



**Estudio de el Boson de Higgs y el
quark Top usando eventos ttH**
(Study of the Higgs boson and the Top
quark properties by using the ttH
production)

Trabajo de Fin de Máster
para acceder al

**MÁSTER EN
FISICA DE PARTICULAS Y DEL COSMOS**

Autor: Rufa Kunnilan Muhammed Rafeek

Director: Javier Brochero
Codirector: Alicia Calderón Tazón

29 - July - 2021

Declaration of Authorship

I, Rufa Kunnilan Muhammed Rafeek, declare that this thesis titled, "Study of the Higgs boson and the Top quark properties by using the ttH production" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:



Date: July 22, 2021

“Thanks to my solid academic training, today I can write hundreds of words on virtually any topic without possessing a shred of information, which is how I got a good job in journalism.”

Dave Barry

UNIVERSITY OF CANTABRIA

Abstract

Science Faculty
Institute of Physics of Cantabria (IFCA)

Masters in Particle Physics and Cosmos

Study of the Higgs boson and the Top quark properties by using the ttH production

by Rufa Kunnilan Muhammed Rafeek

The cross sections for the production of tt pairs, standard model Higgs boson, tt+bb, and the recently observed ttH process are measured using data collected in pp collisions at $\sqrt{s} = 13$ TeV in 2017 by the CMS experiment at the LHC. Events with one lepton (e or μ) and at least two reconstructed jets in the final state are selected. Signal regions for all the signatures are defined based on the number of jets and b jets. The measurements, performed with two different approaches, are compared with the standard model expectations and they are found to be compatible.

En el presente documento presentamos la media de la sección eficaz de la producción de tt, Higgs del modelo standard, tt+bb y el proceso recientemente observado ttH usando datos tomados de colisiones pp a $\sqrt{s} = 13$ TeV en 2017 por el detector CMS del LHC. Eventos con un único leptón (e o μ) y al menos dos jets son seleccionados. Regiones de señal para todos los procesos son definidas con base al número de jets y b jets. Las medidas, que se han hecho con dos métodos distintos, son comparadas con las predicciones del modelo standard, encontrando que son compatibles entre ellos.

Acknowledgements

The project opportunity I had with the Particle Physics Group in the Institute of Physics of Cantabria (IFCA), University of Cantabria was a great chance for learning and experimental research skill development. Therefore, I consider myself as a very lucky individual as I was provided with an opportunity to be a part of it.

Bearing in mind previous, I am using this opportunity to express my deepest gratitude and special thanks to Javier Brochero, Post-Doc Juan de la Cierva Incorporation, University of Cantabria who in spite of being extraordinarily busy with his duties, took time out to hear, guide and keep me on the correct path and allowing me to carry out my project to its completion. It is his precious guidance which was extremely valuable for my study.

I am also really grateful to Prof. Alicia Calderon Tazon, Profesor Contratado Doctor I3, University of Cantabria without whose support I would not able to bring my project to its full fledge completion. It was her timely suggestions which gave me support morally.

I would like to express my gratitude towards the Institute of Physics of Cantabria (IFCA) for supporting me financially through the scholarship provided by the Higher Council for Scientific Research (CSIC) and also for their kind cooperation and encouragement which help me in completion of this project. I also express my gratitude to my teachers of University of Cantabria for their support and timely suggestions which helped me to move forward.

My thanks and appreciation also go to my friends, colleague and people who have willingly helped me out with their abilities academically and mentally.

I perceive as this opportunity as a big milestone in my educational career. I will strive to use the gained skills and knowledge in the best possible way.

Contents

Declaration of Authorship	i
Abstract	iii
Acknowledgements	iv
1 Introduction	1
2 The LHC Accelerator and The CMS Detector	2
2.1 Large Hadron Collider (LHC)	2
2.1.1 Detectors	3
2.1.1.1 ALICE	3
2.1.1.2 LHCb	3
2.1.1.3 ATLAS	3
2.1.1.4 CMS	3
2.2 The CMS Detector	3
2.2.1 Beam pipe	4
2.2.2 Tracker System	4
2.2.2.1 Silicon Pixel Tracker	5
2.2.2.2 Silicon Strip Tracker	5
2.3 Calorimeters	5
2.3.1 ECAL	6
2.3.2 HCAL	6
2.4 Muon System	6
2.4.1 Drift Tube (DT) chambers	6
2.4.2 Cathod Strip Chambers (CSC)	7
2.4.3 Resistive Plate Chambers (RPC)	7
3 The Standard Model	8
4 Top Physics	10
4.1 Top Quark Production	11
4.1.1 Single Top Quark production	11
4.1.2 Top Anti-top pair ($t\bar{t}$) Production	11
4.2 $t\bar{t}$ decays	12
4.3 Semi Leptonic Decay of $t\bar{t}$	13
4.3.1 Main Backgrounds	13
4.4 $t\bar{t}$ +jets Categorization	15
5 Higgs Physics	17
5.1 The Higgs Mechanism	17
5.2 Production Mechanisms at the LHC	18
5.2.1 $t\bar{t}H$	19
5.3 Higgs Decays	19

5.3.1	$H \longrightarrow b\bar{b}$	20
5.3.2	$H \longrightarrow WW$	20
6	Analysis	22
6.1	Object Selection	22
6.1.1	Trigger	22
6.1.2	Identification (ID)	22
6.1.3	Isolation	22
6.1.4	Jets Selection	22
6.1.5	b-jet tagging	23
6.2	Event Selection	23
6.3	Control Plots	24
6.3.1	Regions with (at least) 4j & 6j	25
6.3.2	Signal Region with (at least) 4j & 2b (Inclusive top-like)	25
6.3.3	Signal Region with 2j & 0b ($H \rightarrow WW$ like)	26
6.3.4	Signal Region with (at least) 6j	27
6.3.4.1	At least 6j & 2b ($ttH/ttbb$ like)	27
6.3.4.2	At least 6j & 4b ($ttH/ttbb$ like- tight)	28
7	Estimation of Cross Section	30
7.1	Cut and Count Analysis	30
7.1.1	$t\bar{t}$ Cross Section ($\sigma_{t\bar{t}}$)	31
7.1.2	$H \rightarrow WW$ Cross Section ($\sigma_{H \rightarrow WW}$)	32
7.1.3	$t\bar{t} + b\bar{b}$ Cross Section (σ_{ttbb})	32
7.1.3.1	Signal Region with (at least) 6j & 2b	32
7.1.3.2	Signal Region with (at least) 6j & 4b	33
7.1.4	ttH decay with $H \rightarrow b\bar{b}$ Cross Section (σ_{ttH})	33
7.1.4.1	Signal Region with (at least) 6j & 2b	33
7.1.4.2	Signal Region with (at least) 6j & 4b	34
7.2	Fit Approach	34
7.2.1	Inclusive $t\bar{t}$ Cross Section ($\sigma_{t\bar{t}}$)	35
7.2.2	$H \rightarrow WW$ Cross Section ($\sigma_{H \rightarrow WW}$)	38
7.2.3	$ttH \rightarrow b\bar{b}$ (σ_{ttH}) & $t\bar{t} + b\bar{b}$ (σ_{ttbb}) Cross Section	39
7.2.3.1	SR with (at least) 6j & 2b	40
7.2.3.2	Signal Region with (at least) 6j & 4b	43
8	Conclusions	47
	Bibliography	49

List of Figures

2.1	A transverse slice of the CMS detector	4
2.2	The CMS Tracker	5
2.3	The transverse section of CMS Calorimeters	6
2.4	The longitudinal section of CMS Muon System	7
3.1	Standard Model	8
4.1	Feynman diagrams for single top production: (a) and (b): t-channel, (c): s-channel and (d): tW-channel	11
4.2	Leading Order Feynman diagrams for $t\bar{t}$ production: $q\bar{q}$ annihilation (top image), gg fusion in the s-channel (bottom first), gg fusion in the t-channel (bottom second), gg fusion in the u-channel (bottom third)	12
4.3	$t\bar{t}$ decay products in the semi-lepton channel	13
4.4	Single Top background	14
4.5	W+Jets background	14
4.6	Z+Jets background (one of the lepton is missed while reconstruction)	14
4.7	Associated production of vector boson with $t\bar{t}$	15
4.8	Boson-Boson background	15
4.9	QCD background	15
5.1	Higgs Potential	18
5.2	Higgs Production Mechanisms	19
5.3	Decay Modes of Higgs as a function of its mass	20
5.4	$H \rightarrow b\bar{b}$	20
5.5	$H \rightarrow WW$	21
6.1	Deep CSV	23
6.2	Jet Multiplicity	25
6.3	b-Jet Multiplicity	25
6.4	Lepton Transverse momentum (p_T), $\geq 4j+ \geq 2b$	26
6.5	η of jets in $\ell + jets, \geq 4j+ \geq 2b$	26
6.6	Lepton Transverse momentum (p_T), 2j0b	27
6.7	η of jets in $\ell + jets, 2j0b$	27
6.8	Lepton Transverse momentum (p_T), $\geq 6j+ \geq 2b$	28
6.9	η of jets in $\ell + jets, \geq 6j+ \geq 2b$	28
6.10	Lepton Transverse momentum (p_T), mu+jets	29
6.11	Lepton Transverse momentum (p_T), e+jets	29
7.1	No. of jets in $\mu + jets, \geq 4j+ \geq 2b$	35
7.2	No. of jets in $e + jets, \geq 4j+ \geq 2b$	36
7.3	Invariant mass in $\mu + jets, \geq 4j+ \geq 2b$	37
7.4	Invariant mass in $e + jets, \geq 4j+ \geq 2b$	37
7.5	Invariant Mass in $\ell + jets, 2j0b$	38
7.6	p_T of leading jet in $\ell + jets, 2j0b$	39

7.7	No. of jets in $\ell + jets, \geq 6j+ge2b$	40
7.8	No. of jets in $\ell + jets, \geq 6j+ge2b$	41
7.9	No. of jets in $\mu + jets, \geq 6j+ge2b$	42
7.10	No. of jets in $e + jets, \geq 6j+ge2b$	42
7.11	No. of jets in $\ell + jets, \geq 6j+ \geq 4b$	43
7.12	No. of jets in $\ell + jets, \geq 6j+ \geq 4b$	44
7.13	No. of jets in $\mu + jets, \geq 6j+ \geq 4b$	45
7.14	No. of jets in $e + jets, \geq 6j+ \geq 4b$	45
7.15	Invariant Mass in $\ell + jets, \geq 6j+ \geq 4b$	46

List of Tables

1.1	Results obtained for cross sections of different processes analysed using two different approaches: C & C and fit	1
4.1	W boson decay modes	12
4.2	$t\bar{t}$ decay modes	13
6.1	Signal Regions (SR) j=jets, b=b-jets	24
8.1	Results obtained for cross sections of different processes analysed using two different approaches: C & C and fit	47

Chapter 1

Introduction

The CMS experiment at the LHC has collected a large amount of data at centre of mass energy of $\sqrt{s} = 13\text{TeV}$ during the period of 2016-2018. This thesis is focused on the CMS data taken during the 2017 year that corresponds to an integrated luminosity of $\mathcal{L} = 41.5\text{pb}^{-1}$. In high energy physics, the top quark and Higgs boson are two of the most important particles due to their distinctive and characteristic properties. The top measurements are important to validate the SM predictions with a high precision and the Higgs boson (observed in CMS and ATLAS experiments in 2012) measurements give a better understanding of the particle world and also provides hints to search for new physics.

In this document, we focus in the events with top anti-top pair ($t\bar{t}$), Higgs boson (H) and the associated $t\bar{t}$ production with the Higgs boson ($t\bar{t}H$) in the semi-leptonic decay mode ($\ell + jets$). We present the results of the cross sections of the $t\bar{t}$ pair production ($\sigma_{t\bar{t}}$), Higgs boson decaying to two W bosons ($\sigma_{H \rightarrow WW}$), the production of $t\bar{t}$ associated with b-jets ($\sigma_{t\bar{t}bb}$) and with a Higgs boson decaying to two b-jets ($\sigma_{t\bar{t}H}$).

The cross sections are measured using two different approaches, a cut and count (C & C) analysis and a template fit. Signal regions are defined based on the final states of each process and are chosen such that background contribution is the minimum. The knowledge acquired in the analysis of the $t\bar{t}$ and the $H \rightarrow WW$ processes are combined to prepare and study the $t\bar{t}H$ process. The results obtained are compatible with the SM predictions and are summarised in the table below. In this analysis, the systematic uncertainties are not calculated due to the time constraints.

TABLE 1.1: Results obtained for cross sections of different processes analysed using two different approaches: C & C and fit

Processes	Theoretical [pb]	Observed [pb]	
		C & C	Fit
$t\bar{t}$	831.76	857	855
$H \rightarrow WW$	32.4	~ 0	85
$t\bar{t}bb$	4	6.5	3.7
$t\bar{t}H(H \rightarrow b\bar{b})$	0.29	0.9	1.3

Chapter 2

The LHC Accelerator and The CMS Detector

2.1 Large Hadron Collider (LHC)

The LHC [1] is the world's largest particle accelerator at the European Organization for Nuclear Research (CERN), located near Geneva on the Franco-Swiss border. It consists of a 27 km ring of magnets made up of superconductors with an accelerator complex to boost the energy of the particles as they pass through it and has a depth of 175 metres.

The LHC is a proton-proton and a heavy ion collider. Primarily, it collides beams of protons, but it also can collide beams of heavy ions: proton-lead and lead-lead collisions which are run for a period of one month in a year. It consists of four crossing points, around which there are eight detectors designed for different physics researches. They are,

- ATLAS: A Toroidal LHC Apparatus
- CMS: Compact Muon Solenoid
- LHCb: LHC beauty experiment
- ALICE: A Large Ion Collider Experiment
- TOTEM: TOTAl Elastic and diffractive cross section Measurement [2]
- LHCf: LHC forward [3]
- MoEDAL: Monopole and Exotics Detector At the LHC [4]
- FASER: ForwArd Search ExpeRiment [5]

The four crossing points around the ring correspond to the positions of the four main particle detectors: ATLAS, CMS, ALICE and LHCb.

The LHC is designed to accelerate the proton beams to a centre of mass energy (\sqrt{s}) equal to 14 TeV at a maximum of $10^{34} \text{cm}^{-2}\text{s}^{-1}$ luminosity. This desired energy is reached using the accelerator complex at the CERN. The protons, obtained from extracting orbiting electrons from Hydrogen atoms, are injected and passed through each of the accelerators. The protons in the Linear Accelerator (LINAC 2) reach an energy of 0.12 GeV, from which are injected to the Proton Synchrotron (PS) Booster to reach an energy of 1.4 GeV. Then, the beam is passed to the PS, gets accelerated to 26 GeV and to the Super Proton Synchrotron (SPS) to reach an energy of 450 GeV.

Finally, they are passed to the LHC to reach the desired energy of 14 TeV. There are different kinds of superconducting (SC) magnets and cavities set along the circumference of the LHC ring to obtain this energy value which accelerate protons to the velocity of light and bend trajectories.

2.1.1 Detectors

2.1.1.1 ALICE

The ALICE detector [6] is dedicated to heavy-ion collisions, to study the physics of strong interactions occurring at highest energy densities. The LHC is dedicated to heavy ion collisions, that is, the lead-lead collisions during a part of each year so that the laboratory conditions said to exist just after the Big bang can be recreated to study the phase of matter called the quark gluon plasma and its properties. The ALICE uses a set of 18 detectors for hadron identification which gives the information about the different properties of particles such as electrical charge, mass and the velocity.

2.1.1.2 LHCb

The LHCb experiment [7] is specialized to study heavy flavour physics, primarily to measure the CP violation parameters in the decays of beauty (or b) hadrons. The LHCb detector uses a series of subdetectors with one behind the other, covering only a small angle in one direction unlike with an enclosed detector as the general purpose detectors. It is a 5600 tonne detector made up of a one-arm forward spectrometer and planar detectors.

2.1.1.3 ATLAS

The ATLAS detector [8] is a general purpose detector with a length of 44m and a diameter of 25m making it the largest detector at the LHC. The strong magnetic field is created by a large toroidal magnet and a small solenoidal magnet. It searches for a wide range of physics, from solving the quest in search of the SM Higgs to particles beyond the SM. It has the same scientific goals as the CMS but it is different on the basis of the design of the magnet system and technical solutions.

2.1.1.4 CMS

The CMS detector is also a general purpose detector as ATLAS. In CMS, the momentum resolution of the tracking detectors is much better than ATLAS but the energy resolution of the hadronic calorimeters in the ATLAS is better. More information about the detector is given in the following section 2.2.

2.2 The CMS Detector

The CMS [9] is the general purpose detector with one of the largest international collaboration in scientific history consisting of over 5000 scientists from 200 institutions in 50 countries as of 2019. It is small but compact, weighing about 14,000 tonnes with a solenoid inside its huge superconducting magnet and hence the name. It is 21.5 metres long and has a diameter of about 15 metre. The coil of superconducting cable generates a very strong magnetic field of 4 tesla. The experiment has a wide range of goals including studying the properties of the SM and the Higgs boson, exploring

new physics at the TeV scale, searching for beyond SM physics or extra dimensions and so on.

The CMS experiment has a special interest in studying the Higgs physics and searching for new particles. To achieve this, the detector has a muon system, an electromagnetic calorimeter, a hadronic calorimeter and a tracking system. Thus, it consists of subdetectors which are built to measure the energy and momentum of muons, photons, electrons and hadrons over a large range of energy. The silicon-based tracker is the innermost layer of the detector surrounded by a scintillating crystal electromagnetic calorimeter. The electromagnetic calorimeter is surrounded itself by the hadronic calorimeter. There are the muon detectors located outside the magnet. The interaction of each particle in these components of the detector is in a different way based on the fundamental properties of the particles. The Figure 2.1 depicts a slice of the CMS detector with the behaviour of the particles as they transverse through a section of it.

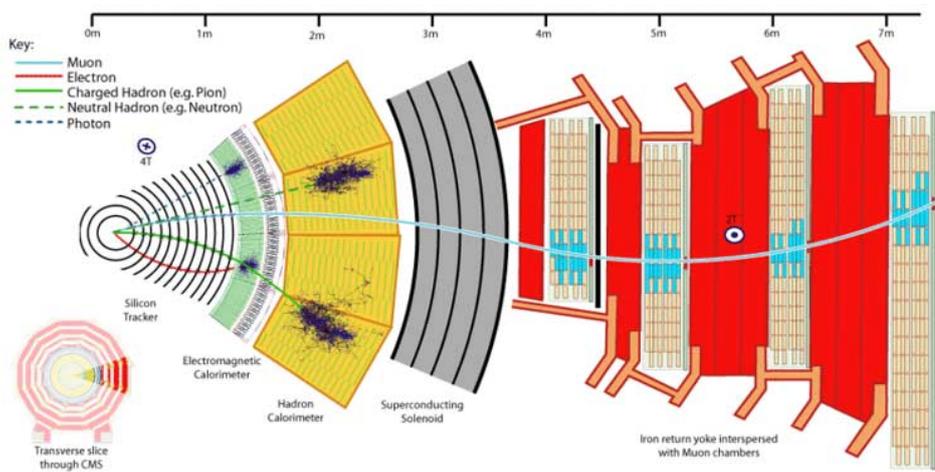


FIGURE 2.1: A transverse slice of the CMS detector

2.2.1 Beam pipe

The beam pipe [10] is a 39 metres long component of the CMS detector in which the beams of the LHC collide at the Interaction Point (IP) or the centre point where the proton-proton collisions occur between the counter rotating beams. Surrounding the beam pipe are the several different layers from which the detector is built.

