Search for the Associated Production of a Z boson with a Single Top Quark using CMS data at 8 TeV and 13 TeV and Performance Studies of the CMS Silicon Tracker

QUAID-I-AZAM UNIVERSITY
ISLAMABAD

By
Muhammad Waqas

Department of Physics
Quaid-i-Azam University
Islamabad, Pakistan
2019

This work is submitted as a thesis in the partial fulfillment of the
requirement for the award of the degree
of

Doctor of Philosophy

In

Physics

Department of Physics
Quaid-i-Azam University
Islamabad, Pakistan

CERTIFICATE

The undersigned declare that the research work presented in this thesis titled, **"Search for the Associated Production of a Z boson with a Single Top Quark using CMS data at 8 TeV and 13 TeV and Performance Studies of the CMS Silicon Tracker"** by **Mr. Muhammad Waqas** is carried out under my supervision in partial fulfillment of the requirements for the degree of **Doctor of Philosophy** in **Physics**.

Supervisor:

Dr. Ashfaq Ahmad
Director, EHEP
National Center For Physics
Quaid-e-Azam University Campus
Islamabad, Pakistan

Submitted Through:

Prof. Dr. Arif Mumtaz
Chairman
Department of Physics
Quaid-e-Azam University
Islamabad, Pakistan

Author's Declaration

The undersigned declare that the research work presented in this thesis titled, **"Search for the Associated Production of a Z boson with a Single Top Quark using CMS data at 8 TeV and 13 TeV and Performance Studies of the CMS Silicon Tracker"** is his own work, and that all the sources he has used or quoted have been indicated or acknowledged by means of complete references.

Muhammad Waqas

Plagiarism Undertaking

I solemnly declare that the research work presented in this thesis titled, "Search for the Associated Production of a Z boson with a Single Top Quark using CMS data at 8 TeV and 13 TeV and Performance Studies of the CMS Silicon Tracker" is solely my work with no significant contribution from any other person. Small contribution/help wherever taken has been dully acknowledged and that complete thesis has been written by me.

I understand the zero tolerance policy of the HEC and Quaid-e-Azam university towards plagiarism. Therefor, I as an author of above titled thesis declare that no portion of my thesis has been plagiarised and any material used as reference is properly referred/cited.

I undertake that if i am found guilty of any formal plagiarism in the above titled thesis even after my Ph.D. degree, the university reserves the rights to withdraw/revoke my Ph.D. degree and that HEC and the University has right to publish my name on the HEC/University website on which names of students are placed who submitted plagirised thesis,

Muhammad Waqas

Abstract

The cross section measurements of a single top quark in association with a Z boson using proton-proton collision data collected by CMS experiment are reported in this thesis. The measurements are performed by analyzing integrated luminosities of 19.7 fb^{-1} and 35.9 fb^{-1} which were collected at 8 TeV and 13 TeV. The measurement are performed within the standard model framework where top quark is produced via the t-channel process. The measurements in three lepton final state, where the W boson from the top quark and the Z boson decay into either electrons or muons, resulting in four possible lepton combinations namely eee, $ee\mu$, $\mu\mu e$, $\mu\mu\mu$, are discussed. The final state electrons or muons can also come from from leptonic τ decays, as they are not specifically excluded. The main sources of background to the tZq process are $t\bar{t}$ production, diboson production (WZ, ZZ), ttV (V = W or Z) and Drell-Yan (DY) production. A simple cut and count technique is used for the cross section measurement. For the $\sqrt{s} = 8$ TeV measurement, a one bin likelihood fit is performed to event yields in each four lepton channels and to the combined channel. The cross section for the combined channel is measured to be $\sigma(t\ell^+\ell^-q) = 18^{+11}_{-9}(stat)^{+4}_{-4}(syst)$ fb, where ℓ stands for electrons, muons and τ leptons, with an observed (expected) significance of 1.81 (0.81) standard deviations. The measured value is compatible with the NLO standard model prediction of 8.2 fb with a theoretical uncertainty of less than 10%. For the second cross section measurement at 13TeV, the signal is extracted by performing a simultaneous binned likelihood fit to yields in the signal and the background enriched control regions. The cross section for the combined channel is measured to be $\sigma(t\ell^+\ell^-q) = 156^{+47}_{-42}(stat)^{+40}_{-34}(syst)$ fb, which is compatible with the NLO standard model prediction of 94.2 ± 3.1 fb. The measurement is found to be in agreement with the standard model with an observed (expected) significance of 2.81(1.95) standard deviations.

Dedicated to my beloved wife
Tehmina and wonderful kids
Aleezeh and Issa.

Acknowledgments

First and foremost I would like to thank my supervisor Dr. Ashfaq Ahmad for his scientific guidance, availability, constructive feed backs, and his open and friendly attitude. I have been extremely lucky to have a supervisor who cared so much about my work, and who responded to my questions and queries so promptly. His positive outlook and confidence in my research inspired me and gave me confidence. I could not have asked for a better Ph.D. supervisor. Thank you for your seemingly unlimited support and kindness.

I want to express my deepest gratitude to the distinguished physicist Dr. Martijn Mulders for all the patience, support and advice he has given me over the time period that I spent at CERN. It was a real privilege and an honor for me to benefit from his exceptional scientific knowledge and also of his extraordinary human qualities. Thank you for the excellent guidance and wisdom you have given me. I would like to express my very great appreciation to Prof. Dr. Hafeez Hoorani for all the support and encouragement he gave me throughout my Ph.D. I am particularly grateful to him for arranging funding for my CERN visits. Thank you sir for trusting me.

I would like to express my gratitude to people in CMS Top group. Most notably, I would like to thank Top physics group convenors Pedro De Silva, Jeremy Andrea, Andrea Giammanco, Andreas Meyer, Rebeca Gonzalez, Orso, Abideh Jafari, Thorsten Chwalek, It was a pleasure contributing to the single top quark efforts of CMS in the friendly and welcoming atmosphere you created. I am also thankful to all the people who worked for the the construction and smooth running the LHC machine. I am also very thankful to the faculty of Quaid-e-Azam University especially Dr. Jamil Aslam and Dr. Mansoor Ahmad for all the wonderful courses they taught me.

I am very thankful to Muhammad Shoaib for all the work we did in collaboration. It was really fun working as a team with you. I would like to express my appreciation to people at NCP: Muhammad Ahmad, Mehar Ali Shah, Qamar ul Hassan, Muhammad Gul, Irfan Asghar, Wajid Ali Khan, Waqar Ahmed, Muhammad Saleh for all the support and discussions. I wish you all a happy life and the best of luck in future.

I am very thankful to my dearest friends Abid Farooq and Muhammad Saqib for encouraging me in down time and partying together. I am thankful to my dearest friends Haider Raza Rizvi, Asif Warsi, Muhammad Nadeem, Tufail Khan, Asad Mehmood, Imran Shamsi, Muhammad Adeel, Sramad Shaheen, Muhammad Ahmad and Priya Sajid Hussain. I have shared with you all kinds of events, both happy and stressful times. Thanks for being available when I needed your support

I am also grateful to all the undergraduate juniors at NCP: Attiqu, Khurram, Binish, Priya, Adil, Nabila, Shahzad, Haseeb, Ammara, Arif, Huda, Khushbakht, Mazhar, Kiran, Amad, Adnan, Iqra, Mubasher, Natasha, Sahar, kubra, saba and Ansar. I spent quality time with you and thanks for arranging all those trips to northen areas and exploring tracks. Wish you all a very bright future.

During my PhD studies I was a recipient of the Higher Education Commission (HEC) scholarship, I would like to thanks HEC for all the financial support.

Last but not least, unlimited thanks to my family: my parents, my brothers and my sister for supporting me throughout my life in general. Special thanks to my Father and elder brother Umer who have done a lot for me. Finally I can not thank enough to my wife Tehmina for all the love, support, encouragement from the day we met. Thank you for bearing with me throughout this work. This work would have not been possible without your endless support. But most of all, thank you for being my best friend. I owe you everything.

Contents

List of Figures

List of Tables

List of Abbreviations

LHC	**L**arge **H**adron **C**ollider
CMS	**C**ompact **M**uon **S**olenoid
PV	**P**rimary **V**ertex
SV	**S**econdary **V**ertex
ML	**M**aximum **L**ikelihood
MLE	**M**aximum **L**ikelihood **E**stimation
pp	**p**roton-**p**roton
IP	**I**mpact **P**arameters
SM	**S**tandard **M**odel
QCD	**Q**uantum **C**hromo**d**ynamics
EW	**E**lectro**w**eak
RH	**R**ight-**h**anded
LH	**L**eft-**h**anded
LO	**L**eading **O**rder
NLO	**N**ext-to-**l**eading-**o**rder
PU	**P**ile**u**p
MET	**M**issing **T**ransverse **E**nergy
ECAL	**E**lectromagnetic **Cal**orimeter
HCAL	**H**adron **Cal**orimeter
ME	**M**atrix **E**lement
PS	**P**arton **S**hower
UE	**U**nderlying **E**vents
MC	**M**onte **C**arlo
PF	**P**article **F**low
CTF	**Combination Track Finder**
SF	**S**cale **F**actor
CSV	**C**ombined **S**econdary **V**ertex
BP	**B**ack**p**lane
PDF	**P**arton **D**istribution **F**unction
JES	**J**et **E**nergy **S**cale
JER	**J**et **E**nergy **R**esolution
NLP	**N**on-prompt **L**epton
DAQ	**D**ata **A**c**q**uisition
WP	**W**orking **P**oint

Chapter 1

Theoretical Framework

Particle physics is the study of the elementary particles and their interactions. The Standard Model (SM) is a mathematical framework that best describes the behaviour of these elementary particles. SM is the most precisely tested scientific model that has ever been constructed. It has passed thousands of tests and verified most of the predictions since its development in 1960's. The final bit of the SM was Higgs boson which was discovered in July 2012 at the Large Hadron Collider. In this chapter, the building blocks of the SM and its shortcomings are discussed.

1.1 The Standard Model

The SM has been tested thoroughly over the years and is outstanding in describing the experimental observations. It is currently our finest theory in understanding the fundamental particle physics. The SM is a non-abelian gauge theory, also called Yang-Mills theory [1], with gauge group

$$\mathcal{G}_{SM} = SU(3)_C \otimes SU(2)_L \otimes U(1)_Y, \tag{1.1}$$

describing all the known elementary particles, their dynamics and interactions, except gravity which is covered by General Relativity (GR) [2]. The particles in the SM are fermions and bosons. Fermions, half integer spin particles, are the constituents of matter while bosons, integer spin particles, are the force mediators. Fermions are subject to the Pauli principle, which forbids that two fermions occupy the same state, namely the same set of quantum numbers.

The bosons in the SM are responsible for three of the four interactions in nature as shown in Table 1. The photons (γ) are the mediators of the electromagnetic interactions, the gluons are of strong interactions and the Z and W bosons are the mediators of the weak interactions. The Z and the W bosons are massive while photons and gluons are massless. Weak force is weak because of these massive mediators. Despite their weakness, the weak interactions have very distinctive features. The symmetries like charge conjugation C, parity P, time-reversal T, the combination of charge conjugation and parity CP are violated in weak interactions. However, all these are the symmetries of strong and electromagnetic interactions.

Table 1.1: The electric charge, mass and interaction of the SM gauge bosons [3].

Name	Interaction	Electric Charge (e)	Mass
Gluon (g)	Strong	0	0
Photon (γ)	Electromagnetic	0	0
Z boson	Weak	0	91.19 GeV
W boson	Weak	±1	80.39 GeV

The fermions in the SM are organized in three families or generations having same quantum numbers but different masses. The 1^{st} family or generation has the lightest particles and the 3^{rd} one has the heaviest. The lightest generations makes up most of the ordinary matter as the heavier families (being unstable) decays into the lighter ones. Each family has four fermions that are identified by their charges under electromagnetic and/or strong interactions. Two of the four fermions in each family are quarks which have color charge under the strong interactions and the other two are leptons which are not charged under strong interaction. The lepton neutrinos feel only weak interactions. The mass of neutrinos are at least six orders of magnitude smaller than any other SM fermion. Additionally, each quark and lepton has an anti-particle, thus doubling the fermion count, with same masses and opposite quantum numbers.

Table 1.2: The fermions in the SM with their electric charge and mass according to the Particle Data Group [3].

Generation	Name	Symbol	Electric Charge (e)	Mass
		Quarks		
1^{st}	Up	u	$+\frac{2}{3}$	2.3 MeV
	Down	d	$-\frac{1}{3}$	4.8 MeV
2^{nd}	Charm	c	$+\frac{2}{3}$	1.275 GeV
	Strange	s	$-\frac{1}{3}$	95 MeV
3^{rd}	Top	t	$+\frac{2}{3}$	173.5 GeV
	Bottom	b	$-\frac{1}{3}$	4.65 GeV
		Leptons		
1^{st}	Electron	e	+1	0.51 MeV
	Electron Neutrino	ν_e	0	$<$ 2 eV
2^{nd}	Muon	μ	+1	105.66 MeV
	Muon Neutrino	ν_μ	0	$<$ 2 eV
3^{rd}	Tauon	τ	+1	1.77 GeV
	Tauon Neutrino	ν_τ	0	$<$ 2 eV

The SM Lagrangian can be studied by splitting the Lagrangian given in the Equation 1.1. $SU(2)_L \otimes U(1)_Y$ symmetry in equation 1.1 corresponds to electroweak (EW) interaction described by the Glashow-Salam-Weinberg (GSW) model [4, 5, 6] combining the electromagnetic and weak interactions. $SU(3)_C$ reflects the symmetry of strong interaction, namely an exact symmetry of three colors of six quarks [7, 8, 9].

This theory of strong force is called Quantum Chromodynamics (QCD). Thus, the SM Lagrangian can be written in these two symmetries as

$$\mathcal{L}_{SM} = \mathcal{L}_{EW} + \mathcal{L}_{QCD} \tag{1.2}$$

1.2 Electroweak Interaction

The charge associated with the weak interaction is the weak isospin T of which the third component T_3 is conserved. The relevant symmetry group for weak interactions is $SU(2)_L$, where the subscript L means that only the left-handed(LH) chiral components of the fields are involved in the interaction. The fermion state can be written as a combination of LH chirality eigenstates ψ_L and right-handed (RH) chirality eigenstates ψ_R,

$$\psi = P_L\psi + P_R\psi = \psi_L + \psi_R \tag{1.3}$$

with the LH and RH projection operators P_L and P_R,

$$P_L = \frac{1-\gamma^5}{2} \quad \text{and} \quad P_L = \frac{1+\gamma^5}{2}. \tag{1.4}$$

The LH fermion fields with $T_3 = \pm\frac{1}{2}$ appear as doublets because the LH chirality eigenstates ψ_L transforms under the weak isospin $SU(2)_L$. The RH chirality eigenstates ψ_R do not transform under the weak isospin $SU(2)_L$ and are therefore singlets ($T_3 = 0$). The electromagnetic interaction, mediated by photon (γ), is described by the symmetry group $U(1)_Q$. The photon does not differentiate between LH and RH fermions but interacts with all the particles carrying electric charge Q. The relationship between electric charge, Q, and quantum number of the other groups is given by,

$$Q = T_3 + \frac{Y}{2}, \tag{1.5}$$

where Y is the weak hypercharge which is conserved in $U(1)_Y$ transformations. The quantum numbers of fermions in the EW sector are summarized in Table 1.3.

The Lagrangian for electroweak theory is given by

$$\mathcal{L}_{SU(2)_L \otimes U(1)_Y} = \mathcal{L}_{free} + \mathcal{L}_{gauge} + \mathcal{L}_{\phi} + \mathcal{L}_{Yuk}. \tag{1.6}$$

The $\mathcal{L}_{free}$, summing over the LH weak iso-doublets ψ_L and RH weak iso-singlets ψ_R,

$$\mathcal{L}_{free} = \sum_{\psi_L} \bar{\psi_L}(x)(i\partial\!\!\!/)\psi_L(x) + \sum_{\psi_R} \bar{\psi_R}(x)(i\partial\!\!\!/)\psi_R(x), \tag{1.7}$$

is invariant under global $SU(2)_L \otimes U(1)_Y$ transformations in flavor space,

Table 1.3: Electric charge Q, weak isospin T_3 and weak hyper charge Y of fundamental fermions in EW sector.

Particle Type	**Q**	**T$_3$**	**Y**
LH quarks $\equiv \mathbf{Q}_L = \left(\begin{pmatrix} u_L \\ d_L \end{pmatrix}\begin{pmatrix} c_L \\ s_L \end{pmatrix}\begin{pmatrix} t_L \\ b_L \end{pmatrix}\right)$	$\begin{pmatrix} \frac{2}{3} \\ -\frac{1}{3} \end{pmatrix}$	$\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$	$\frac{1}{3}$
RH up-type quarks $\equiv \mathbf{u}_R = \left(u_R, c_R, t_R \right)$	$\left(\frac{2}{3}, \frac{2}{3}, \frac{2}{3} \right)$	$\left(0, 0, 0 \right)$	$\frac{4}{3}$
RH down-type quarks $\equiv \mathbf{d}_R = \left(d_R, s_R, b_R \right)$	$\left(-\frac{1}{3}, -\frac{1}{3}, -\frac{1}{3} \right)$	$\left(0, 0, 0 \right)$	$-\frac{2}{3}$
LH leptons $\equiv \mathbf{L}_L = \left(\begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix}\begin{pmatrix} \nu_{\mu L} \\ \mu_L \end{pmatrix}\begin{pmatrix} \nu_{\tau L} \\ \tau_L \end{pmatrix}\right)$	$\begin{pmatrix} 0 \\ -1 \end{pmatrix}$	$\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$	-1
RH leptons $\equiv \boldsymbol{\ell}_R = \left(e_R, \mu_R, \tau_R \right)$	$\left(-1, -1, -1 \right)$	$\left(0, 0, 0 \right)$	-2

$$\psi_L(x) \to \psi'_L(x) \equiv \exp(iY\beta)U_L\psi_L(x) \tag{1.8}$$

$$\psi_R(x) \to \psi'_R(x) \equiv \exp(iY\beta)\psi_R(x) \tag{1.9}$$

with the weak hypercharge Y associated with the field transforming and the non-abelian $SU(2)_L$ transformation

$$U_L \equiv \exp(i\frac{\tau_i}{2}\alpha^i) \tag{1.10}$$

which acts only on the LH doublet fields ψ_L. The lagrangian in Equation 1.7 becomes invariant under local $SU(2)_L \otimes U(1)_Y$ gauge transformation by replacing the derivatives with covariant derivatives. Since the RH particles do not couple to weak isospin, their covariant derivatives takes the simple form

$$D_\mu\psi_R(x) \equiv \left(\partial_\mu + ig'\frac{Y}{2}B_\mu(x)\right)\psi_R(x) \tag{1.11}$$

which serves to define the $U(1)_Y$ coupling g'. The corresponding covariant derivative for the $SU(2)_L$ doublet is given by

$$D_\mu\psi_L(x) \equiv \left(\mathbf{I}\left(\partial_\mu + ig'\frac{Y}{2}B_\mu(x)\right) + ig\frac{1}{2}\boldsymbol{\tau}\cdot\mathbf{W}_\mu(x)\right)\psi_L(x), \tag{1.12}$$

where g represents the $SU(2)_L$ gauge coupling constant, $\mathbf{I}$ denotes 2×2 unit matrix and the vector $\boldsymbol{\tau} = (\tau_1, \tau_2, \tau_3)$ contains the Pauli matrices. The coupling of g' of fermions to B_μ is free and depends on the convention for the weak hypercharge Y.

The gauge part of the Lagrangian is given by:

$$\mathcal{L}_{gauge} = \frac{1}{4}F^i_{\mu\nu}F^{\mu\nu i} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}, \tag{1.13}$$

with field strength tensors are

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \quad \text{and} \tag{1.14}$$

$$F^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu - g\epsilon_{ijk}W^j_\mu W^k_\nu \quad \text{i=1,2,3.} \tag{1.15}$$

$\mathcal{L}_{gauge}$ is not only quadratic in the derivatives of the gauge boson fields but contains as well three- and four-point self-interaction terms proportional to coupling g. B has no self-interaction as the group is abelian. The B and W_3 ultimately becomes the photon and Z boson after mixing.

The scalar part of the electroweak Lagrangian is

$$\mathcal{L}_\phi = (D_\mu\Phi)^\dagger(D^\mu\Phi) - V(\Phi), \tag{1.16}$$

where $\Phi = \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix}$ is the complex Higgs scalar field which is doublet under $SU(2)_L$. $V(\Phi)$ is the Higgs potential given by

$$V(\Phi) = \mu^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2. \tag{1.17}$$

The Higgs potential is restricted to this form because of the invariance and renormalizability of $SU(2)_L \otimes U(1)_Y$. For $\mu^2 < 0$ and $\lambda > 0$, there is spontaneous symmetry breaking that is responsible for giving masses to SM particles. The λ is a quartic self-interaction term between the scalar fields and it has to be greater than zero for Vacuum stability.

$\mathcal{L}_{Yuk}$ term in Equation 1.6 describes the gauge invariant Yukawa interaction that couples the Higgs and fermion field. These couplings are responsible for the fermion mass terms in the SM.

1.2.1 Higgs Mechanism

The gauge principle allows to successfully describe the interactions between particles; however, all fields produced by imposing the gauge invariance are strictly massless. However, the experiment shows that both the chiral fermions and electroweak gauge bosons (W$^\pm$ and Z) are massive. The gauge invariance principle has to be broken spontaneously to give masses to bosons and chiral fermions. With this broken symmetry, the physical states, especially the lowest energy state violates the gauge principle and generate effective masses for the particles passing through it. Therefore a new field, the Higgs field, is introduced that allows other particles to acquire mass by interacting with it [10, 11], while at the same time preserves the gauge invariance. The shape of

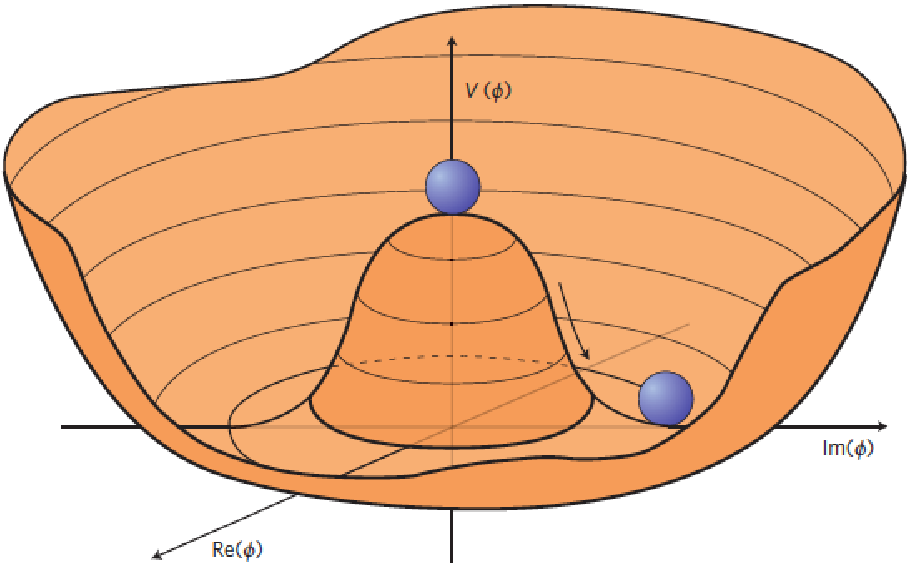


FIGURE 1.1: Higgs potential described from the last two terms of Equation 1.17 with $\mu^2 < 0$ and $\lambda > 0$. The potential minimum is located on a continuous ring. The figure is taken from [12]

Higgs potential given in Equation 1.17 is shown in Figure 1.1. Due to the symmetry of this potential, there are an infinite number of degenerate Higgs ground states configuration, satisfying

$$\langle \Phi^\dagger \Phi \rangle = \frac{\nu^2}{2} \quad \text{with} \quad \nu \equiv \sqrt{\frac{\mu^2}{\lambda}}, \tag{1.18}$$

where ν is the vacuum expectation value (VEV) corresponding to ground state of Φ. It is the choice of this particular ground state configuration that produces the apparent symmetry breaking. Through a suitable rotation and re-parametrization of the field, Φ can be written as:

$$\langle \Phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu + H(x) \end{pmatrix} \tag{1.19}$$

where H is a scalar field with zero expectation value. The Higgs field correspond to the physical Higgs boson. Substituting the expression 1.19 in the scalar part of the electroweak Lagrangian Equation 1.6 give rise to terms proportional to $W_\mu^- W^{+\mu}$ and $Z_\mu Z^\mu$ which introduces explicit mass terms for the vector bosons. The masses for the vector bosons can be expressed in terms of ν and the electroweak coupling constant g:

$$m_W = \frac{1}{2}\nu g \quad , \quad m_Z = \frac{m_W}{\cos\theta_W} = \frac{1}{2}\nu\sqrt{g^2 + g'^2} \tag{1.20}$$

where the Weinberg angle θ_W is defined by:

$$\tan\theta_W \equiv \frac{g'}{g} \implies \sin^2\theta_W = 1 - \frac{m_W^2}{m_Z^2} \tag{1.21}$$

The mass of the Higgs boson particle is given by:

$$m_H = \sqrt{2\lambda}\nu \tag{1.22}$$

By using the same scalar doublet, masses for lepton and quarks can be generated.

1.3 Quantum Chromodynamics (QCD)

QCD is quantum field theory that explains the strong interactions in the SM. QCD has a gauge symmetry group of $SU(3)_C$ where the subscript C is referred to as color charge. This new quantum number distinguish three different possible states of the quarks (namely blue, red and green). $SU(3)_C$ group has eight generators corresponding to eight mediator vector bosons, the gluons, of strong force. The massless gluons allow self-couplings because they are color charged as well. The QCD Lagrangian can schematically be written as

$$\mathcal{L}_{QCD} = \mathcal{L}_{classical} + \mathcal{L}_{guage} + \mathcal{L}_{ghost} \tag{1.23}$$

The term $\mathcal{L}_{classical}$, is the classical Lagrangian density, contains the kinetic terms for the fields, and the mass term for the quark:

$$\mathcal{L}_{classical} = -\frac{1}{4}F^A_{\mu\nu}F^{\mu\nu}_A + \sum_{Flavor} \bar{q}_i(iD_\mu\gamma^\mu - m)_{ij}q_j \tag{1.24}$$

where the indices (i, j) are for quarks, (A, B) are for gluon colors and (μ, ν) are space-time indices. D_μ is the covariant derivative. The Lagrangian will be invariant under local gauge transformation if the gauge invariant derivative is defined as:

$$(D_\mu)_{ij} = \partial_\mu\delta_{ij} + ig_s(t^C A^C_\mu)_{ij} \tag{1.25}$$

$$(D_\mu)_{AB} = \partial_\mu\delta_{ij} + ig_s(T^C A^C_\mu)_{AB}. \tag{1.26}$$

The Equation 1.25 is for the quark fields, using the t^C generators of $SU(3)$ in the fundamental representation, and Equation 1.26 is for gluon fields, with the T^C generators in the adjoint representation. The relationship between the Gell-Mann matrix (λ^a) and these generators is give by:

$$t^a = \lambda^a/2 \tag{1.27}$$

Gell-Mann matrices are Hermitian, traceless and satisfy commutation relation $\left[\lambda^a, \lambda^b\right] = 2if^{abc}\lambda^c$. The generators have the following commutation relations:

$$\left[t^A, t^B\right] = if^{ABC}t^C \tag{1.28}$$

where f^{ABC} form a set called the structure constants of SU(3). They are real numbers and are completely anti-symmetric in the indices. The generators T^A follow the Lie algebra, and are given by $(T^A)_{BC} = -if^{ABC}$.

The generators also obey the normalization condition:

$$Tr(t^A t^B) = T_R \delta^{AB},$$

where $T_R = \frac{1}{2}$.

These generator also follow the identity:

$$\sum_A t^A_{ab} t^A_{bc} = C_F \delta_{ac},$$

where C_F is called the Casimir and is given by:

$$C_F = \frac{4}{3} = \frac{N_C^2 - 1}{2N_C}.$$

N_C is the number of colors. Another Casimir constant C_A can be found to handle the gluon interaction:

$$Tr(T^C T^D) = \sum_{A,B} f^{ABC} F^{ABD} = C_A \delta^{CD},$$

where ,

$$C_A = 3 = N_C.$$

The covariant derivative D_μ can be dotted with the gamma matrices to introduce the slashed notation $\not{D} = D_\mu \gamma^\mu$. The gamma matrices have the anti-commutation relation $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$. So, Equation 1.24 can be written as:

$$\mathcal{L}_{classical} = -\frac{1}{4} F^A_{\mu\nu} F^{\mu\nu}_A + \sum_{Flavors} \bar{q}_i (i\not{D} - m)_{ij} q_j \tag{1.29}$$

where $F^A_{\mu\nu}$ is the field strength tensor and q_i are the quark fields with mass mass m. The $F^A_{\mu\nu}$ is defined as:

$$F^A_{\mu\nu} = \partial_\mu A^A_\nu - \partial_\nu A^A_\mu - g_s f^{ABC} A^B_\mu A^C_\nu \tag{1.30}$$

where (A, B) are the indices for eight color degrees of freedom of the gluon field and g_s is the strength of the coupling interaction. The last term of Equation 1.30 involves the

$SU(3)$ structure constants and two gluon fields. When inserted into the Lagrangian, Equation 1.29 leads to self-interactions between gluons.

$\mathcal{L}_{classical}$ is invariant under local gauge transformations. For co-variant gauges $\partial_\mu A_\mu = 0$, the following term can be added to the Lagrangian:

$$\mathcal{L}_{guage} = -\frac{1}{2\lambda}(\partial^\mu A^A_\mu)^2. \tag{1.31}$$

Although $\mathcal{L}_{gauge}$ breaks gauge invariance but it does not effect physical quantities as they do not depend on the gauge fixing parameter λ and the particular gauge chosen.

Finally, another necessary term is added in non-Abelian theories. It is called ghost term is given by:

$$\mathcal{L}_{ghost} = \partial_{mu}\eta^{A\dagger}(D^{mu}_{AB}\eta^B). \tag{1.32}$$

Ghosts are scalar fields and they cancel nonphysical degrees of freedom that would otherwise propagate in gluon loops.

Now how to apply QCD to hadron collision experiments, we need to understand the quarks and gluons configuration inside the hadron. The QCD potential between two quark is given by:

$$V(r) \approx -\frac{4 \cdot \alpha_s}{3r} + k \cdot r, \tag{1.33}$$

where,

$$\alpha_s = \frac{g_s^2}{4\pi}. \tag{1.34}$$

The 1^{st} term in Equation 1.33 is a linear term that dominates at large distances r or equivalently small exchanged energies. The factor k can be understood as a "gluon-spring" tension similar to that of a harmonic oscillator. New quark-antiquark pairs can be created from the vacuum if the gluon field energy exceeds the mass of the new pair. This can lead to a cascade of particles emerging from the gluon field at sufficiently large energies which bind "free" quarks into color-neutral singlets called hadrons. Such a process is referred to as hadronization. In experiments, the momentum of a quark or gluon candidate is therefore inferred by clustering collimated particles into jets. Hadrons can be either mesons which consist of quark-antiquark pairs $(q\bar{q})$ or baryons consisting of quark triplets (qqq). These constituents of a hadron are referred to as "valence quarks". The valence quarks inside the proton will emit gluons that can further be splitted into quark-antiquark pairs. These additional quarks are called sea quarks. The gluon emissions from sea quarks will tend to be soft with respect to the original quark. This whole group of particles is commonly referred to as partons.

In hadron collision experiments at a sufficiently high momentum transfer, one can approximate all partons as free which allows to treat hadron-hadron scattering as a single parton-parton interaction instead. The momentum of a parton is expressed in terms of the fraction of the hadron momentum $p_{parton} = x \cdot p_{hadron}$, where x is also referred to as "Bjorken scaling variable" [13]. Then, the probability of finding two parton flavors f_I with momentum fraction x_i interacting at an energy scale μ_F in a hadron-hadron collision is given by the parton distribution function (PDF) as PDF(x_1, f_1, μ_F) and $\cdot$ PDF(x_2, f_2, μ_F). The PDFs are normalized such that

$$\sum_{f}^{partons} \int_0^1 dx \cdot x \cdot PDF(x, f, \mu_F) = 1. \tag{1.35}$$

The scale μ_F is called "factorization scale" below which non-perturbative low energy effects such as soft gluon emissions have been absorbed into the PDF. The PDF can not be calculated analytically but can be derived from fits to measured processes like deep inelastic scattering (DIS) or Drell-Yan. There are various groups estimating PDF sets from fits to experimental data using different approaches; e.g. CTEQ, MMHT (formerly MSTW), ABM, HERAPDF, and NNPDF. Exemplary distributions taken from the CT14NNLO are shown in Figure 1.2 for μ_F = 2 GeV and μ_F = 100 GeV. More information on this PDF set can be found in Ref. [14]. At a momentum fraction of x $\sim$ 0.2 one can observe an excess of the up and down quark distributions as expected from the valence quark composition of the proton. These vanish at low momentum fractions at which the contributions from virtual sea quarks and gluons dominate.

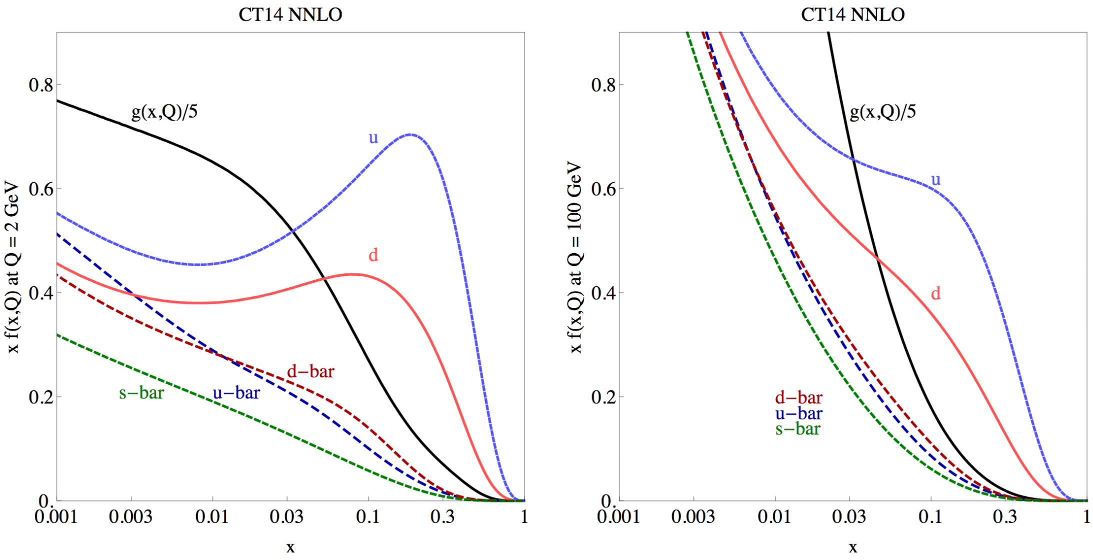


FIGURE 1.2: The CT14 NNLO PDFs at Q = 2 GeV (left) and Q = 100 GeV (right) for $u, d, s = \bar{s}, \bar{u}, \bar{d}$ and gluon (g). The gluon PDF has been scaled down by a factor of 0.2. The figure is taken from [15].

At high energies, quantum fluctuations lead to divergences. In order to let a theory still describe the experimental energy regime, physical quantities are redefined at the so-called "renormalization scale" μ_R. This leads amongst others to a "running" behaviour of the coupling constants as a function of μ_R. Beyond this scale, high energy effects such as loop corrections to propagators ("self-energy") are absorbed in the physical quantities through a renormalization of the fields. In particular, the running behaviour of the strong coupling constant is found to be

$$\alpha_s(\mu_R) = \frac{\alpha_s(\mu_0^2)}{1 + \alpha_s(\mu_0^2) \cdot \frac{33-2.n_f}{12\pi} \cdot \ln\left(\frac{|\mu_R^2|}{\mu_0^2}\right)} \tag{1.36}$$

where n_f denotes the number of quarks and μ_0 is a reference scale where the coupling is known from measurements, e.g. $\mu_0 = m_Z$. The present world average of coupling constant at Z mass is $\alpha_s(m_Z)$=0.1181±0.0011 [3]. Measurement summary of α_s as a function of Q is shown in the Figure 1.3.

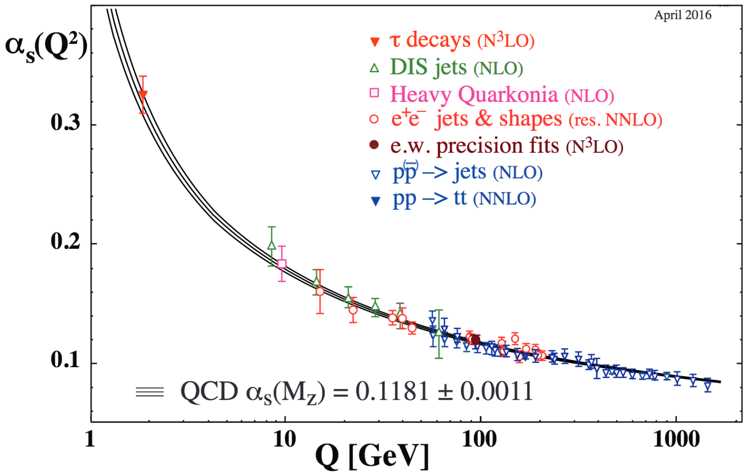


FIGURE 1.3: Measurement summary of α_s as a function of Q. The figure is taken from [16]

1.4 Shortcomings in SM

The SM is believed to be not the final theory describing fundamental TeV interactions. In the following some of its problems are briefly outlined.

- **Gravity:** The SM can not describe gravitational interactions.The mediator of gravity, the graviton, would have to be a spin 2 boson which however leads non-renormalizable divergences. At energy scales reached by colliders, it is fine nonetheless to ignore gravitational interactions since those become only comparable in strength to the electroweak and strong interactions at the Planck scale which is $\mathcal{O}(10^{18}$ GeV) if there are no extra dimensions.
- **Dark Matter:** Cosmological observations of rotation speeds of galaxies, mapping of matter distributions through micro lensing, and acoustic oscillations in the cosmic microwave background suggest that there exists a yet unidentified

kind of matter, called "dark matter". An even more puzzling building block of the universe which is missing from the SM is dark energy. It is a crucial ingredient to describe the expansion of the universe.

- **Naturalness:** A large correction to the bare Higgs mass originates from top quark loops via Yukawa interactions. Those need to be absorbed by the physical Higgs mass through renormalization. The resulting mass can be written as $m_H^2 \approx (m_H^{bare})^2 + \lambda_t^2 \mu^2/(16\pi^2)$. If the SM is valid up to the Planck scale it would require an extraordinary fine-tuning of the bare Higgs mass in order to cancel this large correction. Such a coincidence is considered to be not very "natural".

- **Unification of Forces:** The idea of the grand unification theory (GUT) where the weak, electromagnetic and strong forces are merged into one single force at some high unification or GUT scale, above which the three interactions unify [17]. This is not possible in the SM. Instead the SM prediction is that the three forces do not converge at the Planck scale.

- **Neutrino Oscillation:** According to the SM, neutrinos are massless. However experiments for solar, accelerators and reactors neutrinos have established that neutrinos can oscillate and can change flavor during flight. Only massive neutrinos can have such oscillations. Therefore, the flavor neutrinos $(\nu_e, \nu_\mu, \nu_\tau\ )$ are linear expressions of the fields of at least three mass eigenstates (ν_1, ν_2, ν_3) [18, 19].

- **Matter Antimatter Asymmetry:** It is observed that antimatter content is extremely small as compared to matter content in the universe. This difference in content is referred to as matter-antimatter asymmetry. The SM does not explain this asymmetry nor predict any mechanism that would have generated this asymmetry.

Chapter 2

Top Quark

The top quark, discovered in 1995 by the CDF and D∅ collaborations at the Tevatron accelerator at Fermilab using proton-antiproton collision data [20, 21], is the most massive known particle in SM so far. Top quark mass is 172.44 ± 0.49 GeV [22] and is close to the Higgs VEV ($\nu = 246$ GeV). Top mass gives an estimation for the Yukawa coupling

$$y_t = \frac{\sqrt{2}m_t}{\nu}. \tag{2.1}$$

This may hint at an important role of top quark in the electroweak symmetry breaking. Furthermore, top quarks are excellent probes to search for Beyond Standard Model (BSM) physics. For example, some extensions of the SM predict additional W' and Z' bosons which are heavier versions of their SM counterparts and can decay into top quarks. Other BSM models can have a wide Higgs sector including additional charged and neutral Higgs bosons which may couple to massive top quark.

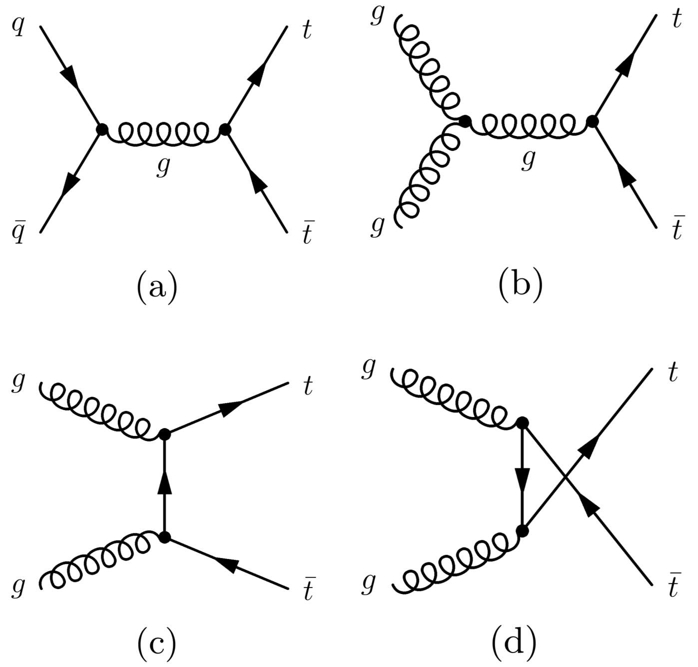
Historically, the top and bottom quarks were proposed in 1973 to make the third family of the quarks. Renormalizability of the SM provides the basis for the existence of top quark. Before the τ lepton discovery, the standard model was a highly successful theory of fundamental interactions. Its success hinged on the facts that it gave a good quantitative predictions for all observed process and the theory was renormalizable. This state of affairs changed with the discovery of tau lepton in 1975 [23]. Detailed studies of the production and decay of τ showed that, except for its mass, it has exactly the same properties as the electron and muon. In other words, it is a third generation lepton. The renormalizability of the standard model, therefore, requires that third generation quark exist. One of these, the bottom quark, was discovered in 1977 with a mass $m_b = 4.2$ GeV. The other 3^{rd} generation quark is, by definition, the top quark, and it ought to exist. The Super Proton Synchrotron (SPS) experiment at CERN discovered W and Z bosons but top quark remained elusive at these energies.

Top quark lifetime (10^{-25} s) is smaller than QCD interactions timescale ($1/\Lambda_{QCD} \sim 10^{-23}$ s) [24]. Therefore, top quark decay before the hadronization process starts. Thus, top quark can not form bound state and its polarization and spin information propagates through the W boson to its decay products.

2.1 Top Quark Pair Production

The dominant process producing top quarks at hadron colliders is top quark pair production through strong interactions via gluons ($gg \rightarrow t\bar{t}$) or quarks fusion ($q\bar{q} \rightarrow t\bar{t}$). The electroweak $t\bar{t}$ production via Z or γ exchange is negligible at hadron colliders because of large cross section of QCD processes. The kinematic threshold for top pair production is $\hat{s} \geq 4m_{top}^2$, where $\hat{s} = x_1 x_2 (P_A + P_B)^2$ is defined as the initial center of mass (CoM) energy. P_A and P_B are the incoming hadrons momenta. For a symmetric collider with both hadrons having the same momenta, $\hat{s} = 4x_1 x_2 P^2 = x_1 x_2 s$, where s is the square of hadron CoM energy. Tevatron is operated at $\sqrt{s} = 1.96$ TeV, hence $x \approx \frac{2m_{top}}{\sqrt{s}} = 0.18$. Therefore, the dominant production process is quark-antiquark annihilation with a contribution of 85% to 90%, as can be seen in the PDF visualised in the Figure 1.2. In case of the LHC with $\sqrt{s} = 14$ TeV, $x \approx 0.025$ is obtained. Consequently, the lower x region is more important, making gluon fusion processes dominant. The Feynman diagrams for the $t\bar{t}$ process at the leading-order (LO) are shown in Figure 2.1.

The $gg \rightarrow t\bar{t}$ channel is the leading production channel in the LHC energy regime [24]. The theoretical $t\bar{t}$ cross sections are given in table 2.1. The cross sections calculated at next-to-next-to-leading-order (NNLO) + next-to-next-to-leading logarithmic (NNLL) accuracy using the Top^{++} 2.0 program [25, 26, 27]. The PDF uncertainty in the Table 2.1 includes also the uncertainty on α_s.



(a) (b)

(c) (d)

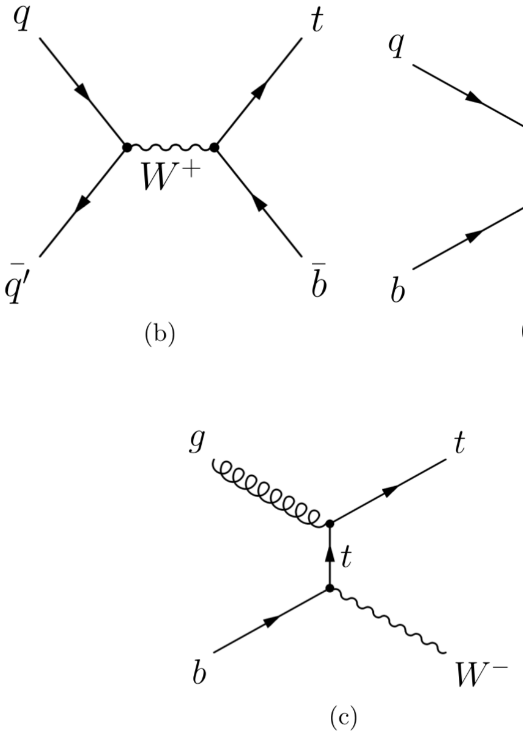
FIGURE 2.1: The figure illustrates the LO Feynman diagrams of the $t\bar{t}$ production through $q\bar{q}$ fusion at (a) and through gg fusion at (b), (c), (d) [28].

TABLE 2.1: Predicted $t\bar{t}$ cross sections calculated in pp collision [25, 26, 27].

CoM energy	$t\bar{t}$ cross section
8 TeV	$252.9^{+6.4}_{-8.6}$ (scale) + 11.7 (PDF) pb
13 TeV	$831.8^{+19.8}_{-29.2}$ (scale) + 35.6 (PDF) pb

2.2 Single Top Quark Production

The $t\bar{t}$ is produced via strong interaction, but it is also possible to produce a single top quark (referred as "single top (t, $\bar{t}$)") through the electroweak interaction. This electroweak production process was discovered in 2009 by the *CDF* and $D\oslash$ experiments [20, 21, 29]. There are three production modes for this kind of top quark production. All three channels are shown in Figure 2.2. All processes include a W_{tb} vertex and therewith a vertex where the CKM matrix element V_{tb} directly occurs. This allows an explicit measurement of $|V_{tb}|$. The single top production channels are s-, t- and tW-associated channels.



(b) (a)

(c)

FIGURE 2.2: The figure illustrates the LO Feynman diagrams of the electroweak production of single top in t-channel at (a), s-channel at (b) and tW-associated channel at (c) [30].

The production channels of single top through W$_{ts}$ and W$_{td}$ vertex are greatly suppressed in SM as compared to W$_{tb}$ vertex. This suppression is because of small off-diagonal CKM matrix elements. CKM matrix elements dealing with top quark can not be measured directly. $|V_{td}|$ is determined indirectly from $B_d^0 - \bar{B}_d^0$ oscillations. In the same way $|V_{ts}|$ is estimated from $B_s^0 - \bar{B}_s^0$ mixing. With both these values and using the unitarity constraint, $|V_{tb}|$ is determined to be close to 1. Current results are [31]:

$$|V_{td}| = 0.0075 - 0.0084 \tag{2.2}$$

$$|V_{ts}| = 0.041 - 0.042 \tag{2.3}$$

$$|V_{tb}| = 0.999096 - 0.0999133 \tag{2.4}$$

Thus, $|V_{tb}|$ is the most precise known CKM element. Cross section measurements of single top quark, allows to extract a limit on the CKM matrix element $|V_{tb}|$. If one assumes $|V_{td}|+|V_{ts}| \ll |V_{tb}|$ then the $t \to bW$ branching ratio can be approximated as

$$R = \frac{\mathcal{B}(t \to Wb)}{\mathcal{B}(t \to Wq)} \tag{2.5}$$

$$R = \frac{|V_{tb}|^2}{|V_{tb}^2| + |V_{ts}^2| + |V_{td}^2|} \approx 100\% \tag{2.6}$$

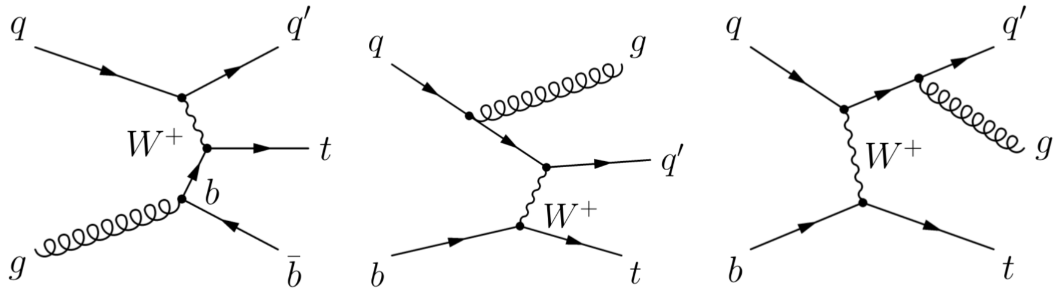
In the following all three channels of t or $\bar{t}$ production via the W_{tb} vertex will be discussed.

2.2.1 t-channel Single Top Production Mode

The production via t-channel (Figure 2.3) has the highest cross section in pp collisions. The virtuality of the W boson is found to be $Q^2 < 0$ and hence it is said to be "space-like". A characteristic feature of this mode is the production of an additional spectator quark (q') which recoils against the W boson and tends therefore to be scattered fairly forward in the CMS detector. In this production mode top quarks are produced roughly twice more often than top antiquarks which is a consequence of the up-over-down valence quark composition of the proton. Hence the production ratio, $\sigma_t/\sigma_{\bar{t}}$, is sensitive to the PDF of the proton. The ratio depends however on the center of mass energy since at higher energies lower momentum fractions are probed at which contributions from the valence quarks become less dominant.

The t-channel cross section is computed at NNLO in QCD [33, 34, 35] and at NLO with NNLL resummation [14]. The calculations performed using $HATHORv2.1$ program at NLO [36, 37], based on $MCFM$ [38]. The NNLO and NLO+NNLL predictions at different LHC energies are compared in the Table 2.2.

The single top t-channel production cross section rises to the same magnitude as $t\bar{t}$ production even though it is a weak interacting process. The reasons is that the cross



(a) (b) (c)

Figure 2.3: Single top production Feynman diagrams at NLO in t-channel [32].

section of a weak interacting process scales like $1/M_W^2$, whereas the strong interaction cross section scales with $1/\hat{s}$. Additionally, the $t\bar{t}$ production via gluon fusion is suppressed by a factor of $\frac{1}{3}$ as the colour combination at the gluon vertex has to match. W-gluon fusion is not colour suppressed. Finally, the production of t or $\bar{t}$ is kinematically enhanced in comparison to the production $t\bar{t}$.

Table 2.2: The t and $\bar{t}$ cross section in the t-channel at LHC energies. The NNLO predictions at 8 TeV and 13 TeV use $\mathrm{m}_{top} = 172.5$ GeV and $\mathrm{m}_{top} = 173.2$ GeV, respectively, and the NLO+NNLL predictions calculated with $\mathrm{m}_{top} = 173$ GeV. The cross-sections are reported for (t), ($\bar{t}$) and (t+$\bar{t}$) [14, 33, 34, 35].

Corrections	cross section	7 TeV	8 TeV	13 TeV
NLO	σ_t	-	$54.2^{+0.5}_{-0.2}$	$134.3^{+1.3}_{-0.7}$
	$\sigma_{\bar{t}}$	-	$29.7^{+0.3}_{-0.1}$	$79.3^{+0.8}_{-0.6}$
	$\sigma_{(\bar{t}+t)}$	-	$83.9^{+0.8}_{-0.3}$	$79.3^{+2.1}_{-1.1}$
NLO + NNLL	σ_t	$43.0^{+1.8}_{-0.9}$	$56.4^{+2.4}_{-1.2}$	$136.0^{+4.0}_{-3.0}$
	$\sigma_{\bar{t}}$	$22.9^{+0.9}_{-1.0}$	$30.7^{+1.5}_{-1.6}$	$82.0^{+3.0}_{-2.0}$
	$\sigma_{(\bar{t}+t)}$	$65.9^{+2.6}_{-1.8}$	$87.2^{+3.4}_{-2.5}$	$218^{+5.0}_{-4.0}$

2.2.2 s-channel Single Top Production Mode

Amongst the three main single top production channels, s-channel (Figure 2.4) has the smallest cross section. This is because of the "time- like" W boson ($Q^2 > 0$) which has to have a large virtuality to produce the heavier top quark. In various BSM scenarios however the cross section of this process is expected to increase due to new heavy particles like W' or charged Higgs bosons that may even be produced on their mass shell, $p_\mu p^\mu - m^2 = 0$, and hence occur as a resonance. The theoretical production cross-section for s-channel obtained at NLO+NNLL [39] and at NLO [36, 38, 40, 37]. The NLO and NNLL predictions at LHC energies have been compared in the Table 2.3. To observe s-channel prediction at Run 2 is even harder than Run 1 as the cross section is increased only by a factor of 2 from 8 TeV to 13 TeV.

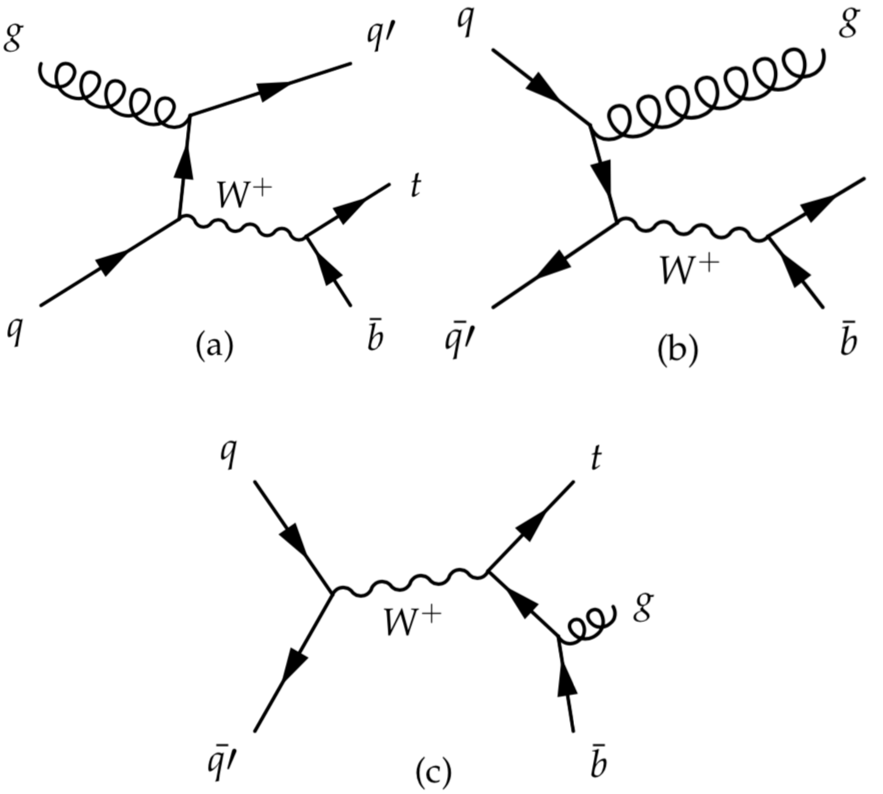


FIGURE 2.4: Single top production Feynman diagrams at NLO in s-channel [41].

TABLE 2.3: The s-channel t and $\bar{t}$ cross section is shown here. NLO+NNLL and NLO predictions have been calculated with m_{top} = 173 GeV and 172.5 GeV respectively. The cross sections are reported for t, $\bar{t}$ and $(t+\bar{t})$ [36, 37, 38, 39, 40].

Corrections	s-channel cross section	7 TeV	8 TeV	13 TeV
NLO	σ_t	2.8 ± 0.1	3.3 ± 0.1	6.3 ± 0.4
	$\sigma_{\bar{t}}$	1.5 ± 0.1	1.9 ± 0.1	4.0 ± 0.2
	$\sigma_{(\bar{t}+t)}$	4.3 ± 0.2	5.2 ± 0.2	10.3 ± 0.2
NLO + NNLL	σ_t	3.1 ± 0.1	3.8 ± 0.1	7.1 ± 0.2
	$\sigma_{\bar{t}}$	1.4 ± 0.1	1.8 ± 0.1	4.1 ± 0.2
	$\sigma_{(\bar{t}+t)}$	4.6 ± 0.2	5.6 ± 0.2	11.2 ± 0.4

2.2.3 tW-channel Single Top Production Mode

The third mode is the tW-associated production where the W boson can be produced on-shell ($Q^2 = m_W^2$). The Feynman diagram of the tW-associated production is shown in the Figure 2.5. This process interferes at NLO with $t\bar{t}$ production which complicates its definition. In the diagram removal (DR) scheme diagrams with two resonant top quarks are subtracted from the common amplitude whereas in the diagram subtraction (DS) scheme the contribution from $t\bar{t}$ is locally eliminated from cross section [42, 43]. The difference between both schemes lies in the treatment of the interference term which is kept in the DS but removed in the DR. A new approach has been developed where the $t\bar{t}$ and tW production and the inference between them are combined into a process with a $W^+W^-b\bar{b}+X$ final state [44] whose simulation is currently being studied [45]. The tW cross section is calculated at NLO+NNLL [46] and at *NLO* with *HATHOR* [36, 37] based on *MCFM* [37]. The Table 2.4 compares the two predictions at LHC energies.

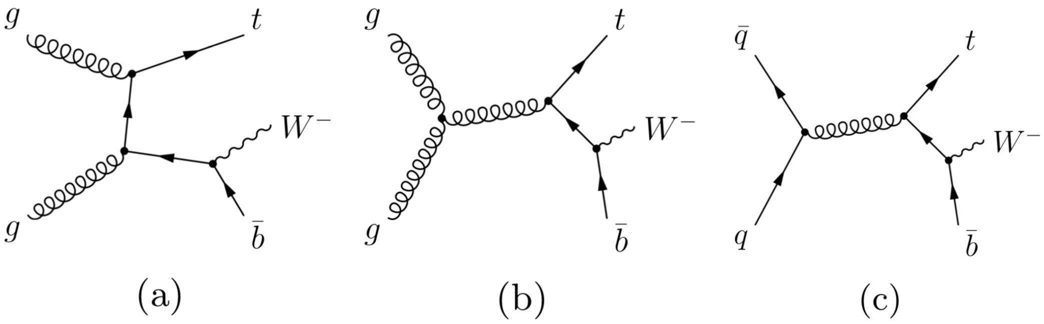


FIGURE 2.5: Leading order Feynman diagram for tW-associated channel production mode of single top quark [47].

TABLE 2.4: The tW-associated t cross section at LHC energies. Both NLO+NNLL and NLO predictions have been calculated with m_{top} =172.5 GeV [46].

Corrections	tW channel cross section	7 TeV	8 TeV	13 TeV
NLO + NNLL	σ_{tW}	17.0 ± 0.7	24.0 ± 1.0	76.2 ± 2.5
NLO	σ_{tW}	13.2 ± 1.4	18.9 ± 1.9	60.0 ± 6.0

2.3 Single Top Quark Production Mode in Association with a Z boson

The t-channel production mode cross section of single top is sufficiently large at LHC to study the coupling of top quark with additional particles. One such example is where t or $\bar{t}$ are produced in association with a Z boson as shown in Figure 2.6. This process is sensitive to the coupling of top quark with Z boson as illustrated in Figure 2.6 (middle-right). In addition to top quark coupling to Z boson, the process is also sensitive to the triple gauge boson coupling WWZ as shown in Figure 2.6 (bottom-left). The process also has a slight input from non-resonant lepton pairs as shown in Figure 2.6 (bottom-right).

