



**W boson helicity measurement in $t\bar{t}$ di-electron
channel with the CMS detector at the LHC
and
the CMS outer tracker upgrade for the HL-LHC**

Hadi Behnamian

Supervisor:
Dr. Mohsen Khakzad

This thesis is presented as part of the requirements for the conferral of the degree:

Doctor of Philosophy in experimental Particle Physics

Institute For Research in Fundamental Sciences (IPM)
School of Particles and Accelerators

December, 2015



Acknowledgements

First and foremost, praises and thanks to the God, the Almighty, for His showers of blessings throughout my research work to complete the research successfully. Secondly, I feel deeply indebted to express my sincere appreciation to everyone who helped me scientifically and emotionally through out this study.

This thesis was carried out at the Institute for Research in Fundamental Sciences (IPM), school of Particles and Accelerators from 2010 to 2015. I would like to thank to Prof. Hessameddin Arfaei, the IPM team leader in CERN program, for providing me the conditions which allowed to fully achieve the goals I was challenged to. I convey my gratitude to my supervisor, Dr. Mohsen Khakzad, for a great deal of useful advice and for his support. It has been my pleasure to collaborate with Dr. Mojtaba Mohammadi Najafabadi and Dr. Mohsen Naseri in data analysis, for the fruitful discussions we had and for all the clarifications they kindly gave me.

I am also very grateful to the CMS Top working group and the Top Properties subgroup members, in particular, to Dr. Andrea Giammanco, Dr. Andreas Meyer, Dr. Rebeca Gonzalez Suarez, Dr. Jacob Thomas Linacre, and Dr. Thorsten Chwalek for all their support, feedback, and direction that has lead to the results discussed in this thesis. Many thanks to the whole CMS tracker CMX group and SSD detector laboratory people, especially to Dr. Anna Elliott-Peisert and Dr. Stefano Mersi for leading me throughout research and development studies on phase II sensor upgrade projects for the CMS tracker outer barrel part.

I have to thank to all my working colleagues and friends, in particular, to Mr. Mohsen Dayyani Kelissani, Dr. Reza Goldouzian, Dr. Morteza Khatiri Yanesari, Dr. Mohsen Naseri, Dr. Ferdos Rezaei Hosseinabadi, and Dr. Shahin Sanaye Hajari, who have shared the experience and effort of these years with me. I am also grateful for all of IPM administrations at Particles and Accelerators school, who helped making my research activities better.

Last, but certainly not least, I would like to thank my parents for their endless encouragement, as well as my wife, whose unfailing patience and love have made setbacks softer and successes sweeter. This work would not have been possible without them.

Abstract

This thesis presents a measurement of the W boson helicity fractions in the $t\bar{t}$ di-electron final state. The top pair events are produced in proton-proton collisions at a center-of-mass energy of 8 TeV, collected by the CMS experiment at the Large Hadron Collider (LHC), corresponding to an integrated luminosity of 19.7 fb^{-1} . Only events with two energetic opposite-sign electrons, at least two b-tagged jets and large missing transverse energy in the final state, are considered as di-electron $t\bar{t}$ signal. The $t\bar{t}$ event is fully reconstructed using the Analytical Matrix Weighting Technique (AMWT).

The helicity fractions of the W boson are estimated from the $\cos\theta_1^*$ distribution using an event-by-event re-weighting technique. Within the quoted statistic and systematic uncertainties, the results are in good agreement with the Standard Model prediction at the 95% confidence level. The results are also consistent with other W boson polarization measurements in the CMS and the ATLAS experiments.

While the W boson helicity fraction measurements are the main topic of this thesis, some important features of new silicon strip sensors for the CMS outer tracker upgrade are also investigated in this study.

Keywords: CMS detector, W boson helicity fractions, $t\bar{t}$, Silicon strip sensors.

Contents

Abstract	iii
1 Introduction	1
1.1 Elementary particles	2
1.2 Fundamental interactions	4
1.2.1 The Electromagnetic interaction	4
1.2.2 The Weak interaction and Electroweak unification	5
1.2.3 The Strong interaction	6
1.2.4 The Gravitation interaction	7
1.3 Theoretical framework of Standard Model	8
1.3.1 Gauge groups	8
1.3.2 Higgs mechanism	12
1.4 Top-quark Physics	15
1.4.1 Top quark production	15
1.4.2 Top quark decay modes	17
1.5 W boson polarization in top-quark decay	19
1.5.1 Anomalous couplings in tWb vertex	23
1.5.2 Measurement of W boson polarization	26
2 Experimental apparatus	29
2.1 The Large Hadron Collider	29
2.2 The Compact Muon Solenoid experiment	33
2.2.1 The CMS coordinate system	35
2.2.2 Tracking system	35
2.2.3 Electromagnetic calorimeter	39
2.2.4 Hadronic calorimeter	42
2.2.5 Solenoid magnet	45
2.2.6 Muon chambers	45
2.2.7 Data acquisition and trigger system	48

3	Sensor R&D for the CMS outer tracker upgrade	53
3.1	The wafer and materials	54
3.1.1	Irradiation	56
3.2	Measurement setup and results	57
3.2.1	Charge collection	57
3.2.2	Study of random ghost hits	60
3.3	Summary of results	61
4	Generation, simulation and object reconstruction	64
4.1	Event generation chain	65
4.1.1	Hard process generation	65
4.1.2	Parton shower	66
4.1.3	Hadronization	66
4.1.4	Underlying event and pileup	67
4.1.5	Detector simulation	67
4.2	Monte Carlo generators	68
4.2.1	MadGraph and MadEvent	68
4.2.2	PowHeg and MC@NLO	68
4.2.3	Pythia and Herwig	69
4.3	Physics object reconstruction	69
4.3.1	Reconstruction of tracks	70
4.3.2	Vertex reconstruction	71
4.3.3	Particle Flow algorithm	71
4.3.4	Reconstruction of electrons	72
4.3.5	Reconstruction of muons	73
4.3.6	Reconstruction of jets	74
4.3.7	b-tagging	77
4.3.8	Missing transverse energy	79
5	Event selection and background estimation	81
5.1	Data and Monte Carlo samples	82
5.2	Event selection	83
5.2.1	Trigger selection	85
5.2.2	Event cleaning	85
5.2.3	Electron selection	86
5.2.4	Jet selection and b-tag requirement	87
5.2.5	Missing transverse energy cut	88
5.2.6	$t\bar{t}$ event selection	88
5.3	Corrections to Monte Carlo events	89
5.3.1	Trigger efficiency	90

5.3.2	Electron identification and isolation efficiencies	91
5.3.3	B-tagging efficiency and scale factor	93
5.4	Drell-Yan background estimation	95
5.5	Data-MC comparison	97
5.6	Full reconstruction of the $t\bar{t}$ events	99
5.6.1	Minimum effective mass of $t\bar{t}$ approach	100
5.6.2	Analytical Matrix Weighting Technique (AMWT)	101
5.6.3	Kinematics method	104
5.6.4	Top quark p_T re-weighting	106
5.7	Event yield and control plots	107
5.8	Fit re-weighting method	109
5.9	Validation of the fitting method	112
5.9.1	Linearity check of method	112
5.9.2	Statistical properties of the estimators	114
6	Systematic uncertainties and results	116
6.1	Electron identification and trigger efficiencies	117
6.2	B-tagging scale factors	117
6.3	Backgrounds normalization	117
6.4	Jet energy resolution and jet energy scale	118
6.5	Pileup effect	118
6.6	Integrated luminosity	118
6.7	Top quark p_T re-weighting	119
6.8	The model dependent uncertainties	119
6.8.1	Top quark mass	119
6.8.2	Factorization and renormalization scales	119
6.8.3	Jet-parton matching scale	119
6.8.4	Hadronization model	120
6.8.5	Parton distribution function uncertainty	120
6.9	Limited statistics of signal sample	120
6.10	Systematics conclusion	121
6.11	Final results	123
6.12	Conclusions	124
	Bibliography	126
A	Amplitude calculation of top quark decay	137
B	Data-MC comparison plots for combination channel	139

Chapter 1

Introduction

One of the most extensively studied areas of physics during the last decades has been the quest for a theory which could describe the entire particle content of our universe together with the interactions amongst them at all possible energies. The Standard Model (SM) of elementary particle physics [1–3], effectively describes all known particles in our universe and is able to include the electromagnetic, the weak, and the strong nuclear forces. Despite success of the SM, there are some puzzling features of observations remaining unexplained. The role of experimental physicist in the continuing pursuit of a complete description of fundamental particles is to test theories by observation, and to cross-check the SM with new observations which could signify new physics. The SM is theoretically inconsistent with general relativity [4], Einstein’s well-tested theory of gravity. Then the SM is not regarded as the ultimate theory of the fundamental particles although gravity is too weak to play a significant role at the masses and energies considered in the particle physics.

This chapter continues with a description of elementary particles, section 1.1. An overview of particle physics interactions are described in section 1.2. Within the SM of elementary particle physics, all mutual interactions among particles are described in the framework of relativistic quantum field theories. In this framework, the fields can interact among each other by requiring the theory to be invariant under local gauge transformations. The theoretical interpretation of particle physics interactions as well as mechanism of spontaneous symmetry breaking, which gives mass to all particles, will be discussed in section 1.3. This chapter continues with discussion of the heaviest particle within the SM, the top quark, in section 1.4. In this section, an overview of the current theoretical and experimental status within the field of top quark physics are given. This chapter will be concluded with studying basic theory on the W boson polarization, introducing anomalous tWb vertex and providing latest results on the W boson polarization measurements in different experiments.

1.1 Elementary particles

The SM of particle physics is a theory which explains the constituents of the matter and their interactions. According to spin property, the elementary particles are divided to two main categories: fermions with a spin of $1/2$ which considered to built up all known matter, and bosons with integer spin values responsible for interactions. An overview of all elementary particles can be found in Figure 1.1.

QUARKS	mass → charge → spin →	$\approx 2.3 \text{ MeV}/c^2$ $2/3$ $1/2$ u up	$\approx 1.275 \text{ GeV}/c^2$ $2/3$ $1/2$ c charm	$\approx 173.07 \text{ GeV}/c^2$ $2/3$ $1/2$ t top	0 0 1 g gluon	$\approx 126 \text{ GeV}/c^2$ 0 0 H Higgs boson
		$\approx 4.8 \text{ MeV}/c^2$ $-1/3$ $1/2$ d down	$\approx 95 \text{ MeV}/c^2$ $-1/3$ $1/2$ s strange	$\approx 4.18 \text{ GeV}/c^2$ $-1/3$ $1/2$ b bottom	0 0 1 γ photon	
		$0.511 \text{ MeV}/c^2$ -1 $1/2$ e electron	$105.7 \text{ MeV}/c^2$ -1 $1/2$ μ muon	$1.777 \text{ GeV}/c^2$ -1 $1/2$ τ tau	$91.2 \text{ GeV}/c^2$ 0 1 Z Z boson	
LEPTONS		$< 2.2 \text{ eV}/c^2$ 0 $1/2$ ν_e electron neutrino	$< 0.17 \text{ MeV}/c^2$ 0 $1/2$ ν_μ muon neutrino	$< 15.5 \text{ MeV}/c^2$ 0 $1/2$ ν_τ tau neutrino	$80.4 \text{ GeV}/c^2$ ± 1 1 W W boson	GAUGE BOSONS

Figure 1.1: A table of all the elementary particles in the SM in association with some of their properties. The faint gray lines indicate which gauge bosons interact with which fermions [5].

Fermions are characterized by Fermi-Dirac statistics and follow the Pauli exclusion principle, which forbids two fermions from occupying the same quantum state. The fermion group consists of 12 particles that are either labelled as leptons or quarks. Since quarks have electric and color charges, they can contribute to all three main interactions. Quarks have fractional electric charges, with the up-types having a $+2/3$ electric charge, and the down-types having charge of $-1/3$ units of elementary charge. They have only been observed in combined or bound states of hadrons, while leptons can exist as free particles. Colorless leptons can exist as charged or neutral, the charged leptons interact both electromagnetically and weakly while neutral ones known as neutrinos, only contribute to weak interaction.

For each of these fermions, there exists a charge-conjugate partner, called an anti-particle with the same quantum numbers but an opposite electrical charge. Quarks and leptons can be grouped into three generations, where each successive generation is a heavier copy of the previous one. Each generation or family consists of eight particles: two quarks, a lepton and its corresponding neutrino, and the

related antiparticles as shown in Figure 1.1. The fermions masses span many orders of magnitude, so that the first generation quarks (up and down) have relatively low masses (2.3 and 4.8 MeV/c², respectively), while the third generation quarks have orders of magnitude higher in mass (the top and bottom quarks have masses of approximately 173.5 and 4.65 GeV/c², respectively) [5].

The elementary particles can interact with each other depending on their quantum numbers. Within these interactions, particles can be converted into different particles or formed as stable states of matter. These processes are followed by energy-momentum, spin, and charge conservation laws. The lightest particles are stable, whereas heavier particles quickly decay to the next most stable level. Therefore, all visible stable matters surrounding us are made out of particles from the first generation. A proton is formed by two up quarks and one down quark, while a neutron is formed by one up quark and two down quarks. Together with the electron, these three particles can bind into all different kinds of atoms. Particles in second and third generations are highly unstable and decay to particles of lower mass states, hence to study the nature of them, they have to be produced under laboratory conditions.

In addition to the fermions, the SM also contains spin-1 gauge bosons, which are the quanta of the fields responsible for the interactions between particles. The photon (γ) as a massless particle, coupled to electric charged particles, mediates the electromagnetic force. The electromagnetic force is responsible for binding electrons in the atom. The other massless gauge boson of the SM is the gluon which carries a color anti-color charge. Various combinations of three different colors of the SM (red, green and blue), result in eight different gluons. The strong force is responsible for binding quarks into hadrons, as well as binding the protons and the neutrons in the nucleus. The massive $W^\pm$ and Z bosons, field carriers of the weak interaction, constitute other group of gauge bosons. The weak force is responsible for switching flavor between quarks and leptons. An overview of all gauge bosons in the SM with their electric charges and masses is given in Figure 1.1.

In addition to the bosons carrying three fundamental forces within the SM, there exists also one more spin-zero boson: the Brout-Englert-Higgs boson. The Higgs boson is the only fundamental scalar particle in the SM which is responsible for giving masses to all other particles via the mechanism of electroweak symmetry breaking [6–8] (see section 1.3.2). The evidence of new boson was recently observed in 2012 by both of the CMS and the ATLAS experiments at the Large Hadron Collider (LHC) [9, 10].

1.2 Fundamental interactions

The SM of elementary particles contains three fundamental interactions: the strong nuclear, the electromagnetic and the weak nuclear forces. The fourth known elementary force is gravitation, which is described by the theory of General Relativity. Up to now, no combination of all four forces to a general universal theory could be achieved. The interactions between fundamental particles are described by the Quantum Field Theories (QFT) which is combination of the Quantum Mechanics and the Special Relativity.

1.2.1 The Electromagnetic interaction

In the SM, the electromagnetic interaction is described by the Quantum Electrodynamics (QED) theory [11], which is a local invariant gauge theory under $U(1)$ rotations. The theoretical predictions of the QED have been confirmed experimentally up to extremely high precision. All electric charged particles interact electromagnetically via a massless and electrically neutral photon, indicating that the range of the interaction is infinite. The power of the electromagnetic force is inversely proportional to square of distance. In QED theory, the strength of interaction can be represented by a dimensionless parameter, the fine structure constant:

$$\alpha = e^2/\hbar c \cong 1/137, \quad (1.1)$$

where c is the velocity of light and $\hbar$ is the Plank constant.

The electromagnetic and the strong forces are both symmetric under charge conjugation (C), which means that the dynamics of those forces are unchanged by swapping every particle with its anti-particle. Figure 1.2 shows an electron scattering off of another electron by exchanging a photon. In the electron scattering process, two electrons reach to interaction point, exchange a virtual photon, and start moving away from each other.

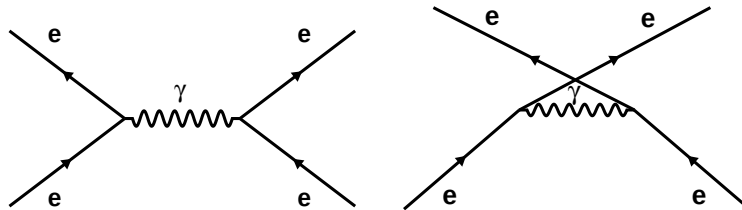


Figure 1.2: Electron-electron scattering Feynman diagram. The solid and wavy lines indicate the electrons and the photon directions, respectively. Time flows upwards.

1.2.2 The Weak interaction and Electroweak unification

The $W^\pm$ and Z bosons as carriers of the weak interaction, couple to all particles with couplings that are proportional to their weak charge. While the $W^\pm$ bosons exist with electric charge, the Z boson is an electrically neutral particle. Their high masses limit the range of the weak force to approximately 10^{-18} m.

The Z boson couples with different strengths to left chiral and right chiral particles while the $W^\pm$ bosons couple only to left chiral particles and right chiral antiparticles. Furthermore, the weak interaction violates parity maximally. This means that it's behaviour is not symmetric under spatial point reflections. The most prominent example for a weak interaction is the β -decay of a neutron transforming into a proton while emitting an electron and an electron anti-neutrino. This happens via the emission of a virtual W^- boson. The Feynman diagram of such decay process is given in Figure 1.3.

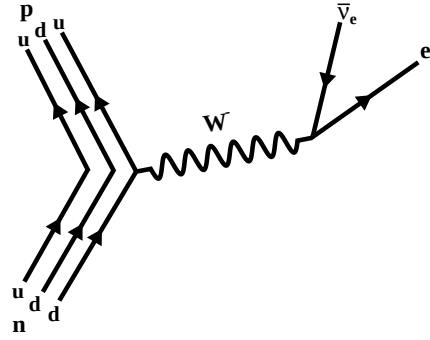


Figure 1.3: β -decay of a neutron to a proton plus electron and related anti-neutrino via weak interaction.

Since weak force eigenstates are not the same as mass eigenstates, then weak decay across generations is possible in the weak interaction. The relation between the two kinds of eigenstate is given by the so called Cabbibo-Kobayashi-Maskawa (CKM) matrix [12, 13]:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad (1.2)$$

In the case of three fermion generations, the CKM matrix is a 3×3 unitary matrix which can be parametrised by three mixing angles and one CP-violating phase. This matrix transforms the mass eigenstates of the down type quarks (d, s, b) into the weak force eigenstates (d', s', b'). The element of this matrix $|V_{ij}|$ represents the probability that a quark type i decays into a quark of j flavour. To check the unitarity of the CKM matrix, many measurements have been performed by different

experiments. The latest results on the CKM matrix elements are given in equation 1.3 [5].

$$V_{CKM} = \begin{pmatrix} 0.97427 \pm 0.00015 & 0.22534 \pm 0.00065 & 0.00351^{+0.00015}_{-0.00014} \\ 0.22520 \pm 0.00065 & 0.97344 \pm 0.00016 & 0.0412^{+0.0011}_{-0.0005} \\ 0.00867^{+0.00029}_{-0.00031} & 0.0404^{+0.0011}_{-0.0005} & 0.999146^{+0.000021}_{-0.000046} \end{pmatrix} \quad (1.3)$$

The nearly diagonal feature of the CKM matrix, implies the preference for flavor-changing processes within one generation. For instance, indirect measurements of $|V_{tb}| \approx 1$ suggest that the top quark will decay predominantly to a bottom quark. Transitions between quarks of different generations are suppressed due to the small magnitude of non-diagonal elements. The quark decay rate to different quark flavors depends upon the interaction strength between two quarks as given by the CKM matrix elements.

Although at first sight, electromagnetic effects and phenomena of the weak interaction seem very different, above certain energy scale, both forces can be explained within the unified electroweak theory. This theory was first introduced by Glashow, Weinberg, and Salam (GWS) [14–16]. The electroweak Lagrangian is required to be invariant under local $SU_L(2) \otimes U_Y(1)$ gauge transformation. The subscript L indicates a coupling only to left handed spinors and Y refers to the weak hypercharge, defined by the Gell-Mann Nishizima formula:

$$Q = T_3 + Y/2, \quad (1.4)$$

where Q and T_3 are the particle charge and third isospin component, respectively. The requirement of invariance under $SU_L(2)$ and $U_Y(1)$ transformations introduces a triplet (referred as $W_\mu^{1,2,3}$) and a single (called B_μ) vector fields. More details on the unification are given in section 1.3.1

1.2.3 The Strong interaction

A similar analogy to the QED and the electroweak interactions lay behind a description of the strong interactions in the SM, known as quantum chromodynamics (QCD) [17] theory. The strong interaction describes the interaction between quarks and gluons. The QCD model is a quantum field theory corresponding to the non-abelian $SU_C(3)$ gauge group, where the C subscript stands for color charge.

QCD introduces internal quantum numbers called color charges as "red", "green" and "blue". Every quark carry a color while every anti-quark carry an anti-color. The strong force carrier, gluons, have one color and one anti-color charges. With three different color charges, there are nine different possible combinations which

constitute a "color octet" or a "color singlet". The singlet is color-neutral, and the color octet states construct gluons, mediating the strong force. Unlike photons in the electromagnetic interaction, gluons have color charge and they not only mediate the interactions between quarks, but also interact strongly with each other. The Feynman diagrams of gluon self interactions are given in the left side of Figure 1.4.

The "running coupling constant" of the strong interaction can be written as:

$$\alpha_s(|q^2|) = \frac{12\pi}{(11n - 2f)\ln(|q^2|/\Lambda_{QCD}^2)} \quad (|q^2| \gg \Lambda_{QCD}^2), \quad (1.5)$$

where q is the momentum of the virtual gluon, $n = 3$ is the number of colors, $f = 6$ is the number of flavors, and Λ_{QCD} as a energy scale constant determined experimentally to be between 100 and 500 MeV. In the right side of Figure 1.4, dependency of the QCD coupling constant on energy scale is shown. The relationship between distance and momentum scale of the QCD interaction is determined by the uncertainty principle, $d = c\Delta t/Q$. The strength decreases at short distances, known as asymptotic freedom, and increases at large distances, which is called "color confinement" and leads to the fact that free quarks and gluons do not exist.