The beams are grouped into bunches of approximately 10^{11} protons. There is one bunch crossing in every 25 ns. When the bunches cross at full design luminosity, there will be a production of an average of 20 proton-proton collisions.

2.2.2 Tracker System

The CMS tracker [11] is the innermost layer of the detector which records and reconstructs the vertices and the paths of the charged particles by taking the position measurements at different key points. The tracker is made entirely of silicon by two types of silicon layers: the silicon pixels and the silicon microstrip detectors. The pixel layers occupy the core of the detector receiving the high volume of particles. The electric signals are produced by the pixels and the microstrips as the particles travel through them and are amplified and detected.

2.2.2.1 Silicon Pixel Tracker

The pixel detector consists of 65 million pixels which allows to reach a very high precision in reconstructing the tracks of high momentum muons, electrons, hadrons and particles of even very short lifetime. Each layer of the detector is a silicon sensor of dimensions, $100\mu\text{m} \times 150\mu\text{m}$ which creates electron-hole pairs as a particle travel through it and collects as a small electric signal.

2.2.2.2 Silicon Strip Tracker

The silicon strip detector consists of 10 layers outside the pixel detector with four different subsystems:

- Tracker Inner Barrel (TIB)
- Tracker Inner Disks (TID)
- Tracker Outer Barrel (TOB)
- Tracker EndCaps (TEC)

There are four TIB layers assembled in shells in the two TIDs which consist of three small discs each and the TOB consists of six concentric layers with two end caps (TEC+ and TEC-) at the end of the tracker. This makes a minimum of 12 measurement points in each charged track over a large range of pseudorapidity (η). The working of the strip tracker is nearly the same as the pixel tracker.

The Figure 2.2 depicts the tracker with its different subsystems.

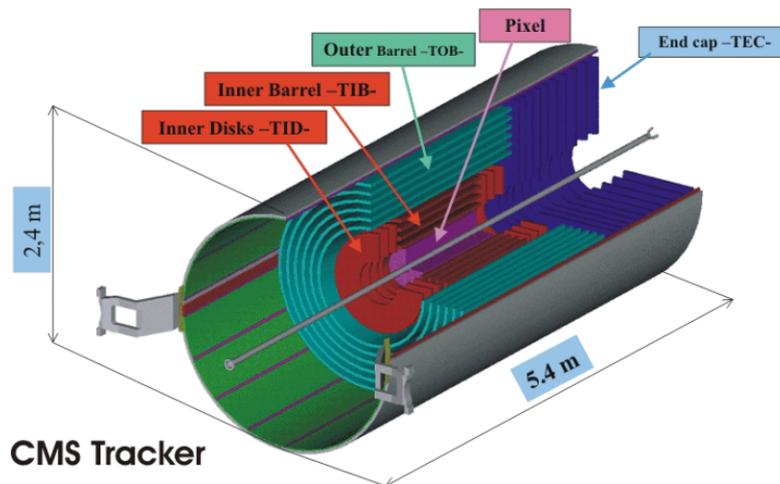


FIGURE 2.2: The CMS Tracker

2.3 Calorimeters

The measurement of energies of different particles emerged as a result of the interactions at the collision point in the CMS detector is made with two types of calorimeters: the inner layer is the Electromagnetic Calorimeter (ECAL) and the outer layer is the Hadronic Calorimeter (HCAL).

2.3.1 ECAL

ECAL [12] is designed to stop the electrons and photons and to measure their energies precisely. It is made of Lead Tungstate ($PbWO_4$) crystals which are highly transparent to the photons and electrons and thus produces light proportional to the energy of the particle as they pass through it. It is constituted of the Electromagnetic Barrel (EB) enclosed by two Electromagnetic Endcaps (EE) and the Electromagnetic preShower detector (ES) covering an area of pseudorapidity, $|\eta| < 3$.

2.3.2 HCAL

HCAL [13] is designed to measure the energies of hadrons (particles made of quarks and gluons) and their decay products. It also indirectly measures the presence of particles which do not interact in the detector such as the neutrinos. It is made of layers of copper and fluorescent scintillators which produces a light pulse as the particles pass through it. It is constituted of the Hadronic Barrel (HB), the Hadronic Endcaps (HE), the outer barrel - Hadronic Outer (HO) and the Hadronic Forward (HF) sections. The HF is located outside the muon system in order to achieve the complete hermeticity.

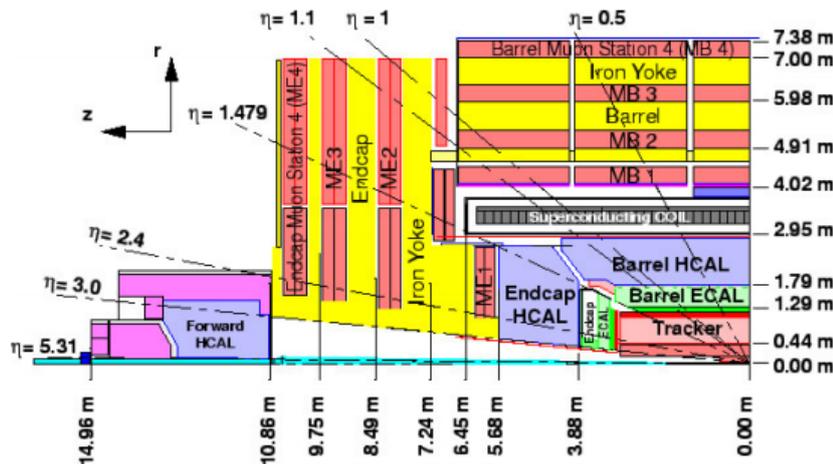


FIGURE 2.3: The transverse section of CMS Calorimeters

2.4 Muon System

Identifying, selecting and reconstructing the muons at high luminosity were important in the CMS as they are the particles mainly involved in the physics processes that the LHC is looking for, such as, Electroweak, top, Higgs and also in search for new physics processes beyond the SM. As the muons can penetrate materials such as iron over several meters, they pass through the calorimeters without getting detected as the neutrinos. Thus, the muon tracking system with four muon stations is located outside the magnetic coil to the very edge of the detector. It consists of three types of sub detectors: Drift tube (DT) chambers, Cathod Strip Chambers (CSCs) and Resistive Plate Chambers (RPCs). The specifications in design of the muon detector and more information is given in Ref. [14].

2.4.1 Drift Tube (DT) chambers

The purpose of the DT system is to measure the positions of the muons in the barrel part of the detector. There are four layers of concentric cylinders surrounding the

beam line where the DTs are installed in and on the wheels of the iron yoke with muon stations on each of the layer. Each DT chamber has a size of $2m \times 2.5m$ comprising of 12 aluminium layers which are arranged in three groups of four layers. The outer layer has 70 DTs where the inner three layers have 60 each, consisting of DT cells with 80% Ar and 20% CO_2 gas mixture.

2.4.2 Cathod Strip Chambers (CSC)

The CSCs are installed in the end caps to detect the muons in the end cap region where there is uneven magnetic field and high rates of particles. The chambers consist of anode wires crossed with copper cathode strips and are filled with six gas layers consisting of 20% CF_4 , 30% Ar and 50% CO_2 . When a muon passes through the layers, it causes gas ionization and as a result, avalanche of electrons are produced which in turn produces the electrical signal. The signal is fast and thus suitable for triggering purposes.

2.4.3 Resistive Plate Chambers (RPC)

The RPCs are gaseous parallel-plate detectors with fast response and excellent time resolution of the order of 25ns, thus used mainly for triggering purposes. It consists of two parallel plates built using electrodes of high volume resistivity and are separated by a gap of gas mixture.

The Figure 2.4 depicts a quarter view of the CMS muon system longitudinally. ME1, ME2, ME3 and ME4 denote the four disks of the CSCs. MB1, MB2, MB3 and MB4 denote the four muon stations on each wheel in the each sector of the iron yoke of the magnet.

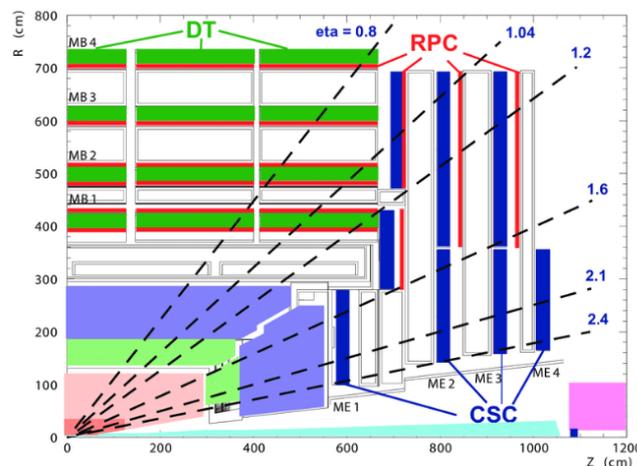


FIGURE 2.4: The longitudinal section of CMS Muon System

Chapter 3

The Standard Model

The Standard Model of Particle Physics [15] describes the construction of the universe by uniting the fundamental building blocks called matter particles, which are the fermions (spin 1/2) and the interactions among them through the three of the four fundamental forces known mediated by carrier particles called gauge bosons (spin 1) as given in Figure 3.1.

The fermions constitute quarks (q) and leptons (l). Quarks occur in six different types based on their flavours: up (u), down (d), charm (c), strange (s), bottom (b) and top (t). The up-type quarks: up (u), charm (c) and top (t) carry a charge of $\frac{2}{3}e$ and the down-type quarks: down (d), strange (s) and bottom (b) carry a charge of $-\frac{1}{3}e$ where e is the electronic charge. Leptons also occur in six types of flavours: electron (e), muon (μ), tau (τ) and their associated neutrinos: electron neutrino (ν_e), muon neutrino (ν_μ), tau neutrino (ν_τ). The charged leptons carry a charge of $-1e$ where the neutrinos are chargeless particles.

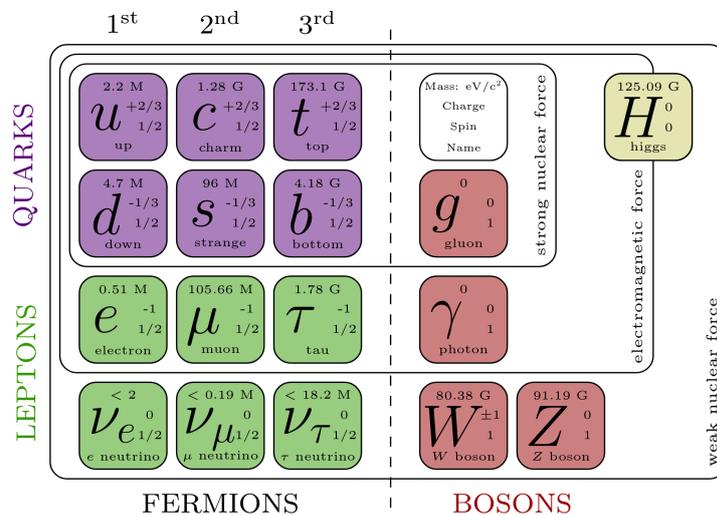


FIGURE 3.1: Standard Model

Quarks and Leptons are grouped into three generations. The second and third generation particles are known as the heavier cousins of the first generation particles. An illustration of this characterisation is given in the Figure 3.1 with masses, spin and charges of the particles. All the fermions described have their own anti particles with same mass but opposite quantum numbers.

The gauge bosons mediate the interaction between the particles through the three

fundamental forces: electromagnetic (EM), weak nuclear, strong nuclear. The fourth fundamental force, the gravitational force, is not described in the standard model as it is negligible for low mass particles as in the Standard Model.

- The electromagnetic force acts between the charged particles through the absorption or emission of virtual photons, described by the gauge theory, Quantum Electrodynamics(QED). The range of this force is infinite as the photons are massless.
- The weak nuclear force and its interactions are responsible for changing the flavour of quarks and it acts between all the fermions through the exchange of charged W^\pm bosons and the neutral Z^0 boson and the corresponding interactions are known as Charged Current(CC) and Neutral Current(NC) interactions respectively. These bosons are massive with Z boson being heavier than the W bosons described by the gauge theory, Quantum Flavourodynamics(QFD).
- Strong nuclear interactions acts between the color charged particles called quarks which are mediated by eight gluons. The gluons are massless and has an effective color charge, explained by the theory of Quantum Chromodynamics (QCD).

The Standard Model explains the theory of electroweak interactions and the quantum chromodynamics. The quantum electrodynamics with the quantum flavourodynamics together explains the united electroweak interactions and the quantum chromodynamics explains the strong interactions of the particles. So the three forces-electromagnetic, weak nuclear and strong nuclear forces are incorporated into a single model.

The whole model is formulated by assigning a set of symmetries of the system which governs the dynamics and kinematics of this theory of interactions between the particles. It follows some general principles, mainly, quantum mechanics and special theory of relativity to explain this theory of interactions which are portrayed beautifully in context of Quantum Field Theory (QFT) where the Lagrangian controls the dynamics and kinematics of the theory [16]. According to the Quantum field theory, everything that we perceive is a collection of vibrations in quanta fields due to the energetic excitations of the fields. These excitations or vibrations of the fields are carried as quanta and interpreted as particles. For more information, refer the Ref. [17].

Chapter 4

Top Physics

The top quark, the heaviest fundamental particle known in the SM and is considered as the isospin partner of the bottom quark. It was predicted by Makoto Kobayashi and Toshihide Maskawa (1973) to explain CP violation by weak interaction and was discovered in 1995 by CDF [18] and DØ [19] collaborations at Tevatron in Fermilab. It belongs to the third generation of quarks with a charge of $\frac{2}{3}e$ and takes part in all the four fundamental interactions known like the other quarks. The anti-particle of top quark is the top antiquark or antitop, denoted as \bar{t} .

Due to the huge mass of the top quark, its properties are different from those of the other quarks in the SM. The mass of the top quark has been estimated by the experiments like CMS and ATLAS and the average of the value is 173 GeV [20]. The top (or the antitop) decays only by the first order weak interaction and gets decayed before hadronisation. It decays into a W boson and a quark.

$$t \rightarrow q + W^+; \quad q = d(0.007\%), s(0.17\%), b(99.8\%) \quad (4.1)$$

where the most frequent decay is to a b-quark as the probability of decay to the other down type quarks are very small, given by $|V_{td}|^2$ and $|V_{ts}|^2$ of the Cabibbo-Kobayashi-Maskawa (CKM) matrix (see Equation 4.3) [21]. The CKM matrix provides the transition probability of a quark q to another q' proportional to $|V_{qq'}|^2$.

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \quad (4.2)$$

The magnitude of CKM elements,

$$|V_{CKM}| = \begin{pmatrix} 0.097446 \pm 0.00010 & 0.22452 \pm 0.00044 & 0.00365 \pm 0.00012 \\ 0.22438 \pm 0.00044 & 0.97359 \pm 0.00011 & 0.04214 \pm 0.00076 \\ 0.00896 \pm 0.00023 & 0.04133 \pm 0.00074 & 0.999105 \pm 0.000032 \end{pmatrix} \quad (4.3)$$

The top-Yukawa coupling [22] dominates over the gauge couplings as the strength of the coupling is proportional to the mass of the top quark. Since it is the heaviest fundamental particle known in the SM, the coupling of the top to the Higgs field is the strongest, ~ 1 . Thus, its properties are extensively studied as a means for the indirect determination of the Higgs boson's mass and other theories of physics beyond the SM.

4.1 Top Quark Production

There are two main SM processes by which the top quarks are produced at the hadron colliders. The most dominant production mechanism is the top anti-top pair production ($t\bar{t}$) and the other mechanism is the single top quark production.

4.1.1 Single Top Quark production

Single top quarks are produced via electroweak processes [23]. On the virtuality of the W boson exchange, there are mainly three production mechanisms at LHC. The most abundant production is through t-channel with a 70% of the total cross section, followed by the associated production with a real W boson (tW) with a 25% and then by the s-channel production with a 5%.

The t-channel production mode of the single top is mediated by a virtual space-like W boson. An initial state gluon, arising from the sea quark inside the proton, gets converted into a $b\bar{b}$ pair and one of the b-quarks interacts with the W boson to produce the top quark. Thus, the process is also referred as W gluon fusion, represented in the Figure 4.1 (a) and (b).

The s-channel production mode of the single top is mediated by a virtual time-like W boson. This W boson produced via the fusion of two quarks ($q\bar{q}$) produce a single top quark along with a \bar{b} quark. Eventhough this is the most challenging channel due to its low statistics, it is sensitive to new physics such as searches for charged Higgs bosons and W' bosons. It is represented in the Figure 4.1 (c).

The tW-channel produces the single top in association with a real W boson. A b-quark arising from the sea quark inside the proton fuses with a gluon to produce this pair. This process is represented in the Figure 4.1 (d).

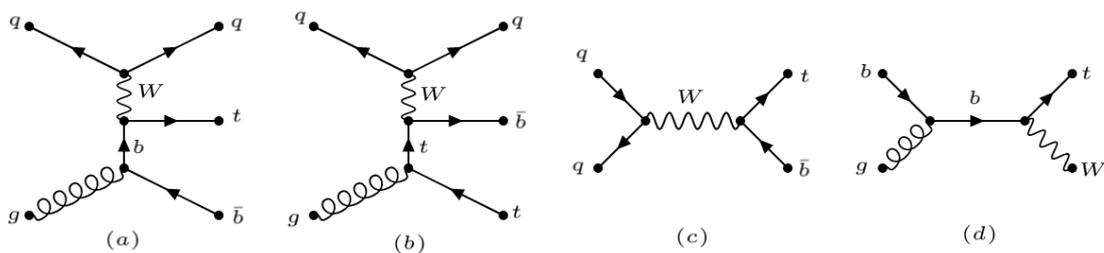


FIGURE 4.1: Feynman diagrams for single top production: (a) and (b): t-channel, (c): s-channel and (d): tW-channel

4.1.2 Top Anti-top pair ($t\bar{t}$) Production

$t\bar{t}$ are produced through strong interaction, either via the gluon-gluon (gg) fusion or the quark anti-quark ($q\bar{q}$) annihilation. At LHC ($p\bar{p}$ collisions), the dominant process is the gluon fusion occurring 83% followed by $q\bar{q}$ annihilation with a 17%. The $q\bar{q}$ annihilation or the QCD pair production is the dominant mechanism at the hadron colliders with $p\bar{p}$ collisions as the valence quarks in protons can annihilate with the

valence anti-quarks in anti-protons.

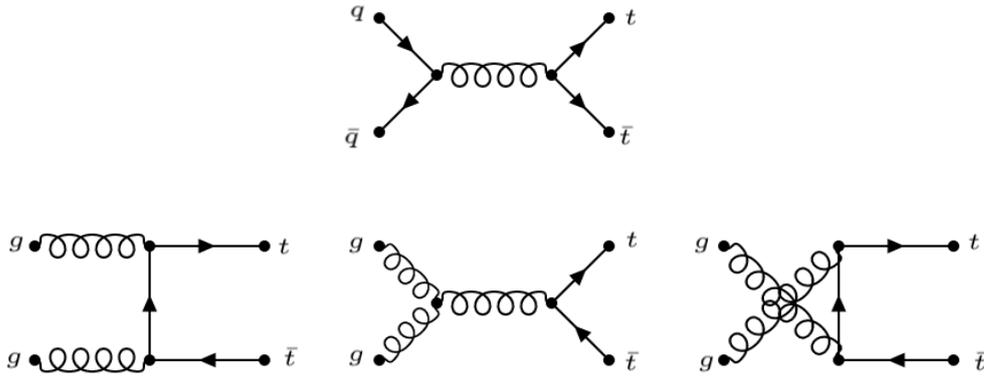


FIGURE 4.2: Leading Order Feynman diagrams for $t\bar{t}$ production: $q\bar{q}$ annihilation (top image), gg fusion in the s-channel (bottom first), gg fusion in the t-channel (bottom second), gg fusion in the u-channel (bottom third)

4.2 $t\bar{t}$ decays

As a top quark decays 99.8% of its times to a W boson and a b-quark, the decay products of a $t\bar{t}$ event are two W bosons and two b-quarks. The b-quarks are seen as jets as the quarks are so energetic and prefer to produce jets whereas the W bosons decay into two leptons or leptons and jets or light flavour quark jets. The table 4.2 gives the information about the decay modes of W boson and their corresponding branching ratios (BR) [24].