The QCD-induced top pair production has higher cross section as compared to electroweak single top production at LHC. Therefore, the top pair associated production channels like $t\bar{t}$Z and $t\bar{t}\gamma$ seem more sensitive than tZ. The production of additional massive Z boson suppresses the phase space of the process and has effectively removed the top pair higher cross section advantage over single top. The Z-associated production cross sections of t, $\bar{t}$ and $t\bar{t}$ are shown in Figure 2.6 [48]. The cross sections of $t\bar{t}$Z and $t\bar{t}\gamma$ calculated by CMS at 8 TeV are shown in the Table 2.5 [49, 50]. The theoretical NLO calculations for tZ and $\bar{t}$Z are also given in Table 2.5. The details of these calculations are given in reference [48].

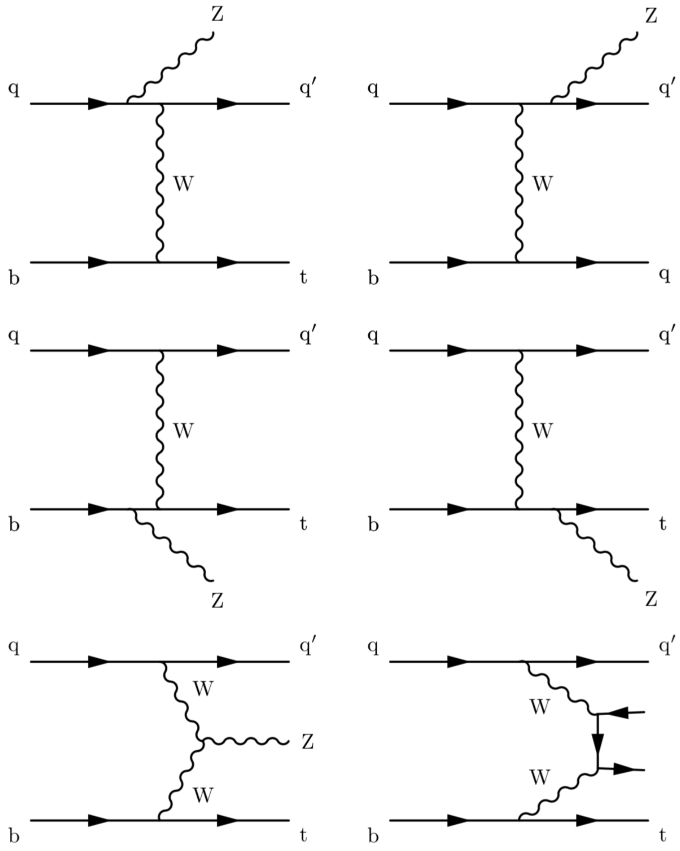


FIGURE 2.6: The LO production diagrams of the tZq process shown. The bottom-right diagram shows the non-resonant input to the process [51].

The signatures of the process are a t or $\bar{t}$, a Z boson and additional quark that is why it is referred as tZq process. The tZq process probes two SM couplings (tZ and triple gauge-boson couplings) in a single process, whereas the similar final state $t\bar{t}Z$ only probes the tZ coupling. These couplings are sensitive to beyond SM physical processes with same experimental signatures. An example of this is flavor changing neutral current (FCNC) which involve the direct coupling of the top quark to a Z boson and an up or charm quark [52, 53]. In SM, FCNC processes are forbidden at LO and are suppressed at higher levels [54]. Any deviation from this expected SM process can be hint of new beyond SM process.

TABLE 2.5: CMS measurement of $\bar{t}Z$ and $t\bar{t}\gamma$ cross sections at 8 TeV are given. Also the NLO measurements for the tZ+$\bar{t}$Z are shown here [49, 50].

Process	cross section (fb)
$t\bar{t}Z$	200^{+80}_{-70}(stat)$^{+40}_{-30}$(syst)
$t\bar{t}\gamma$	2400±200 (stat) ±600(syst)
tZ	160^{+7}_{-2}(scale) + $^{+11}_{-11}$(pdf)
$\bar{t}Z$	76^{+4}_{-1}(scale)$^{+5}_{-5}$(pdf)
tZ + $\bar{t}$Z	236^{+11}_{-4}(scale)± 11(pdf)

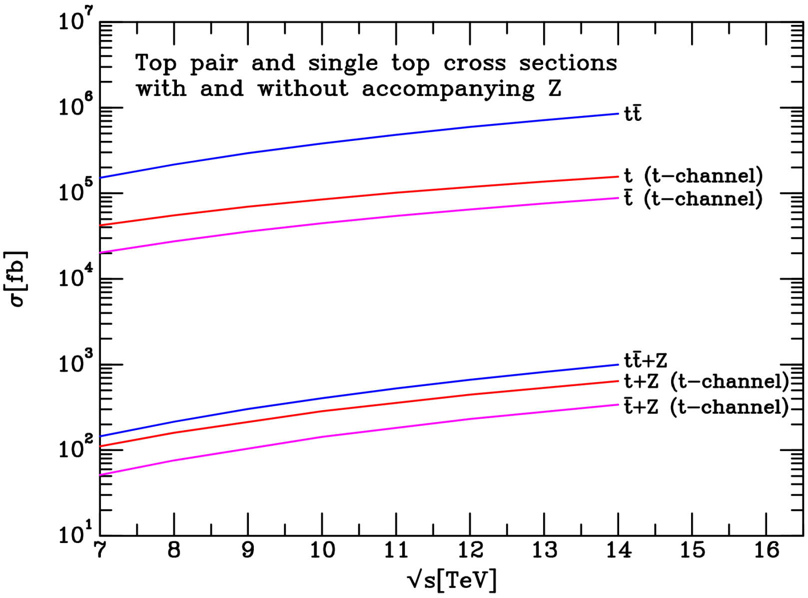


FIGURE 2.7: Production cross sections (at NLO) of $t\bar{t}$, t and $\bar{t}$) with and without an associated Z boson [48].

The tZq process is the main topic of this thesis. The easiest way to find the process is in three lepton final state. For three lepton final state, top quark decays leptonically via a b quark and W boson. The other two leptons come from the leptonic decay of Z boson. The leptons from Z boson should be same flavor and opposite sign and mass compatible with on-shell Z mass. There is a neutrino in the event, referred as missing transverse energy ($\vec{\not{E}}_T^{miss}$) in the accelerator terms. The schematic view of a pp collision producing tZq event is shown in Figure 2.8.

Finally, there are two hadronic jets present in the event. One is b quark which is originated from the top quark decay and other is due to the recoil quark. The recoil quark is usually a light quark in forward region of the detector also known as forward jet. This recoil quark may also come from a b quark; in this case it will be referred as tbZ process. There may also be an additional jet from the gluon splitting that creates the b-quark in the initial state.

The final state of tZq is therefore three leptons, a large amount of $\vec{\not{E}}_T^{miss}$ and two jets. The analysis is performed using 2012 CMS data at $\sqrt{s} = 8$ TeV and 2016 CMS data at $\sqrt{s} = 13$ TeV. The analysis at $\sqrt{s} = 8$ TeV will be discussed in chapter 7 and at $\sqrt{s} = 13$ TeV will be discussed in chapter 8.

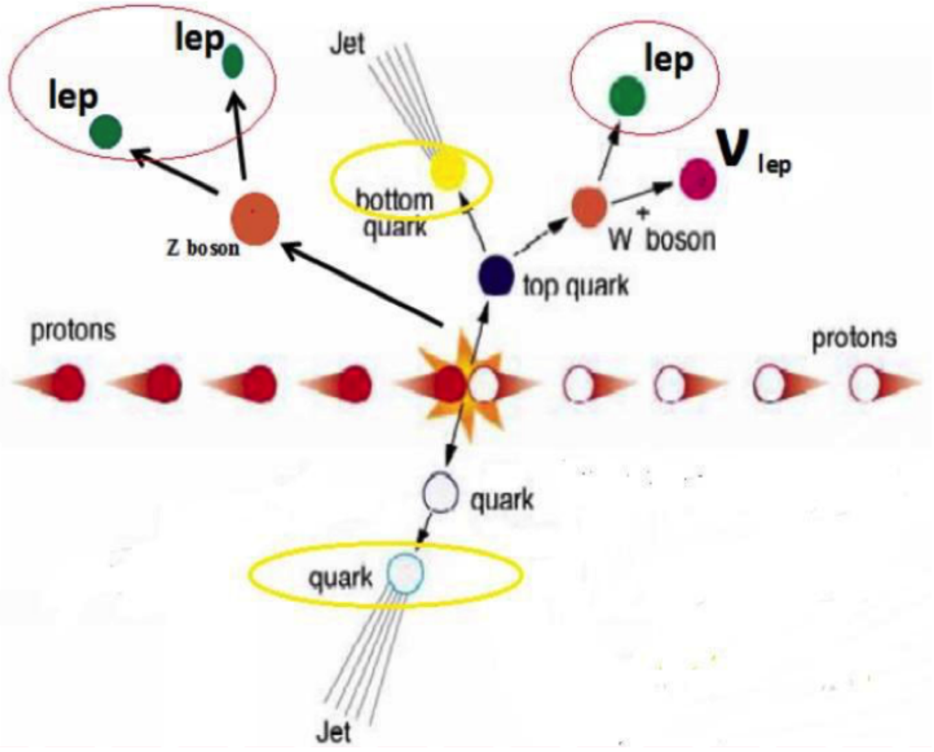


FIGURE 2.8: Schematic diagram of tZq event showing three leptons, one b quark, one light quark and neutrino in the final state.

2.3.1 Experimental Results

To conclude this chapter, a selection of experimental results of the top quark are presented. An overview of inclusive $t\bar{t}$ cross section measurements at $\sqrt{s} = 5$ TeV, 7 TeV, 8 TeV and 13 TeV are shown in Figure 2.9 and compared to the prediction. Measurements of inclusive single top cross section in t-, tW-, and s-channel at $\sqrt{s} =$ 7 TeV, $=$ 8 TeV, and $=$13 TeV are summarized and compared to the theoretical NLO predictions in Figure 2.10. The measurements and SM predictions are in agreement.

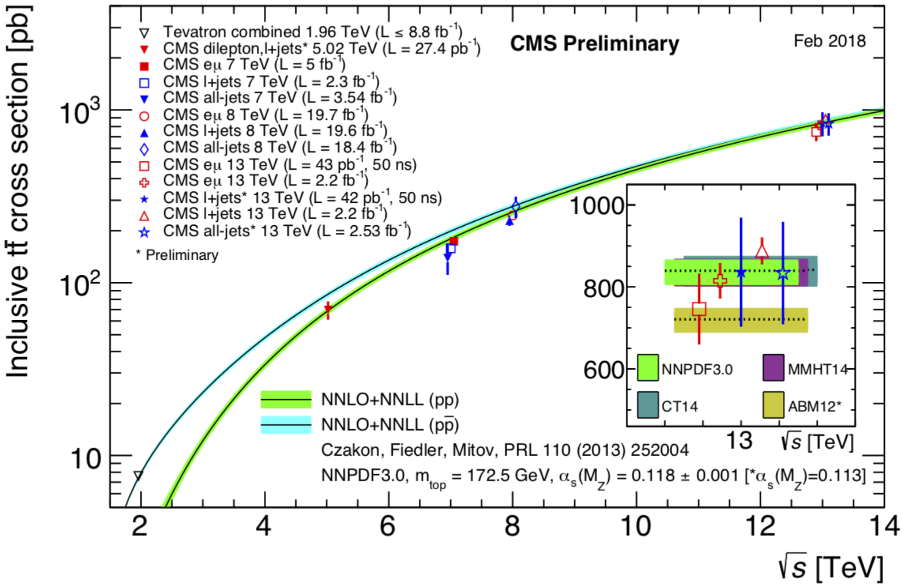


FIGURE 2.9: The CMS summary of $t\bar{t}$ cross section measurements as a function of CoM [55].

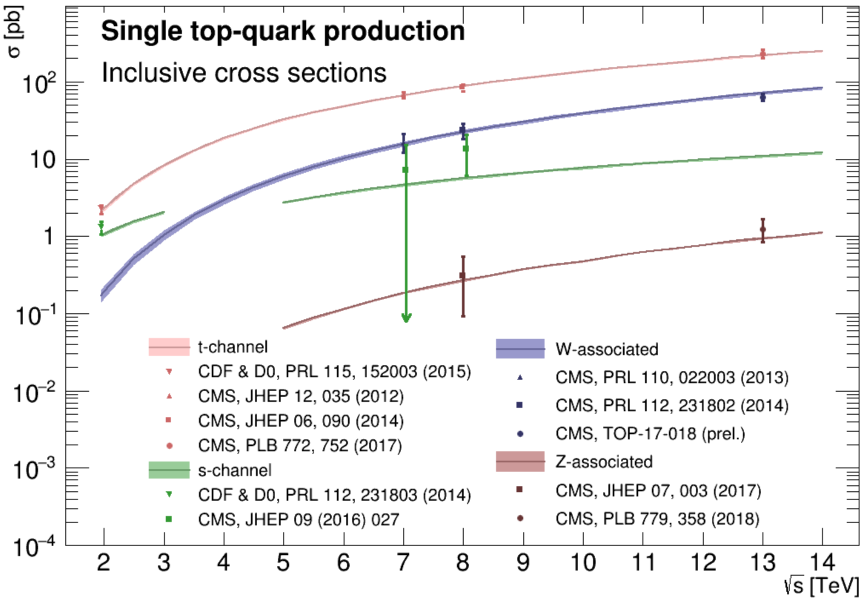


FIGURE 2.10: The CMS summary of single top cross section measurements as a function of $\sqrt{s}$ in different channels. Z-associated productions are explained in this thesis [56].

Chapter 3

Large Hadron Collider

The Large Hadron Collider (LHC) at CERN (the European Organization for Nuclear Research), in Geneva, Switzerland, is the world's largest and most powerful tool for particle physics research. It is installed in the 26.7 km long ring tunnel [57] that was previously housing the LEP machine [58]. The LHC is designed to collide proton beams and heavy ions. The main objective of the LHC was the search for the Brout-Englert-Higgs boson.

The LHC design energy is 14 TeV. The limiting factor for the achievable energy is the magnetic field strength (B) in the dipole magnets that are used to keep the particles on the ring. Particles circulating in the LHC are organized in bunches, due to the acceleration with radio frequency (RF) cavities. The nominal number of bunches in the ring is 2808 and two bunches are separated by 25 ns.

3.1 CERN Accelerator Complex

The schematic view of CERN accelerator complex is shown in Figure 3.1. Before reaching the main LHC tunnel, the protons bunches pass through an injection chain consisting of CERN's acceleration structure that was already present before the LHC was built. The first linear accelerator of the complex, LINAC 2, accelerates the protons upto 50 MeV. With this energy, the protons are then directed to the second accelerator the Proton Synchrotron Booster (PSB). At the PSB (or Booster) proton packs are accelerated up to 1.4 GeV and each pack is split up in twelve bunches with 25 or 50 ns spacing [59]. Then the proton bunches are injected into the Proton Synchrotron (PS) where they gain energy up to 25 GeV before injecting in the Super Proton Synchrotron (SPS). At SPS protons accelerate up to 450 GeV and are then directed into two separate beam pipes of LHC where they travel in opposite direction. Superconducting dipole magnets (1232 in number) are used to keep the beams at fixed orbit. The superconducting dipole magnets ($B = 3.8$ T) are kept at a temperature of 1.9 K using liquid helium. To keep the beams focused, a total of 392 main superconducting quadrupole magnets are used. Particles are accelerated through the electric field generated by 8 superconducting RF cavities. These cavities operate at 400 MHz.

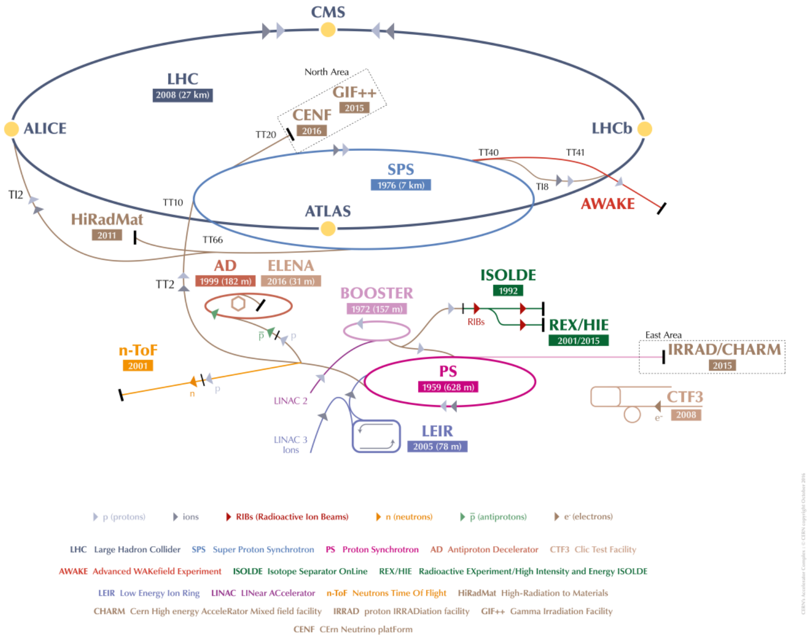


Figure 3.1: Schematic overview of the CERN accelerators complex. The figure is taken from [60].

3.1.1 Experiments at CERN

CERN is capable of colliding protons and heavy ions. Following the LEP tunnel geometry, the LHC has 8 curved and eight straight sections. Detectors that record events from collisions are installed in four of the straight sections numbered clockwise, starting from point-1, close to the CERN Meyrin site.

1. CMS (Compact Muon Solenoid), located at point-5, is a multi-purpose detector with broad physics objectives [61]. In addition to its capability to find the Higgs boson with masses up to 1 TeV, other objectives are defined for example, the search for new physics, dark matter candidates or supersymmetric particles, more details about this detector will be described in Section 2.

2. ATLAS (A Toroidal LHC ApparatuS), located at point-1, is also a multi-purpose detector and has the same objectives as CMS but exploits different technical design, including a large toroidal magnet [62]. The ATLAS is a cylindrical detector, 44 m long and 25 m diameter, around the beam pipe.

3. ALICE (A Large Ion Collider Experiment), located at point-2, focuses on heavy ion collisions in which the properties of quark-gluon plasma can be studied. [63].

4. LHCb (Large Hadron Collider beauty): It aims to study precision measurements of CP-violating processes and searches for rare decays of B hadrons amongst others. [64].

Although ATLAS and CMS analyze the same dataset independently and have similar goals, the technical realization of two experiments is different and their results can verify each other, and can also be merged to improve the precision.

There are several other smaller experiments located near the interaction points. These experiments study the particles emitted in the very forward region. For example MoEDAL [65] is an experiment to search for the magnetic monopoles, TOTEM [66] studies the protons themselves as they emerge from collision and LHCf [67] uses protons in the forward regions as a source to simulate cosmic rays.

3.1.2 Luminosity

One very important parameter to characterize the accelerator is luminosity which is a measure of the collision rate. The number of collisions in a detector is related to instantaneous luminosity L_{ins} as:

$$N = L_{ins} \times \sigma \tag{3.1}$$

in which σ is the cross section of the expected process. L_{ins} is defined as:

$$L_{ins} = \frac{N_{1b} N_{2b} n_b f_{rev}}{4\pi \sigma_x \sigma_y} F \tag{3.2}$$

The values of these parameters are given in Table 3.1.

Table 3.1: Summary of the parameters entering Equation 3.2. The parameters refer to the nominal values during 2015 data taking.

Label	Quantity	Value
N_{1b}, N_{2b}	No. of Particles per bunch	1.2×10^{11}
n_b	No. of bunches per beam	2808
f_{rev}	Rev. frequency	$11245 kHz$
γ_r	Rel. gamma factor	~ 7000
F	Geometrical factor	0.84
σ_x, σ_y	Transverse beam size	$\sim 17 \mu m$

The geometrical factor, F, accounts for reduction in luminosity due to a slight tilting of the beams by the crossing angle θ.

$$F = \frac{1}{\sqrt{1 + \left(\theta \cdot d_z\right) / \left(2 d_{x,y}\right)^2}} \tag{3.3}$$

It is seen in Equation 3.1 that the instantaneous luminosity is hard to be kept stable during the data taking. In fact, it decays mainly due to the loss of protons in the bunches leading to a decrease in collisions. Therefore, another quantity, integrated luminosity (L_{int}), is introduced by integrating the instantaneous luminosity over time, with a unit of inverse of barn denoted by b^{-1} (per cm^2), to measure how much data is recorded by the LHC and CMS. Figure 3.2 shows how the total integrated luminosity accumulated versus time during the LHC data taking.

The amount of data recorded by a detector is usually smaller than the amount of data delivered. It is because the data acquisition system is not always 100 % efficient. In addition, not all the data is good enough for analysis purpose as some of the subsystems may be in an error state that corrupt the data recorded or other unexpected behaviours of the LHC or CMS may have interrupted the data recording. All the data must be certified by a dedicated task force in CMS. As a result, the total amount of data used in an analysis can be much smaller than the total amount of data delivered.

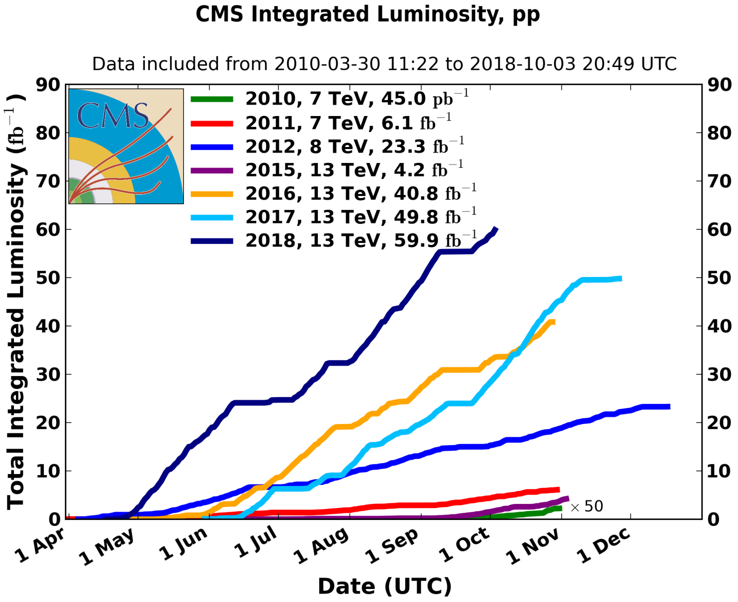


FIGURE 3.2: L_{int} delivered to CMS versus data taking period for pp collisions at different LHC energies. The data collected in the year 2012 and 2016 has been analyzed in this thesis [68].

3.1.3 Pileup (PU)

Another important parameter of the accelerator is pileup. During a bunch crossing multiple proton-proton interactions can occur which are referred to as pileup. Their number on average is proportional to the luminosity times the total inelastic pp cross section. There are two types of pileups:

- in-time PU comes from the collisions happening in the same bunch crossing;
- out-of-time PU comes from the collisions from the previous bunch crossings whose signal has not yet been recorded by the detector.

In-time PU is calculated by the number of primary vertices (N_{PV}) while out-of-time PU is parametrized by the mean number of interactions per bunch crossing (μ). In 2012, an average of 21 pileup interactions has been observed at 8 TeV. This increased in 2016 due to the higher luminosity and cross section at 13 TeV to about 27 interactions on average. The $\langle\mu\rangle$, calculated for 2011-2012 and 2015-2018 CMS data taking period, is shown in the Figure 3.3 .

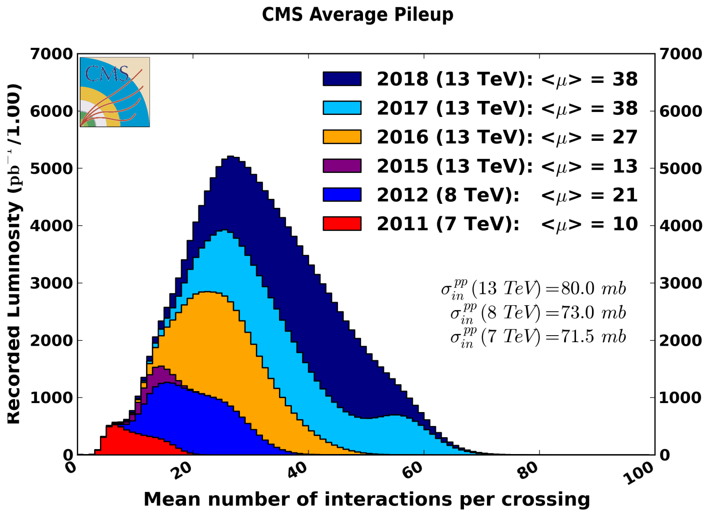


FIGURE 3.3: The distribution of $\langle\mu\rangle$ versus luminosity recorded by CMS for the 2011-2012 and 2015-2018 data taking period [68].

3.2 CMS Experiment

CMS is one of the main LHC experiment, situated in a cavern especially excavated 100 meters underground. The sailent features of CMS that distinguish it from other experiments are the high field superconducting solenoid, a full-silicon inner tracker and a homogeneous electromagnetic calorimeter based on scintillating crystals without non-active absorbers. The main purpose of CMS is the general search for new physics beyond the SM with a special focus on Higgs searches and supersymmetry but it also records data during heavy-ion operation.

3.2.1 CMS Overview

The CMS is designed as a hermetic 4π detector where several subdetectors surround the point of the particle collisions. CMS weighs 14000 tons having a length of 28.7 m and a diameter of 15 m. CMS uses right-hand coordinate system with origin at the collision point, z-axis points counterclockwise along the beam pipe, y-axis points upwards and the x-axis towards the center of the ring. The azimuthal angle ϕ is defined in the transverse plane, spanned by the x- and y-axes, which lies perpendicular to the beam pipe. In this plane, the radial coordinate r is defined as r=$\sqrt{x^2+y^2}$ and the polar angle θ measured from the z-axis, in a plane spanned by r- and z-axis perpendicular to the x- and y-plane. The polar angle θ defines another variable which is called pseudorapidity η as:

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right). \tag{3.4}$$

Pseudorapidity η is widely-used in accelerator physics because it has the advantage of being Lorentz invariant and is used as a measure of the angle of a particle with respect to the z-axis.

Figure 3.4 shows the key elements of the CMS detector. The CMS design is driven by the choice of the magnetic field layout which features a solenoid type field with a field strength of 3.8 Tesla. This immense bending power is necessary to attain a good momentum resolution which is an important indicator for interesting physics processes. Further information can be found in Refs [61, 69]. The individual sub-components of CMS will be presented in the following sections.

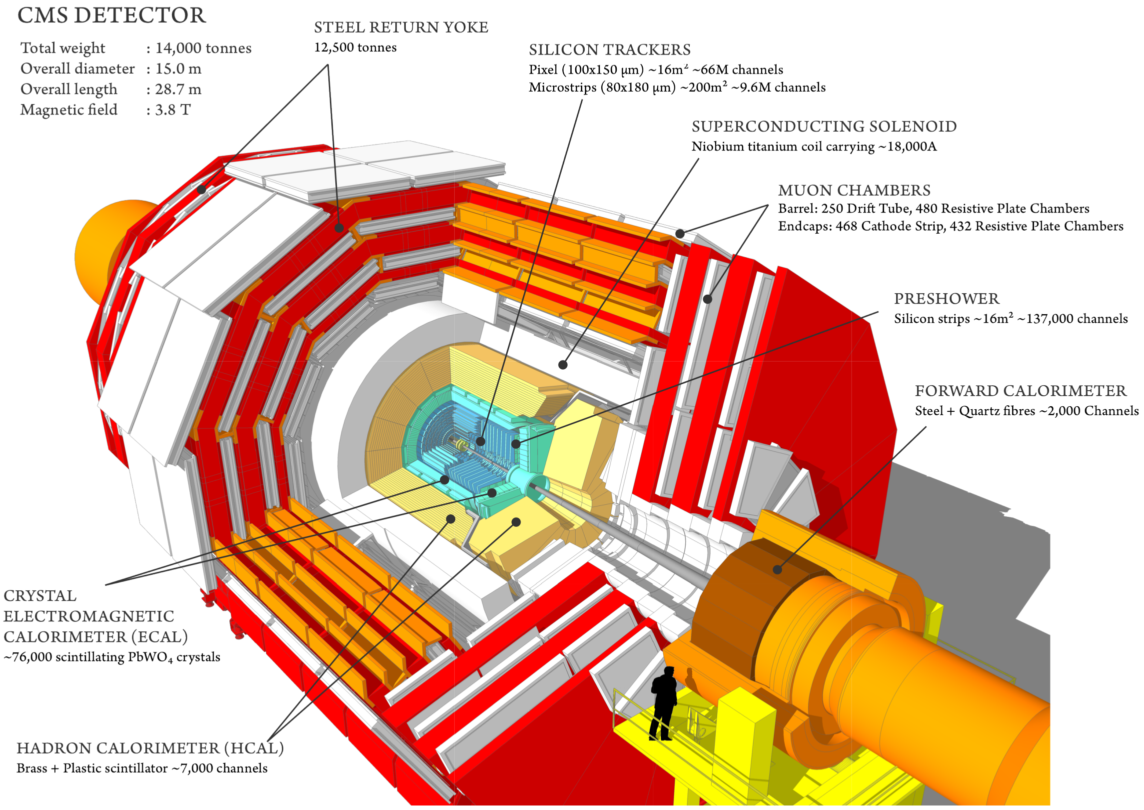


FIGURE 3.4: Schematic drawing of CMS detector with all subsystems. [70].

3.2.2 Tracking System

CMS has an all silicon tracker with a total surface of 210 m^2. Its main purpose is to measure the trajectories of the particles that are created in the collisions and to reconstruct secondary vertices. Details about the tracking system will be presented in chapter 5.

3.2.3 The Electromagnetic Calorimeter (ECAL)

The CMS ECAL is a hermetic and homogeneous calorimeter constructed of 61200 lead-tungsten crystals ($PbWO_4$) in the barrel region and 7324 crystals in each of the end cap regions. This means that the absorber material is also the signal producing material. Its purpose is to measure and absorb electromagnetically interacting particles. The ECAL encloses the inner tracker and covers a range of $|\eta| < 3$. The ECAL layout overview is shown in Figure 3.5.

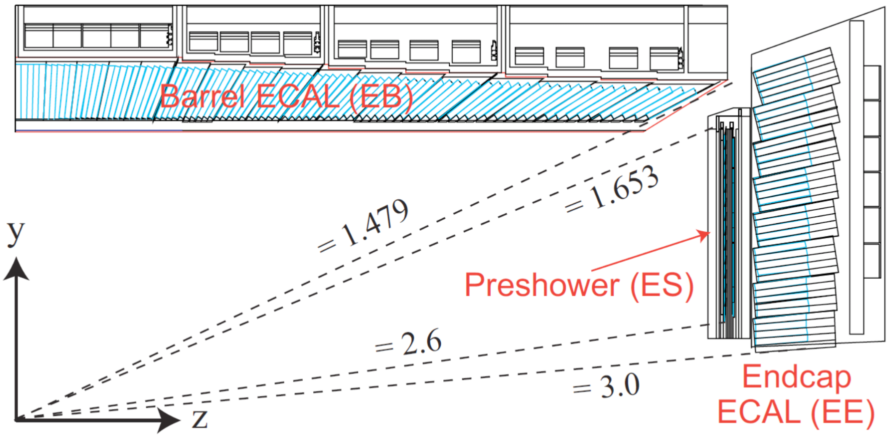


FIGURE 3.5: Sectional view of a quarter of the ECAL. The figure is taken from [69].

$PbWO_4$ was chosen because of its short radiation length ($X_0 = 0.89$ cm) and Molière radius (2.2 cm) and also the crystals are very fast (80 % of the light is emitted within 25 ns) and radiation hard. The crystals are read out by photodiodes. These photodetectors have intrinsic gain and can operate in a magnetic field, which is required by the ECAL's location inside the solenoid magnet and because of the relatively low light yield of the $PbWO_4$ crystals. The size of individual crystals is 22×22 mm^2 in the barrel region and 28.6×28.6 mm^2 in the end caps with respective lengths of 230 mm and 220 mm (~ 25 X_0) [71, 72].

The end cap region also features an additional pre-shower detector in front of the crystal calorimeter. These are 2 layers of silicon strip sensors with a pitch of 1.9 mm, which are mounted behind discs of lead at depths of 2 X_0 and 3 X_0.

The energy resolution as a function of incident particle energy of the calorimeter is given as:

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

with S denoting the stochastic term, N is the noise and C is the constant contribution.

3.2.4 The Hadronic Calorimeter (HCAL)

HCAL, like the ECAL, is mostly situated inside CMS's magnet coil, outside the ECAL, and therefore the properties of the magnet system govern the design of HCAL. The requirements are good containment and hermeticity for missing transverse energy $\vec{\not\!\!E}_T^{miss}$ measurement, which requires a maximum amount of absorber material and thus interaction lengths. For that reason it has been designed as a so-called sampling calorimeter with alternating absorber scintillating material. The CMS HCAL consists of four major sections as shown in Figure 3.6: the HCAL Barrel (HB), the HCAL Outer (HO), the HCAL Endcap (HE) and the HCAL Forward (HF).

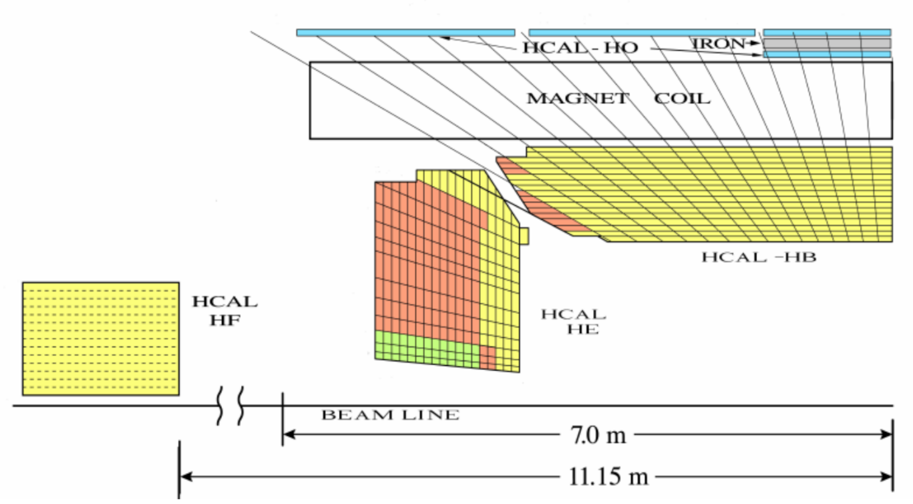


Figure 3.6: Sectional view of a quarter of HCAL. The figure is taken from [73].

The central hadron calorimeter, consisting of a barrel (HB) and two endcap (HE) regions, is located directly after the ECAL and extends up to the solenoid. The HB covers a range of $|\eta| < 1.3$. It consists of brass absorber plates in 14 layers oriented along the z-axis with a thickness of 50.5 mm (first eight) and 56.5 mm (last six), respectively. For structural support two additional layers of 40 mm and 50.5 mm thick steel absorber plates are installed at its inner and outer rim respectively. Between the absorbers 72 azimuthal wedges of plastic scintillators in 17 layers with a thickness of 3.7 mm or 9 mm are installed where each covers a segment of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$. Their emitted light is optically added per tower and guided through wavelength shifting (WLS) fibres to hybrid photodiodes (HPDs) located at the end of the HB structure. The HB has a depth of 5.8 interaction lengths at $\eta = 0$ which increases with pseudorapidity and amounts to 10.4 interaction lengths at $|\eta| = 1.3$.

The same design of alternating brass absorbers (79 mm thick) with 17 scintillator layers in between and WLS fibres for readout is utilized in the HE. It covers a range of $1.3 < |\eta| < 3$ and has a depths of about 10 interaction lengths. Its granularity increases from $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ to $\Delta\eta \times \Delta\phi = 0.17 \times 0.17$ for $|\eta| > 1.6$.

To contain hadronic showers in the barrel region further, HO is placed directly at the outside of the solenoid utilizing it as absorber. It consists of one or two layers of scintillators, depending on the pseudorapidity, which match the granularity of the HB. This extends the depths of the combined HB+HE+HO system to an overall minimum of 11.8 interaction length with the exception of the HB-HE transition region ($|\eta| \sim 1.3$).

An additional Hadron Forward (HF) calorimeter is located more than 11 meters away from the interaction point in deep-forward direction. It is built from steel and quartz fibres, which results in narrower and shorter hadronic showers (a requirement in the forward region). The signal creation mechanism is Cherenkov radiation.

Neutral hadronic interactions are the most important consideration for the energy resolution of the HCAL as, unlike charged particles, no additional information can be obtained from the tracker to combine with the calorimetry measurement to improve the resolution. The energy resolution, based on the measurement in pion test beam in GeV range, is parametrized as:

$$\frac{\sigma_E}{E} = \frac{110\%}{\sqrt{E}} \otimes 9\%. \tag{3.6}$$

3.2.5 The CMS Muon System

It is one of the key components of CMS, since high-energetic muons are important signatures for interesting events. Muons, apart from neutrinos, are the only particles that can traverse the whole detector without being stopped and therefore are very important for triggering and event selection. Good muon identification and momentum resolution are key aspects of the detector. This is accomplished by linking the tracker, where muons are measured for the first time, and muon system information. The momentum resolution for the lower range is governed by the tracker but the muon chambers are crucial for larger momenta.

The muon system uses three type of gaseous detectors adapted to the different radiation environments where they are installed, interleaved inside the hollow parts of the magnet return yoke. In the barrel region drift-tube (DT) chambers are used because the contributions from neutron background, residual magnetic field and muon flux are small.

The end-cap regions, in contrast, use cathode strip chambers (SCS) because of higher particle rates and background. Additional resistive plate chambers (RPC), that operate in avalanche mode, are used all around the detector as they provide good time resolution for event selection and perform well at high rates (Figure 3.7). The combined information of all three systems provides very fast and robust trigger information [74, 75, 76].

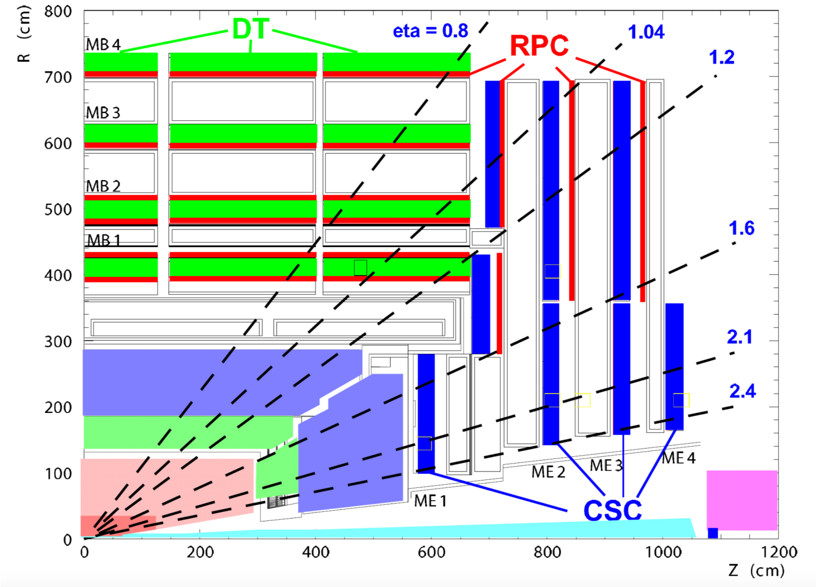


FIGURE 3.7: The layout of the CMS muon system with DTs, CSCs and RPCs [75].

3.2.6 CMS Data Acquisition System (DAQ)

The CMS DAQ [77, 78] deals with the individual readouts of the various subdetector front-end systems. It associates them to an event and finally transmits a file of multiple events to the CERN computing cluster for storage. The CMS trigger system works in two sequential stages: the Level-1 (L1) Trigger and the High-Level Trigger (HLT). A two-staged event triggering system is employed to lower the high data rate and to select only events for storage of certain physics interest.

- L1 Trigger: The L1 trigger system analyzes only the readout of the calorimeter and muon systems per bunch crossing to reach a decision within a maximum latency of 3.2 μs. The L1 Trigger uses ECAL, HCAL and muon data for the decision. The tracker detector is not part of the trigger due to high bandwidth requirements. For safe operation, the Global Trigger electronics are located away (on average $\sim$ 90 m) from the detector, in the counting room adjacent to the experimental underground cavern.
- HLT Trigger: A multi-stage iterative algorithm uses the complete detector information to reproduce the L1 Trigger decision. The algorithm iteratively improves

on the L1 decision by the staged introduction of fine-grained calorimetry and tracking information. The HLT reduces the data rate to approximately 100 Hz to store the data for further physics analyses.

3.2.7 Computing at CMS

The selected data by the above mentioned triggers need to be stored and transferred to research institutes for further processing. This is achieved via the Worldwide Large Hadron Collider Computing Grid (WLCG) [79]. It combines the computing resources of many of the institutes participating to LHC experiments around the world. WLCG consists of three tier-levels that are distributed among worldwide centres as shown in Figure 3.8.

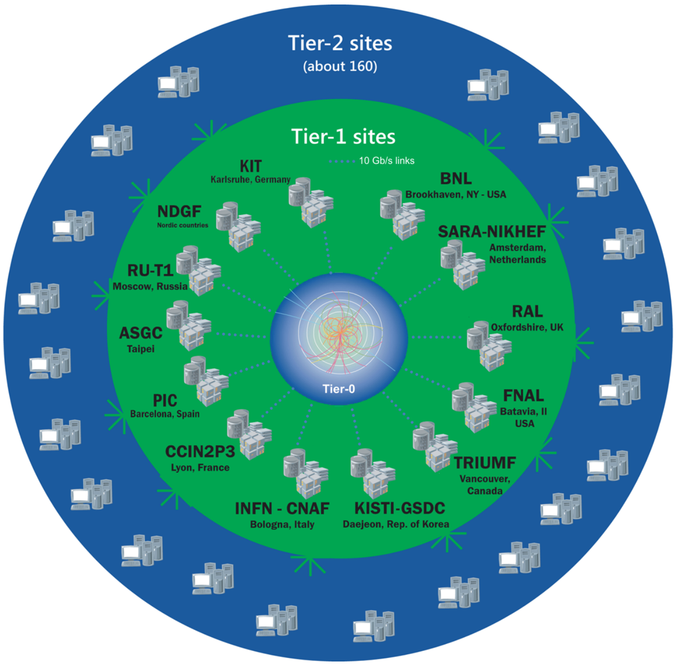


Figure 3.8: A view of WLCG layers, or "tiers"; 0, 1 and 2 [80].

The data collected by the experiments are processed as well as stored at Tier-0. Then the data is distributed to 14 Tier-1 and around 150 Tier-2 computer centres. These centres take care of the data reprocessing and storage. They are also used for the production of simulated data. Physicists access and analyze the data from Tier-2 centres.

For CMS experiment the Tier-0 is located at CERN. The data is then distributed to fourteen other Tier-1 centres. Thereafter, it is further distributed from Tier-1 centres to over 150 Tier-2 centres.

Chapter 4

Event Reconstruction

Many models in particle physics are developed based on fundamental hypotheses. In order to confront any theoretical model with empirical observations its induced phenomena should be predicted in a form that can be compared with experimentally observed features. This is the function of simulation programs. The Monte Carlo (MC) event simulation at colliders is considered as a transformer of the theoretical model into a form that can be analyzed like real data collisions.

In this chapter we will explain how the collision event looks like at the LHC and how the final state particles are reconstructed by interaction with CMS detectors materials (described in Chapter 3). Secondly, the major physics objects used in the analysis will be described. These physics objects are selected by using recommendations of the CMS physics object groups (POG). Details about the reconstruction and identification of muons and electrons will be discussed in Section 4.6 and 4.7 respectively. The charged and neutral hadrons are clustered into jets as explained in Section 4.8. Finally, the missing energy reconstruction will be summarized in Section 4.9.

4.1 Event Generation

To compare reconstructed data with theoretical predictions, samples of simulated collision events are generated and run through a simulation of the CMS detector and an emulation of its readout. For the analyses, the standard so-called "Full Simulation" package [81, 82] is employed for performing the detector simulation. It is based on the Geant4 toolkit [83] that performs a simulation for particles interaction with the detector. A faster alternative, the so-called "Fast Simulation" package [84] exists within CMS as well, but it has not been utilized within this thesis.

The generation of a $\mathrm{pp} \to X$ collision event means generating the whole process starting from the incoming protons at the LHC to the experimentally observable final states X. These final states are detected with different kinds of detectors. The process of event generation takes place through the following sequential steps or subprocesses [85].

- Factorization theorem
- Matrix elements

- Parton Shower
- Hadronization
- Underlying event
- Pile-up

Figure 4.1 illustrates these steps.

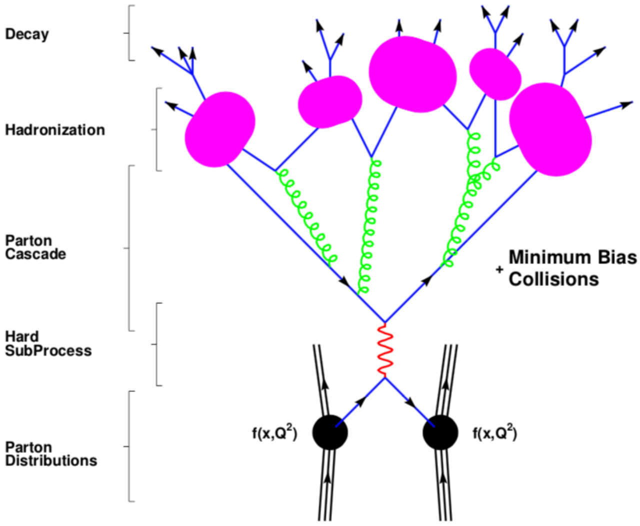


FIGURE 4.1: A simplified representation of the structure of the event generation chain for hadron-hadron collisions [86].

4.1.1 Factorization Theorem

The full proton-proton collision can be divided into various steps. The partons, carrying a fractional momentum of the total proton energy, collide and produce a final state X as shown in Figure 4.2. The kinematical properties of the final state X can be predicted by performing perturbative calculation to the parton collision. The partonic cross section of the final state X can be factorized in terms of factorization scale (μ_F) and renormalization scale (μ_R) as $\hat{\sigma}_{pp\to X}(\mu_F, \mu_R)$. The μ_F takes short and long distance effects into account and a fixed order perturbation theory calculations bring about a dependence on μ_R. The interaction between two hadrons at a collider can be given by factorization theorem:

$$\sigma_{pp\to X} = \sum_{a,b=q,\bar{q},g} \int_0^1 dx_a dx_b f_a(x_a, \mu_F) f_b(x_b, \mu_F) \times \hat{\sigma}_{pp\to X}(\mu_F, \mu_R) \tag{4.1}$$

where the sum is over the possible initial parton flavors. The PDFs, $f_i(x_i, \mu_F)$, determine the probability of finding a parton with fraction x_i of the proton momentum when

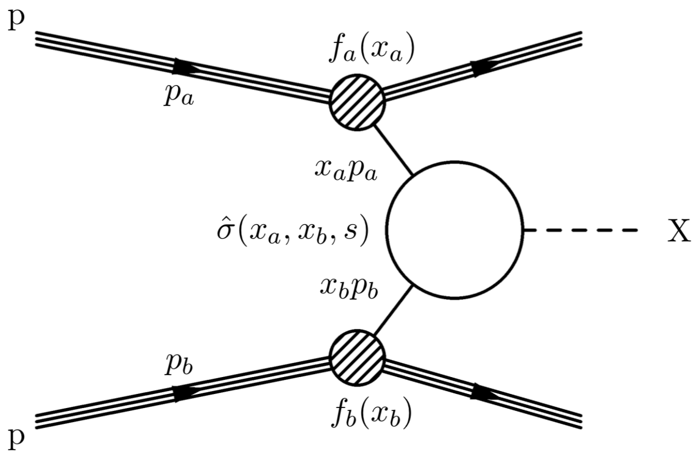


FIGURE 4.2: Diagram of a generic collision of two proton. The partons inside a proton collide and produce the final state X [87].

it is probed at an energy scale μ_F. Perturbative QCD does not work below energy scales of about 1 GeV because at low energy the strong coupling constant α_s becomes larger than one. PDFs are determined experimentally from the data of deep-inelastic collision experiments because they contain a non-perturbative part that can not be determined from first principles.

4.1.2 Matrix Element (ME)

The hadronic cross section as given by equation 4.1 is a function of the differential partonic cross section or the cross section of the hard interaction $\sigma_{ij} \rightarrow X$ between the partons. The differential cross section of this process for specific phase space is given by:

$$d\hat{\sigma}_{ij\rightarrow X} = \frac{1}{2s_{real}} \frac{|\mathcal{M}^2_{ij\rightarrow X}|}{8(2\pi^2)^2} dps \tag{4.2}$$

where s_{real} is real CoM energy, and $\mathcal{M}$ is the ME of the process. $\mathcal{M}$ is calculated from the Lagrangian density of the quantum field theory under consideration. Also the hadronic cross section $d\hat{\sigma}_{ij\rightarrow X}$ can be expanded in the strong coupling constant α_s. A perturbative calculation of the process is possible with sufficiently high energy scale of the interaction or equivalently if α_s is small enough. Hence, the $\mathcal{M}$ can be calculated up to a certain perturbative order in α_s.

The quantum fluctuations can cause ultraviolet divergences at high energies. These divergences can be dealt by choosing a certain scale μ_R (renormalization scale). Typically μ_R is used to redefine physical quantities in the way which makes the theory able to describe the experimental regime. Consequently, coupling constants will run as a function of μ_R and beyond μ_R the divergences due to the high energy corrections, like loop corrections to the propagators are absorbed in the physical quantities through a renormalization of the field. From equation 1.36, it can be noticed that the coupling strength inversely proportional to the renormalization scale, which is called asymptotic freedom. Moreover, by following the trend of $\alpha_s(\mu_R)$ a limit $\Lambda_{QCD} \approx 200$ MeV was found for for which α_s becomes larger than one. Additionally infrared (soft gluons) and collinear divergences (gluons collinear to the partons) are also encountered in the calculation of the matrix element.

4.1.3 Parton Shower

Parton shower (PS) program simulates the hadronization of the partons in final state. It also simulates radiation of soft gluons or quarks from initial or final state partons. Furthermore, contributions from soft secondary interactions, the so-called underlying event, and color reconnection effects are considered in PS simulation as well. A sketch of an exemplary pp collision event after hadronization is shown in Figure 4.3, where various parts of the event simulation are highlighted.

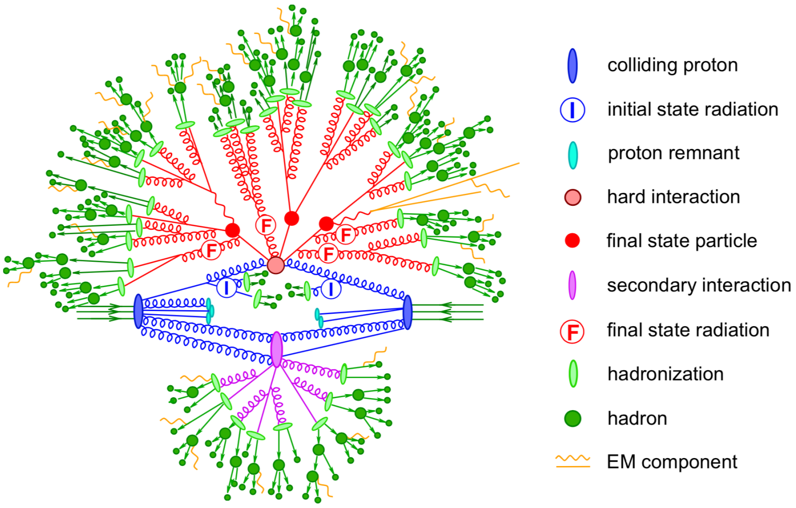


FIGURE 4.3: A sketch of a generated event from the simulation of the hard interaction and subsequent hadronization through a parton shower. The figure is taken from Ref. [88].

PS simulations are based on Altarelli-Parisi splitting functions [89] which allow to calculate the probability that a soft parton splits into two others as q$\rightarrow$ gq, g$\rightarrow$ q$\bar{q}$, or g$\rightarrow$ gg. It is convenient to calculate the "survival" probability, the so-called Sudakov factor, that a parton does not split between two energy scales. The emission of soft partons leads to a complication during the PS simulation since such emissions could be double-counted if the upstream simulation of the hard interaction may have produced a similar soft emission already. The problem of double-counting occurs due to the fact that the phase spaces, where an event generator and a subsequent PS program simulate their products in, are overlapping. This is avoided by applying a dedicated ME-to-PS matching scheme which consists of a method to assign additional emissions exclusively to either the ME or the PS simulation depending on the event kinematics. Thus both simulation phase spaces become orthogonal to each other. Further information on parton shower simulation and matching algorithms can be sought in Refs. [84, 88].

4.1.4 Hadronization

Colored partons can not be seen directly as they exist in bound states only. The transformation of these colored partons into bound states of colorless hadrons is called hadronization. Unstable hadrons will further decay at various timescales until stable or sufficiently long-lived particles are created. Currently there are mainly two phenomenological approaches for hadronizarion process. One based on string fragmentation while the other is based on the cluster model. In PYTHIA, the hadronization is simulated using the Lund string model [90, 91] based on the idea of a physical picture of a colour flux tube "string" connecting a quark q and an anti-quark $\bar{q}$. As the two quarks move apart, the string gets stretched and potential energy $V(r) = kr$ builds up, where r is the distance between q and $\bar{q}$. The string constant k corresponds to the amount of energy per unit length and is estimated to be ≈ 1 GeV/fm. When $V(r)$ becomes high enough a new pair of quarks $q'\bar{q'}$ is produced with a mass m. In this way two new color singlets $q\bar{q'}$ and $q'\bar{q}$ are formed. Further breaking might occur provided that the invariant mass of the new color singlets is high enough [92]. The splitting process continues until only color-neutral hadrons with an on-shell mass remain. Light quarks (u, d and s) are included in this model while heavy quarks (c, b) are ignored as the probability to create a heavy quark pair is heavily suppressed. Gluons in the string model are set as small snags on the string between two quarks. The Lund string model is infrared and collinear safe [93].

4.1.5 Underlying Events (UE)

The two colliding partons in general carries only a small fraction of the total proton energy. Most of the proton energy remains in the beam and continues to travel in the beam direction. These remnants also carry color charges that balance the color charges taken away by the colliding partons. More than one parton from each proton can collide in a pp collision, resulting in multi-parton interaction (MPI). The hadronization of the remaining beam partons and the multi-parton interaction process are known as the underlying event (UE). A more detailed description of the UE is given in [94].

4.1.6 Pileup

In addition, another process, the so-called pileup has to be considered as well. Pileup events occur at high instantaneous luminosities when more than just two protons collide during the same bunch crossing (in time PU). For LHC design luminosity there can be up to 25 pileup events at the same time. MC event simulation takes (in time) PU effects into account.

4.2 MC Generators

There are many MC event generators used to generate full collision steps as described in previous sections. These steps can be considered sequentially in the MC methods. Starting with simple rules at each step, more and more complex states can be constructed by using iterations. This ends with hundreds of particles moving out from the interaction point in different directions. The main event generators used for the production of simulated samples used in the physics analysis in this thesis are the following:

- **MadGraph:** MadGraph is a LO event generator that calculates the $\mathcal{M}$ from tree-level Feynman diagrams. Additional partons are included to get an approximate description of ISR and FSR. It takes PDF sets as input which describe the kinematics of the incoming partons. Based on these generated ME, the hard process events are generated by MadEvent to calculate LO cross sections.

  The **MadGraph5_aMC@NLO** [95] is a merge of the LO MadGraph generator [96] and the aMC@NLO program (Ref. [97]) into a common framework. It supports the generation of samples at LO or NLO accuracy together with a dedicated matching to parton showers using the MLM [98] or MC@NLO [99] schemes respectively. The latter matching scheme produces a certain fraction of events with negative weights (depending on the process) which originate from a subtraction of amplitudes containing additional emissions from the NLO ME to prevent double counting. The MadGraph5_aMC@NLO framework is also capable of producing multiple samples with additional final state partons at ME level that can be merged into a combined sample. The overlap with the PS simulation between the exclusive samples is removed through the MLM [100] or FxFx [101] merging schemes.

- **PYTHIA:** The PYTHIA [102, 103] program produces particularly good agreement with data compared to other PS generators. It takes the parton-level event generated by ME generators to produce the parton shower (PS) after adding soft emissions from ISR and FSR state particles. PYTHIA also performs the fragmentation and hadronization of quarks and gluons.

- **Powheg:** The Powheg box (versions 1,2) [104] is a program that contains predefined implementations of various processes such as t-channel single top production at NLO [105]. It applies the so-called Powheg method [106] for ME-to-PS matching in which the hardest radiation generated from the ME has priority over subsequent PS emissions to remove the overlap with the PS simulation. A small fraction of negatively weighted events may be generated as well in phase space regions where NLO calculations are not feasible.

- **Herwig++:** The Herwig program [107, 108] is an NLO event generator that can also perform a standalone PS simulation which can be interfaced with various other event generators as well. Its hadronization algorithm is based on a model in which color-connected quarks are spatially kept together in clusters [109]. This is motivated by the concept of "pre-confinement" for colored particles [110]. If the mass of a cluster is sufficiently high it can decay into lighter clusters with a certain probability. In the final simulation step, a cluster then decays into hadrons according to its quark content.
- **CompHEP:** The CompHEP program (version 4.5) [111] can perform calculations of cross sections from Lagrangian densities at LO. In addition, generation of events such as single top quark production is also possible [112]. Here, an approximation is used by combining events from the 2 → 2 and 2→3 processes which reproduces NLO corrections in an effective way.
- **Tauola:** Event generators can be interfaced with the Tauola library [113, 114] which is specialized for simulating leptonic and hadronic decays of tau leptons with high accuracy. It accounts for spin polarization effects while the radiation of photons from QED corrections is also included by incorporating the PHOTOS library [115].

4.3 Detector Simulation

After the parton shower and hadronization, the particles interactions with the detector material is simulated. For a detailed detector simulation, GEANT4 is used. This toolkit includes the active and the dead material regions. The active material region detects the traversing particles with the sensitive layers of the detector and the dead region consists of support structures and cables between detector components. GEANT4 also incorporates magnetic field effects needed to simulate charged particles curvature in the presence of B. The particles interaction with the detector material, detector response and its conversion into electric signals are also simulated. Based on well understood underlying interaction mechanisms, the energy losses of the traversing particles and their trajectories are simulated as well. The simulated effects include bremsstrahlung and photon conversions to $e^{-}e^{+}$. The showering of electrons, photons and hadrons in the calorimeters are simulated as well.

4.4 Physics Object Reconstruction

The event reconstruction attempts to build and identify basic analysis objects from the raw detector data. In CMS, basic objects are charged particle tracks, vertices, charged leptons, photons, and jet candidates. During the reconstruction, additional information such as $\vec{\not{E}}_T^{miss}$ and b-tagging is determined. Since tau leptons have not been utilized and photons are not relevant in the presented studies of t-channel single top production, a description of their reconstruction and performance is omitted here yet details can be found in Refs. [116, 117]. Although tau leptons are not used explicitly, but their decays to muons and electrons are included.

A key ingredient in the event reconstruction of CMS is the particle flow (PF) algorithm. It creates particle candidates by combining various subdetector information for a global event interpretation which improves the identification, spatial resolution, and energy measurement of particles.

4.5 The Particle Flow (PF) Algorithm

The function of PF algorithm is to collect and combine all information from the CMS subsystems. The algorithm uses this information to reconstruct the whole event by identifying all final state particles traversing the detector as shown in Figure 4.4. The combined information gives more precise measurements of particle momenta compared to using information from a single sub-detector.