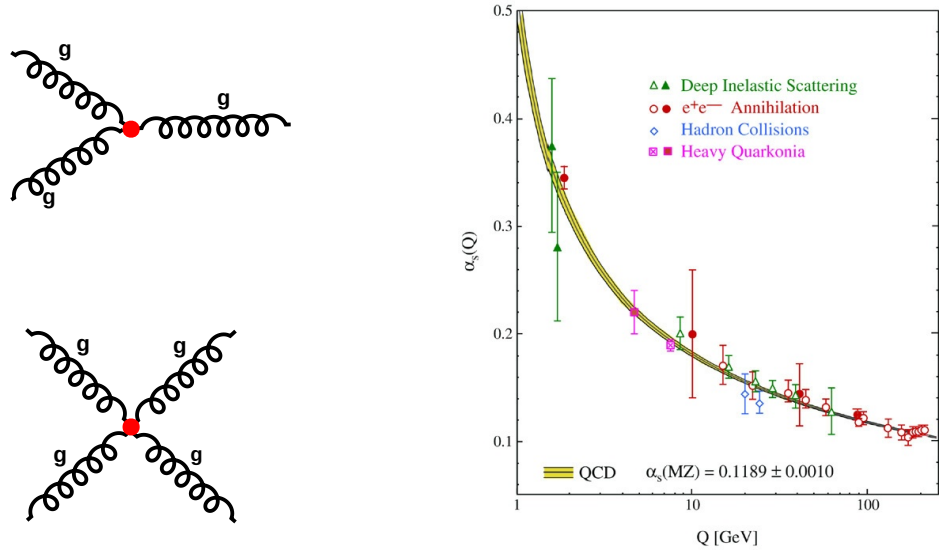


Figure 1.4: Left: Gluon self interactions with three and four vertices. Right: Running the strong coupling constant, α_s in dependence of the energy scale, Q .

1.2.4 The Gravitation interaction

The fourth known elementary force is gravitation, which is described by the theory of the General Relativity. The gravitational force is thought to be mediated by the massless particle with spin two, named graviton. Since the gravitons are massless

and do not themselves carry any charge, they can be exchanged over infinite range. Due to the size of the elementary particle masses, gravity does not play a significant role in comparison to the other interactions and can be safely neglected at the currently achievable energy scales. Apart from the lack of a complete unification of three fundamental forces in the SM, the gravitation force can never included in the SM to make a complete theory explaining all phenomena in the universe.

1.3 Theoretical framework of Standard Model

As mentioned at the beginning of previous section, interactions between fundamental particles are described by the QFT. In this theory, fermions are described by fields which are the solutions of the Euler-Lagrange equations. The interactions between these fields are mediated by the gauge bosons. These interactions are introduced by requiring the theory to be locally invariant under phase transformations. In the SM, these transformations are generated by the gauge groups [1–3]:

$$G_{SM} = SU(3)_C \otimes SU(2)_L \otimes U(1)_Y, \quad (1.6)$$

where the group $SU(3)_C$ introduces the gauge symmetry of the strong force, and the group $SU(2)_L \otimes U(1)_Y$ represents the unified electroweak interaction. The electroweak symmetry is spontaneously broken using the Higgs mechanism which results in one massless and three massive gauge bosons. This mechanism predicts the existence of new spin-0 particle, the Higgs boson. The general form of the Lagrangian density for the SM framework can be written as [1–3]:

$$\mathcal{L}_{SM} = \mathcal{L}_g + \mathcal{L}_f + \mathcal{L}_\phi + \mathcal{L}_{Yukawa}, \quad (1.7)$$

which includes spin-1/2 matter field (the fermions), the spin-1 interaction fields (the gauge bosons), the spin-0 scalar field (the Higgs doublet ϕ), and Yukawa interactions between fermions and Higgs field. In this section, firstly the gauge groups constituting the SM will be explained. Afterwards the higgs mechanism which responsible to give masses to particles will be introduced.

1.3.1 Gauge groups

The interaction between fermions within the SM is derived based on the concept of local gauge invariance. In this framework, the Dirac Lagrangian for a free massive fermion as a Dirac-spinor field ψ is written as:

$$\mathcal{L}_{Dirac} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi, \quad (1.8)$$

In this equation, γ^μ are the Dirac matrices, ψ and $\bar{\psi}$ represent the fermion and anti-fermion Dirac fields and m is the mass of the fermions.

Following the gauge invariance principle, it has to be invariant under a local phase transformation. The transformation is given by:

$$\psi \rightarrow e^{i\vec{\epsilon}(x) \cdot \frac{\vec{\tau}}{2}} \psi, \quad (1.9)$$

where $\vec{\epsilon}(x)$ is space-time dependent rotation vector represented by the generator $\vec{\tau}$.

In the case of rotation which depends on the space-time coordinates, the Dirac Lagrangian described in equation 1.8 is not invariant under transformation defined in equation 1.9. In order to preserve the gauge invariance locally, the normal derivative ∂_μ in equation 1.8 has to be replaced by a so-called covariant derivative D_μ defined as:

$$D_\mu = \partial_\mu - ig \frac{\vec{\tau}}{2} \cdot \vec{A}_\mu, \quad (1.10)$$

where the coupling constant g determines the universal interaction strength, and $\vec{A}_\mu$ is a new interacting vector gauge field which compensates the local gauge transformation if changes as:

$$\vec{A}_\mu \rightarrow \vec{A}_\mu + \frac{1}{g} \partial_\mu \vec{\epsilon}(x). \quad (1.11)$$

By using the covariant derivative, the Dirac Lagrangian is transformed into:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu D_\mu\psi - m\bar{\psi}\psi = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi + g\bar{\psi}\gamma^\mu\frac{\vec{\tau}}{2} \cdot \vec{A}_\mu\psi, \quad (1.12)$$

where the last term in the new Lagrangian illustrates the interaction between fermions and the new vector field. It is found that requiring the theory to be invariant under local gauge transformations, introduces new vector fields called gauge fields in the theory which couple to the fermions.

The gauge transformations can be classified in two classes: Abelian and non-Abelian Lie groups. In the Abelian group, the interaction gauge fields only couple to the fermions, while in the case of the non-Abelian group, the gauge fields interact also with together. The QED and the QCD interactions are samples of the Abelian and the non-Abelian groups, respectively.

Quantum Chromodynamics

The strong interaction is described in terms of the QCD theory. The free Dirac Lagrangian for massless quarks is given by:

$$\mathcal{L} = \sum_f \bar{q}_f i \gamma^\mu \partial_\mu q_f, \quad (1.13)$$

where $q_f^T = (q_f^1, q_f^2, q_f^3)$ is the quark vector with flavor f .

For this interaction, the local phase transformation can be written as:

$$q_f \rightarrow e^{i\epsilon^a(x)\frac{\lambda^a}{2}} q_f, \quad (1.14)$$

where λ^a are the eight generators of the $SU(3)_C$ group, known as Gell-Mann matrices. The Gell-Mann matrices are traceless and satisfy in the following commutation relation:

$$\left[\frac{\lambda^a}{2}, \frac{\lambda^b}{2} \right] = i f^{abc} \frac{\lambda^c}{2}, \quad (1.15)$$

where the structure constants f^{abc} are completely asymmetric. To obtain the local gauge symmetry, covariant derivative and corresponding gauge transformation are defined as:

$$\begin{aligned} D_\mu &= \partial_\mu + i g_s \frac{\lambda^a}{2} G_\mu^a, \\ G_\mu^a &\rightarrow G_\mu^a - \frac{1}{g_s} \partial_\mu \epsilon^a(x) - f^{abc} \epsilon^b(x) G_\mu^c, \end{aligned} \quad (1.16)$$

where g_s is the strong coupling constant and G_μ^a are the eight associated gauge boson fields. The gauge invariant kinematic term responsible for self-interaction gluon fields is:

$$\mathcal{L}_g = -\frac{1}{4} G_a^{\mu\nu} G_{\mu\nu}^a, \quad (1.17)$$

which, $G_a^{\mu\nu} = \partial^\mu G_a^\nu - \partial^\nu G_a^\mu - g_s f^{abc} G_b^\mu G_c^\nu$. Adding the kinematic term of the gauge fields to the massless Dirac Lagrangian, the total QCD Lagrangian is obtained as:

$$\begin{aligned} \mathcal{L}_{QCD} &= \sum_f \bar{q}_f i \gamma^\mu \partial_\mu q_f - g_s G_a^\mu \sum_f \bar{q}_f^\alpha \gamma_\mu \left(\frac{\lambda^a}{2} \right)_{\alpha\beta} q_f^\beta \\ &\quad - \frac{1}{4} (\partial^\mu G_a^\nu - \partial^\nu G_a^\mu) (\partial_\mu G_\nu^a - \partial_\nu G_\mu^a) \\ &\quad + \frac{g_s}{2} f^{abc} (\partial^\mu G_a^\nu - \partial^\nu G_a^\mu) G_\mu^b G_\nu^c - \frac{g_s^2}{4} f^{abc} f_{ade} G_b^\mu G_c^\nu G_\mu^d G_\nu^e. \end{aligned} \quad (1.18)$$

The described Lagrangian density contains interactions between colored massless quarks and gluons as well as gluon self interactions. The quarks mass term of the SM Lagrangian can be added when the Yukawa interaction will be introduced in section 1.3.2.

Electroweak

As mentioned before, the gauge group corresponding to the electroweak is given as the $SU(2)_L \otimes U(1)_Y$. The group consists of the non-Abelian $SU(2)_L$ and the Abelian $U(1)_Y$ symmetry, where the subscript L indicates a coupling only to left handed spinors and Y refers to the defined hypercharge in equation 1.4. The electroweak Lagrangian for free massless fermions is written as:

$$\mathcal{L}_{fermion} = \sum_f [\bar{\psi}_L i\gamma^\mu \partial_\mu \psi_L + \bar{\psi}_R i\gamma^\mu \partial_\mu \psi_R], \quad (1.19)$$

where L and R represent left handed and right handed chirality projections, respectively. Any local phase-space transformation for the electroweak unification can be defined as:

$$\begin{aligned} \psi_L &\rightarrow e^{i(\vec{\alpha}(x) \cdot \frac{\vec{\tau}}{2} + Y_L \beta(x))} \psi_L, \\ \psi_R &\rightarrow e^{iY_R \beta(x)} \psi_R, \end{aligned} \quad (1.20)$$

where $\vec{\tau}$ represents the Pauli matrices. The electric charge is related to hypercharge and the third component of the isospin via equation 1.4. To make the Lagrangian invariant, following covariant derivatives are defined:

$$\begin{aligned} D_\mu \psi_L(x) &= (\partial_\mu - i\frac{g}{2}\vec{\tau} \cdot \vec{W}_\mu - ig'Y_L B_\mu) \psi_L(x), \\ D_\mu \psi_R(x) &= (\partial_\mu - ig'Y_R B_\mu) \psi_R(x). \end{aligned} \quad (1.21)$$

Here g and g' are the corresponding coupling constants for the $SU(2)_L$ and the $U(1)_Y$ gauge groups, respectively. Furthermore, three gauge fields W_μ^a and a field B_μ associated to the defined gauge groups are introduced with the following transformations:

$$\begin{aligned} B_\mu(x) &\rightarrow B_\mu(x) + \frac{1}{g'} \partial_\mu \beta(x), \\ \vec{\tau} \cdot \vec{W}_\mu(x) &\rightarrow U(x) \vec{\tau} \cdot \vec{W}_\mu(x) U^\dagger(x) + \frac{2i}{g} U(x) \partial_\mu U^\dagger(x), \end{aligned} \quad (1.22)$$

where $U(x) = e^{i\vec{\alpha}(x) \cdot \frac{\vec{\tau}}{2}}$. Similarly to the QCD theory, the gauge invariant kinematic terms for the gauge bosons in the electroweak Lagrangian are given by:

$$\mathcal{L}_{gauge} = -\frac{1}{4} B^{\mu\nu} B_{\mu\nu} - \frac{1}{4} \vec{W}^{\mu\nu} \cdot \vec{W}_{\mu\nu}, \quad (1.23)$$

where

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu, \quad W_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a + gf^{abc} W_\mu^b W_\nu^c. \quad (1.24)$$

Here, the structure constants f^{abc} are obtained from commutation relation between Pauli matrices. Finally the electroweak gauge invariant Lagrangian is written as:

$$\mathcal{L} = \sum_f [\bar{\psi}_L i \gamma^\mu D_\mu \psi_L + \bar{\psi}_R i \gamma^\mu D_\mu \psi_R] - \frac{1}{4} B^{\mu\nu} B_{\mu\nu} - \frac{1}{4} \vec{W}^{\mu\nu} \cdot \vec{W}_{\mu\nu}. \quad (1.25)$$

The four described gauge fields can be expressed by the linear combinations of the charged $W_\mu^\pm$ bosons, the neutral Z_μ boson, and the massless photon A_μ fields. These observed gauge fields are defined as:

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}} (W_\mu^1 \mp W_\mu^2), \\ \begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} &= \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix}, \end{aligned} \quad (1.26)$$

where in the above equation, θ_W is known as the Weinberg angle and defined as:

$$\tan \theta = \frac{g'}{g}. \quad (1.27)$$

By substituting the observed gauge bosons in equation 1.25, two charged and neutral current sectors for the electroweak interaction are obtained. The charged sector for three families of quarks and leptons is written as:

$$\mathcal{L}_{CC} = \frac{g}{2\sqrt{2}} W_\mu^\dagger \left(\sum_i \bar{u}_i \gamma^\mu (1 - \gamma_5) d_i + \bar{\nu}_i \gamma^\mu (1 - \gamma_5) l_i \right) + h.c. \quad (1.28)$$

Similarly, the neutral current sector of the electroweak interaction is given by:

$$\begin{aligned} \mathcal{L}_{NC} &= \sum_i \bar{\psi}_i \gamma^\mu \left(\frac{g}{2} \tau_3 \sin \theta_W + g' y_i \cos \theta_W \right) A_\mu \psi_i \\ &+ \sum_i \bar{\psi}_i \gamma^\mu \left(\frac{g}{2} \tau_3 \cos \theta_W - g' y_i \sin \theta_W \right) Z_\mu \psi_i + h.c. \end{aligned} \quad (1.29)$$

By adding mass terms for the gauge bosons, the electroweak Lagrangian would break the gauge invariance. Also the mass terms of fermions must be added to the SM Lagrangian. Both of these issues can be solved by introducing spontaneous symmetry breaking which is explained in the next section.

1.3.2 Higgs mechanism

The described electroweak interaction Lagrangian in the previous section contains only massless gauge bosons and fermions. In order to give masses to gauge bosons and fermions while preserving the gauge invariance, the electroweak symmetry has to be broken spontaneously. The most promising candidate mechanism providing such

a breakdown of the electroweak symmetry is the Brout-Englert-Higgs mechanism, in which generates related masses through spontaneous symmetry breaking. Symmetry breaking happens, if the Lagrangian has symmetries which do not validate for the vacuum state of the system.

In the Higgs mechanism to break the electroweak gauge symmetry, a complex isospin doublet scalar field ϕ , the Higgs field, is introduced:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}, \quad (1.30)$$

where ϕ^+ and ϕ^0 are positively charged and neutral complex scalar fields. While introducing the doublet scalar field can break the symmetry of the vacuum expectation state, the electroweak Lagrangian leaves invariant. The gauge invariant Lagrangian term for the new scalar field can then be added to the SM Lagrangian:

$$\begin{aligned} \mathcal{L}_\phi &= (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi), \\ V(\phi) &= \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2. \end{aligned} \quad (1.31)$$

Here D_μ , μ^2 and λ are the covariant derivative in the electroweak theory, a mass parameter and the strength of the Higgs boson field self-interaction, respectively. Two μ and λ parameters, determine the shape of the Higgs potential. The Higgs potential function is minimized for the case of $\mu^2 < 0$ and $\lambda > 0$. With this selection, there exists a non-vanishing vacuum expectation value which minimise the potential:

$$\phi \phi^\dagger \equiv \frac{v^2}{2} = -\frac{\mu^2}{2\lambda}. \quad (1.32)$$

The minimum value of the Higgs potential is obtained for infinite number of degenerate states shown as a continuous ring in Figure 1.5.

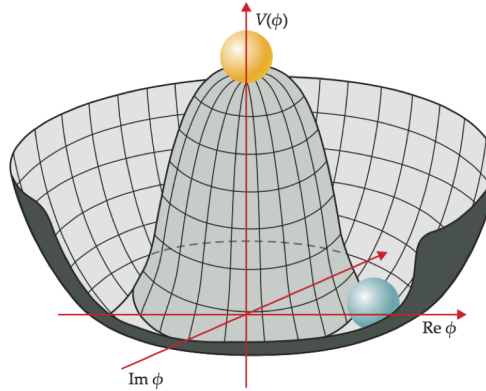


Figure 1.5: Graph of Higgs potential in the complex plane.

It is required that the Higgs Lagrangian is invariant under the reflection sym-

metry, ϕ to $-\phi$. The choice of particular minimum state does not affect on the physics results while by choosing an arbitrary vacuum state, the electroweak symmetry is spontaneously broken in a way that the electromagnetic symmetry remains as a symmetry of the vacuum. Therefore the photon remains massless while the gauge bosons in the weak interaction become massive. To write down the Higgs Lagrangian, it is convenient to work in the unitary gauge, where the only remained degree of freedom in the equation 1.30 is the Higgs field, ϕ_3 . By fixing the vacuum state, the defined scalar field ϕ can be written as a quantum fluctuation around minimum v in the most general form:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}, \quad (1.33)$$

where $H(x)$ denotes the Higgs boson field. By rewriting the Higgs sector, the scalar potential becomes:

$$V(\phi) = \frac{1}{2}(2\lambda v^2)H^2(x) + \lambda v H^3(x) + \frac{\lambda}{4}H^4(x). \quad (1.34)$$

Therefore, the mass of the Higgs boson equals to $m_H = \sqrt{2\lambda}v$. The kinematic term using defined electroweak covariant derivative is also changed to:

$$\frac{1}{2}\partial_\mu H(x)\partial^\mu H(x) + (v + H(x))^2 \left(\frac{g^2}{8} W_\mu^+ W^{+\mu} + \frac{g^2}{8} W_\mu^- W^{-\mu} + \frac{g^2}{8 \cos^2 \theta_W} Z_\mu Z^\mu \right), \quad (1.35)$$

which introduces mass terms for the $W_\mu^\pm$ and the Z_μ gauge bosons by fixing the vacuum expectation value. The masses of these gauge bosons are:

$$M_W = M_Z \cos \theta_W = \frac{gv}{2}. \quad (1.36)$$

In contrast to the mass of the gauge bosons, the fermion masses are not generated by a gauge symmetry breaking. The mass terms of fermions are generated by the Yukawa coupling of the fermions to the Higgs boson. The coupling strength of interaction as a free parameter of the theory is proportional to the fermion mass. The gauge invariant Yukawa terms to be added to the SM Lagrangian for the first generation of leptons and quarks are defined as:

$$\mathcal{L}_{Yukawa} = -\left(\frac{v + H(x)}{\sqrt{2}}\right)(g_\nu \bar{\nu}_e \nu_e + g_e \bar{e} e + g_u \bar{u} u + g_d \bar{d} d). \quad (1.37)$$

where g_f are the Yukawa coupling constants and the fermion masses are included as $M_f = g_f v / \sqrt{2}$. The introduction of the spontaneous symmetry breaking and the Higgs mechanism, allows to have masses without violating any gauge symmetries.

Although the Higgs mechanism has been formulated in the early 1960's, the discovery of the Higgs particle at the LHC was announced in July 2012 simultaneously by the two experiments CMS and ATLAS. This new particle is compatible with the SM Brout-Englert-Higgs boson with the mass of $125.9 \pm 0.4 \text{ GeV}$ [9, 10].

1.4 Top-quark Physics

The current understanding of the top quark physics such as top quark production, decay, and properties are discussed in this section. The existence of a third generation of quarks was predicted in 1973 to explain CP violation in the Kaon decay [13]. The third generation up-type quark of the SM, the top quark, with spin 1/2 and electric charge of $+2/3e$ is the most massive known elementary particle. In 1995, the CDF [18] and the DØ [19] experiments at the Tevatron collider have discovered the top quark via direct top quark pairs production in 1.8 TeV proton-antiproton collisions. This fundamental particle has a mass of $m_{\text{top}} = 173.34 \pm 0.27(\text{stat}) \pm 0.71(\text{syst}) \text{ GeV}$ [20].

The top quark is the only quark decaying via the weak interaction before forming a bound state as a hadron. Therefore, there is a unique opportunity to access the top quark as a bare one, and study precise measurements on its properties. Measurements on the top quark properties can be open a window to test the validity of beyond the SM. In the SM, the top quark interacts through the strong and the electroweak interactions. Due to its large mass, the top quark plays a special role in the electroweak symmetry breaking.

After discovery of the top quark, some different aspects of that have been studied in details. Since 2010, the CMS and the ATLAS experiments have made precise measurements on the top quark properties. Top quark production processes with focus on the pair production will be discussed in the next section. Decay mode of the top quark pairs with different final states are detailed in section 1.4.2.

1.4.1 Top quark production

Since the top quark discovery, it has been produced by the hadron colliders such as the Tevatron and the LHC. There are two main processes for the top quark production, top quark pair production via the strong interaction and single top quark production via the electroweak interaction. To produce a real top quark pair, the energy of system must be at least twice the mass of the top quark. At the LHC with the center-of-mass energy of 8 TeV, the parton energy fractions correspond to $x_1 x_2 \geq 0.04^2$. Because of large value of gluon density at the scale of x , the gluon-gluon fusion is the dominant mechanism for the top quark pair production. Given

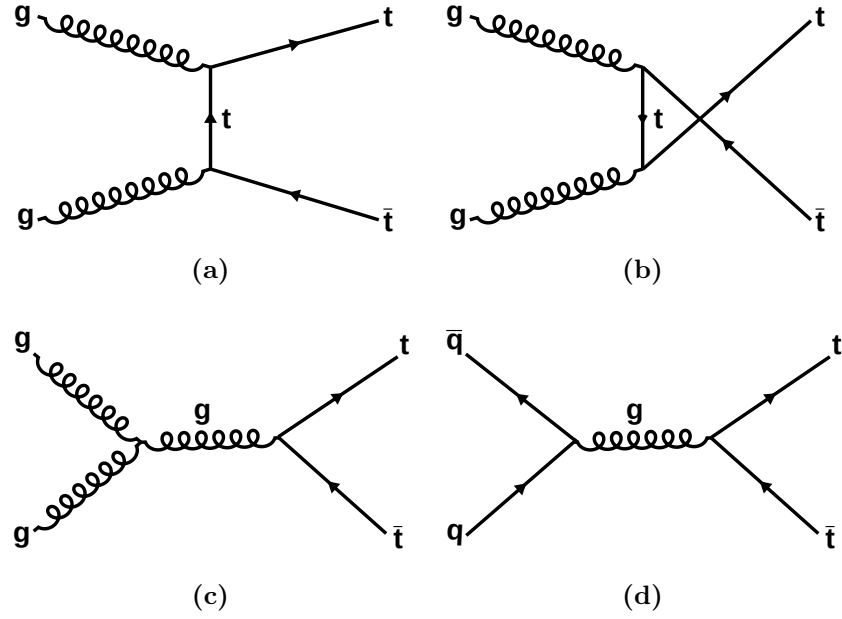


Figure 1.6: Leading-order Feynman diagrams for top-quark pair production in p-p collisions via gluon-gluon fusion (a-c) and quark-antiquark annihilation (d).

the high energy ranges available at the LHC, the gluon-gluon fusion is expected to account about 90% of the $t\bar{t}$ events.

