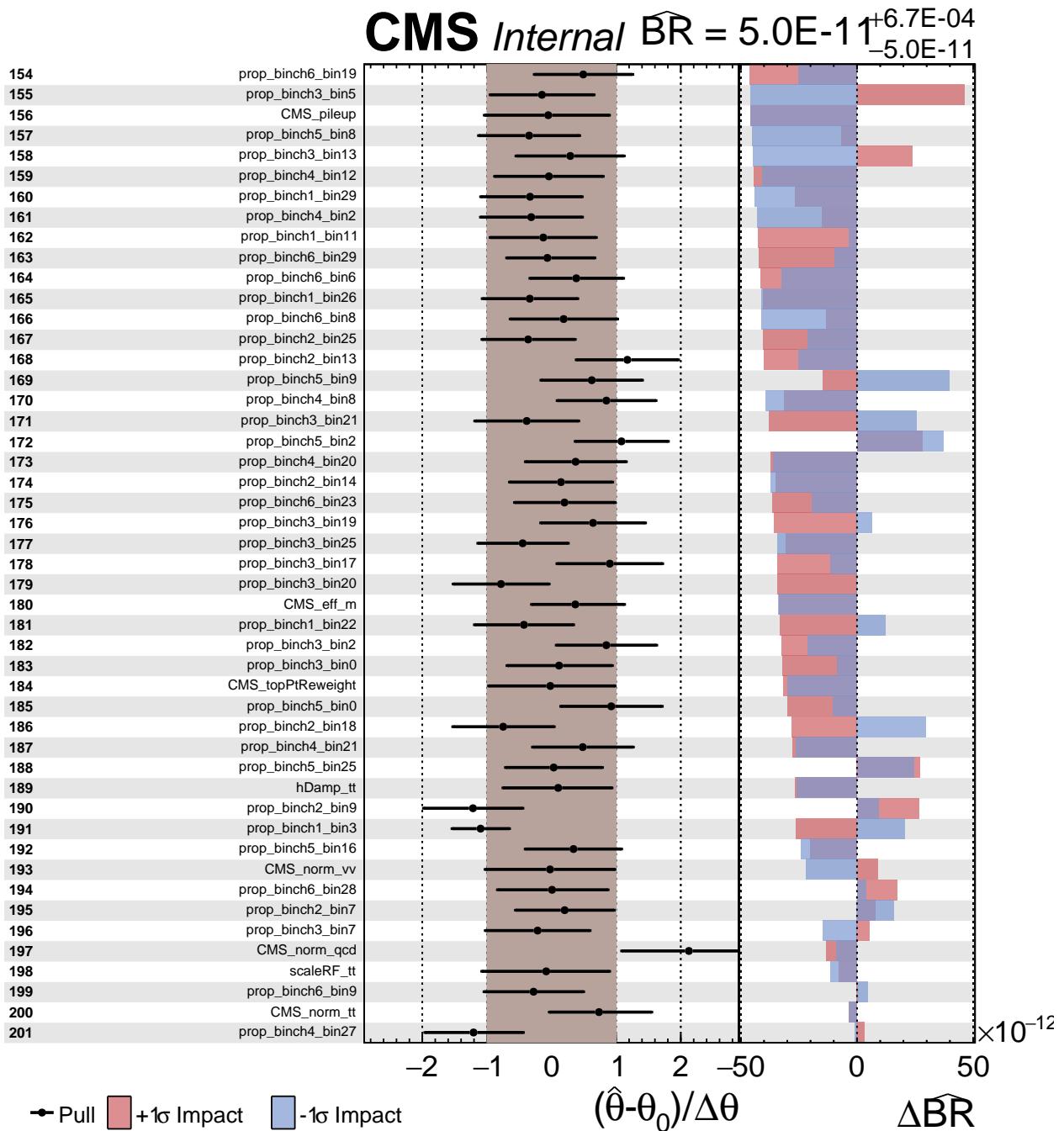


Figure 12.13.: Distribution of post-fit pulls and impacts of nuisance parameters from exclusive event categories based on charm-tagging for  $m_{H^+} = 120$  GeV from  $\ell + \text{jets}$  channel.

KS (Kolmogorov-Smirnov) [141], [142], and AD (Anderson-Darling) [143]. The saturated algorithm is used in this analysis for GOF. The test statistics as a measure of GOF is given as [140]

$$q_{\text{GoF, saturated}} = -2 \ln \left( \frac{L_{\text{nominal}}(n|\mu s(\theta) + b(\theta))}{L_{\text{saturated}}(n|n)} \right) \quad (12.3)$$

The likelihood functions of Equation (12.3) are given by

$$L_{\text{nominal}}(n|\mu s(\theta) + b(\theta)) = \prod_{j=1}^N \frac{(\mu s_j(\theta) + b_j(\theta))^{n_j}}{n_j!} \exp(-(\mu s_j(\theta) + b_j(\theta))) \quad (12.4)$$

and,

$$L_{\text{saturated}}(n|n) = \prod_{j=1}^N \frac{(n_j)^{n_j}}{n_j!} \exp(-n_j) \quad (12.5)$$

where  $n_j$  is the expectation value in  $j^{\text{th}}$  bin of data,  $N$  is the total number of bins. The  $s_j(\theta)$  and  $b_j(\theta)$  are the mean number of events in  $j^{\text{th}}$  bin of signal and background process and depend on the nuisance parameters ( $\theta$ ). The  $\mu$  is the signal strength which is zero for background only hypothesis and non-zero for signal+background hypothesis. In  $j^{\text{th}}$  bin,  $s_j(\theta)$  and  $b_j(\theta)$  are given by [138]

$$s_j(\theta) = s_{\text{tot}} \int_{j^{\text{th}} \text{bin}} f_s(x; \theta_s) dx, \quad (12.6)$$

$$b_j(\theta) = b_{\text{tot}} \int_{j^{\text{th}} \text{bin}} f_b(x; \theta_b) dx. \quad (12.7)$$

The  $f_s(x; \theta_s)$  and  $f_b(x; \theta_b)$  are PDFs for signal and background events. The  $s_{\text{tot}}$  and  $b_{\text{tot}}$  are mean of total number of events from signal and background process. Using Equation (12.4) and (12.5),

$$q_{\text{GoF, saturated}} = 2 \sum_j \left( \mu s_j(\theta) + b_j(\theta) - n_j + n_j \ln \left( \frac{n_j}{\mu s_j(\theta) + b_j(\theta)} \right) \right) \quad (12.8)$$

The goodness of fit ( $q_{\text{GoF, saturated}}$ ) is shown in Figure 12.14 for  $\ell + \text{jets}$  channel from exclusive charm categories for all signal mass points using observed data and generated toys. Lower the value of GOF, the better is the agreement between data and model PDF. For different channels and various event categories, the values of GOF for every category, all the mass of the charged Higgs, and all channels are shown in Tables 12.6, 12.7, and 12.8. If one normalizes Figure 12.14, then the area on the right hand side of the arrow (observed value of GOF) will correspond to the p-value.