TABLE 4.1: W boson decay modes

Modes	BR (%)
$\ell + \nu$	10.8 ± 0.09
$e + \nu$	10.75 ± 0.13
$\mu + \nu$	10.57 ± 0.15
$\tau + \nu$	11.25 ± 0.20
hadrons	67.60 ± 0.27

In case of W boson decaying to a τ lepton and its neutrino, it further decays leptonically to an electron or a muon with their neutrinos or hadronically to produce jets [25]. Thus, there are mainly three modes of decay for $t\bar{t}$ events: dileptonic, semi-leptonic and hadronic modes.

The $t\bar{t}$ signature studied in this project is the semi-leptonic mode, which is described in the section 4.3.

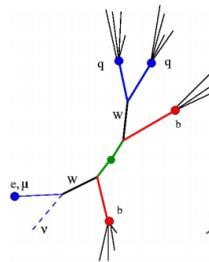
TABLE 4.2: $t\bar{t}$ decay modes

Modes	Products	BR (%)
Dileptonic	$\ell\ell + \nu\nu$	10.5 ± 0.12
	$ee + \nu_e\nu_e$	1.16 ± 0.02
	$\mu\mu + \nu_\mu\nu_\mu$	1.12 ± 0.02
	$\tau\tau + \nu_\tau\nu_\tau$	1.27 ± 0.03
	$e\mu + \nu_e\nu_\mu$	2.27 ± 0.04
	$e\tau + \nu_e\nu_\tau$	2.42 ± 0.05
	$\mu\tau + \nu_\mu\nu_\tau$	2.38 ± 0.05
Semi-leptonic	$\ell\nu + qq'$	43.80 ± 0.40
	$eqq' + \nu_e$	14.53 ± 0.19
	$\mu qq' + \nu_\mu$	14.29 ± 0.21
	$\tau qq' + \nu_\tau$	15.21 ± 0.28
Hadronic	$qq' + q''q'''$	45.70 ± 0.26

4.3 Semi Leptonic Decay of $t\bar{t}$

The semi-lepton state of decay of $t\bar{t}$ includes a single lepton, which can be an electron or a muon, the associated neutrino and four jets [26]. The two jets from the W boson are mainly light flavoured (LF) and the two b-jets directly from the top quarks are heavy flavoured (HF). The lepton has a high p_T and the neutrino and two b-jets are accounted by the missing transverse energy (\cancel{E}_T).

The semi-leptonic channel has a balance in terms of branching ratio and the background contribution when compared to the fully leptonic and hadronic channels. It has more statistics with a branching ratio of about 44% than the dilepton channel and not much contamination from background as in the fully hadronic channel. The Figure 4.3 shows the topology of $t\bar{t}$ signature in this channel.

FIGURE 4.3: $t\bar{t}$ decay products in the semi-lepton channel

4.3.1 Main Backgrounds

The background processes to the semi-leptonic channel of $t\bar{t}$ are the processes which leave a similar signature, with a single lepton and jets, as the final state products. The leptons in the final state are called as *prompt* leptons, if the lepton is coming from the primary interaction vertex from interesting physics (through electroweak) or otherwise, *fake* leptons, coming from meson decays in jets, cosmic rays, unidentified leptons, jets that go through muon chambers etc..

The main SM background processes are as follows:

1. **Single Top** The production of a single top quark with a W boson can decay and produce a single lepton and jets in the final state. It will have missing transverse energy and b-quark along with jets. This is observed when the W boson produced from the top decay and the W boson from the tW vertex decay semi-leptonically. It is represented in the figure 4.4.

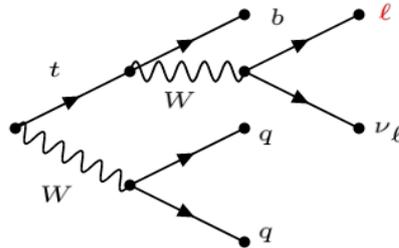


FIGURE 4.4: Single Top background

2. **W+Jets** The W boson produced in association with quarks or gluons can decay and produce signature equivalent to the $t\bar{t}$ semi-leptonic channel. The gluons can produce heavy flavour quarks, thus producing a $Wb\bar{b}$ combination which in turn can produce a lepton, two b-jets and \cancel{E} . It is represented in the figure 4.5.

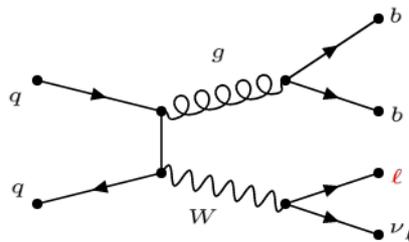
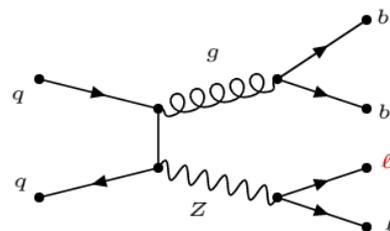


FIGURE 4.5: W+Jets background

3. **Z+Jets** The Z boson produced in association with quarks or gluons can decay and produce a $Zb\bar{b}$ combination which leave a semi-leptonic $t\bar{t}$ signature due to reconstruction mistakes (a lepton is missed while reconstructing the event). Or the Z boson decays into quarks and one of the jet is misidentified as a lepton. This is represented in the figure 4.6.

FIGURE 4.6: Z+Jets background
(one of the lepton is missed while reconstruction)

4. **$t\bar{t}$ dileptonic and hadronic** It is possible to miss one of the two leptons in the fully leptonic channel and misidentify the event as semi-leptonic. Also, it is possible to have a fake lepton in the hadronic channel.
5. **Vector bosons ($t\bar{t}V$)** A W/Z boson produced with $t\bar{t}$ pair can fake a semi-leptonic decay signature of $t\bar{t}$ when the W/Z boson decays hadronically to

produce light flavour jets. It is represented in the figure 4.7.

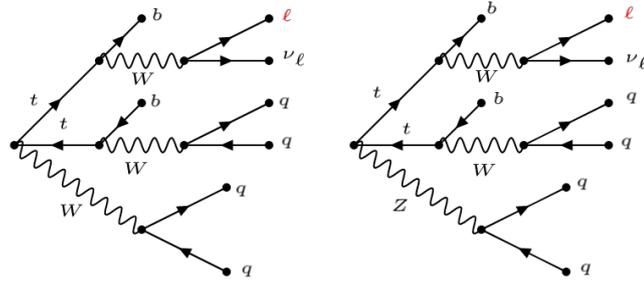


FIGURE 4.7: Associated production of vector boson with $t\bar{t}$

6. **Dibosons (VV)** The vector bosons produced in pairs such as WW , WZ and ZZ can serve as background processes for the semi-leptonic $t\bar{t}$ event. In WW process, one of the W boson can decay semi-leptonically and the other via hadrons. In WZ process, the W boson can decay semi-leptonically with the Z boson decaying into heavy flavour quark jets. And in the ZZ process, one Z boson decay leptonically and the other Z boson into heavy flavour jets with one lepton being missed in reconstruction giving semi-leptonic final state. The processes are represented in the figure 4.8.

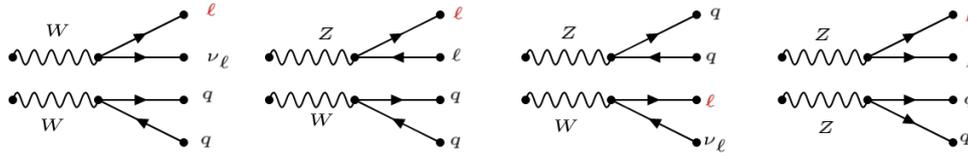


FIGURE 4.8: Boson-Boson background

7. **QCD** The quarks or gluons in the detector background can produce $b\bar{b}$ pair which can further produce a pair of W bosons. one of the W bosons can decay semi-leptonically and the other can decay hadronically giving semi-leptonic final state. One of the possible Feynman diagram is represented in the figure 4.9.

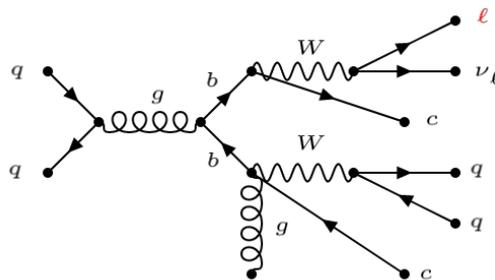


FIGURE 4.9: QCD background

4.4 $t\bar{t}$ +jets Categorization

The top quark pairs can be associated with jets, so that, they cannot be analysed together. The $t\bar{t}$ pairs are thus analysed with a pair of jets, $t\bar{t}jj$, where j is jets produced from quarks (u, d, c, s, b) or gluons. The $t\bar{t} + b\bar{b}$ process is interesting in both theoretical and experimental aspects. It is one of the dominant backgrounds in the study of the $t\bar{t}H$ process.

The flavour of the additional jet is identified using an algorithm known as 'ghost-matching technique' [27]. A b-jet is a particle-level jet which contains bottom hadrons and a c-jet is the one with charm hadrons. All the other jets are considered as light flavour jets, that is, the jets from a up, down, strange or gluon. Particle-level jets accompanying the $t\bar{t}$ decay are classified mainly into four categories:

1. $t\bar{t}b\bar{b}$ This consists of events with at least two b-jets, independent of the number of the bottom hadrons in each of them.
2. $t\bar{t}bj$ This consists of one additional b-jet. It occurs when one of the b-jets is missed due to reconstruction efficiencies or since it was out of the acceptancy of the detector. This classification also consists of jets with one additional particle-level jet containing at least two bottom hadrons ($t\bar{t} + 2b$).
3. $t\bar{t}c\bar{c}$ This consists of jets with at least two c-jets, independent of the number of the charm hadrons in each of them.
4. $t\bar{t}LF$ This $t\bar{t}$ with light flavour jets consists of at least two additional light flavour jets or a light flavour jet and a c-jet. In our analysis, this category also includes the $t\bar{t}$ events without any associated jets.

Chapter 5

Higgs Physics

The Higgs Boson is a chargeless and spinless fundamental particle produced as a result of the excitation of the Higgs field. It is responsible for the generation of mass for the bosons and fermions (except neutrinos). The mass of the SM-like Higgs boson is experimentally found to be 125 GeV. [28].

5.1 The Higgs Mechanism

The SM unites the electromagnetic and weak interactions at higher energy values, above 100 GeV, under a unified single gauge group called $SU_L(2) \times U_Y(1)$, which separate only after the electroweak symmetry breaking [29]. It is experimentally observed that only left-handed particles take part in weak interactions, explained under the gauge group, $SU_L(2)$ where the members of the doublet are distinguished on the basis of a quantity called weak isospin or particularly, the third component of isospin (T_3). This makes the symmetry incomplete and is saved by including the electromagnetic gauge group, $U_1(Y)$ based on the quantity, hypercharge (Y). The hypercharge (Y) and the charge (Q) are related as,

$$Q = T_3 + \frac{Y}{2} \quad (5.1)$$

Both the fermions and bosons acquire their mass through the Higgs mechanism, without which the gauge invariance inhibits the SM particles to have mass [30].

The Higgs field is a complex doublet scalar field with four components, two charged and two neutral components with a Mexican hat-like potential as shown in the figure 5.1.

$$\Phi = \begin{bmatrix} \phi^+ \\ \phi^0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{bmatrix} \quad (5.2)$$

where ϕ^+ denotes the charged field with ϕ_1 and ϕ_2 as its components and ϕ^0 denotes the neutral field with ϕ_3 and ϕ_4 as its components and it turns out that both of the charged components along with one of the neutral component ϕ_4 give rise to massless Goldstone bosons whereas the remaining neutral component due to the vibrations of the side of the valley gives rise to the massive Higgs boson. The goldstone bosons are coupled via a weak hypercharge to the gauge fields of Z and W bosons which provide them an extra degree of freedom to accommodate mass and there are no goldstone bosons to provide mass to the photons.

The mass generation of fermions is explained by the Yukawa coupling term in the Lagrangian, which is the interaction between the Higgs field and the Dirac fields.

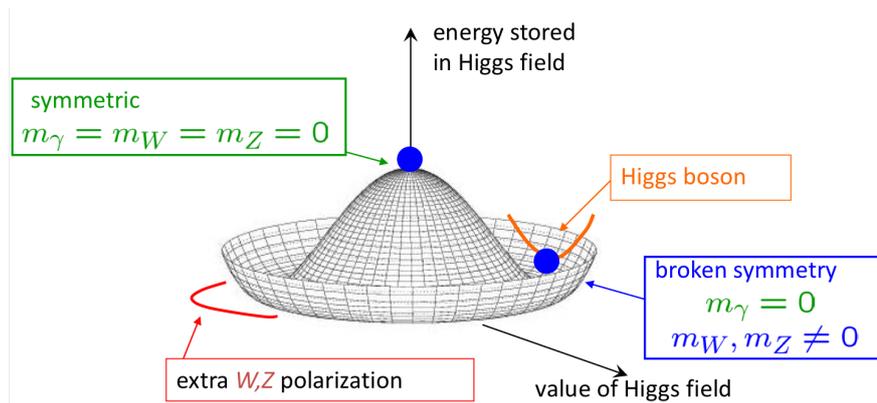


FIGURE 5.1: Higgs Potential

5.2 Production Mechanisms at the LHC

There are different mechanisms by which the Higgs boson is produced at the collider [31]. Each of the mechanism makes use of the preference according the SM Higgs couplings where the coupling of Higgs to other particles is proportional to the mass of the particle. Thus, the coupling to the top quarks and the vector bosons dominate over the couplings to other quarks or bosons.

The four main production mechanisms of the Higgs boson at the LHC [32] are:

1. **gluon-gluon fusion (ggH)** In the proton-proton collisions, the gluons inside the protons collide, combine and produce a Higgs through a heavy quark loop. It is the most dominant mechanism of Higgs production [33]. The Feynman diagram for the process is shown in the figure 5.2 (a).
2. **Vector Boson Fusion (VBF)** Two fermions or anti-fermions collide and produce a Higgs through the exchange of a virtual vector boson, that is, W or Z boson and their fusion. It is the second most dominant mechanism of production after the ggH process [33]. The Feynman diagram for the process is shown in the figure 5.2 (b).
3. **Associated production with vector bosons** Vector bosons such as W and Z bosons are produced in the collision of a fermion with an anti-fermion ($q\bar{q}$). If these bosons carry enough energy to produce the Higgs boson, they emit a Higgs boson. It is the third largest mechanism since the LHC has proton-proton collisions and not proton-anti proton [33]. The Feynman diagram for the process is shown in the figure 5.2 (c).
4. **Associated production with top quarks (ttH)** Higgs boson can be produced in the fusion of a top anti-top or from the radiation of a single top quark. Two gluons collide and produce a $q\bar{q}$ pair which can further combine to form a Higgs boson. It is a very rare process and accounts for only 1% of the total Higgs production [33]. The Feynman diagram for the process is shown in the figure 5.2 (d).

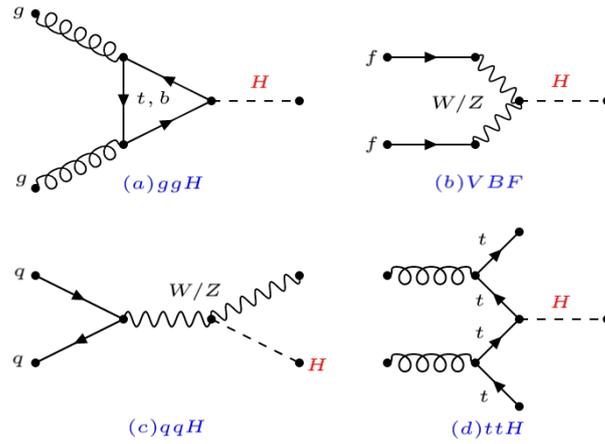


FIGURE 5.2: Higgs Production Mechanisms

5.2.1 ttH

The coupling of Higgs to the top quark is the strongest, ~ 1 and thus, measuring it gives a very important understanding of the particle world. The Higgs boson can decay into a pair of b-quarks but not to a pair of top quarks due to the mass of the top quark being larger than that of the Higgs. The production probability of the Higgs depend on the coupling and thus it is important to measure the coupling of Higgs to the top quark. It is also a way for the measurement of the top-Yukawa coupling.

Higgs boson is produced either in the fusion of a $t\bar{t}$ pair or radiated from a single top quark [34]. Thus, looking for collisions with two top quarks and Higgs through their final decay products, this coupling can be directly probed [35]. But it is very difficult to probe this, since the Higgs decay quickly in different ways as explained in the section 5.3. It is an open window to new physics as the coupling is not only sensitive to the properties of the Higgs boson but also to the properties of the particles taking part in the interaction.

5.3 Higgs Decays

The Higgs boson can decay leptonically to a pair of τ leptons or into a quark anti-quark pair: $b\bar{b}$ and $c\bar{c}$ or into a pair of gluons or into a pair of W bosons etc. Higgs decaying into a pair of b-quarks is the most dominant one followed by τ , then charm (c) and gluons.

The figure 5.3 shows the decay branching ratios and the total width and it can be seen that the decay modes change in prominence on the basis of the value of the mass of the Higgs boson (M_H). The experimentally observed mass is 125GeV.

Also, we can see that at $M_H = 125\text{GeV}$, the prominent decays are $H \rightarrow b\bar{b}$ and $H \rightarrow WW$.

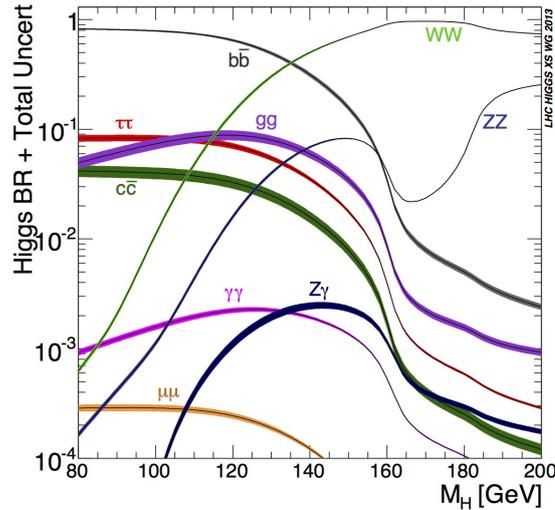
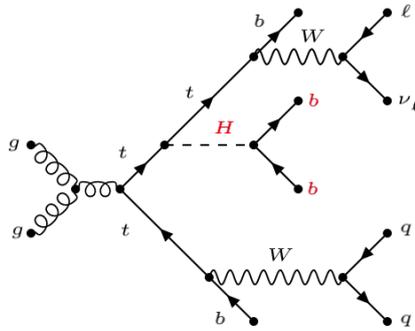


FIGURE 5.3: Decay Modes of Higgs as a function of its mass

5.3.1 $H \longrightarrow b\bar{b}$

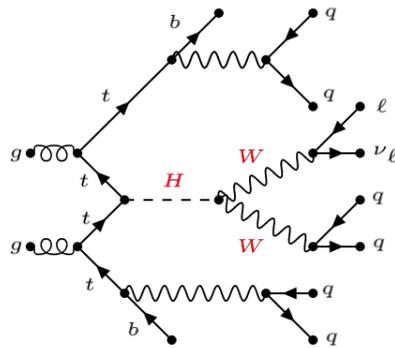
As mentioned above, it is the leading decay mode, accounting for 60% of the decay of the Higgs boson [36]. It is not a clear channel due to the noise from background processes such as top quark decays, W and Z bosons with jets, dibosonic decays, single top decays etc. This problem is managed in the CMS experiment by looking for signatures where the Higgs is produced in association with a vector boson which significantly reduced the background. The Feynman diagram is shown in the figure 5.4 as being Higgs radiated from a single top quark.