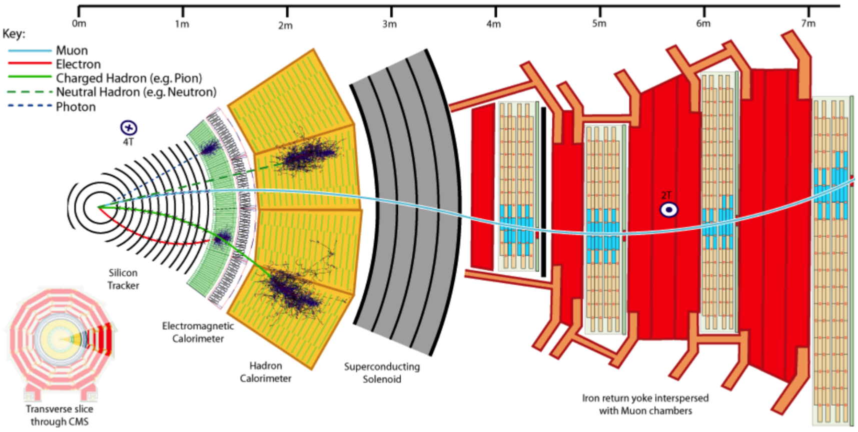


FIGURE 4.4: The sketch presents the behaviour of different particle interactions in a transverse slice of the CMS detector. Charged particles have tracks in the tracker and deposits in calorimeters or muon chambers. Neutral hadrons have no tracks in the tracker but leave deposits in the calorimeters. Photons leave deposits only in the ECAL. The figure is taken from Ref. [118]

The PF algorithm links the elements, i.e. the charged particles tracks and the clusters from calorimeters, into "blocks". Then the algorithm interprets them in terms of muons, electrons, photons and charged and neutral hadrons. The PF algorithm reconstructs objects in a well defined order starting from the muons. Once a block is found, the corresponding elements are removed from the event. Subsequently the next link is sought for, and so on until no more elements are left. From these blocks, PF candidates are identified in the following order:

- **Muons:** PF muons are the first objects to be reconstructed. The muon is identified by its deposits in the muon chambers and the corresponding track in the tracker. Once a muon is identified, its track and deposit is eliminated from the respective collection. The details about muon reconstruction is discussed in Section 4.6.
- **Electrons:** In CMS, an electron radiates more than 70% of its energy with a probability of 35% in the inner tracker through bremsstrahlung before reaching the ECAL. This leads to an increasingly curved electron trajectory in B field as a function of its flight distance. The standard tracking is suboptimal for reconstructing such trajectories because the employed Kalman filtering assumes that the energy loss is Gaussian- distributed. Therefore, a different filtering algorithm the so-called Gaussian sum filter (GSF) [119] is used in the electron tracking reconstruction instead. Further details about electron reconstruction is discussed in Section 4.7.
- **Charged Hadrons:** After electron and muon reconstruction, the remaining tracks in the event are due to PF charged hadrons. These tracks are linked to energy deposits in calorimeters. The energy of hadrons is determined by taking information from calorimeters.
- **Photons and Neutral Hadrons:** Neutral hadrons and photons are reconstructed in the end. PF photons are identified as they leave no tracks in the tracker but have energy deposits in the ECAL. The remaining HCAL deposits are determined to be as PF neutral hadrons.

After the reconstruction of all PF objects in the event, PF jets are identified as discussed in Section 4.8. Missing transverse energy $\vec{\not{E}}_T^{miss}$, has no deposits or tracks in the detector, therefor it is reconstructed in the last and is discussed in Section 4.9

4.5.1 Track Reconstruction

Charged particles create hits in the silicon sensors arranged in the cylindrical layers of the trackers. The trajectory of the charged particle or its track can be reconstructed from these hits. The CMS uses a software called Combination Track Finder (CTF) for track reconstruction. CTF is used for pattern recognition and track fitting and is an extension of the Kalman filter [120, 121, 122]. CTF reconstructs tracks with multiple iterative steps:

- **Seeding:** The first step is seed generation that gives an estimate of trajectory parameters and associated uncertainties. The initial track candidates are identified by only a few (2 or 3) 3D hits on the pixel layers. The low channel occupancy of the pixel layers results in less ambiguity and a higher efficiency for close-by tracks [123]. Seed candidates have to fulfil certain quality criteria like a minimal transverse momentum and compatibility with either the beam spot or a preliminary reconstructed vertex depending on the iteration before they are passed to the trajectory finding stage.

- **Track-Finding Recognition:** CTF algorithm extrapolates the estimated trajectory to find additional hits on subsequent layers that are compatible with the initial track candidates. The algorithm then updates the track parameters and their uncertainties iteratively after adding a hit to the trajectory candidate. If multiple compatible hits are found on a layer, the trajectory is cloned for each of them. If there is no compatible hit in a layer, a ghost hit is created instead.
- **Track-Fitting:** The hits per trajectory are then passed to a KF-based helix fit to estimate the track parameters without utilizing the initial estimate from the seed. In addition, the fit accounts for material effects and the inhomogeneous magnetic field. The fitted tracks have to pass a quality selection to reduce the amount of fake tracks before they are considered in physics analyses.
- **Track-Selection:** This module defines criteria for selecting or rejecting tracks. The criteria reflects the seed requirements and depends additionally on the total number of fitted 2D/3D hits, the χ^2/ndof of the fit, the amount of ghost hits and the amount of shared hits with other tracks amongst others. Track p_T and η dependant requirements are also applied for better performance.

There are fake tracks in the tracker because of some random particle hit. Any charged particles that have deposits in ECAL and HCAL but somehow missed track in the tracker, will be considered as a neutral particle. So the challenge of PF event reconstruction is to keep the fake tracks reconstruction minimum while keeping the track reconstruction efficiency to the maximum.

4.5.2 Primary Vertex (PV) Reconstruction

As discussed in Chapter 3 many interactions happen at each bunch crossing. As a result a large number of vertices is to be reconstructed along the z-direction as illustrated in Figure 4.5. PV reconstruction is used to identify the location of all pp interaction vertices in the event. This is done by using the available reconstructed tracks. It is important to mention that these reconstructed vertices include the signal vertex and pileup collision vertices. The reconstruction consists of three steps:

- Selecting tracks that are consistent with the ones formed promptly in the primary interaction region.
- Clustering tracks which seems to appear from the same interaction point based on their z-coordinates.
- Each vertex and its associated tracks are fitted. The PV reconstruction depends upon the tracks multiplicity used in the fit and the p_T of the tracks.

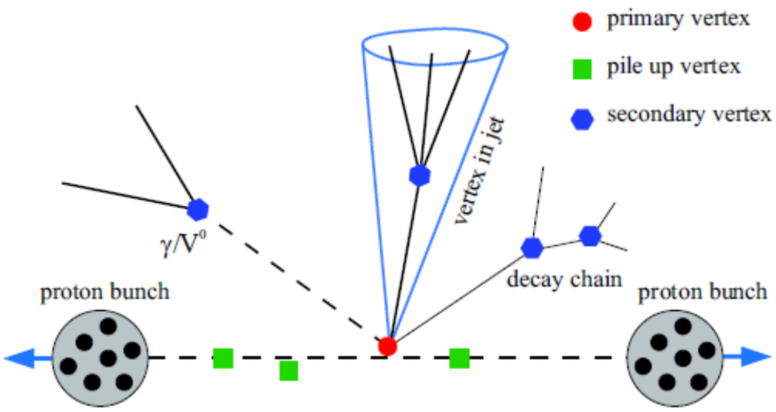


FIGURE 4.5: Schematic diagram showing the different types of reconstructed vertices. This figure is taken from the CERN website [88].

The tracks p_T in a vertex is squared and then summed. PV vertices are then graded with the value of the squared tracks momentum sum. The vertex with the higher sum is identified as signal vertex.

4.6 Muon Reconstruction

Identification of muon candidates for physics analyses is performed by requiring additional selection criteria. A detailed study of muon identification with 7 TeV pp collision data can be found in Ref. [124]. Throughout the analysis within this thesis, muon candidates have to fulfil identification criteria which correspond to a "tight" working point yielding most genuine muons while rejecting falsely reconstructed ones. In the following, the criteria and performance of muon identification employed in the analysis of 8 TeV and 13 TeV pp collision data within this thesis are briefly discussed. Detailed reports on their performance can be found in Ref. [125, 126].

The muons are reconstructed by following a set of selections based on the global and tracker muon properties. In Run II data "tight", "medium" and "loose" ID working points (WP) are used to define three different identification criteria for the isolated muons reconstruction. One chooses a working point depending on the physics analysis envisaged. For each working point different requirements (i.e. cuts) are applied on variables such as those summarized in Table 4.1. The muons are required to be global muons, with small normalized χ^2 (i.e. χ^2/ndof). The transverse impact parameter (IP) d_0 in the $\phi - plane$ and the longitudinal IP d_z of the muon with respect to the leading PV should be small to ensure a good rejection of cosmic muons and those originating from additional proton collisions (pileup) in the event. To ensure that the muon candidate is well separated from a jet, a requirement $\Delta R_{\mu,jet} > 0.4$ is added [127].

Compared to "medium" or "loose" muons, "tight" muons have strict requirements on different quantities including their transverse momentum because as the muon momentum decreases, it becomes harder to distinguish it from other particles. These IDs are in general used when defining selection requirements on the muons. Tight ID is used for identifying muons from the signal process. Loose muons are in general used to reject additional objects which may be muons but could be a misidentified object such as a pion. The loose criteria will capture more real muons in its selection but with a lower purity. Medium muons have the same criteria as loose muons with additional track quality and muon quality requirements. Tight WP is used for both of the analyses in this thesis.

Muons are also required to be spatially isolated from EM and hadronic activity in addition to the tight identification criteria. The relative $\Delta\beta$ based ("delta-beta") isolation for muons is defined as

$$I_{\Delta\beta} = \frac{I_{charge-hadron} + max(I_{neutral-hadron} + I_{\gamma} - \beta \cdot I_{PU}, 0.)}{p_T}, \quad \beta = 1/2, \tag{4.3}$$

where $I_{charge-hadron}$, I_{γ}, and $I_{neutral-hadron}$ denote the summed transverse energies of charged hadrons, photons, and neutral hadrons respectively within a cone of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\Phi)^2} < 0.4$ around a muon candidate. The term I_{PU} is used to correct the amount of considered neutral energy. It denotes the summed transverse energies of charged-particle tracks that are associated to pileup vertices. Hence, the applied correction $\beta \cdot I_{PU}$ can be interpreted as an estimate of the amount of neutral energy from PU interactions within I_{γ} and $I_{neutral-hadron.}$. The chosen value for β is motivated by assuming equal production rates for the (π^+, π^0, π^-) isospin triplet leading to a ratio of 1/2 for the production of neutral pions over charged ones.

Table 4.1: The table shows the muon criteria for each working point. The table is taken from Ref. [126].

Requirements	Tight WP	Medium WP	Loose WP
$PFmuonid$	yes	yes	yes
$Whichmuon$	Global	Global or tracker	Global or tracker
χ^2/ndof of track	< 10	<3	-
RelIso	< 0.15	< 0.25	<0.25
$d_0(cm)$	< 0.2	-	-
$d_z(cm)$	< 0.2	-	-
No. of pixel hits	> 0	-	-
No. of trackers with muon hits	> 5	-	-
No. of hits in muon chamber	> 0	-	-
No. of muon stations	> 1	-	-
$\Delta R_{(\mu,jet)}$	> 0.4	< 0.4	< 0.4

4.7 Electron Reconstruction

Similar to muons, electrons require certain identification criteria as well. Studies of electron reconstruction and identification at 8 TeV and 13 TeV CoM can be found in Refs. [128, 129]. The "tight" identification criteria, employed in this thesis, is elaborated briefly in the following.

A PF electron candidate with a GSF track is required. Candidates within the ECAL region of $1.4442 < |\eta| < 1.5660$ are ignored because of barrel and endcap transition. The electron track has to have a hit on the innermost tracker layer which prevents the selection of electrons from potential photon conversions ($\gamma \rightarrow e^+e^-$). An explicit photon conversion veto is applied by testing if a pair of electron tracks originates from a common displaced vertex. Further selection criteria are combined into a multivariate identification discriminant. It is based on various input observables like the GSF track quality, the ECAL cluster shapes, their energy distribution and the agreement between independent cluster energy and track energy estimates. For 13 TeV data, the discriminant is replaced by a simplified cut-based version where multiple fined-tuned selections on similar observables are applied. A comparison of the efficiency of electron identification in 8 TeV data and simulation is estimated from $Z \rightarrow e^+e^-$ events with Tag and Probe method. For electron $p_T > 30$ GeV, the identification efficiencies saturates around $\sim 95\%$.

An electron candidate is should be isolated from other EM or hadronic activity in its vicinity. The relative A^{ρ}_{eff}-based ("effective area") isolation for electrons is defined as

$$I^e_{rel} = \frac{I_{charge-hadron} + max(I_{neutral-hadron} + I_{photon} - \rho \cdot A_{effe.}(\eta_{SC}), 0.)}{p^e_T} \tag{4.4}$$

where the transverse energies I_X per particle type X are summed in a cone of $\Delta R < 0.3$ around the electron candidate. The amount of neutral energy is corrected by the effective area A_{eff} times the median of the transverse energy density ρ calculated in $\delta\eta \times \delta\phi$ from charged particle tracks that are associated to PU vertices. The effective area is estimated from simulation and denotes the expected amount of neutral energy from PU interactions per ρ within the isolation cone as a function of η of the associated ECAL supercluster. The general idea behind the A^{ρ}_{eff}- based isolation for electrons is motivated by a proposed pileup subtraction method for jets which is detailed in Ref. [130].

4.8 Jet Reconstruction

Jets are the experimental equivalent of energetic partons produced hadron collision. When partons hadronize as discussed in Section 4.1.4, they form a number of charged and neutral hadrons travelling approximately in the original parton direction. These final state particles can be clustered into what is known as a jet. This is achieved by using various reconstruction algorithms. All jet-clustering algorithms look for calorimeter deposits and cluster them together in jets. A jet serves then as the high-level physics object representation of the parton in the final event analysis.

Any jet-clustering algorithm needs to be insensitive to the additional soft radiation (IR safety) and to the collinear splitting of a hard parton (collinear safety). CMS collaboration uses the so-called anti-kT algorithm [131] which satisfies both of these requirements. The anti-kT algorithm starts with a high p_T deposit in the calorimeter and considers it as a seed. Then it uses the distance definition of Equation 4.5 to find the nearest deposit to merge with.

$$d_{ij} = min(k_{T,i}^{2p}, k_{T,j}^{2p})\frac{\Delta_{ij}^2}{R^2}, \tag{4.5}$$

$$d_{i,B} = k_{T,i}^{2p}, \tag{4.6}$$

where

d_{ij} = distance between particle i and particle j, and

$d_{i,B}$ = distance between particle i and the beam.

The values $k_{T,i}$ and $k_{T,j}$ represent the transverse momenta of particles i, j respectively, $\Delta_{ij}^2 = (yi - yj)^2 + (\phi_i - \phi_j)^2$, R is the radius and p is a parameter that govern the transverse momenta relative power with respect to the geometrical Δ_{ij}^2 scale. If d_{iB} is smaller than d_{ij}, the particle is merged with the beam, otherwise the particle is merged with the nearest hit, according to the value of d_{ij}. The algorithm terminates when all particles i have $d_{iB} < d_{ij}$ and the merged particles are considered to be a jet [132].

In the PF event reconstruction for CMS Run II, a cone of radius 0.4 is used for jet reconstruction. For the anti-kT algorithm, the parameter p assumes a value of -1. This choice for $p =$ -1 ensures soft particles are more likely to be associated to a close by hard particle. Using tracker information additional to the calorimeter information improves the jet momentum and spatial resolution. PF jets are reconstructed by applying certain selection cuts to distinguish real jets from fake and noisy jets. These selection cuts are referred as PF Jet ID. Two set of IDs (tight and loose) are defined as described in Table 4.2. These cuts are recommended by the JetMET POG [133] at CMS.

In clustering hadrons into jets, different sources can affect the jet energy scales, so some corrections are applied. Furthermore the jet energy resolution (JER) in simulated samples is different from that in real data, so an additional calibration has to be done for the simulated jet energy resolution.

When comparing energies of the generated jet to the reconstructed jet, it is observed that they are not identical. A fraction of the difference between them is due to the

TABLE 4.2: The table shows the two Jet-Id criteria. The cuts are applied for jets at $|\eta| < 2.4$ while for jets within $2.4 < |\eta| < 2.7$ only the first 3 cuts are applied.

Variable	Tight ID	Loose ID
$NeutralHadronFraction$	< 0.9	< 0.99
$NeutralEMFraction$	< 0.9	< 0.99
$NumberofConstituents$	> 1	> 1
$ChargedHadronFraction$	> 0	> 0
$ChargedMultiplicity$	> 0	> 0
$NeutralEMFraction$	< 0.99	< 0.99

non-linearity of the detector response and the presence of pileup particles. Consequently reconstructed jets are calibrated in order to have the correct energy scale (i.e. reconstructed jet energy matches the generated jet energy). This is done by jet energy corrections (JEC).

In CMS, a factorized approach is used to apply a set of corrections sequentially and with a fixed order. As can be seen in Figure 4.6, at each correction level, the jets four-momentum vector is scaled with a SF. The corrections can be summarized as follows [134]:

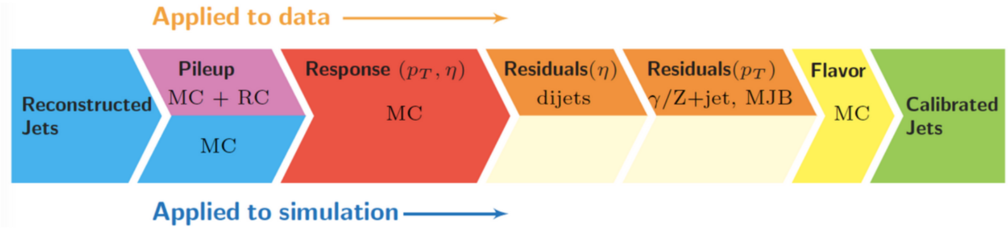


FIGURE 4.6: Jet energy correction levels [134].

- **Level 1:** This offset correction eliminates the jet energy response dependance on the additional PU activity within an event. It is based on the jet area method [134]. The SFs are derived by comparing the jet responses in simulated events with and without pileup events overlaid.
- **Level 2:** The next level of corrections aims to obtain a uniform energy response which is independent of the p_T and η of jet. The corrections are derived from simulated events by matching reconstructed jets to close-by true particle jets and comparing their momenta.
- **Level 3:** In this level, the residual differences between data and simulation are corrected by comparing the p_T balance in various types of events (multijet, Z+jets, γ+jets) where one jet is confined to be within the barrel region ($|\eta| < 1.3$) to provide a reference.

Additional jet energy (optional) corrections like jet-flavor dependencies of the energy response and the underlying event activity has also been developed in the CMS collaboration. These make small corrections compared to the above recommended corrections and therefore are not used in this thesis. The uncertainties associated with the JES are propagated as systematic uncertainties in physics analyses.

Measurements find that the JER in simulation is better than in data. Therefore simulated jets are smeared to match the data. In this analysis, the scaling method is utilized to smear reconstructed simulated jets. As a result their p_T resolution matches that of reconstructed jets observed in data. With this method the corrected four-momentum of a reconstructed jet in simulation is re-scaled with a factor c_{SF} which is defined by the following equation

$$c_{SF} = 1 + (s_{SF} - 1)\frac{p_T^{reco.} - p_T^{true}}{p_T^{reco.}}, \tag{4.7}$$

where $p_T^{reco.}$ is the the reconstructed transverse momentum, p_T^{true} is the jet transverse momentum from generator level particles and s_{SF} is the data to simulation core resolution scale factor which is measured in bins of η as given in Table 4.3.

TABLE 4.3: Jet energy resolution scale factors (s_{SF}) as a function of $|\eta|$ and their uncertainties for 13 TeV (2015 and 2016) data.

$\|\eta\|$	s_{SF}	Uncertainty
0 - 0.05	1.109	0.008
0.5 - 0.8	1.138	0.013
0.8 - 1.1	1.114	0.013
1.1 - 1.3	1.123	0.024
1.3 - 1.7	1.084	0.011
1.7 - 1.9	1.082	0.035
1.9 - 2.1	1.140	0.047
2.1 - 2.3	1.067	0.053
2.3 - 2.5	1.177	0.041

4.8.1 B-tagging

The identification of jets that stem from the hadronization of b quarks, the so-called "b-tagging", is a crucial ingredient in studies of top quark production. It provides discrimination power to single out jets which can be related to b quarks as expected in top quark decays. Many algorithms have been developed at CMS for b-tagging [135, 136] of jets that fall within the pseudorapidity acceptance of the tracker. A common feature of most algorithms is the identification of a secondary vertex (SV) which is reconstructed from displaced tracks within a jet. The general idea is illustrated in Figure 4.7. After hadronization, a final state b quark is encapsulated into a B meson (e.g. $B^{\pm}, B_0, B_s$) which can then travel a measurable distance away from the primary vertex before decaying due to its relatively long lifetime. For example, a $B^{\pm}$ meson, which has a mean lifetime of about 1.6 ps, can travel distances of roughly 4-9 mm for momenta of 40-100 GeV on average before decaying.

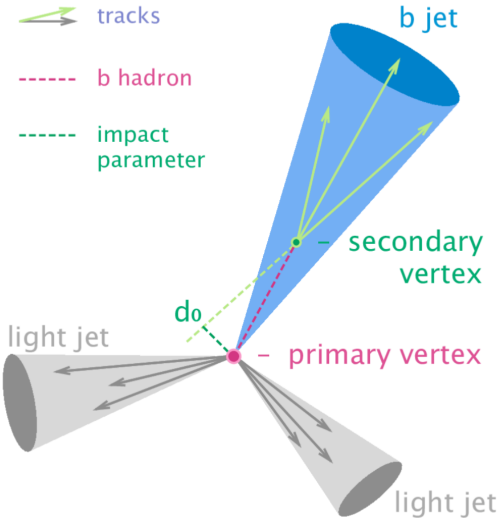


Figure 4.7: Schematic diagram of a b quark jet from the decay of the B meson resulting in charged particle displaced tracks. The distance between PV and the SV is the flight distance of B meson while d_0 is impact parameter. The figure is taken from [137].

After reconstruction, secondary vertices are subjected to pass certain quality criteria to enhance their purity with respect to the B meson hypothesis. These are based on amount of shared tracks with the PV, the invariant vertex mass to reject kaon decays, and the direction of tracks with respect to the jet axis. In this thesis, b-tagging algorithms based on multivariate discriminators are employed. The training of the discriminator covers scenarios where no SV has been reconstructed within a jet. In such cases, the compatibility of tracks with the PV is condensed into the discriminant. A comparison of the performance of various b-tagging algorithms employed within CMS is shown in Figure 4.8. It shows the misidentification probability of falsely tagging charm and light (g,u,d,s) quarks as a function of the efficiency of true b jets identification in simulation.

The combined secondary vertex (CSV) algorithm is employed in the 8 TeV and first 13 TeV (CSV version 2) analyses within this thesis. At their tight working points, a b-tagging efficiency of approximately 50% with a misidentification probability of only 0.1% is achieved. A new algorithm, called combined MVA (cMVA), is utilized in the analysis of 13 TeV data recorded in 2016. It combines the output of various other b-tagging algorithms which includes also the CSV discriminant. Additionally, the results of MVA discriminators trained to identify low-p_T electrons or muons inside a jet for b-tagging are taken as inputs as well. The final cMVA discriminant exhibits

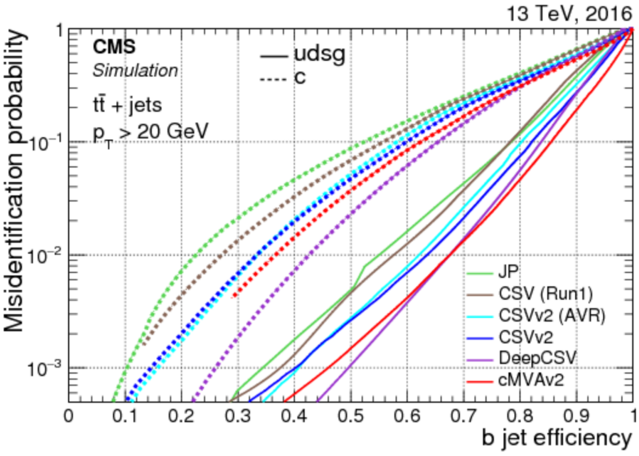


FIGURE 4.8: B-tagging efficiency and misidentification probability for c and light flavor (d,u,s,g) jets for various for various discriminators in 2016 data [136].

an improved rejection of charm and light quark jets compared to the other algorithms as shown in Figure 4.8 [136].

The performance of the taggers was studied on simulated samples. However, small deviations are found in the final discrimination variables between data and simulation. These deviations can cause a small discrepancy in tagging performances between data and simulated samples. In order to overcome these deviations SFs need to be applied to simulated events. The SFs are calculated as a function of p_T and η, and the discriminator value of a jet of flavor f according to the following equation:

$$SF_f(p_T, \eta, disc.) = \frac{\epsilon_f^{data}(p_T, \eta, disc.)}{\epsilon_f^{sim}(p_T, \eta, disc.)}, \tag{4.8}$$

The flavor of jets in simulations is determined from the matched generated hadrons. In data they are measured by selecting a sample enriched in jets with flavor f. Depending on the use of b-tagging in an analysis, the SFs are measured and applied in different ways [138].

4.9 Missing Transverse Energy (MET)

MET, denoted as $\vec{\not{E}}_T^{miss}$, is the imbalance in the p_T of all visible particles in the final state of the collisions. In physics analyses, it is the summed transverse momenta of produced neutrinos which escape the CMS detector without being detected due to their very low interaction probability. In this thesis, one neutrino is expected to be produced (from top decay) which can lead to significant missing transverse energy of about MET $\approx$ 50 GeV on average. MET in the PF event is calculated from the vector sum of $\vec{p}_T$ of all reconstructed particles [139].

$$\vec{\not{E}}_{T,raw} = -\sum_{i=1}^{n} \vec{p}_{T,i} \tag{4.9}$$

where n is multiplicity of reconstructed particles. Since the energy of jets are corrected, as explained in Section 4.8, these corrections have to be taken in account so that the uncorrected transverse momentum of each PF jet with p_T >10 GeV is replaced by the corrected transverse momentum $\vec{p}^{corr}_{T,j}$, so the corrected $\vec{\not{E}}_T^{miss}$ is given by:

$$\vec{\not{E}}_{T,raw} = -\sum_{i=1}^{n} \vec{p}_{T,i} - \sum_{j=1}^{njets} (\vec{p}^{corr}_{T,j} - \vec{p}_{T,j}) \tag{4.10}$$

Chapter 5

Measurement of Backplane Corrections for CMS Silicon Tracker

The inner tracking system of the CMS detector, also called tracker, has to fulfill several tasks. These are the charge and momentum measurement of created particles, the reconstruction of their trajectories and the measurement of the IP and thus the identification of SV. To achieve these goals, the tracker measures space points of the tracks in a high B-field of 3.8 Tesla.

The curvature of a charged particle travelling in a B-field is related to the momentum and charge, which allows for indirectly measuring both parameters according to the formula:

$$p_T = qBR, \tag{5.1}$$

where p_T is transverse momentum, B is magnetic field, q is charge and R is radius of the curvature.

The geometric layout of the tracker was designed such that each track is measured independently in the different subsystems several times. The number of measurement points is typically between 8 and 13, which allows for robust track reconstruction provided that the precision of these measured space points is high and they are unambiguous. The impact-parameter resolution is determined by two contributions. One is the granularity of the sensors and the other is multiple scattering in the detector layers. To achieve a good performance, the granularity of the tracking layers should be high and the distance of the first layer to the beam pipe has to be minimal. Choosing detector technology with thin layers such as silicon reduce the contribution of multiple scattering. The momentum resolution also has two contributions that govern the performance in different regimes of momenta. For high momenta, the resolution depends on the point resolution of the individual tracking layers and the number of measurement points, the arc length of the track in B-field, the measurement points and the momentum itself. For lower momenta, the resolution is dominated by multiple scattering and can be increased to a certain limit by using less material in the tracking volume. The design resolution of the individual layers of the CMS tracker ranges from 15 to 40 μm from the inner to the outer regions and combined with the large tracking volume and strong magnetic field leads to excellent impact parameter and momentum resolution.

This chapter will give a detailed summary of the tracker, its components, the readout system and the tracker performance studies for backplane correction performed using 2017 CMS data.

5.1 CMS Tracker Description

The total surface of the CMS tracker is 210 m^2 divided into 13 barrel layers and 14 forward discs. It is radially divided into three regions according to particle flux and track density. The innermost layers are designed as silicon-pixel detectors with very high granularity and the two outer sections are made of silicon-strip detectors with the strip pitch and detector thicknesses are adapted to the particle rate. The whole tracker is divided into the barrel region, where the layers are situated around the interaction point in a barrel-like geometry, and the end-cap region, where they are parallel circular discs. The tracker extends from as close as four cm to a radius of 110 cm from the interaction point and has an overall length of 540 cm [140]. Figure 5.1 shows one quarter of the tracker and the nomenclature of the individual regions. These individual regions are discussed in the following sections.

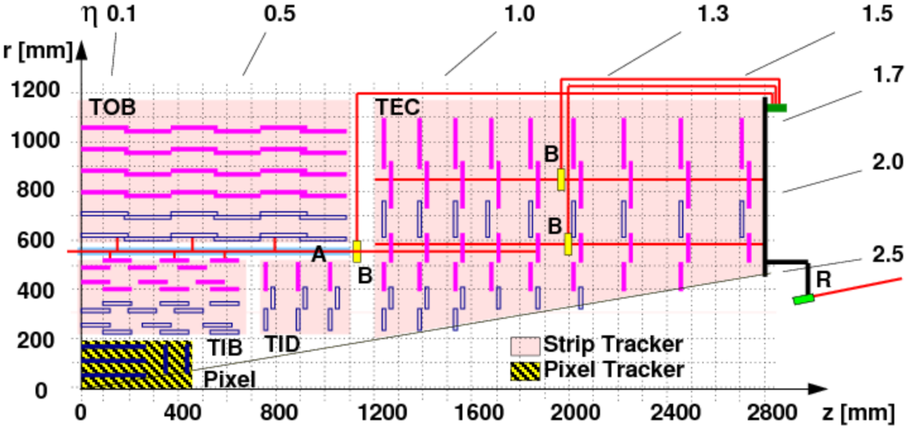


Figure 5.1: One quarter schematic view of the silicon tracker in the r-z plane. The pixel modules are at lower radii and at larger radii are strip modules [141].

5.1.1 CMS Pixel Detector

The part adjacent to the primary vertex, where the particle flux is maximum, is the silicon pixel tracker. It has to meet extreme requirements in terms of rate capability and radiation tolerance whilst providing high precision, three-dimensional space points for track reconstruction and the identification of secondary vertices. When a short-lived particle like a b quark or tau lepton is created in a collision it will not live long enough to be directly measured by one of the detector layers but it is possible to extrapolate the trajectories of the decay products and thus reconstruct the original secondary vertex. This requires an exceptional impact parameter resolution between 100 and 10 μm, depending on the particle momentum, which is achieved by high spatial resolution of the pixel system [142]. The CMS pixel tracker is composed of three barrel layers at radii of 4.4 cm, 7.3 cm and 10.2 cm and two end-cap discs in a slanted geometry (rotated at an angle of 20^o) covering a radius from 6 to 15 cm. The overall length of the barrel ladders is 53 cm and the pixel tracker has a total of 66 million readout channels.

The individual pixel cells are rectangular with a size of 100×150 μm^2, which was found optimal due to the high magnetic field. In CMS, significant Lorentz effects occur, which lead to charge sharing between the pixels and thus improved resolution. Consequently, analog readout of the pixel sensors is necessary to take advantage of these effects. The measured resolution of the CMS pixel tracker is 10 μm in r-ϕ direction and 20 μm along the beam (z-direction) [140].

The 270 and 285 micrometer-thick sensors are based on the n-in-n process, which means that the signal is read out from n^+-doped electrodes in a high resistivity n-doped substrate. The p implant at the back requires a double-sided production process and has the advantage of allowing under depleted operation. The electron readout of the n-in-n concept allows for the large Lorentz angle (LA) due to the increased mobility of the electrons as compared to holes

The front-end chip of a pixel system is one of the key elements in its performance as it has to be highly integrated and fulfil tasks like identifying detector signals created by a passing particle, store time, position and analog signal height and has to send out data if a trigger is received [142]. The chip reads an array of 52 times 80 pixels and also has a peripheral section that houses the buffers for event-data storage and the bond pads for connection to the hybrid. The individual pixel readout cells of the chip consist of a pre-amplifier stage, a shaper and a comparator circuit [143].

5.1.2 Silicon Strip Detector (SSD)

In the outer regions of the central tracking volume of CMS, where the particle rate is considerably lower than in the pixel region and extreme granularity is not required, silicon microstrip detectors are used. They are arranged in cylindrical layers outside the pixel volume and provide r-Φ information, with some layers providing additional z information. The silicon-strip tracker must measure the momenta of particles precisely, separate electrons and photons reliably and localize calorimeter showers exactly while keeping the general occupancy of the readout channels at the percent level. Together with the space points of the pixel tracker the strip tracker provides robust tracking information. The radiation environment in this region of CMS is a lot less hostile than in the pixel tracker, therefore the strip system was designed and built for the full nominal ten years of LHC operation before the high-luminosity upgrade. Some design choices that were made include high modularity for relatively easy mass production of the huge number of modules, simple and relatively cheap detectors and

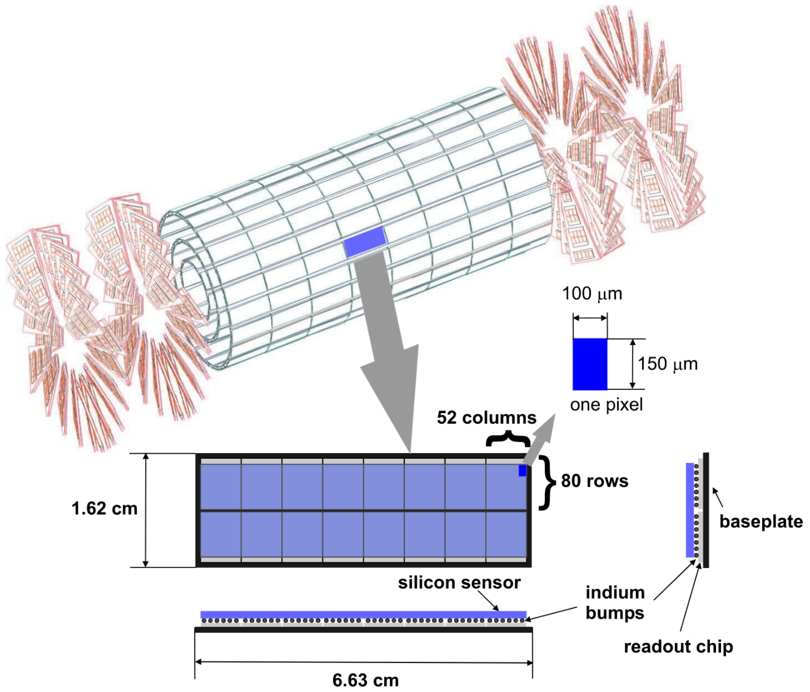


FIGURE 5.2: Schematic diagram of the CMS pixel detector at the top and a zoomed barrel module at bottom [144].

fully analog readout electronics together with low-mass support structures and service infrastructure [145].

The SSD consists of four individual parts that can be seen in Figure 5.1. Two are in the barrel region named as tracker inner barrel (TIB) and tracker outer barrel (TOB) and the remaining two are in the endcap named as tracker endcap (TEC) and tracker inner discs (TID).

TIB and TOB feature different support structures and sensor parameters. TIB has four four layers of silicon microstrip detectors made up of 320 μm thick silicon and a strip pitch between 80 and 120 μm. The two innermost layers are made up of stereo modules where two single-sided sensors are mounted back-to-back, slanted at a stereo angle of 100 mrad to provide two-dimensional position resolution. The resolution that is achieved is between 23 and 34 μm in r-ϕ and 23 μm in z-direction [140]. In the TOB region the radiation and occupancy levels are considerably lower, therefore the granularity can be decreased even further, which means that the strip pitches are wider and the individual strips longer. To maintain the level of signal to noise ratio, 500 μm thick silicon sensors are used. The TOB has six layers with strip pitches between 120 and 180 μm and the two innermost layers again feature a stereo angle. As the strip pitch is wider, the resolution is lower with 35-52 μm in r-ϕ and 52 μm in z direction.

TEC consists of nine discs and TID has three rings that fill the gap between TIB and TEC (see Figure 5.1). The wedge-shaped modules are arranged in rings with the strips pointing towards the beam and the first two discs of the TID and discs one, two and five of the TEC feature double sided modules with a stereo angle. Analog to the barrel region, the thickness of the silicon was chosen to be 320 and 500 μm for TID and TEC respectively. TID and TEC are arranged in a way that they cover the maximum surface with the fewest gaps [145].

Due to the large surface and the large number of sensors, a lot of careful consideration had to go into the selection of the technological baseline. The inner and outer barrel and endcap regions all use AC coupled, poly-silicon biased, single sided p-on-n devices processed on six inch wafers [145]. This means that the readout electrodes (the microstrips) are highly doped p^+ implants in a high resistivity n-doped substrate. This is relatively simple and low cost process widely used in the semiconductor industry and it has been proven to be sufficiently radiation hard for the expected fluence over the nominal tracker lifetime. Such devices can be produced in large quantities and with high yields by the industry. The only difference between inner and outer silicon-strip tracker sensors is the thickness, which is larger for the outer regions (500 vs. 320 μm). Due to the various geometries for stereo modules, barrel- and end-cap regions and the varying pitch, a total number of 15 different sensor layouts were used.

The choice of p-on-n technology for the detectors yields some effects that have to be taken into account when planning the operation. Due to radiation damage with increasing fluence, these detectors undergo a process called type inversion where the effective doping of the substrate changes from n- to p-type. Consequently the depletion region will start growing from the back when the detector is biased which does not allow under depleted-operation any more. Table 5.1 summarizes the important features of the silicon sensors.

Table 5.1: Silicon Strip tracker characteristics.

Part	No. of Detectors	Thickness (μm)
TIB	2,742	320
TOB	5,208	500
TID	816	320
TEC	2,512	320
TEC(2)	3,888	500

5.2 Tracker Readout

Silicon microstrip detectors are usually read out by charge sensitive amplifiers. Due to the large number of ($\approx$ 9 million) channels in the CMS tracker, this amplifier is realized as a highly integrated front-end chip that has to meet very tight requirements in terms of particle rate, magnetic field, material budget and radiation tolerance. Since the event rate in the LHC collisions is 25 ns, fast sampling is absolutely necessary but data has to be stored for at least the duration of the Level-1 trigger decision, which takes 3.7 microseconds, a time equivalent to 150 bunch crossings. The general performance of the tracker in terms of position resolution can benefit from charge sharing effects between the individual strips and therefore it has been decided to read out the full analog signal from each individual strip and transmit it to the front end driver boards, where it is processed. The following section will describe the front end chip of the strip tracker and the general architecture of the readout system.

The readout chip used in the CMS silicon strip tracker is called APV25 and was specifically developed to meet the strict requirements of the experiment. Excellent radiation tolerance was considered very important and consequently iterations of the chip were produced in different radiation hard processes. Finally the implementation in a 0.25 μm CMOS process met the requirements and also provided significant savings of costs. The chips have 128 input channels split into two groups of 64 channels. Each channel consists of a pre-amplifier stage which is coupled to a shaping stage which produces a 50 ns CR-RC pulse. A gain inverter between the pre-amplifier and the shaper ensures same-polarity output independent of the input-signal sign, which allows to read detectors with electron- (n-in-n, n-in-p) and hole (p-in-n) signals [146]. The CR-RC pulse of the shaper of each channel is sampled at a clock rate of 40 MHz and written into a 192 cell deep pipeline, which allows for storage for a maximum time of 4.8 μs and thus longer than the 3.7 μs of a Level-1 trigger decision. Upon the reception of a trigger signal the content of the affected pipeline cell columns are read out by the analog Pulse Shape Processor (APSP), which has two operation modes (Figure 5.3).

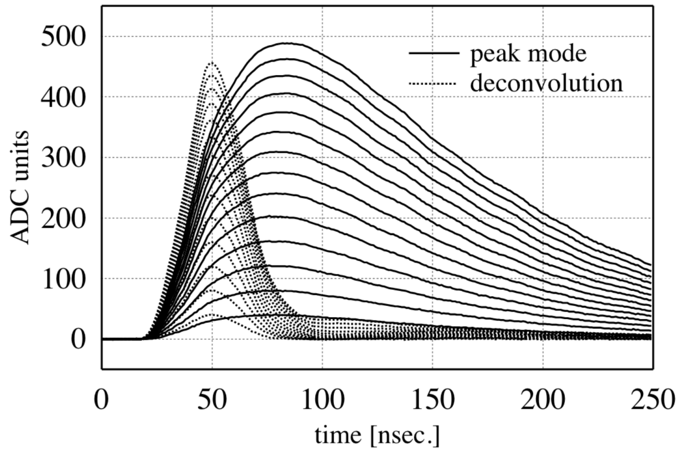


FIGURE 5.3: APV25 pulse shapes for different input-signal heights in peak and deconvolution-mode [146].

- **Peak Mode (PM):** In peak-mode, only one cell per input channel is read out and the output is timed to the peak of the analog pulse with a relatively slow rise time of 50 ns.
- **Deconvolution Mode (DM):** In this mode, three consecutive samples are read and the output is a weighted sum of the three. This results in a reshaped, shorter pulse that is peaked at 25 ns, which corresponds to a single bunch crossing.

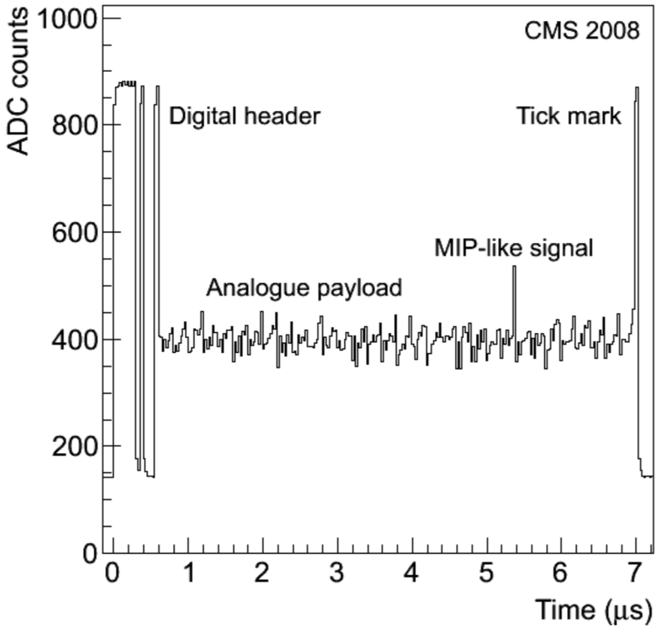


Figure 5.4: Multiplexed signal of two APV25 chips with the different components of the frame [147]

After this stage data of each channel is multiplexed to a differential current with non-consecutive order of channels. The resulting signal is 7 μs long, has a 12-bit digital header that includes address and error information, followed by the analog signal of 128 strips, and is terminated by a tick-mark, which is used for many purposes like for example the calibration of the absolute signal [146, 147]. The more or less constant ADC count of the channels that show no particle hit in Figure 5.4 is called the pedestal value and is unique to each channel. Another contribution that varies with time and is small compared to the pedestals but common to all 128 channels of a chip is called common-mode-shift [147].

5.3 Coordinate System of the CMS Tracker

The global coordinate system tracker is defined as (r|ϕ|z), where r, ϕ and z are explained in Section 3.2.1. For each module, a local coordinate system (u|v|w) is defined. The origin is defined at the center of the module. The Figure 5.5 shows the local coordinate system how they are related to the global coordinate system.

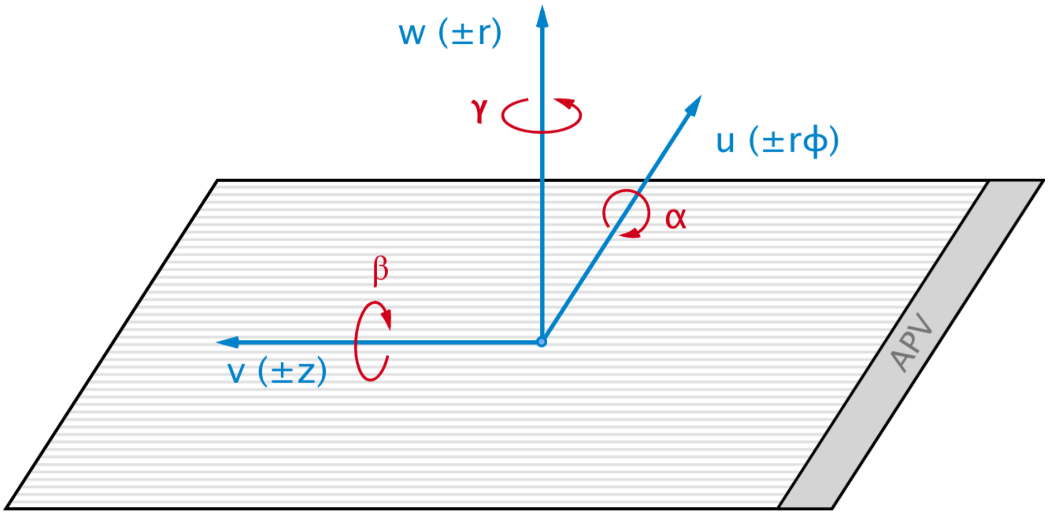


FIGURE 5.5: Schematic diagram of strip module with local coordinate system (u|v|w) and local rotations α, β and γ. The right-handed global coordinates of the tracker (r|ϕ|z) are shown in brackets [141].

5.4 Backplane (BP) Correction

As described in Section 5.2, tracker module operates in two modes. The peak mode is used to take the data from the APV25 chip and then save the amplified signals after every 50 ns. Deconvolution mode consists of a weighted sum of three consecutive peak signals, which efficiently decreased the time up to 25 ns. The PM has a good signal to noise ratio and large integration time, perfect for cosmic data. The PM is not acceptable to the high collision rate of hadron colliders. Therefore to process LHC data, the APV25 chip is operated in DM. Only the DM is suitable for normal data taking during pp collisions because of 25 ns time spacing. Because of this short readout time, the effective depletion area becomes narrow resulting in change of the cluster shape as shown in Figure 5.6. This results in shift in the measured hit position. The shift depends on the incident track angle. A dedicated backplane corrections are therefore used to correct this shift.

In 2009, it was observed by the tracker alignment group that the sensor positions during data collection in deconvolution mode changed with respect to data collection in peak mode. This shifts in the sensor position was in the direction orthogonal to the sensor plane and in the direction of the Lorentz drift. TOB and TEC outer rings sensors are thick and have longer drift time, while the time window is narrower in the deconvolution mode, which produce bias in sensor position. The signal released across the thick sensors is not collected with uniform efficiency introducing a bias on the signal barycenter along the drift direction.

The shift in two different alignment geometries for peak and deconvolution data is avoided by applying a correction to the cluster position in the local reconstruction. The effective reduction of the sensor thickness is parametrized by the parameters that

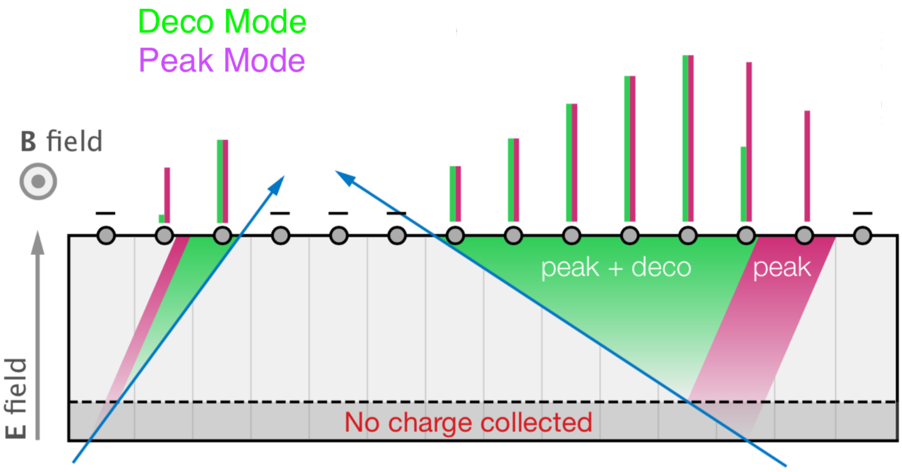


FIGURE 5.6: Schematic diagram of the BP correction in CMS tracker. In PM, complete charge from the sensor backplane is collected. In the DM, cluster size is reduced due to short time spacing of 25 ns. This reduction in cluster size depends on the track incident angle [148].

indicate the fraction of the sensor which has to be considered "inactive" to explain the observed bias in the w-direction. Since, the local reconstruction returns only the local coordinate of the clusters orthogonal to the strip direction, assuming that the barycentre of the cluster is in the w-direction, the bias on the w coordinate is translated into a bias in the u direction.

All charges move towards the plane of the sensor and make a cluster, while they are processing in both modes. The cluster barycentre is hypothesized to the centre point of the sensor plane. It is the original reconstructed hit point. In the DM, the readout scale of time is small and therefore all charges can not accumulate because the drift time for the backplane of the sensor or module is quite large. Therefore, the reconstructed hit point hypothesizes with the cluster barycenter by shifting position of sensor in w-direction such that the reconstructed hit point aligns with the track path.

The Figure 5.7 shows the case when some of the cluster charge is lost in the backplane. The geometrical barycentre of the track in the sensor is the blue point. The green point is the barycentre of the charge due the loss of the charge in the purple region of the track. The charge drifts to the surface creating the green cluster which is then reconstructed as the red hit by the local reconstruction which translates (back) the barycentre of the cluster along the expected drift direction up to the median plane of the sensor. Therefore, there is a mismatch between the track (and in particular the green point) and the reconstructed hit (red point). The alignment procedure can recover this mismatch for all the tracks by shifting the module along the orange arrow which is parallel to the drift direction, introducing a shift (Δw).

In case of magnetic field, the backplane effect is shown in Figure 5.8. Everything is similar to the previous case but the charge drift, the hit translation and the alignment correction occur along a direction which deviates from the perpendicular one by the Lorentz angle.

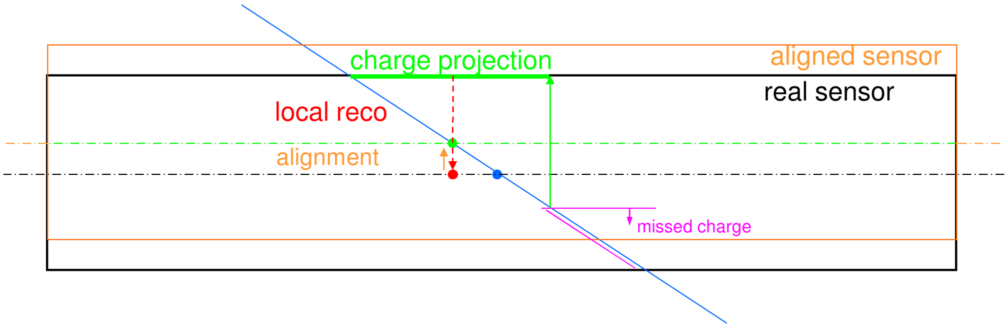


FIGURE 5.7: Measured and true hit positions in a silicon module in the absence of magnetic field. The position of the hit is determined as a position of charge cluster barycenter. Alignment moves the sensor to superimpose charge median plane (green) to geometrical median plane (orange) along the Lorentz drift to superimpose reconstructed hit (red) to the track (green) [149].

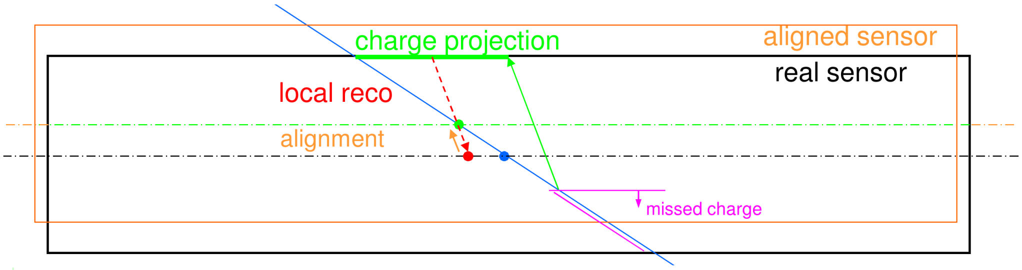


FIGURE 5.8: Measured and true hit positions in a silicon module in the presence of magnetic field. The position of the hit is determined as a position of charge cluster barycenter. Alignment moves the sensor to superimpose charge median plane (green) to geometrical median plane (orange) to superimpose reconstructed hit (red) to the track(green) [149].

The signal lost due to this BP effect, affects the measured hit position. Thus, to compensate this BP effect, a BP correction (Δw) is calculated for the strip tracker, where (Δw) is defined as:

$$\Delta w = \frac{\Delta u}{tan(\theta_{trk})}. \tag{5.2}$$

In the presence of the Lorentz drift, the correlation is given by:

$$\Delta w = \frac{\Delta u}{(tan(\theta_{trk}) - tan(\theta_{LA}))} \tag{5.3}$$

Furthermore, the deconvolution mode causes an effective change of the Lorentz drift direction, that is of the Lorentz angle. This is taken into account by using a different set of Lorentz angle values when running in deconvolution mode. These new values are equal to the values used in peak mode plus an overall correction called $\Delta(\tan(\theta_{LA}))$. The corrections are extracted by plotting track cluster residuals Δu versus difference between track incidence angle and LA $(\tan(\theta_{trk}) - tan(\theta_{LA}))$.

$$\Delta u = \Delta w(tan(\theta_{trk}) - tan(\theta_{LA})) \pm (H - \Delta w)\Delta tan(\theta_{LA}) \tag{5.4}$$

where

$$H = \frac{\text{detector thickness}}{2} \tag{5.5}$$

Equation of line is fitted to find slope Δw and y-intercept. The y-intercept is positive for modules with v parallel to z and negative for v anti-parallel to z.

5.5 Measurement of Backplane Correction

The width of the BP zone is calculated using separate PM and DM data samples collected in 2017. The effect of Lorentz angle (θ_{LA}) is not included in backplane correction measurement because the integration time for the charge accumulation in deco mode is too small that it leads to a minute shift in backplane correction value. The effect of non-zero θ_{LA} can be ignored as it is smaller than 1%, but for a more precise treatment this effect should be considered.

The shift in u-direction for peak and deconvolution modes for V+ and V- modules are shown in Figure 5.9 and Figure 5.10 respectively where V+ modules refers to the modules for which v-direction is parallel to z-direction and V- for v-direction anti-parallel to z-direction. Backplane correction as a function of track incident angle for W+ modules and W- modules orientation is shown in Figure 5.11 and Figure 5.12 respectively where W+ modules refers to the modules with orientation w-direction parallel to r-direction and W- for w-direction anti-parallel to r-direction.

The backplane correction values Δw = 1.15 ± 0.15 μm and $\Delta\theta_{LA}$ = -0.0007 ± 0.0000 for V+ modules are obtained from the fit and for the V- modules the values are Δw = 0.11 ± 0.16 μm and $\Delta\theta_{LA}$ = 0.0006 ± 0.0000. The backplane correction values Δw = -0.28 ± 0.15 μm and $\Delta\theta_{LA}$ = 0.0001 ± 0.0000 for W+ modules are obtained from the fit and for the W- modules the values are Δw = 1.87 ± 0.15 μm and $\Delta\theta_{LA}$ = 0.0012 ± 0.0000.

The sensor position in deconvolution modes are also compared with the peak mode when BP corrections are not applied. It was observed that backplane correction is of the order of 5-6 μm. The backplane corrections reported in this thesis are close to zero. This work is still in progress as the estimated backplane corrections are to be validated with cosmic data.

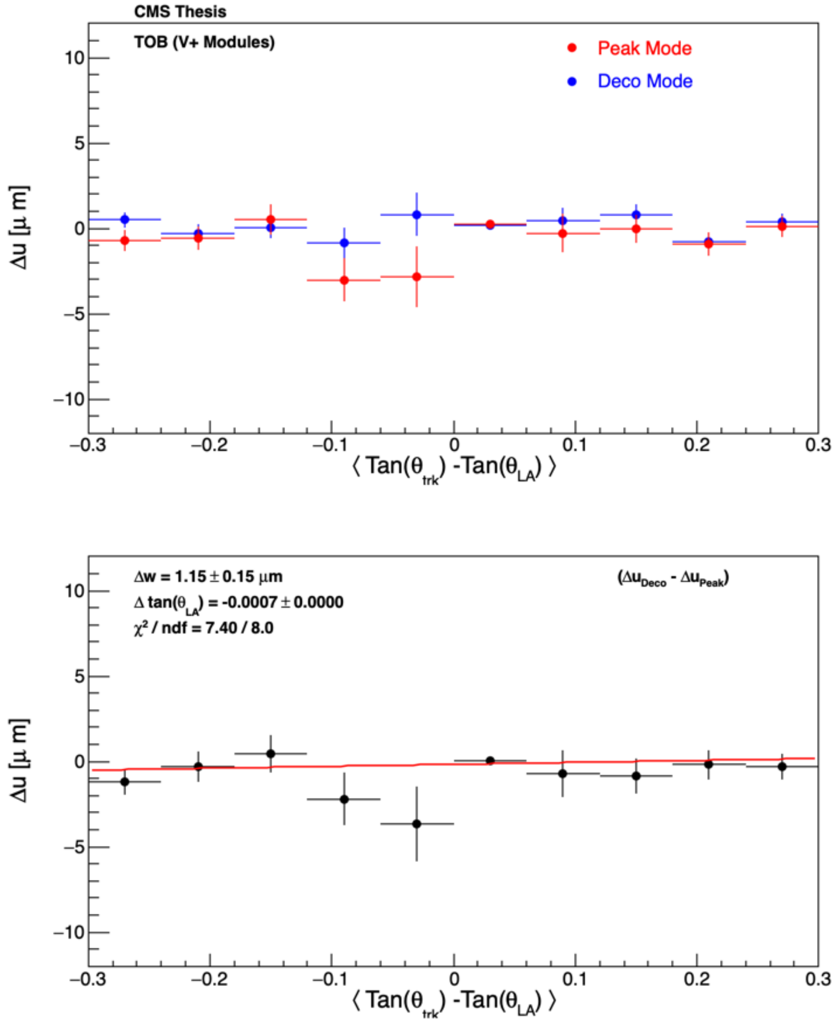


FIGURE 5.9: The shift Δu as a function of track incident angle is plotted for peak and deconvolution readout modes (top). The straight line fit to the difference of Δu in peak and deconvolution readout modes is shown (bottom). The plots are for all six TOB layers summed and the orientation of the modules is V+. V+ modules refers to the modules for which v-direction is parallel to z-direction. The determined backplane correction Δw is 1.15 ± 0.15 μm and Lorentz angle $\Delta\theta_{LA}$ = -0.0007 is extracted from the fit

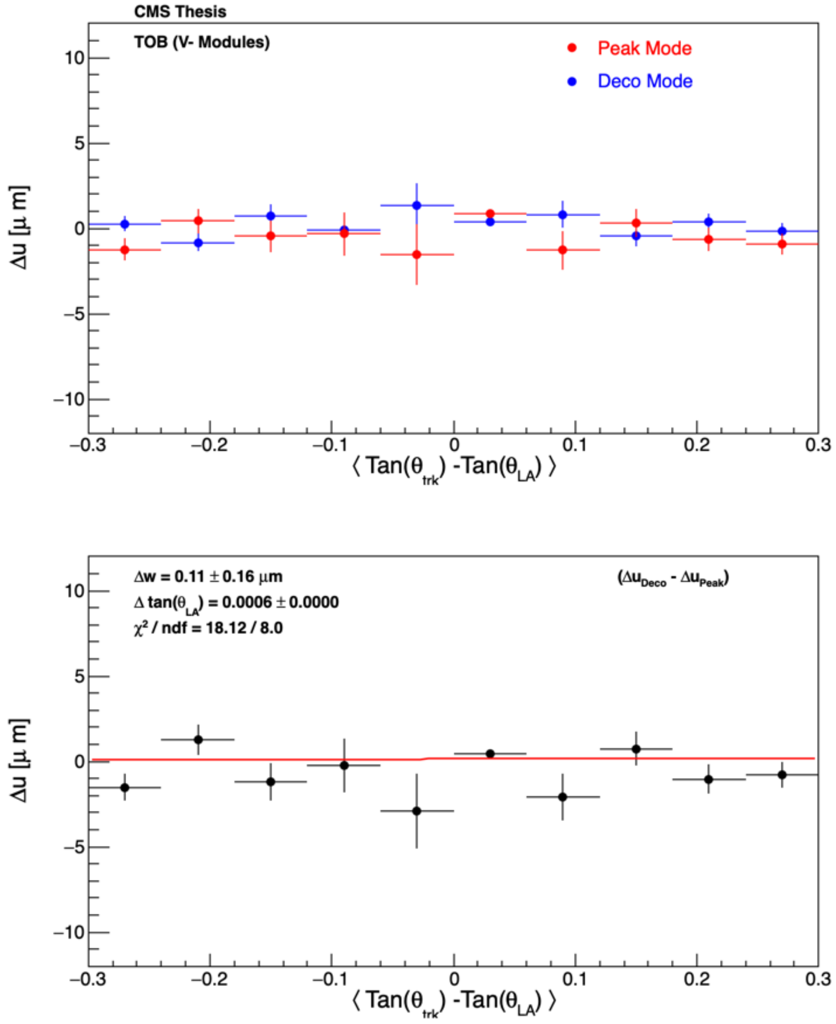


FIGURE 5.10: The shift Δu as a function of track incident angle is plotted for peak and deconvolution readout modes (top). The straight line fit to the difference of Δu in peak and deconvolution readout modes is shown (bottom). The plots are for all six TOB layers summed and the orientation of the modules is V-. V- modules refers to the modules for which v-direction is anti-parallel to z-direction. The determined backplane correction Δw is 0.11 ± 0.16 μm and Lorentz angle $\Delta\theta_{LA} = 0.0006 \pm 0.0000$ is extracted from the fit.