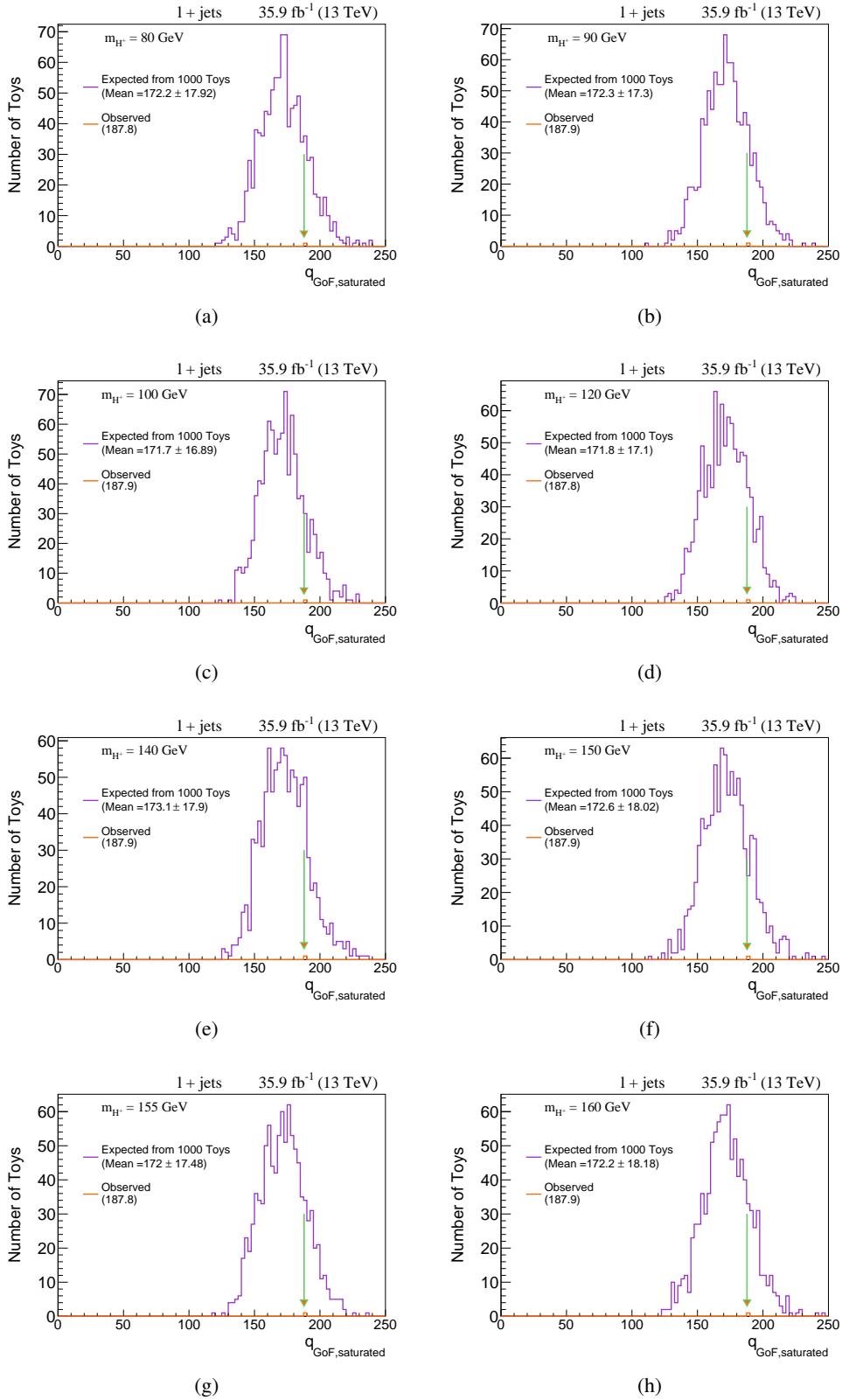
Figure 12.14.: Goodness of fit for  $\ell + \text{jets}$  from exclusive charm categories.

Table 12.6.: Goodness of fit for  $\mu + \text{jets}$  channel, from different event categories.

$m_{H^\pm}$ (GeV)	$m_{jj}(\text{Inc})$		$m_{jj}(\text{Inc CTagL})$		$m_{jj}(\text{Ex CTag})$	
	from toys	from data	from toys	from data	from toys	from data
80	$29.57 \pm 7.284$	49.22	$29.34 \pm 7.946$	52.86	$88.91 \pm 13.35$	108.9
90	$29.39 \pm 7.533$	49.22	$29.14 \pm 7.301$	52.86	$88.86 \pm 12.86$	108.9
100	$29.43 \pm 7.594$	49.22	$29.52 \pm 7.695$	52.86	$88.57 \pm 12.87$	108.9
120	$29.64 \pm 9.094$	49.22	$29.67 \pm 7.547$	52.86	$88.11 \pm 12.68$	108.9
140	$29.45 \pm 7.441$	49.22	$29.89 \pm 7.778$	52.86	$88.83 \pm 13.13$	108.9
150	$29.66 \pm 7.976$	49.22	$29.51 \pm 7.602$	52.86	$88.34 \pm 12.86$	108.9
155	$29.31 \pm 7.22$	49.22	$29.71 \pm 8.658$	52.86	$88.17 \pm 13.02$	108.9
160	$29.45 \pm 7.746$	49.22	$29.7 \pm 8.043$	52.86	$88.38 \pm 13.11$	108.9

Table 12.7.: Goodness of fit for  $e + \text{jets}$  channel, from different event categories.