FIGURE 5.4: $H \rightarrow b\bar{b}$

5.3.2 $H \longrightarrow WW$

The Higgs boson decaying to W^+W^- has a large branching ratio fraction which makes it suitable for precision measurements of Higgs couplings and cross section. The W bosons decaying to a lepton and jets [37] is studied in this document. It has more statistics and thus easily analysed but the contamination from the background processes are large. The main backgrounds to this decay are from processes like W +Jets, Z +Jets, top anti-top decays, hadronic and leptonic decays, single top with a W boson, diboson etc.

In W+Jets, one of the W boson can decay semi-leptonically and the other hadronically leaving the same signature of $H \rightarrow WW$. Also, $t\bar{t}$, single top produced with a W boson decay in a similar manner. The Z+Jets can go a pair of leptons where one of them can go unidentified or into a pair of quarks. The Feynman diagram is shown in the figure 5.5 as being Higgs produced in the fusion of a $t\bar{t}$ pair.

FIGURE 5.5: $H \rightarrow WW$

Chapter 6

Analysis

6.1 Object Selection

The analysis of data/MC involves several steps of processing the raw data obtained from the detector and undergone many steps by the time the data is ready for the physics measurements. This skimming of data mainly involves the following selections:

6.1.1 Trigger

For selection of muon and jets channel ($\mu + jets$), the trigger 'IsoMu27' [38] is used. It implies that isolated muons with $p_T > 27\text{GeV}$ are selected. It is important to isolate muons and confirm that they are not from b-jets, thus selecting the prompt muons.

For the selection of electron and jets channel ($e + jets$), the trigger used is 'Ele32_WPTight_Gsf_L1DoubleEG' which implies prompt electrons of $p_T > 32\text{GeV}$ are selected. The tight working point (WPTight) means that there is a low rate of fake electrons and thus more probability for the electron selected to be a prompt electron and not from jets.

6.1.2 Identification (ID)

In addition to the trigger, cut based ID is used in selecting muons [39] and electrons [40] in order to reduce the fake muons passing the selection criteria. Additionally, we also require the following selection cuts: For electrons, $p_T^e > 35\text{GeV}$ and $|\eta^e| < 2.1$. For muons, $p_T^\mu > 30\text{GeV}$ and $|\eta^\mu| < 2.4$.

6.1.3 Isolation

The electrons already include a selection criteria by applying the tight WP in the identification as given in the section 6.1.1. The selection criteria of the tight WP is applied to muons separately and thus, prompt muons are selected. As a result, it guarantees both electrons and muons are surrounded by an empty cone of particles.

6.1.4 Jets Selection

The jets selection criteria is collected in the Ref. [41] which is the ID recommended by the CMS collaboration. The sample is selected with at least 2 jets. Thus, our sample contains at least 2 jets selected with the leading jet with $p_T > 40\text{GeV}$ and the second leading jet with $p_T > 30\text{GeV}$.

6.1.5 b-jet tagging

It is important to identify the heavy flavour b-jets from the light flavour jets for the categorization of the events [42]. The algorithm used to identify the heavy flavour b-jets is called as the DeepCSV algorithm [43]. It is generated using the deep learning technique and CSV stands for Combined Secondary Vertex. The output is the value between 0 to 1 and thus there are three working points (WP) defined. For the 2017 CMS data, they are

- Loose WP : DeepCSV = 0.1522
- Medium WP: DeepCSV = 0.4941
- Tight WP : DeepCSV = 0.8001

The figure 6.1 represents the selection of b-jets based on the WPs where the signal denotes the b-jets and the background denotes the light flavour jets.

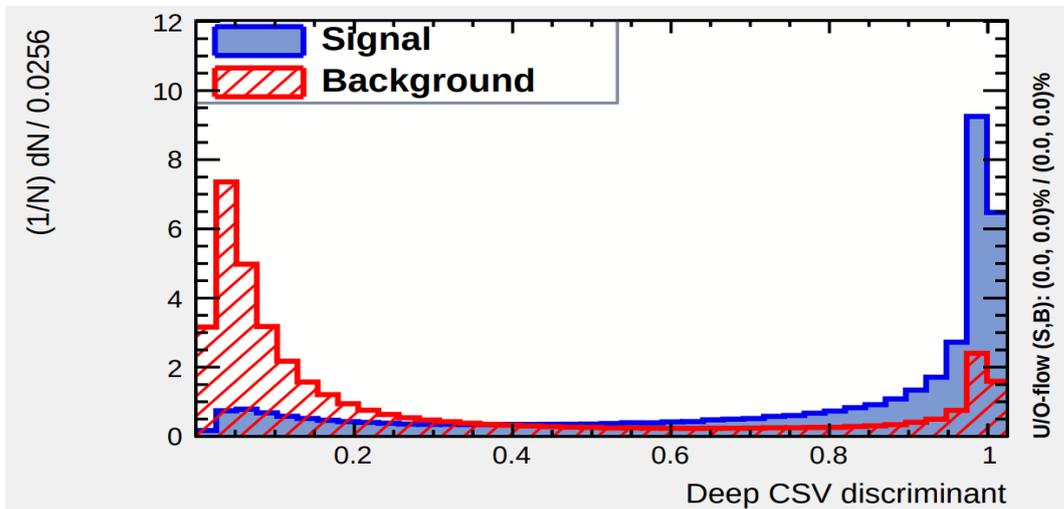


FIGURE 6.1: Deep CSV

For our analysis, we use Deep CSV algorithm with Medium WP.

6.2 Event Selection

The events of interest in the current study is the $t\bar{t}$, $H \rightarrow WW$, $t\bar{t} + b\bar{b}$ and $t\bar{t}H$ with Higgs decaying to two b-quarks. The processes are studied in the channel with one lepton and jets. The signature of each event and the regions selected are explained below.

In the semi-leptonic channel of decay, a single lepton (an electron or a muon), the associated MET (neutrino) and atleast two jets are expected for $t\bar{t} \rightarrow l\nu + jets$, $H \rightarrow WW$, $t\bar{t} + jets$ and $t\bar{t}H$. The $t\bar{t}H$ channel where the $H \rightarrow b\bar{b}$ is analysed. Since the $t\bar{t}$ decay is always associated with jets, the analysis of $t\bar{t} + jets(t\bar{t}jj)$ where $jj = b\bar{b}$ is equivalent to the analysis of $t\bar{t}H \rightarrow b\bar{b}$. A brief explanation about the $t\bar{t} + jets$ categorization is given in the section 4.4.

In the top anti-top decay into two b-jets and two W bosons, where one of the W bosons decay into a lepton (ℓ) and its neutrino (ν) and the other decay hadronically to produce two light flavour (LF) jets, the signal events will be accumulated in the

region of atleast 4 jets where atleast two of them are b-jets. Thus, the signal region (SR) is selected as the region consisting of number of jets > 4 and number of b-jets > 2 . The neutrino associated with the lepton is accounted by the $MET > 30\text{GeV}$.

In the decay channel of Higgs to a pair of W bosons with a W boson decaying to a lepton and the other to LF jets, the SR is selected with 2 LF jets and 0 b-jets. The neutrino associated with the lepton here is also accounted by the $MET > 30\text{GeV}$.

In the ttH channel with the Higgs decaying to a pair of b quarks similar to the signature of $t\bar{t}$ event accompanied with a pair of b-quarks, atleast 4 jets and atleast 2 b-jets are expected from $t\bar{t}$ decay with extra two b-jets, thus leading to the selection of the SR with atleast 6 jets and atleast 2 and atleast 4 b-jets.

The information explained above is summarised in the table 6.1 for the each decay channel analysed.

TABLE 6.1: Signal Regions (SR)
j=jets, b=b-jets

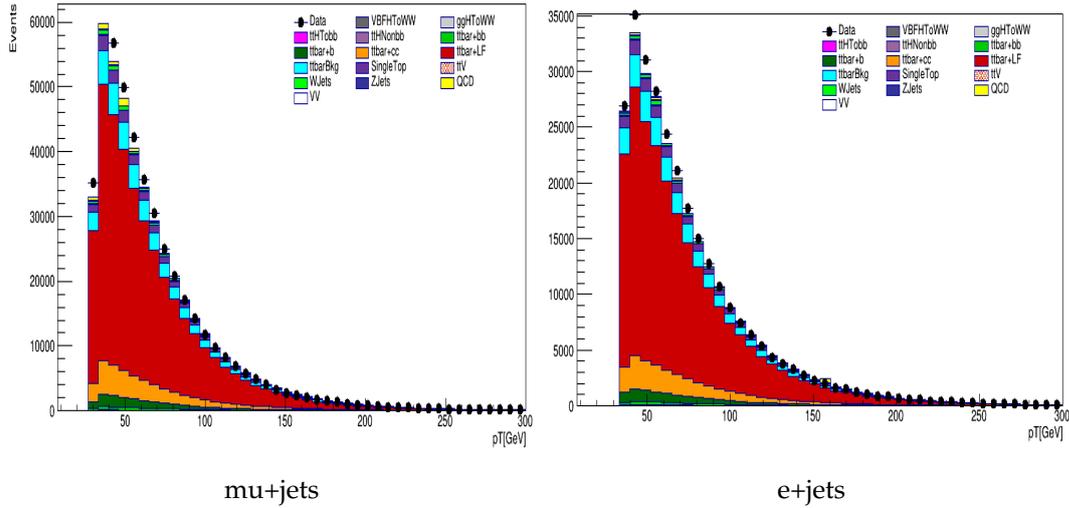
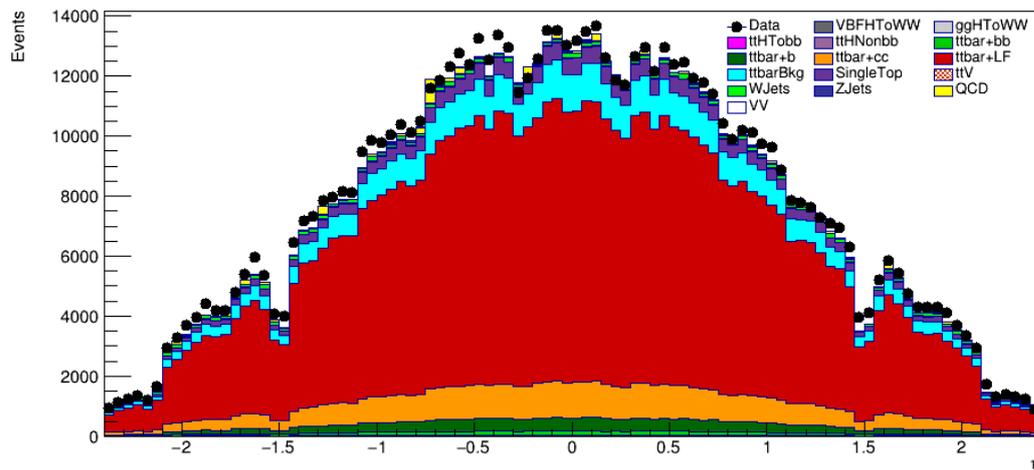
Channels	Products	SR
$t\bar{t} \rightarrow l\nu + \text{jets}$	$l\nu + qq + b\bar{b}$	$(\geq)4j + (\geq)2b$
$H \rightarrow WW$	$l\nu + qq$	$2j + 0b$
$ttH \rightarrow l\nu + \text{jets}$ ($t\bar{t} + b\bar{b}$)	$l\nu + qq + b\bar{b}b\bar{b}$	$(\geq)6j + (\geq)2b$ $(\geq)6j + (\geq)4b$

In the ttH process, looser selection is applied at first, with at least 2-b jets region to have increased efficiency and a tighter selection with at least 4-b region so that there is more purity in the events selected. This helps to compare and analyse the differences while estimating cross section.

The $t\bar{t}$ and the $H \rightarrow WW$ are separately studied to prepare a ttH analysis. Exploring these channels without analysing the $ttH \rightarrow b\bar{b}$ alone is done so that the cross-section and properties of $t\bar{t}$ and Higgs boson can be studied with channels having a balance between statistics and background and can be compared here. The ttH process is very rare accounting for only 1% of the total Higgs production. Among that, $ttH \rightarrow b\bar{b}$ is studied in case of ttH and $H \rightarrow WW$ is studied when H decay channel alone is considered because of the very low statistics of the process as mentioned above and $ttH \rightarrow b\bar{b}$ is the leading decay mode.

6.3 Control Plots

Since the measurements are done in the semi-leptonic channel, the objects such as a lepton, its associated neutrino and jets are being dealt with. The plots are given in each of the SR chosen as given in the section 6.2. Note that the QCD calculations are deducted from Monte Carlo (MC) due to the limitation of time.

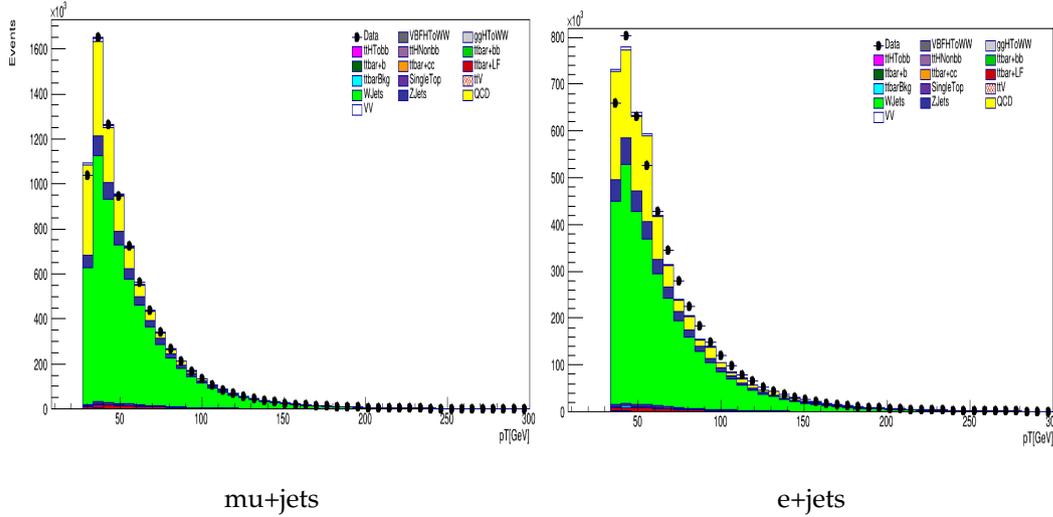
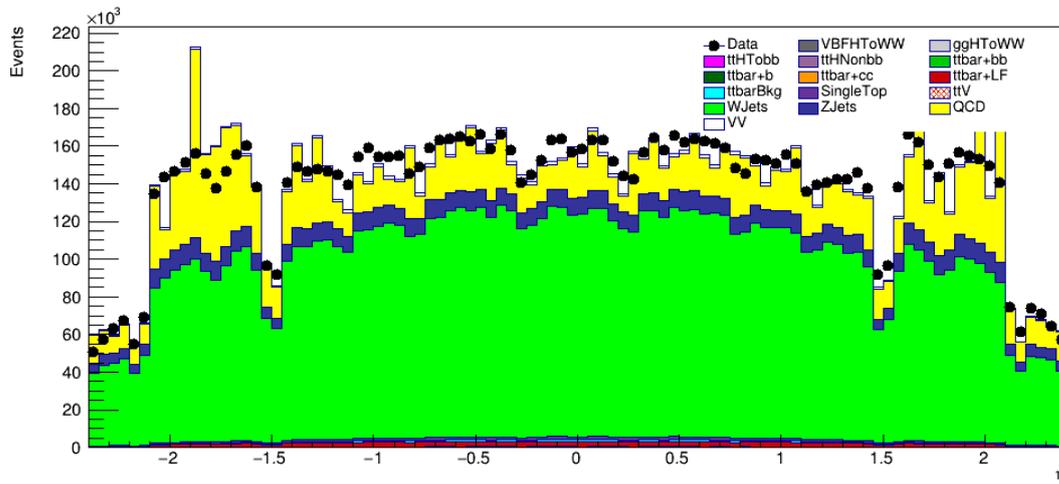
FIGURE 6.4: Lepton Transverse momentum (p_T), $\geq 4j + \geq 2b$ FIGURE 6.5: η of jets in $\ell + jets, \geq 4j + \geq 2b$

The pseudorapidity (η) distribution of the lepton is shown below in the above figure 6.5 and is seen that it is almost symmetric as expected and shows a good agreement between the data and MC.

6.3.3 Signal Region with $2j$ & $0b$ ($H \rightarrow WW$ like)

In the lepton p_T distribution (Figure. 6.6), the data to MC agreement is fair enough, but the signal events of $H \rightarrow WW$ are too few making high background contamination. Thus the background process, particularly the QCD can significantly affect the analysis of Higgs events and its cross section estimation.

The η distribution of the lepton shown in the figure 6.7 is almost symmetric. There are much fluctuations due to the QCD which affects the agreement between the data and MC.