The $t\bar{t}$ production is also possible through photon or Z boson exchange but they are negligible in comparison with the other processes. The leading-order Feynman diagrams for the top quark pairs production via gluon-gluon fusion and quark-antiquark annihilation are shown in Figure 1.6.

As discussed before, the top quarks can also be produced as single top quarks. The single top quark production has been first observed by the CDF [21] and the DØ [22] collaborations. The Feynman diagrams for the three different configurations of the single top quark production: t-channel, s-channel and associated production with on-shell W boson are given in Figure 1.7.

At a high energy hadron collider, collisions can be classified into soft and hard interactions. At sufficiently high energies, proton-proton collisions can be regarded as interactions between free proton constituents, called partons. In this limit, these hard interactions are described by the perturbation theory [23, 24]. The total $t\bar{t}$ cross section, $\sigma(pp \rightarrow t\bar{t})$, is approximated by summing over all possible partonic interactions for given subprocess cross sections, $\hat{\sigma}(ij \rightarrow t\bar{t})$, with related PDFs:

$$\sigma^{pp \rightarrow t\bar{t}}(p_1 p_2) = \sum_{i,j} \int dx_i dx_j f_i(x_i, \mu_F^2) f_j(x_j, \mu_F^2) \hat{\sigma}_{i,j}(p_1, p_2, \alpha_s(\mu_R^2)), \quad (1.38)$$

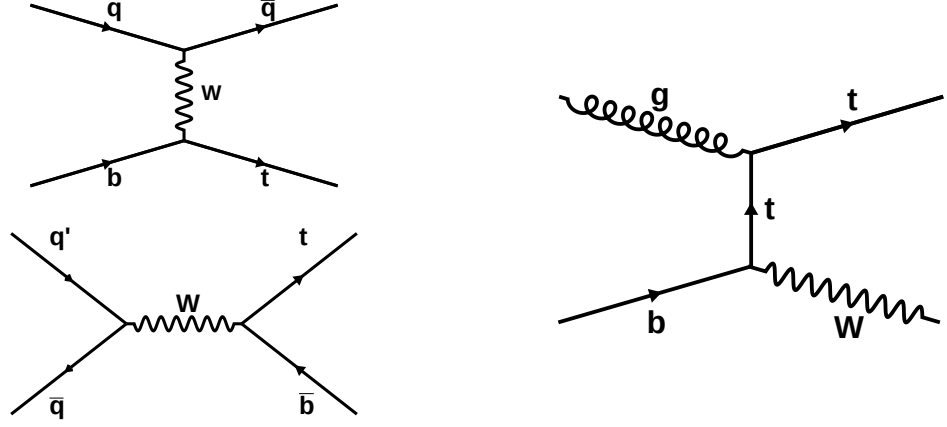


Figure 1.7: Single top quark production Feynman diagrams: t-channel (top left), s-channel (bottom left) and associate production with on-shell W boson (right).

where i and j are the possible combinations of incoming gluons or quark anti-quark pairs, and integration over the parton momenta $x_i = p_i/p_1$ and $x_j = p_j/p_2$. The probability for a parton to carry momentum fraction x_i with factorization scale μ_F is described by the parton density function (PDF), f_i , which is determined empirically. To avoid the ultraviolet divergences appearing in loop diagrams, the theory has to be renormalised at a certain energy scale. The factorization and renormalization scales, μ_F and μ_R , are an arbitrary parameters which often set to the top quark mass.

The theoretical calculations of the $t\bar{t}$ cross section with the highest accuracy at the next-to-next leading-order (NNLO) are described in [25]. Figure 1.8 summarizes recent $t\bar{t}$ cross section measurements from the CMS and the ATLAS collaborations as a function of the center-of-mass energy, compared to the latest theoretical predictions. The comparison with the predicted values shows that the measurements are in good agreement with the SM expectation.

1.4.2 Top quark decay modes

As discussed before, top quarks can decay to down-type quarks via the weak interaction before hadronization. Since the CKM matrix element $|V_{tb}|$ is close to one, the top quark decays almost exclusively into a W boson and a b-quark. Due to the small CKM matrix element values for other down-type quarks, the related decay processes are suppressed.

It is convenient to classify final states of the $t\bar{t}$ events based on the decay products of the W bosons. Figure 1.9 shows the Feynman diagrams for the top quark decaying into a b quark and a W^+ boson, and subsequently decaying the W boson

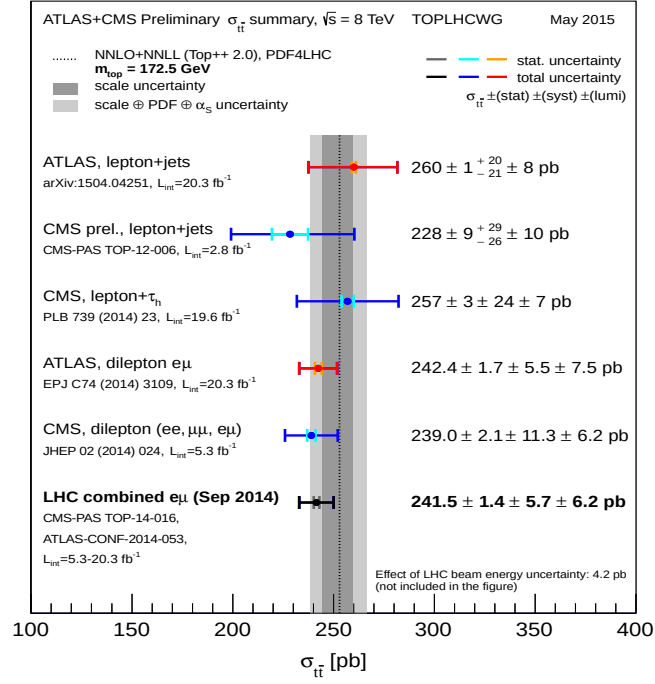


Figure 1.8: The summary of the top pair production cross section measurements for the CMS and ATLAS experiments at 8 TeV, compared to the corresponding theoretical expectation.

either into lepton-neutrino or quark-antiquark pairs.

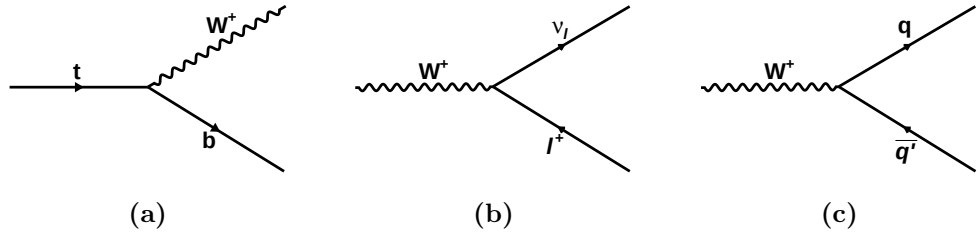


Figure 1.9: Decays of top quark into a b quark and a W^+ boson (a). The W^+ boson decays into a charged lepton and the corresponding neutrino (b) or into a quark-antiquark pair (c).

Considering the mass of the W boson and neglecting flavour mixing, there is two hadronic and three leptonic possibilities for the final state. Depending on the decay modes of the W bosons, decaying of the top quark pairs can then be classified into three different channels:

- The all-hadronic final state, in which both W bosons decay hadronically. This decay mode has the largest branching ratio, 45.5% of the total, but because of the QCD multi-jet processes, it suffers from a large background. Therefore

making it difficult to distinguish signal events from the SM QCD multi-jet backgrounds.

- The lepton+jets final state, in which one W boson decays leptonically and the other one decays hadronically. This channel has a relatively large branching ratio, about 43.8% of the total $t\bar{t}$ events. The lepton+jets channel suffers from a larger background rate than that of the dilepton channel but it has a comparative small background contamination relative to all-hadronic final state.
- The dileptonic final state, in which both W bosons decay leptonically. Conventionally, the presence of the electrons or the muons at the dilepton final states, e^+e^- , $\mu^+\mu^-$ and $e^\pm\mu^\mp$ are known as di-electron, di-muon and electron-muon channels, respectively. The dileptonic decay mode has the smallest branching ratio among all other decay channels, nearly 10.5%. Due to presence two high transverse momentum leptons and large missing transverse energy at the final state, the dilepton channel has the advantage of the cleanest signature. Since there are two undetectable neutrinos which contribute to the missing transverse energy, the reconstruction of the dileptonic channel is more complex than the other described decay modes.

The relative contributions of these different $t\bar{t}$ decay modes are shown in Figure 1.10. The analysis described in this thesis was carried out in the di-electron channel so that only electrons, including those from τ decay, are taken into account as signal events. The signature of the $t\bar{t}$ di-electron events in the detector is characterized by two highest p_T oppositely charged electrons, large missing transverse energy E_T^{miss} from two undetected neutrinos and two jets coming from b-quarks.

1.5 W boson polarization in top-quark decay

As described before, the W boson is the mediator of the charged-current weak interaction. The vertex factor for this part of the weak interaction has a following form:

$$i\frac{g}{\sqrt{2}}\gamma^\mu\frac{1}{2}(1-\gamma^5). \quad (1.39)$$

This factor is known as a vector ($\bar{\psi}\gamma^\mu\psi$) minus axial-vector ($\bar{\psi}\gamma^\mu\gamma^5\psi$) which is often called V-A structure. The term $\frac{1}{2}(1-\gamma^5)$ is introduced as the left handed projection operator, when operates on four-components spinors, remains only left handed part of them. This means that the W boson in the weak charged-current vertex, only couples to left handed particles or right handed antiparticles.

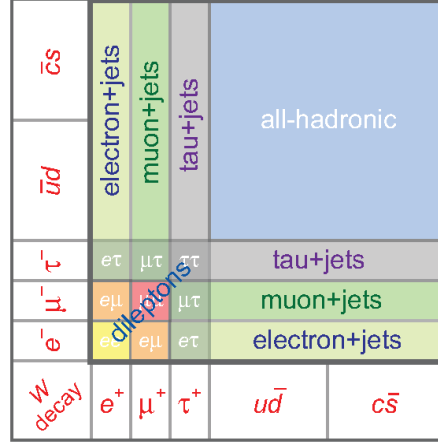


Figure 1.10: The relative contributions of different $t\bar{t}$ decay modes. The sizes of the different areas reflect the branching ratios of the different decay channels.

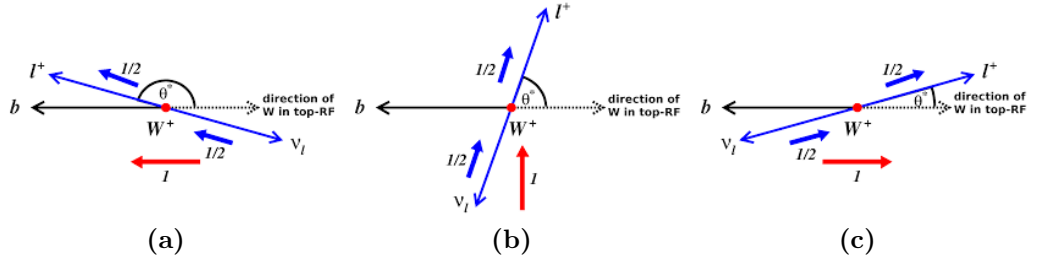


Figure 1.11: Different helicity configurations of (a) a left handed, (b) a longitudinally and (c) a right handed polarized W boson into a charged lepton and a neutrino in the W boson rest frame. The direction of motion of the W boson in the top-quark rest frame mentioned with the dotted black arrow in this figure.

The W boson polarization measurement in the top quark decay is important from the point that it could reveal new physics beyond the SM. We can write down the mathematical representation of the top quark decay similar to the described W boson decay vertex form. Neglecting the b -quark mass in comparison to the top quark and the W boson masses, the b -quark can only be left handed in the tWb vertex. The helicity configurations of the W boson decay in the top quark rest frame are shown in Figure 1.11.

The direction of z -axis is adjusted so that the spin- $\frac{1}{2}$ of the top quark is aligned in the positive side. The spin projections on the direction of motion (helicity) for the massive W bosons can be $-1, 0, +1$, called as left handed, longitudinal and right handed helicity states, respectively. Since the W boson only couples to the left handed b -quark, angular momentum conservation suppresses presence of the right handed W boson helicity state.

SM gives specific prediction about helicity fractions of the W boson, where a measurement of non-zero value for the right handed helicity fraction would be evidence of new physics beyond the SM. The goal of this analysis is to make precise measurements on the W boson helicity fractions in order to determine if there is evidence of new physics in the tWb vertex.

The corresponding helicity fractions are defined as a ratio of the relative top quark decay width for specific helicity state to the total width:

$$F_{L,0,R} = \frac{\Gamma_{L,0,R}(t \rightarrow W^+b)}{\Gamma_L(t \rightarrow W_L^+b) + \Gamma_0(t \rightarrow W_0^+b) + \Gamma_R(t \rightarrow W_R^+b)}. \quad (1.40)$$

To determine quantitative expression of the top quark decay width to the various W boson helicity states, following calculations are done. The Lorentz invariant amplitude of the top quark decay can be written in terms of elements shown in Figure 1.12 as:

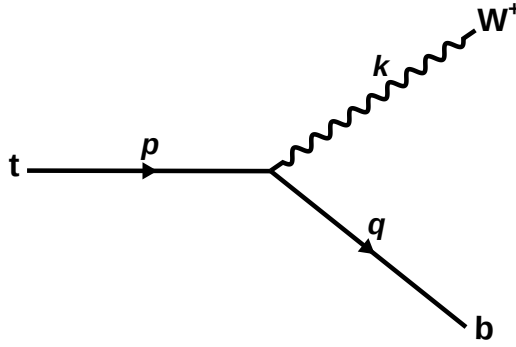


Figure 1.12: Feynman diagram for the top quark decay $t \rightarrow W^+b$.

$$i\mathcal{M} = \frac{ig}{\sqrt{2}} \bar{b}(q) \gamma^\mu \left(\frac{1 - \gamma^5}{2} \right) t(p) \epsilon_\mu^*(k). \quad (1.41)$$

After squaring the amplitude and averaging over the top, bottom and W boson polarization states, we find:

$$\begin{aligned} \langle |\mathcal{M}|^2 \rangle &= \frac{1}{2} \frac{1}{3} \sum_{\text{quark spins}} \sum_{W \text{ polarizations}} |\mathcal{M}|^2 \\ &= g^2 [p^\mu q^\nu + p^\nu q^\mu - g^{\mu\nu} p \cdot q - i \epsilon^{\mu\nu\alpha\beta} p_\alpha q_\beta] \times \frac{1}{3} \sum_{W \text{ polarizations}} \epsilon_\mu^*(k) \epsilon_\nu(k). \end{aligned} \quad (1.42)$$

The details to extract equation 1.42 can be found in appendix A. The W boson polarization vector with different helicity configurations and the four-momentum vector of particles in the top quark rest frame can be written as:

$$\begin{aligned}
\epsilon_L^\mu &= (0, \frac{1}{\sqrt{2}}, -\frac{i}{\sqrt{2}}, 0), \\
\epsilon_0^\mu &= (\frac{p}{m_W}, 0, 0, -\frac{E_W}{m_W}), \\
\epsilon_R^\mu &= (0, -\frac{1}{\sqrt{2}}, -\frac{i}{\sqrt{2}}, 0).
\end{aligned} \tag{1.43}$$

$$\begin{aligned}
p^\mu &= (m_t, 0, 0, 0), \\
q^\mu &= (E_b, 0, 0, p_b), \\
k^\mu &= (E_W, 0, 0, -p_W),
\end{aligned} \tag{1.44}$$

where in this case, $p_W = p_b \equiv p$. Inserting these expressions into equation 1.41, we find the contribution of different helicity configurations separately as:

$$\begin{aligned}
\text{Left handed W boson: } \quad \langle |\mathcal{M}|^2 \rangle_L &= \frac{g^2}{3} m_t (E_b + p), \\
\text{Longitudinal W boson: } \quad \langle |\mathcal{M}|^2 \rangle_0 &= \frac{g^2}{3} (2m_t^2 \frac{p^2}{m_W^2} + m_t E_b), \\
\text{Right handed W boson: } \quad \langle |\mathcal{M}|^2 \rangle_R &= \frac{g^2}{3} m_t (E_b - p).
\end{aligned} \tag{1.45}$$

Since all three polarization states have the same phase space factor, equation 1.40 can be written as the ratio of the relative spin-averaged square of the matrix element to the total one:

$$F_{L,0,R} = \frac{\langle |\mathcal{M}|^2 \rangle_{L,0,R}}{\langle |\mathcal{M}|^2 \rangle_L + \langle |\mathcal{M}|^2 \rangle_0 + \langle |\mathcal{M}|^2 \rangle_R}. \tag{1.46}$$

By substituting equations 1.45 into above expression, the W boson helicity fractions are extracted as:

$$\begin{aligned}
F_L &= \frac{1 + \frac{p}{E_b}}{1 + 2(\frac{E_W}{m_W})^2(1 + \frac{p^2}{E_b E_W})}, \\
F_0 &= \frac{1 + 2(\frac{m_t}{m_W})^2 \frac{p^2}{m_t E_b}}{1 + 2(\frac{E_W}{m_W})^2(1 + \frac{p^2}{E_b E_W})}, \\
F_R &= \frac{1 - \frac{p}{E_b}}{1 + 2(\frac{E_W}{m_W})^2(1 + \frac{p^2}{E_b E_W})}.
\end{aligned} \tag{1.47}$$

In the rest frame of the top quark, the quantities of E_b and E_W can be expressed in terms of the top quark, the W boson and the b-quark masses as:

$$\begin{aligned}
E_b &= \frac{m_t^2 + m_b^2 - m_W^2}{2m_t}, \\
E_W &= \frac{m_t^2 + m_W^2 - m_b^2}{2m_t}.
\end{aligned} \tag{1.48}$$

Finally, by using above formula and massless b-quark assumption, we can write the W boson helicity fractions at the tree level of the SM as:

$$\begin{aligned} F_L &= \frac{2x^2}{1+2x^2}, \\ F_0 &= \frac{1}{1+2x^2}, \\ F_R &= 0, \end{aligned} \tag{1.49}$$

where x is defined as $x = m_W/m_t$. It can be concluded that in the limit of a vanishing b-quark mass, right handed b-quarks are forbidden at the leading order. Taking into account also the b-quark mass, the helicity fractions can be expressed in leading order as [26]:

$$\begin{aligned} F_L &= \frac{x^2(1-x^2+y^2+\sqrt{\lambda})}{(1-y^2)^2+x^2(1-2x^2+y^2)} = \frac{2x^2}{1+2x^2} + \dots, \\ F_0 &= \frac{(1-y^2)^2-x^2(1+y^2)}{(1-y^2)^2+x^2(1-2x^2+y^2)} = \frac{1}{1+2x^2} + \dots, \\ F_R &= \frac{x^2(1-x^2+y^2-\sqrt{\lambda})}{(1-y^2)^2+x^2(1-2x^2+y^2)} = y^2 \frac{2x^2}{(1-x^2)^2(1+2x^2)} + \dots, \end{aligned} \tag{1.50}$$

where y and λ are defined as $y = m_b/m_t$ and $\lambda = 1+x^4+y^4-2x^2y^2-2x^2-2y^2$. The right handed component of the W boson is suppressed by the factor of m_b^2/m_t^2 . Using the current measurements of $m_t = 173.34(0.76) \text{ GeV}/c^2$ [20], $m_W = 80.401(43) \text{ GeV}/c^2$ [5] and $m_b = 4.65(3) \text{ GeV}/c^2$ [5], the helicity fractions in the leading order equal to: $F_L = 0.301$, $F_0 = 0.699$ and $F_R = 0.0004$.

In order to draw any conclusions on new physics models, these fractions have to be calculated at higher order with large precision. By including the NNLO QCD and electroweak corrections, the W boson helicity fractions are changed slightly to [27]:

$$\begin{aligned} F_L &= 0.311(5), \\ F_0 &= 0.687(5), \\ F_R &= 0.0017(1). \end{aligned} \tag{1.51}$$

1.5.1 Anomalous couplings in tWb vertex

At higher energies, new physics beyond the SM can contribute to the tWb vertex by including radiative corrections [28]. Above the electroweak symmetry breaking scale, these corrections can be implemented into the Lagrangian in terms of effective operators [29]. The effective Lagrangian describing physics processes at the high

energy scales is written as:

$$\mathcal{L}_{eff} = \mathcal{L}_{SM} + \sum \frac{C_x^{(6)}}{\Lambda^2} O_x^{(6)} + \dots, \quad (1.52)$$

where $O_x^{(6)}$ are the operators of dimension six are constrained to be invariant under the SM gauge transformations. The dimensionless coefficients constant, $C_x^{(6)}$, represent the impact of operators in the effective Lagrangian. At very high energy scale Λ , high-order operators are suppressed and they are not considered here. The gauge invariant operator of dimension five contributes at very high energy scales ($\Lambda = 10^{15}$ GeV), responsible for generating Majorana masses for neutrinos [30], not relevant in this study.

The most general Lagrangian for the tWb interaction containing terms up to dimension six can be expressed as:

$$\mathcal{L}_{eff} = -\frac{g}{\sqrt{2}} \bar{b} \gamma^\mu (V_L P_L + V_R P_R) t W_\mu^- - \frac{g}{\sqrt{2}} \bar{b} \frac{i\sigma^{\mu\nu} q_\nu}{M_W} (g_L P_L + g_R P_R) t W_\mu^- + h.c., \quad (1.53)$$

where $q_\nu = p_\nu^t - p_\nu^b$, $P_L = \frac{1}{2}(1 - \gamma^5)$, $P_R = \frac{1}{2}(1 + \gamma^5)$ and $i\sigma^{\mu\nu} = \frac{1}{2}[\gamma^\mu, \gamma^\nu]$. The Lagrangian contains two vector couplings V_L and V_R as well as two tensor couplings g_L and g_R , representing decays from a top quark to a bottom quark with different chiralities. In the SM, the V_L parameter is the CKM matrix element $V_{tb} \simeq 1$ whereas the other couplings are vanished at tree level. In case if there is any new physics, a deviation of effective Lagrangian from the SM values will be seen. The four couplings can be measured indirectly from experimental measurements of radiative B meson decays ($b \rightarrow s\gamma$) and $B\bar{B}$ mixing [31].