$m_{H^\pm}$ (GeV)	$m_{jj}(\text{Inc})$		$m_{jj}(\text{Inc CTagL})$		$m_{jj}(\text{Ex CTag})$	
	from toys	from data	from toys	from data	from toys	from data
80	$29.58 \pm 7.663$	38.24	$29.15 \pm 7.537$	37.44	$88.16 \pm 12.64$	81.26
90	$29.9 \pm 7.966$	36.8	$29.21 \pm 7.647$	37.3	$87.68 \pm 12.23$	81.26
100	$29.24 \pm 7.653$	38.06	$29.61 \pm 7.781$	37.44	$87.8 \pm 12.83$	81.26
120	$29.22 \pm 7.394$	37.15	$29.18 \pm 7.542$	37.44	$87.66 \pm 13.02$	81.26
140	$29.74 \pm 7.669$	38.24	$29.3 \pm 7.607$	37.44	$87.68 \pm 12.94$	81.26
150	$29.79 \pm 7.717$	38.24	$29.43 \pm 7.645$	37.44	$87.81 \pm 12.66$	81.26
155	$29.31 \pm 7.612$	38.24	$29.88 \pm 8.07$	37.44	$88.07 \pm 12.6$	81.26
160	$29.1 \pm 7.687$	38.24	$29.57 \pm 7.951$	37.44	$87.92 \pm 12.51$	81.26

Table 12.8.: Goodness of fit for  $\ell + \text{jets}$  channel, from different event categories.

$m_{H^\pm}$ (GeV)	$m_{jj}(\text{Inc})$		$m_{jj}(\text{Inc CTagL})$		$m_{jj}(\text{Ex CTag})$	
	from toys	from data	from toys	from data	from toys	from data
80	$58.85 \pm 10.63$	87.27	$58.96 \pm 10.77$	89.81	$169.6 \pm 15.31$	187.8
90	$58.6 \pm 10.34$	87.27	$58.94 \pm 10.44$	89.81	$170.2 \pm 15.18$	187.9
100	$58.62 \pm 10.75$	87.27	$58.84 \pm 10.92$	89.81	$169.5 \pm 14.53$	187.9
120	$58.97 \pm 10.8$	87.27	$58.77 \pm 10.46$	89.81	$170.1 \pm 15.47$	187.8
140	$58.89 \pm 10.34$	87.27	$58.52 \pm 10.55$	89.79	$170.4 \pm 14.99$	187.9
150	$58.65 \pm 10.65$	87.27	$58.48 \pm 10.34$	89.79	$170 \pm 15.4$	187.9
155	$58.57 \pm 10.9$	87.27	$58.98 \pm 10.85$	89.79	$169.9 \pm 15.45$	187.8
160	$58.63 \pm 10.95$	87.27	$59.24 \pm 11.97$	89.79	$169.7 \pm 15.69$	187.9



# 13 | Summary and Future Outlook

In order to explain some failures of the standard model, the 2HDM postulated existence of an additional Higgs doublet. It leads to two charged and three neutral physical scalar particles. A charged Higgs boson decays to various channels depending on the parameter space corresponding to its mass and  $\tan\beta$ . A search for the charged Higgs in the  $H^+ \rightarrow c\bar{s}$  channel has been performed in this thesis with the data collected by the CMS experiment at the center-of-mass energy of 13 TeV with an integrated luminosity of  $35.9\text{ fb}^{-1}$ . The observed data and standard model predictions are in agreement within the statistical and systematic uncertainties. That is, no signal has been observed in the data for the mass of charged Higgs in the range 80 to 160 GeV. In the absence of an excess, a 95% CL upper limit has been put on  $\mathcal{B}(t \rightarrow H^+ b)$ , assuming  $\mathcal{B}(H^+ \rightarrow c\bar{s}) = 100\%$ . The kinematic fitting and a new technique which uses c tagging has been intensively exploited to improve the limits. The final limits from combined charm categories are shown in Tables 13.1, 13.2, and 13.3. The observed limits are in the range 0.29–2.12%, 0.27–3.29%, and 0.20–1.65% for  $\mu + \text{jets}$ ,  $e + \text{jets}$ , and  $\ell + \text{jets}$  channel, respectively. The observed limits at 13 TeV are better by an average factor of 7.6 as compared to the earlier CMS results at 8 TeV [63]. The improvement in the limit is partly due to higher luminosity and partly due the new technique used in this thesis.

As a continuation of the analysis presented in this thesis, here are a few points that can/will be done in the near future:

- As of now, the search for charged Higgs in  $\mathcal{B}(H^+ \rightarrow c\bar{s})$  channel at 13 TeV has not been made public from the ATLAS experiment. It would very interesting to compare our results with that from the ATLAS once they become available. Currently, we have analysed only a fraction of data recorded during Run-II (year 2016, 2017, and 2018). Analysing proton-proton collision data recorded by the CMS experiment from full Run-II, which correspond to an integrated luminosity of  $160\text{ fb}^{-1}$ , would put tighter constraints on  $\mathcal{B}(t \rightarrow H^+ b)$ . The new limit should be further constrained by about a factor of  $\sqrt{160/36} \approx 2$  if one goes by luminosity only. However, advanced techniques such as multivariate analysis, deep b/c tagging, etc, will further constrain the upper limits. Finally, combining the final limits with  $160\text{ fb}^{-1}$  data from the

CMS and ATLAS experiments will put even tighter constraints on  $\mathcal{B}(t \rightarrow H^+ b)$ .