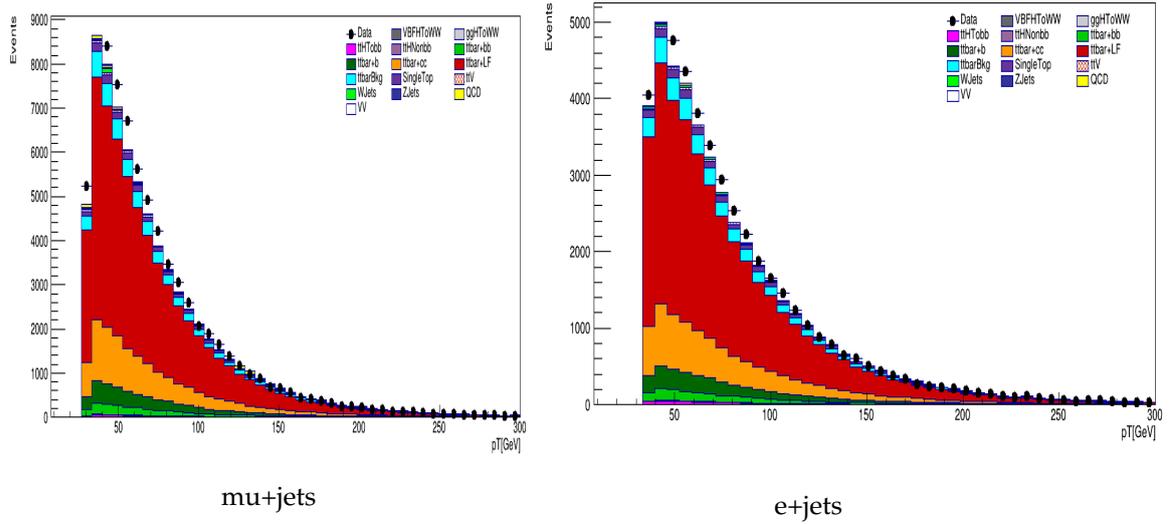
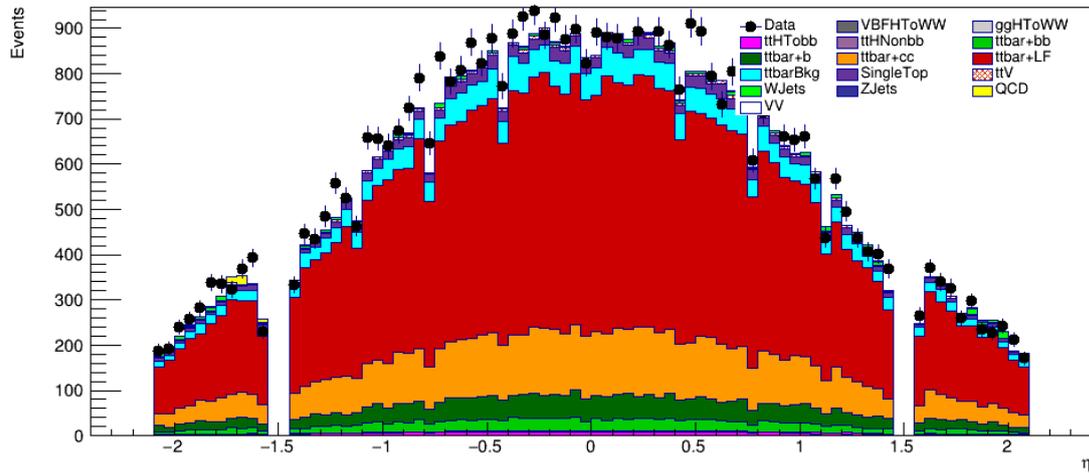
FIGURE 6.6: Lepton Transverse momentum (p_T), 2j0bFIGURE 6.7: η of jets in $\ell + jets$, 2j0b

6.3.4 Signal Region with (at least) 6j

6.3.4.1 At least 6j & 2b ($ttH/ttbb$ like)

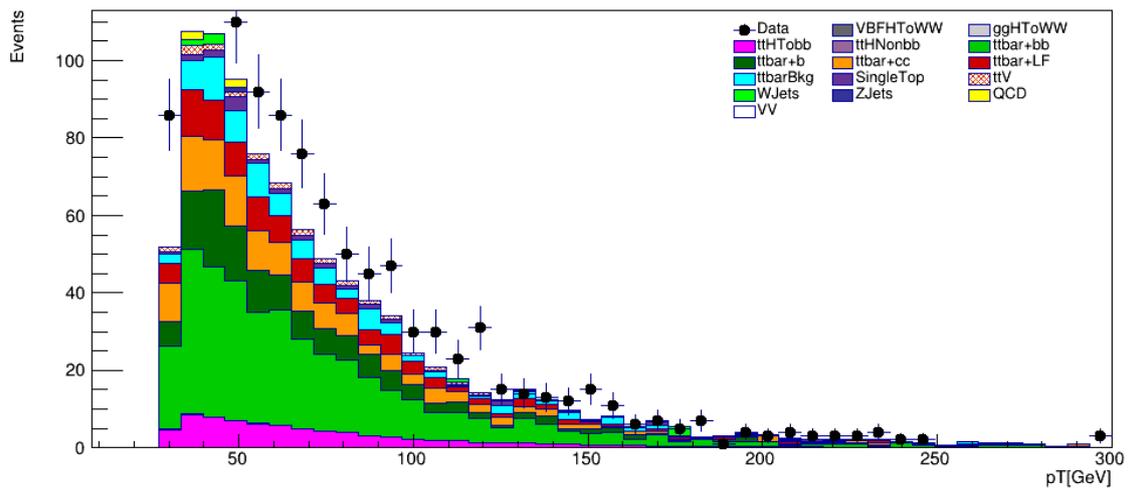
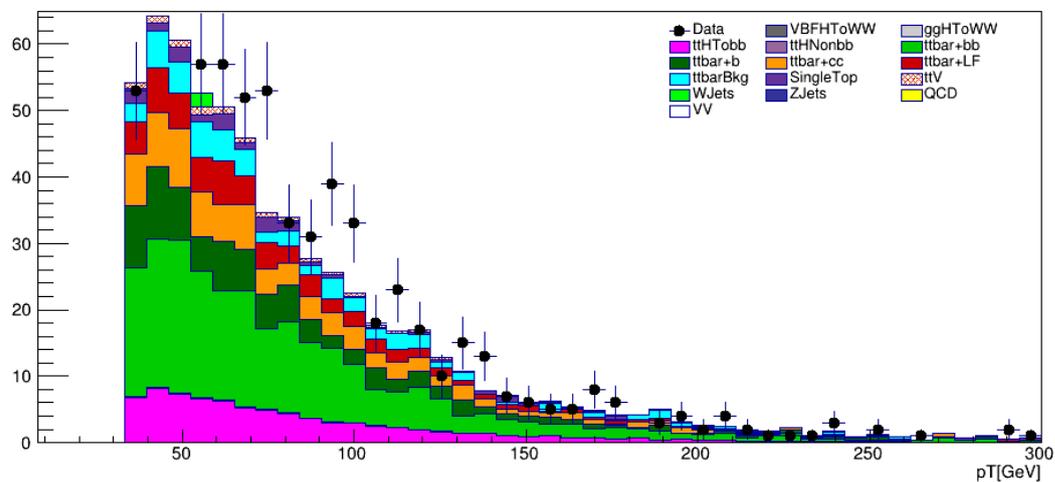
In the lepton p_T distribution (Figure. 6.8), the data to MC agreement is fair enough, but the signal events of ttH are a few and there is significant contributions from the background processes.

The η distribution plot of the lepton depicted in the figure 6.9 is almost symmetric.

FIGURE 6.8: Lepton Transverse momentum (p_T), $\geq 6j + \geq 2b$ FIGURE 6.9: η of jets in $\ell + jets, \geq 6j + \geq 2b$

6.3.4.2 At least 6j & 4b ($ttH/ttbb$ like- tight)

Lets also analyse the p_T plots in region of at least 6 jets with 4 b-jets. In the following plots, the $ttH \rightarrow bb$ can seem to dominate over the light flavour decays as expected in this region. The data events are more than the MC events which leads to an increase in the measured cross section as compared to the theoretical value at the defined regions.

FIGURE 6.10: Lepton Transverse momentum (p_T), μ +jetsFIGURE 6.11: Lepton Transverse momentum (p_T), e +jets

In each of these plots, the number of data events dominate a little over the MC events which will be reflected in the estimation of cross sections. Thus, a cross section greater than the theoretical cross section is expected at 13 TeV for different decay channels.

Chapter 7

Estimation of Cross Section

The chapter deals with the extraction of cross section using the 2017 CMS data at an integrated luminosity of $\mathcal{L} = 41.5\text{fb}^{-1}$ and at a centre of mass energy of $\sqrt{s} = 13\text{TeV}$. The cross section is calculated using the cut and count method and the fit method for

- Top anti-top pair production ($t\bar{t}$)
- Higgs boson decay to two W bosons ($H \rightarrow WW$)
- Top anti-top pair production with a pair of b-jets ($t\bar{t} + b\bar{b}$)
- ttH decay with $H \rightarrow b\bar{b}$

The last two processes share the same signature. Thus, the semi-leptonic channels for each of the above decay are analysed and results are presented for both muon+jets ($\mu + jets$) and electron+jets ($e + jets$).

For the present analysis, the systematic uncertainties are not calculated and QCD calculations are only deducted from Monte Carlo (MC) due to the limitation of time.

To extract the cross section, two different approaches are used,

1. Cut and Count (C & C) method
2. Fit Approach: Template fit over a discriminant variable

7.1 Cut and Count Analysis

A region of optimal signal events with the least background contribution is selected for each decay process to be analysed. Then, the number of signal events and background events are counted after applying the selection cuts. At the present defined value of \mathcal{L} and \sqrt{s} , the cross section is estimated using the expression as explained in the following section.

Cross Section (σ)

Cross section is the probability that a particular interaction will occur, that is, in simple terms, the counting of the desired number of events that were produced when a given number of protons collided at each other, in case of the LHC [44].

The number of events per second (N_{evt}) occurred in a particular interaction can be extracted from the product of integrated luminosity (\mathcal{L}) and cross section (σ). Thus,

$$N_{evt} = \sigma \cdot \mathcal{L} \quad (7.1)$$

where the luminosity is the integrated luminosity. Another important quantities in the accurate measurement of cross section are the efficiencies associated with the detector performance and geometry.

- Efficiency (ϵ) It is related to the performance of the detector. It is the ratio of the number of signal events that pass the selection cuts to the number of generated signal events expected in the same region.
- Acceptancy (\mathcal{A}) It is the fraction of signal events in the vicinity of the detector. So, it is the ratio of the events that the detector received to the total number of generated signal events.

Thus, the full space cross section equation becomes,

$$\sigma = \frac{N_{data} - N_{bkg}}{\mathcal{L} \cdot (\epsilon \cdot \mathcal{A} \cdot Br)} \quad (7.2)$$

where N_{data} is the number of data events, N_{bkg} is the sum of events due to the contribution from all background sources and Br is the fraction of signal events that decay semi-leptonically (accounts for 43%).

And the product of efficiency and the acceptancy,

$$\epsilon \cdot \mathcal{A} \cdot Br = \frac{N_{reco}}{N_{gen}} \quad (7.3)$$

where N_{reco} is the number of reconstructed signal events from MC simulation. The information of the total number of generated signal events (N_{gen}) is obtained from the MC simulation and is the product of the theoretical cross section (σ_{th}) and \mathcal{L} as given in equation 7.4. Thus, it is an independent quantity for a given interaction and detector.

$$N_{gen} = \sigma_{th} \cdot \mathcal{L} \quad (7.4)$$

Therefore, the expression of the σ inferred using the equations 7.2, 7.3 and 7.4 is,

$$\sigma = \frac{N_{data} - N_{bkg}}{N_{reco}} \cdot \sigma_{th} \quad (7.5)$$

In the ideal case, the ratio between the measured and theoretical cross section is 1.

7.1.1 $t\bar{t}$ Cross Section ($\sigma_{t\bar{t}}$)

The theoretical cross section value for $t\bar{t}$ production at $\sqrt{s} = 13\text{TeV}$ is $\sigma_{t\bar{t}}^{th} = 831.76\text{pb}$ [45]. The SR selected for this decay channel is at least 4 jets with at least 2 b-jets (section 6.2). From the control plot (Figure 6.4), for $\mu + jets$,

$$N_{reco}^{t\bar{t}MC} = 377080 \quad (7.6)$$

$$N_{data} - N_{bkg} = 464705 - 74404.8 \quad (7.7)$$

$$= 390300.2 \quad (7.8)$$

$$\sigma_{t\bar{t}}^{\mu j} = \frac{390300.2}{377080} \cdot \sigma_{t\bar{t}}^{th} \quad (7.9)$$

$$\approx 861\text{pb} \quad (7.10)$$

Using the same approach, for $e + jets$ and thus for $\ell + jets$,

$$\sigma_{t\bar{t}}^{ej} \approx 851\text{pb} \quad (7.11)$$

$$\sigma_{t\bar{t}}^{all} \approx 857\text{pb} \quad (7.12)$$

7.1.2 $H \rightarrow WW$ Cross Section ($\sigma_{H \rightarrow WW}$)

The theoretical cross section value is $\sigma_{H \rightarrow WW}^{th} = 32.371\text{pb}$, which is the sum of the contributions from the vector boson fusion (VBF) and gluon-gluon (gg) fusion production mechanisms [46]. The value of the cross section times the branching ratio of gg Higgs decaying to two W bosons, $\sigma_{ggH \rightarrow WW}^{th} = 28.87\text{pb}$ and of VBF Higgs is $\sigma_{VBFH \rightarrow WW}^{th} = 3.447\text{pb}$. Since it is mostly impossible to calculate VBF Higgs decay cross section separately with this method due to its very low statistics, two can be merged and analysed. The SR selected for this decay channel is with exactly 2 jets and 0 b-jets (section 6.2).

From the plots displayed in the section 6.3.3, the signal events are negligible when compared to the background contributions. Due to this high contamination, the cross section observed is compatible with 0pb.

7.1.3 $t\bar{t} + b\bar{b}$ Cross Section ($\sigma_{t\bar{t}b\bar{b}}$)

The theoretical cross section value is $\sigma_{t\bar{t}b\bar{b}} \approx 4\text{pb}$. The SRs selected for this process are same as that of $t\bar{t}H$ decaying to a $b\bar{b}$ pair since they share the same signature (section 6.2).

7.1.3.1 Signal Region with (at least) 6j & 2b

For $\mu + jets$,

$$N_{reco}^{t\bar{t}b\bar{b}MC} = 2241.07 \quad (7.13)$$

$$N_{data} - N_{bkg} = 76800 - 70088.2 \quad (7.14)$$

$$= 6711.79 \quad (7.15)$$

$$\sigma_{t\bar{t}b\bar{b}} = \frac{6711.79}{2241.07} \cdot \sigma_{t\bar{t}b\bar{b}}^{th} \quad (7.16)$$

$$\approx 11.979\text{pb} \quad (7.17)$$

Using the same approach, for $e + jets$,

$$\sigma_{t\bar{t}b\bar{b}}^{ej} \approx 8.984\text{pb} \quad (7.18)$$

and thus for $\ell + jets$,

$$\sigma_{t\bar{t}b\bar{b}}^{all} \approx 10.482\text{pb} \quad (7.19)$$

7.1.3.2 Signal Region with (at least) 6j & 4b

For $\mu + jets$,

$$N_{reco}^{ttbb_{MC}} = 359.764 \quad (7.20)$$

$$N_{data} - N_{bkg} = 1179 - 557.385 \quad (7.21)$$

$$= 621.615 \quad (7.22)$$

$$\sigma_{ttbb} = \frac{621.615}{359.764} \cdot \sigma_{ttbb}^{th} \quad (7.23)$$

$$\approx 6.911\text{pb} \quad (7.24)$$

Using the same approach, for $e + jets$,

$$\sigma_{ttbb}^{ej} \approx 5.648\text{pb} \quad (7.25)$$

and thus for $\ell + jets$,

$$\sigma_{ttbb}^{all} \approx 6.473\text{pb} \quad (7.26)$$

There is a small discrepancy in the value of cross section calculated in both the channels in this decay too. By the comparison between the regions of at least 2b and 4b, it can be seen that the deviation from σ_{ttbb}^{th} is less for the region of 4b as expected.

7.1.4 ttH decay with $H \rightarrow b\bar{b}$ Cross Section (σ_{ttH})

The theoretical cross section value is $\sigma_{ttH} = 0.2934\text{pb}$ [46]. Two SRs are selected for this process, one is the looser selection with at least 6 jets and 2 b-jets and the other is the tighter with at least 6 jets and at least 4 b-jets (section 6.2).

7.1.4.1 Signal Region with (at least) 6j & 2b

For $\mu + jets$,

$$N_{reco}^{ttH_{MC}} = 447.359 \quad (7.27)$$

$$N_{data} - N_{bkg} = 76800 - 71881.9 \quad (7.28)$$

$$= 4918.08 \quad (7.29)$$

$$\sigma_{ttH} = \frac{4918.08}{447.359} \cdot \sigma_{ttH}^{th} \quad (7.30)$$

$$\approx 3.226\text{pb} \quad (7.31)$$

Using the same approach, for $e + jets$,

$$\sigma_{ttH}^{ej} \approx 1.475\text{pb} \quad (7.32)$$

and thus for $\ell + jets$,

$$\sigma_{ttH}^{all} \approx 2.350\text{pb} \quad (7.33)$$

7.1.4.2 Signal Region with (at least) 6j & 4b

For $\mu + jets$,

$$N_{reco}^{ttH_{MC}} = 75.41 \quad (7.34)$$

$$N_{data} - N_{bkg} = 1179 - 841.74 \quad (7.35)$$

$$= 337.26 \quad (7.36)$$

$$\sigma_{ttH} = \frac{337.26}{75.41} \cdot \sigma_{ttH}^{th} \quad (7.37)$$

$$\approx 1.312\text{pb} \quad (7.38)$$

Using the same approach, for $e + jets$,

$$\sigma_{ttH}^{ej} \approx 0.637\text{pb} \quad (7.39)$$

and thus for $\ell + jets$,

$$\sigma_{ttH}^{all} \approx 0.978\text{pb} \quad (7.40)$$

There is a discrepancy in the value of cross section calculated in $\mu + jets$ and $e + jets$ channel. But it will be covered when the systematic uncertainties are considered as we can expect nearly 50% uncertainty here due to high contribution from background sources and low statistics of signal events.

The cross section calculated with the region with at least 6 jets and 4 b-jets has less deviation from the σ_{ttH}^{th} than the region with 2 b-jets. Even though the signal events are a few in this region, a closer value of σ is obtained as we expect at least 2 b-jets from top decays and other two from Higgs decays in this particular channel.

7.2 Fit Approach

A fit is an efficient method to explore the shape/structure of the physics models, measurement and prediction. It adjusts the theoretical prediction to the experimental data by finding values for its parameters that minimize the differences between the experimental data and predicted MC events. With the observables calculated within the model, likelihoods can be calculated with a given measurement and prediction such as for the cross section of an interaction. A fit improves the cross section extraction by exploiting the different Signal/Background shapes for some observables.

In high energy physics (HEP), the one of the widely used methods is the RooFit [47]. In particle physics analyses, we need to use all the statistical information in the data and combination of the background suppression along with the multidimensional fitting. Also, the probability density functions (PDF) do not have explicit formulas for our analysis models, hence RooFit is the best choice.

The RooFit library of ROOT [48] provides toolkit to model the expected distribution of events in HEP analysis. It is a powerful tool for creating PDFs and fitting them to the binned data. It is primarily designed as a particle physics data analysis tool but has wide use in other fields too.

An optimal signal region (SR) is selected for each decay channel of our interest of

study as done in cut and count analysis. It is explained in the section 6.2. Then a variable which provides the best discrimination between signal and background is chosen and this variable can be Number of jets, Number of b-jets, Invariant mass etc. Using the PDF of each of the signal and backgrounds, a physics model dependent on the cross section of the interaction is constructed and the fit is performed. Thus the cross section can be extracted and the post fitted model to the data is plotted for the chosen discriminant variable. The systematic uncertainties are not considered.

7.2.1 Inclusive $t\bar{t}$ Cross Section ($\sigma_{t\bar{t}}$)

A physics model based on the PDFs of the $t\bar{t}$ signal and background is constructed which has the following form,

$$\text{Data} = k \cdot \text{PDF}_{t\bar{t}} + k \cdot \text{PDF}_{t\bar{t}Bkg} + \text{PDF}_{SMBkg} \quad (7.41)$$

where k , the signal strength is the ratio of the measured value to theoretical value of cross section,

$$k = \frac{\sigma_{t\bar{t}}}{\sigma_{t\bar{t}}^{th}} \quad (7.42)$$

$\text{PDF}_{t\bar{t}Bkg}$ denotes the background from the hadronic and dileptonic decays of $t\bar{t}$ which also has the effect of $\sigma_{t\bar{t}}$ and $\text{PDF}_{t\bar{t}}^{SMBkg}$ denotes the background contribution due to all other SM processes other than $t\bar{t}$ background as explained in the section 4.3.1.