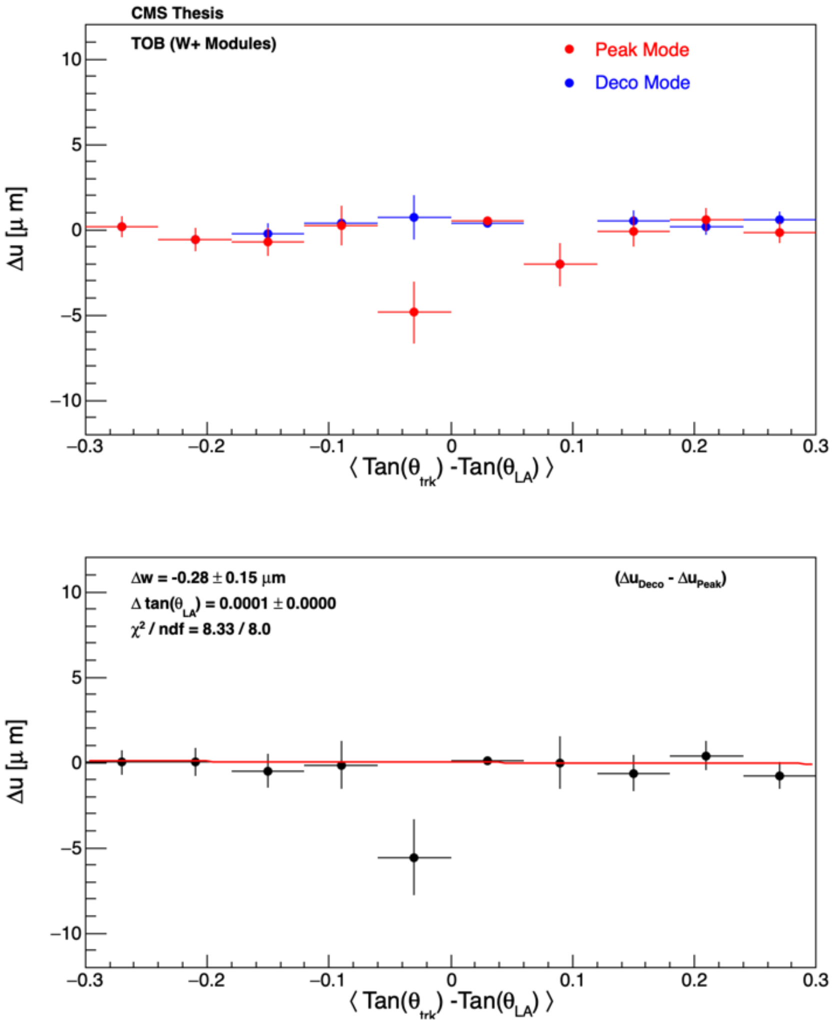


Figure 5.11: The shift Δu as a function of track incident angle is plotted for peak and deconvolution readout modes (top). The straight line fit to the difference of Δu in peak and deconvolution readout modes is shown (bottom).The plots are for all six TOB layers summed and the orientation of the modules is W+. The determined backplane correction Δw is -0.28 ±0.15 μm and Lorentz angle $\Delta\theta_{LA} = 0.0001 \pm 0.0000$ is extracted from the fit.

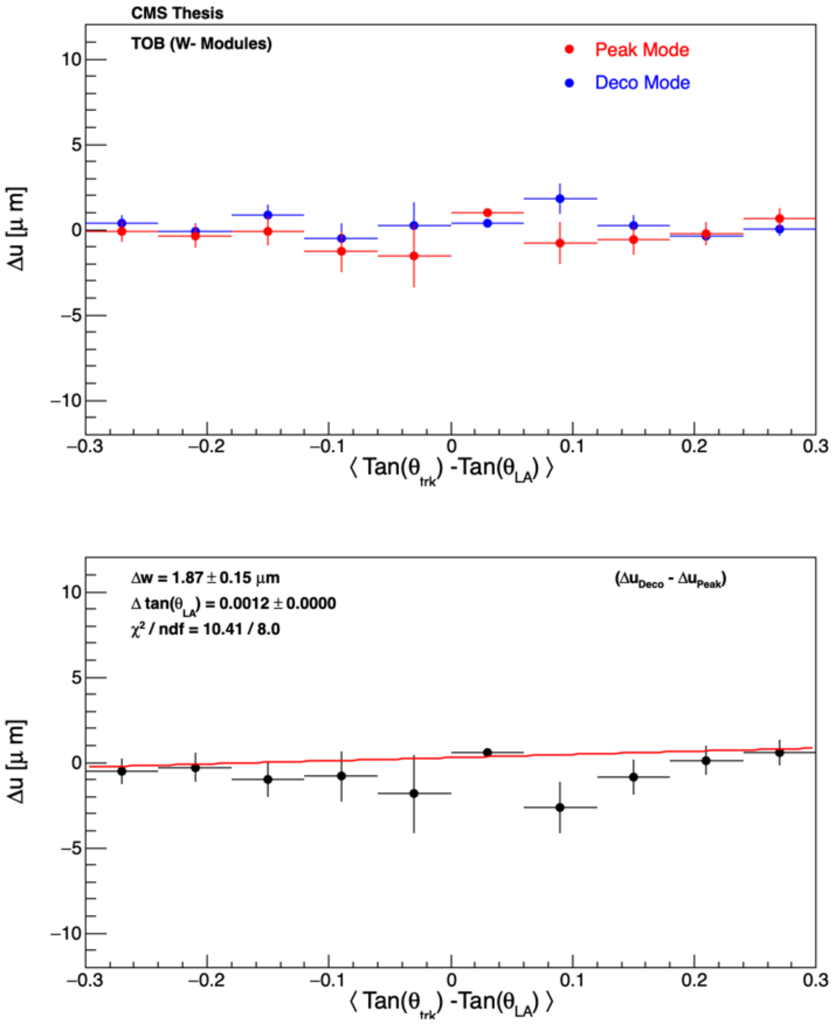


FIGURE 5.12: The shift Δu as a function of track incident angle is plotted for peak and deconvolution readout modes (top). The straight line fit to the difference of Δu in peak and deconvolution readout modes is shown (bottom). The plots are for all six TOB layers summed and the orientation of the modules is W-. The determined backplane correction Δw is 1.87 ± 0.15 μm and Lorentz angle $\Delta\theta_{LA} = 0.0012 \pm 0.0000$ is extracted from the fit.

Chapter 6

Statistical Analysis

The statistical analysis techniques and tools used to perform the measurements within this thesis are introduced in this chapter. The estimation of the amount of signal and background events in data through maximum-likelihood (ML) fits and to compare the results with theoretical predictions, statistics tools and methodology used at LHC are described. In Section 6.5, the reconstruction method of the top quark is described.

6.1 Hypothesis Testing

In particle physics, in order to claim the existence or absence of a new physics phenomenon a statistical test is performed involving two hypotheses hypotheses, the so called null and alternative-hypothesis. Clearly the null-hypothesis is tested against the alternative-hypothesis. In the physics analysis described in this thesis, the statistical method is implemented in the so called "Higgs Combine Tool" [150]. This tool is developed by the ATLAS and CMS collaborations specifically for the benefit of the LHC Higgs Combination working group (LHC-HCG) [151, 152]. It provides a command line interface for various statistical methods available inside RooFit/RooStats that are used widely inside CMS [153]. For purposes of excluding a signal process (i.e. setting limits) the two hypotheses are defined as:

- Null hypothesis H_0: The model which describes the background processes only and so is termed as the background-only (B) hypothesis.
- Alternative hypothesis H_1: The model which describes all processes including both the background and the new physics signal. This is also knows as the signal-plus-background (S+B) hypothesis.

A new parameter signal strength μ can be introduced to understand these two hypothesis. Signal strength is used as multiplicative factor to estimate signal yield. So for the (S+B) hypothesis μ should be 1 in the presence of signal and for the (B) hypothesis μ should be 0.

6.2 Likelihood Function and Likelihood Ratio

The probability for any observation to be commenced from a model is given by the likelihood function. A simplified version of a binned likelihood function in terms of the parameter of interest (POI), μ, and nuisance parameters, θ, can be written as the product of the probability of observing n_k data events when ($\mu\mathrm{s}_k$+ b_k) events are expected, for each bin k:

$$L(Data|\mu,\theta) = \prod_{k\in\text{All bins}} P(n_k|\mu s_k(\theta) + b_k(\theta)) \cdot pdf(\tilde{\theta}|\theta) \tag{6.1}$$

where s_k is the expected signal and b_k is expected background events in the k^{th} bin. The parameters $\theta = (\theta_1,\ \theta_2\ ...)$ are nuisance parameters (NP) used to model the variation of the signal $\mathrm{s}_k(\theta)$ and background $\mathrm{b}_k(\theta)$ contributions due to systematic uncertainties. Often, these nuisance parameters are constrained by external measurements, $\hat{\theta}$, which are encoded in the probability density function, $pdf(\tilde{\theta}|\theta)$. Assuming a Poisson distribution in each bin, the binned likelihood can be written as the product of the Poisson likelihoods of the individual bins of the variable

$$L_{Pois}(n_k|\mu,\theta) = \prod_{j}^{regions} \prod_{k}^{bins(j)} \frac{(\mu s_{jk}(\theta) + b_{jk}(\theta))^{n_{jk}}}{n_{jk}!} \exp\left(-(\mu s_{jk}(\theta) + b_{jk}(\theta))\right) \cdot pdf(\tilde{\theta}|\theta) \tag{6.2}$$

where $\mathrm{s}_{jk}(\theta)$ and $\mathrm{b}_{jk}(\theta)$ are the expected signal and background events respectively in the k^{th} bin of the distribution. The fit has been performed in the region j. The nuisance parameters, which will be discussed in Section 7.7 and 8.7, are the systematic uncertainties estimated from auxiliary measurements. Systematic uncertainties are considered as 100% uncorrelated or correlated. In case the uncertainty is partially correlated it is splitted into sub-components to fit the requirements. The morphing technique [154] is used for modelling the shape uncertainties.

The Bayesian and modified classical frequentist statistical approaches are used in HEP to characterise the absence of a signal. The level of incompatibility of data with a signal hypothesis is quantified in terms of confidence levels (CLs). The convention is to require a 95% CL for excluding a signal. In general, limits are not set on the signal cross section directly, but are set on μ. In the physics analyses of this thesis, the modified frequentist approach is used for confidence level approximation [155].

The hypothesized value of μ can be tested by using the profiled likelihood ratio which is defined as

$$\lambda(Data|\mu) = \frac{L(Data|\mu,\hat{\hat{\theta}})}{L(Data|\hat{\mu},\hat{\theta})} \tag{6.3}$$

where $\hat{\theta}$ and $\hat{\mu}$ are the maximum likelihood estimators (MLE). It means that both $\hat{\theta}$ and $\hat{\mu}$ maximises the likelihood in denominator. The $\hat{\hat{\theta}}$ is the value of θ that (conditionally) maximizes the likelihood for a specified μ for a given (pseudo/real) data. For many analyses, the signal contribution is assumed to be non negative which means μ must be positive. However, the definition of $\hat{\mu}$ should be the value of μ that maximized the likelihood even if this gives a negative value for $\hat{\mu}$, the Poisson mean which is represented by $\mu s_i + b_i$, remains positive.

6.3 Test Statistic

At the LHC, the test statistic, t_μ, is given by:

$$t_\mu = -2 \ln \lambda(\mu) \tag{6.4}$$

The distribution of t_μ is sampled for the (S+B) and (B) hypotheses by means of simulated pseudo-experiments. The calculation of t_μ can be CPU intensive, so to reduce computing time the Asymptotic CL method is used. In this method, Monte Carlo samples are replaced by one representative dataset, called the Asimov dataset which is constructed such that all observed quantities are set equal to their MLE values. More details about this procedure can be found in Ref. [156].

The test statistic is used to compute the p-value. The p-value is the probability of obtaining an outcome of a statistical test at least as extreme as the observed outcome. The p-value is given as:

$$p_\mu = \int_{t_{\mu,obs}}^{\infty} f(t_\mu|\mu) dt_\mu \tag{6.5}$$

where $f(t_\mu|\mu)$ is the PDF of test statistic t_μ ,assuming signal strength μ and $t_{\mu,obs}$ is the observed value of the test statistic in data. The integral extends from this observed value to infinity. The (B) hypothesis p-value (p_0) can be calculated with special case $\mu = 0$. The p-value and significance Z are related as:

$$Z = \Phi^{-1}(1 - p) \tag{6.6}$$

where Φ is the quantile of the standard Gaussian distribution.

The relation between the observed test statistic t_μ, the p-value and significance Z is shown in the Figure 6.1. In particle physics, a very stringent criteria of significance Z $>$ 5 is set for the discovery of new physics. This value of Z is equivalent to the p-value of 2.87 $\times 10^{-7}$. The significance Z $>$ 3 is required for the evidence of new physics.

6.4 Confidence Levels (CLs) Procedure

In most of the physics analysis the CLs method is used to exclude regions of phase space with CLs $<$ (1- α), where α is the desired confidence level. For the real data, the

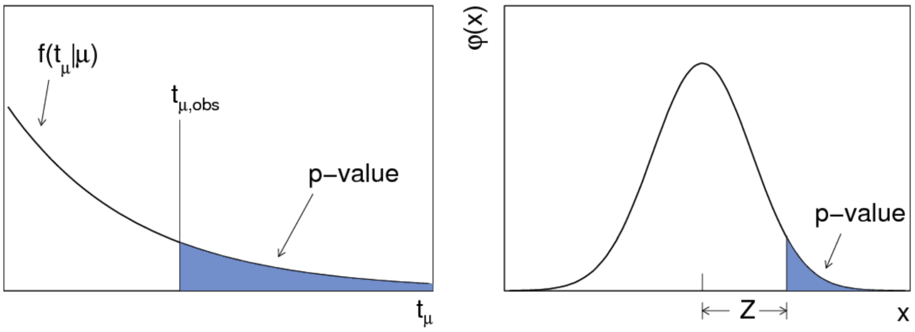


FIGURE 6.1: The relation between test statitic t_μ, the p-value, and the significance Z is shown. The figure is taken from Ref. [88].

probability of finding a data set with equal or greater incompatibility to the (S+B) hypothesis is determined by p-value associated to (S+B) hypothesis which is given by P_{S+B}:

$$P_{S+B} = \int_{t_{\mu,obs}}^{\infty} f_{S+B}(t_\mu|\mu, \hat{\theta_\mu})dt_\mu \tag{6.7}$$

where $t_{\mu,obs}$ is test statistic observed vaue for real measured data, $f_{S+B}(t_\mu|\mu)$ is the probability density function describing the distribution of t_μ for the (S+B) hypothesis which is constructed by pseudo-data with nuisance parameters fixed to $\hat{\theta_\mu}$. Similarly, for the background-only hypothesis:

$$1 - P_{S+B} = \int_{t_{\mu,obs}}^{\infty} f_B(t_\mu|\mu = 0, \hat{\theta_\mu} = 0)dt_\mu \tag{6.8}$$

The signal is excluded at 95% confidence level (CL) of

$$CLs = \frac{P_{S+B}}{1 - P_B} \leq 0.05 \tag{6.9}$$

The observed limit is defined as the upper limit on μ at the 95% CL which is obtained by adjusting μ until CL = 0.05. The expected median upper limit and the $\pm 1\sigma$ and $\pm 2\sigma$ band are generated by a set of pseudo-data, the CLs and the value of at 95% CL for each of them. The median expected value is obtained by replacing the $t_{\mu,obs}$ with the value of the test statistic that represents the 50% quantile of the cumulative distribution function. The $\pm 1\sigma$ and $\pm 2\sigma$ bands on the expected limit are obtained when replacing $t_{mu,obs}$ with the 68% and 95% quantiles of the cumulative distribution function respectively.

6.5 Top Quark Reconstruction

In the presented analyses, a top quark candidate is reconstructed in data and simulated events. Assuming that the top quark decays leptonically as t→bW→blν in selected events, its energy and momentum is calculated by summing the 4-momenta of a selected lepton candidate (muon or electron), a b-tagged jet, and a neutrino candidate ($\vec{\not{E}}_T^{miss}$). The x- and y- components of the $\vec{\not{E}}_T^{miss}$ originate from the neutrino in the transverse plane. The neutrino z-component ($\mathscr{P}_{z,\nu}$) is reconstructed using $\vec{\not{E}}_T^{miss}$ and the lepton momentum by imposing a W boson mass constraint i.e. satisfying the following equation .

$$M_W^2 = (80.4)^2 = \left(E_e + \sqrt{\vec{\not{E}}_T^{miss^2} + \mathscr{P}_{z,\nu}^2}\right)^2 - (\vec{P}_{T,e} + \vec{\not{E}}_T^{miss})^2 - (\mathscr{P}_{z,e} + \mathscr{P}_{z,\nu})^2 \quad (6.10)$$

There are in general two solutions for $\mathscr{P}_{z,\nu}$:

$$\mathscr{P}_{z,\nu}^{A,B} = \frac{\mu \cdot \mathscr{P}_{z,e}}{P_{T,e}^2} \pm \sqrt{\underbrace{\frac{\mu^2 \cdot \mathscr{P}_{z,e}^2}{P_{T,e}^4} - \frac{E_e^2 \cdot \vec{\not{E}}_T^{miss^2} - \mu^2}{P_{T,e}^2}}_{\equiv \xi}} \quad (6.11)$$

where

$$\mu = \frac{M_W^2}{2} + \vec{P}_{T,\mu} \cdot \vec{E}_T^{miss}. \quad (6.12)$$

A detailed derivation of this result can be found in Ref. [157]. In simulated t-channel single top events, this procedure leads to two real solutions in about 65% of all selected events. The solution which has the smallest absolute $|P_{z,\nu}|$ value yields on average the solution which is significantly closer to the true neutrino p_z.

Complex solutions are obtained if the radicand xi in Equation 6.12 becomes negative. This happens in about 35% for selected signal events. Such solutions occur mostly due to the finite $\vec{\not{E}}_T^{miss}$ resolution whereas negligence of off-shell W bosons and the resolution of the lepton momentum are found to be minor effects. The imaginary part of the solutions is removed by requiring that the discriminant vanishes. This is equivalent to setting $m_T = m_W$ which leads to a quadratic equation for $\mathscr{P}_{y,\nu}$ with dependance on $\mathscr{P}_{x,\nu}$. Assuming that the true values of the neutrino momentum components are not too far away from the measured $\vec{\not{E}}_T^{miss}$ components, the distance (δ) between the neutrino $\vec{\not{E}}_T^{miss}$ and the modified $\vec{\not{E}}_T^{miss}$ is minimized with respect to the solutions $\mathscr{P}_{y_{1,2},\nu}$:

$$\delta_{1,2}(\mathscr{P}_{x,\nu}) = \sqrt{\left(\mathscr{P}_{x,\nu} - (\vec{\not{E}}_T^{miss})_x\right)^2 + \left(\mathscr{P}_{y_{1,2},\nu}(\mathscr{P}_{x,\nu}) - (\vec{\not{E}}_T^{miss})_y\right)^2}. \quad (6.13)$$

The smallest δ value is chosen to keep the transverse energy closer to $\vec{\not{E}}_T^{miss}$. After removing the imaginary part, the $p_{z,\nu}$ solution is on average closer to the true neutrino $\mathscr{P}_{z,\nu}$ momentum compared to cases with real solutions.

The top quark can be reconstructed once neutrino $\mathscr{P}_{z,\nu}$ component is known. The jet assignment, which jet originates from which quark, depends on the multiplicity of b-tagged jets in the event. In this thesis, there is only one b-tag jet in the event therefore it is assumed to be the b-jet from top decay (correct in 92.7% of the cases [158]).

Chapter 7

Search for tZq Process at $\sqrt{s} = 8$ TeV

A first search of single top quark production in association with a Z boson is presented. The search is performed within standard model framework using the CMS data at $\sqrt{s} = 8$ TeV. This data corresponds to a total L_{int} of 19.7 fb^{-1} . The search is performed in three lepton final state. The dominant backgrounds are estimated using data driven techniques. The cross section and significance of results are extracted by performing a one bin likelihood fit.

The predicted SM inclusive production cross section of tZ at NLO [48] for pp collisions at $\sqrt{s} = 8$ TeV is:

$$\sigma(pp \to tZ) = 236 \pm 11(scale) \pm 11(PDF) fb. \tag{7.1}$$

The top quark in tZ can be either a top or anti-top quark. The scale uncertainty on the predicted cross section is due to the renormalization and factorization scales, and the second one is associated with the choice of parton distribution functions (PDFs). The CTEQ6M set of PDFs [159] are used to determine the predicted cross section.

The predicted cross section of the tZq process for three lepton final state, $\sigma(\mathrm{pp}{\to}\mathrm{t}\ell^+\ell^-\ \mathrm{q})$ $\mathscr{B}(\mathrm{t}{\to}\ \ell\nu\mathrm{b})$, where ℓ denotes a charged lepton (electron, muon, or tau), is calculated to be

$$\sigma(pp \to t\ell^+\ell^- q)\mathscr{B}(t \to \ell\nu b) = 8.2 fb \tag{7.2}$$

with a theoretical uncertainty of less than 10%. The calculation is made in the five flavor scheme, with MADGRAPH5aMC@NLO [95], using the NNPDF (version 2) PDF set [160, 161]. In five flavor scheme, b quarks are considered as coming from the interacting protons. This includes lepton pairs from off-shell Z bosons with an invariant mass $m_{\ell^+\ell^-} > 50$ GeV. The cross section given in Equation 7.2 is used as reference in this search.

7.1 Analysis Strategy

The analysis strategy is discussed in this section. The analysis is performed in three lepton final state, where one lepton comes from leptonic decay of top quark and the remaining two leptons from Z boson. The Z boson leptons should have same flavor and opposite sign. The two lepton flavors i.e electrons and muons are considered. Contributions from the leptonic taus (τ) decay also considered, as they are not specifically excluded. Taus (τ) decaying to electron or muon are considered. With this selection, there are four possible lepton channels i.e all three electrons (eee), all three muons ($\mu\mu\mu$), one muon and two electrons (μee) and two muons one electron ($\mu\mu e$). The analysis is performed using simple and robust cut and count technique. Major backgrounds were estimated using data driven techniques. Cross section and significance are calculated by performing likelihood fit. The fit is performed individually in each of the four lepton channels and for the inclusive channel. Inclusive channel is obtained by performing a fit simultaneously in all four lepton channels.

7.2 Data Samples

The data used in the analysis was collected at $\sqrt{s} = 8$ TeV in 2012 during Run I of CMS operation. This data corresponds to pp collision with L_{int} of 19.7 fb^{-1} . The pp beams collide in bunches, and each bunch cross each other after 50 ns. Only the events where CMS detector was fully operational are used. These events have been certified by the CMS data validation group as "good" for physics analysis. The certified JSON file (given in Table 7.1) with all good runs have been used [162].

TABLE 7.1: JSON file used for 2012 run conditions.

Run period	Golden JSON	Int. Lumi.
2012	`Cert_190456_208686_8TeV_22Jan2013ReRco_Collisions12_JSON.txt`	19.7 fb^{-1}

The di-lepton data streams including double-muon ($\mu\mu$), double-electron (ee) and muon-electron (μe) were used in the analysis. The data samples used are shown in the Table 7.2.

TABLE 7.2: The table shows the di-lepton($\mu\mu$, μe, ee) datasets used in the analysis. Also the period of data taking and the specific runs taken during the period are mentioned in the last two columns of the table.

Data Sets	Data Period	Data Runs
$DoubleMuParked$	$Run2012A - *$	$190456 - 193621$
$DoubleElectron$	$Run2012B - *$	$193834 - 196531$
and	$Run2012C - *$	$198022 - 203742$
$MuEG$	$Run2012D - *$	$203777 - 208686$

where * is 22Jan2013-v1

7.2.1 Trigger Selection and Efficiency

Since the LHC is a hadron machine with an extremely high event rate and high occupancy, the trigger system is a key component of the whole detector. The trigger system of CMS works in two sequential stages: L1 trigger and the HLT trigger as described in Section 3.2.6. The trigger paths that are used in the 8 TeV analysis are based on the online trigger objects with at least two leptons and are summarized in Table 7.3. The di-lepton triggers used in this analysis requires that one or leading lepton should have $p_T > 17$ GeV and the other lepton must satisfy $p_T > 8$ GeV condition.

Double electron triggers have been named as (EE), where double muon triggers are named as (MM) and at least one muon and one electron are named as (ME).

Table 7.3: The table shows the strings of trigger paths for di-lepton dataset.

Process	Trigger path	Type
Di-electrons (ee)	$HLT\_Ele17\_CalIdT\_CalIsoVL\_TkIdVL\_TrkIsoVL$	EE
	$HLT\_Ele8\_CalIdT\_CalIsoVL\_TkIdVL\_TkIsoVL$	EE
Di-muons ($\mu\mu$)	$HLT\_Mu17Mu8$	MM
	$HLT\_Mu17TkMu8$	MM
Muon-Electron(eμ)	$HLT\_Mu17\_El8\_CalIdT\_CalIsoVL\_TrkIdL\_TrkIsoVL$	ME
	$HLT\_Mu8-El17\_CaloIdT\_CalIsoVL\_TrkIdL\_TrkIsoVL$	ME

Since different data streams and triggers are being analyzed simultaneously, there is a possibility of counting same events twice. To avoid this double counting, a trigger logic is adopted as shown in Table 7.4. The procedure to avoid double counting consists of vetoing a given dataset where the events are already selected in another.

Table 7.4: The table shows the trigger logic used in the analysis.

Dataset	Trigger fired
DoubleElectron	EE
DoubleMuon	MM && !EE
MuEG	ME && !EE && ! MM

The trigger selection efficiency is estimated using a sample selected with missing transverse energy (MET) triggers that are weakly correlated to di-lepton triggers. Trigger efficiency is estimated for both data and MC. In data, the di-lepton trigger efficiency is calculated by taking the ratio of events triggered by MET and di-lepton triggers to the events triggered by MET triggers only. For MCs, the trigger efficiency is the ratio of the di-lepton triggered events to the total number of events. The resulting trigger efficiencies are 99% for eee and $ee\mu$, 98% for $\mu\mu\mu$ and 89% for $\mu\mu e$.

7.3 Simulated Samples

The NLO simulated signal sample (tZq) is produced using MadGraph5_aMC@NLO v5.1.3.30 generator [163, 95]. For parton showering and hadronization, Pythia with version 8.212 is used [102, 103]. The dominant production mode t-channel is used with four-flavor scheme. The four- and five-flavor scheme differ in how many partons are included in the proton PDFs. In the four-flavour scheme, there are no b quarks in the proton PDFs and the b quarks needed for single top production in the t-channel are generated dynamically from gluon splitting.

The data with prompt leptons are contaminated by genuine leptons from hadron decays (usually referred to as "non-prompt leptons") and by hadrons or jets misidentified as leptons (usually referred to as "fake leptons"). In addition, non-prompt isolated electrons can arise from the conversion of photons. For simplicity of notation, and given that these background sources are evaluated similarly, by making control samples in data, all such sources are referred to as "non-prompt leptons", or simply "NPL", in this thesis. So simulated samples are divided in two categories. one is prompt lepton backgrounds and other is non-prompt leptons.

7.3.1 Prompt Lepton Backgrounds

The standard model backgrounds which can mimic the signature of the signal process with prompt leptons are:

- WZ and ZZ: Among the diboson processes WZ is the main background in the analysis. It has three charged leptons coming from the leptonic decays of W and Z boson (i.e. Z $\rightarrow \ell^+\ell^-$ and W $\rightarrow \ell^\pm\nu$) in the final state. ZZ decays to four charged leptons, with both Z bosons decaying leptonically, and is a background to our three lepton final state when one of the charged lepton is not detected. All these processes are produced at LO using the MadGraph5_aMC@NLO generator interfaced with Pythia version 6.425.
- $t\bar{t}Z$ and $t\bar{t}W$: The top quark pair production in association with a vector boson (V = W or Z) is a reducible background. The $t\bar{t}Z$ imitates signal signatures when the top pair decays semi-leptonically. It means one top quark decays hadronically (to a b quark and two light jets) and the other leptonically(to a lepton, a neutrino and b quark). The remaining two charged leptons come from Z boson. For $t\bar{t}W$, the top pair decays in di-lepton channel i.e both the quarks decays leptonically. And the third charged lepton is from W boson decay. Both of these samples are produced with MadGraph5_aMC@NLO generator interfaced with Pythia version 6.425 at LO.

7.3.2 Non-prompt Lepton (NPL) backgrounds

The following background processes have different signatures than that of our signal process but they come in our selection because of the non-prompt lepton. The non-prompt lepton simulated samples are used only for validation in the control region, as the contribution from NPL backgrounds is extracted from the data for the final results.

- Drell Yan (Z+Jets): This is a process where Z boson are produced along with multiple jets. This is the dominant process in NPL backgrounds. Two charged leptons from Z are prompt leptons and a third lepton is a NPL that passes our selection. Two Drell Yan samples are produced, one with two charged leptons mass ranging from 10 GeV to 50 GeV and the other with mass greater than 50 GeV. Both of these are produced with MadGraph5_aMC@NLO interfaced with Pythia.
- $t\bar{t}$: Top quark pair production ($t\bar{t}$) has small contribution as a background in our process. The di-lepton decay mode of $t\bar{t}$ is considered, where both top decays leptonically producing a b quark, a charged lepton and neutrinos. The contribution occurs because the third lepton is a non-prompt lepton that passes our selection. The fully leptonic sample is generated with MadGraph5_aMC@NLO.
- WW: The background due to WW has two leptons in the final state and it contributes when a third non-prompt lepton passes our analysis. The contribution from WW background is small in comparison to Z+Jets and $t\bar{t}$.

A full list of MC samples used in this analysis is given in Table 7.5. All of these are Summer12 analysis object data (AOD) simulated MC samples.

7.4 Event Selection

The signatures of tZq signal contains a Z boson, a top quark and an additional ("recoiling") jet. The leptonic decays of both top quark and Z boson lead to three lepton (electrons and/or muons, including those coming from tau leptons) final state process plus a neutrino that goes undetected and is reconstructed from an imbalance in the transverse momentum. The signatures also include bottom quark (b-tag jet) that arises from the hadronization of the b quark produced in the top quark decay. In the final state there is additional jet arising from the hadronization of a light or a charm quark. The pp collision data used in the analysis is collected by CMS detector during the year 2012.

The Particle Flow (PF) event reconstruction algorithm, discussed in detail in Section 4.5, reconstructs global particle candidates like electrons, muons, photons, and charged and neutral hadrons in an event by tracing the flow of particles through the various subdetectors. The combination of the individual subdetector information leads to an improved spatial resolution, energy measurement and identification of particles

Table 7.5: Simulated samples used in this analysis are shown here. Also the total number of events in these samples and the cross section times branching ratio of the samples shown here.

Process	Sample	σ x BR (pb)	No. of Events
tZq	$tZq\_4f\_3leptons - TeV\_amcatnlo\_pythia8\_TuneCUETP8M1$	0.00816	1498500
$t\bar{t}$	$TTJets\_FullLeptMGDecays - 8TeV\_madgraph$	17.2	12119013
$t\bar{t}W$	$TTWJets - 8TeV\_madgraph$	0.232	196046
$t\bar{t}Z$	$TTZJets - 8TeV\_madgraph\_v2$	0.2057	229952
WZ	$WZJetsTo3LNu\_TuneZ2 - 8TeV\_madgraph\_tauola$	1.058	2133868
Zγ	$ZGToLLG - 8TeV\_\_madgraph$	132.6	6588161
ZZ	$ZZTo4e - 8TeV\_powhg\_pythia6$	0.07691	1499093
	$ZZTo4mu - 8TeV\_powhg\_pythia6$	0.07691	1499064
	$ZZTo2e2mu - 8TeV\_powhg\_pythia6$	0.1767	1497445
Z+jets	$DYJetsToLL\_M\_10To50filter - 8TeV\_madgraph$	860.5	37835275
	$DYJetsToLL\_M\_50\_TuneZ2Star - 8TeV\_madgraph\_tarball$	3532.8	30459503
WW	$WW\_TuneZ2star - 8TeV\_pythia6\_tauola$	5.817	10000431

while avoiding double counting. The energy of photons is obtained from ECAL. The electrons are reconstructed by combination of tracker and ECAL information. Hadrons are constructed by combination of informations from the tracker and the matching ECAL and HCAL energy deposits. The neutral hadrons are determined from ECAL and HCAL deposits.

The event is selected if the following selection criteria for primary vertex, electrons, muons, jets, b-tag jet and $\vec{\not{E}}_T^{miss}$ is met.

Primary Vertex (PV): Each event must have at least one good PV as discussed in Section 4.5.2. The vertex reconstruction tries to locate points of pp interactions which are determined by sets of close-by charged particle tracks in the interaction region. The association of tracks to vertices allow to separate tracks belonging to the hard scattering from additional tracks which originate from PU interactions instead. The vertices are ordered by the summed p_T^2 of their associated tracks where the leading vertex is assigned to mark the hard interaction of interest while the others are treated as PU interactions. The vertex will be the PV if it is with the largest sum and satisfies following conditions:

- reconstructed using five or more tracks fitting
- it lies within 24 cm of the nominal interaction point in the z-direction($|z|$)
- within 2 cm in the transverse plane(d_0).

Electron Selection: The electrons selected for the analysis have to fulfil certain criteria. The selected electrons are reconstructed by PF algorithm. These electrons have been reconstructed by GSF tracks. The electron transverse momentum, p_T , has to be larger than 20 GeV, they have to pass the tracker within $|\eta| < 2.4$ and their supercluster must be located outside the region $1.4442 < |\eta_{sc}| < 1.5660$. The later region defines the ECAL endcap-barrel transition region and is excluded due to the large amount of material an electron has to pass, which increases the probability to emit a large amount of energy via bremsstrahlung photons. Selected electron has to be isolated, which means only a small amount of energy is found around the track and cluster of electron in calorimeter. Electron isolation (I_{rel}, given in equation 4.3) is determined by calculating the sum of p_T of all the other reconstructed particles that lie within a cone of fixed radius $\Delta\text{R} = \sqrt{(\Delta\eta)^2 + (\Delta\Phi)^2} = 0.3$ around the electron, correcting for the expected contribution from pileup and dividing the corrected sum by the p_T of the electron. The selected tight electrons must have $I_{rel} < 0.12$.

Furthermore multivariate identification (MVA) discriminator with values between 0.5 and 1.0 is used. Loose electron are identified as GSF electrons having p_T of at least 10 GeV and $|\eta|$ not greater than 2.4. Less stringent criteria on relative isolation is applied for loose electrons with $I_{rel} < 0.20$. The loose electrons are used vetoing the event if there is any additional loose electron present in the event.

Muon Selection : Muons are more clean objects than electrons in the CMS. PF muons are both global and tracker muons. These should have $p_T > 20$ GeV and $|\eta| < 2.4$. The normalised χ^2 to number of degrees of freedom (n_{dof}) ratio for the muon track fitting should be less than 10. Muons are isolated using $I_{\Delta\beta}$ corrected definition of isolation (given in equation 4.1). A muon is said to be isolated, if it satisfy $I_{\Delta\beta}$ < 0.12 using a cone of ΔR = 0.4. To differentiate between prompt and fake muons, several additional identification criteria are applied like impact parameter cuts (d_0 and d_z). Loose muons are PF muons with p_T more than 10 GeV and $|\eta|$ less than 2.4. Loose isolation cut is applied to loose muons with a value $I_{\Delta\beta}$ less than 0.20. The detailed muon selection criteria both for tight and loose muon is shown in the Table 7.6.

TABLE 7.6: The table shows muon selection criteria for tight and loose muon.

Requirements	Tight Muon selection	Loose Muon selection
Muon-id	PF Muon	PF Muon
Which Muon	Global and Tracker	Global or Tracker
$\chi^2/ndof$	< 10	-
$I_{\Delta\beta}$	< 0.12	≤ 0.20
d_0 (cm)	< 0.2	-
d_z (cm)	< 0.5	-
No. of pixel hits	> 0	-
No. of tracker with muon hits	> 5	-
No. of hits in the muon chamber	> 0	-
No. of muon-stations with Muon segments	> 1	-

Jet and b-tag Jet Selection: The selected jets are PF objects reconstructed with anti-kt algorithm using cone size of ΔR = 0.5. All jets are corrected with first three levels of the jet energy corrections discussed in Section 4.8. In order to avoid contributions from misidentified photons, electrons, pions or calorimeter noise, the fractions of charged electromagnetic, neutral electromagnetic and neutral hadronic energies must be less than 0.99, respectively. Further, the jet must have at least two constituents. PF candidates belonging to preselected muon or electron candidates that are loosely isolated are not clustered into jets to prevent double counting. In addition, jets that are within ΔR < 0.5 to the selected tight lepton are ignored in the analysis. The reconstructed jets are corrected by dedicated jet energy scale factors (JES and JER) as explained in Section 4.8. Jets with p_T more than 30 GeV and $|\eta|$ less than 4.5 are selected. The light jet in t-channel process is usually in the forward direction, therefore relaxing $|\eta|$ value to 4.5 improves the signal acceptance. Jets that originate from b quark are reconstructed (tagged) using CSV algorithm [164]. The CSV discriminant is used to distinguish between b jets and non-b jets. These jets are accepted within $|\eta| < 2.4$. For the results presented here, the so-called tight WP for discriminator is used. This tight WP has a b-tagging efficiency of about 50% and a misidentification probability of 0.1%.

Missing Transverse Energy (MET): MET, denoted as $\vec{\not{E}}_T^{miss}$, is reconstructed using PF objects as input (see Section 4.9). For small MET values, Z+jets is a dominant process that is why a cut of MET greater than 30 GeV is applied to reject non-prompt lepton background in certain analysis regions.

7.4.1 Selection Region

The base lepton selection of tZq final state is given as:

- Exactly three tight leptons in the event and vetoing the event if there is any additional loose or tight lepton.
- Two of the three leptons should be same flavor opposite sign (SFOS) leptons reconstructing the Z boson mass within window of [78-102] GeV. And there should be a separation between Z boson leptons i.e $\Delta\Phi_{l^+l^-} > 0.5$. If there is more than one such lepton pair, then the pair with Z mass value close to 91.2 GeV is used.
- Remaining third lepton and $\vec{\not{E}}_T^{miss}$ should reconstruct W boson mass. The lepton and $\vec{\not{E}}_T^{miss}$ should also be separated in $\Delta\Phi_{l\nu_l} > 0.5$.

The analysis is further divided into different regions depending upon the number of light jets and b-tag jets in the event. These regions are given as:

WZ Control Region (CR) : The WZ region has same base lepton selection plus additional jet requirements i.e. zero b-tag jet and one or more light jet in the event. This region is called WZ region because it is enriched with WZ background. No cut on $\vec{\not{E}}_T^{miss}$ is applied in this region.

Preselection (PS) region : This region is defined just for validating our selection. No cut on light jet and b-tag jet is applied here. Since after the b-tag jet selection we have very limited statistics, so this region is used to plot the distributions of different variables in all the four lepton channels and in the combined channel to demonstrate the good data to MC comparison. MET cut of 30 GeV is also applied in this region to suppress NPL background.

Signal Region (SR) : In this region, the target events are those arising from the tZq process, where two jets, one of which a b-tag jet, are expected. In order to increase the signal acceptance for the cases where one additional jet is produced by radiation, signal region is defined as containing events with exactly three leptons and at least 2 jets, one of which must be a b-tag jet. Also the MET should be greater than 30 GeV.

The details of the selection regions and cuts applied are given in Table 7.7.

Table 7.7: Selection criteria for different control regions and the signal region. The three tight leptons selection, including the requirement of two leptons compatible with Z boson, is applied for all regions. The regions differ according to their light jet and b-tagged jet multiplicities and $\vec{\not{E}}_T^{miss}$ cut.

Selection Regions	light Jets	b-tag Jets	$\vec{\not{E}}_T^{miss}$ (GeV)
WZ CR	≥ 1	$= 0$	≥ 0
Pre-selection Region	No requirement	No requirement	≥ 30
Signal Region	≥ 1	=1	≥ 30

7.5 Non-prompt Lepton Estimation

The simulated samples are normalized corresponding to the total recorded integrated luminosity of the data. In this analysis the integrated luminosity adds up to $L_{int} = 19.7\ fb^{-1}$. For each simulated sample a SF is calculated as:

$$SF_{lumi} = \frac{\sigma_{process} \times L_{int}}{N_{events}}. \tag{7.3}$$

where $\sigma_{process}$ is the cross section of a process involved and N_{events} is the number of events in the simulated sample corresponding to the process.

For the backgrounds with limited statistics MC samples, it is not possible to estimate their impact on the analysis. For such backgrounds, a data driven approach is used. In this analysis, the background due to NPL are estimated using data driven technique. In NPL background, it is assumed that leptons coming from Z boson decay are more likely to be real prompt leptons if they pass the Z boson mass cut. The third lepton candidate coming from W boson can be a fake or non-prompt lepton. The events with two fake leptons are considered as negligible. The Z+Jets, $t\bar{t}$ and WW backgrounds populate the signal region if they contain a reconstructed NPL that passes our selection criteria. The contributions from NPL background that survive the event selection are not simulated but estimated from data.

The shape of NPL events per channel is extracted from a sideband WZ control region for which the isolation of the third lepton (coming from W boson) is inverted. It means that selection cuts of light jets, b-tag jets and $\vec{\not{E}}_T^{miss}$ are same as in WZ control region but the isolation of the third lepton is inverted. The selection criteria for NPL data sample is given as:

- Two SFOS selected leptons compatible with Z boson mass
- Isolation of the remaining third lepton is inverted. The isolation is inverted as $I_{rel} > 0.25$ for the electron candidate and $I_{\Delta\beta} > 0.25$ for the muon case.
- light jet and b-tag jet multiplicities are as usual depending upon the selection regions defined in Table 7.7

The distribution of data after the above selection criteria serves as a template to model the shape of non-prompt leptons production. And the normalisation of this data sample is calculated by performing a maximum likelihood fit (described in Section 6.2). The event yield of simulated MC samples after applying this selection criteria is shown in Table 7.8 for both muon and electron case:

Table 7.8: The table shows the percent yield of simulated signal and background samples after the inverted isolation of electrons ($I_{rel} > 0.25$) and muons($I_{\Delta\beta} > 0.25$).

Samples	Muon Iso Inverted Yield (%)	Electron Iso Inverted Yield (%)
$Z + Jets$	77	80
$t\bar{t}$	20	18
WW	1	1
WZ	< 1	< 1
tZq	< 1	< 1
ZZ	< 1	< 1

This shows that Z+Jets sample has dominant contribution in the non-prompt lepton template. The W boson transverse mass (m_T) distribution is used as discriminator in the background-enriched region to estimate the backgrounds related to non-prompt leptons, as well as the dominant WZ background. Both the shape and normalisation of the other backgrounds are estimated from simulation. The normalisation of the non-prompt lepton is estimated by fitting the m_T distribution. The m_T distribution peaks around the W transverse mass for a lepton and $\vec{\not{E}}_T^{miss}$ from a W boson decay, while for non-prompt lepton backgrounds it peaks close to zero and falls rapidly. The amount of non-prompt leptons events in signal and control regions is estimated by fitting the extracted template to data. The distribution of the W boson transverse mass in WZ control region for the case of muon isolation inversion is shown in left of the Fig 7.1 while the resulting non-prompt muon template after its yield has been estimated from the data is shown in right of the Fig 7.1. The same distributions for the case of electron are shown in Fig 7.2.

7.6 Prediction and Data Agreement

Preselection Region: In this section, a preselection is described to demonstrate the agreement between the predictions and data. This region is composed of the full signal selections with no jet requirements. This is done to be able to see the comparison of data and prediction for our selection. No requirement on jet selection is applied because after asking for exactly one b-tag jet in the event we have very low statistics. This results in a region with more events of the signal region and enhanced backgrounds, both being useful to help illustrate the agreement of the predictions and data. The WZ is a dominant background and thus the reconstructed masses of Z and W bosons are good variables for comparison. The figure 7.3 shows combined

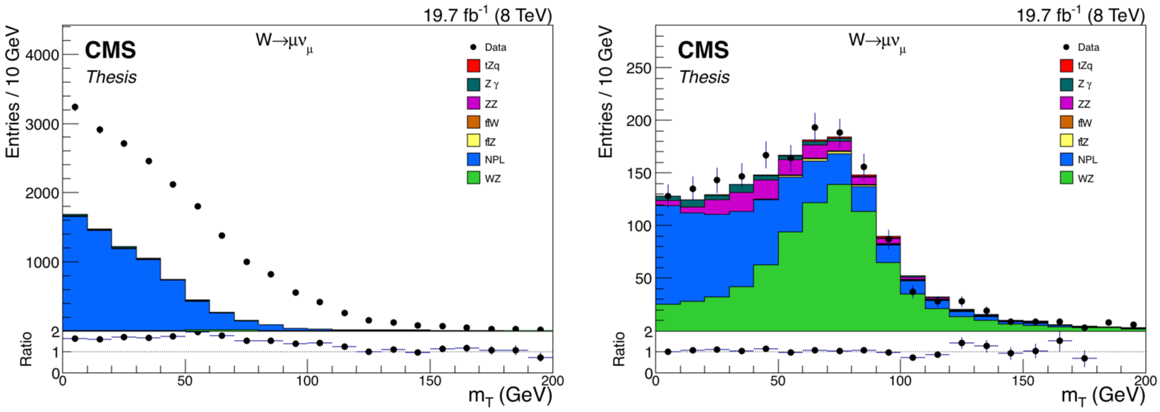


FIGURE 7.1: Distributions of the W boson transverse mass (m_T) for the muon case in WZ control region: (left) third lepton isolation inverted for extracting the non-prompt lepton template; (right) resulting distribution after scaling the non-prompt lepton background with the extracted SF from the ML fit. The m_T distribution in the left peaks at zero (instead of on-shell W boson mass), showing that the region is enriched with NPL.

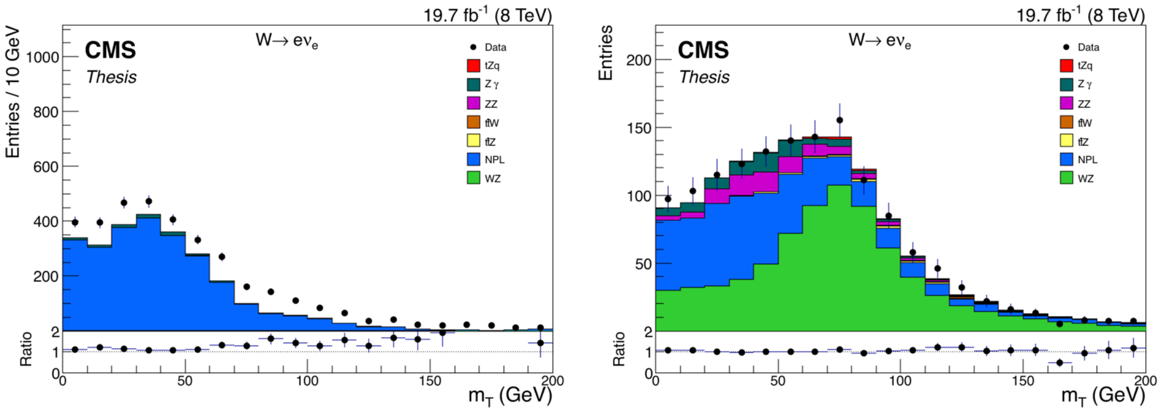


FIGURE 7.2: Distributions of the transverse W boson mass for the electron case in WZ control region: (left) third lepton isolation inverted for extracting the non-prompt lepton template; (right) resulting distribution after scaling the non-prompt lepton background with the extracted SF from the ML fit. The m_T distribution in the left peaks around zero (instead of on-shell W boson mass), showing that the region is enriched with NPL.

plots (all four lepton channels summed up) of reconstructed Z boson invariant mass, W transverse mass (m$_T$), $\vec{\not{E}}_T^{miss}$ and p_T of Z boson. These plots shows that WZ is a dominant background in this region. The channel wise plots of Z boson invariant mass and W boson transverse mass are shown in Figure 7.4 and Figure 7.5 respectively. The distribution of reconstructed Z boson p_T is shown in Figure 7.6. The $\vec{\not{E}}_T^{miss}$ distribution is shown in Figure 7.7. Additionally, in the preselection region, other distributions like the jet multiplicity and p_T of the leading jet is plotted in Figure 7.8. Other distributions such as the muon p_T is shown in left of Figure 7.9 and electron p_T is plotted in the right of Figure 7.9. The simulated prompt leptons backgrounds are normalized to data luminosity while the normalization of non-prompt lepton backgrounds is extracted from the data driven estimation of NPL as described in Section 7.5.

Signal Region: The signal region is defined by requiring exactly one b-tag jet and at least one light jet in the event. With one b-tag jet requirement, there is limited statistics but still the comparison between data and simulation is within the uncertainties. The distributions for Z boson mass and W transverse mass are show in Figure 7.10 and Figure 7.11 respectively. Jet multiplicity is also shown in Figure 7.12. As can be seen from the plots, the dominant backgrounds are due to $t\bar{t}Z$ and WZ.

7.7 Systematics Uncertainties

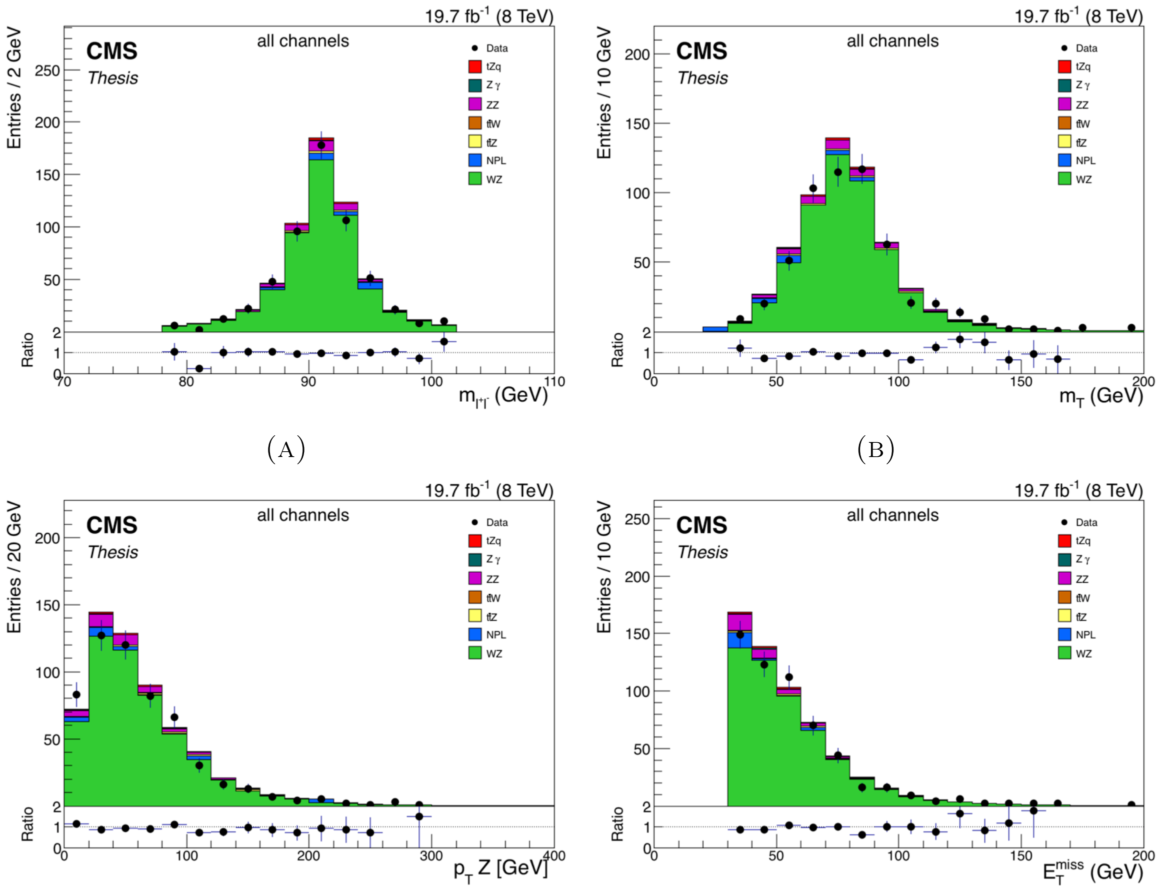
In high energy physics experiments, estimation of both statistical and systematic uncertainties play a key role. Statistical uncertainties are due to stochastic fluctuations arising from the fact that number of events in a sample are finite. Systematics uncertainties arises due to the nature of the measurement apparatus, inaccuracies in signal and background modelling and assumptions made by the experimenter. The main sources of systematics uncertainties studied in the analysis are discussed in the following sections.

7.7.1 Luminosity

The uncertainty on L$_{int}$ is estimated to be ± 2.6 at $\sqrt{s}$ =8 TeV. This measurement is done using pixel cluster counting method [165]. The uncertainty due to luminosity affects event yield of all simulation processes.

7.7.2 Pileup

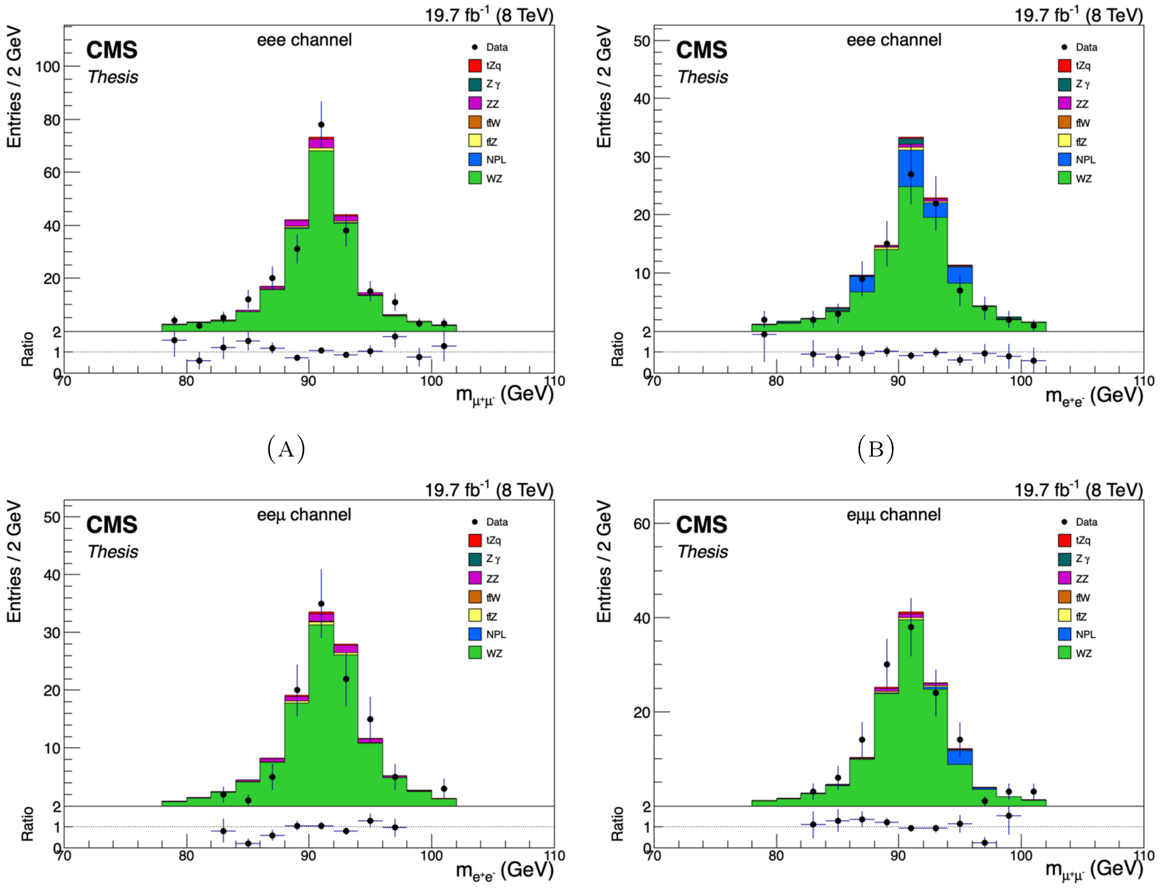
The systematic shift in the number of PU events is calculated by varying the minimum bias cross section by ±5 % [166].



(A) (B)

(C) (D)

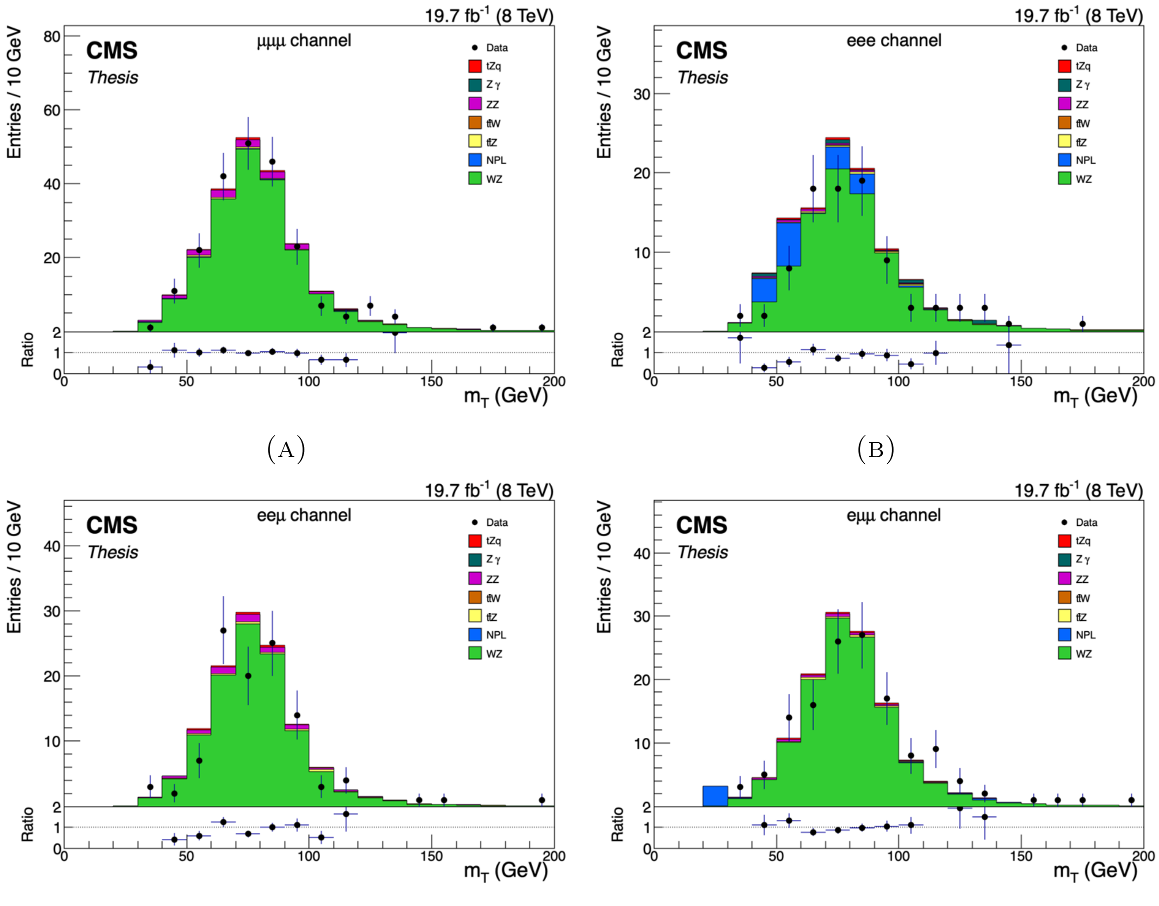
FIGURE 7.3: The distributions of Z boson invariant mass at (A), transverse mass of W boson at (B), p_T of Z boson at (C) and $\not\vec{E}_T^{miss}$ at (D) are shown in the pre-selection region with the contribution of all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) summed up. WZ is the dominant background in green with smaller contribution from other backgrounds such as ZZ and NPL. The signal process tZq has some contribution in red at top of other processes.The data is visible as black dots. The bottom panel of the plot displays data to MC ratio.