- Since we already know the allowed region of  $\mathcal{B}(t \rightarrow H^+ b)$  at 13 TeV with  $35.9 \text{ fb}^{-1}$  data, we can use these experimental results to see what value of  $\tan\beta$  is allowed for a different mass of the charged Higgs. Note that for the signal hypothesis, during the generation of simulated signal samples, we have not put any constraint on the value of  $\tan\beta$ . We have kept it floating so that we can restrict it once we know the limit on  $\mathcal{B}(t \rightarrow H^+ b)$ . However, in the context of the 2HDM, we will have to restrict ourselves to those Types where  $\mathcal{B}(H^+ \rightarrow c\bar{s})$  is close to 100%. For this, we have to interpret the current experimental results in the Type-II for lower values of  $\tan\beta$ . Using the formula for  $\mathcal{B}(t \rightarrow H^+ b)$ , given in Equation (2.7), we can impose a limit on  $\tan\beta$ . In fact, such a phenomenological study has already been carried out using the  $\tau\nu, t\bar{b}$  channels [144], where they have used HDECAY package [145] to generate theoretical values of the branching fractions for different charged Higgs masses. We hope a similar study can/will be carried out for the  $c\bar{s}$  channel also.
- So far in the experimental analysis we have considered  $\mathcal{B}(H^+ \rightarrow c\bar{s}) = 100\%$  and put upper limits on  $\mathcal{B}(t \rightarrow H^+ b)$  at 95% CL. However, we can relax this assumption and do a similar analysis where we will put limits on  $x = \mathcal{B}(t \rightarrow H^+ b) \times \mathcal{B}(H^+ \rightarrow c\bar{s})$ , where  $\mathcal{B}(H^+ \rightarrow c\bar{s}) \leq 1$ . For this, we will have to modify the Equation (12.2) so that the signal and background process are scaled appropriately under the new parameter of interest. The relaxation of the present assumption ( $\mathcal{B}(H^+ \rightarrow c\bar{s}) = 100\%$ ) would be interesting for the phenomenology of charged Higgs in models like the 2HDM and MSSM. As the Higgs sector of the MSSM corresponds to the Type-II of the 2HDM, and as shown in Figure 2.2, the  $\mathcal{B}(H^+ \rightarrow c\bar{s})$  is dominant (i.e. close to 100%) only when  $\tan\beta < 1$ . In Type-X also, we see a similar trend. For the other two types of 2HDM (Type-I and Y), the  $\mathcal{B}(H^+ \rightarrow c\bar{s})$  could not be realised for any choice of  $\tan\beta$ . However, relaxing the assumption  $\mathcal{B}(H^+ \rightarrow c\bar{s}) = 100\%$  would definitely help in imposing limits on  $m_{H^+}$ – $\tan\beta$  parameter space for any choice of  $\tan\beta$  for all types of 2HDM and MSSM.
- Currently, different channels such as  $\tau^+\nu_\tau, c\bar{s}, t\bar{b}, c\bar{b}$ , etc of the charged Higgs decays are examined separately. It would be very interesting if all results from all the channels are combined and a final limit on  $\mathcal{B}(t \rightarrow H^+ b)$  is set under the assumption that the charged Higgs decays to everything ( $H^+ \rightarrow q\bar{q}$  and  $l^+\nu_l$ ).

Table 13.1.: 95% CL exclusion limits in % for the  $\mu +$  jets from combined c tagging categories.

$m_{H^+}$ (GeV)	Expected					Observed
	$-2\sigma$	$-1\sigma$	median	$+1\sigma$	$+2\sigma$	
80	1.62	2.18	3.06	4.31	5.85	2.12
90	0.72	0.96	1.34	1.87	2.49	0.79
100	0.36	0.49	0.68	0.94	1.25	0.38
120	0.25	0.33	0.46	0.64	0.86	0.29
140	0.22	0.30	0.41	0.57	0.75	0.33
150	0.21	0.28	0.39	0.54	0.72	0.37
155	0.22	0.30	0.41	0.58	0.77	0.45
160	0.23	0.31	0.44	0.62	0.84	0.44

Table 13.2.: 95% CL exclusion limits in % for the  $e +$  jets from combined c tagging categories.

$m_{H^+}$ (GeV)	Expected					Observed
	$-2\sigma$	$-1\sigma$	median	$+1\sigma$	$+2\sigma$	
80	1.97	2.63	3.66	5.14	6.89	3.29
90	0.80	1.07	1.48	2.07	2.77	1.45
100	0.43	0.56	0.78	1.09	1.45	0.61
120	0.29	0.38	0.53	0.73	0.97	0.53
140	0.24	0.32	0.45	0.62	0.83	0.35
150	0.24	0.31	0.44	0.61	0.81	0.27
155	0.23	0.31	0.43	0.61	0.81	0.27
160	0.26	0.35	0.49	0.69	0.94	0.33

Table 13.3.: 95% CL exclusion limits in % for the  $\ell +$  jets from combined c tagging categories.

$m_{H^+}$ (GeV)	Expected					Observed
	$-2\sigma$	$-1\sigma$	median	$+1\sigma$	$+2\sigma$	
80	1.33	1.77	2.46	3.45	4.63	1.65
90	0.56	0.74	1.03	1.43	1.90	0.68
100	0.29	0.38	0.53	0.73	0.98	0.30
120	0.20	0.26	0.36	0.51	0.67	0.25
140	0.18	0.23	0.32	0.45	0.59	0.21
150	0.17	0.22	0.31	0.43	0.57	0.20
155	0.17	0.22	0.31	0.44	0.58	0.22
160	0.18	0.24	0.34	0.47	0.64	0.25