First, let's consider 'Number of Jets' as the discriminant variable and perform the fitting. The value of k , thus the $\sigma_{t\bar{t}}$ is calculated and the post-fitting plot is drawn.

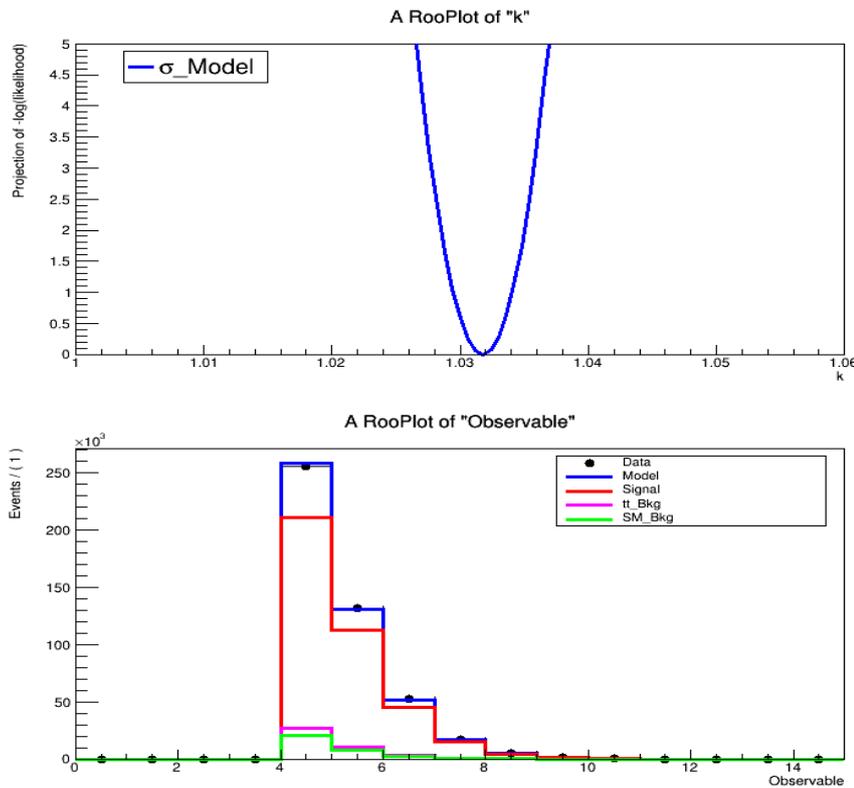


FIGURE 7.1: No. of jets in $\mu + jets, \geq 4j + \geq 2b$

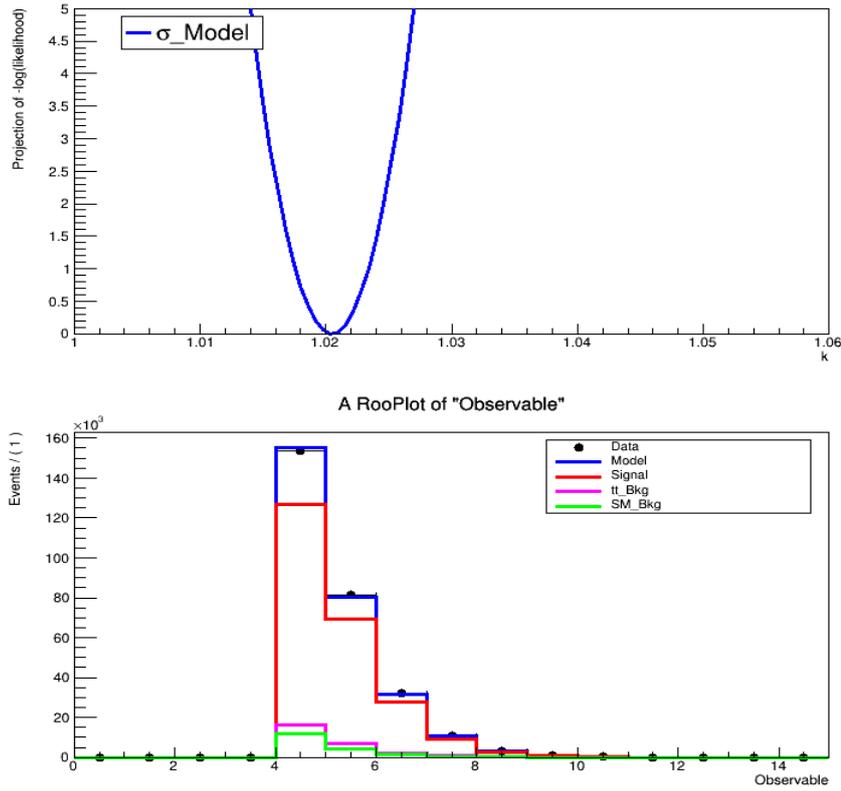


FIGURE 7.2: No. of jets in $e + jets, \geq 4j + \geq 2b$
 From the figure 7.1, for $\mu + jets$ channel, the value of $k \approx 1.03$ and thus,

$$\sigma_{t\bar{t}}^{\mu j} = 1.03 \cdot \sigma_{t\bar{t}}^{th} \quad (7.43)$$

$$= 1.03 \cdot \sigma_{t\bar{t}}^{th} \quad (7.44)$$

$$\approx 858\text{pb} \quad (7.45)$$

Similarly, from the figure 7.2, for $e + jets, k \approx 1.02$ and thus,

$$\sigma_{t\bar{t}}^{ej} \approx 849\text{pb} \quad (7.46)$$

and thus for $\ell + jets,$

$$\sigma_{t\bar{t}}^{all} \approx 854\text{pb} \quad (7.47)$$

The number of jets may not be the best discriminant variable. Lets consider the discriminant variable to be the invariant mass of the whole system (sum of the invariant mass of the dijet system and lepton-neutrino system) and reanalyse the plots.

From the figure 7.3, for $\mu + jets$ channel, the value of $k \approx 1.03$ and thus,

$$\sigma_{t\bar{t}}^{\mu j} = 1.03 \cdot \sigma_{t\bar{t}}^{th} \quad (7.48)$$

$$\approx 859\text{pb} \quad (7.49)$$

Similarly, from the figure 7.4, for $e + jets, k \approx 1.02$ and thus,

$$\sigma_{t\bar{t}}^{ej} \approx 850\text{pb} \quad (7.50)$$

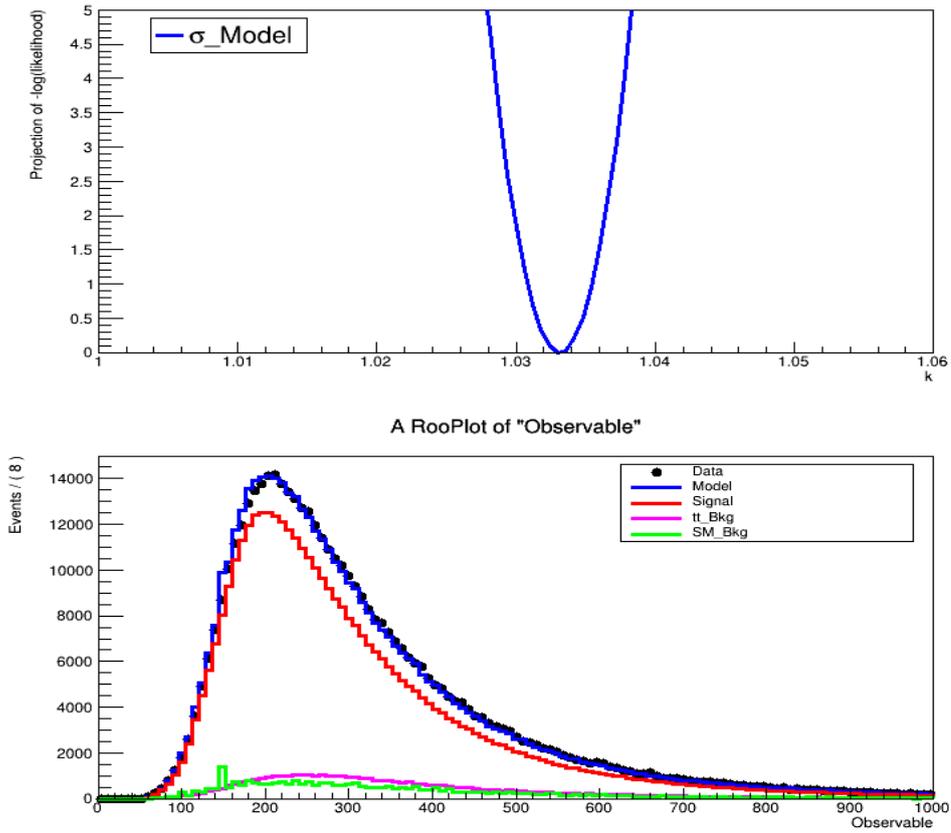


FIGURE 7.3: Invariant mass in $\mu + jets, \geq 4j + \geq 2b$

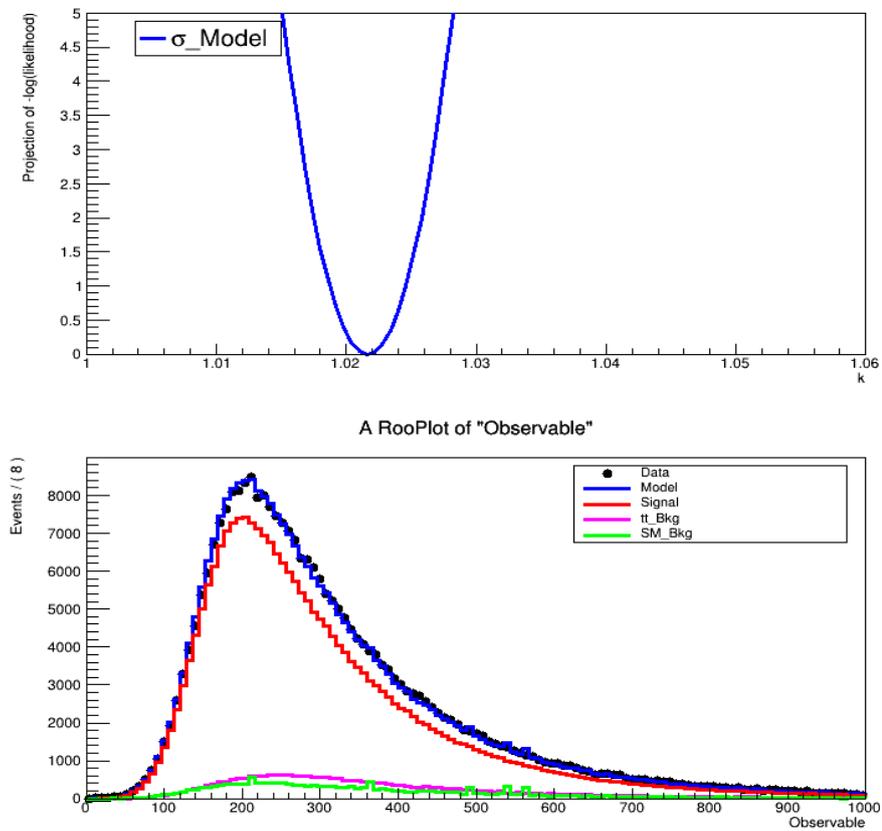


FIGURE 7.4: Invariant mass in $e + jets, \geq 4j + \geq 2b$

The values of cross sections when the discriminant variable is the number of jets or the invariant mass do not seem to have much difference.

7.2.2 $H \rightarrow WW$ Cross Section ($\sigma_{H \rightarrow WW}$)

The physics model created with signal as sum of the contributions of Higgs from gluon-gluon and Vector Boson fusion and the other SM background processes including $t\bar{t}$ and ttH events.

$$\text{Data} = k \cdot \text{PDF}_{HWW} + \text{PDF}_{Bkg} \quad (7.51)$$

where k is the ratio of the measured value to theoretical value of cross section of $H \rightarrow WW$ and PDF_{Bkg} is the background contribution due to all SM processes including $t\bar{t}$ background.

From the figures displayed in the section 6.3.3, the Higgs events are very less when compared to the background processes. Thus, any small fluctuation in the background, particularly QCD since the QCD that is dealt here is only from MC simulation.

Lets consider the invariant mass of the whole system as the discriminant variable for $\ell + jets$. Here, the number of jets as the variable does not make any sense as we create a region of exactly two number of jets.

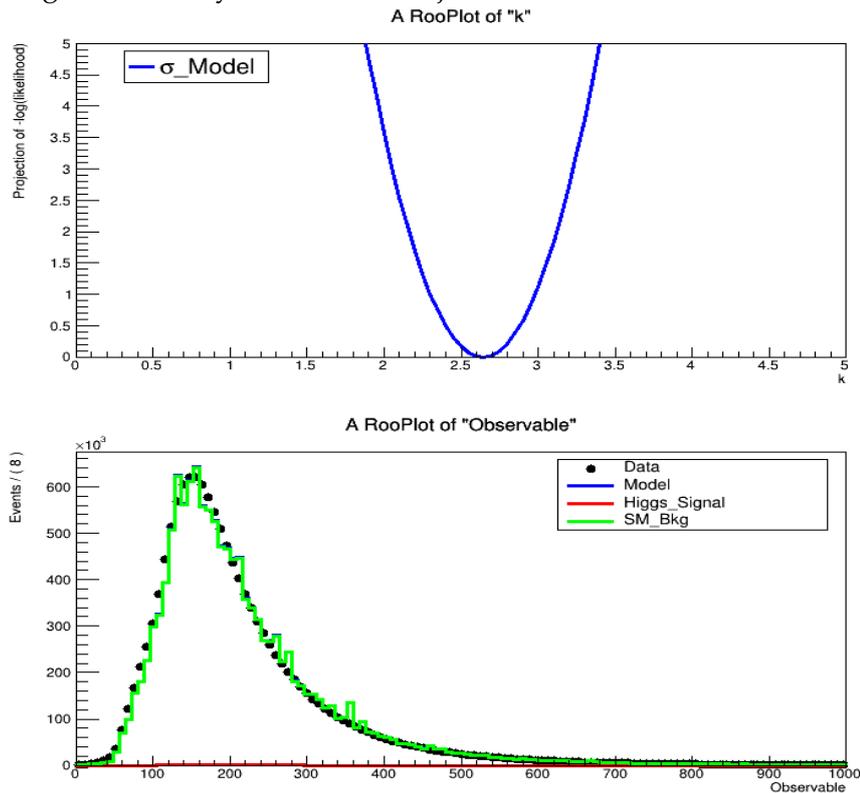


FIGURE 7.5: Invariant Mass in $\ell + jets$, 2j0b

value of $k = 2.64$ and thus

$$\sigma_{HWW}^{all} = 2.64 \cdot \sigma_{H \rightarrow WW}^{th} \quad (7.52)$$

$$\approx 85\text{pb} \quad (7.53)$$

This value makes sense as it is of the same order of the theoretical value unlike using cut and count analysis.

If the discriminant variable is changed to the p_T of the leading jet, the post fit plot is drawn as in the figure 7.6.

It can be seen that the value is very sensitive to the variable chosen and it should not be the case. The cross section will mostly be relied on W +Jets as it is concentrated in this region and QCD as mentioned before. Thus, this approach is not sufficient to calculate the cross section of the decay of Higgs to two W bosons and the main reason is the very low statistics of the channel.

$$\sigma_{HWW}^{all} \approx 223\text{pb} \quad (7.54)$$

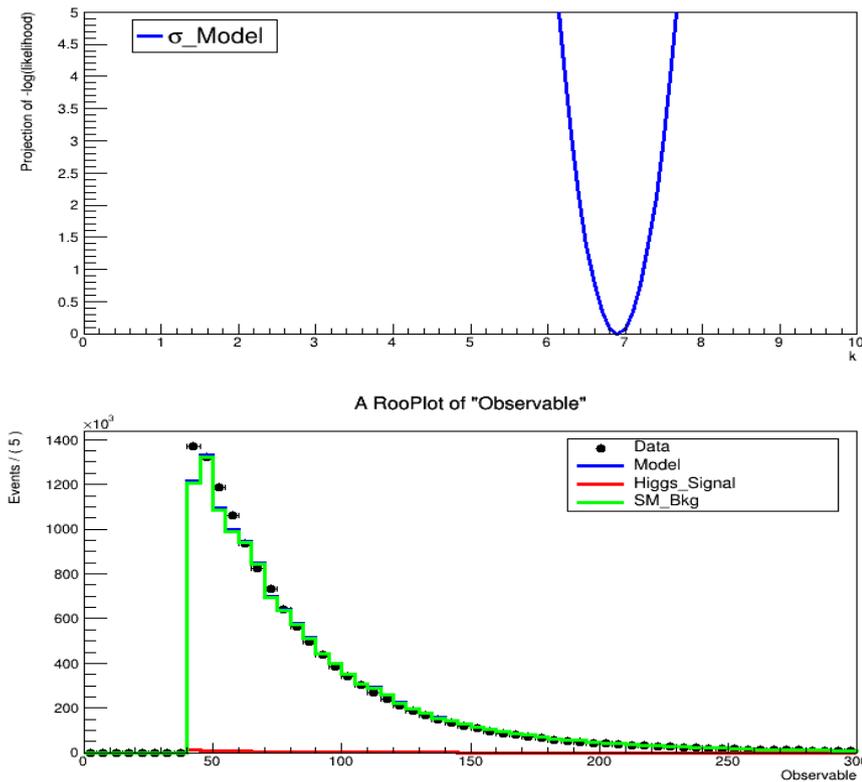


FIGURE 7.6: p_T of leading jet in $\ell + jets, 2j0b$

7.2.3 $t\bar{t}H \rightarrow b\bar{b}$ ($\sigma_{t\bar{t}H}$) & $t\bar{t} + b\bar{b}$ ($\sigma_{t\bar{t}b\bar{b}}$) Cross Section

Two SRs are selected for these decay channels with the same signature, looser selection (at least 6 jets and 2 b-jets) and the other is a tighter selection (at least 6 jets and 4 b-jets) (section 6.2). The physics model created out of the PDFs of the signal and

background has the following general form,

$$\text{Data} = k \cdot \text{PDF}_{\text{signal}} + \text{PDF}_{\text{Bkg}} \quad (7.55)$$

where k is the ratio of the measured value to theoretical value of cross section, $\text{textPDF}_{\text{signal}}$ denotes the PDF of the signal event chosen and $\text{textPDF}_{\text{Bkg}}$ is that of the background corresponding to the signal.

The cross sections of $ttH \rightarrow bb$ and $ttbb$ are extracted in different ways on the basis of how k affects the signals and background and are compared. Let k_{ttbb} be the ratio associated with $ttbb$ and k_{ttH} with the $ttH \rightarrow bb$. When $t\bar{t}$ is divided on the basis of the pair of jets associated with it (refer section 4.4), there are mainly four categories: $ttbb$, ttb , $ttcc$ and $ttLF$. The discriminant variable is number of jets and plots are drawn for $\ell + jets$ for first two models considered and for both $\mu + jets$ and $e + jets$ for the final model.

7.2.3.1 SR with (at least) 6j & 2b

Model with k_{ttbb} affecting $ttbb$ only

The model is defined as,

$$\text{Data} = k \cdot \text{PDF}_{ttbb} + \text{PDF}_{ttb} + \text{PDF}_{ttcc} + \text{PDF}_{ttLF} + \text{PDF}_{ttH} + \text{PDF}_{\text{SMBkg}} \quad (7.56)$$

where $\text{PDF}_{\text{SMBkg}}$ denotes the background contribution due to all other SM processes.