(A) (B)

(C) (D)

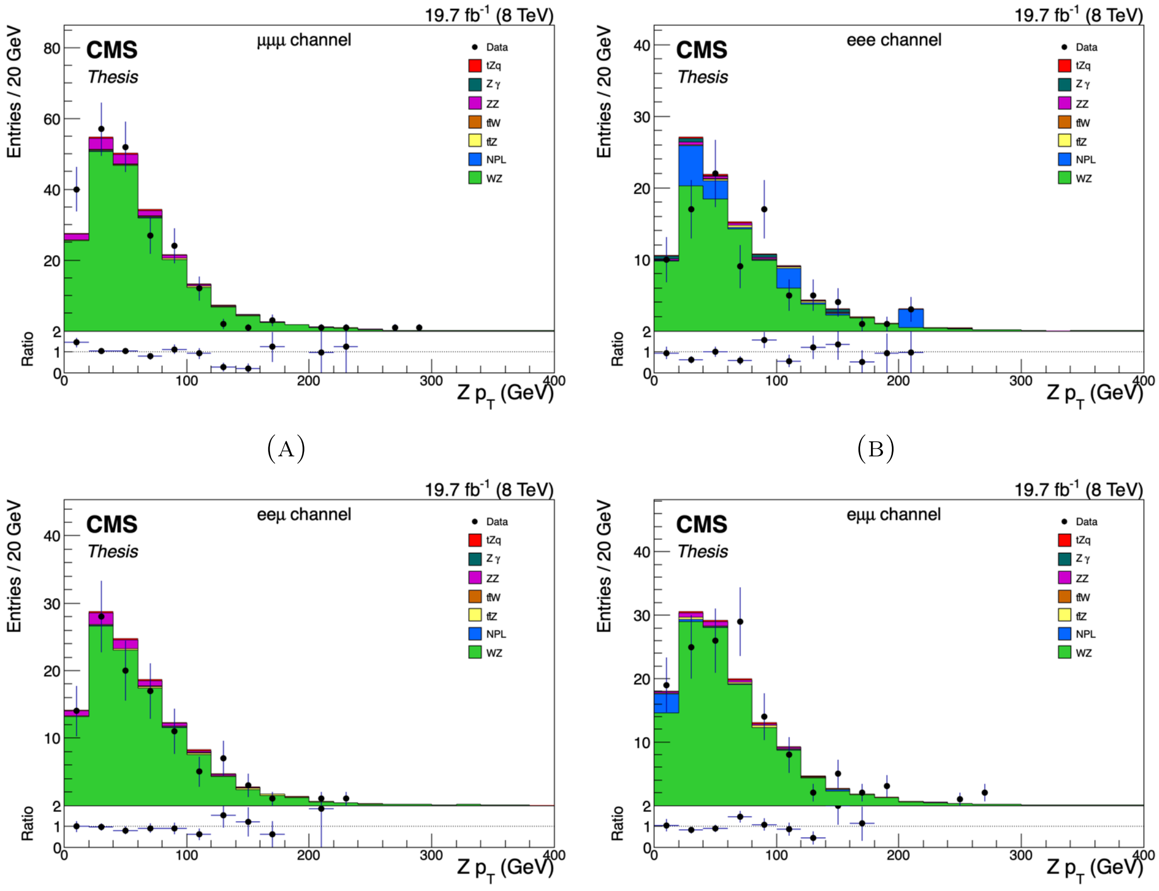
Figure 7.4: Distributions of Z boson invariant mass for $\mu\mu\mu$ channel at (A), eee channel at (B), $ee\mu$ channel at (C) and $\mu\mu e$ channel at (D) are shown in the pre-selection region. WZ is the dominant background in all channels with smaller contribution from other backgrounds such as ZZ and NPL. NPL background has more contribution in eee and $\mu\mu e$ channels where third lepton is electron as compared to $\mu\mu\mu$ and $ee\mu$ channels where third lepton is muon.



(A) (B)

(C) (D)

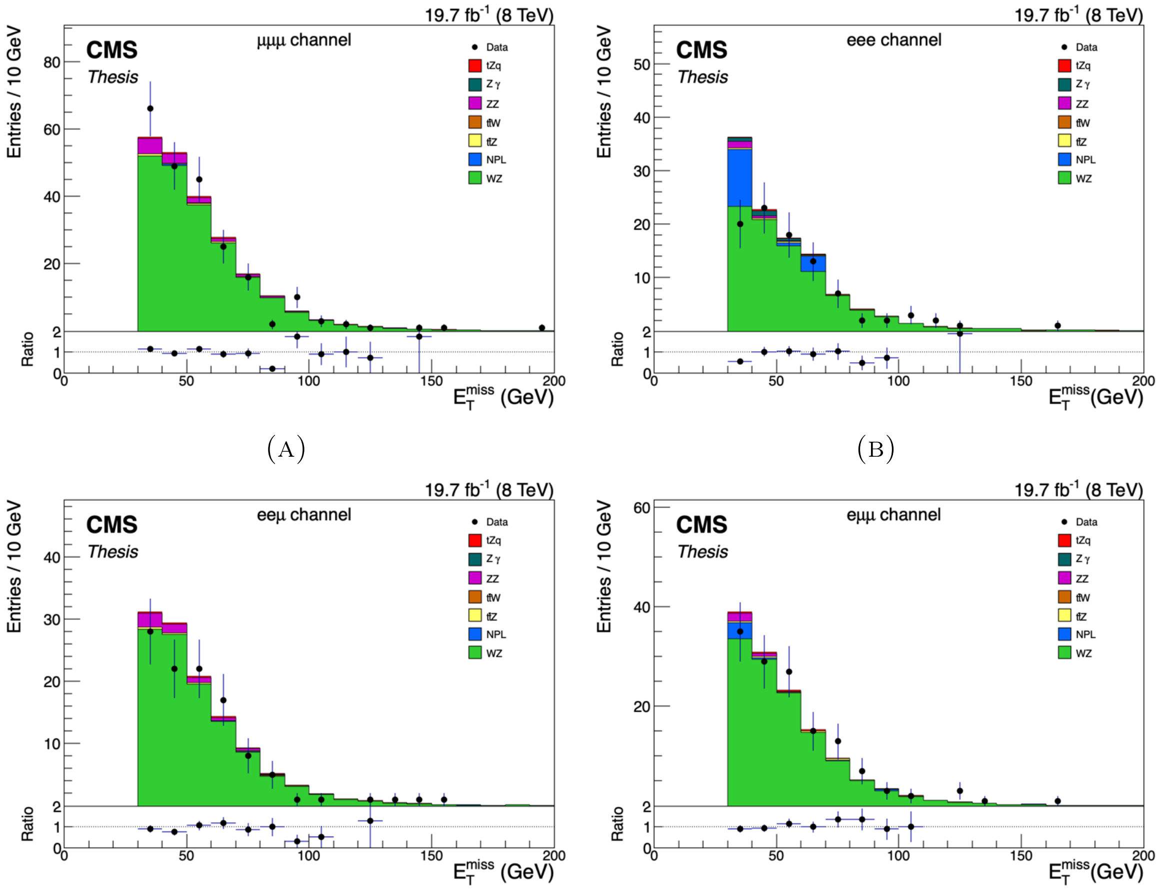
Figure 7.5: Distributions of W boson transverse mass for $\mu\mu\mu$ channel at (A), eee channel at (B), $ee\mu$ channel at (C) and $\mu\mu e$ channel at (D) are shown in the pre-selection region. WZ is the dominant background in all channels with smaller contribution from other backgrounds such as ZZ and NPL. NPL background has more contribution in eee and $\mu\mu e$ channels where third lepton is electron as compared to $\mu\mu\mu$ and $ee\mu$ channels where third lepton is muon.



(A) (B)

(C) (D)

Figure 7.6: Distributions of Z boson transverse momentum for $\mu\mu\mu$ channel at (A), eee channel at (B), $ee\mu$ channel at (C) and $\mu\mu e$ channel at (D) are shown in the pre-selection region. WZ is the dominant background in all channels. NPL background has more contribution in eee and $\mu\mu e$ channels where third lepton is electron as compared to $\mu\mu\mu$ and $ee\mu$ channels where third lepton is muon.



(A) (B) (C) (D)

FIGURE 7.7: Distributions of $\vec{\not{E}}_T^{miss}$ for $\mu\mu\mu$ channel at (A), eee channel at (B), $ee\mu$ channel at (C) and $\mu\mu e$ channel at (D) are shown in the pre-selection region. WZ is the dominant background in all channels. NPL background has more contribution in eee and $\mu\mu e$ channels where third lepton is electron as compared to $\mu\mu\mu$ and $ee\mu$ channels where third lepton is muon.

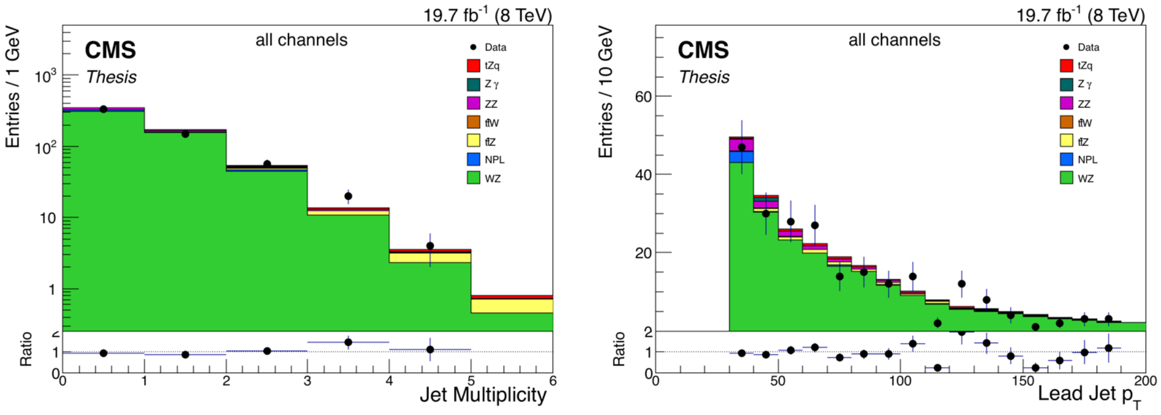


FIGURE 7.8: Distributions jet multiplicity (left) and p_T distribution for leading jet (right) in the pre-selection region with the contribution of all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) summed up. WZ is the dominant background. The data is visible as black dots. The bottom panel of the plot displays data to MC ratio.

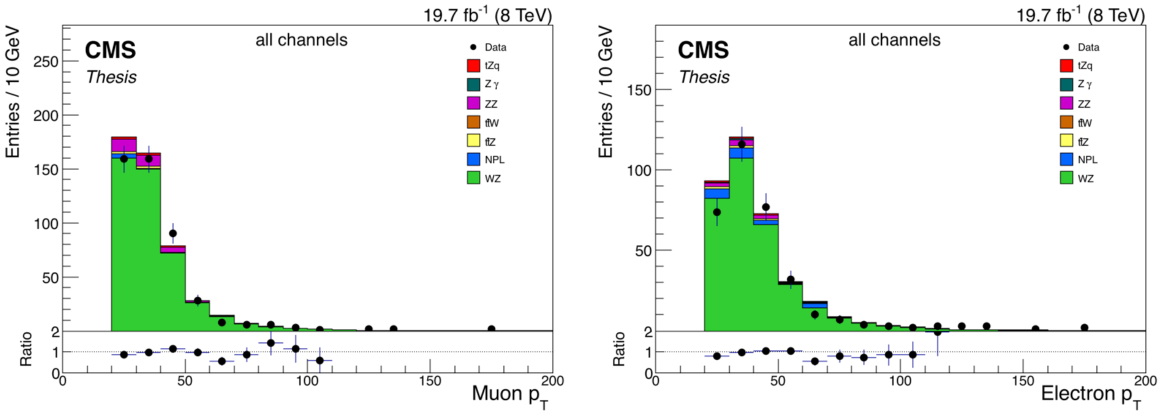


FIGURE 7.9: The p_T distributions of lead muon (left) and lead electron (right) in the pre-selection region with the contribution of all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) summed up. The data is visible as black dots. The bottom panel of the plot displays data to MC ratio.

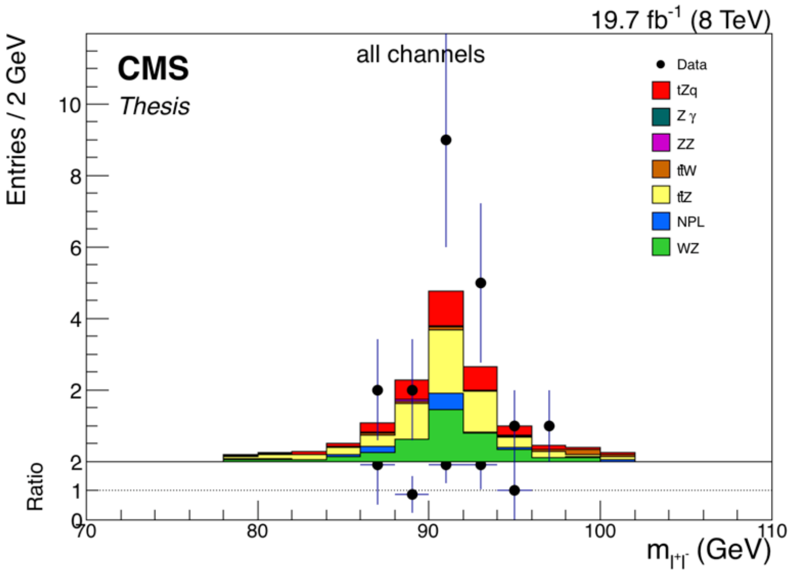


FIGURE 7.10: Distributions of Z boson invariant mass in the signal region for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) summed up. The distribution shows nice peaks around Z boson pole mass with the contributions from tZq signal in red. The contribution from dominant backgrounds NPL (blue), WZ (green) and $t\bar{t}Z$ (yellow) is also visible. The data is visible as black dots. The bottom panel of the plot displays data to MC ratio.

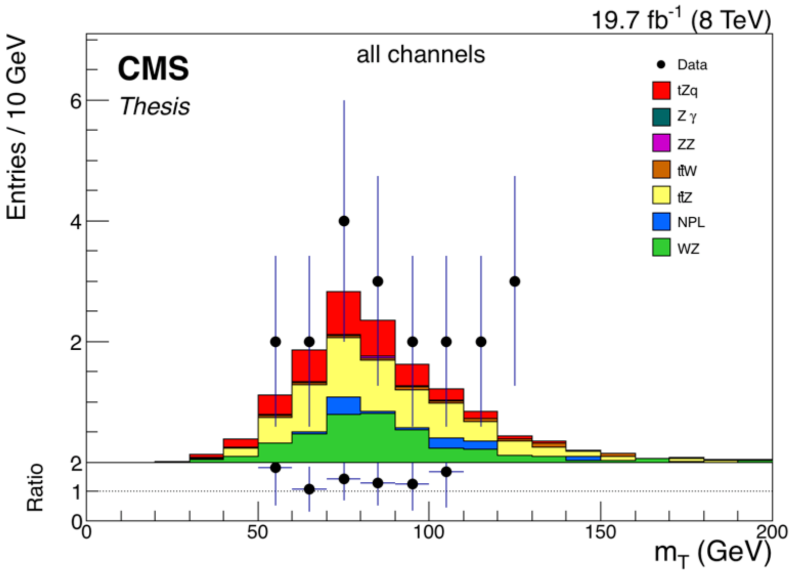


FIGURE 7.11: Distributions of the W boson transverse mass for the four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) summed up in the signal region. The distribution shows nice peaks around Z mass with the contributions from tZq signal in red. The contribution from dominant backgrounds NPL (blue), WZ (green) and $t\bar{t}Z$ (yellow) is also visible. The data is visible as black dots. The bottom panel of the plot displays data to MC ratio.

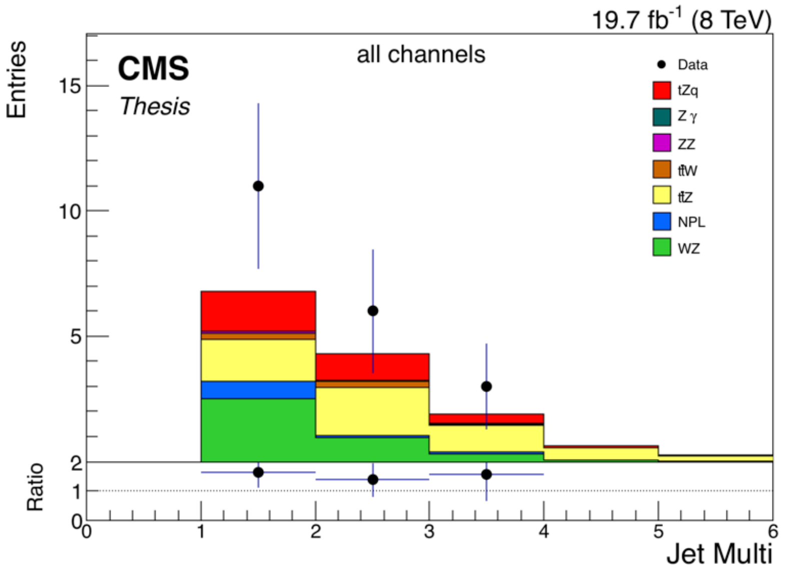


FIGURE 7.12: Distribution for light jet multiplicity for the four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) summed up in the signal region.

7.7.3 Lepton Trigger, Reconstruction and Identification Efficiencies

To ensure that the efficiency of di-lepton triggers observed in data is properly reproduced, a set of data to simulation corrections is applied to all simulated events; likewise, an additional set of corrections (p_T and $|\eta|$ dependent) is used to ensure that the efficiency for reconstructing and identifying leptons observed in the data is correctly reproduced in the simulation. The corrections are varied by their corresponding uncertainties, which amounts to about 4% per event for the trigger selection and 2% per event for the lepton selection.

7.7.4 Jet Energy Scale (JES)

There are discrepancies in the energies of detector level and generator level jets. Jet energy corrections are needed to remove these discrepancies. Therefore, correction factors (p_T - η dependant) are applied to each simulated jet in the event. The effect of JES uncertainty is then calculated by varying the correction factor by $\pm 1\sigma$ for up and and down variations. This change in jet p_T effects the total momenta in transverse plane effecting the MET as well. Therefore, this change is transversed to MET also.

7.7.5 Jet Energy Resolution (JER)

The jets in the event are smeared up or down with respect to central value measured in the data. Smearing depends upon jet p_T and η. The recommended prescription for determining JER is outlined in [167]. Using signal Monte Carlo, the prescription was applied as follows:

- each reconstructed jet is matched to the closest generator jet within a cone of ΔR < 0.5,
- if a match is found the p_T of the reconstructed jet is scaled by $p_T \rightarrow \max[0, p_T^{gen} + c \times (p_T - p_T^{gen})]$ where p_T^{gen} is matched generator jet transverse momentum, and c is the data/MC scale factor between the measured and expected particle flow jet resolution in Table 7.9,
- if no match is found a gaussian having center at one and width at c is used to smear the momentum.

7.7.6 B-tagging Uncertainty

B-tagging is crucial for analysis involving top quark. In CMS, b-tag and vertexing (BTV) group is responsible for providing scale factors to take care of discrepancies (if any) between simulation and data. These scale factors are calculated using control samples [168]. B-tagging decision of each jet is updated on event by event basis depending upon the SF.

7.7.7 Unclustered $\vec{\not{E}}_T^{miss}$

This is the energy which is not reconstructed as lepton and neither is part of jet reconstruction. This energy remains unclustered and is known as unclustered MET. It is calculated by varying the nominal energy by ± 10%.

7.7.8 Background Normalization

The normalization of the NPL background processes are estimated from data while performing the final fit. The normalisation uncertainties in the backgrounds estimated from simulation are taken as ±30%.

7.7.9 Theory Uncertainties

These are the uncertainties associated with the production of simulated samples. Though the standard model is very accurate and precise measurements of many of its parameters have been done, there are still some gaps that can alter the simulated samples. Varying these parameters during production of the samples can effect analysis.

- **PDFs:** The nominal PDF sets used for the analyses described in this chapter are quoted in Section 7.3. In order to compute the corresponding uncertainty, simulated events are re-weighted by using the eigenvalues associated to each PDF set. The corresponding variations are summed in quadrature and the results are compared with the nominal prediction. Uncertainties estimated from different PDF sets are also compared and the largest uncertainty is taken.
- **Physics Process Modelling:** The renormalisation and factorisation scales used in the simulated samples, are multiplied or divided by a factor of two, and the corresponding variations are considered systematic uncertainties. The procedure

Table 7.9: Data/MC scale factors used in determining the Jet Energy Resolution.

Jet Pseudorapidivity	Scale Factor
0.0 - 0.5	1.052
0.5 - 1.1	1.057
1.1 - 1.7	1.096
1.7 - 2.3	1.134
2.3 - 5.0	1.288

used in PYTHIA to match the partons in the matrix-element calculation with those in the parton showering includes a number of scale thresholds. These are varied in the simulated WZ sample and the resulting variation is taken as the associated systematic uncertainty.

The systematic uncertainties are taken as nuisance parameters in the final fit to extract tZq cross section. There is 100% correlation between the four lepton channels for every systematics except for NPL where the $\mu\mu\mu$ and $ee\mu$ channels are 100% correlated and the $\mu\mu e$ and eee channels are 100% correlated. The nominal and varied distributions of some of these systematics are given in Appendix A for the tZq signal sample and dominant backgrounds process.

7.8 Results

The event yields of data and MC in the signal region is given in Table 7.10. All the MCs are normalized to data luminosity and scaled by different SFs due to lepton efficiency, lepton trigger, pileup, jet energy corrections and b-tagging. The data yields are shown along side the MC yields to help demonstrate the contributions of individual backgrounds. The tZq cross section is calculated by performing a maximum likelihood fit to the number of events counted in signal region. The likelihood function takes event yields of the data and simulated processes as input. RooStats package [169] used for calculating cross section and significance of the process. The analysis is performed as binned counting experiment. The binned likelihood function (see Section 6.2 for details) is given as

$$\mathscr{L}(data|\mu,\theta) = \frac{(\mu s_i(\theta) + b_i(\theta))^{N_i}}{N_i!} \exp^{-(\mu s_i(\theta)+b_i(\theta))}, \tag{7.4}$$

where N_i is the observed number of events and $s_i(\theta)$ and $b_i(\theta)$ are the expected signal and background events respectively in the i^{th} bin. All the systematics uncertainties, described in Section 7.7, are taken into account as nuisance parameters (θ).

The fit is performed to maximises $\mathscr{L}(data|\mu,\theta)$, from which the measured cross section σ is extracted according to its relation to the signal strength (μ) as:

$$\mu = \frac{\sigma(\ell^+\ell^- q)}{\sigma^{SM}(\ell^+\ell^- q)}. \tag{7.5}$$

The systematic uncertainties explained in Section 7.7 are set as free parameters in the fit. Each systematics have different effect on the result, therefore, the impact of each systematic uncertainty on μ is calculated. The impact is defined as the shift $\Delta\mu$ when θ is fixed and brought to its $\pm 1\sigma$ post-fit values while other systematics are treated as usual. The impact is given by:

$$\Delta\mu = \hat{\mu}(\hat{\theta} \pm \Delta_\theta) - \hat{\mu}(\hat{\theta}) \tag{7.6}$$

Table 7.10: The various column represents post-fit event yield of tZq signal sample and dominant backgrounds for data and MC simulation in the signal region for all four lepton channels of tZq process.

	total	$\mu\mu\mu$	eee	$ee\mu$	$e\mu\mu$
WZ	3.82 ± 0.21	1.38 ± 0.12	0.64 ± 0.08	0.81 ± 0.15	0.99 ± 0.10
$t\bar{t}Z$	5.23 ± 0.32	1.71 ± 0.18	1.18 ± 0.15	1.15 ± 0.15	1.19 ± 0.15
$t\bar{t}W$	0.50 ± 0.11	0.12 ± 0.05	0.10 ± 0.05	0.12 ± 0.10	0.16 ± 0.06
ZZ	0.14 ± 0.01	0.06 ± 0.01	0.02 ± 0.00	0.03 ± 0.05	0.03 ± 0.01
NPL	2.57 ± 0.25	0.64 ± 0.07	0.86 ± 0.22	0.43 ± 0.06	0.64 ± 0.07
Total Bkg MC	12.02 ± 0.47	3.92 ± 0.23	2.79 ± 0.28	2.54 ± 0.20	3.00 ± 0.20
Signal tZq	3.40 ± 0.01	1.22 ± 0.01	0.63 ± 0.00	0.70 ± 0.00	0.85 ± 0.01
Data	20.00 ± 4.47	7.00 ± 2.65	5.00 ± 2.24	3.00 ± 1.73	5.00 ± 2.24

The impact these uncertainties have on signal strength is reported in Table 7.11.

Table 7.11: Impact of nuisance parameters and their correlation with the signal strength.

Source	$\Delta\ \mu$	$\rho(\mu)$
Fake Electron	0.08	−0.16
Fake Muon	0.05	−0.11
JER	0.01	−0.03
JES	0.03	−0.07
Lepton Eff.	0.01	−0.09
Unclustered $\not{\vec{E}}_T^{miss}$	0.02	−0.14
PU	0.01	−0.08
Trigger	0.02	−0.05
Btagging	0.01	−0.13
Luminosity	0.01	−0.04
WZ rate	0.02	+0.03
$t\bar{t}Z$ rate	0.01	−0.09
WZ Scale and Matching	0.04	+0.12
$t\bar{t}Z$ Scale	0.02	+0.05

The fit is performed for each of the four lepton channels ($\mu\mu\mu$, $\mu\mu e$, $ee\mu$ and eee) independently to calculate the cross section for that channel. For the combined channel cross section measurement, a simultaneous fit to data in all four channels is performed. The cross section numbers calculated through these individual independent lepton channel fits and simultaneous fit in all for channel are shown in Table 7.12. The systematics and statistics uncertainties are also reported on the cross section.

The tZq measured and SM predicted cross sections are compared in Figure 7.13. The SM predicted cross section with its theoretical uncertainty is shown in blue and the black dots represent measured cross section for the individual and combined channel. The uncertainties on the measured cross section are categorised as statistical and

TABLE 7.12: The measured cross sections for each of the four lepton channels and the combined channel for the tZq process.

Channel	Cross section (fb)
eee	29^{+32}_{-24} (stat)$^{+8}_{-7}$(syst)
$ee\mu$	6^{+23}_{-6} (stat)$^{+4}_{-3}(syst)$
$\mu\mu e$	19^{+24}_{-18} (stat)±5 (syst)
$\mu\mu\mu$	20^{+11}_{-9} (stat)$^{+4}_{-3}$(syst)
Combined fit	18^{+11}_{-9} (stat) ±4 (syst)

systematical ones. Statistical uncertainties are shown as the red line on measured cross section value while systematics are overlaid as green line in the figure. The measured and predicted SM cross sections are in agreement within the uncertainties. The statistical uncertainties are dominant over systematical uncertainties because tZq is very rare SM and with complete 2012 CMS data, only a few events were identified as tZq event.

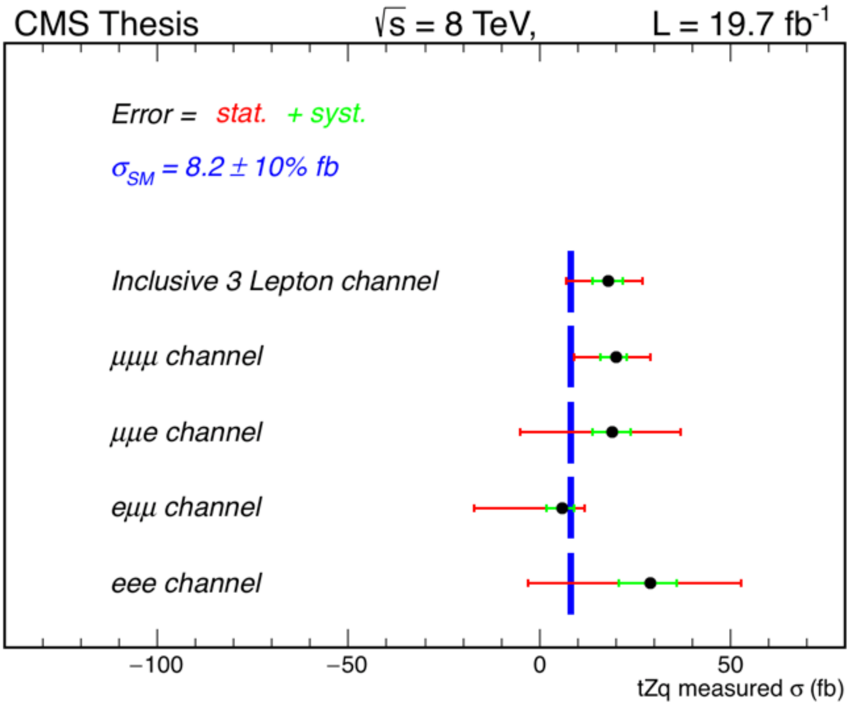


FIGURE 7.13: Measured cross section of tZq process in each of four lepton channels ($\mu\mu\mu$, $\mu\mu e$, $ee\mu$ and eee) and the combined channel at L$_{int}$ of 19.7 fb^{-1} recorded by CMS in the year 2012. The standard model expected cross section with 10 % theoretical uncertainty (thickness of line) is shown in blue line. Black dot represents value of measured cross section for each individual three lepton channel and combined channel. Statistical error is on the measured cross section is shown in red while systematical error is in green line.

The expected and observed significance of cross section is also reported for individual lepton channels and combined channel in Table 7.13.

Table 7.13: The expected and measured significance for each of four lepton lepton channels and the combined channel of tZq cross section (σ) are given. The 68 % (1 σ) and 95 % (2 σ) confidence level on either side of the expected mean significance are also shown.

Channel	Expected Significance			Observed Significance
-	Mean	68% CL	95% CL	-
$\mu\mu\mu$	0.659 ±0.033	[0-1.451]	[0-2.482]	1.352
eee	0.462 ±0.032	[0-1.154]	[0-2.183]	1.202
$\mu\mu e$	0.614 ±0.032	[0-1.352]	[0-2.281]	1.041
$ee\mu$	0.532 ±0.021	[0-1.162]	[0-2.203]	0.322
Combined Channel	0.811 ± 0.041	[0-1.592]	[0-2.681]	1.803

The expected significance is calculated using inputs yields from simulation only. Toy simulations are produced to calculate expected significance assuming the signal strength value $\mu = 1$. Also the 68 % and 95 % confidence levels are reported on either side of the expected mean. The observed significance is calculated by taking into account the data. The significance for the combined channel is 1.8 standard deviation, while the expected significance is 0.8 standard deviations, with a 68% CL range of [0-1.59] and 95% CL range [0-2.68]. The signal significance for the $\mu\mu\mu$ channel is observed to be 1.35± 0.03 standard deviations, while the expected significance is 0.8 standard deviations, with a 68% CL range of [0-1.45] and 95% CL range [0-2.68].

7.9 Conclusions

The associated production cross section of a single top quark with a Z boson was measured using data from pp collisions at 8 TeV collected by the CMS experiment, corresponding to an integrated luminosity of 19.7 fb^{-1} . The measurement uses events containing three charged leptons in the final state. The cross section is measured to be $\sigma(t\ell^+\ell^-q) = 18^{+11}_{-9}$(stat) ± 4 (syst) fb, for $m_{\ell^+\ell^-} > 50$ GeV. where ℓ stands for electrons, muons and τ leptons. This value is compatible with the NLO standard model prediction of 8.2 fb with a theoretical uncertainty of less than 10 %. The tZq cross section measurement is found to be in agreement with SM with an observed (expected) significance of 1.8 (0.8) standard deviations.

7.10 Shape Analysis Results

The results reported in Section 7.8 are calculated by simple and robust cut and count technique. These results are reported in the paper [170] as a cross check to the shape analysis result. In the shape analysis a simultaneous fit on the BDT outputs in the signal region is performed. The number of excess events in the background only hypothesis is compared to the SM expectation for tZq production in order to measure the cross section. The combined measured cross section of tZq signal is found to be $10^{+8}_{-7} fb$. The corresponding observed and expected significances are 2.4 and 1.8 standard deviations, respectively, with the expected significance having a one standard deviation range of [0.4 - 2.7] at 68 % CL.

Chapter 8

The tZq cross section Measurement at $\sqrt{s}$ =13 TeV

The cross section of single top quark in association with a Z boson has been measured at $\sqrt{s}$=8 TeV using CMS data with a significance of 1.8 σ in previous chapter. In this chapter, the cross section measurement of single top quark production in association with a Z boson in three lepton(e, μ) final state is presented using data corresponding to L$_{int}$ of 36 fb^{-1} recorded by the CMS detector at an unprecedented centre of mass energy of 13 TeV. The cross section is measured using a simple cut and count technique. The cross section is extracted by performing a simultaneous likelihood fit in signal and background enriched control regions. Theoretical cross section for tZq calculated at NLO is $\sigma(tZq \rightarrow Wb\ell^{+}\ell^{-}q) = 94.2 \pm 3.1$ fb calculated for pp collisions at $\sqrt{s} = 13$ TeV. The NLO cross section is calculated, considering only the leptonic decays of Z bosons (to electrons, muons, or tau leptons, generically denoted by ℓ), using the Monte Carlo (MC) generator MADGRAPH5 aMC@NLO 2.2.2 [95]. The calculation is performed using the NNPDF 3.0 set of parton distribution functions (PDFs) [161] in the five flaor scheme. The cross section calculation requires that lepton pair from off-shell Z boson has invariant mass $m_{\ell^{+}\ell^{-}} > 30$ GeV. The expected cross section is,

$$\sigma^{SM}(t\ell^{+}\ell^{-}q) = 94.2^{+1.9}_{-1.8}(scale) \pm 2.5(PDF)fb. \tag{8.1}$$

The scale and PDF uncertainties are estimated by changing the QCD renormalization and factorization scales by factors of 0.5 and 2, and by using the 68 % confidence level (CL) uncertainty on the NNPDF3.0 PDF set. This cross section is used as a reference in this analysis. Another calculation, including all Z boson decays, gives a compatible cross section when the branching fraction to charged leptons is taken into account [48].

8.1 Analysis Strategy

Like 8 TeV analysis in the previous chapter, the measurement of production cross section of tZq in three lepton final state is reported by using the simple cut and count technique. The signatures of tZq production consists of a single top production in t-channel, a Z boson and an additional recoiling jet emitted at pseudorapidivity $|\eta| < 4.5$. The Z boson decays in two opposite sign electrons(e^+e^-) or muons ($\mu^+\mu^-$). The top quark decays via a W boson and a b quark, while W boson decays to a neutrino and an electron or muon. With the three leptons in the final state, there are four possible leptonic combinations $\mu\mu\mu$, eee, $\mu\mu e$ and $ee\mu$. There is also small contribution from tau leptons (τ) decaying into electrons or muons.

The tZq cross section is obtained by performing a maximum likelihood fit simultaneously on the signal and the background enriched control regions. After the baseline selection of three leptons, background enriched regions are chosen according to jet multiplicities. Like 8 TeV analysis, the major backgrounds to tZq process are due to WZ, $t\bar{t}Z$ and NPL. The three regions used in the analysis are the signal tZq region and the background enriched control regions (CRs) due to background WZ and $t\bar{t}Z$. These regions are defined in Table 8.1. Background enriched control regions are defined to constrain backgrounds due to NPL, WZ and $t\bar{t}Z$ processes. The maximum likelihood fit is performed simultaneously in signal region and background enriched CRs.

8.2 Datasets and MC Samples Analysed

The tZq cross section measurement is performed using the full Run II data collected by the CMS during 2016 at 13 TeV corresponding to 36 fb^{-1} of integrated luminosity. The primary datasets, consisting of single and double lepton data streams, used in the analysis are shown in Table 8.2. For single lepton data streams, at least one lepton is ensured to be part of the event. Similarly for the double lepton data streams, at least two leptons are ensured to be present in the event. In 8 TeV analysis, only double lepton data streams were used. In this analysis, single lepton data stream is used in addition to double lepton because it might add some additional events in this statistically limited rare process. The results reported in this chapter are based on the $03Feb2017ReReco$ datasets. Events of interest are selected by using the *golden json* file given in Table 8.3. This json file is applied to ensure high data quality with good functioning of all subdetectors. This data corresponds to an integrated luminosity of 35.9 fb^{-1} with uncertainty on luminosity to be 2.6%.

Data is compared to MC simulated events. MC simulated events are used extensively in the measurement to evaluate detector resolution, the efficiencies and acceptance, and

Table 8.1: Selection criteria defining control regions and signal region. These regions are divided according to jet multiplicities after the baseline selection of three leptons.

Regions	untag jets	btag jets
WZ Control Region	>= 1	=0
$t\bar{t}Z$ Control Region	>= 1	>= 2
tZq Signal Region	= 1 or 2	= 1

to estimate the contribution from background processes that have topologies similar to the tri-lepton tZq final state.

The tZq signal samples are generated at NLO precision using the MADGRAPH5 aMC@NLO 2.2.2 package. The two main background processes, WZ+jets and top quark pair production in association with vector bosons ($t\bar{t}Z$ and $t\bar{t}W$), are also simulated with the same event generator, with upto one additional hadronic jet at NLO. For other minor background such as ZZ, NLO generators MADGRAPH5 aMC@NLO and POWHEG v2.0 were used [171, 172, 173, 174]. The PDF set NNPDF 3.0 is used in all generators. The simulated samples are interfaced to PYTHIA 8.205 [175] with the CUETP8M1 tune [176] for the parton shower and hadronization. The detector response is simulated using the GEANT4 package [177]. The events are simulated in final states that include decays to electrons, muons, and tau leptons. A top quark mass of 172.5 GeV is assumed. Multiple minimum bias events generated with PYTHIA are added to each simulated event to mimic the presence of PU, with weights that reproduce the measured distribution of the number of PU vertices. The list of simulated MC samples is given in Table 8.4.

Table 8.2: Primary Datasets used in this analysis are listed, with their corresponding Runs taken during collisions. PD is an abbreviation for SingleElectron, SingleMuon, MuonEG, DoubleEG or DoubleMuon.

Datasets	Data Runs
/PD/Run2016B-*_ver2-v2/MINIAOD	272007 - 275376
/PD/Run2016C-*-v1/MINIAOD	275657 - 276283
/PD/Run2016D-*-v1/MINIAOD	276315 - 276811
/PD/Run2016E-*-v1/MINIAOD	276831 - 277420
/PD/Run2016F-*-v1/MINIAOD	277772 - 278808
/PD/Run2016G-*-v1/MINIAOD	278820 - 280385
/PD/Run2016H-*_ver2-v1/MINIAOD	280919 - 284044
/PD/Run2016H-*_ver3-v1/MINIAOD	
where PD is SingleElectron, SingleMuon, MuonEG, DoubleEG, DoubleMuon and * stands for 03Feb2017	

Table 8.3: Golden JSON file used for 2016 run conditions.

Run period	Golden JSON	Int. Lumi.
2016	`ReReco/Final/Cert_271036-284044_13TeV_23Sep2016ReReco_Collisions16_JSON.txt`	35.9 fb^{-1}

8.3 Trigger Strategy

The events in this analysis are selected online by two stage CMS triggers, the so called L1 and HLT triggers as discussed in Section 3.2.6. The L1 trigger is pure hardware

Table 8.4: MC samples used in the analysis are listed, with their corresponding cross sections.

Dataset	$\sigma\times$ BR (pb)
tZq_ll_4f_13TeV-amcatnlo-pythia8	0.09418
WZTo3LNu_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	4.42965
ZZTo4L_13TeV_powheg_pythia8	0.0374
TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	0.2043
TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.2529
WWTo2L2Nu_13TeV-powheg	3.84
DYJetsToLL_M-10to50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	18610
DYJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	6025.2
TT_TuneCUETP8M2T4_13TeV-powheg-pythia8	831.76

trigger and cut down the rate from 40 MHz to 100 kHz. Events which pass the L1 decision are transferred to DAQ, a complete and more accurate reconstruction is performed at the HLT level. Events are selected using single, double and tri-lepton triggers. The trigger paths which are basically lists of strings that identifies the list of selection performed at HLT level are given in Table 8.5.

Table 8.5: High Level Trigger triggers used are shown here.

HLT path	Trigger type
HLT_Mu8_DiEle12_CaloIdL_TrackIdL_v	ME
HLT_Mu8_TrkIsoVVL_Ele17_CaloIdL_TrackIdL_IsoVL_v	ME
HLT_DiMu9_Ele9_CaloIdL_TrackIdL_v	ME
HLT_Mu17_TrkIsoVVL_Ele12_aloIdL_TrackIdL_IsoVL_v	ME
HLT_Ele16_Ele12_Ele8_CaloIdL_TrackIdL_v	EE
HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v	EE
HLT_TripleMu_12_10_5_v	MM
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v	MM
HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v	MM
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_v	MM
HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_v	MM
HLT_soMu22_v	M
HLT_IsoTkMu22_v	M
HLT_Ele27_WPTight_Gsf_v	E

A simple 'OR' logic is followed for the online event selection for all triggers in the MC samples and an event is accepted if it passes the prescribed HLT paths. For each dataset an extra logic is applied to avoid double counting of events in the data as described in Table 8.6.

Table 8.6: HLT trigger logic used in the analysis.

Dataset	Trigger fired
MuonEG	EM
DoubleMu	MM && !EM
DoubleEG	EE && !MM && !EM
SingleMuon	M && !EE && !MM && !EM
SingleElectron	E && !M && !EE && !MM && !EM

8.3.1 Trigger Efficiency

The trigger efficiency is calculated for both data and simulation. For simulation, the trigger efficiency is straight forward as there is no trigger bias in simulated events. To measure the trigger efficiency in data, the procedure described in Ref. [178] is followed. An unbiased sample collected using MET triggers. Only the events that passes following trigger paths were recorded on this unbiased sample.

- HLT_PFHT300_PFMET110_v
- HLT_MET200_v
- HLT_PFMET300_v
- HLT_PFMET120_IDTight_v
- HLT_PFMET170_HBHECleaned_v

These triggers are completely uncorrelated with respect to the lepton triggers given in Table 8.5. Trigger efficiency studies in MC were based on the the signal tZq sample, corrected for PU and luminosity scale factors. For this analysis, the events that passes 3 lepton selection are used without applying any jet requirement. The following equations are used to calculate the trigger efficiencies in data and MC

$$\epsilon_{Data} = \frac{\text{No. of events passing Lepton and } \vec{\not{E}}_T^{miss} \text{ triggers}}{\text{No. of events passing } \vec{\not{E}}_T^{miss} \text{ triggers}},$$

$$\epsilon_{MC} = \frac{\text{No. of events passing Lepton triggers}}{\text{Total No. of events}}.$$

The efficiencies and corresponding scale factors are given in Table 8.7. Since the trigger efficiency is measured to be nearly 100% for both data and MC, so no scale factor is applied. Assigning a large uncertainty to the trigger efficiency based on the MET dataset statistics is over conservative. We, therefore, account for an uncertainty of

1 % on the trigger selection for $ee\mu$ and $\mu\mu\mu$ final states, and 2 % for the eee and $\mu\mu e$ final states.

Table 8.7: Trigger efficiencies after the three lepton selection, for data and MC. The uncertainties displayed on data and MC are statistical only. The uncertainties of the last column were estimated using the largest of the asymmetric uncertainties of the individual efficiencies.

Channel	ϵ_{Data}	ϵ_{MC}	$\epsilon_{Data}/\epsilon_{MC}$
$\mu\mu\mu$	$1.0000^{+0.0000}_{-0.0169}$	$0.9998^{+0.0001}_{-0.0002}$	1.000 ± 0.017
$\mu\mu e$	$0.9907^{+0.0077}_{-0.0210}$	$0.9992^{+0.0003}_{-0.0004}$	0.991 ± 0.021
$ee\mu$	$0.9915^{+0.0071}_{-0.0194}$	$0.9988^{+0.0003}_{-0.0004}$	0.993 ± 0.019
eee	$0.9833^{+0.0138}_{-0.0373}$	$0.9974^{+0.0006}_{-0.0008}$	0.986 ± 0.037

8.4 Physics Object Selection

The tZq cross section measurement is performed in three lepton final state with two leptons coming from the Z boson decay while the third lepton from top quark decay via a W boson and b quark. Therefore in this analysis exactly three leptons (e or μ) are selected and the event is vetoed if there is additional fourth lepton (e, μ) is the event. Two out of the three leptons must be same flavored and oppositely charged compatible to Z boson mass. In addition to three leptons, event should have at least two jets one of which must be a b-tag jet and MET from W boson decay. The selection criteria of each of these objects (leptons, jets, and neutrinos) is described in the following.

8.4.1 Muon Selection

In this analysis particle flow muons (PFMuons) have been used from the collection of global muons. In an event, selection criteria for muons follow the recommendation by the MUON POG [179]. In this analysis, two working points are considered for the identification of muons, the loose working point and the tight working point as described in [180]. Only the muons with $|\eta| < 2.4$ and having p_T greater than 25 GeV are considered. The detailed selection criteria for tight as well as loose muons is illustrated in Table 8.8.

The isolation is calculated using the transverse energy sum of the neutral particles, charged particles and photons inside a cone with respect to the direction of the target lepton as described in Section 4.6. The isolation selection is based on the ratio of the above energy sum to the target muon p_T . Tight muon selection is applied together with a cut of 0.15 on particle flow combined isolation. The working point of tight isolation value is recommended by the CMS MUON POG [179]. The recommended cone size used in this analysis is $\Delta R = 0.4$.

To minimize the contribution from PU, the $\Delta\beta$ correction method has been used. Within the cone size, the charged candidates which were not compatible with the primary vertex have been removed from the sum. Furthermore, the contribution of neutral hadrons from PU is estimated to be half the one coming from charged candidates, and this quantity is subtracted from the total. Finally, the scalar sum is divided by the muon p_T. The general formula for the PU corrected isolation is as follow,

$$I_{rel} = \frac{\sum p_T^{charge-hadron} + max(\sum p_T^{neutral-hadron} + \sum p_T^{\gamma} - 0.5 * \sum p_T^{PU}, 0.)}{p_T^{\mu}}$$

where $p_T^{charge-hadron}$ is the sum of the p_T of the charged hadrons, $\sum p_T^{neutral-hadron}$ is the sum of transverse energies of the neutral hadrons, $\sum p_T^{\gamma}$ is the sum of the transverse energy of photons and $\sum p_T^{PU}$ is the sum of transverse momenta of charged particle in the cone of interest but which were not originated from the primary vertex i.e. for pileup corrections.

TABLE 8.8: Summary of muon identification cuts used in the analysis [179]. In table, there are two columns: "tight WP" and "loose WP". The "Tight" column refers to the criterion implemented for signal selection while the "loose" refers to the loose muon selection which is useful for background prediction.

Cut	Tight WP	Loose WP
Particle Flow Muon ID	isPFMuon	isPFMuon
Transverse momentum	> 25 GeV	$>= 10$ GeV
$\|\eta\|$	< 2.4	$<= 2.4$
I_{rel}	0.15	$<= 0.25$
Cone size	0.4	0.4
Cut based ID	tight	veto

8.4.2 Electron Selection

The electrons used in this analysis are PF electrons. PF electrons are reconstructed using GSF [119] tracks reconstruction of electrons. Electrons are identified on the basis of "cut-based" ID as defined by the EGamma POG [181]. Tight and loose cut-based IDs are used in the analysis. All the recommended cuts for tight electron selection together with those used to veto additional electrons are mentioned in Table 8.9, also with additional cuts on the distance of track from primary vertex in the transverse plane ($d_{xy} < 0.05$), distance along z-axis ($d_z < 0.1$) and significance of the impact parameter SIP$_{3D} < 4$.

For electrons, the ρ corrected isolation with cone size of 0.3 is used to reduce the pileup effects. Inside the cone size, the expected pileup energy is calculated from mean energy density per area of pileup impurity (ρ), that is calculated event by event, and the effective area (A_{eff}). The relative electron isolation has been calculated as

$$I_{rel} = \frac{\sum p_T^{charge-hadron} + max(\sum p_T^{neutral-hadron} + \sum p_T^{\gamma} - \rho * A_{eff}, 0.)}{p_T^{electron}}.$$

The value of I_{rel} is set to 0.0588 (0.0571) for tight electrons in the barrel (endcap) regions, and 0.175 (0.159) for veto electrons in the barrel (endcap) regions.

TABLE 8.9: Summary of cut based electron IDs used in the analysis. The criterion used for signal selection is referred in column "tight" while the column "loose" refers to the selection of electrons for background prediction.

Cut	Tight ID	Loose ID
Cut based ID	tight	loose
p_T	>25 GeV	$> =$ 10 GeV
$\|\eta\|$	$<$ 2.4	$< =$ 2.5
$\|\Delta\eta_{in}^{seed}\|$	$<$ 0.00308	$<$ 0.00477
$\|\Delta\phi_{in}\|$	$<$ 0.0816	$<$ 0.222
$\|\sigma_{i\eta i\eta}\|$	<0.00998	$<$ 0.011
H/E	<0.0414	<0.298
$\|1/E - 1/p\|$	<0.0129	<0.241
conversion veto	Yes	Yes
relIso (endcap)	0.0571	0.159
relIso (barrel)	0.0588	0.175

8.4.3 Jets Selection

The jets are reconstructed by using anti-kT jet-clustering algorithm as discussed in Section 4.8. FastJet package [182, 183] with cone size of 0.4 is used. The charged hadrons that does not originate from PV, are excluded from jet-clustering PF candidates. The vertex with the maximum sum of p_T^2 of its constituent tracks, is selected as the PV. In addition, a multivariate discriminator is applied to distinguish between jets coming from the PV and jets coming from pileup vertices. Also the jets are required to be separated from lepton by requiring a cut on $\Delta R_{(\ell,jet)}$ defined by:

$$\Delta R_{(\ell,jet)} = \sqrt{(\eta^{\ell} - \eta^{jet})^2 - (\phi^{\ell} - \phi^{jet})^2}.$$

Loose jet IDs defined in Section 4.8 used to select the jets. The selection criteria of jets is summarised as:

- PF candidate with anti-kT algorithm
- Loose PF Jet ID
- $p_T > 30$ GeV

- $|\eta| < 4.5$
- $\Delta R_{(\ell,jet)} > 0.4$

8.4.4 B-tag Jet Selection

The b-tagged jets are identified using CSV algorithm [184, 185]. In this analysis, the "tight" working point for the CSV output discriminator is used. The tight WP has approximately 45% efficiency for tagging jets with b quarks and 1% efficiency to tag jets with only light quarks or gluons. B-tag jet selection cuts are given as:

- CSV discriminator $>$ 0.94 (tight WP)
- $|\eta|< 2.4$
- $p_T > 30$ GeV
- $\Delta R_{(\ell,jet)} > 0.4$

8.4.5 Missing Transverse Energy ($\vec{\not{E}}_T^{miss}$)

$\vec{\not{E}}_T^{miss}$ is the summed transverse momenta of produced neutrinos which are not detected by the CMS detector. $\vec{\not{E}}_T^{miss}$ reconstruction has been discussed in Section 4.9 .

8.4.6 Pileup re-weighting

The simulated datasets have been generated using the best knowledge of pileup in the foreseen pp collisions but still lack the reality. For this reason, the multiplicity of primary vertices in simulation are different from the actual ones in the data. The simulations are therefore corrected to match the data. The pileup is estimated using the minimum bias cross section of 69 mb [186]. The multiplicity of PV is used to show the effects of PU before and after re-weighting as shown in Figure 8.1. The scale factors derived by using the pileup reweighting procedure are then applied to all distributions as a weight.

8.5 Non-prompt Lepton Background

Like 8 TeV analysis, the dominant background contributions to the tZq process are emerging from WZ, $t\bar{t}Z$, and NPL processes. In case of WZ and $t\bar{t}Z$, prompt leptons are in the final state. Therefore both of these are irreducible backgrounds to the tZq process. Background due to NPL is a reducible background where jets are faking as leptons. These non-prompt leptons are either from heavy hadrons decay or jets with

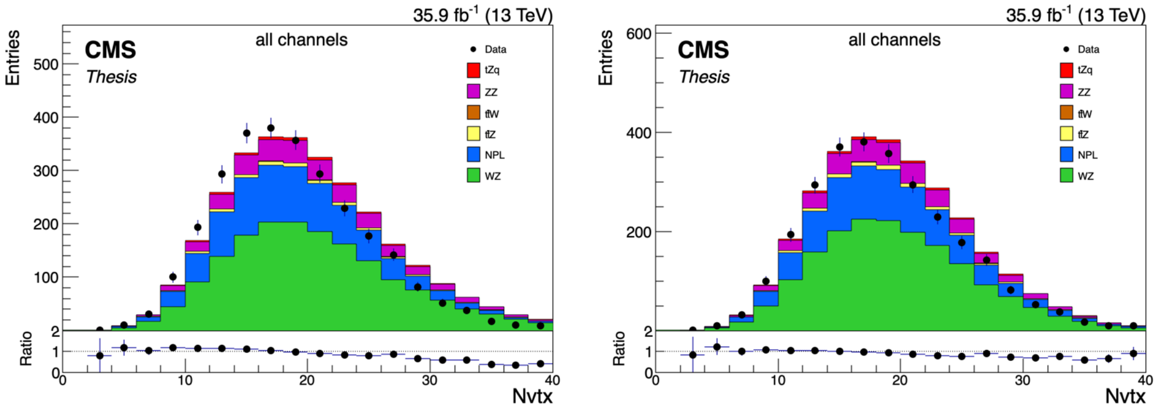


FIGURE 8.1: Plots of the primary vertex multiplicities in the WZ control region combined in all four channels. The number of reconstructed primary vertices are plotted before (left) and after (right) pileup reweighting.

such high EM fraction that they propagate through HCAL and leave tracks in the muon system. These tracks mimic lepton signatures. NPL contribution is mainly from Z+jets background process. We assume that the Z boson leptons are prompt leptons, while the third lepton which comes from W boson decay can be a non-prompt-lepton. There is also little contribution from $t\bar{t}$ and WW processes to the NPL background.

Data driven estimation of NPL is done due to limited MC statistics. Therefore, the NPL sample is taken from the data by selecting three tight leptons but the isolation of the third lepton is inverted. For the case when third lepton is muon the isolation is inverted with a value greater than 0.25 (i.e. muon iso > 0.25). For the case of the third electron, the isolation is inverted to a value greater than 0.175 and less than 1 (i.e. $0.175 <$ electron iso < 1).

For the purpose of non-prompt lepton sample normalization, a separate Z+jets control region is defined by requiring three tight leptons (e or μ) and at least one light jet in the event. The isolation of the third lepton (lepton from W) is inverted. Z+jets CR is described in Table 8.10. MCs used for Z+jets, $t\bar{t}$ and WW are shown in Table 8.11. NPL sample is taken from data, these MCs are used only for data to MC comparison in Z+jets CR and to get idea of individual contribution sof Z+Jets, $t\bar{t}Z$ and WW in the NPL sample.

The distribution for W boson transverse mass is shown in Figure 8.2. The distribution is plotted for the case when third lepton is muon (left of Figure 8.2) and for the case when third lepton is electron (right of Figure 8.2).

The distribution peaks around zero showing that it is fake lepton enriched region. Similarly the $\vec{\not{E}}_T^{miss}$ is plotted for the case of electron (left of Figure 8.3) and muon (right of Figure 8.3). The major contribution is from the Z+jets background and relative smaller contributions from $t\bar{t}$ and WW backgrounds. Also the shape comparison of dominated background and data in NPL enriched region is done in Figure 8.4. Major background Z+jets and data both have same shapes in transverse mass distribution peaking at zero. The distributions for η and p_T of the third lepton are also shown in Figure 8.5 and Figure 8.6 respectively. The validity of Z+jets region is evident because the Z+jets process is dominant while other NPL processes are suppressed by the selection criteria applied in this control region.

TABLE 8.10: Selection criteria for Z+jets CR to estimate contribution due to the fake leptons.

Non-prompt Lepton Sample Selection
Exactly three tight leptons.
RelIso > 0.25 for third muon coming from W boson.
RelIso > 0.175 and RelIso < 1.0 for third electron coming from W boson.
At least one light jet.
no b-tag jet.

TABLE 8.11: MCs used for Z+jets, $t\bar{t}$ and NPL for data to MC comparison in Z+jets CR.

Dataset	$\sigma\times$ BR (pb)
$DYJetsToLL\_M-10to50\_TuneCUETP8M1\_13TeV-madgraphMLM-pythia8$	18610
$DYJetsToLL\_M-50\_TuneCUETP8M1\_13TeV-madgraphMLM-pythia8$	6025.2
$TT\_TuneCUETP8M2T4\_13TeV-powheg-pythia8$	831.76
$WWTo2L2Nu\_13TeV-powheg$	3.84

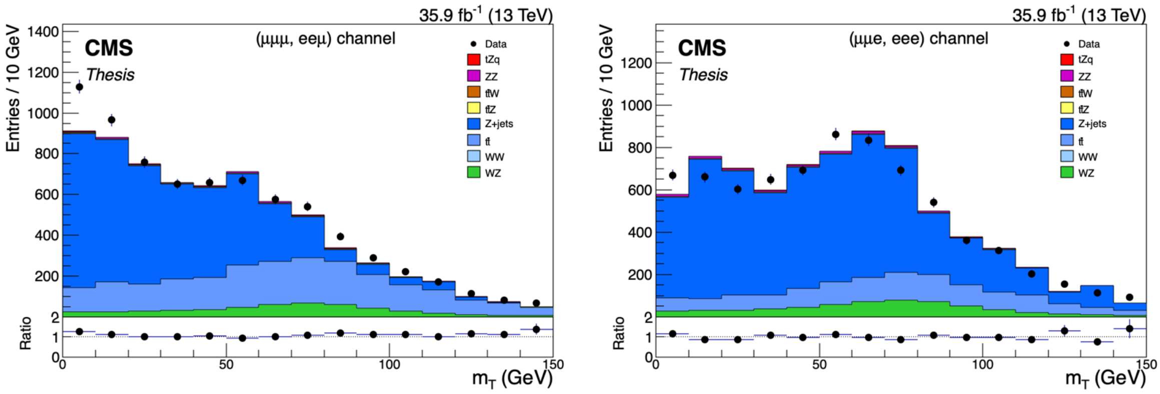


FIGURE 8.2: The W transverse mass in the Z+jets CR is plotted. Left: The distribution is for the case when third lepton is muon ($ee\mu$ and $\mu\mu\mu$ channels summed up). Right: The distribution is for the case when third lepton is electron ($\mu\mu e$ and eee channels summed up). The dominant contribution in NPL sample is from Z+jets background with smaller contributions from $t\bar{t}$ and WW processes..

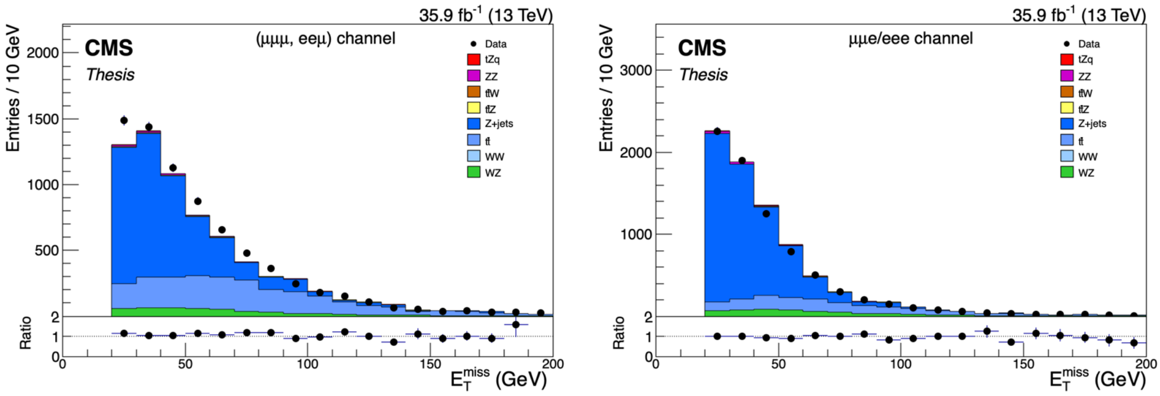


FIGURE 8.3: The distribution of $\vec{\not\!E}_T^{miss}$ in the Z+jets CR. Left: The distribution is for the case when third lepton is muon ($ee\mu$ and $\mu\mu\mu$ channels summed up). Right: The distribution is for the case when third lepton is electron ($\mu\mu e$ and eee channels summed up). The dominant contribution is from Z+Jets sample with smaller contributions from $t\bar{t}$ and WW processes.