# Bibliography

- [1] Ravindra Kumar Verma. <http://inspirehep.net/author/profile/R.K.Verma.1>.
- [2] What Is The Smallest Particle We Know? <https://www.scienceabc.com/nature/universe/what-is-the-smallest-particle-we-know.html>.
- [3] Standard Model. [https://en.wikipedia.org/wiki/Standard\\_Model](https://en.wikipedia.org/wiki/Standard_Model).
- [4] Bora Isildak. *Measurement of the differential dijet production cross section in proton-proton collisions at  $\sqrt{s} = 7$  tev.* PhD thesis, Bogazici U., 2011.
- [5] Mayumi Aoki, Shinya Kanemura, Koji Tsumura, and Kei Yagyu. Models of yukawa interaction in the two higgs doublet model, and their collider phenomenology. *Phys. Rev. D*, 80:015017, Jul 2009.
- [6] Cecilia E. Gerber and Costas Vellidis. Review of Physics Results from the Tevatron: Top Quark Physics. *Submitted to: Int. J. Mod. Phys. A*, 2014.
- [7] Luca Fiorini. Top-Quark Physics Results From LHC. In *Helmholtz Alliance Linear Collider Forum: Proceedings of the Workshops Hamburg, Munich, Hamburg 2010-2012, Germany*, pages 38–45, Hamburg, 2013. DESY, DESY.
- [8] Wikimedia Commons. File:cern-accelerator-complex.svg — wikimedia commons, the free media repository, 2019. [Online; accessed 1-August-2019].
- [9] R. Bruce, R.W. Assmann, V. Boccone, G. Bregliozi, H. Burkhardt, F. Cerutti, A. Ferrari, M. Huhtinen, A. Lechner, Y. Levinson, A. Mereghetti, N.V. Mokhov, I.S. Tropin, and V. Vlachoudis. Sources of machine-induced background in the atlas and cms detectors at the cern large hadron collider. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 729:825 – 840, 2013.
- [10] R. Schmidt. Machine Protection and Interlock Systems for Circular Machines - Example for LHC. (arXiv:1608.03087):23 p, Aug 2016. Comments: 23 pages, contribution to the 2014 Joint International Accelerator School: Beam Loss and Accelerator Protection, Newport Beach, CA, USA , 5-14 Nov 2014.
- [11] Introduction to LHC. <https://indico.cern.ch/event/386181/contributions/1816638/attachments/1143882/1639362/20150825-CMSDAS-LHC-v3.pdf>.
- [12] LHC Report: playing with angles. <https://home.cern/news/news/accelerators/lhc-report-playing-angles>.
- [13] Werner Herr and B Muratori. Concept of luminosity. 2006.
- [14] Public CMS Luminosity Information. <https://twiki.cern.ch/twiki/bin/view/>

[CMS Public/Lumi Public Results](#).

- [15] CMS Wiki Pages. [https://wiki.physik.uzh.ch/cms/latex:example\\_spherical\\_coordinates](https://wiki.physik.uzh.ch/cms/latex:example_spherical_coordinates).
- [16] Thomas Lenzi. Development and Study of Different Muon Track Reconstruction Algorithms for the Level-1 Trigger for the CMS Muon Upgrade with GEM Detectors. Master's thesis, U. Brussels (main), 2013.
- [17] The CMS Collaboration and S Chatrchyan et. all. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08004–S08004, aug 2008.
- [18] Drawing of the CMS experiment. <http://phenix.vanderbilt.edu/>.
- [19] A.M. et al. Sirunyan. Particle-flow reconstruction and global event description with the CMS detector. Particle-flow reconstruction and global event description with the CMS detector. *JINST*, 12(CMS-PRF-14-001. CMS-PRF-14-001-004. 10):P10003. 82 p, Jun 2017. Replaced with the published version. Added the journal reference and DOI. All the figures and tables can be found at <http://cms-results.web.cern.ch/cms-results/public-results/publications/PRF-14-001> (CMS Public Pages).
- [20] High Energy Physics Department of Firenze. Silicon Detector.gif. [http://hep.fi.infn.it/CMS/sensors/Silicon\\_Detector.gif](http://hep.fi.infn.it/CMS/sensors/Silicon_Detector.gif).
- [21] Serguei Chatrchyan et al. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *JINST*, 9(10):P10009, 2014.
- [22] ECAL Preshower removed from detector for repairs. <http://cms.web.cern.ch/news/ecal-preshower-removed-detector-repairs>.
- [23] Serguei Chatrchyan et al. Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in  $pp$  Collisions at  $\sqrt{s} = 7$  TeV. *JINST*, 8:P09009, 2013. [JINST8,9009(2013)].
- [24] Wikimedia Commons. File:pistates.svg — wikimedia commons, the free media repository, 2015. [Online; accessed 1-August-2019].
- [25] J Mans, J Anderson, B Dahmes, P de Barbaro, J Freeman, T Grassi, E Hazen, J Mans, R Ruchti, I Schimdt, T Shaw, C Tully, J Whitmore, and T Yetkin. CMS Technical Design Report for the Phase 1 Upgrade of the Hadron Calorimeter. Technical Report CERN-LHCC-2012-015. CMS-TDR-10, Sep 2012. Additional contact persons: Jeffrey Spalding, Fermilab, spalding@cern.ch, Didier Contardo, Universite Claude Bernard-Lyon I, contardo@cern.ch.
- [26] CMS-solenoid. [http://irfu.cea.fr/en/Phocea/Vie\\_des\\_labos/Ast/ast\\_sstechnique.php?id\\_ast=2261](http://irfu.cea.fr/en/Phocea/Vie_des_labos/Ast/ast_sstechnique.php?id_ast=2261).
- [27] A. M. Sirunyan et al. Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at  $\sqrt{s} = 13$  TeV. *JINST*, 13:P06015, 2018.
- [28] Vardan Khachatryan et al. The CMS trigger system. *JINST*, 12(01):P01020, 2017.
- [29] Computing Model. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookComputingModel>.
- [30] CMSSW Application Framework. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookCMSSWFframework>.
- [31] Top group. Measurement of jet activity in ttbar events using lepton + jets final states in proton-proton