In order to extract the W boson helicity fractions, similar method discussed in section 1.5 has been applied and the following values are obtained. The details of partial widths calculations for the effective tWb vertex can be found in [32].

$$\begin{aligned}
\Gamma_0 &= \frac{g^2|q|}{32\pi} \left(\frac{m_t^2}{M_W^2} [|V_L|^2 + |V_R|^2] (1 - x_W^2 - 2x_b^2 - x_W^2 x_b^2 + x_b^4) - 4x_b \text{Re}(V_L V_R^*) \right. \\
&\quad + [|g_L|^2 + |g_R|^2] (1 - x_W^2 + x_b^2) - 2 \frac{m_t}{M_W} \text{Re}[V_L g_R^* + V_R g_L^*] (1 - x_W^2 - x_b^2) \\
&\quad \left. - 4x_b \text{Re}(g_L g_R^*) + 2 \frac{m_t}{M_W} x_b \text{Re}[V_L g_L^* + V_R g_R^*] (1 + x_W^2 - x_b^2) \right), \\
\Gamma_{R,L} &= \frac{g^2|q|}{32\pi} \left(\frac{m_t^2}{M_W^2} [|g_L|^2 + |g_R|^2] (1 - x_W^2 - 2x_b^2 - x_W^2 x_b^2 + x_b^4) - 4x_b \text{Re}(g_L g_R^*) \right. \\
&\quad + [|V_L|^2 + |V_R|^2] (1 - x_W^2 + x_b^2) - 2 \frac{m_t}{M_W} \text{Re}[V_L g_R^* + V_R g_L^*] (1 - x_W^2 - x_b^2) \\
&\quad \left. - 4x_b \text{Re}(V_L V_R^*) + 2 \frac{m_t}{M_W} x_b \text{Re}[V_L g_L^* + V_R g_R^*] (1 + x_W^2 - x_b^2) \right) \\
&\quad \pm \frac{g^2}{64\pi} \frac{m_t^3}{M_W^2} (-x_W^2 [|V_L|^2 - |V_R|^2] + (1 - x_b^2) [|g_L|^2 - |g_R|^2] \\
&\quad + 2x_W \text{Re}[V_L g_R^* - V_R g_L^*] + 2x_W x_b \text{Re}[V_L g_L^* - V_R g_R^*]) \\
&\quad \times (1 - 2x_W^2 - 2x_b^2 + x_W^4 - 2x_W^2 x_b^2 + x_b^4),
\end{aligned} \tag{1.54}$$

where $x_W = m_W/m_t$, $x_b = m_b/m_t$ and,

$$|\vec{q}| = \frac{1}{2m_t} (m_t^4 + M_W^4 + m_b^4 - 2m_t^2 M_W^2 - 2m_t^2 m_b^2 - 2m_b^2 M_W^2)^{1/2}. \tag{1.55}$$

The total top width is calculated by summing over partial widths:

$$\begin{aligned}
\Gamma &= \frac{g^2|q|}{32\pi} \frac{m_t^2}{M_W^2} ([|V_L|^2 + |V_R|^2] (1 + x_W^2 - 2x_b^2 - 2x_W^4 + x_W^2 x_b^2 + x_b^4) - 12x_W^2 x_b \text{Re}(V_L V_R^*) \\
&\quad + 2[|g_L|^2 + |g_R|^2] (1 - \frac{x_W^2}{2} - 2x_b^2 - \frac{x_W^4}{2} - \frac{x_W^2 x_b^2}{2} + x_b^4) - 12x_W^2 x_b \text{Re}(g_L g_R^*) \\
&\quad - 6x_W \text{Re}[V_L g_R^* + V_R g_L^*] (1 - x_W^2 - x_b^2) + 6x_W x_b \text{Re}[V_L g_L^* + V_R g_R^*] (1 + x_W^2 - x_b^2)).
\end{aligned} \tag{1.56}$$

The resulting helicity fractions of the W boson can then be calculated using equation 1.40. By substituting the values of g_L , g_R and V_R to zero and at the limit where $m_b \rightarrow 0$, the SM results shown in previous section are obtained exactly. In order to search new physics beyond the SM in a model-independent way, several measurements on the W boson helicity fractions were performed which described in the next section.

1.5.2 Measurement of W boson polarization

The W boson polarization can be measured experimentally by studying the angular distribution of the W boson decay products in its rest frame. Considering the decay mode of the top quark, $t \rightarrow W^+ b \rightarrow l^+ \nu_l b$, polar angle θ_l^* is defined as the angle between the charged lepton three-momentum in the rest frame of the W boson and the W boson direction in the top quark rest frame. The decay of differently polarized W bosons in the top quark rest frame with the defined decay angle have been illustrated in Figure 1.11.

As described before, since neutrinos are nearly massless, the W bosons always couple to the left handed neutrinos. Charged leptons from the left handed W boson decays are likely to lie in the direction of the b quark, leading to $\cos \theta_l^*$ near to -1, while the charged leptons coming from the right handed W bosons fly mostly in the opposite direction, leading to $\cos \theta_l^* \approx 1$. From the longitudinally polarized W boson decays, most of times, charged leptons are emitted perpendicular to the direction of the b-quark. By ignoring some common factors, the helicity amplitudes for the process $t \rightarrow W^+ b \rightarrow l^+ \nu_l b$ can be written as [33]:

$$\begin{aligned}\mathcal{M}_L &\propto -e^{-i\phi_l^*} \left(\frac{1 - \cos \theta_l^*}{2} \right), \\ \mathcal{M}_0 &\propto -\frac{\sin \theta_l^*}{\sqrt{2}}, \\ \mathcal{M}_R &\propto -e^{i\phi_l^*} \left(\frac{1 + \cos \theta_l^*}{2} \right),\end{aligned}\tag{1.57}$$

where ϕ_l^* is defined as the azimuthal angle of the charged lepton in the rest frame of the W boson.

The most general form of the differential decay width as a function of both the polar (θ_l^*) and the azimuthal (ϕ_l^*) angles of the charged leptons is given by [34]:

$$\begin{aligned}\frac{1}{\Gamma} \frac{d\Gamma}{d\cos \theta_l^* d\phi_l^*} &= \frac{3}{16\pi} [(1 + \cos^2 \theta_l^*) + A_0 \frac{1}{2} (1 - 3\cos^2 \theta_l^*) + A_1 \sin 2\theta_l^* \cos \phi_l^* \\ &+ A_2 \frac{1}{2} \sin^2 \theta_l^* \cos 2\phi_l^* + A_3 \sin \theta_l^* \cos \phi_l^* + A_4 \cos \theta_l^* + A_5 \sin \theta_l^* \sin \phi_l^* \\ &+ A_6 \sin 2\theta_l^* \sin \phi_l^* + A_7 \sin^2 \theta_l^* \sin 2\phi_l^*],\end{aligned}\tag{1.58}$$

where the A_i ($i = 0, 1, \dots, 7$) coefficients depend on the charge, the transverse momentum, and the rapidity of the W boson. Coefficients A_5 , A_6 , and A_7 vanished because of asymmetry under time-reversal and parity operator. By integrating equation 1.58 over the azimuthal angle, we obtain:

$$\frac{1}{\Gamma} \frac{d\Gamma}{d\cos\theta_l^*} = \frac{3}{8}[(1 + \cos^2\theta_l^*) + A_0\frac{1}{2}(1 - 3\cos^2\theta_l^*) + A_4\cos\theta_l^*]. \quad (1.59)$$

By considering $\cos\theta_l^*$ dependency for different helicity amplitudes in equation 1.57, one obtains the following parametrization of the W boson cross section:

$$\frac{1}{\Gamma} \frac{d\Gamma}{d\cos\theta_l^*} = \frac{3}{8}F_L(1 - \cos\theta_l^*)^2 + \frac{3}{4}F_0(1 - \cos^2\theta_l^*) + \frac{3}{8}F_R(1 + \cos\theta_l^*)^2, \quad (1.60)$$

where, as defined before, the three parameters F_L , F_0 , and F_R determine the values of the left handed, the longitudinal, and the right handed helicity fractions, respectively. An integration over $\cos\theta_l^*$ leads to the expected constraint $F_L + F_0 + F_R = 1$. By comparing the two equations 1.59 and 1.60, the defined helicity fractions in terms of the coefficients, A_i are given as:

$$F_L = \frac{1}{4}(2 - A_0 - A_4), \quad F_0 = \frac{1}{2}A_0, \quad F_R = \frac{1}{4}(2 - A_0 + A_4). \quad (1.61)$$

Figure 1.13 shows the angular distribution of the charged leptons for different W boson helicity configurations which implies a good separation power between the three helicity states.

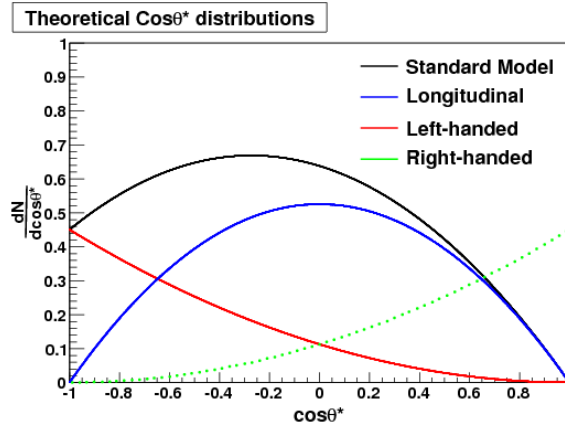


Figure 1.13: The differential $\cos\theta_l^*$ distributions for each W boson polarization state as well as the SM prediction.

Experimental results of W boson polarization

Several measurements on the W boson helicity fractions in the top-quark decays have been performed by the CDF, the DØ, the ATLAS, and the CMS experiments. The table 1.1 compares the most recently published results of different experiments as

well as our measurement. It can be concluded that all measurements are compatible with each other, and also compatible with the expectations from the SM [27].

Table 1.1: Summary of recent W boson helicity measurements of different experiments.

Experiments	$\sqrt{s}$ [TeV]	$\int \mathcal{L} dt [\text{fb}^{-1}]$	Channel	Nominal $\pm$ stat. $\pm$ syst.	Ref.
CDF	1.96	5.1	dilepton	$F_0 = 0.7^{+0.18}_{-0.17} \pm 0.06$ $F_R = -0.09 \pm 0.09 \pm 0.03$	[35]
CDF	1.96	8.7	l+jets	$F_0 = 0.73 \pm 0.07 \pm 0.07$ $F_R = -0.05 \pm 0.04 \pm 0.06$	[36]
DØ comb.	1.96	5.4	l+jets dilepton	$F_0 = 0.67 \pm 0.08 \pm 0.07$ $F_R = 0.02 \pm 0.04 \pm 0.03$	[37]
Tevatron comb.	1.96	2.7-5.4	l+jets dilepton	$F_0 = 0.72 \pm 0.06 \pm 0.05$ $F_R = -0.03 \pm 0.03 \pm 0.03$	[38]
ATLAS comb.	7	1.04	l+jets dilepton	$F_L = 0.32 \pm 0.02 \pm 0.03$ $F_0 = 0.67 \pm 0.03 \pm 0.06$ $F_R = 0.01 \pm 0.01 \pm 0.04$	[39]
CMS	7	5.0	l+jets	$F_L = 0.31 \pm 0.02 \pm 0.02$ $F_0 = 0.68 \pm 0.03 \pm 0.03$ $F_R = 0.01 \pm 0.01 \pm 0.01$	[40]
CMS	8	19.6	l+jets	$F_L = 0.35 \pm 0.01 \pm 0.02$ $F_0 = 0.66 \pm 0.02 \pm 0.02$ $F_R = -0.01 \pm 0.01 \pm 0.02$	[41]
CMS	8	19.7	single top	$F_L = 0.30 \pm 0.03 \pm 0.03$ $F_0 = 0.72 \pm 0.04 \pm 0.04$ $F_R = -0.02 \pm 0.02 \pm 0.01$	[42]
CMS	8	19.7	dilepton	$F_L = 0.33 \pm 0.01 \pm 0.03$ $F_0 = 0.65 \pm 0.02 \pm 0.02$ $F_R = 0.02 \pm 0.01 \pm 0.03$	[43]

Chapter 2

Experimental apparatus

In order to check the SM validity as well as to search for possible beyond the SM phenomena, high energy particle colliders were employed. The main task of a particle collider is to accelerate two charged particle beams, collide them and provide enough energy to produce final states with new particles. To study physical processes on the final state particles, an appropriate particle detectors could be served around the interaction point. One of large and energetic colliders in the world is the Large Hadron Collider (LHC) [44] which provides scientists with tools to verify the SM.

The present work in this thesis is based on proton-proton collisions data produced by the LHC at a center-of-mass energy $\sqrt{S} = 8 \text{ TeV}$, which recorded by the Compact Muon Solenoid (CMS) detector [45, 46] during the 2012. In the first part of this chapter, description of the LHC machine as well as all installed detectors will be briefly discussed. In the section 2.2, the CMS experiment will be described in more details and it's main components will be outlined.

2.1 The Large Hadron Collider

The LHC machine is currently the most powerful circular hadron accelerator in the world. It is located at the European Organization for Nuclear Research (CERN) near Geneva in the border region of France and Switzerland. The LHC was constructed in the 27 km tunnel previously occupied by the Large Electron Positron (LEP) accelerator which has been operating until it's decommissioning in 2000. It has been designed to accelerate two circular, opposite way beams of protons up to center-of-mass energy of $\sqrt{s} = 14 \text{ TeV}$. In addition, the LHC offers the possibility to study heavy ion collisions by accelerating lead ions at the center-of-mass energy of $\sqrt{s} = 5.5 \text{ TeV}$.

The most important parameter of the LHC besides the centre-of-mass energy is the number of proton-proton collisions produced per second. In a typical particle collider, the rate of events for any process is given by:

$$N = \sigma L, \quad (2.1)$$

where σ is the cross section for the studied process which represents the probability for a certain process to happen. In the above formula, L is the instantaneous luminosity derived entirely from machine parameters and depends only on the beam parameters. The instantaneous luminosity L for two colliding bunches a and b consisting of N_a and N_b particles can be written as:

$$L = \frac{n_b N_a N_b f_{rev} \gamma_{rel}}{4\pi \epsilon_{norm} \beta^*} F. \quad (2.2)$$

In this equation, n_b is the total number of bunches, f_{rev} and $\gamma_{rel} = E_P/m_P$ are the revolution frequency of the bunches within the orbit and the Lorentz factor, respectively. β^* and ϵ_{norm} denote the beta function and the normalised beam emittance which describe the beam focusing at the interaction point. The geometric luminosity reduction factor F accounts for the crossing angle, θ_c , between the two beams:

$$F = 1 + \frac{\sigma_z \theta_c}{2\sigma^*}. \quad (2.3)$$

By considering two symmetric beams, σ^* is the transverse quadratic mean beam size in x and y directions. To derive the above equation, the longitudinal quadratic mean beam size in z direction, σ_z , is assumed to be much smaller than β^* .

Since the number of protons in the beam decreases during a run, the instantaneous luminosity is varied time function. The integrated luminosity, $L_{int} = \int L dt$ is a common parameter to measure the recorded dataset size over a specific time period.

The LHC machine is started to produce proton-proton collisions at the end of 2009. In the first two years of the LHC operation, collisions have been recorded at a center-of-mass energy of $\sqrt{s} = 7$ TeV which was half of the design value. During this time, the instantaneous luminosity increased continuously reached to $3.5 \times 10^{33} \text{cm}^{-2}\text{s}^{-1}$ at the end of 2011. The protons beam energy was raised during 2012 to 4 TeV per beam corresponding to $\sqrt{s} = 8$ TeV center-of-mass energy. The progress of the accelerator in terms of integrated luminosity is shown in Figure 2.1. Currently the LHC is starting a new data tacking run at a center-of-mass energy of $\sqrt{s} = 13$ TeV with a maximum luminosity of $10^{34} \text{cm}^{-2}\text{s}^{-1}$.

Before the proton beams are ready to be injected into the LHC machine, they first need to have an initial energy of 450 GeV by using a number of smaller particle accelerators. The brief description of a proton accelerated processes through the accelerator complex of CERN is illustrated in Figure 2.2.

These pre-acceleration devices increase the proton energy gradually and make

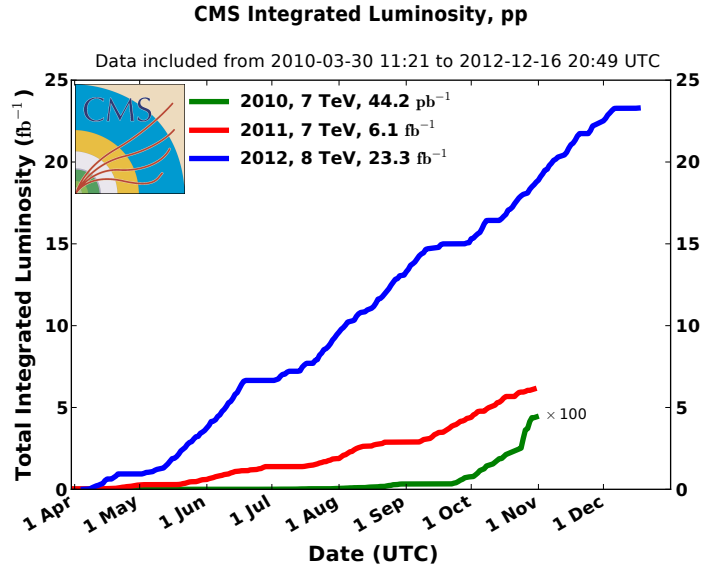


Figure 2.1: LHC integrated luminosity delivered to CMS during 2010-2012 as a function of time, shown separately for the 2010 (green), the 2011 (red) and the 2012 (blue) data-taking periods [47].

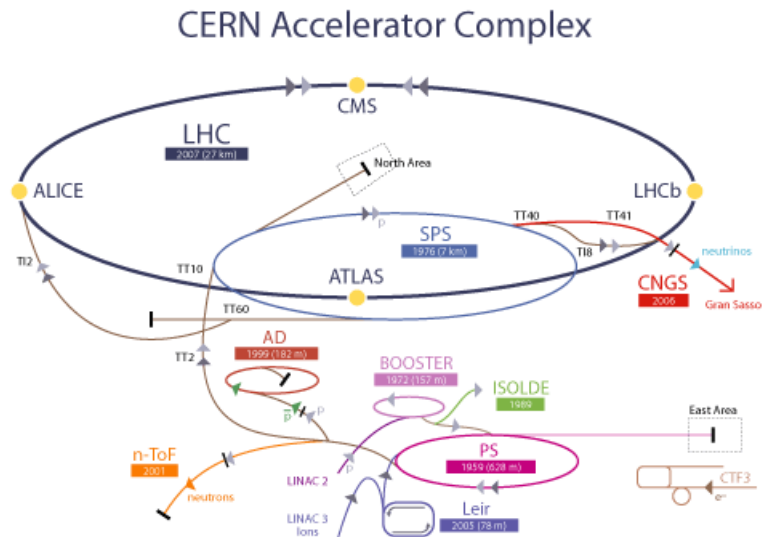


Figure 2.2: Sketch of the CERN Accelerator Complex.

the spacing between two bunches 25 ns. Protons are obtained by removing electrons from hydrogen atoms. The protons are passed to the first accelerator in the chain, LINAC2 linear accelerator, where they are accelerated to 50 MeV in energy. From there they are passed to the Proton Synchrotron Booster (PSB) to be accelerated to 1.4 GeV and then are injected into the Proton Synchrotron (PS), bunched and accelerated to 25 GeV. After that, protons are passed to the Super Proton Synchrotron (SPS) which increases the energy to 450 GeV, and then injected to the LHC. Finally, the beam splits up into two beams and are injected in opposite directions into the LHC via the TI2 and TI8 transfer lines as shown in Figure 2.2.

These two beams are not continuous, but consist of bunches of protons at an interval of 25 ns, accumulating to as many as 2808 per beam. Upon completion of the fill, the bunched beams are focused and collided at the different interaction points where detectors are positioned. In order to keep the protons in the LHC beam pipes, high dipole magnets with a field strength of 3.8 T are used. To achieve such a high magnetic field, the LHC uses superconducting magnets operating at a temperature of 1.9 K. Additionally, around each of the four interaction points, a few quadrupole and sextupole magnets for collimating, focusing and orbital corrections of the beams are installed.

Along the four interaction regions, six large experiments were installed. They are dedicated to perform the precise measurement of the particle collision events delivered by the LHC. Two very big general purpose experiments installed at high luminosity points are the ATLAS (A Toroidal LHC ApparatuS) [48] and the CMS [49]. They are designed to be usable for the search of new phenomena beyond the SM as well as high-precision measurement of physics processes such as the top quark. More information about the CMS detector will be discussed in section 2.2. Two smaller experiments, ALICE (A Large Ion Collider Experiment) [50] and LHCb (Large Hadron Collider beauty) [51] are specialized in different phenomena. The ALICE detector is specialised to study heavy-ion physics with investigation of the quark-gluon plasma. Only one month per year is devoted for data taking of this detector, when the LHC is operated with lead ions instead of protons. The LHCb detector focuses on physics with hadrons originating from the bottom quarks. One of the main physics objectives of this experiment is to find out matter-antimatter asymmetry puzzle in physics.

In addition to these four main experiments, there are also two small experimental facilities located few hundreds of meters away from the ATLAS and the CMS interaction points. The LHCf (LHC-forward) [52] experiment uses forward particles created inside the LHC as a source to simulate cosmic rays in laboratory conditions. The TOTEM (TOTAl Elastic and diffractive cross section Measurement) [53] detector studies the total proton-proton cross section and elastic proton scattering at

very low angles with respect to the colliding beams.

2.2 The Compact Muon Solenoid experiment

One of two general purpose detectors at the LHC, the CMS experiment, is designed to perform a large variety of measurements. The main task is to investigate wide physics programs through identifying different kinds of particles by measuring their properties like energy, momentum and charge. The data taken by this experiment is used to perform precise measurements of the SM parameters as well as to search for evidence of new physics beyond the SM.

The CMS detector is placed 100 m underground at an interaction point of the LHC ring. The CMS experiment is arranged in onion-like structure with different layers of materials and detection equipments. It has a cylindrical symmetry shape around the direction of the beam axis, with an overall length of 22 m, 15 m in diameter, and a total weight of 12500 tons. The overall geometry consists out of a central barrel part which closed by two endcaps parts, as shown in Figure 2.3.