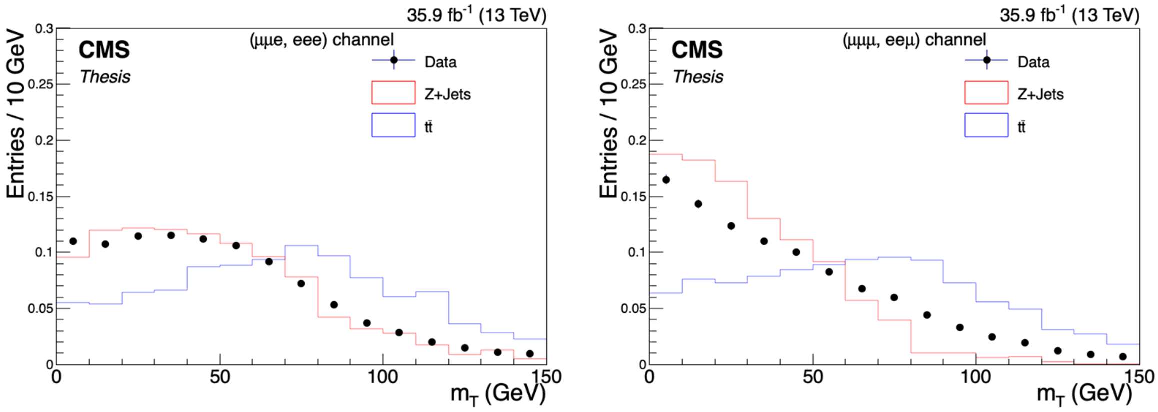


FIGURE 8.4: Shape comparison of fake MCs (Z+jets, $t\bar{t}$) and data when third lepton is an electron (left) and the case when the third lepton is a muon (right). The distributions are made in the Z+jets control region. The histograms are normalized to unit area to compare shapes.

8.5.1 Normalization of the NPL Background

MC simulated samples are normalized to the total recorded L_{int} of the data as described in Equation 7.3. All the prompt lepton background are normalized this way. But NPL background sample is taken from data, and therefore the normalization is extracted from data as well. The normalization is obtained by using the W boson transverse mass templates in all four (eee, $\mu\mu\mu$, $\mu\mu e$, $ee\mu$) channels in WZ CR region. A maximum likelihood fit is performed independently in all four channels with WZ as a signal. Background due to NPL is left unconstrained in the fit. The fit gives four normalization scale factors i.e. one scale factor for each channel. NPL background samples are scaled

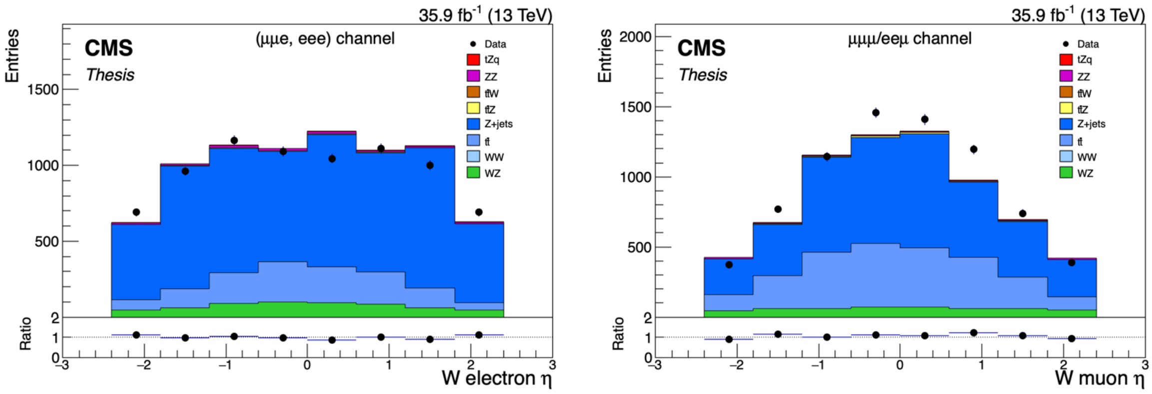


FIGURE 8.5: The distribution of lepton η when the third lepton is an electron (left) and the case when the third lepton is a muon (right). The distribution are obtained with an inverted isolation for the third lepton coming from W decay.

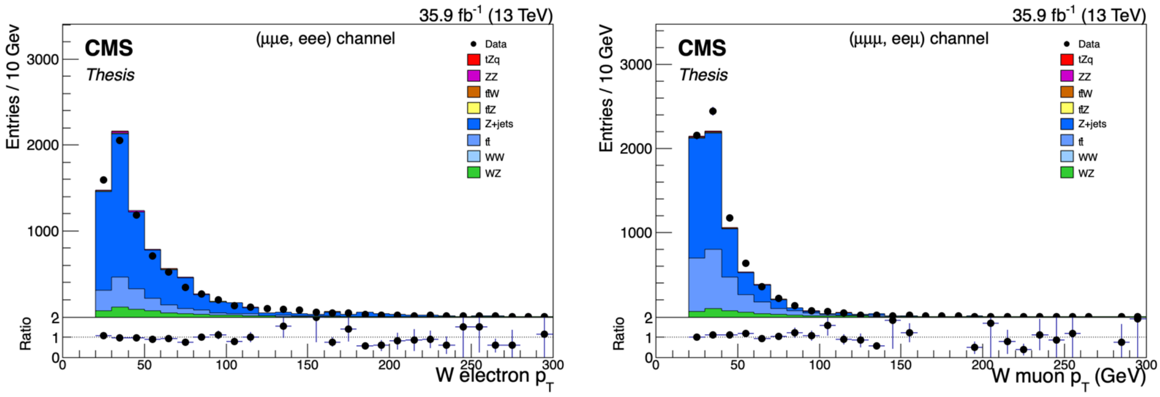


FIGURE 8.6: The distribution of lepton p_T when the third lepton is an electron (left) and the case when the third lepton is a muon (right). The distribution are obtained with an inverted isolation for the third lepton coming from W decay.

with these scales factors throughout the analysis. Same normalization scale factors are used in the tZq SR and $t\bar{t}Z$ CR. A final fit is performed simultaneously in WZ CR, $t\bar{t}Z$ CR and tZq signal region. Since these scale factors are obtained in zero b-tag jet region (WZ) and are propagated in one b-tag (tZq signal region) and a two b-tag region ($t\bar{t}Z$ control region), therefore a conservative uncertainty of 50 % is taken on these SFs in the final fit.

Without this prefit SFs, dealing with NPL background is very difficult. The W boson transverse mass distribution corrected with these scale factors and uncorrected background for eee, $\mu\mu\mu$, $\mu\mu e$ and $ee\mu$ channels are shown in Figure 8.7, Figure 8.8, Figure 8.9 and Figure 8.10 respectively.

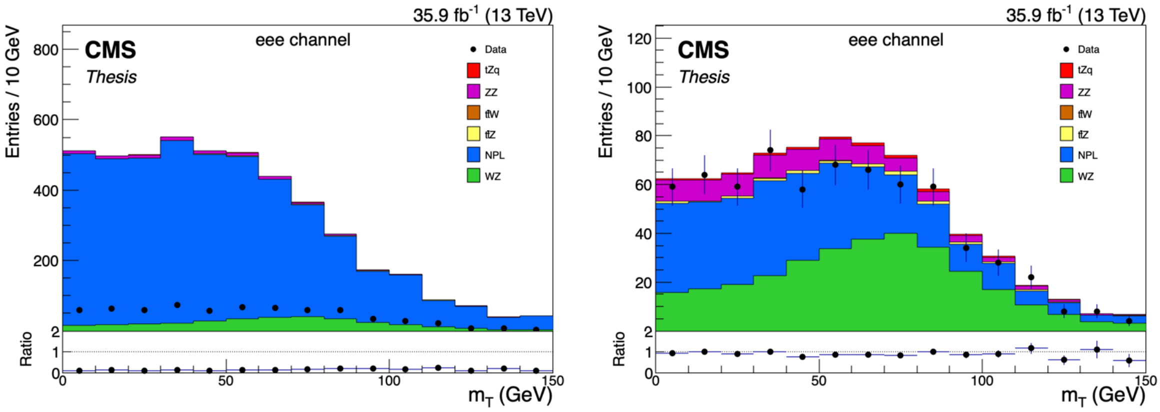


FIGURE 8.7: Distributions of the W boson transverse mass in the WZ CR for *eee* channel. Left: without normalization of NPL background. Right: with normalization of NPL background.

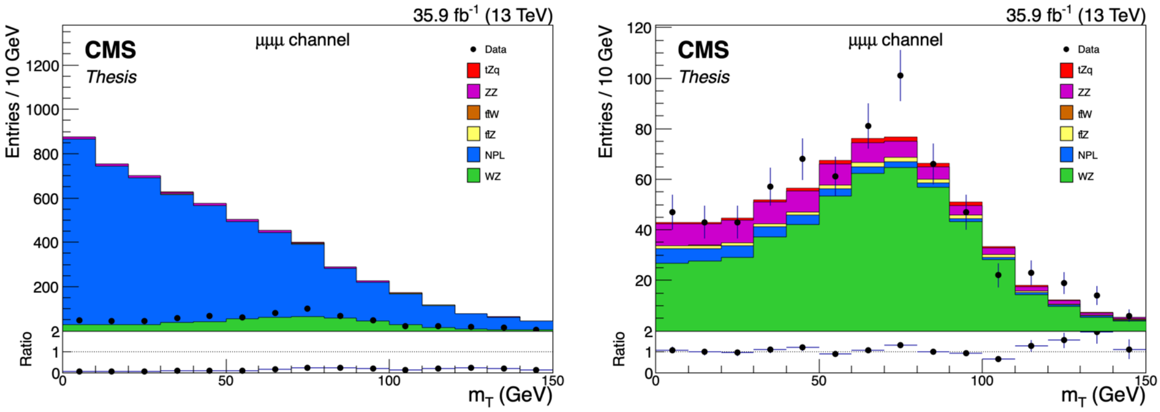


FIGURE 8.8: Distributions of the W boson transverse mass in the WZ CR for $\mu\mu\mu$ channel. Left: without normalization of NPL background. Right: with normalization of NPL background.

8.6 Event Selection

The basic selection for the analysis is exactly three tight leptons of which two leptons are compatible with the Z boson decay and the third lepton is from the top quark decay. Events with additional fourth leptons are vetoed. In the three lepton topology, the other main backgrounds apart from fakes are WZ and $t\bar{t}Z$. WZ has exactly three tight leptons making it an irreducible background. The $t\bar{t}Z$ has both 3 lepton and 4 lepton decay modes but the cross section for the three lepton final state is higher than that of the four lepton decay mode. In three lepton decays mode of $t\bar{t}Z$, one top quark decays hadronically. To study these backgrounds in detail we divide our event selection in the following three main regions as described in Section 8.1:

- WZ Control Region,

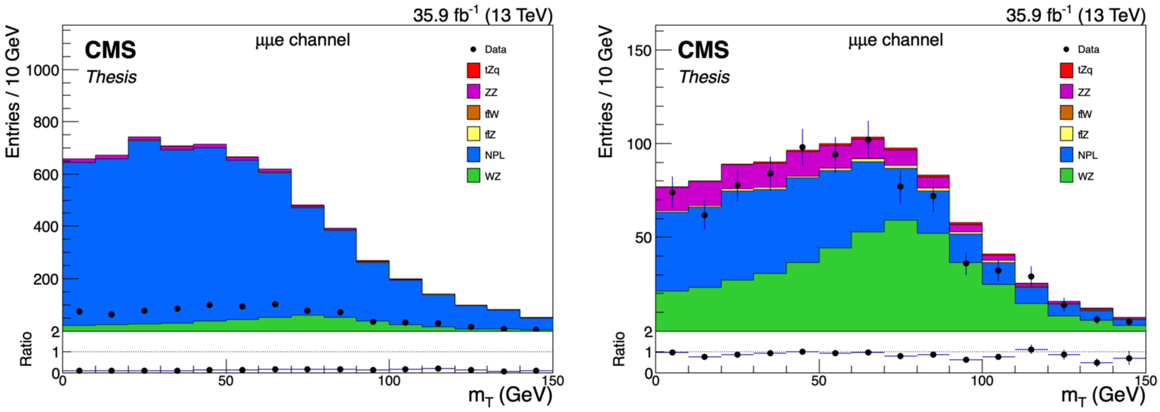


FIGURE 8.9: Distributions of the W boson transverse mass in the WZ control region for $\mu\mu e$ channel. Left: without normalization of NPL background. Right: with normalization of NPL background.

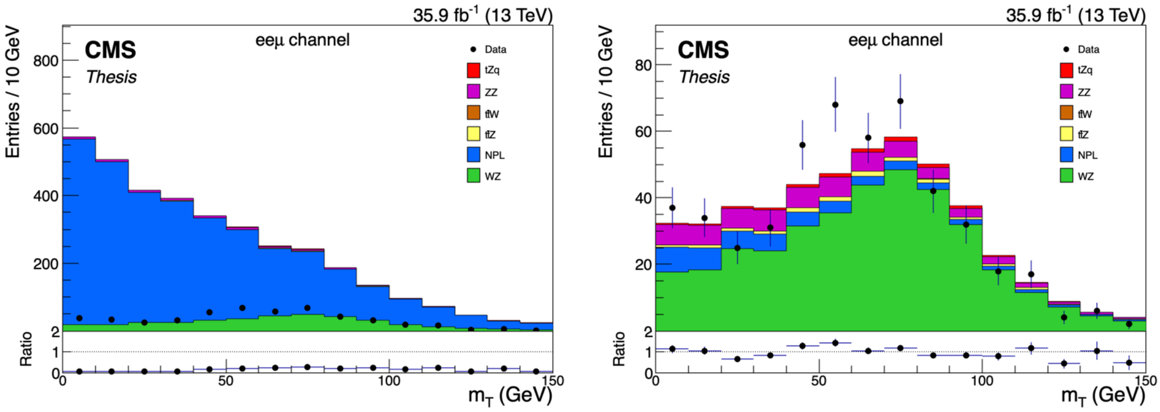


FIGURE 8.10: Distributions of the W boson transverse mass in the WZ control region for $ee\mu$ channel. Left: without normalization of NPL background. Right: with normalization of NPL background.

- $t\bar{t}Z$ Control Region,
- Signal Region.

8.6.1 WZ Control Region

A comparison of pre-selected distributions between the data and simulation is shown in this section. The pre-selected events are defined in WZ CR as described in Table 8.12. In this region, light jet greater than zero and zero b-tag jet are required.

The muon and electron multiplicities are shown in Figure 8.11 with data and MC. The p_T of the leading electron combined in four lepton channels is shown in Figure 8.12. The muon p_T for the combined channel is shown in Figure 8.13. Leading light jet p_T

Table 8.12: Selection criteria for WZ control region.

Background Selection for WZ control region.
Exactly three tight leptons. Veto any additional loose lepton. Two must be SFOS with Z mass window of 15 GeV. At least one light jet. No b-tag jet requirement.

distribution is shown in Figure 8.14. W transverse mass and $\vec{\not{E}}_T^{miss}$ in all four lepton channels and for the inclusive channel are shown in Figure 8.16 and Figure 8.15.

The distributions of di-leptons invariant mass in four possible decay channels of the tZq process and for the inclusive channel are shown in Figure 8.17 where inclusive means the sum of the four lepton channels. The invariant mass of the Z bosons shows a very nice peak with good data to MC agreement. The p_T of the Z bosons is shown in Figure 8.18. The event yield of various channels of tZq measurement in the WZ CR is shown in Figure 8.19. All these plots have very good data to MC agreement and validate the object selection performed in the analysis.

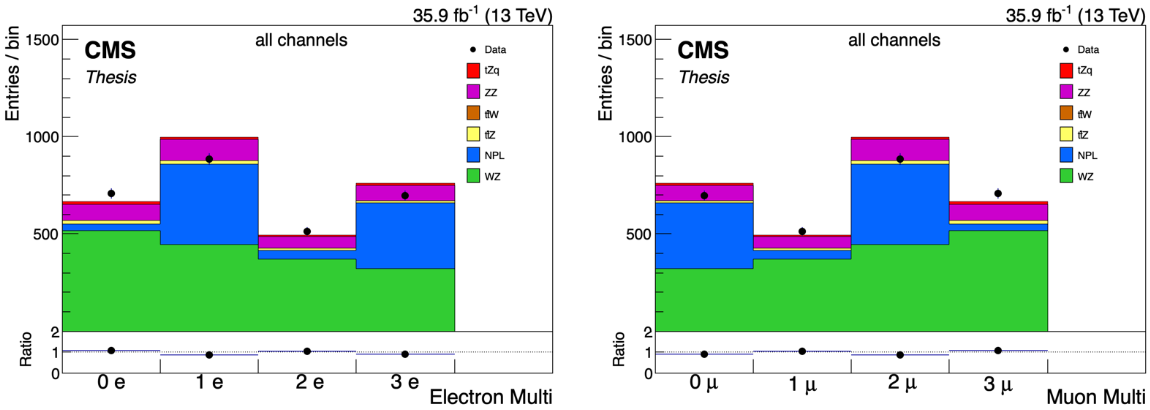


Figure 8.11: Electron (left) and muon (right) multiplicities in the WZ CR. The distributions are for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) summed up. WZ (green) and NPL (blue) are the dominant backgrounds in this region. The data is visible as black dots. The bottom panel of the plot displays data to MC ratio.

8.6.2 $t\bar{t}Z$ Control Region

The region enriched with $t\bar{t}Z$ background is called $t\bar{t}Z$ control region. The selection cuts of this region are given in Table 8.13. There are two top quarks in $t\bar{t}Z$ process which results in two b-tag jets in the event. The event yield distribution in $t\bar{t}Z$ control region is shown in Figure 8.20. The distribution for Z boson invariant mass is given in Figure 8.21.

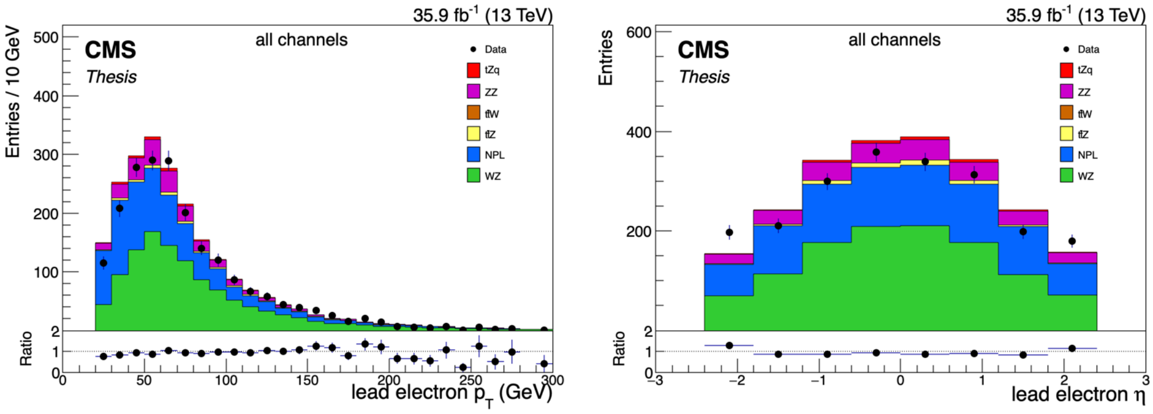


FIGURE 8.12: The electron p_T (left) and η (right) distributions in the WZ CR. The distributions are for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) summed up. WZ (green) and NPL (blue) are the dominant backgrounds in this region.The data is visible as black dots. The bottom panel of the plot displays data to MC ratio.

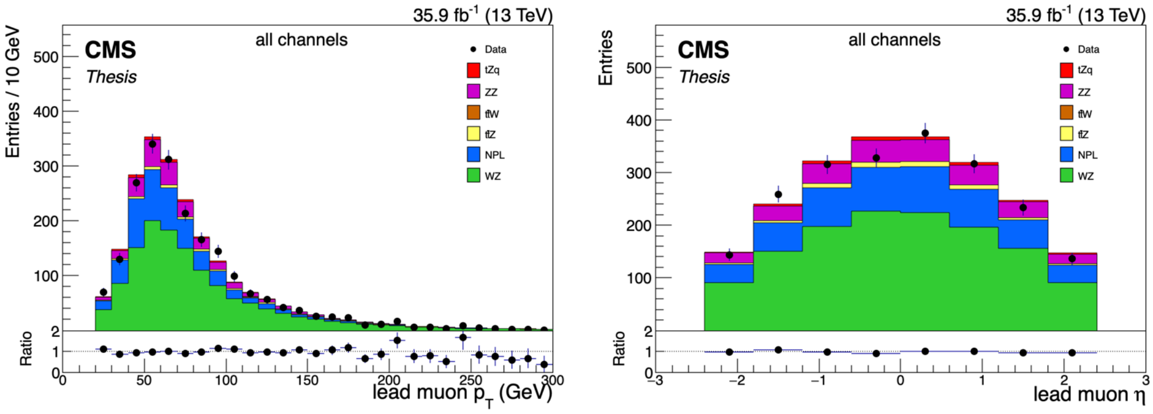


FIGURE 8.13: The muon p_T (left) and η (right) distributions in the WZ control region. The distributions are for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) summed up. WZ (green) and NPL (blue) are the dominant backgrounds in this region. The data is visible as black dots. The bottom panel of the plot displays data to MC ratio.

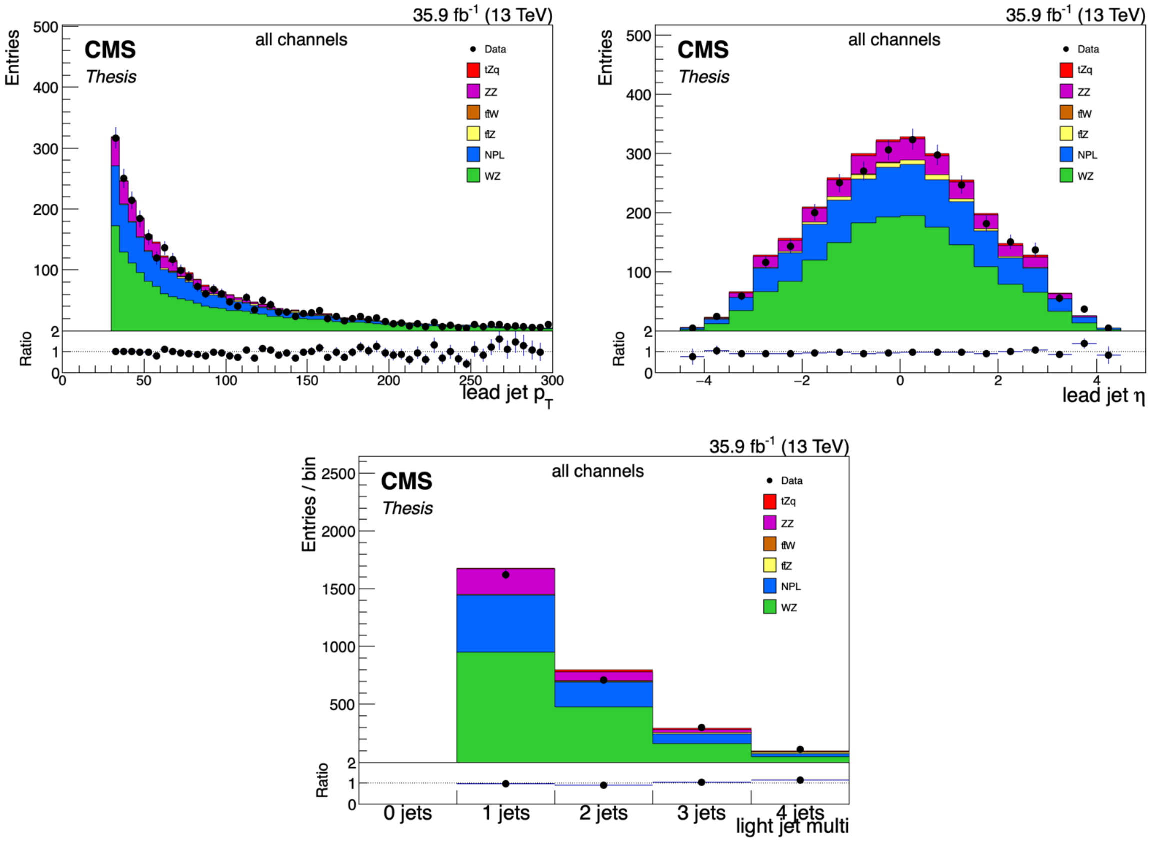


FIGURE 8.14: Leading light jet p_T (top-left), η (top-right) and multiplicity(bottom) distributions for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) summed up. WZ (green) and NPL (blue) are the dominant backgrounds in this region. The data is visible as black dots. The bottom panel of the plot displays data to MC ratio.

8.6.3 Signal Selection

In this region of phase space, the tZq signal is extracted. After the validation of the selection cuts in the pre-selection region, some additional cuts are applied to select signal events and to suppress the background events as much as possible. Thus, the region of phase space defined with additional cuts is called signal region. For the signal selection, we require three tight leptons for which two must be of same flavor opposite sign and at least two jets and one of which must be a b-tagged jet. The complete signal selection criteria is defined in Table 8.14.

The invariant mass distributions for Z boson and top quark for the combined channel (all four lepton channels summed up) are shown in Figure 8.22. Top quark is reconstructed by adding four momentum vectors of lepton, neutrino and b-tag jet (see Section 6.5 for details). The combined channel distribution for p_T of the Z boson and the top quark are shown in Figure 8.23. The events yield of tZq process in the SR is shown in Figure 8.24. The distributions of W boson transverse mass and $\vec{\not{E}}_T^{miss}$ for

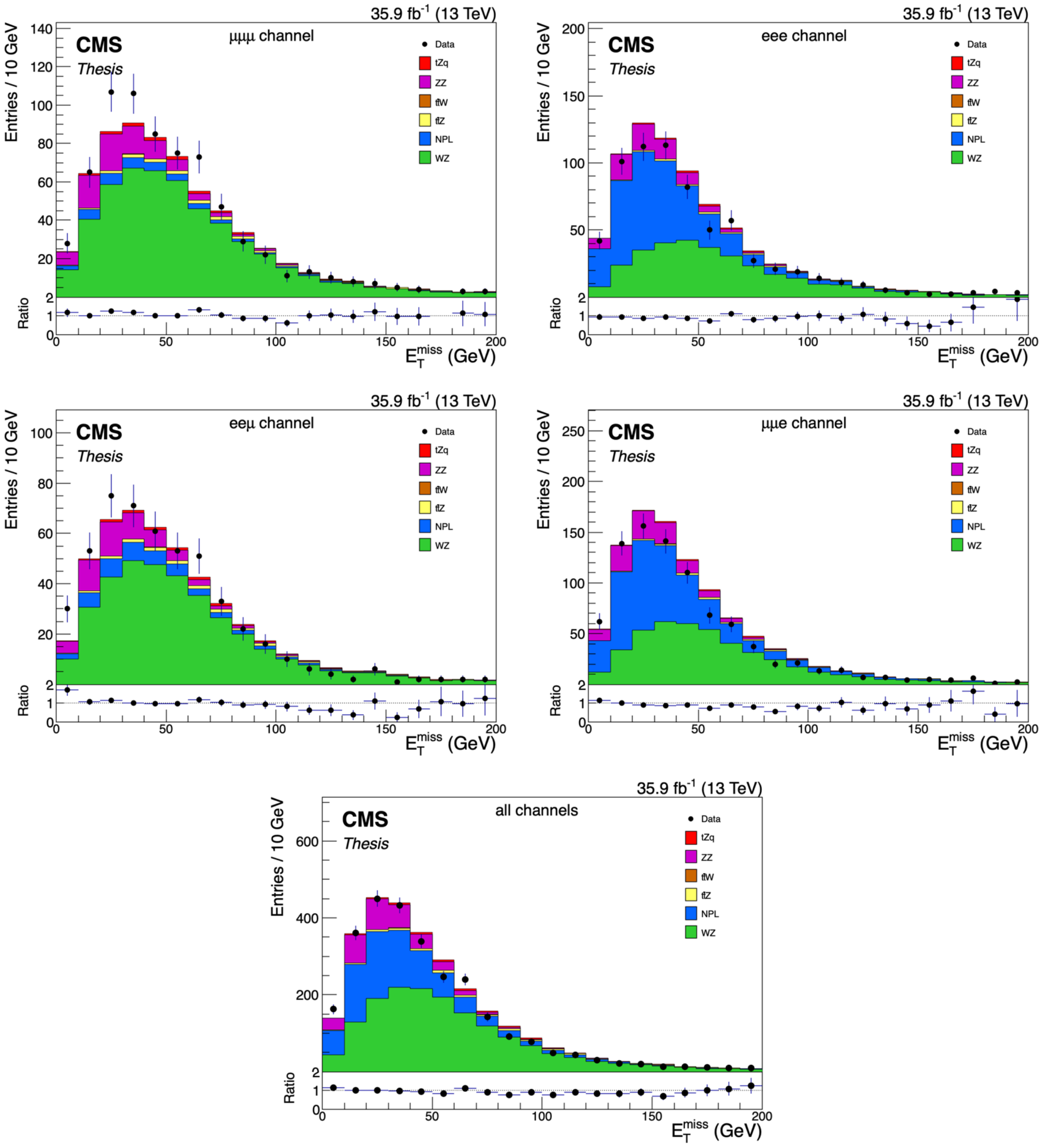


FIGURE 8.15: Distributions of the $\vec{\not{E}}_T^{miss}$ in the WZ CR region for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) and the combined channel. WZ (green) is the dominant background in all channels. NPL (blue) background has more contribution in eee and $\mu\mu e$ channels where third lepton is electron as compared to $\mu\mu\mu$ and $ee\mu$ channels. The data is visible as black dots. The bottom panel of the plot displays data to MC ratio.

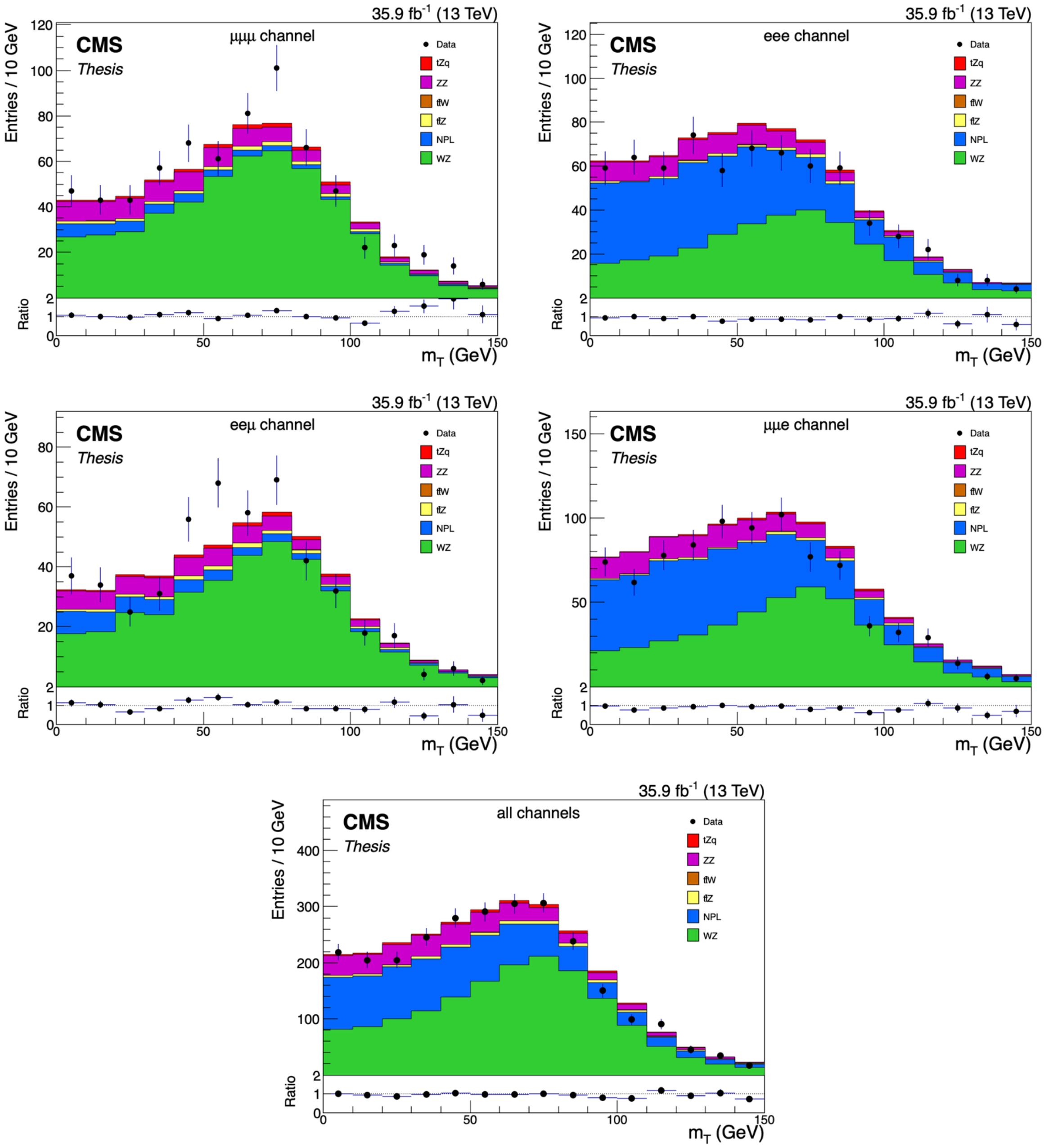


FIGURE 8.16: Distributions of the W transverse mass in the WZ CR region for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) and the combined channel. WZ (green) is the dominant background in all channels. NPL (blue) background has more contribution in eee and $\mu\mu e$ channels where third lepton is electron as compared to $\mu\mu\mu$ and $ee\mu$ channels. The data is visible as black dots. The bottom panel of the plot displays data to MC ratio.

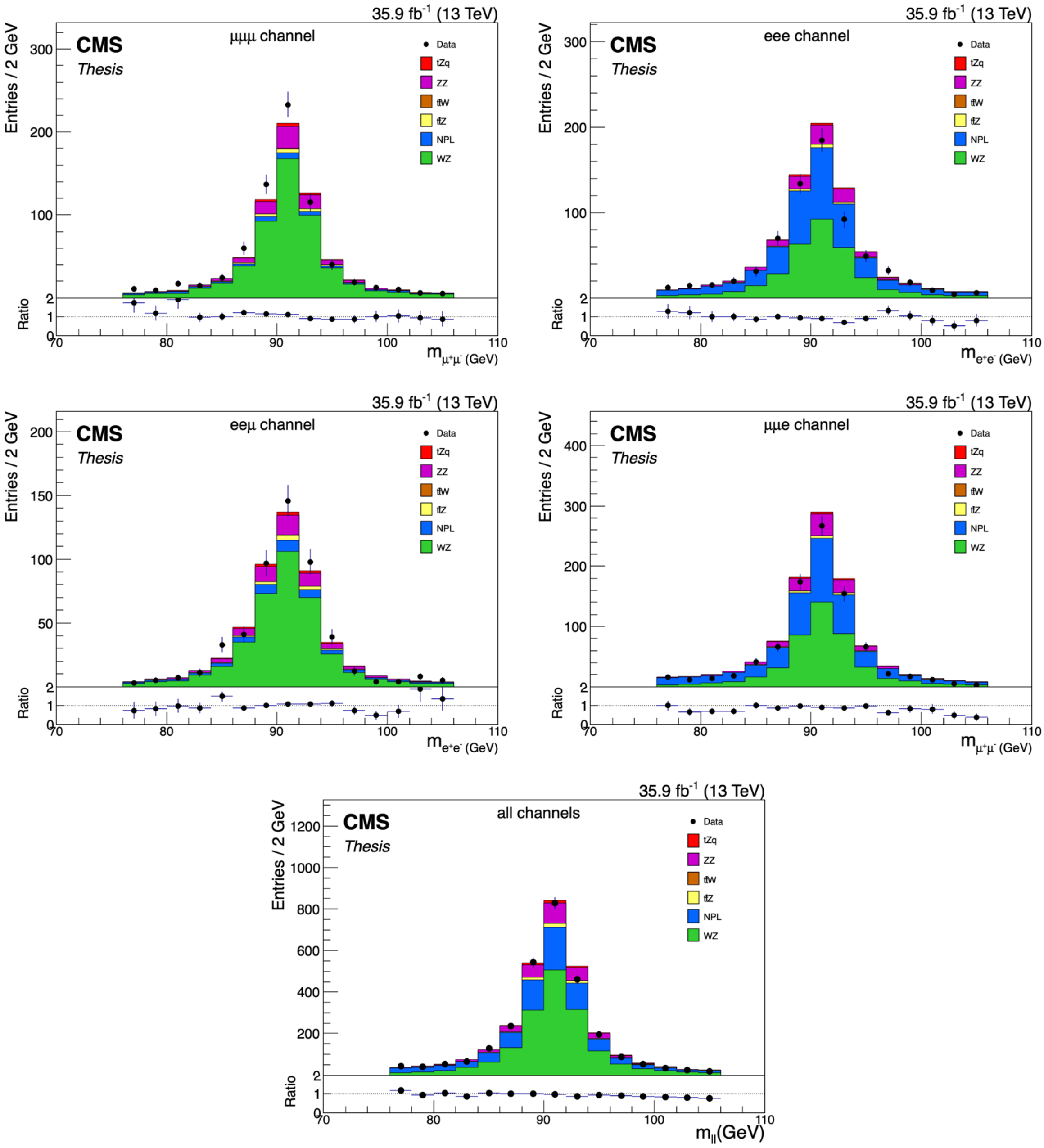


Figure 8.17: Distributions of the Z invariant mass in the WZ CR region for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) and the combined channel. WZ (green) is the dominant background in all channels. NPL (blue) background has more contribution in eee and $\mu\mu e$ channels where third lepton is electron as compared to $\mu\mu\mu$ and $ee\mu$ channels. The data is visible as black dots. The bottom panel of the plot displays data to MC ratio.

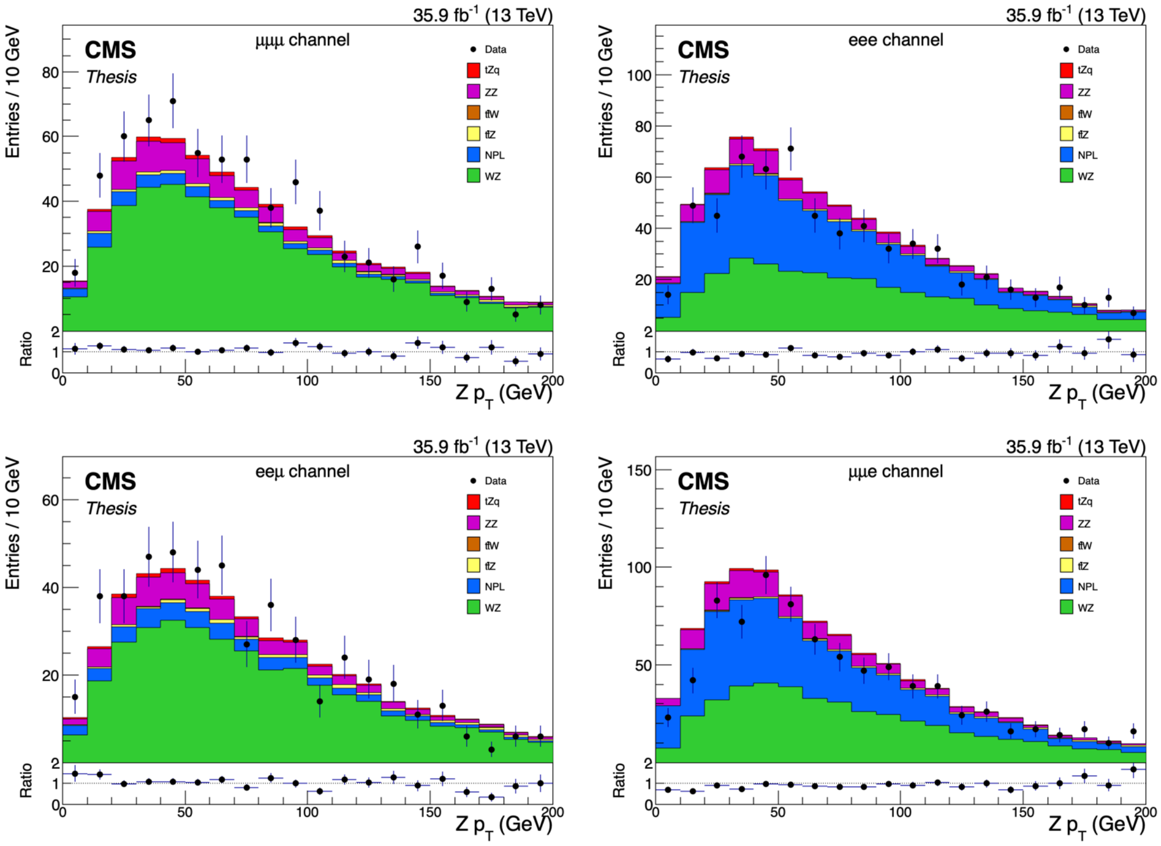


FIGURE 8.18: Distributions of the Z p_T in the WZ CR region for all four lepton channels ($\mu\mu\mu$, eee, $\mu\mu e$, $ee\mu$). WZ (green) is the dominant background in all channels. NPL (blue) background has more contribution in eee and $\mu\mu e$ channels where third lepton is electron as compared to $\mu\mu\mu$ and $ee\mu$ channels. The data is visible as black dots. The bottom panel of the plot displays data to MC ratio.

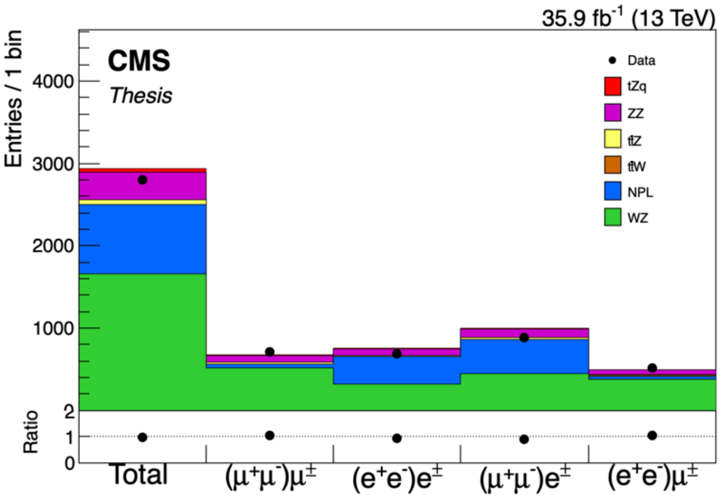


FIGURE 8.19: Event yield in the WZ control region. 1^{st} bin shows the total number of events (sum of all four lepton channels) in this region while the 2^{nd}, 3^{rd}, 4^{th} and 5^{th} bins are for each of the four lepton $\mu\mu\mu$, eee, $\mu\mu e$ and $ee\mu$ channels respectively. WZ (green) and NPL (blue) are the dominant backgrounds in this region.

TABLE 8.13: Selection criteria for $t\bar{t}Z$ control region.

Background Selection for $t\bar{t}Z$ control region
Exactly three tight leptons and veto fourth loose or tight lepton. Two must be SFOS within Z boson mass window of 15 GeV. At least one untagged jet. At least two b tagged jet.

the combined channels are shown in Figure 8.25. These distribution show that signal region has limited statistics but the data and simulation agrees within uncertainties.

8.7 Systematic Uncertainties

The following sources of systematical uncertainties are considered in the final fit as nuisance parameters..

- Jet energy corrections (JER/JES)
- b-tagging
- Lepton Selection (Identification and Isolation)
- Pileup re-weighting,
- Un-clustered $\vec{\not{E}}_T^{miss}$

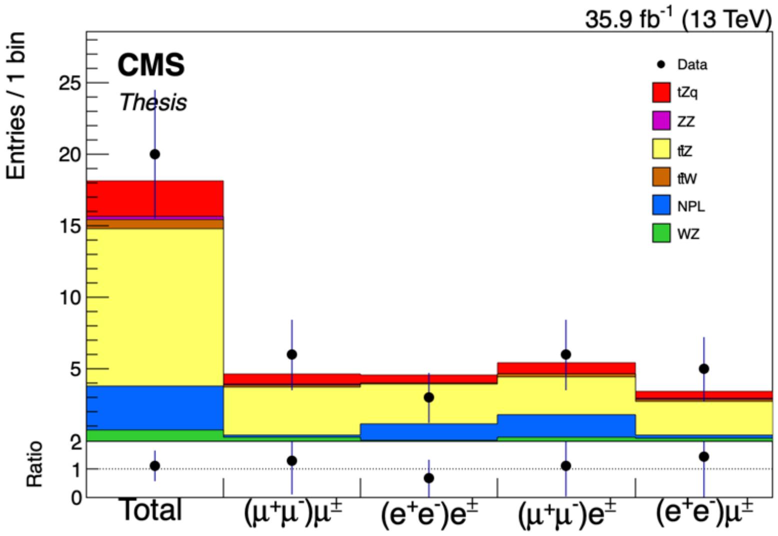


FIGURE 8.20: Event yield in the $t\bar{t}Z$ control region. 1^{st} bin shows the total number of events (sum of all four lepton channels) in this region while the 2^{nd}, 3^{rd}, 4^{th} and 5^{th} bins are for each of the four lepton $\mu\mu\mu$, eee, $\mu\mu e$ and $ee\mu$ channels respectively. $t\bar{t}Z$ is the dominant background in this region.

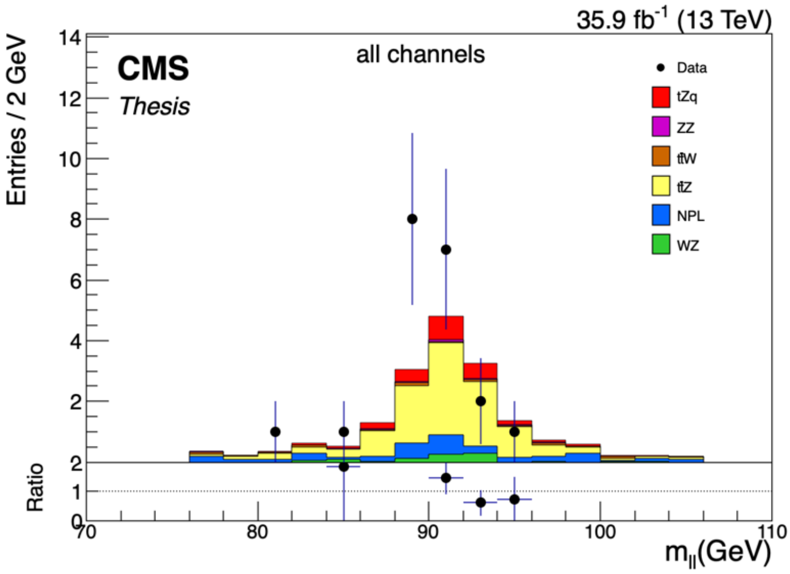


FIGURE 8.21: Distributions of the Z invariant mass in the $t\bar{t}Z$ CR region for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) summed up. $t\bar{t}Z$ (yellow) is the dominant background with smaller contribution from NPL (blue) background and signal (red). Because of two b-tag selection, the distribution has limited statistics.

- Luminosity

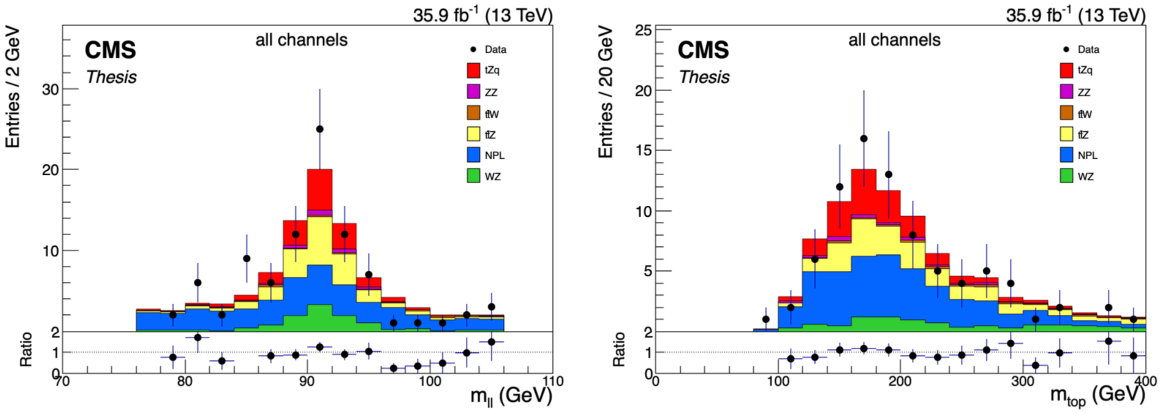


FIGURE 8.22: Distributions of the Z boson invariant mass (left) and top quark invariant mass (right) in the SR for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) summed up. The distribution shows nice peaks around Z mass(left) and top mass (right) with contributions from tZq signal (red) and NPL (blue) and $t\bar{t}z$ (yellow) as dominant backgrounds. The data is visible as black dots. The bottom panel of the plot displays data to MC ratio.

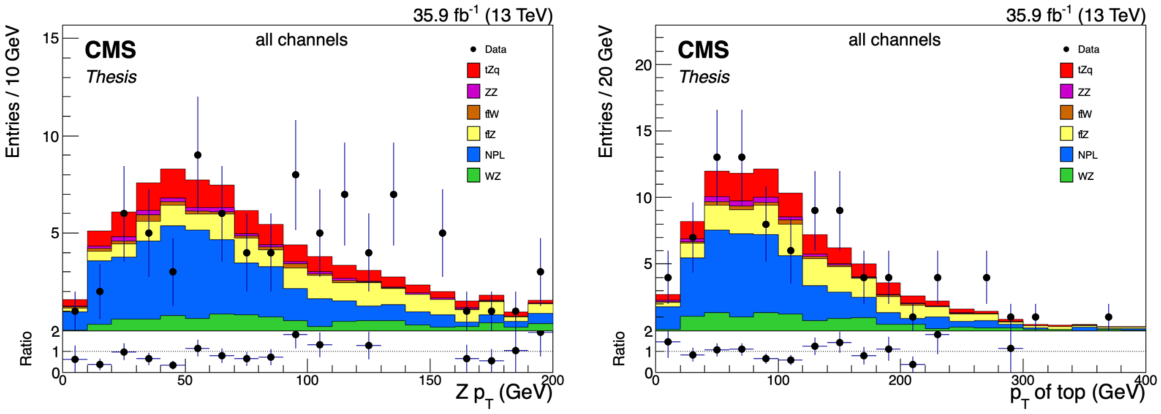


FIGURE 8.23: Distributions of the p_T of Z boson (left) and top quark quark (right) in the SR for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) summed up. The data is visible as black dots. The bottom panel of the plot displays data to MC ratio.

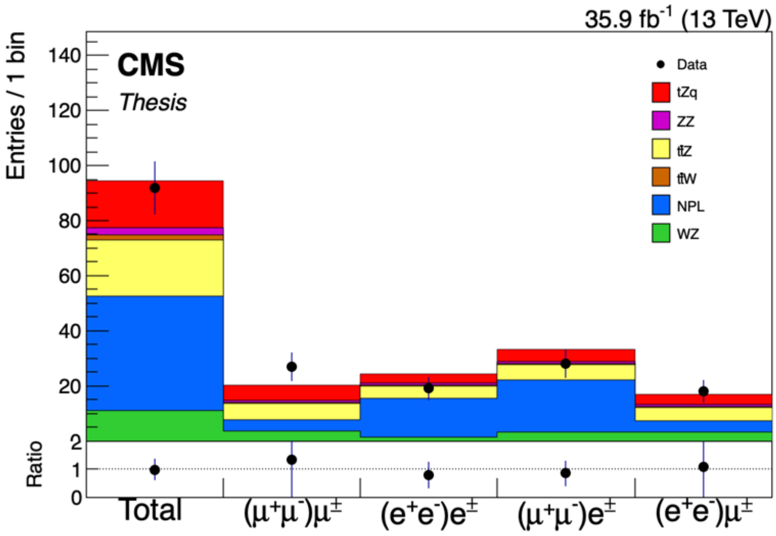


FIGURE 8.24: Event yield in the siganl region. 1st bin shows the total number of events (sum of all four lepton channels) in this region while the 2nd, 3rd, 4th and 5th bins are for each of the four lepton $\mu\mu\mu$, eee, $\mu\mu e$ and $ee\mu$ channels respectively. Only 92 data events are found after applying full signal selection cuts with $t\bar{t}Z$ (yellow) and NPL (blue) as dominant backgrounds.

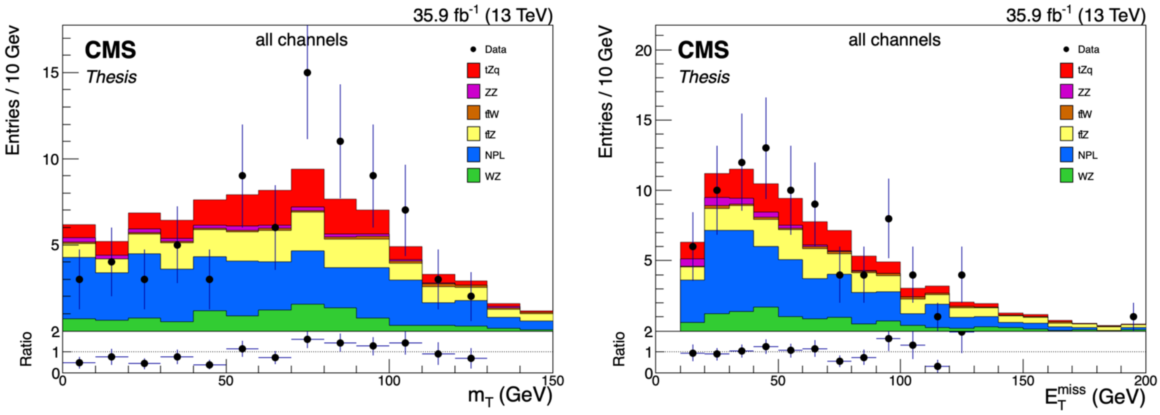


FIGURE 8.25: Distributions of the W boson transverse mass (left) and $\vec{\not\!E}_T^{miss}$ (right) in the SR for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) summed up. NPL (blue) and $t\bar{t}z$ (yellow) are dominant backgrounds with significant contribution from tZq simulated signal (red). The data is visible as black dots. The bottom panel of the plot displays data to MC ratio.

Table 8.14: Analysis signal selection criteria

Signal Selection
Exactly three tight leptons and vetoing fourth tight as well as loose lepton
Two must be SFOS with in Z mass window of 15 GeV
At least one or two untagged jet
Exactly one b-tag jet

- Trigger
- Normalization of simulated backgrounds
- Theory Systematics
- PDF
- Factorization and Renormalization

These systematics sources and their contribution to the final result are discussed in the following subsections.

8.7.1 Jet Energy Corrections (JER/JES)

Corrections and smearing are applied on jets to match the overall energy scale and resolution of simulation and data as described in Section 4.8. Jet energy corrections as suggested by POG for Summer16_23Sep2016V4 are applied. The systematic uncertainties on signal cross section are estimated by varying these SFs within their uncertainties.

The estimation of the systematic effect of JES/JER uncertainties require repeating the measurements of all jet and $\vec{\not{E}}_T^{miss}$ related kinematical observables for up and down variations. For the evaluation of uncertainties due to the JES the recipe documented in [187] is used. Systematic differences in the JER are treated by scaling the difference of the reconstructed jet p_T to the matched generator jet p_T as described in Section 7.7.5. The scale factor are given in Table 4.3.

8.7.2 Unclustered Missing Transverse Energy

Since the main uncertainty in the MET is due to the jet energy calibration which is already covered in JES uncertainty where the MET is modified according to the variation of the jet energy calibration. The remaining uncertainties in the MET are due to electron and muon energy uncertainties and unclustered energy. The contribution due to unclustered missing energy to MET is calculated by varying the unclustered energy by 10%.

8.7.3 Lepton Selection

For the muons, the scale factors recommended by muon POG for "tight" ID and tight isolation are applied. The uncertainty on these SFs is propagated as systematic uncertainty. The electron SFs are taken from [188, 189] and the uncertainty on these SFs is propagated as systematics.

8.7.4 Luminosity

The overall uncertainty of the luminosity for the 2016 data is estimated to be 2.5% [190].

8.7.5 b-tagging

There are eight sources of systematic uncertainties for b-tagging that affect the shape of the CSVv2 distribution, which are all considered as uncorrelated nuisance parameters in this physics analysis. These include two types of statistical uncertainties on the b-,c-, and light-flavor components of the MC event samples, light-flavor contamination of the b tagging scale factors, and b quark contamination in the mis-tag scale factors. The b-tagging SFs are varied by $\pm 1\ \sigma$ of their nominal value.

8.7.6 Pileup

Pileup re-weight SFs are recalculated by varying the minimum bias cross section (69 mb) by ± 4.6% [186]. These recalculated SFs are then applied to simulation. The difference of nominal and varied samples is calculated and propagated to the fit as log normally distributed nuisance parameter.

8.7.7 Trigger

Trigger efficiencies are estimated to be 100% in data and simulations. An uncertainty of 1% has been included for trigger selection for $ee\mu$ and $\mu\mu\mu$ and 2% for the eee and $\mu\mu e$ final states.

8.7.8 Non-prompt Lepton Estimation

The uncertainty on the normalisation of the overall non prompt electrons and muons yield is conservatively assumed to be 50%. The uncertainty obtained by varying the isolation inversion cut values with respect to nominal for electrons and muons are found to be negligible.

8.7.9 Normalization of Simulated Backgrounds

The cross section of the simulated backgrounds is varied by ± 30% to cover all the uncertainties at NLO accuracy. This uncertainty is not applied to NPL background as it is estimated from data.

8.7.10 PDF Uncertainty

Theory uncertainties due to parton density functions used for the modelling of the backgrounds are estimated only for the tZq process. This is done using the PDF4LHC recipe described in [191], which combines the MMHT14, CT14, and NNPDF 3.0 PDF sets. The nominal pdf set NNPDF30_nlo_as_0118 used for signal sample tZq while the set of error PDFs are taken from PDF4LHC15_nlo_100.

8.7.11 Renormalization and Factorization Scales

The effect of theory uncertainties due to the factorization scale μ_F and renormalization scale μ_R are estimated by varying the μ_F by a factor 0.5 and 2 at fixed μ_R then varying μ_R by a factor 0.5 and 2 at fixed μ_F. Two more variations at matrix element-level are retrieved by varying μ_F and μ_R simultaneously by a factor 2 and 0.5. The envelope of these variations is used as an final uncertainty. This uncertainty is only considered for the tZq sample.

8.7.12 ME-PS Matching Scale

Other systematic uncertainties from the modelling of simulated events arise from the matching of matrix-element partons and parton showers in Powheg. The high-p_T radiations during this matching process are regulated by damping real emissions with a factor of $\frac{h^2_{damp}}{(p^2_T+h^2_{damp})}$. The nominal value for the h$_{damp}$ parameter was calculated to be $1.58^{+0.66}_{-0.59} \times m_{top}$ using m_{top} = 172.5 GeV [192]. The differences between the nominal and the up/down h_{damp} variations are estimated using dedicated weighted samples for which the information are stored in the simulated samples produced by Powheg. In this analysis, the ME-PS systematic uncertainty is applied for tZq simulated samples.

All these systematics are accounted for nuisance parameters in the fit. For a given source of systematic uncertainty there is 100% correlation between the four lepton channels, except for the lepton misidentification where the $\mu\mu\mu$ and $ee\mu$ channels are 100% correlated and the $\mu\mu e$ and eee channels are 100% correlated, due to the isolation inversion of the lepton candidate from the W decay. The nominal and varied distributions of some of these systematics are given in Appendix A for the tZq signal sample and dominant backgrounds process.

8.8 Results

The analysis is done in three lepton final state with several jet and b-tagged jet multiplicities. In this physics analysis, it is assured that no bias is introduced into the measurements. Therefore, the analysis strategy has been established using a blinded methodology where the cross section measurement procedure has been validated by using a pseudo dataset. Then the analysis is unblinded to use the real data.

The maximum likelihood procedure explained in Section 6.2 is applied for each lepton channel ($\mu\mu\mu$, eee, $ee\mu$ and $\mu\mu e$) separately and for the combination of the four channels as well. The tool used for fitting is based on the RooStats framework [193, 169]. The results are obtained in terms of signal strength (μ) defined in Equation 7.5. The expected and observed μ as a function of test statistic is shown in Figure 8.26. The measured signal sterngth μ is compatible with the expected SM value. The width of the likelihoods follow very closely the expectations obtained from the Asimov datasets.