- collisions at 13 tev. *CMS Analysis Note*, CMS AN-17-003, 2017.
- [32] Top group. Measurement of the top quark decay width in the dilepton channel at  $\sqrt{s} = 13$  tev with cms. *CMS Analysis Note*, CMS AN-17-116, 2017.
- [33] Cristina Ferro. B-tagging in CMS. *EPJ Web Conf.*, 28:12055, 2012.
- [34] CMS Collaboration. Identification of c-quark jets at the CMS experiment. CMS Physics Analysis Summary CMS-PAS-BTV-16-001, 2016.
- [35] A. M. Sirunyan et al. Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV. *JINST*, 13(05):P05011, 2018.
- [36] P. A. R. Ade et al. Planck 2013 results. XVI. Cosmological parameters. *Astron. Astrophys.*, 571:A16, 2014.
- [37] P. A. R. Ade et al. Planck 2013 results. I. Overview of products and scientific results. *Astron. Astrophys.*, 571:A1, 2014.
- [38] David J Griffiths. *Introduction to elementary particles; 2nd rev. version.* Physics textbook. Wiley, New York, NY, 2008.
- [39] Emmy Noether. Invariant Variation Problems. *Gott. Nachr.*, 1918:235–257, 1918. [Transp. Theory Statist. Phys.1,186(1971)].
- [40] Jeffrey Goldstone, Abdus Salam, and Steven Weinberg. Broken symmetries. *Phys. Rev.*, 127:965–970, Aug 1962.
- [41] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964.
- [42] P.W. Higgs. Broken symmetries, massless particles and gauge fields. *Physics Letters*, 12(2):132 – 133, 1964.
- [43] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.
- [44] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global conservation laws and massless particles. *Phys. Rev. Lett.*, 13:585–587, Nov 1964.
- [45] Peter W. Higgs. Spontaneous symmetry breakdown without massless bosons. *Phys. Rev.*, 145:1156–1163, May 1966.
- [46] T. W. B. Kibble. Symmetry breaking in non-abelian gauge theories. *Phys. Rev.*, 155:1554–1561, Mar 1967.
- [47] Nicola Cabibbo. Unitary symmetry and leptonic decays. *Phys. Rev. Lett.*, 10:531–533, Jun 1963.
- [48] Makoto Kobayashi and Toshihide Maskawa. CP-Violation in the Renormalizable Theory of Weak Interaction. *Progress of Theoretical Physics*, 49(2):652–657, 02 1973.
- [49] Georges Aad et al. Combined Measurement of the Higgs Boson Mass in  $pp$  Collisions at  $\sqrt{s} = 7$  and 8 TeV with the ATLAS and CMS Experiments. *Phys. Rev. Lett.*, 114:191803, 2015.
- [50] Stephen P. Martin. A Supersymmetry primer. pages 1–98, 1997. [Adv. Ser. Direct. High Energy Phys.18,1(1998)].

- [51] V. M. Abazov et al. Precision measurement of the ratio  $\mathcal{B}(t \rightarrow wb)/\mathcal{B}(t \rightarrow wq)$  and extraction of  $V_{tb}$ . *Phys. Rev. Lett.*, 107:121802, Sep 2011.
- [52] D. Acosta et al. Measurement of  $\mathcal{B}(t \rightarrow wb)/\mathcal{B}(t \rightarrow wq)$  at the collider detector at fermilab. *Phys. Rev. Lett.*, 95:102002, Aug 2005.
- [53] Vardan Khachatryan et al. Measurement of the ratio  $\mathcal{B}(t \rightarrow Wb)/\mathcal{B}(t \rightarrow Wq)$  in pp collisions at  $\sqrt{s} = 8$  TeV. *Phys. Lett.*, B736:33–57, 2014.
- [54] Abdesslam Arhrib, Rachid Benbrik, Hicham Harouiz, Stefano Moretti, and Abdessamat Rouchad. A Guidebook to Hunting Charged Higgs Bosons at the LHC. 2018.
- [55] A. G. Akeroyd et al. Prospects for charged Higgs searches at the LHC. *Eur. Phys. J.*, C77(5):276, 2017.
- [56] P. Achard et al. Search for charged Higgs bosons at LEP. *Phys. Lett. B*, 575:208, 2003.
- [57] A. Heister et al. Search for charged Higgs bosons in  $e^+e^-$  collisions at energies up to  $\sqrt{s} = 209$  GeV. *Phys. Lett. B*, 543:1, 2002.
- [58] J. Abdallah et al. Search for charged higgs bosons at LEP in general two Higgs doublet models. *Eur. Phys. J. C*, 34:399, 2004.
- [59] G. Abbiendi et al. Search for charged Higgs bosons in  $e^+e^-$  collisions at  $\sqrt{s} = 189 - 209$  GeV. *Eur. Phys. J. C*, 72:2076, 2012.
- [60] T. Aaltonen et al. Search for charged Higgs bosons in decays of top quarks in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV. *Phys. Rev. Lett.*, 103:101803, 2009.
- [61] V.M. Abazov et al. Search for charged Higgs bosons in top quark decays. *Phys. Lett. B*, 682:278, 2009.
- [62] Georges Aad et al. Search for a light charged Higgs boson in the decay channel  $H^+ \rightarrow c\bar{s}$  in  $t\bar{t}$  events using pp collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector. *Eur. Phys. J. C*, 73:2465, 2013.
- [63] Vardan Khachatryan et al. Search for a light charged Higgs boson decaying to  $c\bar{s}$  in pp collisions at  $\sqrt{s} = 8$  TeV. *JHEP*, 12:178, 2015.
- [64] Georges Aad et al. Search for charged Higgs bosons through the violation of lepton universality in  $t\bar{t}$  events using  $pp$  collision data at  $\sqrt{s} = 7$  TeV with the ATLAS experiment. *JHEP*, 03:076, 2013.
- [65] Georges Aad et al. Search for charged Higgs bosons decaying via  $H^\pm \rightarrow \tau^\pm\nu$  in fully hadronic final states using  $pp$  collision data at  $\sqrt{s} = 8$  TeV with the ATLAS detector. *JHEP*, 03:088, 2015.
- [66] Morad Aaboud et al. Search for charged Higgs bosons produced in association with a top quark and decaying via  $H^\pm \rightarrow \tau\nu$  using  $pp$  collision data recorded at  $\sqrt{s} = 13$  TeV by the ATLAS detector. *Phys. Lett.*, B759:555–574, 2016.
- [67] The ATLAS collaboration. Search for charged Higgs bosons in the  $\tau+\text{jets}$  final state using  $14.7\text{ fb}^{-1}$  of pp collision data recorded at  $\sqrt{s} = 13$  TeV with the ATLAS experiment. 2016.
- [68] Morad Aaboud et al. Search for charged Higgs bosons decaying via  $H^\pm \rightarrow \tau^\pm\nu_\tau$  in the  $\tau+\text{jets}$  and  $\tau+\text{lepton}$  final states with  $36\text{ fb}^{-1}$  of  $pp$  collision data recorded at  $\sqrt{s} = 13$  TeV with the ATLAS experiment. *JHEP*, 09:139, 2018.
- [69] Albert M Sirunyan et al. Search for a charged Higgs boson decaying to charm and bottom quarks in