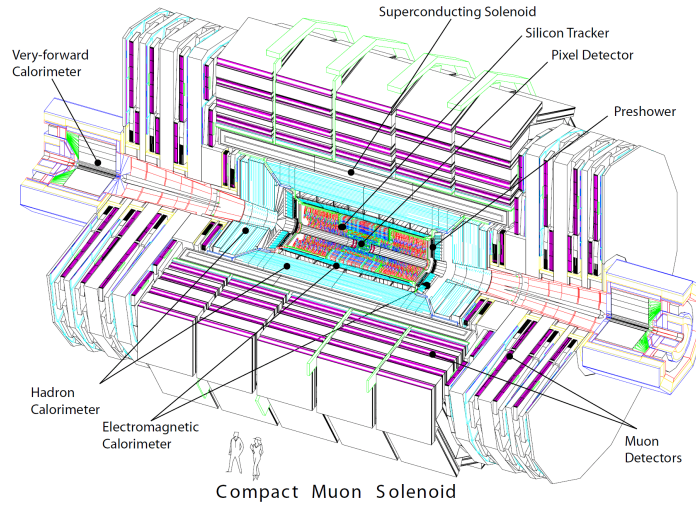


Figure 2.3: Perspective view of the subdetectors of the CMS-detector.

In the CMS design, subdetectors are arranged such that a good coverage of the solid angle is achieved. Although the CMS detector is significantly smaller than the very large ATLAS detector, but the weight is nearly doubled. The larger weight is mostly due to the iron return yoke which is constructed around the very large superconducting solenoid magnet. The main feature of the CMS detector is the superconducting solenoid, which allows to bend the high energetic particles in a compact design by providing a strong magnetic field.

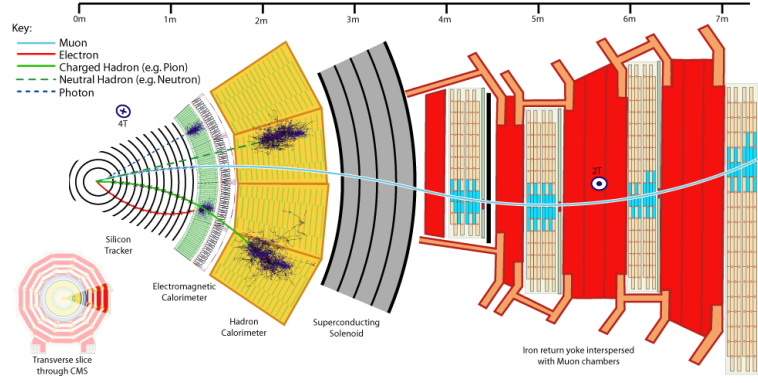


Figure 2.4: A schematic overview of a slice of the CMS detector with the components in the transverse plane.

At the center of the CMS and closest to the beam axis and the nominal interaction point, a silicon-based tracking detector is installed. The inner tracking system consists of an inner pixel and silicon-strip detectors allows to reconstruct charged particle trajectories as well as to measure their momentum in the magnetic field. More details about the CMS tracking system can be found in section 2.2.2. The tracking detector is surrounded by the electromagnetic and hadronic calorimeters. The energy of electrons and photons are measured by the electromagnetic calorimeter (ECAL) by absorbing these particles. Similarly, the hadronic calorimeter (HCAL) is designed to measure the energy of hadrons. More information about the calorimeter systems are given in sections 2.2.3 and 2.2.4. The tracking and calorimeter systems are located inside a superconducting solenoid generates a magnetic field of 3.8 T used to bends the charged particle tracks. In the outermost layer of detector, several muon chambers are installed to measure momentum and energy of the muons with high resolution, traversing all detector layers. In section 2.2.6, the CMS Muon system with more details are described.

The interaction of various particles with the different subdetectors is shown in Figure 2.4. The photon is a neutral particle and hence does not leave any track in the tracking system. Since the photon interacts through the electromagnetic force with the detector material, it can be detected by the ECAL subdetector. An electron particle interacts with the inner tracker part, leaving a track due to ionisation and deposits all of it's energy on the ECAL through electromagnetic interaction. The proton as a charged hadron, leaves track on the inner tracker, passes from the ECAL and interacts with the HCAL subdetector. The neutral hadrons such as neutrons pass from the tracker and the ECAL detectors without any response, while loose their energy in the HCAL system. Muons are much more penetrating than electrons

because of their lower charge to mass ratio. Then, they travel across the whole detector without significant energy loss. A muon leaves a track in the tracker and the muon chambers and it deposits small fraction of its energy on the ECAL and the HCAL. Neutrinos have no charge so leave no detectable tracks in the tracking system. Due to very low interaction cross section with ordinary matter, they escape direct detection. The neutrino energies are obtained by summing the energies of the other observed particles as the amount of missing energy in the detector.

In the following sections, after introducing the CMS coordinate system, different CMS subdetectors will be presented going from the centre to the outside. A more detailed information about the CMS subdetectors can be found in Ref [49].

2.2.1 The CMS coordinate system

To describe event reconstruction in the detector, it is useful to introduce a coordinate system centred at the nominal proton-proton interaction point. The x-axis is horizontal points towards to the center of the LHC ring, while the y-axis points vertically upward. The z-axis is set along the anticlockwise beamline in such a way that the coordinate system is right-handed. To characterise the location of detector material, and even particles, two angular quantities are defined with respect to the coordinate axes. The azimuthal angle ϕ is defined as the angle between the object direction and the x-axis in the transverse plane. Furthermore, the polar angle θ is measured from the z-axis and usually expressed by the pseudorapidity η given by:

$$\eta \equiv -\ln \tan \frac{\theta}{2}, \quad (2.4)$$

which has the advantage that are invariant under Lorentz boosts along the beam direction. Another useful quantity is the rapidity defined as:

$$y \equiv \frac{1}{2} \ln \frac{E + P_z}{E - P_z}. \quad (2.5)$$

This parameter is determined such as the particle flux per rapidity interval is constant for hadron-hadron colliders. For massless particles or in the high energy limit, where $E \approx |P|$, the rapidity and the pseudorapidity are equal.

2.2.2 Tracking system

The central part of the CMS detector is the silicon tracker [54, 55]. It is one of the most important subdetector of the CMS which uses semi-conducting silicon as active material. The tracker is designed to provide a precise measurement of the charged particles trajectories emerging from LHC collisions. It is also used to reconstruct the primary vertex as a origin of charged particle tracks. A tracker has a cylindrical

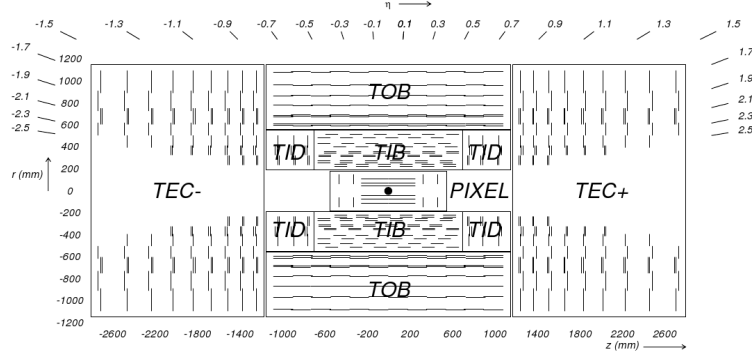


Figure 2.5: Schematic overview of the CMS tracking system. Each line represents a detector module. Double lines indicate back-to-back modules.

shape with a sensitive silicon area of about 210 m^2 , length of 5.8 m and diameter of 2.5 m. The CMS solenoid provides a homogeneous magnetic field of 3.8 T allowing tracking system to accurately measures charged particle momenta. Due to presence high track multiplicity in the tracker, the system needs to withstand the radiation from this flux of particles for a long time period.

The structure of the tracker is based on a doped semiconductor which operated in reverse-biasing voltage mode. As charged particles passing through the semiconductor, due to ionisation they create electron-hole pairs resulting in an electronic signals which are interpreted as hits. The trajectories of charged particles can be reconstructed by using a large number of hits in different layers of tracking system. In order to reduce the electronic noise from the silicon dark current, the whole tracking system is cooled down at a constant temperature of -20°C . Due to increasing intensity of particles flux with decreasing distance from the interaction point, the tracking system can be divided into two subdetectors. Each part uses a different technology with the required granularity and radiation hardness from the incident flux of particles. A schematic overview of the tracking system is given in Figure 2.5.

The CMS tracking system consists of two main components: the silicon pixel detector in the innermost region and silicon strip tracker in the outer region. The whole structure of tracking system provide coverage up to $|\eta| \approx 2.5$.

Silicon pixel detector

The pixel detector is the innermost part of the CMS detector. It is the most important detector for determining the impact parameter used for secondary vertex reconstruction by giving very precise measurement of the particle locations. The pixel detector consists of three barrel layers (BPIX) and four forward disks in the endcap (FPIX) as shown in Figure 2.6.

Three cylindrical barrel layers are 53 cm long which are parallel to the beam

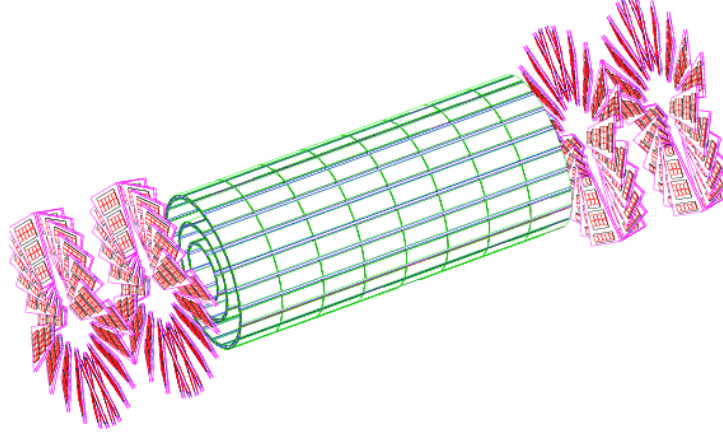


Figure 2.6: Three dimensional sketch of the CMS pixel detector showing the BPIX (green) and FPIX (red) modules.

axis and centred on the interaction point. They are installed respectively at a radius of 4.4 cm, 7.3 cm and 10.2 cm and contain 48 million pixels with a total active area equals to 0.78 m^2 . There are also two endcap disks at each side extending from 6 cm to 15 cm in radius and located at longitudinal distances of $z = \pm 34.5 \text{ cm}$ and $z = \pm 46.5 \text{ cm}$. The FPIX discs contain 18 million pixels and cover an active area of 0.28 m^2 . Each pixel cell of the detector has a size of $100 \mu\text{m} \times 150 \mu\text{m}$. The small size of cells enables the pixel detector to achieve good level of tracking resolution even in events with high track multiplicity. These set of pixel cells provide two-dimensional measurements of the track with an average spatial resolution of $10 \mu\text{m}$ in the $r - \phi$ plane, and $15 \mu\text{m}$ in the z direction. This high granularity is needed for an efficient reconstruction of charged particle tracks with at least 3 tracking points over the pseudorapidity range coverage of $|\eta| < 2.5$.

Silicon strip detector

Outside the pixel detector, the particle flux gets low enough to be able to use a silicon strip detector rather than pixel cells. The advantage is a lower number of readout channels as well as amount of material, power consumption and cost but with a small loss of accuracy. This subdetector consists of 9.3 million silicon strips, arranged in 15148 detector modules which is located at the radial region between 20 cm and 116 cm. The type of installed strip sensors are p-on-n, the formed strips

by p^+ implants in an n type silicon bulk. The silicon strip sensors are mounted on modules, which also contain the read-out and control chips for the sensors. The strip detector is subdivided into four subsystems: Tracker Inner Barrel (TIB), Tracker Inner Disk (TID), Tracker Outer Barrel (TOB) and Tracker End Cap (TEC).

The TIB consists of four concentric cylindrical layers located at radii of 0.255, 0.339, 0.419 and 0.498 m from the beam axis. These layers with a length of 1.4 m are parallel to the beamline extend from -0.7 m to +0.7 m in the z direction. The two inner layers of the TIB use double sided modules, while the outer two layers contain single sided modules. Since the amount of particles flux is decreased by increasing distance from interaction point, the pitch strip of TIB layers are varied between $80\text{ }\mu\text{m}$ and $120\text{ }\mu\text{m}$ from inner to outer parts. The occupancy for these subdetector is typically about 2 to 3% and spatial resolutions are $23\text{ }\mu\text{m}$ to $34\text{ }\mu\text{m}$ along $r - \phi$ plane and $230\text{ }\mu\text{m}$ in the z direction.

The TIB subdetector is complemented with three additional disks at each side called the TID system. These endcap disks are located between 0.8 m and 0.9 m from the interaction point along the z axis. Each disk is made of three rings of modules, spanning a radius between 0.20 m and 0.5 m centred on the z axis. Similar to the TIB, the inner two rings on each disk are made of double sided modules while outer ring hosts single sided ones. The strips of TID modules have a pitch size of $80\text{ }\mu\text{m}$ to $141\text{ }\mu\text{m}$ for different rings and disks. The TIB and TID use strips with a length of 10 cm and a sensor thickness of $320\text{ }\mu\text{m}$.

The TOB is located outside the TIB and composed of six cylindrical layers with radii of 0.608, 0.692, 0.780, 0.868, 0.965 and 1.080 m. Each layer with a sensor thickness of $500\text{ }\mu\text{m}$ extends 1.18 m from the interaction point in both directions along the z axis. The first two layers of the TOB contain double sided modules, while single sided modules are used for the four outer ones. This leads to an occupancy at the percent level with the the spatial resolution between $35\text{ }\mu\text{m}$ and $52\text{ }\mu\text{m}$ in the $r - \phi$ plane and about $530\text{ }\mu\text{m}$ along the z direction.

The TEC with nine disks at each side, encloses the rest of the tracker. These disks are installed at a radius of 0.22 m to 1.14 m which extend from $\pm 1.24\text{ m}$ to $\pm 2.8\text{ m}$ along the z direction. Each disk contains up to 7 rings of modules, where all the modules are placed in concentric rings with respect to the beam line. The inner two rings and the fifth ring from the center made of double sided modules while the rest of the TEC uses single sided ones. The four innermost rings of the TEC use thinner sensors with a thickness of $320\text{ }\mu\text{m}$, while the ones in the three outer rings are $500\text{ }\mu\text{m}$ thick. For the TOB and the TEC, strips with a length of 25 cm and the varied pitch sizes from $97\text{ }\mu\text{m}$ to $184\text{ }\mu\text{m}$ accordingly to the distance from the interaction point are used. The main objective of the tracking system is the precise reconstruction of the charged particle trajectories. Then it is used to identify and

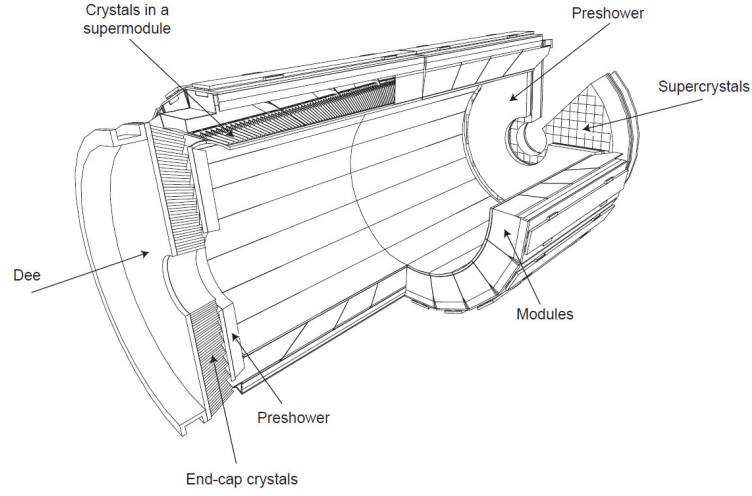


Figure 2.7: Layout of the CMS electromagnetic calorimeter.

to reconstruct primary vertices as well as other vertices originating from the decay of long lived particles.

At the high luminosity phase of the LHC machine (HL-LHC), the tracking system of the CMS will be exposed to harsher conditions than the current installed system. Therefore, the new radiation hard sensors to cope with higher radiation and occupancy levels are needed to be installed. The CMS tracker collaboration has been performed a comparative survey of large silicon material and technologies to identify the baseline material and design that fulfills all requirements of the future tracker [56]. In this thesis, in addition to the data analysis topic, some studied sensor R&D for the CMS outer tracker upgrade project have been done, presented with details in next chapter and reference [56].

2.2.3 Electromagnetic calorimeter

The ECAL [57] is the next subdetector enclosing the tracking system which is located inside the solenoid. The ECAL system is a nearly hermetic and homogeneous calorimeter designed to measure the energy of electrons, positrons and photons. In order to stop high energetic particles by producing a cascade of photons and electron-positron pairs within the material, the very dense material has to be used. Since the calorimeter consists of a scintillating material, the showering induces photons and the number of emitted photons is a direct tool to measure the energy of the original particle. The ECAL shown in Figure 2.7, is built out of lead tungstate crystals (PbWO_4) that are transparent and therefore operate as scintillator and absorber at the same time.

Due to the relatively short radiation length ($X_0 = 0.89 \text{ cm}$) of these crystals,

the ECAL can stop high energetic particles within short distance. These radiation hardness crystals with fine granularity are capable of emitting 80% of the total scintillation energy within 25 ns which is a time length corresponding to the bunch crossing frequency of the LHC. They also have a small Moliere radius of 2.2 cm which allows the ECAL calorimeter to be kept compact. The transparency of the crystals has a major influence on the amount of light detected by the read-out electronics. In order to achieve a reliable energy measurement, amount of detected light has to be precisely measured for every crystal. The environment temperature can be affect on light yield of lead tungstate crystals. For this reason, a cooling system has been installed within the ECAL which keeps the whole system at a constant temperatures of 18°C. Totally the ECAL of the CMS detector is composed of 75848 scintillating lead tungstate crystals.

The ECAL calorimeter can be segmented into three parts, the ECAL Barrel (EB), the ECAL Endcaps (EE) and the ECAL Preshower (ES). The barrel part covers the pseudorapidity range of $|\eta| < 1.479$ which contains 61200 PbWO₄ crystals. The endcap region extends the pseudorapidity coverage to $|\eta| < 3.0$ with the number of 7324 crystals in each of two sides. The preshower detector is located in front of the endcaps. The mechanical structure of crystals is designed with the shape of a truncated pyramid in a way that is most suitable for the read-out electronics. The layout of the ECAL with the introduced subdetectors is shown in Figure 2.7.

ECAL barrel

The cylindrical EB calorimeter is located at a distance of 1.3 m from the beam pipe. The crystals populating the barrel have a front plane, the one facing towards the interaction point, of $22 \times 22 \text{ mm}^2$ and a backward plane of $26 \times 26 \text{ mm}^2$. In terms of pseudorapidity coverage, each crystals corresponds approximately to 0.0174×0.0174 in $\eta - \phi$ phase space. These crystals are mounted such that their surfaces are facing the interaction region though they are tilted by 3° to reduce the effect of particles crossing the boundary in between adjacent crystals. The length of the EB calorimeter crystals is 230 mm corresponding to $25.8 X_0$. Scintillation light is measured with photodetectors attached to the rear face of the crystal. In the barrel, silicon avalanche photodiodes (APDs) are used, since they are insensitive to the high axial magnetic field. There are two APDs on the back of each crystal which read out in parallel and each pair has a mean gain of 50.

ECAL endcap

The ECAL endcaps are located on either side of the barrel. Each endcap is 3.15 m away from the interaction point and together they cover the pseudorapidity range

of $1.479 < |\eta| < 3.0$. The crystals in the endcaps are a little bit larger than barrel region with a front face and rear face cross sections of $28 \times 28 \text{ mm}^2$ and $30 \times 30 \text{ mm}^2$, respectively and a length of 220 mm corresponding to $24.7 X_0$. Similar to the EB calorimeter, the crystals are not facing exactly towards the interaction point, but they are slightly tilted. For the end caps, vacuum phototriodes (VPTs) are used because they withstand the high radiation in this region. In the back of each endcap crystals, one VPT with a mean gain of 10.2 at zero magnetic field is attached. The value of VPT gain is decreases by about a factor of 2 when the CMS magnet is operational.

ECAL preshower

In front of each end caps, there is a two-layer sampling calorimeter with high spatial resolution in the region $1.653 < |\eta| < 2.6$. This subdetector has a much greater spatial resolution than the ECAL itself, making it an ideal complement for the system in the high pseudorapidity region where the flux of crossing particles is stronger. The ES subdetector consists of a double-layer lead-silicon calorimeter, with the lead initiating the shower and the silicon strip detector placed after each lead layer to measure both the position and deposited energy of particles traversing the detector with high granularity. The main goal of the preshower is to identify neutral pions, in the mentioned pseudorapidity region, where the angle between the two photons from the decay of the neutral pion is small enough to fake the signal of just one photon with larger energy. It also improve the identification and position resolution for electrons and photons.

ECAL resolution

The ECAL allows a very precise energy measurement of electrons, positrons, and photons. The energy resolution for the ECAL can be characterized as a function of the energy by the following formula:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2, \quad (2.6)$$

where the energy of a particle E is given in units of GeV. The first term describes stochastic effects which depends on event-to-event fluctuation in lateral shower containment, photo-statistics and photo-detector gain. The second term arises from electronic noise which is given in units of MeV. The last constant term is driven by the non-uniformity of the longitudinal light collection, calibration errors and leakage of energy from the back of the crystals. By fitting test beam data to this function, the parameters can be extracted. The ECAL energy resolution as a function of

the energy for a test beam is shown in Figure 2.8 [49]. A typical energy resolution extracted from a test beam data is found to be:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right)^2 + \left(\frac{0.12}{E}\right)^2 + (0.30\%)^2, \quad (2.7)$$

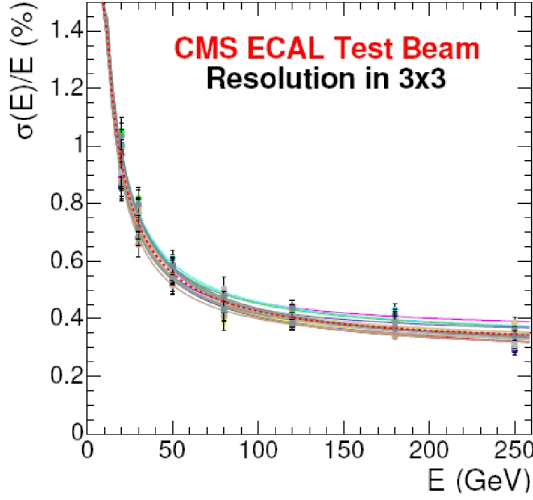


Figure 2.8: ECAL energy resolution of a matrix of 3×3 crystals as a function of energy.