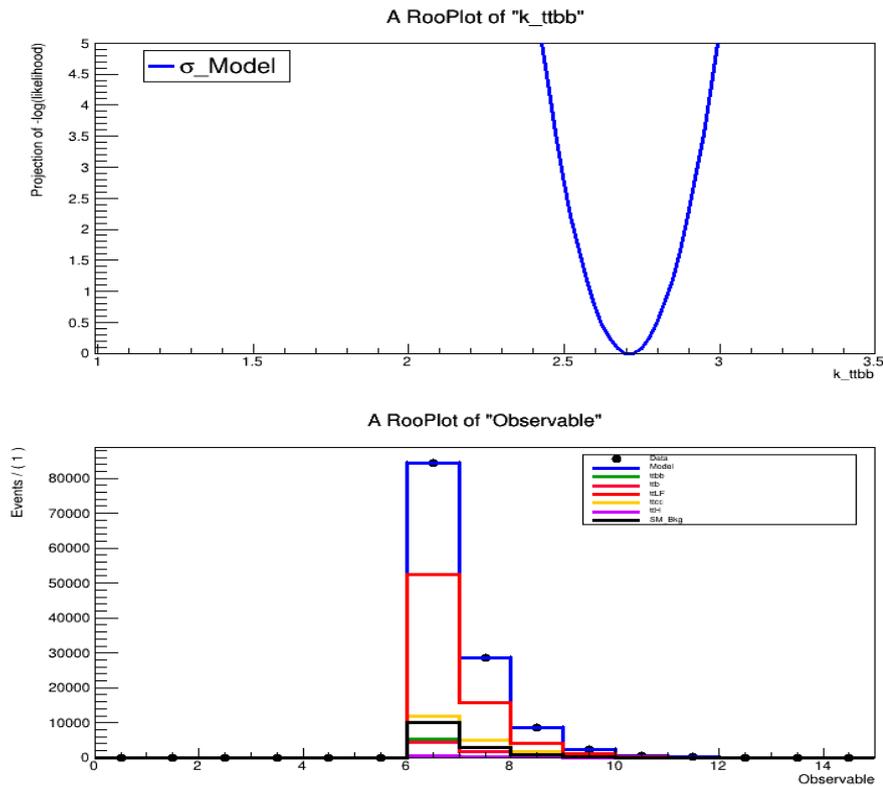


FIGURE 7.7: No. of jets in $\ell + jets$, $\geq 6j+ge2b$

From the above figure, $k_{ttbb} = 2.71$ and thus

$$\sigma_{ttbb}^{all} = 2.71 \cdot \sigma_{ttbb}^{th} \quad (7.57)$$

$$\approx 11\text{pb} \quad (7.58)$$

Model with k_{ttbb} affecting $ttbb$ & ttb

The model is defined as,

$$\text{Data} = k \cdot \text{PDF}_{ttbb} + k \cdot \frac{N_{ttb}}{N_{ttbb}} \text{PDF}_{ttb} + \text{PDF}_{ttcc} + \text{PDF}_{ttLF} + \text{PDF}_{ttH} + \text{PDF}_{SMBkg} \quad (7.59)$$

Since ttb events and $ttbb$ are related, there should be a factor relating these decays, which is the ratio of N_{ttb} to N_{ttbb} where N_{ttbb} and N_{ttb} are the number of reconstructed $ttbb$ and ttb events respectively.

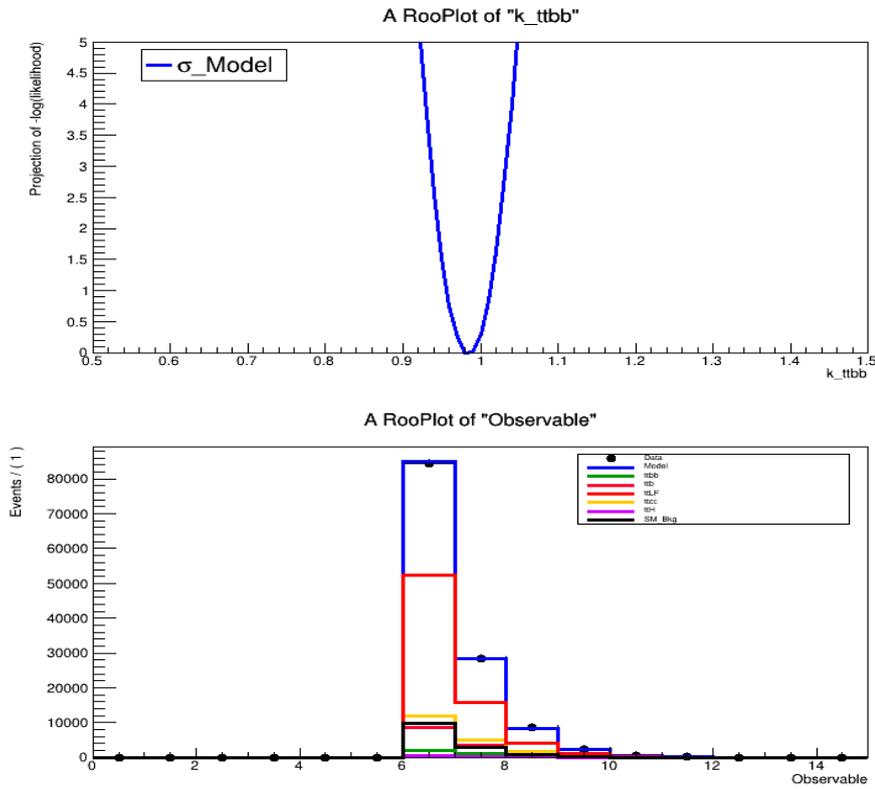


FIGURE 7.8: No. of jets in $\ell + jets, \geq 6j+ge2b$

From the above figure, $k_{ttbb} = 0.98$ and thus

$$\sigma_{ttbb}^{all} = 0.98 \cdot \sigma_{ttbb}^{th} \quad (7.60)$$

$$\approx 3.93\text{pb} \quad (7.61)$$

Model with k_{ttbb} & k_{ttH}

The model is defined as,

$$\text{Data} = k \cdot \text{PDF}_{ttbb} + k \cdot \frac{N_{ttb}}{N_{ttbb}} \cdot \text{PDF}_{ttb} + \text{PDF}_{ttcc} + \text{PDF}_{ttLF} + k_{ttH} \cdot \text{PDF}_{ttH} + \text{PDF}_{SMBkg} \quad (7.62)$$

The discriminant variables chosen is the number of jets. Thus, the value of k_{ttbb} and k_{ttH} are extracted simultaneously and the σ_{ttbb} and σ_{ttH} are calculated. The post-fitting plots are then drawn.

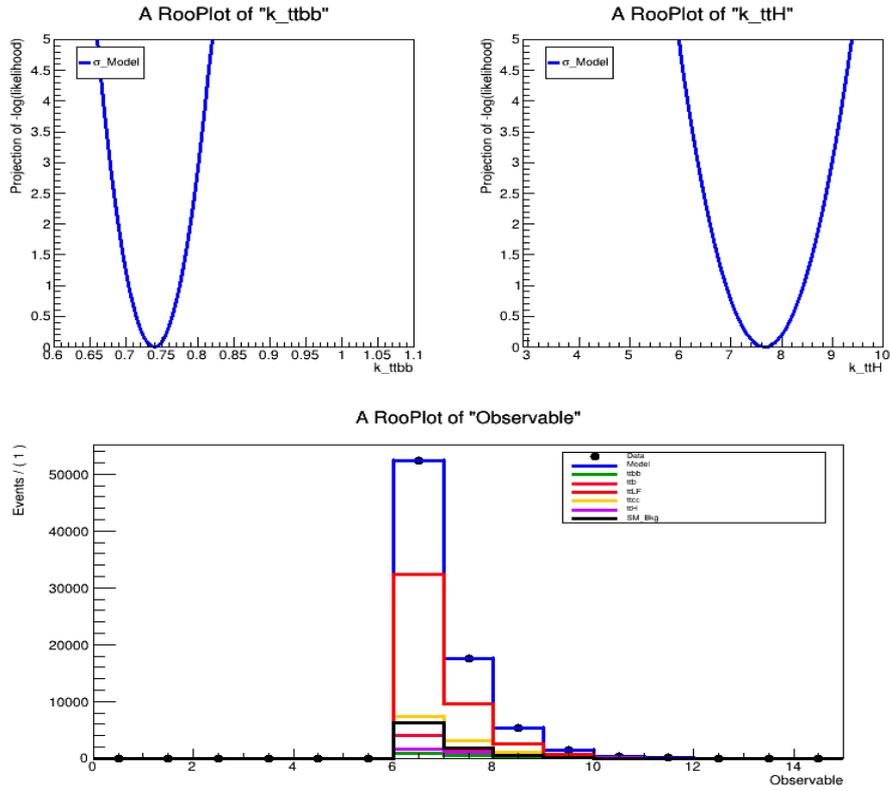


FIGURE 7.9: No. of jets in $\mu + jets, \geq 6j+ge2b$

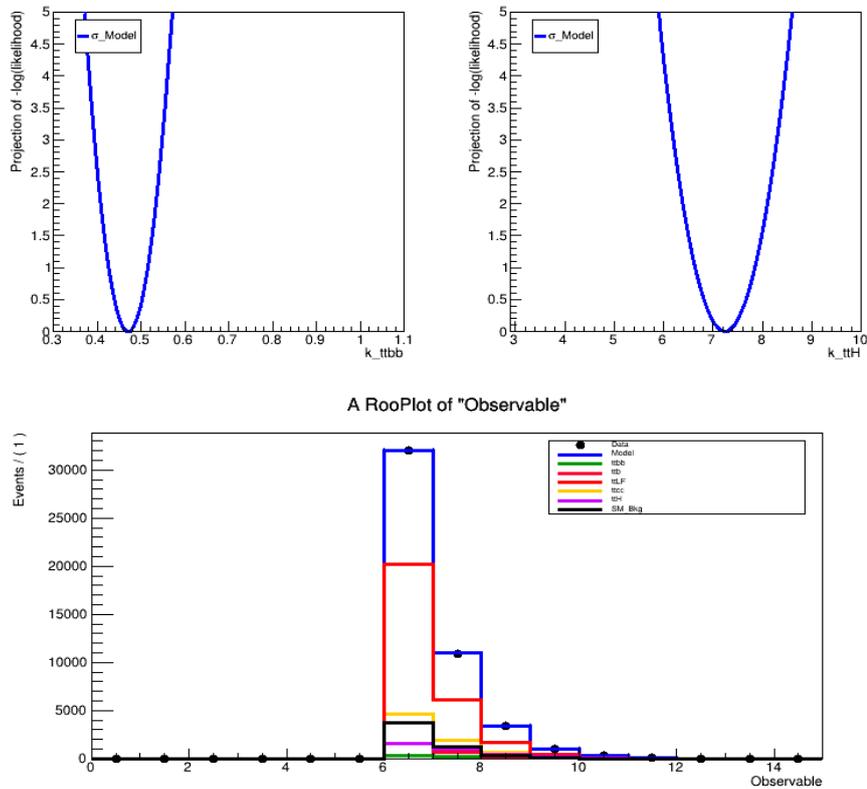


FIGURE 7.10: No. of jets in $e + jets, \geq 6j+ge2b$

In the figure 7.9, for $\mu + jets$ channel,

$$\sigma_{ttbb}^{\mu j} = 0.74 \cdot \sigma_{ttbb}^{th} \quad (7.63)$$

$$\approx 3\text{pb} \quad (7.64)$$

$$\sigma_{ttH}^{\mu j} = 7.67 \cdot \sigma_{ttH}^{th} \quad (7.65)$$

$$\approx 2\text{pb} \quad (7.66)$$

Similarly, from the figure 7.10, for $e + jets$,

$$\sigma_{ttbb}^{ej} = 0.47 \cdot \sigma_{ttbb}^{th} \quad (7.67)$$

$$\approx 2\text{pb} \quad (7.68)$$

$$\sigma_{ttH}^{ej} = 7.24 \cdot \sigma_{ttH}^{th} \quad (7.69)$$

$$\approx 2\text{pb} \quad (7.70)$$

7.2.3.2 Signal Region with (at least) 6j & 4b

The models explained above are constructed for at least 4 jets and 4 b-jets region too and the results are compared.

Model with k_{ttbb} affecting $ttbb$ only

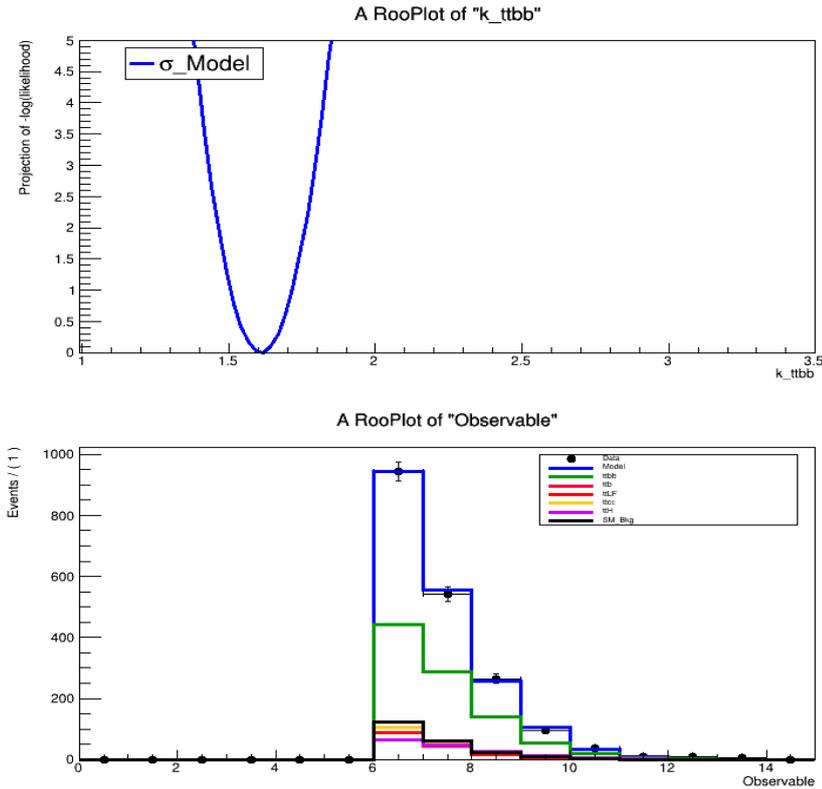


FIGURE 7.11: No. of jets in $\ell + jets, \geq 6j+ \geq 4b$
The value of $k_{ttbb} = 1.64$ and thus,

$$\sigma_{ttbb}^{all} \approx 6.5\text{pb} \quad (7.71)$$

Model with k_{ttbb} affecting $ttbb$ & ttb

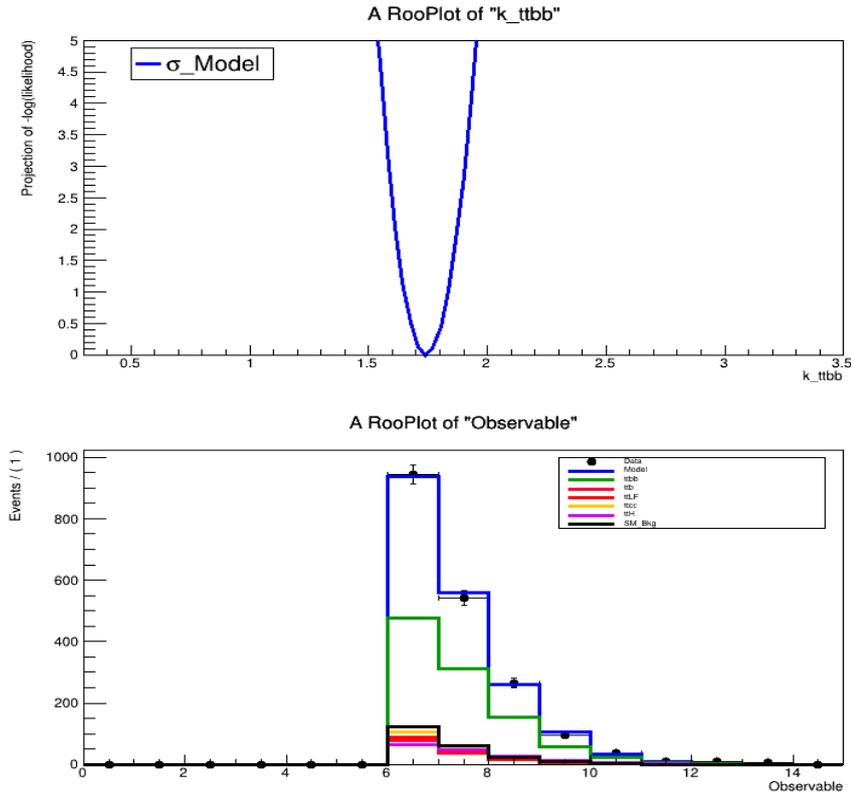


FIGURE 7.12: No. of jets in $\ell + jets, \geq 6j + \geq 4b$
The value of $k_{ttbb} = 1.74$ and thus,

$$\sigma_{ttbb}^{all} \approx 6.9\text{pb} \quad (7.72)$$

Model with k_{ttbb} & k_{ttH}

It is the best model as both the processes, $ttbb$ and $ttH(H \rightarrow bb)$ share the same signature.