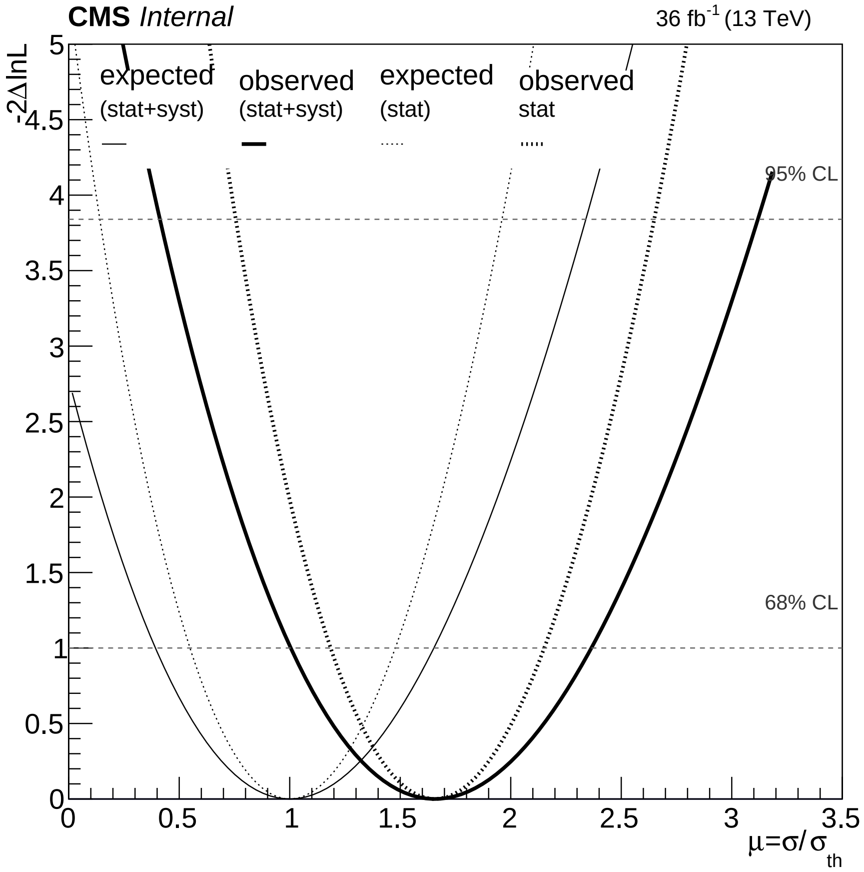


FIGURE 8.26: The variation of likelihood as a function of the μ for the observation and expectation. The solid curves include both statistical and systematical errors while dashed curve includes statistical error only. The horizontal lines are for 68% and 95% confidence level (CL) intervals for the μ.

The systematic uncertainties explained in Section 8.7 are set as free parameters in the fit. Each systematic uncertainty has different effect on the result. Therefore, the impact of each systematic uncertainty on μ is defined as the shift of $\Delta\mu$ when θ is fixed and brought to its $\pm 1\sigma$ post-fit values while other systematics are treated as usual:

$$\Delta\mu = \hat{\mu}(\hat{\theta} \pm \Delta_{\theta}) - \hat{\mu}(\hat{\theta})$$

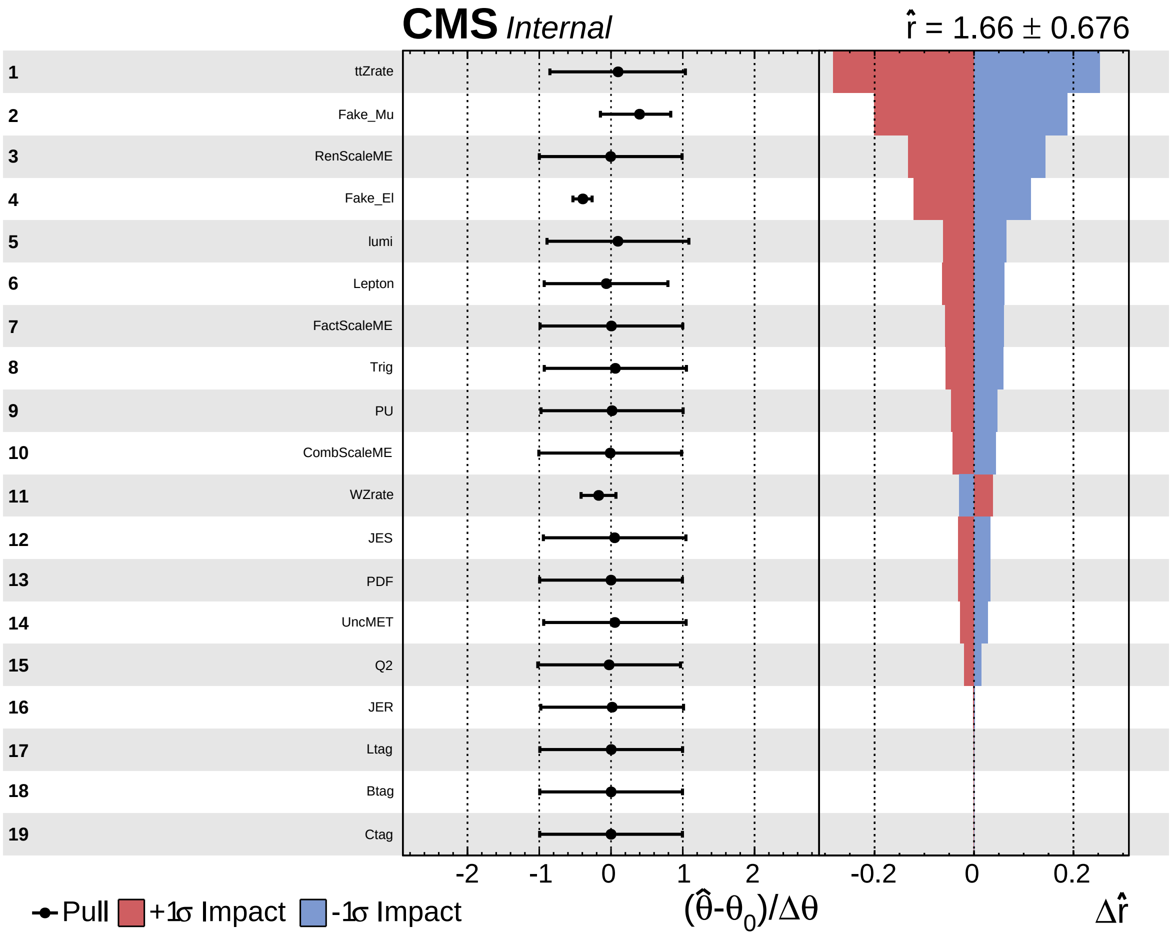


FIGURE 8.27: The pulls of the nuisance parameters and the influence of their uncertainty on the maximum likelihood estimation of the tZq signal strength μ.

The impact of each systematic source in ascending order is shown in Figure 8.27. It can be see that the uncertainties due to the $t\bar{t}Z$ background cross section and non-prompt muon have the largest effect followed by those from the renormalization and factorization scales. In addition to the impact of the uncertainties, the best fit value of the nuisances after the fit (post-fit) can be different from their initial (pre-fit) values given as a free parameter in the fit. Some of the nuisances can be pulled up (down) having larger (smaller) post-fit values while some may not get constrained in the fit. This difference can be described as $(\hat{\theta} - \theta_0)/\Delta\theta$ where $\hat{\theta}$ and θ_0 are the post-fit and the pre-fit values of θ respectively and $\Delta\theta$ is its uncertainty. The pull distribution of the nuisances are shown in Figures 8.27, where it can be seen that some statistical uncertainties are pulled up and down by less than 1 σ. Furthermore, the nuisance parameters related to the non-prompt electron and WZ background cross section uncertainty are shifted with respect to their initial values.

The target of the analysis is to measure the cross section of the tZq process with its statistical significance. In order to extract the the signal strength modifier μ with its associated 68% CL interval, a simultaneous ML fit is performed in the signal and control regions in all four lepton channels individually and the combined channel. In Table 8.15, the event yields for each process after the fit are given. The last column displays the ratio of the post-fit to pre-fit predictions, $N^{post-fit}/N^{pre-fit}$, accounting for the systematic uncertainties. The post-fit background normalizations are close to the pre-fit values for most of the background processes.

TABLE 8.15: The table shows postfit event yields for SM backgrounds simulated samples as well as Run II 2016 data after the baseline selection cuts in the tZq signal region

	$\mu\mu\mu$ Channel	eee Channel	$ee\mu$ Channel	$e\mu\mu$ Channel	Combined Channel	$N^{postfit}/N^{prefit}$
tZq	9.1 ± 3.3	5.4 ± 1.9	5.9 ± 2.1	7.7 ± 2.8	28.1 ± 5.2	-
WZ	3.5 ± 0.2	1.4 ± 0.1	2.7 ± 0.1	2.9 ± 0.1	10.5 ± 0.2	0.95
$t\bar{t}Z$	6.0 ± 1.1	4.3 ± 0.5	4.7 ± 0.6	5.7 ± 1.1	20.7 ± 1.7	0.98
$t\bar{t}W+ZZ$	1.3 ± 0.1	1.0 ± 0	1.0 ± 0	1.2 ± 0	4.5 ± 0.1	1.01
Fakes	3.9 ± 1.2	8.3 ± 0.9	5.3 ± 1.6	14.2 ± 1.2	31.7 ± 2.5	0.81
Total	23.8 ± 4.1	20.1 ± 2.3	19.6 ± 3.3	31.7 ± 3.7	95.2 ± 6.8	1.02
Data	27	19	18	28	92	

The combined channel best fit value of the signal strength modifier that describes the data is found to be

$$\mu = 1.66^{+0.50}_{-0.45}(stat)^{+0.43}_{-0.36}(syst).$$

From this signal strength using the reference NLO cross section from Equation 8.1, the measured cross section is found to be

$$\sigma_{tZq}(t\ell^+\ell^-q) = 156^{+47}_{-42}(stat)^{+40}_{-34}(syst)fb$$

for $m_{\ell^+\ell^-} > 30$ GeV, where ℓ stands for electron, muon, and tau lepton. The quoted systematic uncertainty is then calculated as the difference in quadrature between the 68% CL intervals obtained in the nominal fit and in the fit without systematic uncertainties. The precision of the measurement is limited by the statistical uncertainty.

The expected and measured significance calculated using the background-only hypothesis are shown in Table 8.16. The expected significance is estimated from an Asimov toy dataset [194]. The observed significance for inclusive cross section is 2.81 σ compared to the expected significance of 1.95 ± 0.09. The most sensitive channel is $\mu\mu\mu$ and $ee\mu$ i.e. when the lepton from W decays to a muon and neutrino. While the third electron channel eee and $\mu\mu e$ are affected more by fake lepton.

Table 8.16: The table shows the expected and observed significance of all four lepton and combined channel. The expected significance is estimated from an Asimov toy dataset. The significance is calculated using background-only hypothesis.

Channel	Expected Significance			Observed Significance
-	Mean	68% CL	95% CL	
$\mu\mu\mu$	1.27 ± 0.06	0 - 2.04	0 - 3.01	2.68
eee	0.77 ± 0.03	0 - 1.69	0 - 2.61	0.44
$\mu\mu$e	0.96 ± 0.04	0 - 1.93	0 - 2.77	0.42
eeμ	0.90 ± 0.04	0 - 1.82	0 - 2.79	1.28
inclusive	1.95 ± 0.09	0.95 - 3.02	0 - 4.27	2.81

8.9 Conclusions

The associated production cross section of a single top quark with a Z boson was measured using data from pp collisions at 13 TeV collected by the CMS experiment, corresponding to an integrated luminosity of 35.9 fb^{-1} . The measurement uses events containing three charged leptons in the final state. The tZq cross section measurement is found to be in agreement with SM with an observed (expected) significance of 2.81(1.95) standard deviation. The cross section is measured to be $\sigma(t\ell^+\ell^-q) = 156^{+47}_{-42}$ (stat)$^{+40}_{-34}$ (syst) fb, for $m_{\ell^+\ell^-} > 30$ GeV. where ℓ stands for electrons, muons and τ leptons. This value is compatible with the NLO standard model prediction of 94.2 ± 3.1 fb.

8.10 Shape Analysis Results

The results reported in Section 8.8 are calculated by simple and robust cut and count technique. These results are reported in the paper [51] as a cross check to the shape analysis result. In the shape analysis, a simultaneous ML fit is performed on the W transverse mass distribution in the background enriched sample and on the BDT outputs in the signal region. The excess of events in the background only hypothesis is compared to the SM expectation for tZq production in order to measure the cross section. Evidence for tZq production is found with an observed (expected) significance of 3.7(3.1) standard deviations. The cross section is measured to be $\sigma(t\ell^+\ell^-q) = 123^{+33}_{-31}(stat)^{+29}_{-23}(syst)$ fb, for $m_{\ell^+\ell^-} > 30$ GeV, where ℓ stands for electrons, muons and *tau* leptons. This value is compatible with the NLO standard model prediction of 94.2 ± 3.1 fb. The latest observation results of tZq process (with significance greater than five standard deviation) calculated by using CMS data of 2016 and 2017 are reported in the reference [195].

Chapter 9

Conclusion

A first cross section measurement of single top quark production in association with a Z boson (tZq), using the data collected by CMS experiment at LHC, is performed. The data sample corresponds to an integrated luminosity of 19.7 fb^{-1} at $\sqrt{s} = 8$ TeV collected during the year 2012. The measurement is performed within the standard model framework where top quark is produced via the t-channel process. The signature for tZq production consists of a single top quark, a Z boson, and an additional ("recoiling") jet. The analysis uses events where the Z boson decays to e^+e^- or $\mu^+\mu^-$, while the W boson, produced in the decay of top quark, decays to a neutrino and an electron or a muon, resulting in four possible lepton combinations in the final state namely: eee, $ee\mu$, $\mu\mu e$, and $\mu\mu\mu$. There is also a small contribution from τ leptons decaying into electrons or muons. A simple cut and count method is used for the cross section measurement. The main sources of background to the tZq process are $t\bar{t}$ production, diboson production, ttV (V = W or Z) and Drell-Yan (DY) production. The backgrounds due to $t\bar{t}$ and Drell-Yan are estimated from the data, whereas other backgrounds are estimated from simulation using constraints from data. The cross section is measured for each lepton channel and the combined channel (sum of leptonic channels) by performing a one bin likelihood fit to the event yields in the signal region. The cross section for the combined channel is measured to be $\sigma(t\ell^+\ell^-q) = 18^{+11}_{-9}$(stat) ± 4 (syst) fb, for $m_{\ell^+\ell^-} > 30$ GeV, where ℓ stands for electrons, muons and τ leptons, with an observed and expected significance of 1.81 and 0.81 standard deviation respectively. This measured value is compatible with the NLO standard model prediction of 8.2 fb with a theoretical uncertainty of less than 10%.

With the high centre of mass proton-proton collision energy of 13 TeV at the LHC, together with large integrated luminosities, the cross section measurement of tZq process is repeated to improve upon the previous 8 TeV results. The dataset used in the cross section measurement is collected during the year 2016 and corresponds to an integrated luminosity of 35.9 fb^{-1} . The measurement is performed for each of four lepton channels ($\mu\mu\mu$, eee, $\mu\mu e$, $ee\mu$) and the combined channel using a simple cut and count method. The dominant backgrounds to the tZq process are due to non-prompt lepton ($t\bar{t}$, Drell-Yan), WZ and $t\bar{t}Z$. To constrain the backgrounds, two control regions are defined, each containing different contributions from signal and background processes. A simultaneous fit to the data in the signal and control regions is performed to extract the cross section. The cross section is measured to be $\sigma(t\ell^{+}\ell^{-}q) = 156^{+47}_{-42}(stat)^{+40}_{-34}(syst)$ fb, where ℓ stands for electrons, muons and τ leptons. This value is compatible with the NLO standard model prediction of 94.2 ± 3.1 fb. The measurement is in agreement with the SM with an observed (expected) significance of 2.81(1.95) standard deviation.

The CMS Silicon tracker module operates in peak and deconvolution readout modes. Because of the short readout time in deconvolution mode, the alignment procedure introduces a shift in silicon sensor position. A backplane correction is measured to correct the misalignment of the sensor position. It is validated that the backplane corrections works fine with 2017 CMS data.

Appendix A

Systematics Uncertainties Plots

A.1 Systematics plots for 13 TeV Analysis

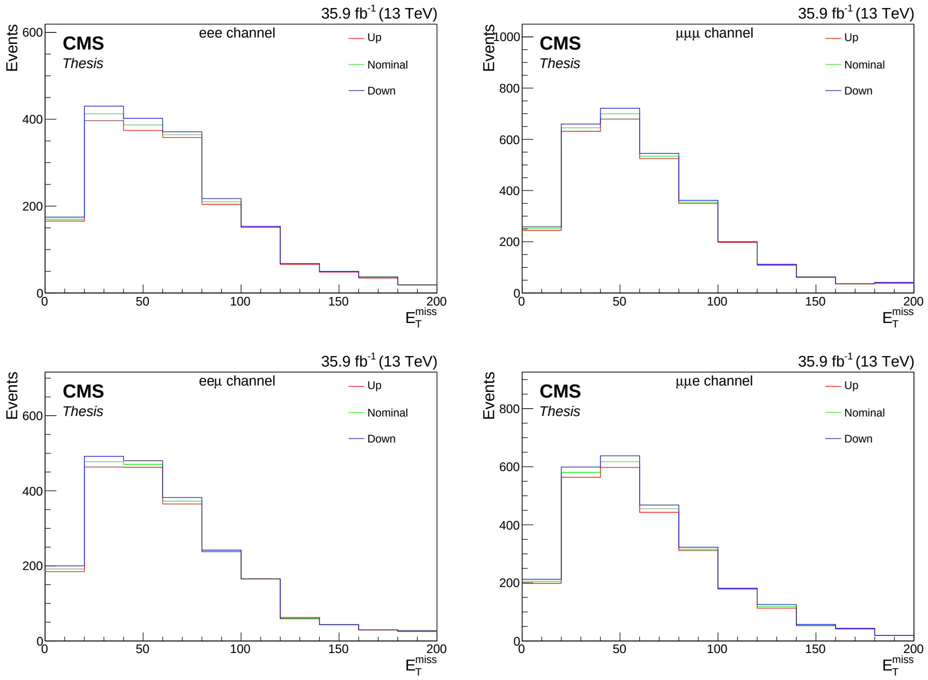


FIGURE A.1: Systematic uncertainty on the $\vec{\not\!\!E}_T^{miss}$ distribution due to nominal and varied PU SFs for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) on the tZq signal sample.

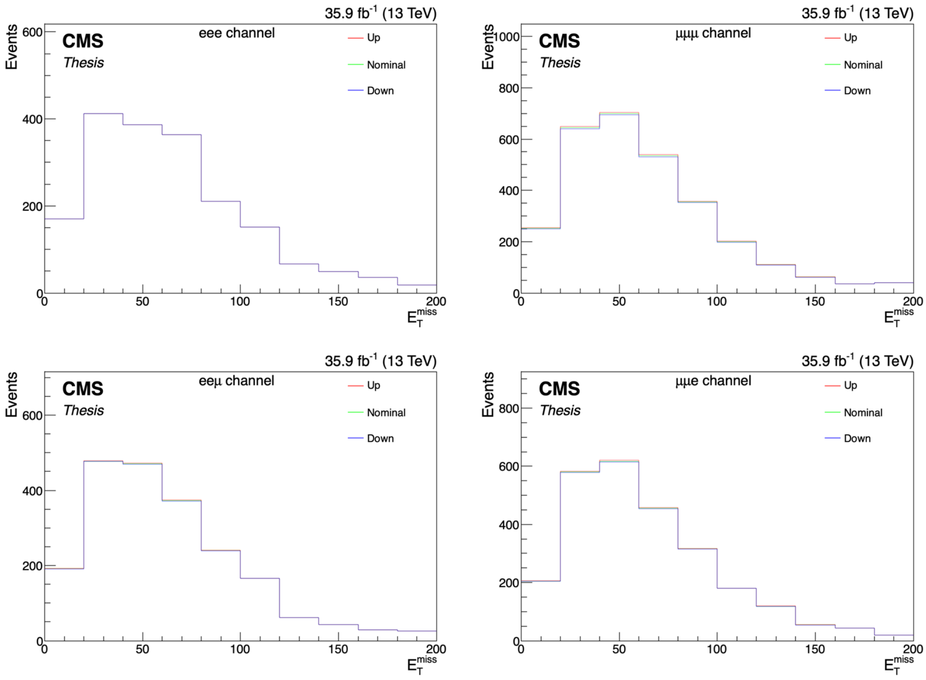


FIGURE A.2: The effect of the muon efficiency systematic uncertainty on the $\vec{\not{E}}_T^{miss}$ distribution for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) on the tZq signal sample.

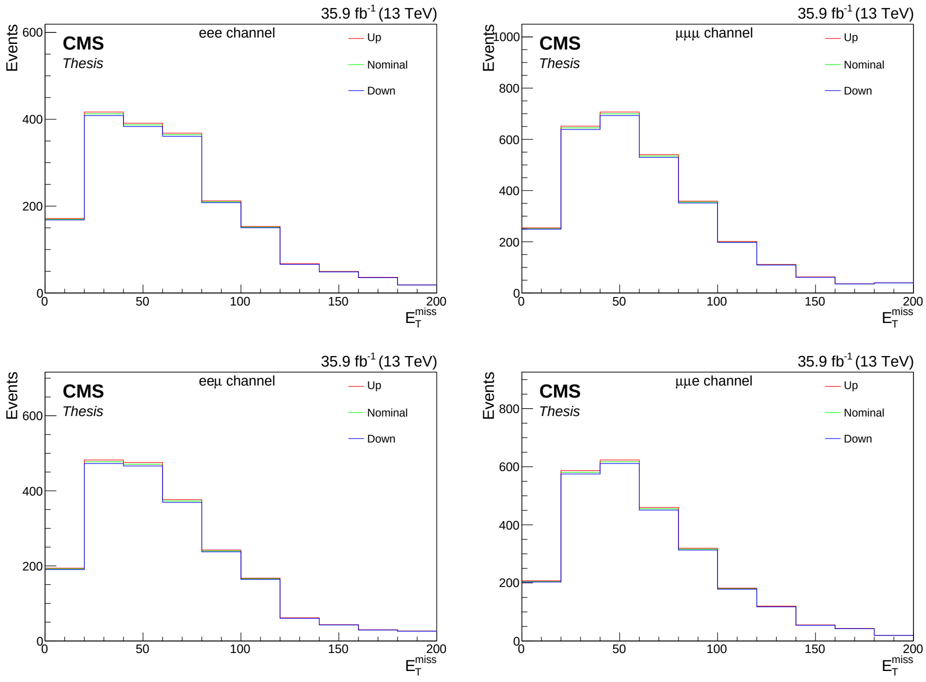


FIGURE A.3: Systematic uncertainty on the $\vec{\not{E}}_T^{miss}$ distribution due to nominal and varied lepton trigger SFs for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) on the tZq signal sample.

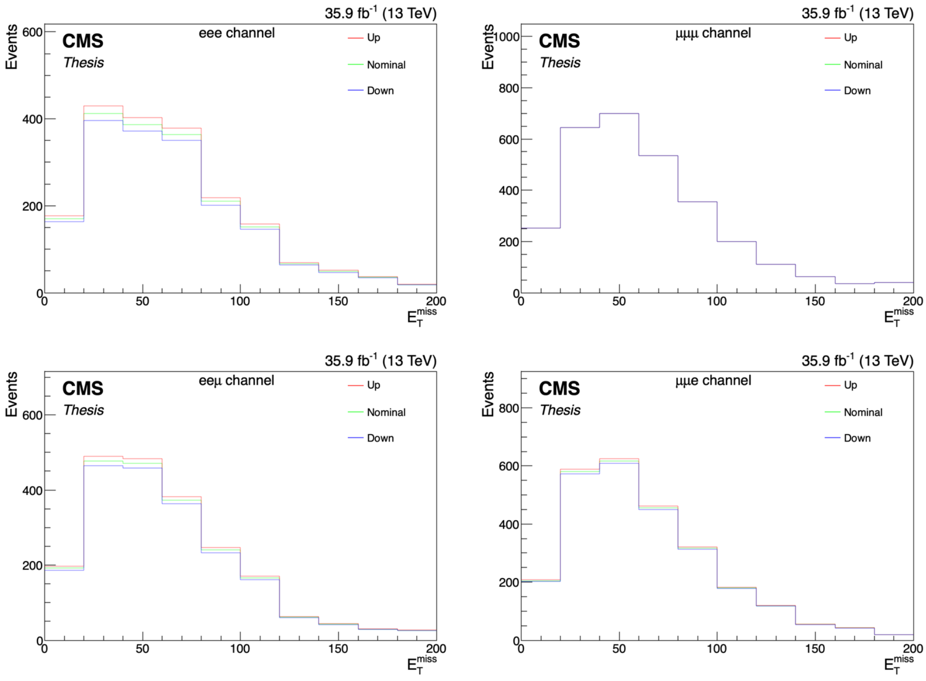


FIGURE A.4: The effect of the electron efficiency systematic uncertainty on the $\vec{\not{E}}_{T}^{miss}$ distribution for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) on the tZq signal sample.

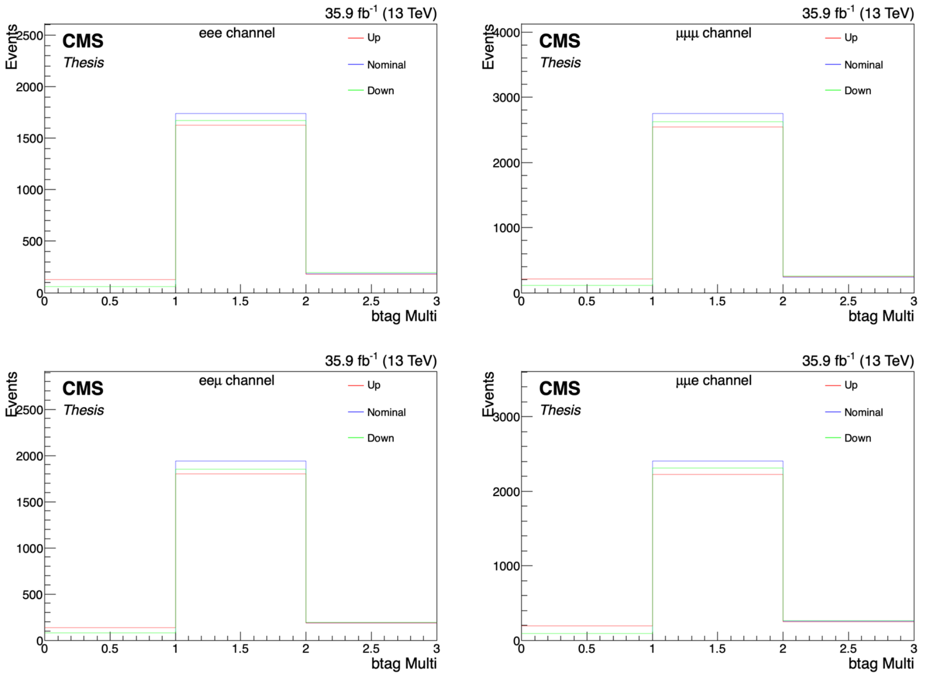


FIGURE A.5: Systematic uncertainty on the b-tag multiplicity due to nominal and varied b-tagging SFs distribution for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) on the tZq signal sample.

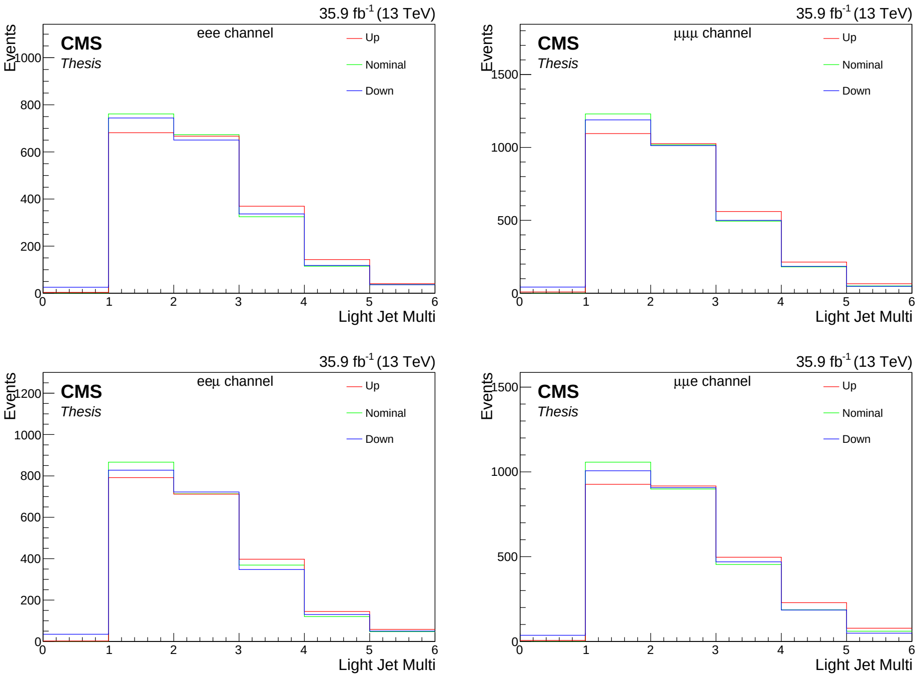


FIGURE A.6: Systematic uncertainty on the jet multiplicity distribution due to nominal and varied jet energy scales for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) on the tZq signal sample.

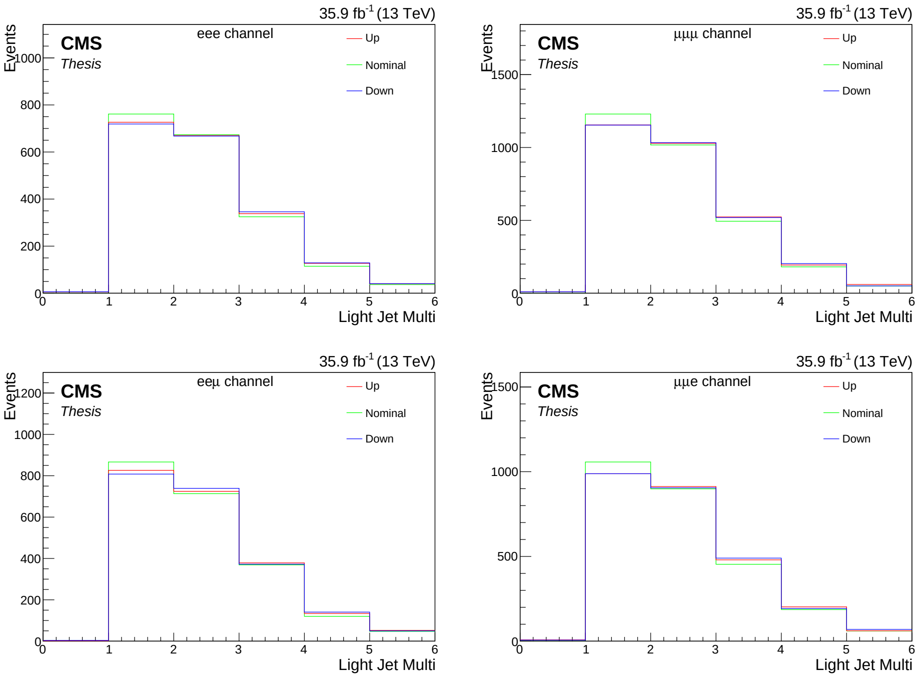


FIGURE A.7: The effect of the JER systematic uncertainty on the jet multiplicity distribution for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) on the tZq signal sample.

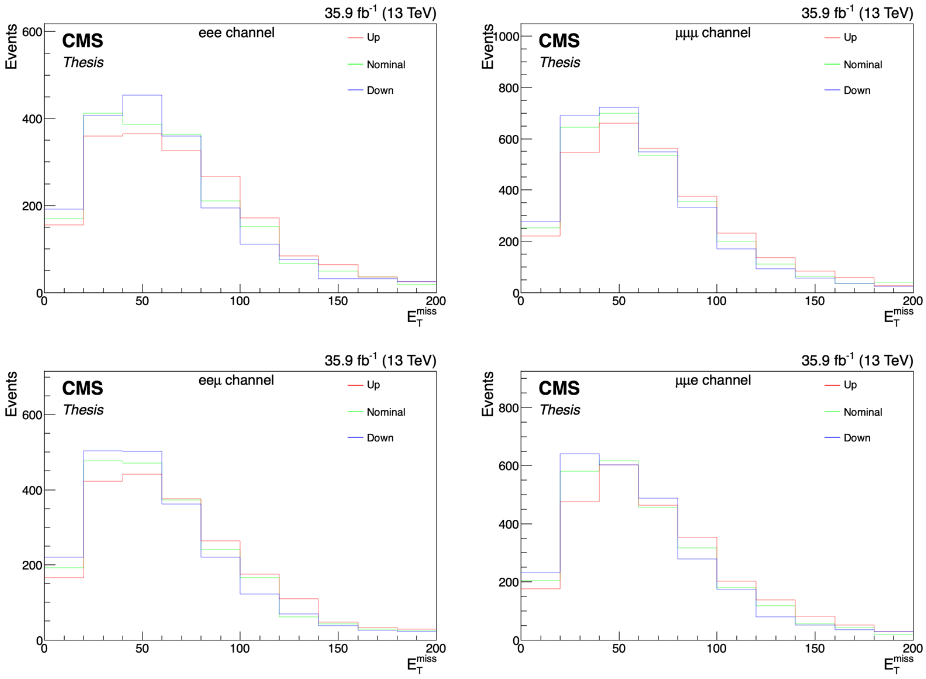


FIGURE A.8: Systematic uncertainty on the $\vec{\not E}_T^{miss}$ distribution due to nominal and varied unclustered missing transverse energy for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) on the tZq signal sample.

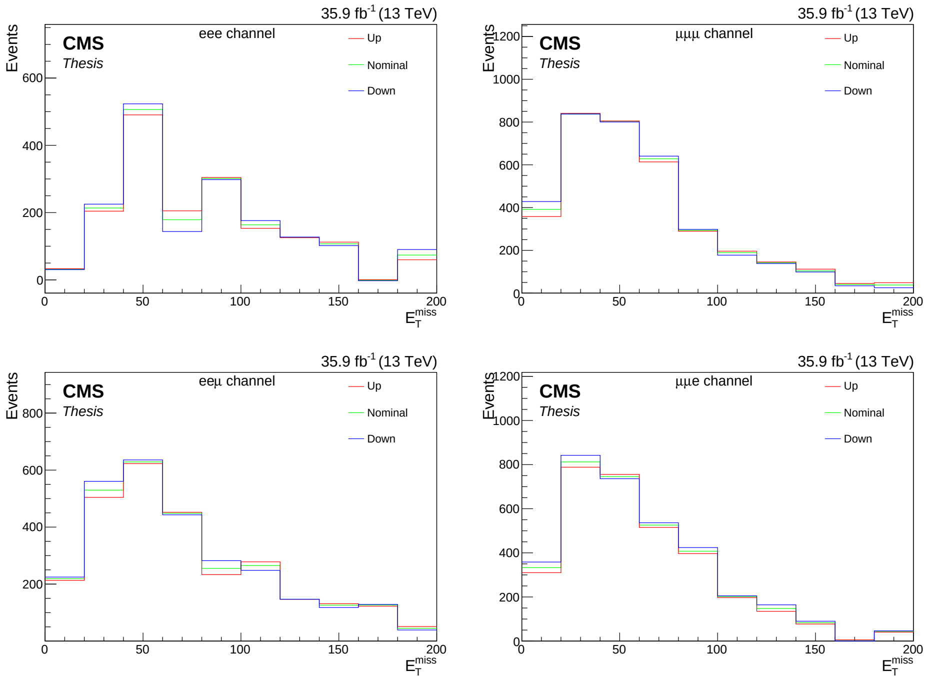


FIGURE A.9: The effect of the PU systematic uncertainty on the $\vec{\not E}_T^{miss}$ distribution for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) on the WZ background sample.

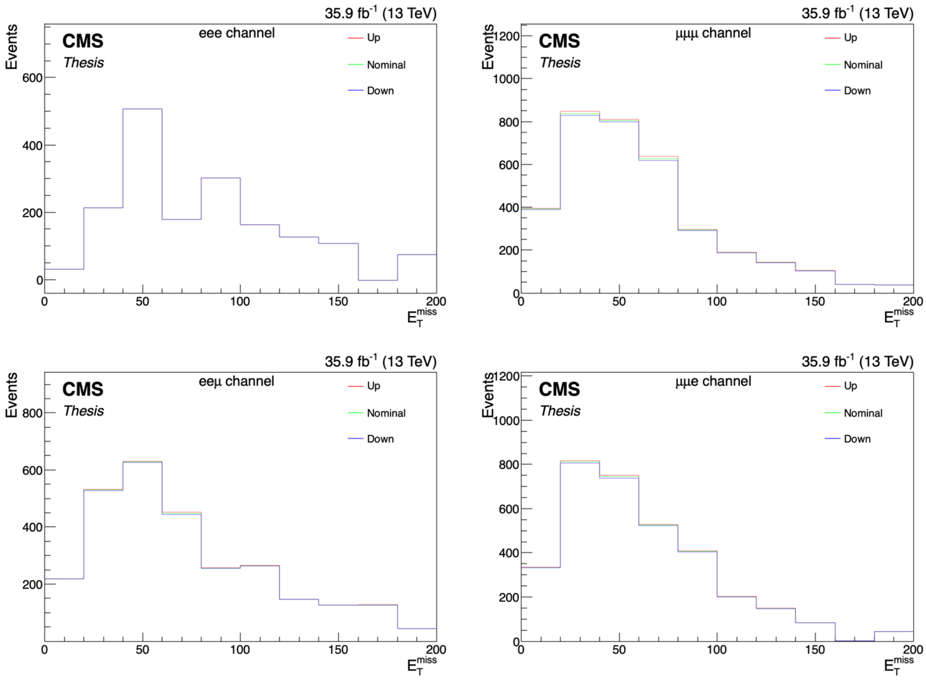


FIGURE A.10: Systematic uncertainty on the $\vec{\not{E}}_{T}^{miss}$ distribution due to muon efficiency SFs for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) on the WZ background sample.

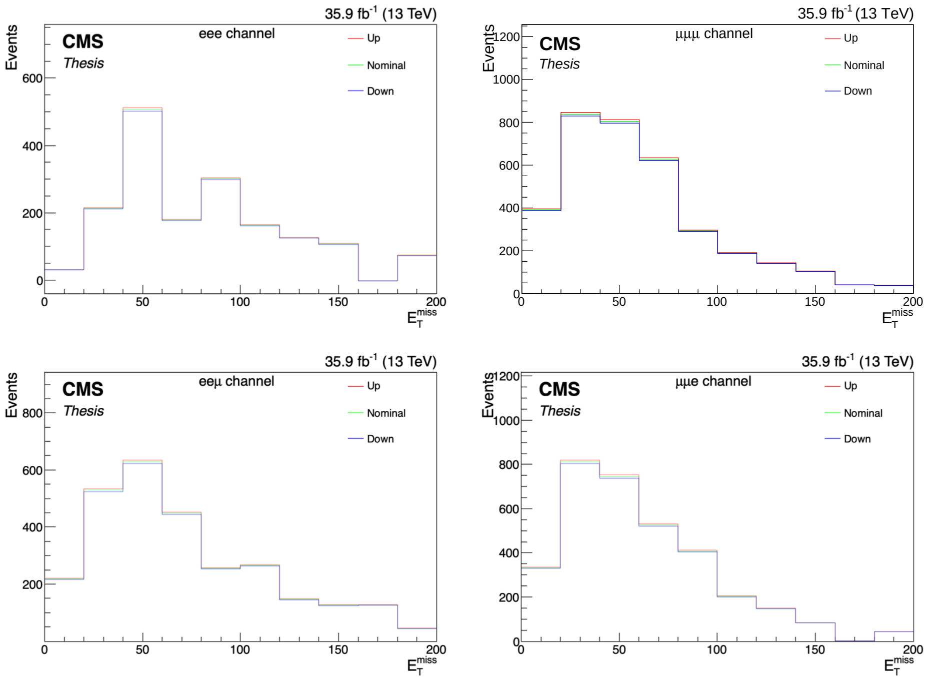


FIGURE A.11: The effect of the lepton trigger systematic uncertainty on the $\vec{\not{E}}_{T}^{miss}$ distribution for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) on the WZ background sample.

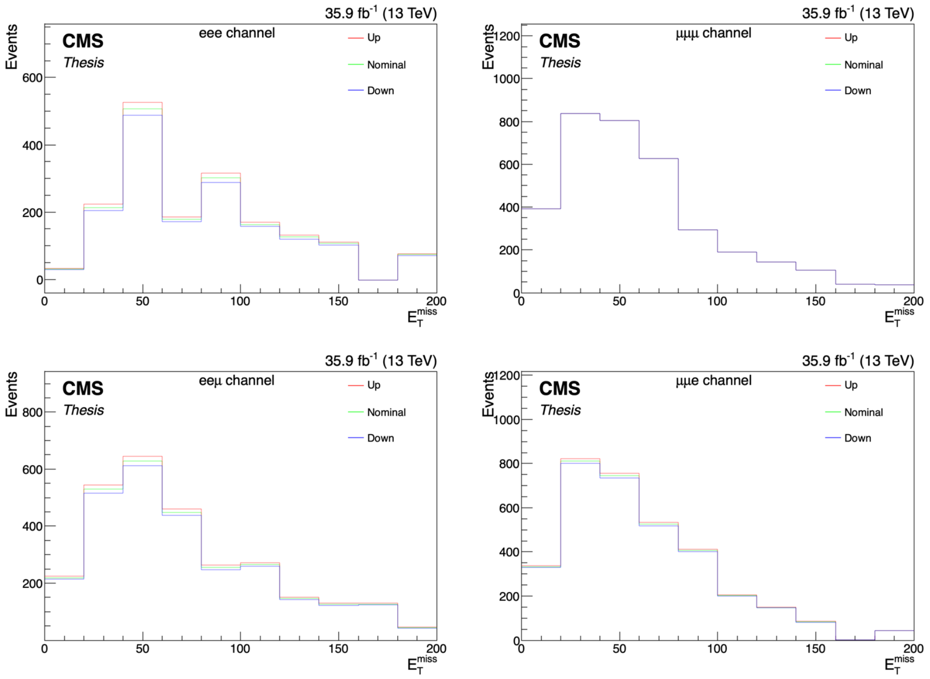


FIGURE A.12: The effect of the electron efficiency systematic uncertainty on the $\vec{\not{E}}_T^{miss}$ distribution for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) on the WZ background sample.

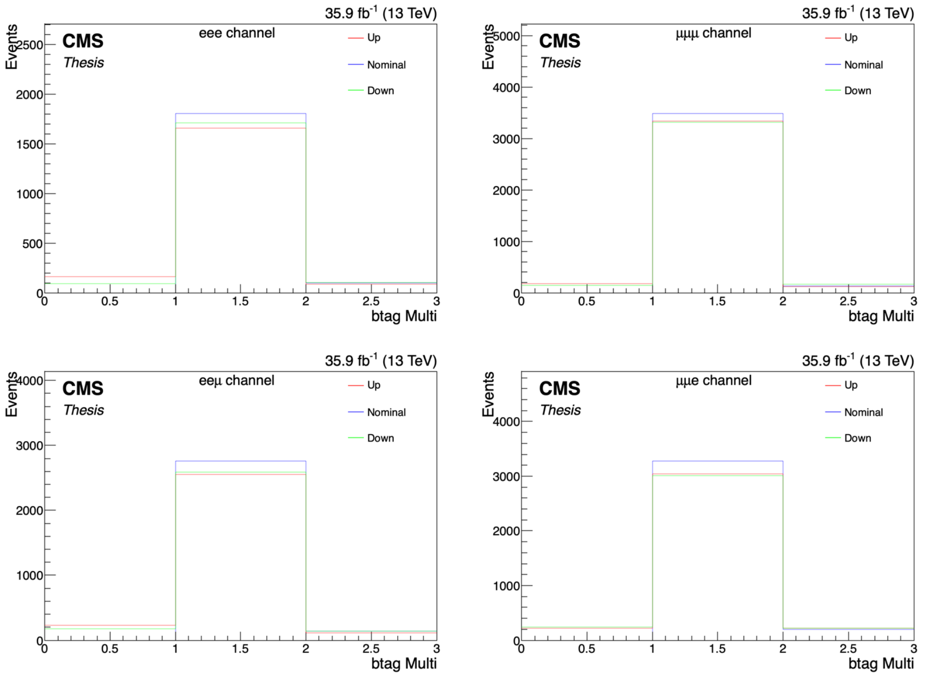


FIGURE A.13: B-tagging systematic uncertainty on the b-tag multiplicity distribution due to nominal and varied b-tagging SFs for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) on the WZ background sample.

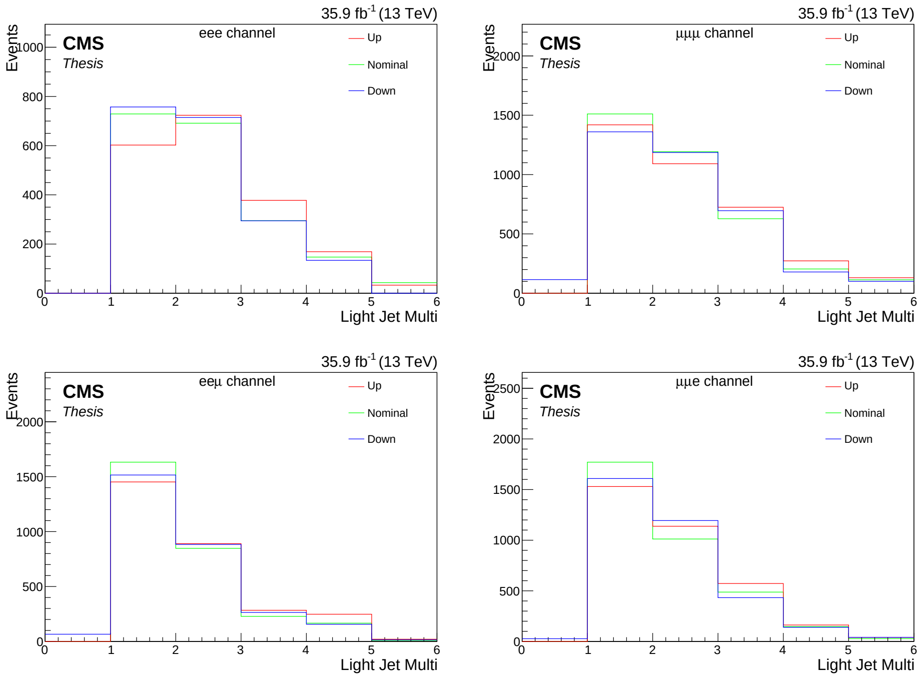


FIGURE A.14: Systematic uncertainty on the jet multiplicity distribution due to nominal and varied JES SFs for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) on the WZ background sample.

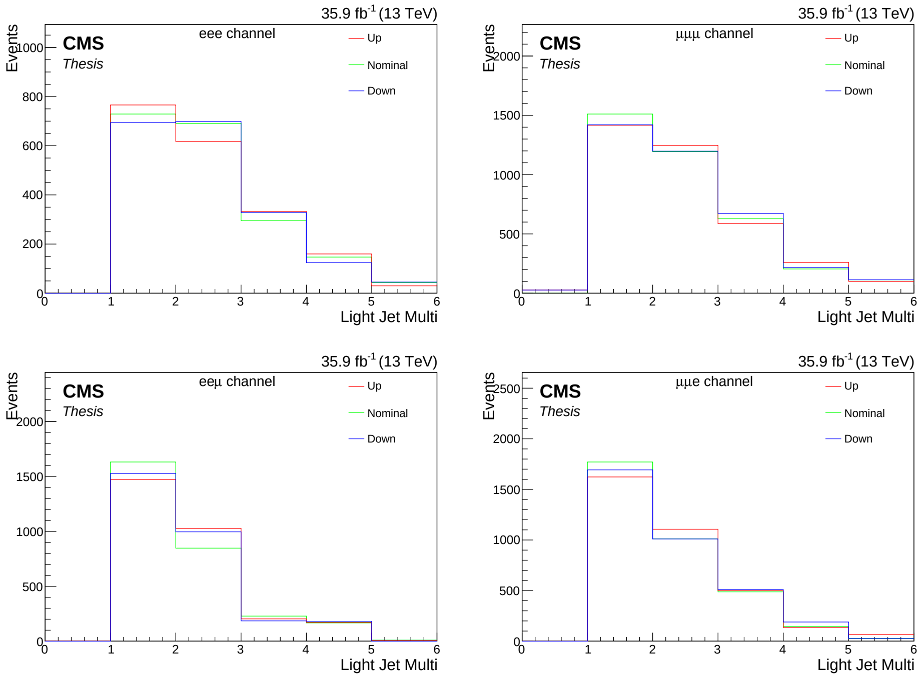


FIGURE A.15: Systematic uncertainty on the jet multiplicity distribution due to nominal and varied JER SFs for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) on the WZ background sample.

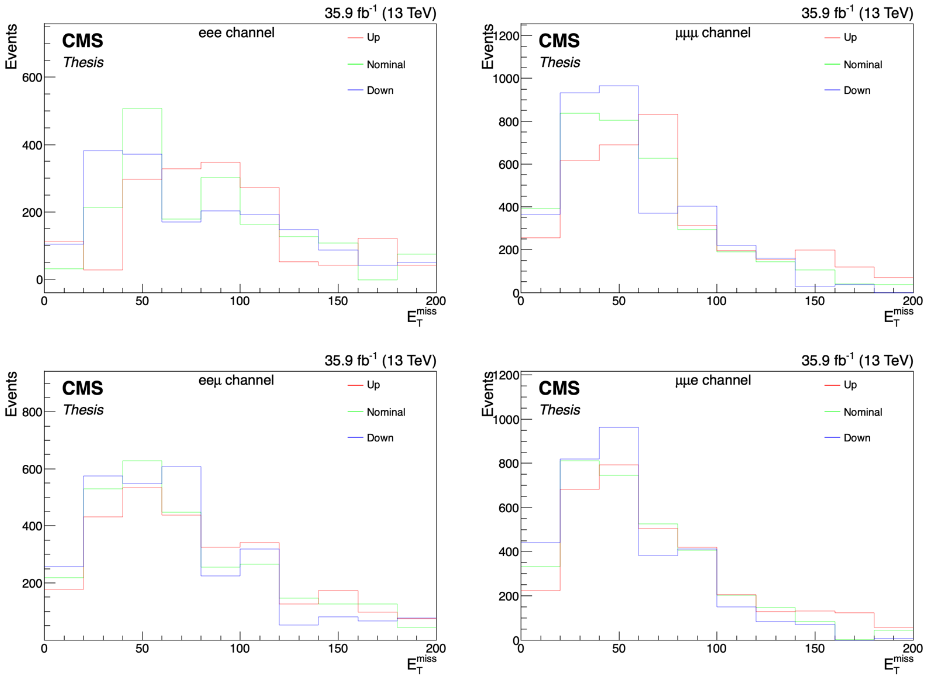


FIGURE A.16: The effect of the unclustered MET systematic uncertainty on the $\vec{\not\!E}_T^{miss}$ distribution for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) on the WZ background sample.

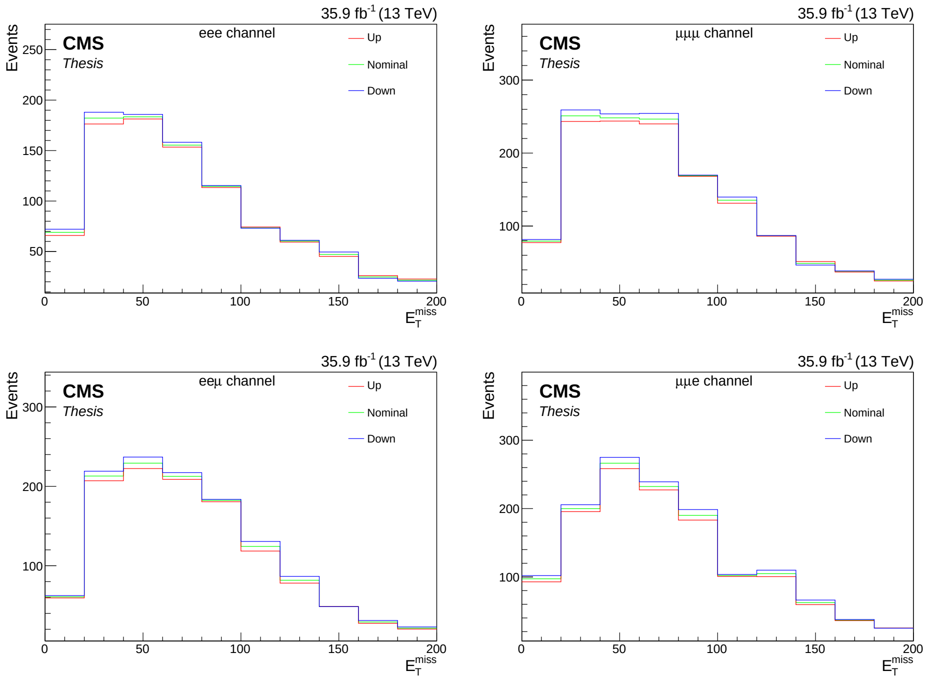


FIGURE A.17: PU systematic uncertainty on the $\vec{\not\!E}_T^{miss}$ distribution due to nominal and varied PU minimum bias cross section for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) on the $t\bar{t}Z$ sample.

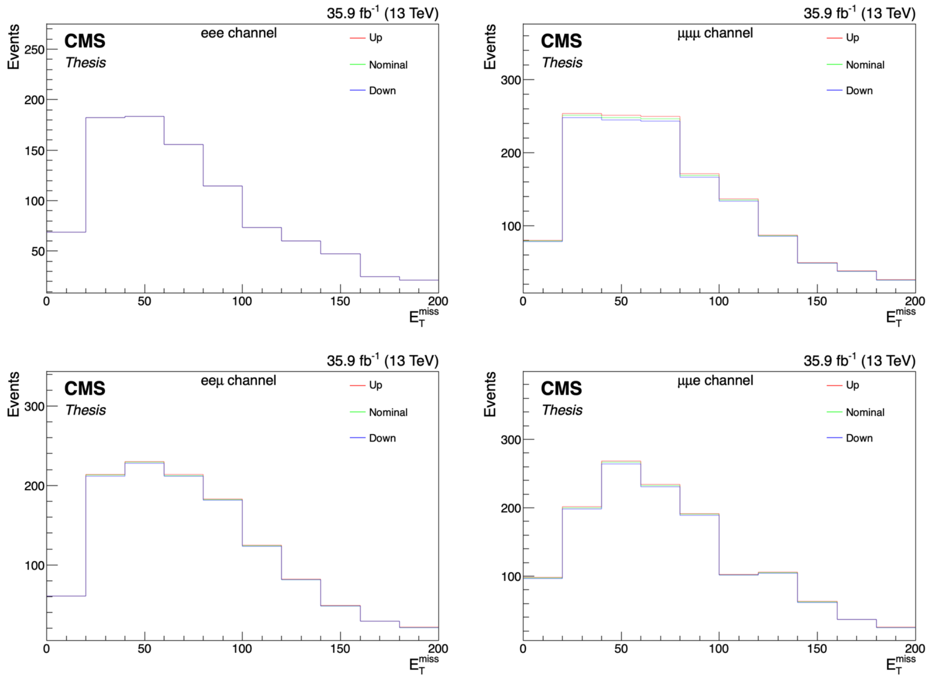


FIGURE A.18: Systematic uncertainty on the $\vec{\not{E}}_T^{miss}$ distribution due to nominal and varied muon efficiency SFs for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) on the $t\bar{t}Z$ background sample

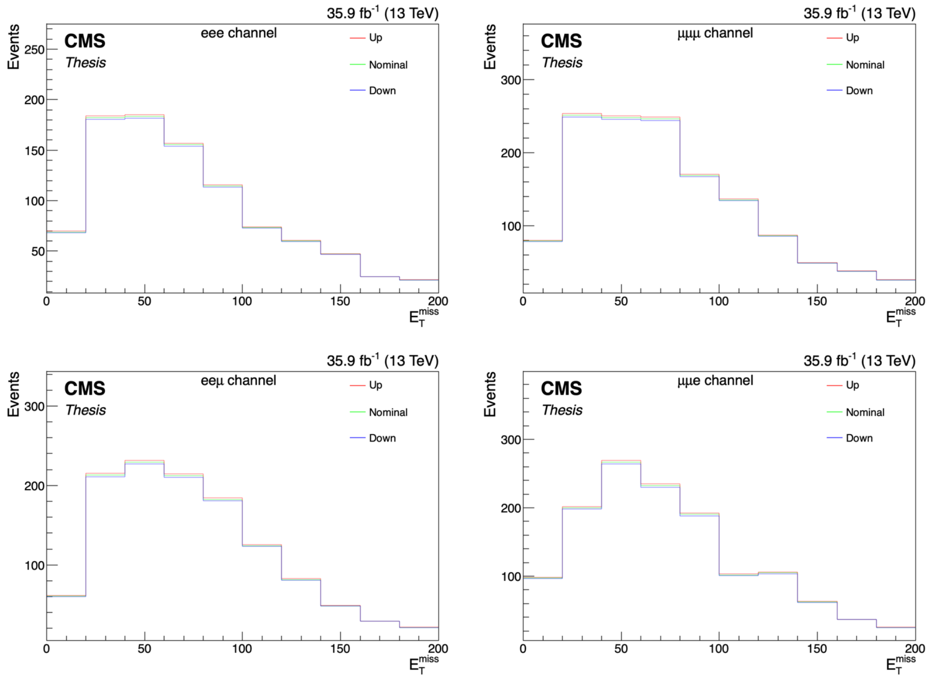


FIGURE A.19: The effect of the lepton trigger systematic uncertainty on the $\vec{\not{E}}_T^{miss}$ distribution for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) on the $t\bar{t}Z$ background sample.

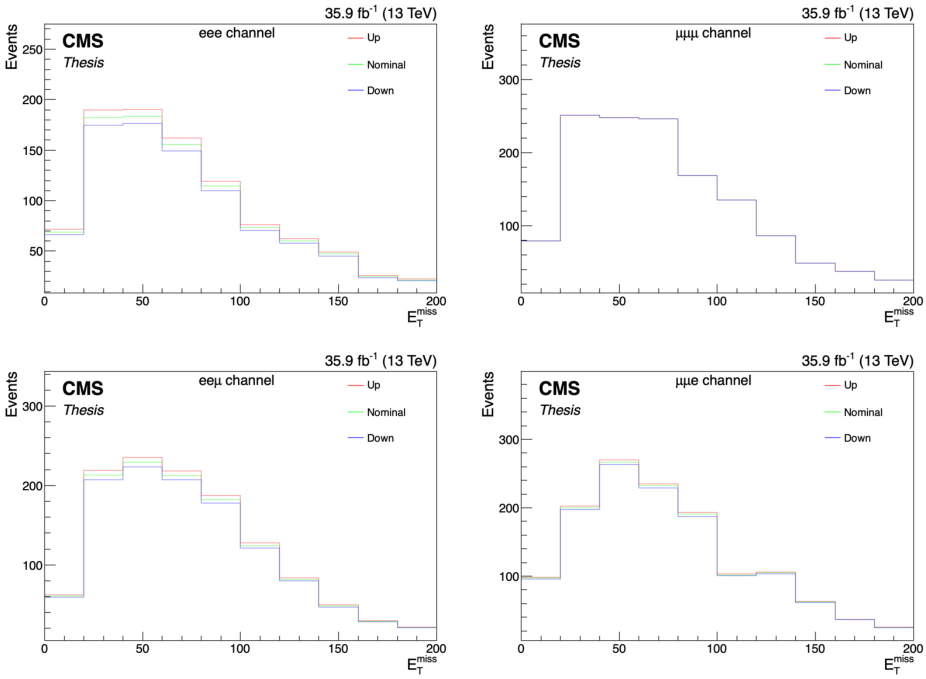


FIGURE A.20: The effect of the electron efficiency systematic uncertainty on the $\vec{\cancel{E}}_T^{miss}$ distribution for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) on the $t\bar{t}Z$ background sample.

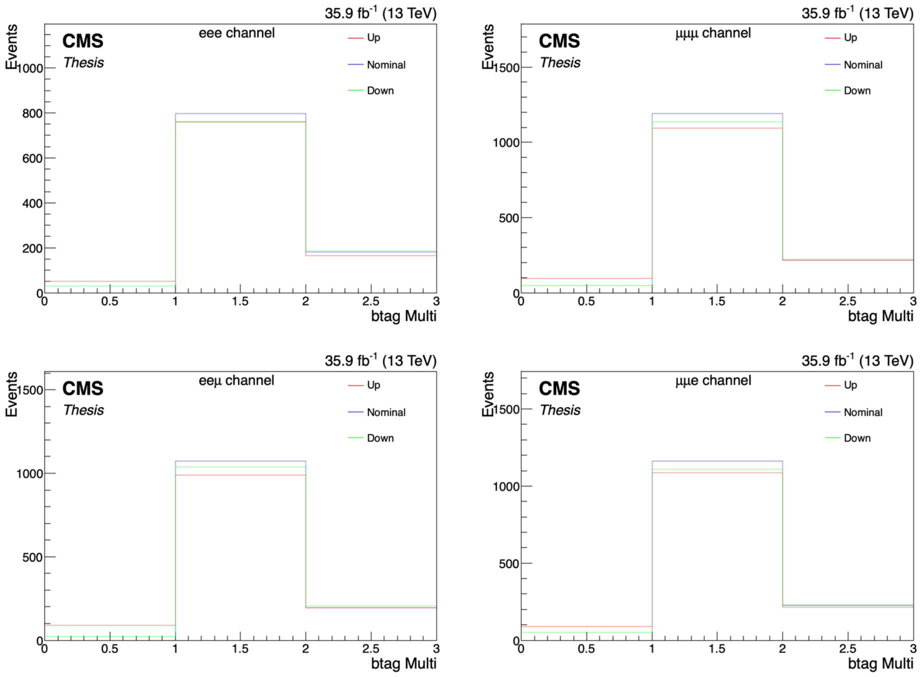


FIGURE A.21: B-tagging systematic uncertainty on the b-tag multiplicity distribution due to nominal and varied b-tagging SFs for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) on the $t\bar{t}Z$ background sample.