2.2.4 Hadronic calorimeter

All hadrons pass the ECAL without much energy loss, then a second calorimeter has to be installed just after the ECAL which is dedicated to hadrons. The HCAL [58] is designed to measure the energy of hadrons with enough accuracy to allow for the precise determination of jets and missing transverse energy. Unlike the ECAL, the HCAL is a sampling calorimeter consisting of non-magnetic brass absorber plates interleaved with plastic scintillators to measure the deposited energy. The number of absorber and plastic scintillator layers are not constant within the calorimeter depending on the installed position in the detector. In order to make precise measurements on missing transverse energy, the HCAL must be cover the whole solid angle (η, ϕ) as well as providing good spatial resolution. The pseudorapidity coverage is limited because of the beam-pipe which has to traverse the HCAL.

The HCAL subdetector is split into four subsystems: the hadron barrel (HB), hadron endcap (HE), hadron outer (HO), and hadron forward (HF) calorimeters. The HB and the HE subdetectors surround the ECAL and are located inside the magnet whereas the HF and the HO are outside of it which together cover the pseudorapidity range of $|\eta| < 5$. A schematic overview of the HCAL subdetectors is given in Figure 2.9.

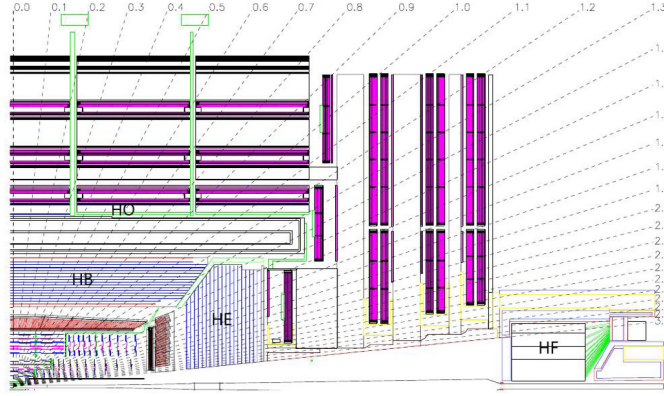


Figure 2.9: Longitudinal view of the CMS hadronic calorimeter, showing the locations of different subdetectors.

HCAL barrel

The HB extends from an inner radius of 1.77 m (the outer radius of the ECAL) to an outer radius of 2.95 m (the inner radius of the magnet) which covers the pseudorapidity range of $|\eta| < 1.3$. It is made of 36 identical azimuthal wedges, arranged into two groups of 18 to form the two halves of the barrel. These wedges are made of 16 layers of absorbers followed by plastic scintillating tiles. The innermost one is a 4 cm thick steel plate, the next eight brass plates are 5.05 cm thick while the other six brass plates made of 5.65 cm thick and finally followed by a 7.5 cm thick steel plate. The effective thickness of the absorbers in terms of interaction length (λ_I) increasing from $5.82 \lambda_I$ at $\eta = 0$ to $10.6 \lambda_I$ at $|\eta| = 1.3$.

HCAL endcap

The HE calorimeter on both sides of the barrel overlaps slightly with the HB sub-detector and covers the pseudorapidity range of $1.3 < |\eta| < 3$. The HE layers are made out of 7.9 cm thick brass absorber plates surrounded by plastic scintillator.

HCAL outer

The size of the barrel of HCAL is limited by the radius of the solenoid. Due to the limited amount of absorber material in the HB, hadronic shower leakage can sometimes happen. As a consequence, an extra layers are added just outside the magnet called the HO subdetector which collect the energy coming from residual particles that crossed all the material inside the magnet. The HO system with the same segmentation of the HB, except iron absorbers instead of brass, covers pseudorapidity range of $|\eta| < 1.26$. Two scintillators are used for the central ring (at $\eta = 0$) while there is only one scintillator for the rest of pseudorapidity regions.

The HO subdetectors extends the total thickness of the HCAL to a minimum of $11 \lambda_I$.

For each of these three subdetectors, the scintillator is segmented into HCAL towers both in η and ϕ to provide the necessary granularity. For the HB and the HO subsystems, the calorimeter tower segmentation is 0.087×0.087 in $\eta - \phi$ phase space. The tower size in the HE part is the same as for the HB up to $|\eta| < 1.74$. Beyond this region, the tower sizes for the HE, are incrementally enlarged to a maximal size of $(\Delta\eta, \Delta\phi) = 0.35 \times 0.175$. The produced light from the plastic scintillators is collected by a set of wavelength shifting optical fibers and finally read out by using hybrid photodiodes (HPD). Each HPD consist of a photocathode kept at high voltage of about 8 kV and a pixelated silicon photodiode. These devices were chosen due to their high dynamic range, their high gain (~ 2000) and their capability to work in high magnetic field.

HCAL forward calorimeter

The HF calorimeter is placed along the beam line at a distance of 11.2 m from the interaction point. This subdetector by using a Cerenkov-based radiation-hard technology, extends the angular coverage to the high pseudorapidity region $2.9 < |\eta| < 5.2$. The large η range of the HF allows to measure hadronic activity in the very forward region by detecting forward jets. The HF is segmented in towers of 0.175×0.175 in $\eta - \phi$ except for $|\eta| > 4.7$, where the segmentation is 0.175×0.35 . Due to the high flux of particles in the forward regions, radiation-hard materials are chosen. The absorber material of the HF is made out of 5 mm thick steel plates.

The optical signal generated by charged shower particles, is produced by a set of Cherenkov-light-emitting quartz fibers instead of scintillation. The emitted Cherenkov light can be detected by photomultipliers. There are two different kinds of fibers, short fibers with a length of 22 cm and long fibers which run the entire 1.65 m length of the HF. These two fibres are read out separately which makes it possible to distinguish electromagnetic from hadronic jets because only the electromagnetic one can deposit a large fraction of their energy in the short fibers.

HCAL resolution

The HCAL has a much worse resolution in comparison with the ECAL which makes it necessary to do calibration run on that to retrieve the Jet Energy Correction (JEC) information. The resolution of the HCAL system was studied in a test-beam with single pions crossing the prototype detector [45]. When the information of ECAL and HB are combined, an energy resolution of 20 – 30% is obtained for pions below 50 GeV while for 300 GeV ones, the energy resolution is further improved to

less than 10%.

2.2.5 Solenoid magnet

The CMS detector is designed to achieve a momentum resolution of $\Delta p/p \sim 10\%$ for the particles with extreme high momenta of $p \approx 1 \text{ TeV}/c$. In order to fulfil this requirement, determine charge of energetic particle with good accuracy, and also keep the CMS detector reasonably compact without losing momentum resolution, the CMS detector has to be located in a strong solenoid magnet. This is accomplished using a superconducting solenoid capable of producing a 3.8 T magnetic field in the barrel region providing the bending power needed to achieve a good resolution for energetic charged particles. Additionally the iron return yoke bends the magnetic flux on the solenoids outside and leads to a field of 1.9 T in the opposite direction.

The CMS solenoid magnet has a longitudinal size of 12.5 m and an internal diameter of 6 m. It is made from a 4-layer winding of NbTi coils and has a mass of 220 tons. These coils are contained in an aluminium coating providing thermal and mechanical stability to the magnet. In order to reach the superconducting phase, the magnet coils are cooled up to 1.9 K by using a ^3He system. The produced magnetic flux is returned through a 10000 ton iron yoke surrounding the solenoid. The return yoke is made in 11 parts, five wheels and two endcaps made of three disks each. Each of five barrel wheels is 13 m long and has an outer radius of 7 m, while each of endcaps with same radius as barrel part is 4 m thick. The CMS magnet system with stored energy of 2.6 GJ, carry a total electric current of 20 kA.

2.2.6 Muon chambers

The muon detection and triggering are critical for physics goals of the CMS, as many new physics processes may manifest themselves via decay chains involving muons. Muons as minimum-ionising particles (except at very high energies) can pass through the calorimeters without depositing a large fraction of their energy and thus can be efficiently identified by dedicated muon detectors placed outside the calorimeters. The muon system [59] ensures a very precise reconstruction of the muons complementing the measurement of their momentum performed by the inner tracker using the reflux magnet.

For low p_T muons up to 200 GeV/c, momentum resolution achieved in the tracking system satisfies the CMS requirements while for higher momentum values up to 1 TeV/c, the muon systems are useful in keeping resolution of muons at approximately 5%. The muon system is embedded in the iron magnet return yoke of the solenoid magnet in the outermost region of the CMS detector. Geometry of the Muon system with pseudorapidity coverage range of $|\eta| < 2.4$ is shown in Figure

2.10.

There are three different gaseous detector technologies: Drift Tube (DT) in the barrel region, Cathode Strip Chamber (CSC) in the end cap and Resistive Plate Chamber (RPC) in both regions. The reason to use different detector technologies is mainly lies in the different particle rates and occupancies.

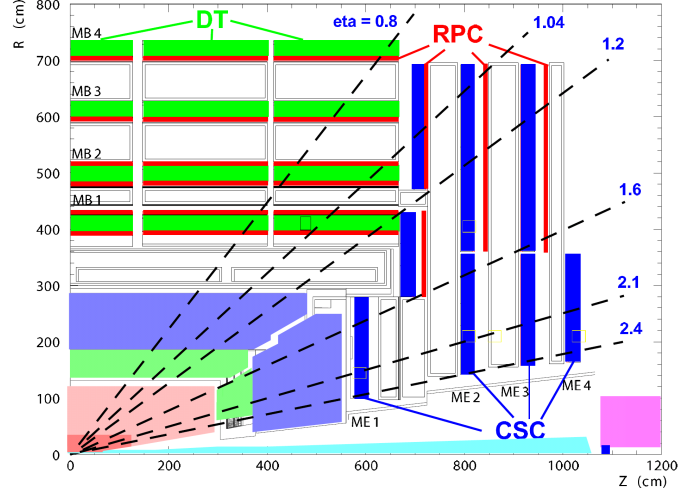


Figure 2.10: Location of the CMS muon detector in the r-z plane, showing the relative locations of the DT in green, RPC in red, and CSC in blue.

Drift tube

In the barrel region up to $|\eta| < 1.2$, four layers of DT chambers are interleaved with the iron return yoke which provides a magnetic field of 1.9 T for the outer part of the CMS detector. In this region the background (mainly neutron-induced) as well as the muon rate are small, allowing for the usage of drift tubes being relatively cheap and easy to handle. Four stations of the DT system are located in the spacing between the layers of the iron return yoke and labelled as MB1-MB4 in Figure 2.10.

Each of first three inner stations has 60 drift chambers which is equipped with two oriented super-layers to measure the muon position in the $r - \phi$ plane and one super-layer to measure the position in the z-direction. The fourth outer station has 70 drift chambers which contains only the two super-layers measuring the position in the $r - \phi$ plane. Each of these super-layers composed of four layers of drift tubes. The drift tubes are constructed of long rectangular chambers with cross section of $13 \times 42 \text{ mm}^2$ and 2.4 m in length, filled with a mixture of 85% Ar and 15% CO_2 gas. In this configuration, cathodes are located along the outside of the chambers, with an anode wire running along the center of the chamber.

As muons passing through the gas, molecules become ionized and the total charge is collected on the wire gives rise to an electronic pulse. The muon position

in each DT is reconstructed by measuring the arrival drift time of the charge to the anode wire and by converting it into a distance from the wire. The maximum drift distance is 2.1 cm corresponding to a drift time of 380 ns. These drift tubes have a very good single-hit spatial resolution of around $100\ \mu\text{m}$ in the $r - \phi$ plane and up to $150\ \mu\text{m}$ in the z direction, and an angular resolution of 1 mrad.

Cathode strip chamber

Each endcap of muon system is equipped with four stations of CSC, covering the $0.9 < |\eta| < 2.4$ pseudorapidity range. There is endcap-barrel overlap region between the barrel DTs and endcap CSCs, $0.9 < |\eta| < 1.2$, which muons are detected with both of them. In the forward region with non uniform and non zero magnetic field, the rate of muons are much higher than what experienced in the barrel. Hence, in this region the CSC chambers with a fast response time, a fine segmentation, and a better robustness against radiation damage are used. Four stations of the CSC detector called as ME1-ME4 are shown in Figure 2.10.

The inner-most disk ME1 is made of three rings of 18 chambers each, while ME2 and ME3 are composed by two rings of 18 and 36 chambers and finally the outermost one has only one ring of 18 chambers. The CSCs are multiwire proportional chambers composed of six anode wire planes interspersed with seven cathode panels. The nominal gas mixture used to fill the CSC chambers is 40% Ar, 50% CO₂ and 10% CF₄. The chambers are trapezoidal and cover either 10° or 20° intervals in ϕ . Each of these chambers has the cathode segmented radially in strips, providing a ϕ measurement, and the anode wires running perpendicular in the azimuthal direction, measuring the position in η .

By passing muons through the cathodes and producing an ionization trail, the electrons are accelerated to the wire where an avalanche occurs which induce charges on the cathode. Totally, there are 468 CSC chambers, with 220,000 cathode strips and 180,000 anode read-out channels. These chambers also provide a good single-hit position resolution varies from 75 to $200\ \mu\text{m}$ from the first to the last station with an angular resolution in the order of 10 mrad.

Resistive plate chamber

In order to ensure the accurate timing required for triggering, the muon system is also equipped with a set of the RPC chambers. The RPCs are located both in the barrel and on the endcaps and cover the $|\eta| < 1.6$ pseudorapidity range. These detectors are dedicated to trigger purpose due to their very fast response time with a good time resolution of the order of 1 ns. Thus, they are mainly used to improve the bunch-crossing assignment when triggering on muons. A total of six layers of RPCs

are embedded in the muon barrel system, two in each of the first two stations and one in each of the last two stations. In the endcap region, the RPCs are trapezoidal arranged in three rings in each of the first three stations. Each of RPC chambers is operated in avalanche mode and made with a double-gap structure in which the two gaps are filled with gas and have a single read-out strip in between them. The final signal is summed from the output of two gaps.

2.2.7 Data acquisition and trigger system

At the design luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$, the LHC will deliver up to 10^9 inelastic collisions per second. During nominal operation, the LHC operated at bunch crossing period of 25 ns leading to a 40 MHz frequency. This leads to a huge amount of processing and data storage which is several orders of magnitude larger than what can be stored with the most modern technologies. Because of these limitations, it is necessary to discard most of the events during the data taking by a factor of 10^6 and select only a subset of all events which are likely to contain physics processes of interest. This is achieved with a trigger system designed to take a decision on each event to be stored on tape or not.

The trigger system must be quick in the reconstruction and identification of the objects to be able to consider every aspects of the events within the very high frequency of the bunch crossing. The decision is divided into two sequential levels: the Level-1 (L1) Trigger [60] and the High-Level-Trigger (HLT) [61, 62] as illustrated in Figure 2.11. The first selection, L1 trigger, is based on hardware systems which roughly read out only a few subdetectors with simple selection criteria and a very fast decision system. Events passing this step are fed to the HLT which is a software based trigger system, using more detailed information of all subdetectors to partially reconstruct the events.

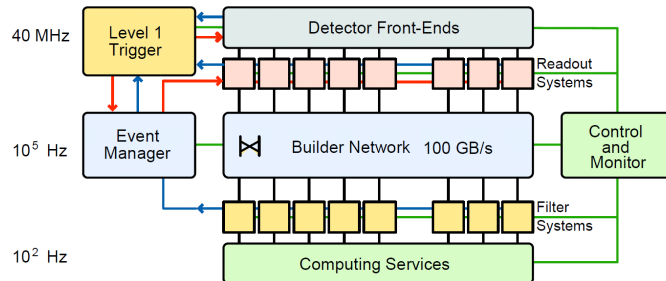


Figure 2.11: Architecture of the CMS Data Acquisition System (DAQ).

Level-1 Trigger

The L1 trigger is performed by custom electronics that is located partly on the detector and partly in the Underground Service Cavern (USC) adjacent to the experimental hall. The programmable hardware system consists of FPGAs (Field Programmable Gate Arrays) and ASICs (Application Specific Integrated Circuits), designed to take a fast accept/reject decision for every bunch crossing. The L1 trigger with a very fast response, is designed to reduce the rate of events accepted for further processing from 40 MHz to the less than 100 kHz.

The trigger must take a decision every 25 ns but the maximum processing time from generated electrical signals to a L1 trigger decision is estimated about $3.2\mu\text{s}$. Then in order to avoid any data loss within an allowed latency, the information of 128 bunch crossings recorded by each detector are pipelined, waiting for the final decision of the trigger system. Due to this short time scale, the L1 trigger uses only the information provided by the segmentation of calorimetric and muon system which are the two CMS subdetectors with the fastest time response. This information is used to identify electrons, muons, photons and jets over a certain threshold as well as rough estimation of the missing transverse energy E_T^{miss} , based on reduced granularity. These objects are only partially reconstructed without exploiting the full detector resolution. The L1 trigger can be divided into three majors subsystems: the Global Muon Trigger (GMT), the Global Calorimeter Trigger (GCT), and the Global Trigger (GT), as shown in Figure 2.12.

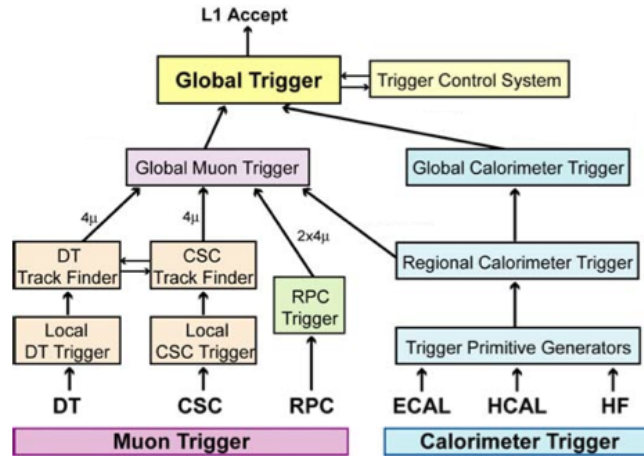


Figure 2.12: Flow of the L1 Trigger decision before data is being transferred to the DAQ.

As described in the last section, muon information is provided by DT, CSC and RPC subdetectors. The DT and the CSC systems can provide the muon tracks and hit patterns, used to reconstruct muon transverse momentum. The RPC is treated separately because it has a good timing resolution which allows it to clearly distin-

guish muons among different bunch crossings. The output of each muon subsystems are sent to the GMT where the final decision is taken.

The calorimeter segmentation trigger is based on trigger towers where the transverse energy from these towers is summed and associated with a bunch crossing. The calorimeters provide trigger system with a set of relevant physics variables like the E_T^{miss} , the scalar sum of the transverse hadronic activity (H_T), the number of jets above different energy thresholds as well as information of particle isolation. Also electron, photon and tau candidates can be identified by using the geometry of the energy distribution in the towers at this trigger level. These information is passed to the GCT to roughly reconstruct particles, calculates the E_T^{miss} and H_T parameters.

In order to make the final decision on whether to accept or reject the event at the L1 trigger, the GT uses a multi level architecture based on the GMT and GCT as described in Figure 2.12. Trigger primitive generators perform a local reconstruction using a small fraction of the interested subdetector, producing basic objects. Regional triggers match the information coming from different primitives, identify higher level objects and build physical quantities, such as p_T . Then described GMT and GCT systems combine information coming from different regional triggers. Finally the GT uses all of objects from the GMT and the GCT and selects events on the basis of programmable trigger conditions corresponding to predefined energy or momenta thresholds in advance. Events passing the L1 trigger are then passed to the HLT, while the others are discarded.

High Level Trigger

The HLT trigger system is designed to further reduce the rate of accepted events of 100 kHz to approximately 100 – 400 Hz which is required limit by the CMS computing infrastructure. Since the HLT uses information from the entire detector including the inner tracking system, it takes more time than the L1 trigger to make a decision. Unlike the L1 trigger, the HLT system is a software based trigger which has access to complete reconstructed event data exploiting the full resolution of most subdetectors and performs a simplified version of the CMS offline event reconstruction on these events. This computational complexity is done on a thousand commercial processors working in parallel.

The HLT makes a decision based on predefined sets of algorithms called paths which contain several trigger filters in sequence. The selection criteria for each of these trigger filters can be updated frequently according to the LHC beam conditions and some needed cuts dictated by the CMS Physics Analysis Groups. The HLT keeps only passed events from all of defined filters, then only certain fraction of events will be stored at the end. At the HLT, different systems are designed for the identification of different kinds of objects like electrons, photons, muons, taus,

jets (including b-quark jets) and E_T^{miss} . When one event passed by the HLT, it is transferred to the CERN computing centre which store the recorded data and the necessary simulated collision events.

Data storage and processing

Event information from each subdetectors containing digitalized output, the L1 and the HLT trigger are recorded in the RAW data formats. The reconstruction of collected and simulated data implies the necessities of a lot of computing resources. In order to make all the collected data available to scientists worldwide, a new distributed computing system has been created, involving different computing facilities in various parts of the world which is often referred to as the Worldwide LHC Computing Grid (WLCG) [63]. This system with fast data transfer ability is a powerful computing tool to process reconstruction data and simulation events throughout the world.

Since the scale of this project is very large, computing facilities are organized in a hierarchical tiered structure. Figure 2.13 gives a schematic overview of the CMS computing grid. The raw data passing through the HLT system is stored at a single Tier-0 computing centre, hosting at CERN. The data recorded by the CMS with the prompt reconstruction is registered here at the Tier-0 centre. One of this permanent mass storage duties is transfer all of collected and simulated data to at least two independent Tier-1 computing centres.

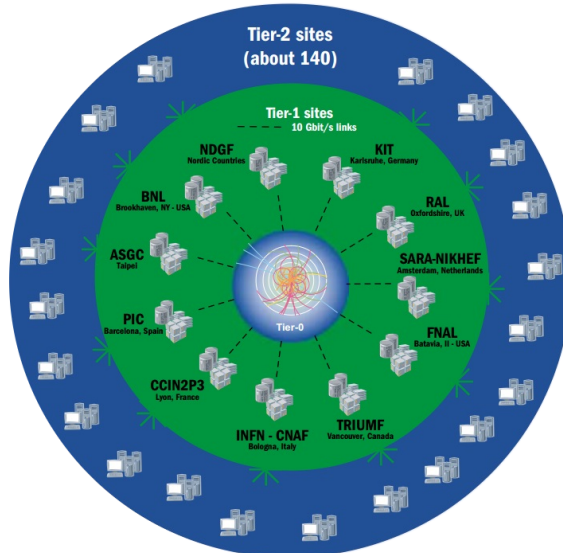


Figure 2.13: Schematic overview of the CMS computing grid.