- proton-proton collisions at  $\sqrt{s} = 8$  TeV. *JHEP*, 11:115, 2018.
- [70] Albert M Sirunyan et al. Search for charged Higgs bosons in the  $H^\pm \rightarrow \tau^\pm \nu_\tau$  decay channel in proton-proton collisions at  $\sqrt{s} = 13$  TeV. 2019.
- [71] CMS Collaboration. Search for a charged Higgs boson decaying into top and bottom quarks in proton-proton collisions at 13TeV in events with electrons or muons. 2019.
- [72] The ATLAS Collaboration and G Aad et. all. The ATLAS experiment at the CERN large hadron collider. *Journal of Instrumentation*, 3(08):S08003–S08003, aug 2008.
- [73] Georges Aad et al. Observation of a new particle in the search for the standard model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716:1, 2012.
- [74] The LHCf Collaboration and O Adriani et. all. The LHCf detector at the CERN large hadron collider. *Journal of Instrumentation*, 3(08):S08006–S08006, aug 2008.
- [75] The ALICE Collaboration and K Aamodt et. all. The ALICE experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08002–S08002, aug 2008.
- [76] K. Aamodt et al. Two-pion Bose-Einstein correlations in central Pb-Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV. *Phys. Lett.*, B696:328–337, 2011.
- [77] The TOTEM Collaboration and G Anelli et. all. The TOTEM experiment at the CERN large hadron collider. *Journal of Instrumentation*, 3(08):S08007–S08007, aug 2008.
- [78] The LHCb Collaboration and A Augusto Alves et. all. The LHCb detector at the LHC. *Journal of Instrumentation*, 3(08):S08005–S08005, aug 2008.
- [79] R Aaij et al. First Evidence for the Decay  $B_s^0 \rightarrow \mu^+ \mu^-$ . *Phys. Rev. Lett.*, 110(2):021801, 2013.
- [80] B. Acharya et al. The Physics Programme Of The MoEDAL Experiment At The LHC. *Int. J. Mod. Phys.*, A29:1430050, 2014.
- [81] S Baird. Accelerators for pedestrians; rev. version. Technical Report AB-Note-2007-014. CERN-AB-Note-2007-014. PS-OP-Note-95-17-Rev-2. CERN-PS-OP-Note-95-17-Rev-2, CERN, Geneva, Feb 2007.
- [82] CMS luminosity measurement for the 2015 data-taking period. Technical Report CMS-PAS-LUM-15-001, CERN, Geneva, 2017.
- [83] Lecture 7 - Rapidity and Pseudorapidity. [http://www.hep.shef.ac.uk/edaw/PHY206/Site/2012\\_course\\_files/phy206rlec7.pdf](http://www.hep.shef.ac.uk/edaw/PHY206/Site/2012_course_files/phy206rlec7.pdf).
- [84] CMS Information. [https://cmsinfo.web.cern.ch/cmsinfo/Media/Publications/CMStimes/2006/11\\_06/index.html](https://cmsinfo.web.cern.ch/cmsinfo/Media/Publications/CMStimes/2006/11_06/index.html).
- [85] S Chatrchyan et al. Commissioning and Performance of the CMS Pixel Tracker with Cosmic Ray Muons. *JINST*, 5:T03007, 2010.
- [86] Viktor Vespremi. Radiation experience with the CMS pixel detector. *JINST*, 10(04):C04039, 2015.
- [87] Electromagnetic Calorimeter. <http://cms.web.cern.ch/news/electromagnetic-calorimeter>.
- [88] Electronic Light Detectors: Photomultipliers. <http://olympus.magnet.fsu.edu/primer/techniques/confocal/pmtintro.html>.

- [89] Detectors: Guideposts on the Road to Selection. [https://www.photonics.com/Articles/Detectors\\_Guideposts\\_on\\_the\\_Road\\_to\\_Selection/a25535](https://www.photonics.com/Articles/Detectors_Guideposts_on_the_Road_to_Selection/a25535).
- [90] <https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEffsRun2>. CMS Twiki.
- [91] Electron Triggers Efficiencies with 2016 data. [https://indico.cern.ch/event/570560/contributions/2389651/attachments/1380332/2098249/EGMTriggers\\_Approval\\_v2.pdf](https://indico.cern.ch/event/570560/contributions/2389651/attachments/1380332/2098249/EGMTriggers_Approval_v2.pdf).
- [92] Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET. Technical Report CMS-PAS-PFT-09-001, CERN, Geneva, Apr 2009.
- [93] Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector. Technical Report CMS-PAS-PFT-10-001, CERN, 2010.
- [94] Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer. MadGraph 5: Going beyond. *JHEP*, 06:128, 2011.
- [95] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014.
- [96] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with parton shower simulations: the POWHEG method. *JHEP*, 11:070, 2007.
- [97] Paolo Nason. A new method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004.
- [98] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06:043, 2010.
- [99] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 05:026, 2006.
- [100] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. A brief introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178:852, 2008.
- [101] Investigations of the impact of the parton shower tuning in Pythia 8 in the modelling of  $t\bar{t}$  at  $\sqrt{s} = 8$  and 13 TeV. Technical Report CMS-PAS-TOP-16-021, CERN, Geneva, 2016.
- [102] S. Agostinelli et al. GEANT4—a simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250, 2003.
- [103] Richard D. Ball et al. Parton distributions for the LHC Run II. *JHEP*, 04:040, 2015.
- [104] M. Beneke, P. Falgari, S. Klein, and C. Schwinn. Hadronic top-quark pair production with NNLL threshold resummation. *Nucl. Phys. B*, 855:695, 2012.
- [105] M. Aliev, H. Lacker, U. Langenfeld, S. Moch, P. Uwer, and M. Wiedermann. HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR. *Comput. Phys. Commun.*, 182:1034–1046, 2011.
- [106] P. Kant, O. M. Kind, T. Kintscher, T. Lohse, T. Martini, S. Mäßlitz, P. Rieck, and P. Uwer. HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions. *Comput. Phys. Commun.*, 191:74–89, 2015.
- [107] Johan Alwall et al. Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions. *Eur. Phys. J. C*, 53:473–500, 2008.