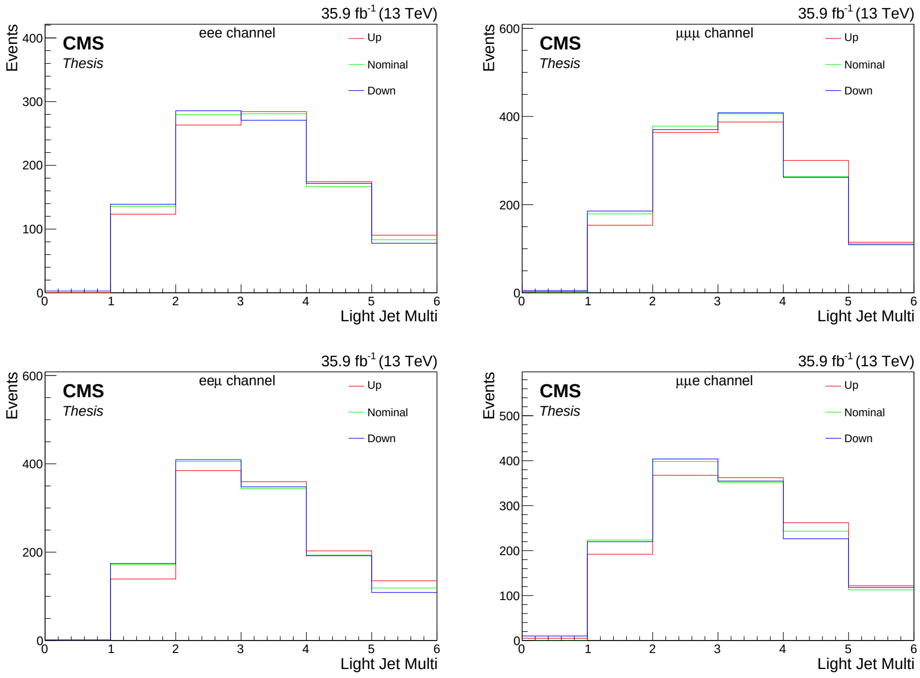


FIGURE A.22: Systematic uncertainty on the jet multiplicity distribution due to nominal and varied jet energy scale for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) on the $t\bar{t}Z$ background sample.

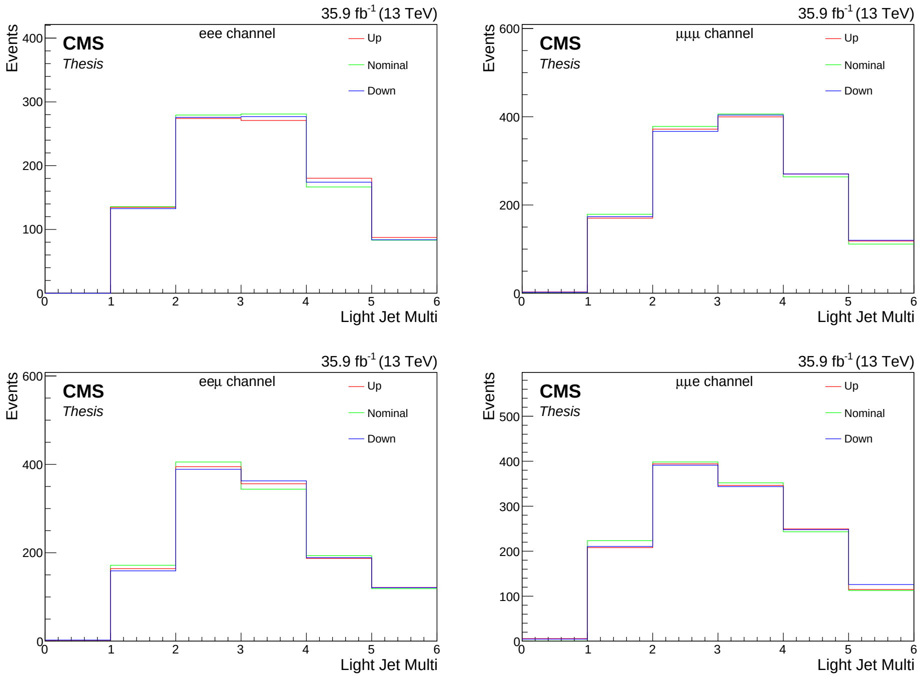


FIGURE A.23: The effect of the JER systematic uncertainty on the jet multiplicity distribution for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) on the $t\bar{t}Z$ background sample.

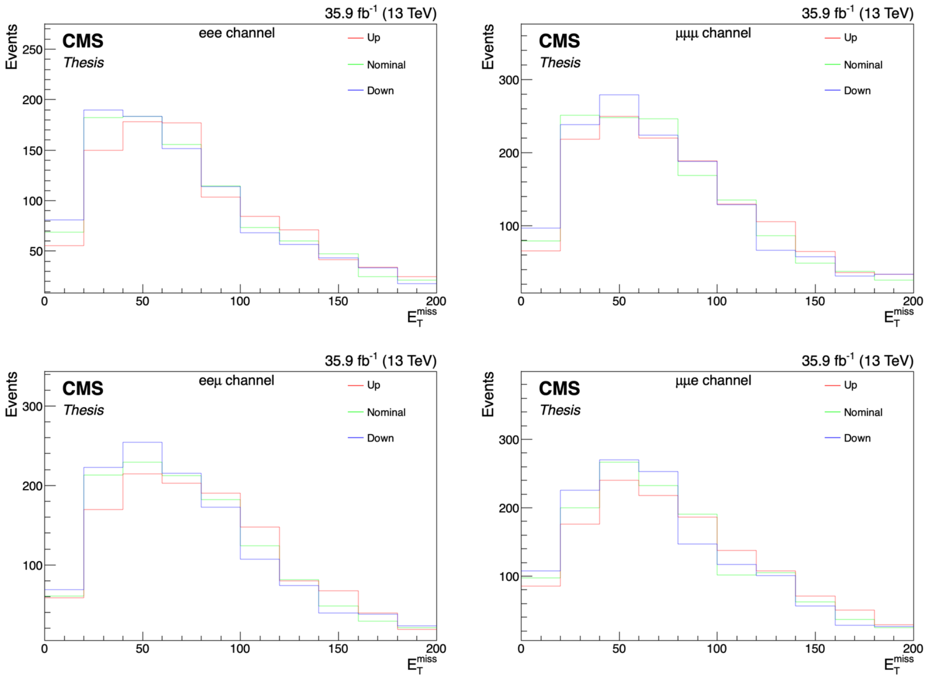


FIGURE A.24: Systematic uncertainty on the $\vec{\not{E}}_T^{miss}$ distribution due to nominal and varied unclustered missing transverse energy for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) on the $t\bar{t}Z$ background sample.

A.2 Systematics plots for 8 TeV Analysis

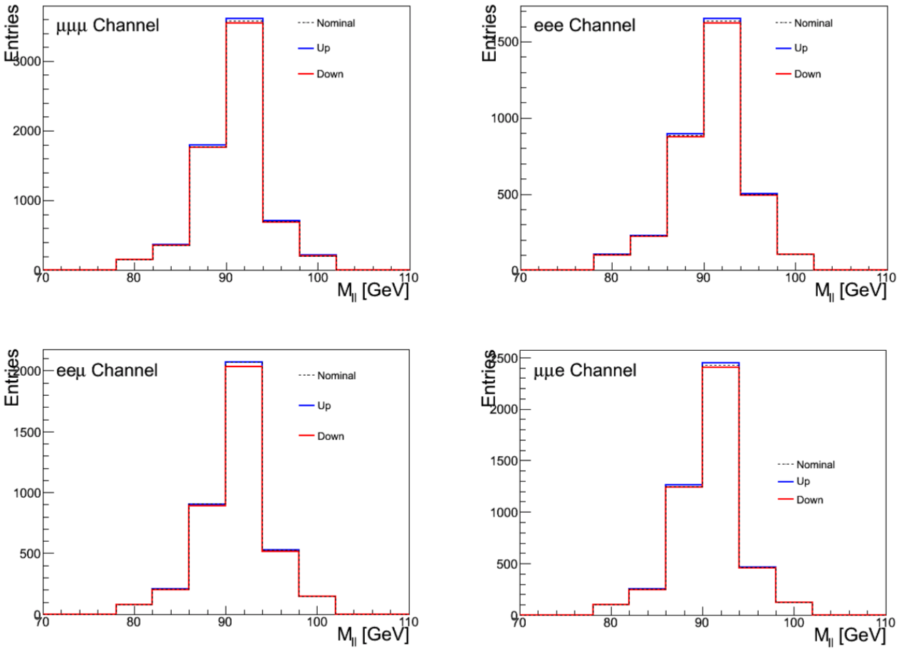


Figure A.25: Systematic uncertainty on the Z mass distribution due to nominal and varied PU minimum bias cross section for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) on the tZq signal sample.

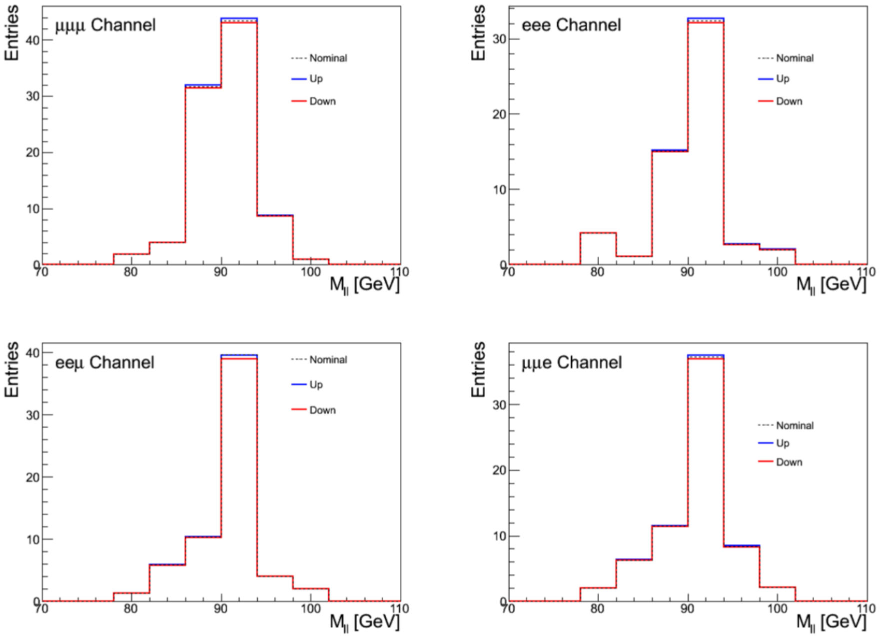


FIGURE A.26: Systematic uncertainty on the Z mass distribution due to nominal and varied PU minimum bias cross section for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) on the $t\bar{t}Z$ background sample.

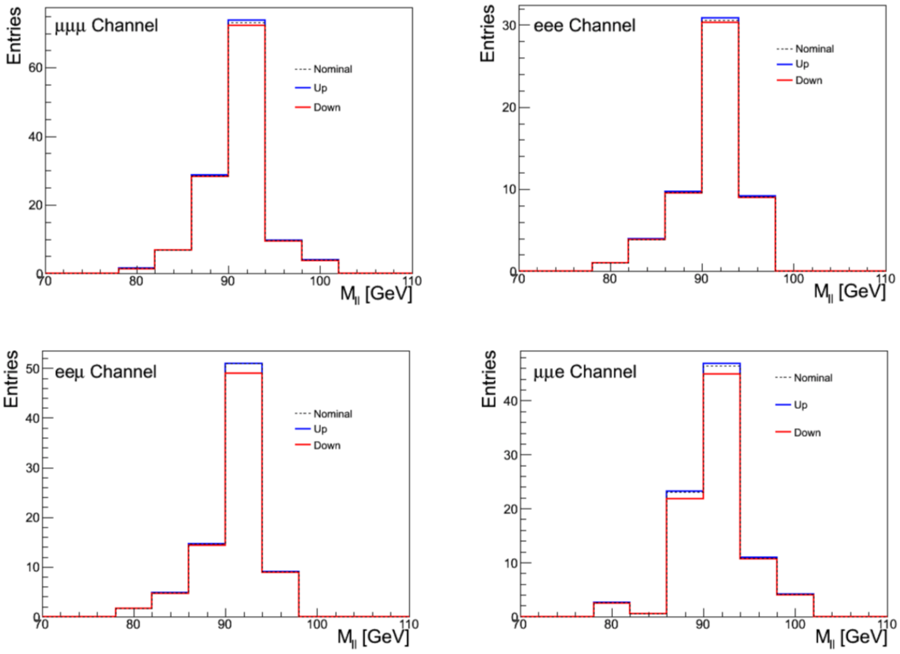


FIGURE A.27: Systematic uncertainty on the Z mass distribution due to nominal and varied PU minimum bias cross section for all four lepton channels ($\mu\mu\mu$, eee, $ee\mu$, $\mu\mu e$) on the WZ background sample.

Bibliography

[1] C. N. Yang and R. L. Mills. "Conservation of Isotopic Spin and Isotopic Gauge Invariance". In: *Phys. Rev.* 96 (1 1954), pp. 191–195. DOI: `10.1103/PhysRev.96.191`. URL: `https://link.aps.org/doi/10.1103/PhysRev.96.191`.

[2] A. Einstein. "Die Grundlage der allgemeinen Relativitätstheorie". In: *Annalen der Physik* 354 (1916), pp. 769–822. DOI: `10.1002/andp.19163540702`.

[3] J. Beringer. "Review of Particle Physics". In: *Phys. Rev. D* 86 (1 July 2012), p. 010001. DOI: `10.1103/PhysRevD.86.010001`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.86.010001`.

[4] S. L. Glashow. "Partial Symmetries of Weak Interactions". In: *Nucl. Phys.* 22 (1961), pp. 579–588. DOI: `10.1016/0029-5582(61)90469-2`.

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

[6] Steven Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (21 1967), pp. 1264–1266. DOI: `10.1103/PhysRevLett.19.1264`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.19.1264`.

[7] H. Fritzsch, Murray Gell-Mann, and H. Leutwyler. "Advantages of the Color Octet Gluon Picture". In: *Phys. Lett.* 47B (1973), pp. 365–368. DOI: `10.1016/0370-2693(73)90625-4`.

[8] H. David Politzer. "Reliable Perturbative Results for Strong Interactions?" In: *Phys. Rev. Lett.* 30 (26 1973), pp. 1346–1349. DOI: `10.1103/PhysRevLett.30.1346`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.30.1346`.

[9] David J. Gross and Frank Wilczek. "Ultraviolet Behavior of Non-Abelian Gauge Theories". In: *Phys. Rev. Lett.* 30 (26 1973), pp. 1343–1346. DOI: `10.1103/PhysRevLett.30.1343`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.30.1343`.

[10] 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`.

[11] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (9 1964), pp. 321–323. DOI: `10.1103/PhysRevLett.13.321`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.13.321`.

[12] John Ellis, Mary K. Gaillard, and Dimitri V. Nanopoulos. "A Historical Profile of the Higgs Boson". In: *The standard theory of particle physics: Essays to celebrate CERN's 60th anniversary.* Ed. by Luciano Maiani and Luigi Rolandi. 2016, pp. 255–274. DOI: `10.1142/9789814733519_0014`. arXiv: `1504.07217 [hep-ph]`.

[13] Pieter Taels. "Quantum Chromodynamics at small Bjorken-x". PhD thesis. Antwerp U., 2017. arXiv: `1711.03928 [hep-ph]`.

[14] Sayipjamal Dulat et al. "New parton distribution functions from a global analysis of quantum chromodynamics". In: *Phys. Rev. D* 93 (3 2016), p. 033006. DOI: `10.1103/PhysRevD.93.033006`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.93.033006`.

[15] Sayipjamal Dulat et al. "New parton distribution functions from a global analysis of quantum chromodynamics". In: *Phys. Rev.* D93.3 (2016), p. 033006. DOI: `10.1103/PhysRevD.93.033006`. arXiv: `1506.07443 [hep-ph]`.

[16] David d'Enterria et al. *High-precision α_s measurements from LHC to FCC-ee.* Workshop Proceedings, CERN, Geneva, Dec. 2015.

[17] Paul Langacker. "Grand unified theories and proton decay". In: *Physics Reports* 72.4 (1981), pp. 185–385. DOI: `https://doi.org/10.1016/0370-1573(81)90059-4`. URL: `http://www.sciencedirect.com/science/article/pii/0370157381900594`.

[18] Super-Kamiokande Collaboration. "Evidence for Oscillation of Atmospheric Neutrinos". In: *Phys. Rev. Lett.* 81 (8 1998), pp. 1562–1567. DOI: `10.1103/PhysRevLett.81.1562`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.81.1562`.

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

[20] CDF Collaboration. "Observation of top quark production in $\bar{p}p$ collisions". In: *Phys. Rev. Lett.* 74 (1995), pp. 2626–2631. DOI: `10.1103/PhysRevLett.74.2626`. arXiv: `hep-ex/9503002 [hep-ex]`.

[21] D0 Collaboration. "Search for high mass top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV". In: *Phys. Rev. Lett.* 74 (1995), pp. 2422–2426. DOI: `10.1103/PhysRevLett.74.2422`. arXiv: `hep-ex/9411001 [hep-ex]`.

[22] CMS Collaboration. "Measurement of the top quark mass using proton-proton data at $\sqrt{s}$=7 and 8 TeV". In: *Physical Review D* 93.7 (Apr. 2016), pp. 072004–. DOI: `10.1103/PhysRevD.93.072004`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.93.072004`.

[23] M. L. Perl et al. "Evidence for Anomalous Lepton Production in $e^+ - e^-$ Annihilation". In: *Phys. Rev. Lett.* 35 (22 1975), pp. 1489–1492. DOI: `10.1103/PhysRevLett.35.1489`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.35.1489`.

[24] C. Patrignani and Particle Data Group. "Review of Particle Physics". In: *Chinese Physics C* 40.10 (2016), p. 100001. URL: `http://stacks.iop.org/1674-1137/40/i=10/a=100001`.

[25] Michał Czakon and Alexander Mitov. "Top++: A program for the calculation of the top-pair cross-section at hadron colliders". In: *Computer Physics Communications* 185.11 (2014), pp. 2930–2938. DOI: `https://doi.org/10.1016/j.cpc.2014.06.021`. URL: `http://www.sciencedirect.com/science/article/pii/S0010465514002264`.

[26] Michal Czakon and Alexander Mitov. "Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders". In: *Comput. Phys. Commun.* 185 (2014), p. 2930. DOI: `10.1016/j.cpc.2014.06.021`. arXiv: `1112.5675 [hep-ph]`.

[27] Michał Czakon, Paul Fiedler, and Alexander Mitov. "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\pm \frac{4}{S})$". In: *Phys. Rev. Lett.* 110 (2013), p. 252004. DOI: `10.1103/PhysRevLett.110.252004`. arXiv: `1303.6254 [hep-ph]`.

[28] Kevin Lannon, Fabrizio Margaroli, and Chris Neu. "Measurements of the Production, Decay and Properties of the Top Quark: A Review". In: *Eur. Phys. J.* C72 (2012), p. 2120. DOI: `10.1140/epjc/s10052-012-2120-0`. arXiv: `1201.5873 [hep-ex]`.

[29] S. Abachi et al. "Observation of the top quark". In: *Phys. Rev. Lett.* 74 (1995), pp. 2632–2637. DOI: `10.1103/PhysRevLett.74.2632`. arXiv: `hep-ex/9503003 [hep-ex]`.

[30] Andrea Giammanco and Reinhard Schwienhorst. "Single top-quark production at the Tevatron and the LHC". In: *Rev. Mod. Phys.* 90.3 (2018), p. 035001. DOI: `10.1103/RevModPhys.90.035001`. arXiv: `1710.10699 [hep-ex]`.

[31] J. Beringer. "Review of Particle Physics". In: *Physics Letters B* 592 (1 2004), p. 010001. URL: `http://pdg.lbl.gov`.

[32] Nils Faltermann. "Measurement of the t-channel single-top quark production cross section at 13 TeV with the CMS detector". In: *Proceedings, 9th International Workshop on Top Quark Physics (TOP 2016): Olomouc, Czech Republic, September 19-23, 2016.* 2016. arXiv: `1611.08443 [hep-ex]`.

[33] Edmond L. Berger et al. "NNLO QCD corrections to t-channel single top quark production and decay". In: *Phys. Rev. D* 94 (7 Oct. 2016), p. 071501. DOI: `10.1103/PhysRevD.94.071501`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.94.071501`.

[34] Edmond L. Berger, Jun Gao, and Hua Xing Zhu. "Differential distributions for t-channel single top-quark production and decay at next-to-next-to-leading order in QCD". In: *Journal of High Energy Physics* 2017.11 (Nov. 2017), p. 158. ISSN: 1029-8479. DOI: `10.1007/JHEP11(2017)158`. URL: `https://doi.org/10.1007/JHEP11(2017)158`.

[35] Mathias Brucherseifer, Fabrizio Caola, and Kirill Melnikov. "On the NNLO QCD corrections to single-top production at the LHC". In: *Physics Letters B* 736 (2014), pp. 58 –63. ISSN: 0370-2693. DOI: `https://doi.org/10.1016/j.physletb.2014.06.075`. URL: `http://www.sciencedirect.com/science/article/pii/S0370269314004808`.

[36] M. Aliev et al. "HATHOR – HAdronic Top and Heavy quarks crOss section calculatoR". In: *Computer Physics Communications* 182.4 (2011), pp. 1034 – 1046. ISSN: 0010-4655. DOI: `https://doi.org/10.1016/j.cpc.2010.12.040`. URL: `http://www.sciencedirect.com/science/article/pii/S0010465510005333`.

[37] P. Kant et al. "HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions".

In: *Computer Physics Communications* 191 (2015), pp. 74 –89. ISSN: 0010-4655. DOI: `https://doi.org/10.1016/j.cpc.2015.02.001`. URL: `http://www.sciencedirect.com/science/article/pii/S0010465515000454`.

[38] John Campbell, R. K. Ellis, and Francesco Tramontano. "Single top-quark production and decay at next-to-leading order". In: *Phys. Rev. D* 70 (9 Nov. 2004), p. 094012. DOI: `10.1103/PhysRevD.70.094012`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.70.094012`.

[39] Nikolaos Kidonakis. "Next-to-next-to-leading logarithm resummation for s-channel single top quark production". In: *Phys. Rev. D* 81 (5 2010), p. 054028. DOI: `10.1103/PhysRevD.81.054028`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.81.054028`.

[40] Sarah Heim et al. "Next-to-leading order QCD corrections to s-channel single top quark production and decay at the LHC". In: *Phys. Rev. D* 81 (3 2010), p. 034005. DOI: `10.1103/PhysRevD.81.034005`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.81.034005`.

[41] Vardan Khachatryan et al. "Search for s channel single top quark production in pp collisions at $\sqrt{s} = 7$ and 8 TeV". In: *JHEP* 09 (2016), p. 027. DOI: `10.1007/JHEP09(2016)027`. arXiv: `1603.02555 [hep-ex]`.

[42] Tim M. P. Tait. "tW^- mode of single top quark production". In: *Phys. Rev. D* 61 (3 1999), p. 034001. DOI: `10.1103/PhysRevD.61.034001`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.61.034001`.

[43] Chris D. White et al. "Isolating Wt production at the LHC". In: *Journal of High Energy Physics* 2009.11 (2009), p. 074. URL: `http://stacks.iop.org/1126-6708/2009/i=11/a=074`.

[44] F. Cascioli et al. "A unified NLO description of top-pair and associated Wt production". In: *The European Physical Journal C* 74.3 (Mar. 2014), p. 2783. ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-014-2783-9`. URL: `https://doi.org/10.1140/epjc/s10052-014-2783-9`.

[45] Tomáš Ježo et al. "An NLO+PS generator for $t\bar{t}$ and Wt production and decay including non-resonant and interference effects". In: *The European Physical Journal C* 76.12 (2016), p. 691. ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-016-4538-2`. URL: `https://doi.org/10.1140/epjc/s10052-016-4538-2`.

[46] Nikolaos Kidonakis. "Soft-gluon corrections for tW production at N^{3}LO". In: *Phys. Rev. D* 96 (3 2017), p. 034014. DOI: `10.1103/PhysRevD.96.034014`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.96.034014`.

[47] CMS Collaboration et al. "Evidence for Associated Production of a Single Top Quark and W Boson in pp Collisions at $\sqrt{s} = 7$". In: *Physical Review Letters* 110.2 (Jan. 2013), pp. 022003–. DOI: `10.1103/PhysRevLett.110.022003`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.110.022003`.

[48] John Campbell, R. Keith Ellis, and Raoul Röntsch. "Single top production in association with a Z boson at the LHC". In: *Phys. Rev.* D87 (2013), p. 114006. DOI: `10.1103/PhysRevD.87.114006`. arXiv: `1302.3856 [hep-ph]`.

[49] V. Khachatryan et al. "Measurement of top quark–antiquark pair production in association with a W or Z boson in pp collisions at $\sqrt{s} = 8$ TeV". In: *The*

European Physical Journal C 74.9 (2014), p. 3060. DOI: `10.1140/epjc/s10052-014-3060-7`. URL: `https://doi.org/10.1140/epjc/s10052-014-3060-7`.

[50] The CMS Collaboration. *Measurement of the inclusive top-quark pair+photon production cross section in the muon + jets channel in pp collisions at* $\sqrt{s} = 8$ *TeV*. Tech. rep. CMS-PAS-TOP-13-011. Geneva: CERN, 2014. URL: `http://cds.cern.ch/record/1644573`.

[51] CMS Collaboration. "Measurement of the associated production of a single top quark and a Z boson in pp collisions at $\sqrt{s} = 13$ TeV". In: *Physics Letters B* 779 (2018), pp. 358–384. DOI: `https://doi.org/10.1016/j.physletb.2018.02.025`. URL: `http://www.sciencedirect.com/science/article/pii/S0370269318301278`.

[52] Jean-Laurent Agram et al. "Probing top anomalous couplings at the LHC with trilepton signatures in the single top mode". In: *Physics Letters B* 725.1 (2013), pp. 123–126. DOI: `https://doi.org/10.1016/j.physletb.2013.06.052`. URL: `http://www.sciencedirect.com/science/article/pii/S0370269313005273`.

[53] J. A. Aguilar-Saavedra. "Top flavor-changing neutral interactions: Theoretical expectations and experimental detection". In: *Acta Phys. Polon.* B35 (2004), pp. 2695–2710. arXiv: `hep-ph/0409342 [hep-ph]`.

[54] S. L. Glashow, J. Iliopoulos, and L. Maiani. "Weak Interactions with Lepton-Hadron Symmetry". In: *Phys. Rev. D* 2 (7 1970), pp. 1285–1292. DOI: `10.1103/PhysRevD.2.1285`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.2.1285`.

[55] Eric Ballabene. "Measurement of the $\bar{t}t$ production cross section at 13 TeV in the all-jets boosted regime with CMS at LHC". In: (Sept. 2017). DOI: `10.13140/RG.2.2.10544.20488`.

[56] *CMS Top Quark Physics Summary Figures*. 2019. URL: `https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsTOPSummaryFigures`.

[57] Lyndon Evans and Philip Bryant. "LHC Machine". In: *Journal of Instrumentation* 3.08 (2008), S08001. URL: `http://stacks.iop.org/1748-0221/3/i=08/a=S08001`.

[58] Stephen Myers. *The LEP Collider, from design to approval and commissioning*. John Adams' Lecture. Delivered at CERN, 26 Nov 1990. Geneva: CERN, 1991. URL: `https://cds.cern.ch/record/226776`.

[59] "Performance of the LHC Pre-Injectors". In: CERN-PS-2001-011-DR (2001), 6 p. URL: `https://cds.cern.ch/record/494746`.

[60] Esma Mobs. In: *Journal of Instrumentation* (2016). General Photo. URL: `https://cds.cern.ch/record/2197559`.

[61] CMS Collaboration. "The CMS experiment at the CERN LHC". In: *Journal of Instrumentation* 3.08 (2008), S08004. URL: `http://stacks.iop.org/1748-0221/3/i=08/a=S08004`.

[62] The ATLAS Collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *Journal of Instrumentation* 3.08 (2008), S08003. URL: `http://stacks.iop.org/1748-0221/3/i=08/a=S08003`.

[63] The ALICE Collaboration. “The ALICE experiment at the CERN LHC”. In: *Journal of Instrumentation* 3.08 (2008), S08002. URL: `http://stacks.iop.org/1748-0221/3/i=08/a=S08002`.

[64] The LHCb Collaboration. “The LHCb Detector at the LHC”. In: *Journal of Instrumentation* 3.08 (2008), S08005. URL: `http://stacks.iop.org/1748-0221/3/i=08/a=S08005`.

[65] J L Pinfold. “The MoEDAL Experiment”. In: *Journal of Physics: Conference Series* 631.1 (2015), p. 012014. URL: `http://stacks.iop.org/1742-6596/631/i=1/a=012014`.

[66] The TOTEM Collaboration. “The TOTEM Experiment at the CERN Large Hadron Collider”. In: *Journal of Instrumentation* 3.08 (2008), S08007. URL: `http://stacks.iop.org/1748-0221/3/i=08/a=S08007`.

[67] The LHCf Collaboration. “The LHCf detector at the CERN Large Hadron Collider”. In: *Journal of Instrumentation* 3.08 (2008), S08006. URL: `http://stacks.iop.org/1748-0221/3/i=08/a=S08006`.

[68] Saranya Samik Ghosh and on behalf of the CMS Collaboration. “Highlights from the Compact Muon Solenoid (CMS) Experiment”. In: *Universe* 5.1 (2019). ISSN: 2218-1997. DOI: `10.3390/universe5010028`. URL: `https://www.mdpi.com/2218-1997/5/1/28`.

[69] CMS Collaboration. *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*. Technical Design Report CMS. There is an error on cover due to a technical problem for some items. Geneva: CERN, 2006. URL: `http://cds.cern.ch/record/922757`.

[70] CMS Collaboration. “Detector Drawings”. CMS Collection. Mar. 2012. URL: `https://cds.cern.ch/record/1433717`.

[71] Cristina Biino. “The CMS Electromagnetic Calorimeter: overview, lessons learned during Run 1 and future projections”. In: *Journal of Physics: Conference Series* 587.1 (2015), p. 012001. URL: `http://stacks.iop.org/1742-6596/587/i=1/a=012001`.

[72] L. Brianza. “Precision crystal calorimetry in LHC Run II with the CMS ECAL”. In: *Journal of Instrumentation* 12.01 (2017), p. C01069. URL: `http://stacks.iop.org/1748-0221/12/i=01/a=C01069`.

[73] CMS Collaboration. “Identification and filtering of uncharacteristic noise in the CMS hadron calorimeter”. In: *Journal of Instrumentation* 5.03 (2010), T03014. URL: `http://stacks.iop.org/1748-0221/5/i=03/a=T03014`.

[74] The CMS collaboration. “The performance of the CMS muon detector in proton-proton collisions at $\sqrt{s} = 7$ TeV at the LHC”. In: *Journal of Instrumentation* 8.11 (2013), P11002. URL: `http://stacks.iop.org/1748-0221/8/i=11/a=P11002`.

[75] The CMS collaboration. “Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV”. In: *Journal of Instrumentation* 7.10 (2012), P10002. URL: `http://stacks.iop.org/1748-0221/7/i=10/a=P10002`.

[76] M S Kim. “CMS reconstruction improvement for the muon tracking by the RPC chambers”. In: *Journal of Instrumentation* 8.03 (2013), T03001. URL: `http://stacks.iop.org/1748-0221/8/i=03/a=T03001`.

[77] Sergio Cittolin, Attila Rácz, and Paris Sphicas. *CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and data-acquisition project.* Technical Design Report CMS. Geneva: CERN, 2002. URL: `http://cds.cern.ch/record/578006`.

[78] T. Bawej et al. "The New CMS DAQ System for Run-2 of the LHC". In: *IEEE Transactions on Nuclear Science* 62.3 (June 2015), pp. 1099–1103. ISSN: 0018-9499. DOI: `10.1109/TNS.2015.2426216`.

[79] Claudio Grandi et al. *CMS Computing Model: The "CMS Computing Model RTAG"*. Tech. rep. CMS-NOTE-2004-031. CERN-LHCC-2004-035. LHCC-G-083. Geneva: CERN, 2004. URL: `http://cds.cern.ch/record/814248`.

[80] *The Large Hadron Collider's worldwide computer.* 2013. URL: `https://sciencenode.org/feature/large-hadron-colliders-worldwide-computer.php`.

[81] S Banerjee. "CMS Simulation Software". In: *Journal of Physics: Conference Series* 396.2 (2012), p. 022003. URL: `http://stacks.iop.org/1742-6596/396/i=2/a=022003`.

[82] M Hildreth et al. "CMS Full Simulation for Run II". In: *Journal of Physics: Conference Series* 664.7 (2015), p. 072022. URL: `http://stacks.iop.org/1742-6596/664/i=7/a=072022`.

[83] S. Agostinelli et al. "Geant4—a simulation toolkit". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. DOI: `https://doi.org/10.1016/S0168-9002(03)01368-8`. URL: `http://www.sciencedirect.com/science/article/pii/S0168900203013688`.

[84] Rahmat Rahmat, Rob Kroeger, and Andrea Giammanco. "The Fast Simulation of The CMS Experiment". In: *Journal of Physics: Conference Series* 396.6 (2012), p. 062016. URL: `http://stacks.iop.org/1742-6596/396/i=6/a=062016`.

[85] Andy Buckley et al. "General-purpose event generators for LHC physics". In: *Physics Reports* 504.5 (2011), pp. 145–233. DOI: `https://doi.org/10.1016/j.physrep.2011.03.005`. URL: `http://www.sciencedirect.com/science/article/pii/S0370157311000846`.

[86] Andy Buckley et al. "General-purpose event generators for LHC physics". In: *Physics Reports* 504.5 (2011), pp. 145–233. DOI: `https://doi.org/10.1016/j.physrep.2011.03.005`. URL: `http://www.sciencedirect.com/science/article/pii/S0370157311000846`.

[87] Adrian Signer. *Low-energy Precision Physics and the High-energy Frontier.* Vol. 51. Dec. 2014, pp. 25–30. DOI: `10.1016/j.phpro.2013.12.007`.

[88] J. Alwall et al. "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions". In: *The European Physical Journal C* 53.3 (2008), pp. 473–500. ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-007-0490-5`. URL: `https://doi.org/10.1140/epjc/s10052-007-0490-5`.

[89] G. Altarelli and G. Parisi. "Asymptotic freedom in parton language". In: *Nuclear Physics B* 126.2 (1977), pp. 298 –318. ISSN: 0550-3213. DOI: `https://doi.org/`

10.1016/0550-3213(77)90384-4. URL: http://www.sciencedirect.com/science/article/pii/0550321377903844.

[90] B. Andersson et al. "Parton fragmentation and string dynamics". In: *Physics Reports* 97.2 (1983), pp. 31 –145. ISSN: 0370-1573. DOI: https://doi.org/10.1016/0370-1573(83)90080-7. URL: http://www.sciencedirect.com/science/article/pii/0370157383900807.

[91] Torbjörn Sjöstrand. "Jet fragmentation of multiparton configurations in a string framework". In: *Nuclear Physics B* 248.2 (1984), pp. 469–502. DOI: https://doi.org/10.1016/0550-3213(84)90607-2. URL: http://www.sciencedirect.com/science/article/pii/0550321384906072.

[92] Bo Andersson, Sandipan Mohanty, and Fredrik Soderberg. "Recent developments in the Lund model". In: *36th Annual Winter School on Nuclear and Particle Physics (PINP 2002) and 8th St. Petersburg School on Theoretical Physics St. Petersburg, Russia, February 25-March 3, 2002*. 2002. arXiv: hep-ph/0212122 [hep-ph].

[93] Torbjörn Sjöstrand. "Jet fragmentation of multiparton configurations in a string framework". In: *Nuclear Physics B* 248.2 (1984), pp. 469–502. DOI: https://doi.org/10.1016/0550-3213(84)90607-2. URL: http://www.sciencedirect.com/science/article/pii/0550321384906072.

[94] Rick D. Field. "The Underlying event in hard scattering processes". In: *eConf* C010630 (2001), P501. arXiv: hep-ph/0201192 [hep-ph].

[95] J. Alwall et al. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations". In: *Journal of High Energy Physics* 2014.7 (2014), p. 79. DOI: 10.1007/JHEP07(2014)079. URL: https://doi.org/10.1007/JHEP07(2014)079.

[96] Johan Alwall et al. "MadGraph 5: going beyond". In: *Journal of High Energy Physics* 2011.6 (2011), p. 128. DOI: 10.1007/JHEP06(2011)128. URL: https://doi.org/10.1007/JHEP06(2011)128.

[97] R. Frederix et al. "Scalar and pseudoscalar Higgs production in association with a top–antitop pair". In: *Physics Letters B* 701.4 (2011), pp. 427–433. DOI: https://doi.org/10.1016/j.physletb.2011.06.012. URL: http://www.sciencedirect.com/science/article/pii/S0370269311006332.

[98] Michelangelo L. Mangano et al. "Matching matrix elements and shower evolution for top-pair production in hadronic collisions". In: *Journal of High Energy Physics* 2007.01 (2007), p. 013. URL: http://stacks.iop.org/1126-6708/2007/i=01/a=013.

[99] Stefano Frixione and Bryan R. Webber. "Matching NLO QCD computations and parton shower simulations". In: *Journal of High Energy Physics* 2002.06 (2002), p. 029. URL: http://stacks.iop.org/1126-6708/2002/i=06/a=029.

[100] J. Alwall et al. "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions". In: *The European Physical Journal C* 53.3 (2008), pp. 473–500. DOI: 10.1140/epjc/s10052-007-0490-5. URL: https://doi.org/10.1140/epjc/s10052-007-0490-5.

[101] Rikkert Frederix and Stefano Frixione. "Merging meets matching in MC@NLO". In: *Journal of High Energy Physics* 2012.12 (2012), p. 61. DOI: `10.1007/JHEP12(2012)061`. URL: `https://doi.org/10.1007/JHEP12(2012)061`.

[102] Torbjörn Sjöstrand et al. "High-energy-physics event generation with Pythia 6.1". In: *Computer Physics Communications* 135.2 (2001), pp. 238–259. DOI: `https://doi.org/10.1016/S0010-4655(00)00236-8`. URL: `http://www.sciencedirect.com/science/article/pii/S0010465500002368`.

[103] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. "A brief introduction to PYTHIA 8.1". In: *Computer Physics Communications* 178.11 (2008), pp. 852–867. DOI: `https://doi.org/10.1016/j.cpc.2008.01.036`. URL: `http://www.sciencedirect.com/science/article/pii/S0010465508000441`.

[104] Simone Alioli et al. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX". In: *Journal of High Energy Physics* 2010.6 (2010), p. 43. DOI: `10.1007/JHEP06(2010)043`. URL: `https://doi.org/10.1007/JHEP06(2010)043`.

[105] Simone Alioli et al. "NLO single-top production matched with shower in POWHEG: s- and t-channel contributions". In: *Journal of High Energy Physics* 2009.09 (2009), p. 111. URL: `http://stacks.iop.org/1126-6708/2009/i=09/a=111`.

[106] Stefano Frixione, Paolo Nason, and Carlo Oleari. "Matching NLO QCD computations with parton shower simulations: the POWHEG method". In: *Journal of High Energy Physics* 2007.11 (2007), p. 070. URL: `http://stacks.iop.org/1126-6708/2007/i=11/a=070`.

[107] Manuel Bähr et al. "Herwig++ physics and manual". In: *The European Physical Journal C* 58.4 (2008), pp. 639–707. DOI: `10.1140/epjc/s10052-008-0798-9`. URL: `https://doi.org/10.1140/epjc/s10052-008-0798-9`.

[108] Johannes Bellm et al. "Herwig 7.0/Herwig++ 3.0 release note". In: *The European Physical Journal C* 76.4 (2016), p. 196. DOI: `10.1140/epjc/s10052-016-4018-8`. URL: `https://doi.org/10.1140/epjc/s10052-016-4018-8`.

[109] B. R. Webber. "A QCD model for jet fragmentation including soft gluon interference". In: *Nuclear Physics B* 238.3 (1984), pp. 492–528. DOI: `https://doi.org/10.1016/0550-3213(84)90333-X`. URL: `http://www.sciencedirect.com/science/article/pii/055032138490333X`.

[110] D. Amati and G. Veneziano. "Preconfinement as a property of perturbative QCD". In: *Physics Letters B* 83.1 (1979), pp. 87–92. DOI: `https://doi.org/10.1016/0370-2693(79)90896-7`. URL: `http://www.sciencedirect.com/science/article/pii/0370269379908967`.

[111] E. Boos et al. "CompHEP 4.4—automatic computations from Lagrangians to events". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 534.1 (2004), pp. 250–259. DOI: `https://doi.org/10.1016/j.nima.2004.07.096`. URL: `http://www.sciencedirect.com/science/article/pii/S0168900204015402`.

[112] E. E. Boos et al. "Method for simulating electroweak top-quark production events in the NLO approximation: SingleTop event generator". In: *Physics of*

Atomic Nuclei 69.8 (2006), pp. 1317–1329. DOI: `10.1134/S1063778806080084`. URL: `https://doi.org/10.1134/S1063778806080084`.

[113] S. Jadach et al. “The τ decay library TAUOLA, version 2.4”. In: *Computer Physics Communications* 76.3 (1993), pp. 361–380. DOI: `https://doi.org/10.1016/0010-4655(93)90061-G`. URL: `http://www.sciencedirect.com/science/article/pii/001046559390061G`.

[114] N. Davidson et al. “Universal interface of TAUOLA: Technical and physics documentation”. In: *Computer Physics Communications* 183.3 (2012), pp. 821–843. DOI: `https://doi.org/10.1016/j.cpc.2011.12.009`. URL: `http://www.sciencedirect.com/science/article/pii/S0010465511003973`.

[115] P. Golonka and Z. Was. “PHOTOS Monte Carlo: a precision tool for QED correctionsin Z and W decays”. In: *The European Physical Journal C - Particles and Fields* 45.1 (2006), pp. 97–107. DOI: `10.1140/epjc/s2005-02396-4`. URL: `https://doi.org/10.1140/epjc/s2005-02396-4`.

[116] The CMS collaboration. “Reconstruction and identification of τ lepton decays to hadrons and ν_τ at CMS”. In: *Journal of Instrumentation* 11.01 (2016), P01019. URL: `http://stacks.iop.org/1748-0221/11/i=01/a=P01019`.

[117] The CMS collaboration. “Performance of photon reconstruction and identification with the CMS detector in proton-proton collisions at $\sqrt{s}$= 8 TeV”. In: 10.08 (2015), P08010–P08010. DOI: `10.1088/1748-0221/10/08/p08010`. URL: `http://dx.doi.org/10.1088/1748-0221/10/08/P08010`.

[118] CERN. “Detector-taking a closer look at lhc”. In: (2017).

[119] W Adam et al. “Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC”. In: *Journal of Physics G: Nuclear and Particle Physics* 31.9 (2005), N9. URL: `http://stacks.iop.org/0954-3899/31/i=9/a=N01`.

[120] R. E. Kalman. “A New Approach to Linear Filtering and Prediction Problems”. In: *Journal of Basic Engineering* 82.1 (Mar. 1960), pp. 35–45. URL: `http://dx.doi.org/10.1115/1.3662552`.

[121] Pierre Billoir. “Progressive track recognition with a Kalman-like fitting procedure”. In: *Computer Physics Communications* 57.1 (1989), pp. 390–394. DOI: `https://doi.org/10.1016/0010-4655(89)90249-X`. URL: `http://www.sciencedirect.com/science/article/pii/001046558990249X`.

[122] E. J Wolin and L. L Ho. “Covariance matrices for track fitting with the Kalman filter”. In: *Nuclear Instruments and Methods in Physics Research* 329.3 (1993), pp. 493–500. DOI: `https://doi.org/10.1016/0168-9002(93)91285-U`. URL: `http://www.sciencedirect.com/science/article/pii/016890029391285U`.

[123] S. Cucciarelli et al. “Track reconstruction, primary vertex finding and seed generation with the pixel detector”. In: *CERN-CMS-NOTE-2006-026* (2006).

[124] The CMS collaboration. “Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV”. In: *Journal of Instrumentation* 7.10 (2012), P10002. URL: `http://stacks.iop.org/1748-0221/7/i=10/a=P10002`.

[125] “Single Muon efficiencies in 2012 Data”. In: (2013). URL: `https://cds.cern.ch/record/1536406`.

[126] "Muon Identification and Isolation efficiency on full 2016 dataset". In: (2017). URL: `http://cds.cern.ch/record/2257968`.

[127] The CMS collaboration. "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV". In: *Journal of Instrumentation* 7.10 (2012), P10002. URL: `http://stacks.iop.org/1748-0221/7/i=10/a=P10002`.

[128] "Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV". In: *Journal of Instrumentation* 10.06 (2015), P06005. URL: `http://stacks.iop.org/1748-0221/10/i=06/a=P06005`.

[129] "Electron and photon performance in CMS with the full 2016 data sample." In: (2017). URL: `http://cds.cern.ch/record/2255497`.

[130] Matteo Cacciari and Gavin P. Salam. "Pileup subtraction using jet areas". In: *Physics Letters B* 659.1 (2008), pp. 119–126. DOI: `https://doi.org/10.1016/j.physletb.2007.09.077`. URL: `http://www.sciencedirect.com/science/article/pii/S0370269307011094`.

[131] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The anti-kt jet clustering algorithm". In: *Journal of High Energy Physics* 2008.04 (2008), p. 063. URL: `http://stacks.iop.org/1126-6708/2008/i=04/a=063`.

[132] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The anti- k t jet clustering algorithm". In: *Journal of High Energy Physics* 2008.04 (2008), p. 063. URL: `http://stacks.iop.org/1126-6708/2008/i=04/a=063`.

[133] *TOP JetMET Analysis*. 2017. URL: `https://twiki.cern.ch/twiki/bin/viewauth/CMS/TopJME`.

[134] CMS Collaboration. "Jet energy scale and resolution in the CMS experiment in pp collisions at $\sqrt{s} = 8$ TeV". In: *Journal of Instrumentation* 12.02 (2017), P02014. URL: `http://stacks.iop.org/1748-0221/12/i=02/a=P02014`.

[135] The CMS collaboration. "Identification of b-quark jets with the CMS experiment". In: *Journal of Instrumentation* 8.04 (2013), P04013. URL: `http://stacks.iop.org/1748-0221/8/i=04/a=P04013`.

[136] *Identification of b quark jets at the CMS Experiment in the LHC Run II*. Tech. rep. CMS-PAS-BTV-15-001. Geneva: CERN, 2016. URL: `http://cds.cern.ch/record/2138504`.

[137] P. Bites. *Diagram showing the common principle of identification of jets initiated by b-hadron decays*. 2017. URL: `https://i1.wp.com/www.particlebites.com/wp-content/uploads/2016/ 08/B-tagging_diagram.png`. (visited on 2017).

[138] Vardan Khachatryan et al. "Search for the associated production of the Higgs boson with a top-quark pair". In: *JHEP* 09 (2014). [Erratum: JHEP10,106(2014)], p. 087. DOI: `10.1007/JHEP09(2014)087, 10.1007/JHEP10(2014)106`. arXiv: `1408.1682 [hep-ex]`.

[139] The CMS collaboration. "Missing transverse energy performance of the CMS detector". In: *Journal of Instrumentation* 6.09 (2011), P09001. URL: `http://stacks.iop.org/1748-0221/6/i=09/a=P09001`.

[140] The CMS Collaboration. “CMS Physics CMS Physics : Technical Design Report Volume 1: Detector Performance and Software ”. In: *CERN-LHCC-2006-001, CMS-TDR-8-1* (2006).

[141] The CMS collaboration. “Alignment of the CMS tracker with LHC and cosmic ray data”. In: *Journal of Instrumentation* 9.06 (2014), P06009. URL: `http://stacks.iop.org/1748-0221/9/i=06/a=P06009`.

[142] W. Erdmann. “The CMS pixel detector”. In: *Int. J. Mod. Phys.* A25 (2010). [,41(2010)], pp. 1315–1337. DOI: `10.1142/S0217751X10049098`.

[143] H. Chr. Kästli et al. “Design and performance of the CMS pixel detector readout chip”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 565.1 (2006), pp. 188–194. DOI: `https://doi.org/10.1016/j.nima.2006.05.038`. URL: `http://www.sciencedirect.com/science/article/pii/S0168900206007674`.

[144] János Karancsi. *Operational Experience with the CMS Pixel Detector*. Vol. 10. Nov. 2014. DOI: `10.1088/1748-0221/10/05/C05016`.

[145] CMS Collaboration. *CMS, tracker technical design report*. 1998. URL: `http://weblib.cern.ch/abstract?CERN-LHCC-98-6`.

[146] M. J. French et al. “Design and results from the APV25, a deep sub-micron CMOS front-end chip for the CMS tracker”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 466.2 (2001), pp. 359–365. DOI: `https://doi.org/10.1016/S0168-9002(01)00589-7`. URL: `http://www.sciencedirect.com/science/article/pii/S0168900201005897`.

[147] CMS Collaboration. “Commissioning and performance of the CMS silicon strip tracker with cosmic ray muons”. In: *Journal of Instrumentation* 5.03 (2010), T03008. URL: `http://stacks.iop.org/1748-0221/5/i=03/a=T03008`.

[148] *A model to explain Deco vs Peak changes in alignment geometries*. 2015. URL: `http://venturia.home.cern.ch/venturia/GlobalRuns/PeakVsDeco/PeakVsDeco.html`.

[149] *SiStrip Back Plane Correction Analysis*. 2015. URL: `https://twiki.cern.ch/twiki/bin/viewauth/CMS/BackPlaneCorrections`.

[150] Higgs Combine working group. *Documentation of the RooStats-based statistics tools for Higgs PAG*. Accessed: 2018-06-30. URL: `https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideHiggsAnalysisCombinedLimit`.

[151] CMS Collaboration. “Combined results of searches for the standard model Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV”. In: *Physics Letters B* 710.1 (2012), pp. 26–48. DOI: `https://doi.org/10.1016/j.physletb.2012.02.064`. URL: `http://www.sciencedirect.com/science/article/pii/S0370269312002055`.

[152] *Procedure for the LHC Higgs boson search combination in Summer 2011*. Tech. rep. CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11. Geneva: CERN, 2011. URL: `https://cds.cern.ch/record/1379837`.

[153] L. Moneta et al. “The RooStats project”. In: *Proceedings of the 13th International Workshop on Advanced Computing.* 2010, p. 57. arXiv: `1009.1003 [physics.data-an]`.

[154] Harrison B. Prosper and Louis Lyons, eds. *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues.* CERN. Geneva: CERN, 2011. ISBN: 9789290833673. DOI: `10.5170/CERN-2011-006`.

[155] Thomas Junk. “Confidence level computation for combining searches with small statistics”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 434.2 (1999), pp. 435–443. DOI: `https://doi.org/10.1016/S0168-9002(99)00498-2`. URL: `http://www.sciencedirect.com/science/article/pii/S0168900299004982`.

[156] Glen Cowan et al. “Asymptotic formulae for likelihood-based tests of new physics”. In: *Eur. Phys. J.* C71 (2011). [Erratum: Eur. Phys. J.C73,2501(2013)], p. 1554. DOI: `10.1140/epjc/s10052-011-1554-0, 10.1140/epjc/s10052-013-2501-z`. arXiv: `1007.1727 [physics.data-an]`.

[157] Thorsten Chwalek. “Messung der $W - Boson - Helizittsanteile$ in Top Quark $Zerfllen$ mit dem CDF II Experiment und Studien zu einer frühen Messung des ttbar-Wirkungsquerschnitts mit dem CMS Experiment”. Presented 12 Feb 2010. 2010. URL: `https://cds.cern.ch/record/1416031`.

[158] Jens Hansen. “Study of Single Top Quark Production at the CMS Experiment”. PhD thesis. KIT, Karlsruhe, 2012.

[159] J. Pumplin et al. “New generation of parton distributions with uncertainties from global QCD analysis”. In: *JHEP* 07 (2002), p. 012. DOI: `10.1088/1126-6708/2002/07/012`. arXiv: `hep-ph/0201195 [hep-ph]`.

[160] Richard D. Ball et al. “Parton distributions with LHC data”. In: *Nucl. Phys.* B867 (2013), pp. 244–289. DOI: `10.1016/j.nuclphysb.2012.10.003`. arXiv: `1207.1303 [hep-ph]`.

[161] The NNPDF collaboration et al. “Parton distributions for the LHC run II”. In: *Journal of High Energy Physics* 2015.4 (2015), p. 40. ISSN: 1029-8479. DOI: `10.1007/JHEP04(2015)040`. URL: `https://doi.org/10.1007/JHEP04(2015)040`.

[162] *CMS TWIKI.* 2016. URL: `https://cms-service-dqm.web.cern.ch/cms-service-dqm/CAF/certification/Collisions12/8TeV/Prompt/Cert_190456-208686_8TeV_PromptReco_Collisions12_JSON.txt`.

[163] Johan Alwall et al. “MadGraph 5: going beyond”. In: *Journal of High Energy Physics* 2011.6 (2011), p. 128. DOI: `10.1007/JHEP06(2011)128`. URL: `https://doi.org/10.1007/JHEP06(2011)128`.

[164] Wolfgang Waltenberger, Rudolf Frühwirth, and Pascal Vanlaer. “Adaptive vertex fitting”. In: *Journal of Physics G: Nuclear and Particle Physics* 34.12 (2007), N343. URL: `http://stacks.iop.org/0954-3899/34/i=12/a=N01`.

[165] *CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update.* Tech. rep. CMS-PAS-LUM-13-001. Geneva: CERN, 2013. URL: `http://cds.cern.ch/record/1598864`.

[166] *Utilities for Accessing Pileup Information for Data.* 2016. URL: `https://twiki.cern.ch/twiki/bin/view/CMS/PileupJSONFileforData`.

[167] The CMS collaboration. "Determination of jet energy calibration and transverse momentum resolution in CMS". In: *Journal of Instrumentation* 6.11 (2011), P11002. URL: `http://stacks.iop.org/1748-0221/6/i=11/a=P11002`.

[168] Serguei Chatrchyan et al. "Identification of b-quark jets with the CMS experiment". In: *JINST* 8 (2013), P04013. DOI: `10.1088/1748-0221/8/04/P04013`. arXiv: `1211.4462 [hep-ex]`.

[169] Lorenzo Moneta et al. "The RooStats Project". In: *PoS* ACAT2010 (2010), p. 057. DOI: `10.22323/1.093.0057`. arXiv: `1009.1003 [physics.data-an]`.

[170] A. M. Sirunyan et al. "Search for associated production of a Z boson with a single top quark and for tZ flavour-changing interactions in pp collisions at $\sqrt{s} = 8$ TeV". In: *Journal of High Energy Physics* 2017.7 (2017), p. 3. DOI: `10.1007/JHEP07(2017)003`. URL: `https://doi.org/10.1007/JHEP07(2017)003`.

[171] Stefano Frixione, Paolo Nason, and Carlo Oleari. "Matching NLO QCD computations with parton shower simulations: the POWHEG method". In: *Journal of High Energy Physics* 2007.11 (2007), p. 070. URL: `http://stacks.iop.org/1126-6708/2007/i=11/a=070`.

[172] Paolo Nason. "A new method for combining NLO QCD with shower Monte Carlo algorithms". In: *Journal of High Energy Physics* 2004.11 (2004), p. 040. URL: `http://stacks.iop.org/1126-6708/2004/i=11/a=040`.

[173] Simone Alioli et al. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX". In: *Journal of High Energy Physics* 2010.6 (2010), p. 43. DOI: `10.1007/JHEP06(2010)043`. URL: `https://doi.org/10.1007/JHEP06(2010)043`.

[174] Emanuele Re. "Single-top Wt-channel production matched with parton showers using the POWHEG method". In: *The European Physical Journal C* 71.2 (2011), p. 1547. DOI: `10.1140/epjc/s10052-011-1547-z`. URL: `https://doi.org/10.1140/epjc/s10052-011-1547-z`.

[175] Torbjörn Sjöstrand et al. "An introduction to PYTHIA 8.2". In: *Computer Physics Communications* 191 (2015), pp. 159–177. DOI: `https://doi.org/10.1016/j.cpc.2015.01.024`. URL: `http://www.sciencedirect.com/science/article/pii/S0010465515000442`.

[176] V. Khachatryan et al. "Event generator tunes obtained from underlying event and multiparton scattering measurements". In: *The European Physical Journal C* 76.3 (2016), p. 155. DOI: `10.1140/epjc/s10052-016-3988-x`. URL: `https://doi.org/10.1140/epjc/s10052-016-3988-x`.

[177] S. Agostinelli et al. "Geant4—a simulation toolkit". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. DOI: `https://doi.org/10.1016/S0168-9002(03)01368-8`. URL: `http://www.sciencedirect.com/science/article/pii/S0168900203013688`.

[178] *CMS Top Trigger (Run2)*. 2018. URL: `https://twiki.cern.ch/twiki/bin/view/CMS/TopTrigger`.

[179] *Baseline muon selections for Run-II*. 2017. URL: `https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2#Muon_Identification`.

[180] *TOP Muon Information for Analysis (Run2)*. 2018. URL: `https://twiki.cern.ch/twiki/bin/view/CMS/TopMUO`.

[181] *CMS Top EGamma Coordination (Run2)*. 2017. URL: `https://twiki.cern.ch/twiki/bin/view/CMS/TopEGM`.

[182] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. “The anti-kt jet clustering algorithm”. In: *Journal of High Energy Physics* 2008.04 (2008), p. 063. URL: `http://stacks.iop.org/1126-6708/2008/i=04/a=063`.

[183] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. “FastJet user manual”. In: *The European Physical Journal C* 72.3 (2012), p. 1896. DOI: `10.1140/epjc/s10052-012-1896-2`. URL: `https://doi.org/10.1140/epjc/s10052-012-1896-2`.

[184] CMS Collaboration. “Identification of b quark jets at the CMS Experiment in the LHC Run 2”. In: (2016).

[185] The CMS collaboration. “Identification of b-quark jets with the CMS experiment”. In: *Journal of Instrumentation* 8.04 (2013), P04013. URL: `http://stacks.iop.org/1748-0221/8/i=04/a=P04013`.

[186] *Utilities for Accessing Pileup Information for Data*. 2018. URL: `https://twiki.cern.ch/twiki/bin/view/CMS/PileupJSONFileforData`.

[187] *Jet energy scale uncertainty sources*. 2016. URL: `https://twiki.cern.ch/twiki/bin/view/CMS/JECUncertaintySources`.

[188] *E/Gamma ID Tasks*. 2017. URL: `https://twiki.cern.ch/twiki/bin/view/CMS/EgammaIDRecipesRun2#Electron_efficiencies_and_scale`.

[189] *Muon reconstruction, identification, isolation and trigger efficiencies*. 2018. URL: `https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonReferenceEffsRun2#Results_for_2016_data`.

[190] CMS Collaboration. “CMS Luminosity Measurements for the 2016 Data Taking Period”. In: (2017).

[191] Andy Buckley et al. “LHAPDF6: parton density access in the LHC precision era”. In: *Eur. Phys. J.* C75 (2015), p. 132. DOI: `10.1140/epjc/s10052-015-3318-8`. arXiv: `1412.7420 [hep-ph]`.

[192] CMS Collaboration. “Investigations of the impact of the parton shower tuning in Pythia 8 in the modelling of $t\bar{t}$ at $\sqrt{s} = 8$ and 13 TeV”. In: (2016).

[193] *Procedure for the LHC Higgs boson search combination in summer 2011*. Tech. rep. ATL-PHYS-PUB-2011-011, CMS-NOTE-2011-005. ATLAS, CMS, LHC Higgs Combination Group, 2011.

[194] Glen Cowan et al. “Asymptotic formulae for likelihood-based tests of new physics”. In: *The European Physical Journal C* 71.2 (2011), p. 1554. DOI: `10.1140/epjc/s10052-011-1554-0`. URL: `https://doi.org/10.1140/epjc/s10052-011-1554-0`.

[195] Albert M Sirunyan et al. “Observation of single top quark production in association with a Z boson in proton-proton collisions at $\sqrt{s} = 13$ TeV”. In: *Submitted to: Phys. Rev. Lett.* (2018). arXiv: `1812.05900 [hep-ex]`.