From the figures 7.13 and 7.14 below,

$$\sigma_{ttbb}^{\mu j} = 0.98 \cdot \sigma_{ttbb}^{th} \quad (7.73)$$

$$\approx 3.90\text{pb} \quad (7.74)$$

$$\sigma_{ttH}^{\mu j} = 5.73 \cdot \sigma_{ttH}^{th} \quad (7.75)$$

$$\approx 1.68\text{pb} \quad (7.76)$$

$$\sigma_{ttbb}^{ej} = 0.89 \cdot \sigma_{ttbb}^{th} \quad (7.77)$$

$$\approx 3.57\text{pb} \quad (7.78)$$

$$\sigma_{ttH}^{ej} = 3.19 \cdot \sigma_{ttH}^{th} \quad (7.79)$$

$$\approx 0.93\text{pb} \quad (7.80)$$

Therefore,

$$\sigma_{ttbb}^{all} \approx 3.68\text{pb} \quad (7.81)$$

$$\sigma_{ttH}^{all} \approx 1.33\text{pb} \quad (7.82)$$

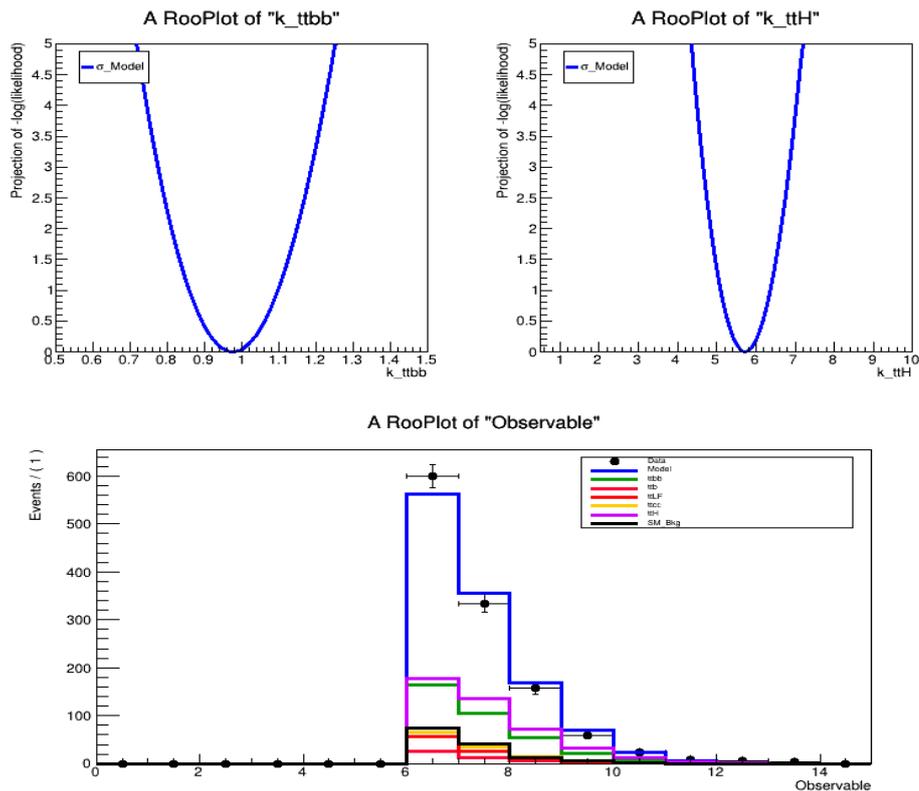


FIGURE 7.13: No. of jets in $\mu + jets, \geq 6j \geq 4b$

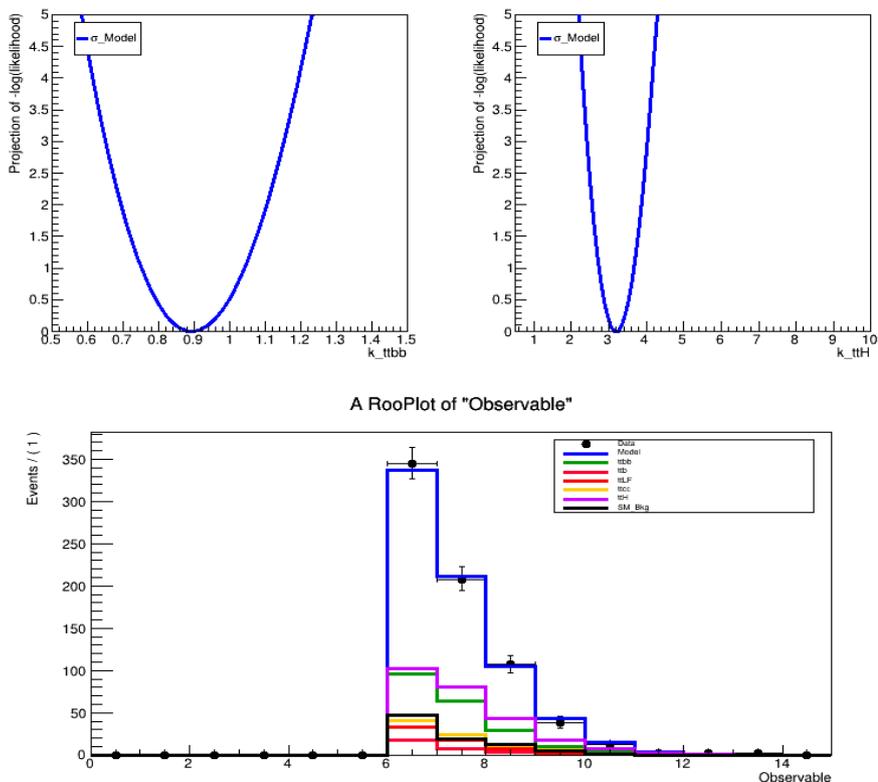


FIGURE 7.14: No. of jets in $e + jets, \geq 6j \geq 4b$

Thus, it is clearly seen that for the region of at least $6j$ and $4b$, the deviation of cross section measured to the theoretical value is the least.

For the best model which extracts both the σ_{ttbb} and σ_{ttH} simultaneously, let's also consider the invariant mass of the whole system as the discriminant variable and calculate the cross section. Thus, from the figure 7.15, the cross-section for $ttbb$ is approximately 4.69pb and the cross-section for ttH is 1.05pb.

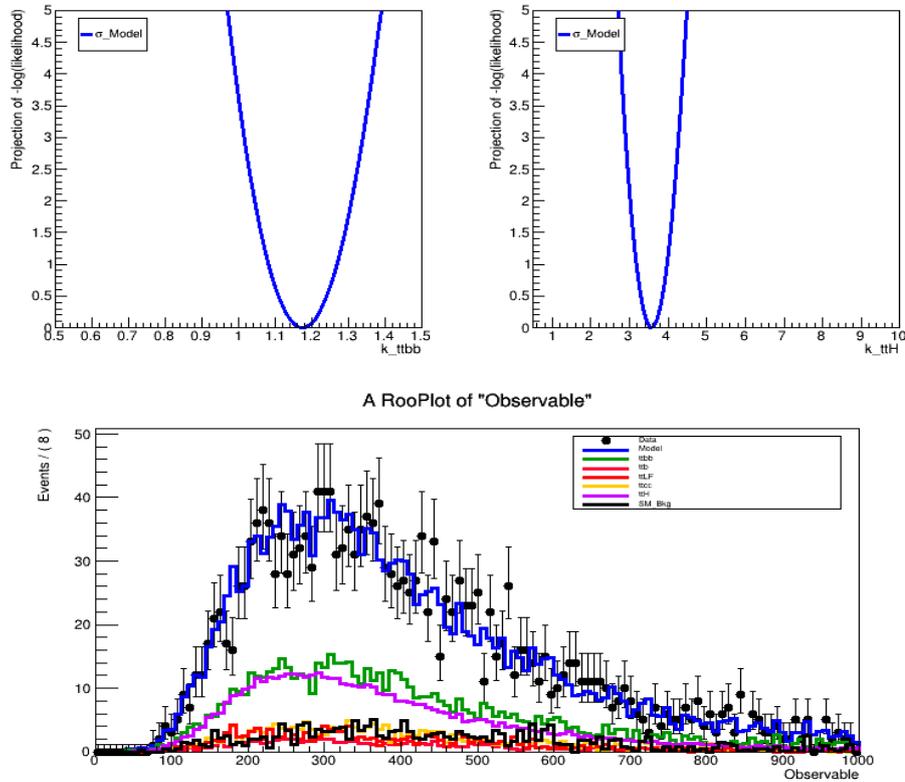


FIGURE 7.15: Invariant Mass in $\ell + jets, \geq 6j+ \geq 4b$

In this case with no systematic uncertainties considered, number of jets seems to be a good discriminant variable.

Chapter 8

Conclusions

The cross sections of the $t\bar{t}$, $H \rightarrow WW$, $t\bar{t}b\bar{b}$ and $t\bar{t}H$ with Higgs decaying to b-jets in the semi-leptonic decay mode are measured using two different approaches: a cut and count analysis (C & C) and a template fit. The signal region selected for these events contain a single prompt lepton (e or μ). Additional requirements in the number of jets and b-jets are different for each of the process. The signal region selected for $t\bar{t}$ includes at least 4 jets where at least 2 of them are identified as b-jets and $H \rightarrow WW$ includes 2 jets without any b-jets. For $t\bar{t}b\bar{b}$ and $t\bar{t}H$ processes that share the same signature, two signal regions are analysed with at least 6 jets where 2 of them are b-jets and a tighter region requiring at least 4 b-jets.

The results are summarised in the table 8.1. For $t\bar{t}b\bar{b}$ and $t\bar{t}H$, the results are provided for 6 jets and 4 b-jets region. The results presented in the table do not show any systematic uncertainties because of the limitation in time we had. We expect that the systematic uncertainties will cover the discrepancies between the theoretical and observed values when considered.

TABLE 8.1: Results obtained for cross sections of different processes analysed using two different approaches: C & C and fit

Processes	Theoretical [pb]	Observed [pb]	
		C & C	Fit
$t\bar{t}$	831.76	857	855
$H \rightarrow WW$	32.4	~ 0	85
$t\bar{t}b\bar{b}$	4	6.5	3.7
$t\bar{t}H(H \rightarrow b\bar{b})$	0.29	0.9	1.3

From the analysis, we infer that the cut and count method is quite good for the extraction of $t\bar{t}$ cross section but not for the $H \rightarrow WW$ process. The results obtained using both the methods work well for $t\bar{t}$ cross section extraction as the signal region selected is concentrated with the signal events with less background contamination. The $H \rightarrow WW$ process is expected to have a few events and those events are mitigated with higher background contamination from QCD and W-Jets. Thus, the cross section obtained is adversely affected by the mis-modelling of these contributions and thus it is compatible with the value equal to 0pb. The cross section calculated from the fit approach in the table 8.1 used the invariant mass of the whole system as the discriminant variable. When the p_T of the lepton is chosen as the variable, it is seen that the cross section jumps to 223pb which implies the value is very sensitive to the discriminant variable chosen due to the same reasons mentioned as before.

The extraction of the cross section of the $t\bar{t}$ associated with the b-jets and with the Higgs boson decaying to two b-jets using the C & C method is possible despite of

its low statistics as the $t\bar{t}$ region can be well defined and its MC simulation is well modelled. The values are compatible with the theoretical predictions. If we choose a looser selection (at least 6j with at least 4b), the results are still compatible but a little higher than the predicted value.

It is also observed from the table that the cross section measured is a little higher than the expected results in most of the cases. It is because the MC expectation is lower than the data events and thus, we have lower cross section by the definition.

It can be seen that the fit approach is better than the C & C as the fit includes more information of the shapes. Even in cases where the cross section shows much deviation from the expected values, the values are reasonable taking into consideration that the systematic uncertainties are not calculated. These uncertainties can be from the uncertainties associated with b-tagging, jets selection, luminosity, lepton identification, pileup etc. We consider that these will cover the discrepancy in the values of the cross section expected and observed.

Bibliography

- [1] Oliver Sim Brüning et al. *LHC Design Report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. DOI: 10.5170/CERN-2004-003-V-1. URL: <https://cds.cern.ch/record/782076>.
- [2] Giovanni Anelli et al. "The TOTEM experiment at the CERN large hadron collider". In: *Journal of Instrumentation* 3 (Aug. 2008). DOI: 10.1088/1748-0221/3/08/S08007.
- [3] The Collaboration et al. "The LHCf detector at the CERN Large Hadron Collider". In: *Journal of Instrumentation* 3 (Aug. 2008), S08006. DOI: 10.1088/1748-0221/3/08/S08006.
- [4] Michael Staelens. *Recent Results and Future Plans of the MoEDAL Experiment*. 2019. arXiv: 1910.05772 [hep-ex].
- [5] FASER Collaboration et al. *FASER: ForwArd Search ExpeRiment at the LHC*. 2019. arXiv: 1901.04468 [hep-ex].
- [6] K. Aamodt et al. "The ALICE experiment at the CERN LHC". In: *JINST* 3 (2008), S08002. DOI: 10.1088/1748-0221/3/08/S08002.
- [7] A. Augusto Alves Jr. et al. "The LHCb Detector at the LHC". In: *JINST* 3 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005.
- [8] The ATLAS Collaboration et al. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *Journal of Instrumentation* 3.08 (2008), S08003–S08003. DOI: 10.1088/1748-0221/3/08/s08003. URL: <https://doi.org/10.1088/1748-0221/3/08/s08003>.
- [9] G Bayatian. "CMS Physics Technical Design Report, Volume II: Physics Performance". In: *J.Phys.* G34 (June 2007), pp. 995–1579. DOI: 10.1088/0954-3899/34/6/S01.
- [10] Jacques Herve Fichet. "CMS beam pipe installation". In: (2021). General Photo. URL: <https://cds.cern.ch/record/2764445>.
- [11] V. Karimäki. "The CMS tracker system project: Technical Design Report". In: (1997). Ed. by M. Mannelli et al.
- [12] *The CMS electromagnetic calorimeter project: Technical Design Report*. Technical design report. CMS. Geneva: CERN, 1997. URL: <https://cds.cern.ch/record/349375>.
- [13] *The CMS hadron calorimeter project: Technical Design Report*. Technical design report. CMS. The following files are from <http://uscms.fnal.gov/pub/hcal-tdr> and may not be the version as printed, please check the printed version to be sure. Geneva: CERN, 1997. URL: <https://cds.cern.ch/record/357153>.
- [14] J. G. Layter. *The CMS muon project: Technical Design Report*. Technical design report. CMS. Geneva: CERN, 1997. URL: <https://cds.cern.ch/record/343814>.

- [15] S.F. Novaes. "Standard Model: An Introduction". In: (2000). URL: arXiv:hep-ph/0001283.
- [16] Matthew D. Schwartz. *Quantum Field Theory and the Standard Model*. Cambridge University, 2014.
- [17] F. Halzen and Alan D. Martin. *QUARKS AND LEPTONS: AN INTRODUCTORY COURSE IN MODERN PARTICLE PHYSICS*. 1984. ISBN: 978-0-471-88741-6.
- [18] F. Abe et al. "Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab". In: *Phys. Rev. Lett.* 74 (14 1995), pp. 2626–2631. DOI: 10.1103/PhysRevLett.74.2626. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.74.2626>.
- [19] S. Abachi et al. "Observation of the Top Quark". In: *Phys. Rev. Lett.* 74 (14 1995), pp. 2632–2637. DOI: 10.1103/PhysRevLett.74.2632. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.74.2632>.
- [20] The ATLAS et al. *First combination of Tevatron and LHC measurements of the top-quark mass*. 2014. arXiv: 1403.4427 [hep-ex].
- [21] Z. Ligeti (LBNL) A. Ceccucci (CERN) and Y. Sakai (KEK). "CKM Quark-Mixing Matrix". In: *Phys. Rev. Lett.* (revised on 2018).
- [22] K. Sasaki. "Renormalization Group Equations for the Kobayashi-Maskawa Matrix". In: *Z. Phys. C* 32 (1986), pp. 149–152. DOI: 10.1007/BF01441364.
- [23] V. Abazov et al. "Evidence for production of single top quarks and first direct measurement of $|V_{tb}|$." In: *Physical review letters* 98 18 (2007), p. 181802.
- [24] J. Beringer et al. "Review of Particle Physics". In: *Phys. Rev. D* 86 (1 2012), p. 010001. DOI: 10.1103/PhysRevD.86.010001. URL: <https://link.aps.org/doi/10.1103/PhysRevD.86.010001>.
- [25] Cristina Ferro. "Measurement of the $t\bar{t}$ production cross section in the tau+jets channel in pp collisions at $\sqrt{s}=7\text{TeV}$ ". PhD thesis. Strasbourg, IPHC, 2012.
- [26] A. M. Sirunyan et al. "Measurement of jet substructure observables in $t\bar{t}$ events from proton-proton collisions at $\sqrt{s} = 13\text{TeV}$ ". In: *Phys. Rev. D* 98.9 (2018), p. 092014. DOI: 10.1103/PhysRevD.98.092014. arXiv: 1808.07340 [hep-ex].
- [27] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. "The catchment area of jets". In: *Journal of High Energy Physics* 2008.04 (2008), 005–005. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2008/04/005. URL: <http://dx.doi.org/10.1088/1126-6708/2008/04/005>.
- [28] S. Dawson. "Introduction to the physics of Higgs bosons". Lectures given at TASI 94. URL: arXiv:hep-ph/9411325.
- [29] S. Dawson. "Introduction to Electroweak Symmetry Breaking". In: *AIP Conf. Proc.* 1116.1 (2009). Ed. by Alejandro Ayala et al., pp. 11–34. DOI: 10.1063/1.3131544. arXiv: 0812.2190 [hep-ph].
- [30] Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (16 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.13.508>.
- [31] Jun Gao. "Differentiating the production mechanisms of the Higgs-like resonance using inclusive observables at hadron colliders". In: *Journal of High Energy Physics* 2014.2 (2014). ISSN: 1029-8479. DOI: 10.1007/jhep02(2014)094. URL: [http://dx.doi.org/10.1007/JHEP02\(2014\)094](http://dx.doi.org/10.1007/JHEP02(2014)094).

- [32] M. Spira et al. "Higgs boson production at the LHC". In: *Nuclear Physics B* 453.1-2 (1995), 17–82. ISSN: 0550-3213. DOI: 10.1016/0550-3213(95)00379-7. URL: [http://dx.doi.org/10.1016/0550-3213\(95\)00379-7](http://dx.doi.org/10.1016/0550-3213(95)00379-7).
- [33] J. Baglio and A. Djouadi. "Higgs production at the LHC". In: *Journal of High Energy Physics* 2011 (2011), pp. 1–53.
- [34] Milos Dordevic. "ttH production at 13 TeV (CMS)". In: *EPJ Web Conf.* 164 (2017). Ed. by L. Bravina, Y. Foka, and S. Kabana, p. 07055. DOI: 10.1051/epjconf/201716407055.
- [35] Jehad Mousa et al. "Observation of ttH Production". In: *Physical Review Letters* 120 (June 2018).
- [36] Abhisek Datta. "Higgs Boson Production in Association with a Top Quark-Antiquark Pair and Decaying into a Bottom Quark-Antiquark Pair at the CMS Experiment". PhD thesis. Cornell U., 2020. DOI: 10.7298/7cwk-2494.
- [37] "Search for the Standard Model Higgs boson in the H to WW to lνjj decay channel". In: (2012).
- [38] "Isolated Muon". In: (2017). URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonIsolationForRun2>.
- [39] "Cut Based Muon ID". In: (2017). URL: https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideMuonIdRun2#Tight_Muon.
- [40] "Cut Based Electron ID". In: (2017). URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/CutBasedElectronIdentificationRun2>.
- [41] "Jets Selection criteria CMS". In: (2017). URL: <https://twiki.cern.ch/twiki/bin/view/CMS/JetID13TeVRun2017>.
- [42] A.M. Sirunyan et al. "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV". In: *Journal of Instrumentation* 13.05 (2018), P05011–P05011. ISSN: 1748-0221. DOI: 10.1088/1748-0221/13/05/p05011. URL: <http://dx.doi.org/10.1088/1748-0221/13/05/P05011>.
- [43] "Performance of the DeepJet b tagging algorithm using 41.9/fb of data from proton-proton collisions at 13TeV with Phase 1 CMS detector". In: (2018). URL: <http://cds.cern.ch/record/2646773>.
- [44] Giulia Pancheri and Yogendra N. Srivastava. "Introduction to the physics of the total cross section at LHC". In: *The European Physical Journal C* 77.3 (2017). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-016-4585-8. URL: <http://dx.doi.org/10.1140/epjc/s10052-016-4585-8>.
- [45] "NNLO+NNLL top-quark-pair cross sections". In: (CERN). URL: <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TtbarNNLO>.
- [46] "C. Patrignani et al. (Particle Data Group)". In: ((2016) and 2017 update). *Chin. Phys. C*, 40, 100001. URL: https://pdg.lbl.gov/2017/html/authors_2017.html.
- [47] Wouter Verkerke and David P. Kirkby. "The RooFit toolkit for data modeling". In: *eConf* C0303241 (2003). Ed. by L. Lyons and Muge Karagoz, MOLT007. arXiv: physics/0306116.
- [48] R. Brun and F. Rademakers. "ROOT: An object oriented data analysis framework". In: *Nucl. Instrum. Meth. A* 389 (1997). Ed. by M. Werlen and D. Perret-Gallix, pp. 81–86. DOI: 10.1016/S0168-9002(97)00048-X.