Totally, seven Tier-1 centres are located around the world which store also the simulated data. Some high priority analysis jobs are run on these sites but the

main tasks are carrying out data reprocessing and storing reconstructed data and simulated samples events in reliable mass storage systems. They also provide very high speed international network links to distribute data to the Tier-2 sites connected to them. Multiple Tier-2 centres are also distributed around the world. These centres store copies of MC samples and data streams to do physics data analysis on them. The Tier-2 centres are accessible for users to submit analysis jobs on them. At the end, Tier-3 centres are local analysis centres which provide limited computing and storage for local groups.

Chapter 3

Sensor R&D for the CMS outer tracker upgrade

The upgrade project of the LHC, known as the High Luminosity-LHC (HL-LHC) [64], is scheduled for 2022. In the HL-LHC phase, LHC has planned to deliver an integrated luminosity of up to 3000 fb^{-1} at an instantaneous luminosity of $5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. From simulations of the radiation environment at the CMS, one expects a maximum particle fluence of $1.5 \times 10^{15} \text{ n}_{\text{eq}}/\text{cm}^2$ for the silicon strip tracker regions (shown in figure 3.1).

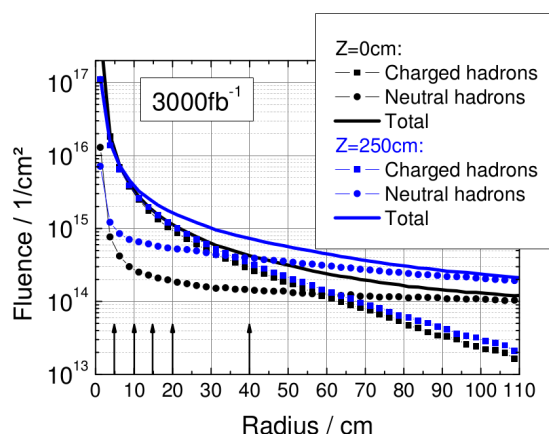


Figure 3.1: Simulated fluences by FLUKA for the CMS tracker, as a function of the radius from the interaction point [65].

Within these conditions, the radiation environment of the tracking system will be severe, required to install a new tracker detector, built with a material able to withstand the radiation rate, and a layout that guarantees a granularity higher than the actual CMS tracker one. To identify the future silicon sensor technology baseline for the phase II upgrade of tracker, CMS started a so called Hamamatsu Photonics K. K. (HPK) measurement campaign. In order to find the best candidate, the CMS Tracker Collaboration has ordered a large number of 6 inch wafers to

the producer HPK to investigate possible sensor technologies and materials. The used wafer layouts with different substrates, thicknesses, implants, and geometries include a variety of devices to investigate different aspects of sensor properties like pad diodes, test-structures, small strip sensors and a strip sensor array with varying strip pitch and strip width.

To find radiation hard materials, an extended mixed irradiation program with protons and neutrons has been set up for the materials of interest. These sensors have been irradiated to various fluences corresponding to the ones expected at different radii in the future tracker, searching for the best material, manufactured by a widely recognised producer. The measurements performed on the structures include electrical sensor characterization, collected charge measurements, and bulk defect characterization. In this part, the current status and results of described campaign will be presented.

3.1 The wafer and materials

The wafer has been packed with a variety of structures, each intended for a different purpose (figure 3.2). A detailed overview of some important parts can be find below [66].

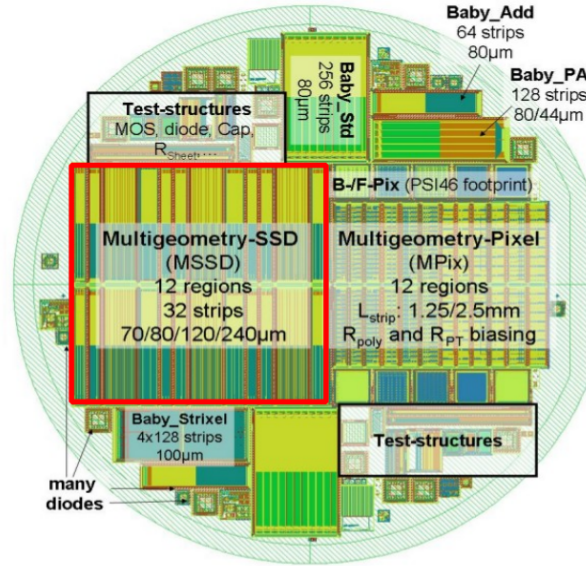


Figure 3.2: The HPK campaign wafer with the different structures.

The material properties and response to irradiation are investigated with different diodes. Dark current value for the detector can be measured by using diode structures. They are used to investigate the material through electric measurements (IV- and CV-curves). There are also test structures to qualify and monitor the quality of the production process as well as determining operating parameters. Several

different measurements such as interstrip parameters (capacitance and resistance) and dielectric break down voltage can be performed with this structure consisting of strips.

In order to save materials in the standard baby sensors, the pitch adapter as connection between sensor and readout system is implemented directly on the sensor. They are produced by converting the about 80 μm pitch of the strips to about 45 μm pitch of the read-out chip pads [67]. The baby structure is used for beam tests and source measurements to evaluate radiation hardness of the different materials. The other sensor with very short strips (1-2 mm), multi-geometry pixel sensor, is used especially for the capacitance measurements, as it leads to conclusions on the noise behaviour.

Our measurements focus on the Multi-Geometry Silicon-Strip Detector (MSSD), which is highlighted with a red square in figure 3.2. This structure is used for test beam measurements in presence of radioactive sources. It contains 12 different regions with different strip pitches and width-to-pitch ratios. Strips are arranged in groups of 32 strips with different pitches from 70 to 240 μm . These kind of sensors could be used in the inner region of the Phase II strip tracker. These sensors allow for the actual assembly of modules and therefore for the investigation of the influence of geometry and materials on resolution, noise and signal-to-noise ratio. In our study, charge collection and a new effect, namely random ghost hits, will be investigated.

This layout with different active thicknesses has been processed by different technologies for the substrate production. Previous work done at CERN [68] has shown that silicon substrates with a high oxygen content show a higher degree of radiation tolerance. In the HPK campaign, three different processes with different oxygen contents for growing silicon sensor silicon are compared. Float Zone (FZ), which is the best known material, serves as a reference with a high quality in terms of the high resistance values. Magnetic Czochralski (MCZ), as a growth technique results in a high oxygen content in the silicon, which has been shown to be beneficial in terms of radiation hardness. Epitaxial (Epi), known as a method allows to produce very thin (25-100 μm) active sensors by chemical vapour deposition of silicon on a carrier wafer.

All described materials are available as p-strip implant in n-bulk material (n-type) and n-strip in p-bulk (p-type) versions. In order to avoid low interstrip resistances due to inversion layers, the p-type sensors require additional isolation which is done with two different techniques. The p-type sensors labelled with the letter Y (p-spray) are produced under a homogeneous p-doping of lower concentration all over the strip side of the sensor. The sensor types named with letter P (p-stop) use a higher doping concentration p+ between the n+ readout strips to intercept the

electron accumulation layer.

Modules with several various thicknesses have been ordered to study the properties of different sensors. The deep diffused FZ materials with a physical thickness of 320 and 200 μm are available. The MCZ material is physically 200 μm thick and the epitaxial wafers come in 50, 75 and 100 μm versions. An advantage of thinner sensors is the lower depletion voltage compared to thicker sensors. Also due to the reduced thickness and a higher electrical field, the trapping of charge carriers is reduced after irradiation for the thinner versions. In our study, labels for the materials are composed of silicon type (FZ, MCZ, FTH: thinned float zone), nominal active thickness (320, 200 μm) and doping type (N: n-type, P: p-type with p-stop, Y: p-type with p-spray).

3.1.1 Irradiation

The radial dependence of neutral and charged hadron fluences differs significantly (figure 3.1), which leads to a variation of the ratio of charged to neutral hadrons as a function of radius. It was found that neutrons, protons and mixed irradiations lead to different effects in different materials. Therefore we need to probe these effects on our materials in a consistent way. Since we cannot survey the conditions in the entire tracker volume, we picked some representative points as summarized in first two lines of table 3.1. The selected fluences reflect the same conditions in the tracker in view of total fluence and mixture of charged and neutral particles.

Table 3.1: Summary of the foreseen particle fluences at $z = 250$ cm. The maximum fluences for the CMS sensors discussed in this paper correspond to $r = 20$ cm.

Radius	Protons, $10^{14} \text{ n}_{\text{eq}}/\text{cm}^2$	Neutrons, $10^{14} \text{ n}_{\text{eq}}/\text{cm}^2$	Total, $10^{14} \text{ n}_{\text{eq}}/\text{cm}^2$	Ratio p/n
40 cm	3	4	7	0.75
20 cm	10	5	15	2.0
15 cm	15	6	21	2.5
10 cm	30	7	37	4.3
5 cm	130	10	140	13

The studied sensors in the HPK campaign are then irradiated under several irradiation steps with protons and neutrons. Irradiations are performed with 1 MeV neutrons from the reactor in Ljubljana, Slovenia, and with 25 MeV protons from the cyclotron in Karlsruhe, Germany. The annealing effects at room temperature due to the transport from the irradiation facilities to the individual institutes, where the measurements are performed, have to be considered.

3.2 Measurement setup and results

In order to understand the sensor properties and their influence on the performance fully, a variety of tests like signal (charge collection) and noise using the CMS readout system have been done. The sensors are mounted on an aluminium plate and equipped with the CMS APV25 readout chips and electronics. For the signal and noise measurements, the MSSD structures with various strip pitch and width ratios in combination with different thicknesses and materials are exposed to particles from a radioactive beta source (^{90}Sr). In the next two sections, the measurements on the varied bias voltage, fluence, annealing time and temperature for different geometries will be reported.

3.2.1 Charge collection

The most important measurements for hit identification is the signal to noise ratio. The irradiated sensors are exposed to the beta source and the collected charges are measured for many times. The distributions of cluster charge are built and fitted with the convoluted Landau and Gaussian functions. The most probable value (MPV) of the fit function is extracted which known as cluster charge in the following figures.

In order to make a high speed off-detector digital links at the HL-LHC, the digital readout chip with a comparator to select "hit" or "no hit" on each strip will be used at CMS. Therefore, it is planned to substitute the existing CMS tracker readout chip (APV25) with the analogue pipeline storage and readout system with the CMS binary chip (CBC) [69]. The CBC noise threshold is expected to be between 4000 and 6000 electrons, thus in this work, the 8000 collected electrons are considered as as signal. In this section, the collected charge values at some fixed applied bias voltage as a function of fluence and annealing time for different materials are presented.

Figure 3.3 shows the degradation of the collected charge for FZ320 μm materials at 600 V and 900 V reverse bias (n-type on the left and p-type on the right). These measurements have been done for different irradiations up to $1.5 \times 10^{15} \text{ n}_{\text{eq}}/\text{cm}^2$, which is the highest level of expected fluence. The initially higher signal of the thick materials drops fast with increasing fluence. To make a comparison between CMS and ATLAS results, the results of ATLAS [70, 71] are also put in the same plot with equivalent doses of neutron and proton irradiations for both p and n-type materials separately.

In general, it has been found that for n-type materials (left side of figure 3.3), CMS measured higher signal than ATLAS. Also there are some differences between vendors and irradiation types (ATLAS results are obtained with only neutron irra-

diations while CMS used mixed ones). Obviously, the n-type strip sensors could be used for fluences below $7 \times 10^{14} \text{ n}_{\text{eq}}/\text{cm}^2$. At the higher level of radiations, a strong decrease in cluster charge below the CBC threshold is observed. Then this kind of sensors would not be useful for higher irradiation scenarios of phase II tracker upgrade.

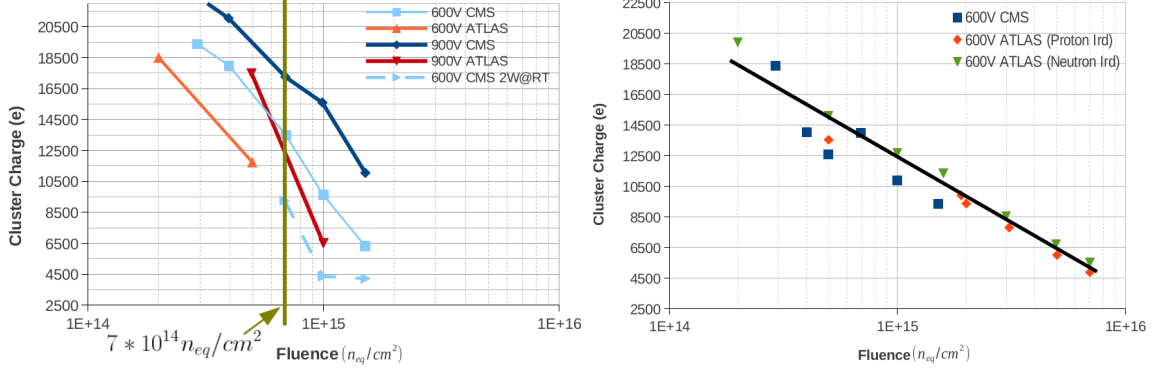


Figure 3.3: The collected charge vs. irradiation fluence for FZ320 μm (n-type on the left and p-type on the right). The ATLAS data were extracted from references [70, 71]. The CMS measurements have been taken at -20°C with a ^{90}Sr beta source by using APV25 read-out system. A cluster has been defined so that a seed and neighbours signals are bigger than three and two times the strip noise, respectively.

For the p-type materials, in the case of FZ320 μm (right side of figure 3.3), the behaviour of collected charge after irradiation is consistent for all thick sensors with a uniform drop trend. The measurements on these strip sensors show a signal higher than 8000 electrons even after high irradiation fluences of $1 \times 10^{15} \text{ n}_{\text{eq}}/\text{cm}^2$. The seed signal is also defined as the sum of collected charge by strips which have signal to noise ratio bigger than three. The left part of figure 3.4 shows the seed signal for all thin materials with 200 μm thickness at fixed bias voltage of 600 V without any annealing.

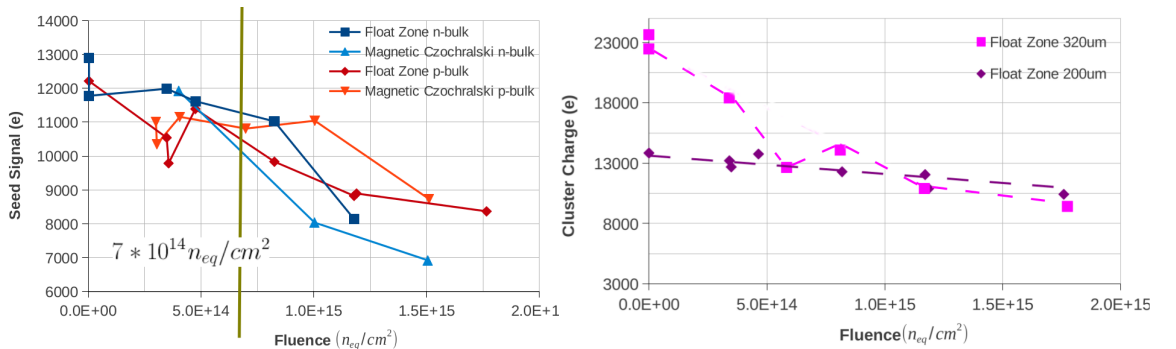


Figure 3.4: Left: Seed signal vs. irradiation fluence for FZ200 μm p and n-type materials. Right: Cluster charge comparison between thick (320 μm) and thin (200 μm) FZ sensors. Both of them are measured at fixed bias voltage of 600 V without any annealing.

It seems that for fluence below $7 \times 10^{14} \text{ n}_{\text{eq}}/\text{cm}^2$, n-type sensors have higher seed signals with respect to p-type ones. Then n-type strip sensors could still be used in this range of fluence. In the case of higher irradiation levels, the collected charge for n-type materials were not high enough to be taken into account as signal. The results of the comparison for the FZ material with different thicknesses (200 and 320 μm) are shown on the right side of figure 3.4. It can be seen that after higher fluence of $1 \times 10^{15} \text{ n}_{\text{eq}}/\text{cm}^2$, the physically thin materials collect comparable or even slightly more charge than the thick sensors. Thus 200 μm thick strip sensors allow to reduce the material budget and keep the charge collection at the same level as 320 μm thick sensors at high fluence.

One of the important parameters to study irradiated sensors is the annealing effect on charge collection. Figure 3.5 shows dependency of the seed signal on annealing time at room temperature (RT), for fixed fluences and different materials and thicknesses. These two plots are related to medium fluence ($7 \times 10^{14} \text{ n}_{\text{eq}}/\text{cm}^2$) and high irradiation level ($1 \times 10^{15} \text{ n}_{\text{eq}}/\text{cm}^2$) for devices that were temperature annealed. For both of them, signal is more stable for p-type sensors than n-type ones. In the case of medium level, all thin p and n-type samples work well and show seed signals higher than 8000 electrons even with high amount of annealing (20 weeks at room temperature). For high irradiation (right side of figure 3.5), signal on n-type sensor decreases after a few weeks at room temperature to a low level. This is one reason why p-type will be the best choice for the outer part of the tracker at the phase II upgrade project. This effect is also most pronounced for MCZ material, which also has a higher signal with annealing than FZ as shown at higher fluences.

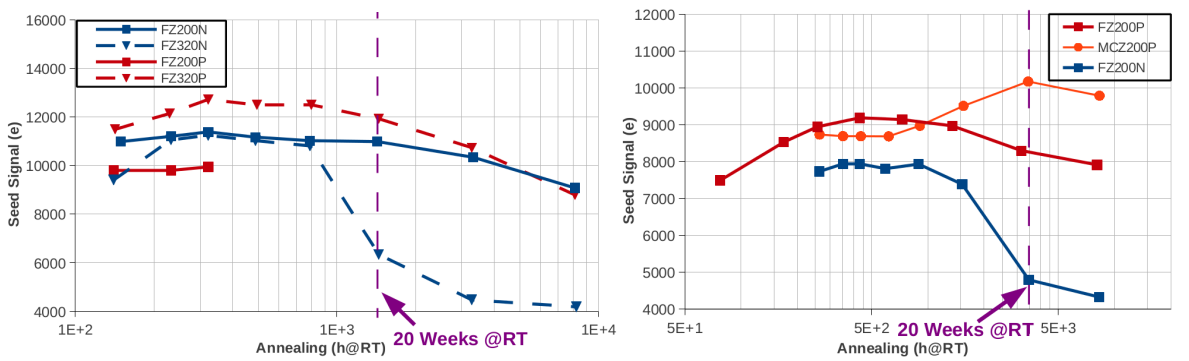


Figure 3.5: Annealing behaviour of seed signal at fixed bias voltage equals to 600 V for two different irradiation levels as a function of annealing time at room temperature. Left: medium fluence ($7 \times 10^{14} \text{ n}_{\text{eq}}/\text{cm}^2$) and Right: high fluence: ($1 \times 10^{15} \text{ n}_{\text{eq}}/\text{cm}^2$).

3.2.2 Study of random ghost hits

High fluences could damage sensors by increasing their leakage current, the depletion voltage and charge carrier trapping which leads to the higher noise and more heat in studied modules. Therefore, it is necessary to study noise behaviours for different materials at higher fluences. It is expected that the noise follows a Gaussian distribution before and after irradiation, but CMS has observed that the noise distributions have large non-Gaussian tails (see figure 3.6). The effects have been observed especially for irradiated n-type sensors and they can lead to strip occupancies of more than 1% fake hits even if no particles are crossing the sensor. By increasing the value of high voltages, significant noise tails are appeared. By taking noise measurements without particles crossing and fitting them with Gaussian function, we can define fake hits which exceed the threshold of 5σ . The real signal is then measured by eliminating these obtained noise hits from signal runs in presence of radioactive beta source.

For fake hit rate computations, two different methods are defined, detailed in figure 3.7. The 5σ cuts are applied on the by fitted noise distributions with Gaussian functions. In the first method, the strip occupancy is defined as the number of signals exceeding 5σ cuts at positive side band divided by the number of strips and the number of events (left plot at figure 3.7). In the second definition, the fake hits are defined as subtraction of signals exceeding 5σ cuts at two side bands divided by the number of strips and the number of events (right plot at figure 3.7). The strip occupancy due to non-Gaussian noise hits depends on bias voltage, fluence and annealing time. To study this effect, a scan in the phase space of these parameters are made.

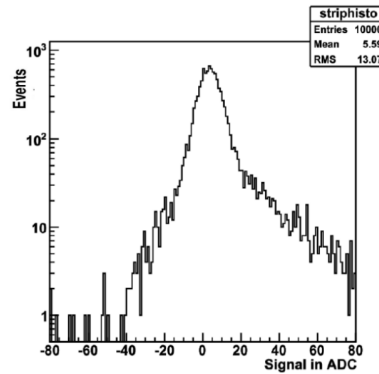


Figure 3.6: Example of pedestal-subtracted noise distribution in irradiated n-type strip sensors at 900 V taken from reference [66].

The strip occupancy due to the noise hits (by using the first method) as a two dimensional function of bias voltage and annealing time is shown in Figure 3.8. From this plot, it can be seen that no effect appeared in p-type sensors even after mixed

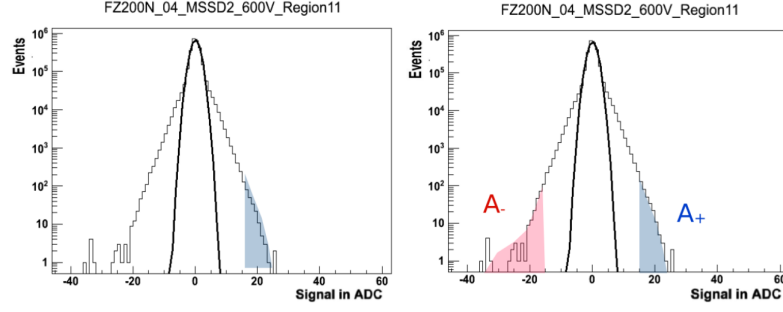


Figure 3.7: Two different definitions of the fake hits rate.

irradiation (top-left part), while these measurements show visible effects for n-type materials for every irradiation type. The rate of strip occupancy is going down, when sensors are annealed for a long time of about 90 days. In case of only neutron irradiated n-type sensors (top-right side), higher bias voltages of about 800 V are needed to produce fake hits. This effect is less pronounced for neutron irradiation compared to irradiation with charged hadrons (see bottom side of figure 3.8). For higher mixed irradiation levels, this effect starts even at low voltages of about 300 V. The different features of charged and neutral hadron irradiations, indicates that ionising energy loss in the SiO_2 interface might be relevant. This dependence on the type of ionizing radiation hints towards a combined effect of bulk damage and surface charge. The simulation on surface and bulk damages can help to study this effect in more detail [72].