- [108] J. Beringer et al. Review of Particle Physics (RPP). *Phys. Rev.*, D86:010001, 2012.
- [109] W. Erdmann. and others. Offline primary vertex reconstruction with deterministic annealing clustering. *CMS Internal Note*, IN-11-014, 2011.
- [110] K. Rose. Deterministic annealing for clustering, compression, classification, regression, and related optimization problems. *Proceedings of the IEEE*, 86(11):2210–2239, Nov 1998.
- [111] R Fruhwirth, Wolfgang Waltenberger, and Pascal Vanlaer. Adaptive Vertex Fitting. Technical Report CMS-NOTE-2007-008, CERN, Geneva, Mar 2007.
- [112] Matteo Cacciari and Gavin P. Salam. Dispelling the  $N^3$  myth for the  $k_t$  jet-finder. *Phys. Lett.*, B641:57–61, 2006.
- [113] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet user manual. *Eur. Phys. J. C*, 72:1896, 2012.
- [114] Matteo Cacciari and Gavin P. Salam. Pileup subtraction using jet areas. *Phys. Lett.*, B659:119–126, 2008.
- [115] R. Fruhwirth. Application of kalman filtering to track and vertex fitting. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 262(2):444 – 450, 1987.
- [116] Serguei Chatrchyan et al. Performance of CMS muon reconstruction in  $pp$  collision events at  $\sqrt{s} = 7$  TeV. *JINST*, 7:P10002, 2012.
- [117] J. Goh et al. CMS RPC tracker muon reconstruction. *JINST*, 9(10):C10027, 2014.
- [118] A. Bodek, A. van Dyne, J. Y. Han, W. Sakumoto, and A. Strelnikov. Extracting Muon Momentum Scale Corrections for Hadron Collider Experiments. *Eur. Phys. J.*, C72:2194, 2012.
- [119] Muon efficiencies and scale factors. <https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonWorkInProgressAndPagResults>.
- [120] Vardan Khachatryan et al. Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at  $\sqrt{s} = 8$  TeV. *JINST*, 10:P06005, 2015.
- [121] Electron efficiencies and scale factors. [https://twiki.cern.ch/twiki/bin/view/CMS/EgammaIDRecipesRun2#Efficiencies\\_and\\_scale\\_factors](https://twiki.cern.ch/twiki/bin/view/CMS/EgammaIDRecipesRun2#Efficiencies_and_scale_factors).
- [122] Alexander M. Polyakov. Quark confinement and topology of gauge groups. *Nucl. Phys. B*, 120:429, 1977.
- [123] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti- $k_t$  jet clustering algorithm. *JHEP*, 04:063, 2008.
- [124] Pileup Removal Algorithms. Technical Report CMS-PAS-JME-14-001, CERN, Geneva, 2014.
- [125] Serguei Chatrchyan et al. Identification of b-quark jets with the CMS experiment. *JINST*, 8:P04013, 2013.
- [126] Identification of b quark jets at the CMS Experiment in the LHC Run 2. Technical Report CMS-PAS-BTV-15-001, CERN, Geneva, 2016.
- [127] Wolfgang Waltenberger. Adaptive Vertex Reconstruction. Technical Report CMS-NOTE-2008-033, CERN, Geneva, Jul 2008.

- [128] Vardan Khachatryan et al. Measurement of  $B\bar{B}$  Angular Correlations based on Secondary Vertex Reconstruction at  $\sqrt{s} = 7$  TeV. *JHEP*, 03:136, 2011.
- [129] Vardan Khachatryan et al. Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV. *JINST*, 12(02):P02014, 2017.
- [130] Methods to apply b-tagging efficiency scale factors. <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BTagSFMethods>.
- [131] MET Filter Recommendations for Run II. [https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETOptionalFiltersRun2#2016\\_data](https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETOptionalFiltersRun2#2016_data).
- [132] Jorgen D'Hondt, Steven Lowette, O L BuchmÃŒller, Susanna Cucciarelli, Frank-Peter Schilling, Maria Spiropulu, S Pakhtinat-Mehdiabadi, Daniele Benedetti, and Luc Pape. Fitting of Event Topologies with External Kinematic Constraints in CMS. Technical Report CMS-NOTE-2006-023, CERN, Geneva, Jan 2006.
- [133] Application of kinematic fitting to top quark mass reconstruction in the  $\mu + \text{jets}$  channel at  $\sqrt{s} = 7$  TeV with the CMS detector. [https://lib.ugent.be/fulltxt/RUG01/001/892/467/RUG01-001892467\\_2012\\_0001\\_AC.pdf](https://lib.ugent.be/fulltxt/RUG01/001/892/467/RUG01-001892467_2012_0001_AC.pdf).
- [134] CMS Collaboration. CMS luminosity measurements for the 2016 data taking period. CMS Physics Analysis Summary CMS-PAS-LUM-17-001, 2017.
- [135] Measurement of the differential cross section for  $t\bar{t}$  production in the dilepton final state at  $\sqrt{s} = 13$  TeV. Technical Report CMS-PAS-TOP-16-011, CERN, Geneva, 2016.
- [136] Measurement of the inclusive and differential  $t\bar{t}$  production cross sections in lepton + jets final states at 13 TeV. Technical Report CMS-PAS-TOP-16-008, CERN, Geneva, 2016.
- [137] Thomas Junk. Confidence level computation for combining searches with small statistics. *Nucl. Instrum. Meth. A*, 434:435, 1999.
- [138] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J. C*, 71:1554, 2011.
- [139] Procedure for the LHC Higgs boson search combination in summer 2011. 2011.
- [140] Steve Baker and Robert D. Cousins. Clarification of the Use of Chi Square and Likelihood Functions in Fits to Histograms. *Nucl. Instrum. Meth.*, 221:437–442, 1984.
- [141] A. Kolmogorov. Sulla determinazione empirica di una legge di distribuzione. *G. Ist. Ital. Attuari.*, 4:83–91, 1933.
- [142] N. Smirnov. Table for estimating the goodness of fit of empirical distributions. *Ann. Math. Statist.*, 19(2):279–281, 06 1948.
- [143] T. W. Anderson and D. A. Darling. Asymptotic theory of certain goodness of fit criteria based on stochastic processes. *Ann. Math. Statist.*, 23(2):193–212, 06 1952.
- [144] Prasenjit Sanyal. Limits on the Charged Higgs Parameters in the Two Higgs Doublet Model using CMS  $\sqrt{s} = 13$  TeV Results. 2019.
- [145] A. Djouadi, J. Kalinowski, and M. Spira. HDECAY: A Program for Higgs boson decays in the standard model and its supersymmetric extension. *Comput. Phys. Commun.*, 108:56–74, 1998.