Additionally, the contributions of fake hits for different geometries (different strip width and pitches) have been investigated. To study random ghost hits on different geometries, the MSSD structure with 12 regions depicted on the left side of figure 3.9 is used again. The voltage for which the random ghost hits appeared (called turn-on voltage) has been measured for all of the materials in different regions. In this study, the second definition of fake hits rate is assumed. Figure 3.9 shows the turn-on voltage for a variety of materials at different irradiation levels with different strip width and pitches. It can be seen that this effect has not appeared for p-type sensors with 200 μm thicknesses even at higher bias voltage. The other consequence is that regions with high strip pitches are affected earlier.

3.3 Summary of results

To find the best material candidate for the CMS outer tracker at the HL-LHC, signal and noise were studied on sensors produced within the HPK campaign. To do that, sensors were irradiated to fluences expected at radii of 20 and 40 cm. The output of the measurements can be summarized as follows:

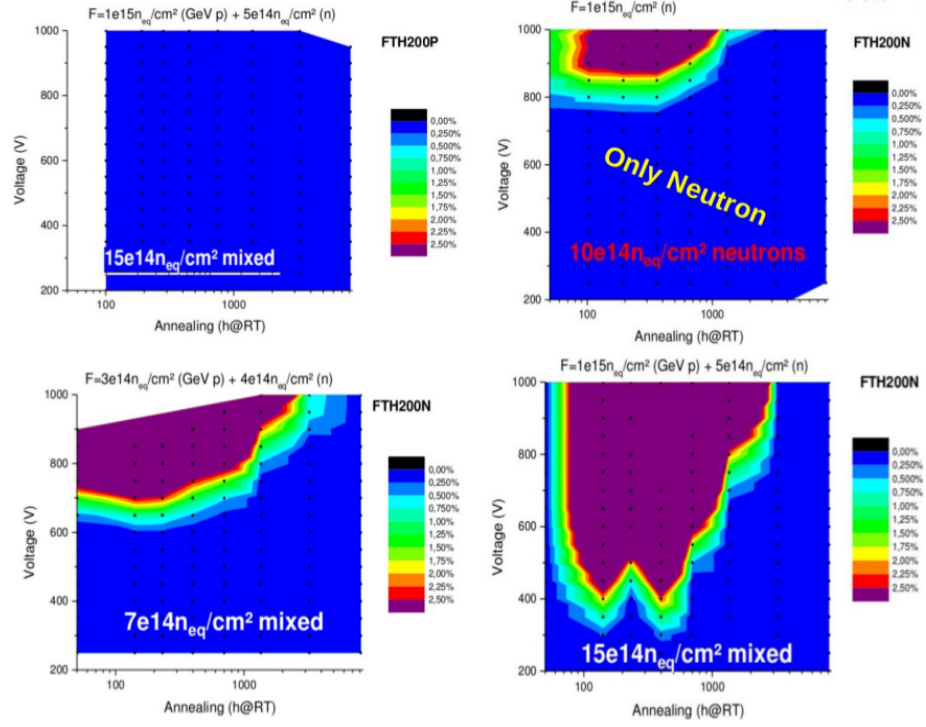


Figure 3.8: Strip occupancy due to random noise hits as a function of bias voltage and annealing time after irradiation (irradiation type and levels have been labelled on each plots). Top-left: p-type with mixed irradiation, Top-right: n-type with only neutron irradiation, Bottom: n-type with two different mixed irradiation levels.

- n-type and p-type materials: For medium irradiation, n-type materials give higher signals, but at higher fluences the p-type sensors collect more charge carriers. While charge collection is rather stable for p-type sensors, signal on n-type materials is decreased after a couple of weeks at room temperature. The new effect of fake hits is appeared for the n-type sensors by increasing the strip occupancy. This new effect could be explained with the simulation of the electric field of the device and needs to be studied further.
- FZ and MCZ: Sensors made of MCZ silicon show an increased signal with annealing. Operation of MCZ sensors would enable annealing scenarios with longer annealing times and hence with lower dark current, less power consumption and lower heat generation.
- 200 and 320 μm thickness: The collected charge was comparable for both of them especially for sensors that will be located at a radius of 20 cm. Thinner strip sensors allow to reduce the material budget and keep the charge collection at the same level as 320 μm thick sensors especially at high fluences.

Finally due to the good charge collection performance of the p-type materials after irradiation with the expected fluences, and since non-Gaussian noise hits were

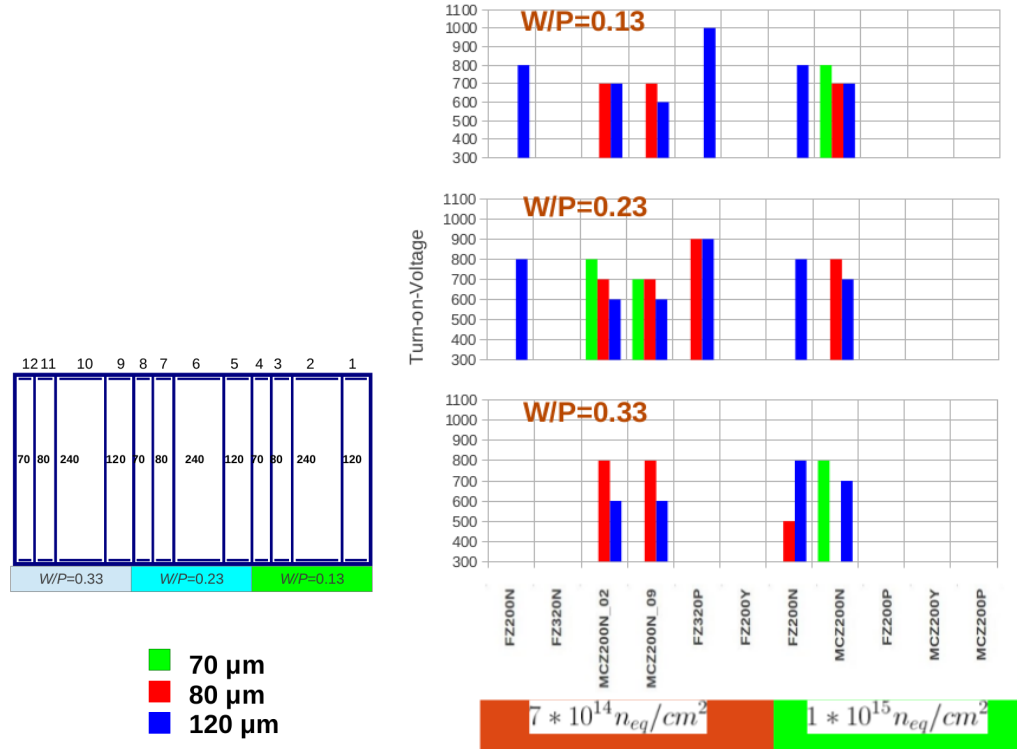


Figure 3.9: Left: Feature of MSSD with 12 regions. Every region has a specific combination of strip width and pitch, and is shown with different color labels. Right: Turn-on-voltage for present materials with different fluence, strip width and width-to-pitch ratios.

not observed for them, CMS decided to use 200 μm p-type sensors for the tracker and concentrate the future work on optimizing the technology and the geometry.

Chapter 4

Generation, simulation and object reconstruction

In proton-proton collisions at the LHC, a huge number of elementary particles are produced due to additional radiation and the process of hadronization and decay. These particles go through different subdetector parts, interact with the detector materials and finally are observed as electronic signals from the read out devices. In order to compare recorded events taken by the CMS experiment with the theoretical expectations from the SM as well as estimate efficiencies and acceptances of analysis algorithms, simulated events from Monte Carlo (MC) generators are used. Simulated events are generated randomly according to the probability of a given process based on the so-called MC methods [73, 74]. The hard process in the proton-proton collisions and the evolution of the generated particles are provided by the MC generators which is given in section 4.1.1. The detector response is simulated using dedicated detector simulations. More information on detector simulation can be found in section 4.1.5.

The detection of produced particles is only the first step of the data taking process. It is essential to translate this raw detector information into physically interpretable quantities. To do that, charged tracks or neutral energy depositions are made and the physical objects like electrons, muons, and jets originating from quarks and gluons are reconstructed. This procedure is carried out by different reconstruction algorithms in the CMS offline software (CMSSW) [75]. The generated and recorded events with the same data formats are reconstructed in the same way. The only difference between them is that in each generated event, the true particle information at the generation level related to each processes of the hard interaction are also accessible. In section 4.3 and followings, the reconstruction of physical objects are explained with details.

4.1 Event generation chain

As explained before, simulated events are produced by the MC generators which use random numbers according to the desired probability distributions at various stages of the generation process. The hard parton interaction in the $pp \rightarrow X$ collisions is the complex process can be subdivided into a number of sequential steps, roughly corresponding to increasing scales of distance and time. These simulation steps are illustrated in Figure 4.1 and are summarised in the following sections.

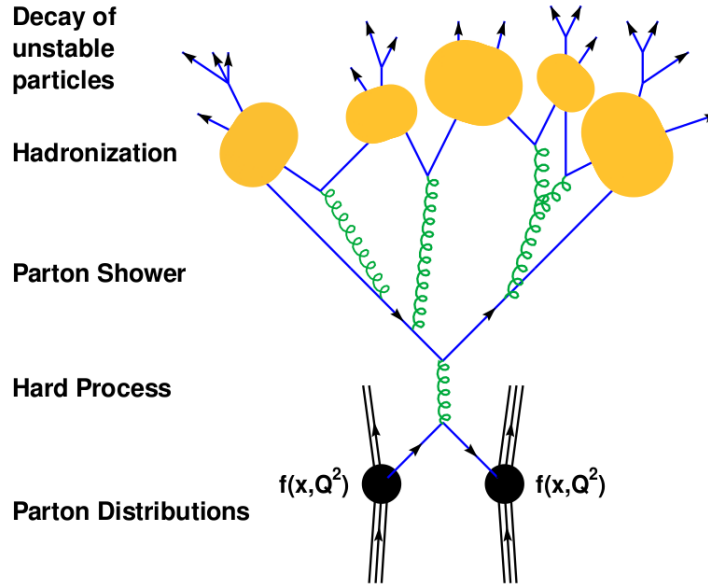


Figure 4.1: Illustration of different steps of the MC event generation processes, starting from the initial interaction between the protons down to the decays of unstable particles in the experiment. This plot is given from [76].

4.1.1 Hard process generation

The interaction between two incoming protons is often soft leading to elastic collisions which are not interesting in the framework of this research. More interesting ones are the hard interactions where the desired final state is produced from the initial partons as described before in equation 1.38. The hard parton interaction is calculated with the MC generators with respect to the corresponding matrix elements of the processes and the probability density of the phase space. In the case of hard process, since the strong coupling constant α_s is small for large Q^2 , the cross section can be calculated with the help of perturbation theory. As a consequence,

this calculation is done by using perturbation matrix element (ME) method and the integration over the PDFs to address all possible initial partonic states.

4.1.2 Parton shower

The incoming and outgoing partons undergo branching, giving rise to so called parton showers. Radiation produced by the incoming partons is called Initial State Radiation (ISR) while the latter is known as Final State Radiation (FSR). These additional radiations lead to possibly large corrections to the event topology of the hard process. Two different methods, matrix element and parton shower, are used to generate such radiations. In the first method, the matrix elements of hard processes with these additional radiations are calculated based on the perturbation theory. As dedicated before, this method is well suited to simulate particle interactions with high momentum transfer, where QCD coupling α_s is small, and then only diagrams corresponding to radiation of hard gluons are taken into account.

In the parton shower method, higher order corrections are approximated by combining an arbitrary number of branchings of single partons into two or more other partons. Cascade of successive emissions of quarks and gluons in each branch are stopped when the energy scale drops below a lower energy scale where $\alpha_s \approx 1$ and the non-perturbative regime begins. The evolution of the parton shower can be described mathematically by the DGLAP QCD evolution equations [77–79]. Also the probability for a gluon radiation is calculated via the Altarelli-Parisi splitting functions. When both methods are applied to simulate events, to avoid double counting of parton emission in the overlapping phase space, an appropriate matching is performed. After the parton shower, the final state consists of colored partons which then evolve into colorless hadrons.

4.1.3 Hadronization

As discussed in section 1.2.3, at short distance the strong interaction can be well described by the QCD perturbation theory. At large distance, the strong coupling becomes very large, thus confinement effect occurs. Since there is no theoretical framework to explain the confinement mechanism, phenomenological hadronization models are then employed to describe how the partons give rise to color neutral baryons and mesons. The two successive models to explain hadronization are the Lund string model [80] and cluster fragmentation [74, 81]. Hadronization is the combination of fragmentation, and the subsequent decay of the unstable hadrons into other hadrons. In the Lund string model, the color field between final state quarks and anti-quarks is modelled as a color flux string. The potential energy stored in the field between them grows linearly with the separation between quark

and anti-quark. When this potential energy exceeds the $q\bar{q}$ production threshold, then the string can break resulting in the creation of a new $q\bar{q}$ string. This process repeats itself until the energy is lower than the threshold for quark pair production.

In the second hadronization algorithm, the cluster fragmentation model, a special property of QCD perturbation theory, the so-called preconfinement [82] is being used. Here all secondary gluons produced during the showering are split non-perturbatively into quark anti-quark pairs resulting in color-singlet clusters. Then these clusters are considered as the basic units, and depending on their invariant mass, can form neutral hadrons. At the end of the hadronization process, decay modelling of the unstable colorless mesons and baryons is simulated randomly by using the known decay branching ratio of the hadrons.

4.1.4 Underlying event and pileup

As described in section 4.1.1, only one parton from each proton contributes to the hard interaction. Since two partons undergoing the hard interaction carry a color charge, the remainder of the incoming protons is also color charged and will be subjected to hadronization as well. In addition, partons in these beam remnants can interact with each other, producing multi-parton interactions. Both the hadronization of the beam remnant and the multi-parton interactions are labeled as the underlying events [73, 83]. This process is simulated properly by the phenomenological models and implemented in the MC simulation. Furthermore, in the case of high instantaneous luminosities, multiple hadron collisions can happen during one bunch crossing. The influence of these events, called pileup, is taken into account in the MC event simulation.

4.1.5 Detector simulation

Most hadrons produced in the hadronization step are unstable and subsequently decay into observable stable particles on the basis of decay modes with related branching ratios. The interaction between decay products with detector material is simulated and implemented within the CMSSW package [75]. This simulation is based on the GEANT4 simulation toolkit [84] which includes the full geometry of the detector and its materials, precise map of the magnetic field, and an approximation to the electronic readout. As the final step, generated events from proton-proton collisions are passed to the described simulation program for the CMS detector resulting in the ultimately measured digital signals. These readout electronic signals should be similar to the output of the real detector which can then be used in the original event reconstruction algorithms.

4.2 Monte Carlo generators

For a full event generation with all steps from the hard scattering up to the decay, different types of MC generators are used. Firstly, the matrix element event generators like MadGraph/MadEvent [85] use perturbative QCD calculations to simulate ISR and FSR in LO matrix elements without hadronization modelling. The second group of MC generators like Pythia [86] and Herwig [87] can model the parton showering and hadronization very well. The former one performs simulation of hard interaction while the latter gives a good description of soft parton emissions. The most common matching between these two simulation steps, are the MLM scheme [88] and the CKKW [89] algorithm. In the following, some different MC event generators used in this analysis are briefly introduced.

4.2.1 MadGraph and MadEvent

MadGraph is a matrix element generator that can produce the leading order Feynman diagrams for desired process and compute the squared amplitudes of the related diagrams. In addition to determine the amplitude for a given process, MadGraph also produces the mapping for the phase space integration. Since the generation of all diagrams is CPU intensive, MadGraph can manage processes with a limited number of final state partons. Then MadEvent event generator processes the information produced by the MadGraph, calculates cross sections and finally generates MC events of hard interaction. As already mentioned, this event generator does not produce any particle showering, therefore the output has to be passed to a shower and hadronization generator like Pythia or Herwig.

4.2.2 PowHeg and MC@NLO

In contrast to the MadGraph approach, the Positive Weight Hardest Emission Generator (PowHeg) [90] and the MC@NLO [91] generators provide matrix element calculations with NLO precision. Since these two generators with NLO calculations only calculate diagrams up to one additional parton, these generators establish interfacing with a parton shower program to generate additional radiations. The main approach of these programs is to generate the hardest radiation first and then feed the generated event to a shower and hadronization generator for softer radiations. The difference between them lies in the way that these theoretical NLO calculations are matched to the parton showers.

4.2.3 Pythia and Herwig

These generators are multi-purpose MC event generator which describes very well possible particle collisions at high energies. They contain wide range of processes from hard scattering to soft ISR/FSR emission generation as well as many models beyond the SM in own packages. They also include initial and final state parton showers, contributions of underlying events, hadronization modelling and decay processes. In Pythia, the hadronization process is implemented following the Lund string model where the linear confinement in QCD is the main concept to model fragmentation.

In contrast to Pythia, Herwig uses cluster fragmentation instead of the string model for jet hadronization via non-perturbative gluon splitting into quark-antiquark pairs and grouping all color connected quarks into clusters. In this analysis, most of the MC simulated samples are showered with the Pythia approach, while some other samples are produced with the Herwig generator.

4.3 Physics object reconstruction

The data format of the CMS detector readout for simulation and real data can not be easily used in physics analysis, then it is essential to translate the readout electronic signals from all subdetectors to physical objects with kinematics required in data analysis. This is accomplished by different steps of reconstruction algorithm embedded in the CMSSW framework aiming at a full reconstruction of all particles from the hard scattering event. Firstly, tracks of charged particles are constructed by converting the electronic signals into hits in the tracking system and clusters of energy deposition in the calorimeters. In a second step, these reconstructed tracks are used to identify physics object candidates.

The object reconstruction in the CMS experiment is described in this section. The most important physics objects to reconstruct the $t\bar{t}$ events in dilepton channel are electrons, muons, jets, and neutrinos. After a brief description of the tracks structure in section 4.3.1, the particle flow (PF) [92] algorithm which is standard muon and electron reconstruction approach will be explained in section 4.3.3. The identification of electrons and muons are presented in sections 4.3.4 and 4.3.5, respectively. The reconstruction of jets from hadrons is discussed in section 4.3.6, while section 4.3.7 is devoted to the b jets identification in the CMS. The other quantity to explain weakly interaction neutrinos, the Missing Transverse Energy E_T^{miss} , will be introduced in section 4.3.8.

4.3.1 Reconstruction of tracks

In order to reconstruct the trajectories of the produced charged particles in the collisions, CMS provides a very precise tracking system. Charged particles go through the tracking volume and produce some hits in the tracking devices along their trajectories. In the CMS detector, the trajectories of these particles are curved due to the magnetic field and the Lorentz force. The curvature of the trajectory is related to the electric charge and the momentum of the particle, so measurement on the charged particle tracks is essential for the precision measurement of its momentum. To reconstruct this trajectory, the positions of isolated hits have to be combined correctly.

There are several methods for track reconstruction by fitting the positions of hits in the tracker layers. In the CMS, tracks are reconstructed with the Combined Track Finder (CTF) method [93], by using a Kalman Filter (KF) [94]. In this method, tracks are constructed from the inner pixel layers to the outer layers of the silicon strip detector. The track reconstruction in an event is done through multiple iterations. Firstly, seeds as track candidates are generated from a group of two or three hits in the inner most tracker layers. These hits are used to construct the trajectories layer by layer using the KF algorithm. In order to find additional tracker hits associated with the track candidate, the curved paths of the charged particle in the magnetic field, the momentum loss in the tracker material and also multiple scattering are taken into account in this algorithm. For each layer, new trajectory candidate is created with respect to the updated parameters for corresponding hit in the layer. The procedure of fitting all trajectory candidates is repeated over layers until the outermost layer is reached.

In case no hit is found for two consecutive layers, the trajectory candidate is rejected. Since for each seed there might be several hits in the subsequent layers, a large number of trajectory candidates are present. To avoid any overlaps between different candidates, the trajectories are compared with each other and the best one is selected based on the number of common hits and the track quality. Firstly the ones with the worst fit result is removed from the list of comparison. In the case of two tracks sharing at least half of the hits, the algorithm removed one with the lower number of hits or if the number of hits is equal, one with the lower χ^2 value of the fit is rejected. Finally, in order to improve the precision of the track parameters, an additional global fit is performed by using all measurement points associated to the track. The whole sequence from local to global fitting is repeated iteratively for many times. At each iteration, the used hits in the last iteration are removed, different track selection criteria are applied and the algorithm runs again to find the best track candidate.

4.3.2 Vertex reconstruction

By reconstruction of track trajectories, we are able to measure the location of vertices. In order to define characteristic of each event, exact position determination of the primary interaction point of the hard elementary process is essential. Generally a primary vertex (PV) is the interaction point of the hard scattering process while secondary vertices usually indicate decay points of long lived particles. The main goal of the primary vertex reconstruction algorithm is finding a common origin of reconstructed tracks and performing precise measurements on the founded vertex parameters. In the CMS, primary vertices are reconstructed by fitting the position of clustering tracks which have a same interaction point.

To make accurate search of the PV, described tracks have to fulfil some quality requirements contain number of hits in the pixel and silicon strip detector as well as cut on the normalized χ^2 value of the track fit. In order to fit the position of the vertex, each track receives a weight which depends on its compatibility χ^2 with the vertex candidate. Like track reconstruction, the fit procedure is performed iteratively, so after each iteration a set of weight values are obtained. In a last step, the list of primary vertices is sorted by the sum of weighted p_T^2 of all associated tracks. For pileup events with multiple primary vertices, one with the highest sum of weighted p_T^2 is considered as the selected PV.

4.3.3 Particle Flow algorithm

Now, there is capability to full reconstruct single particles by combining information of reconstructed tracks and vertices, energy deposits in the calorimeters and signals from the muon chambers. For the identification and reconstruction of stable particles like electrons, muons, jets as well as missing transverse energy, the PF algorithm is used. The main idea behind the PF method is to combine all the information of all CMS subdetectors in order to improve quality of object reconstruction. This leads to estimate much more precisely the momentum of a particle since the information from multiple subdetectors is combined.

In general, the PF algorithm consists of two main steps. At the beginning, the combination of fundamental charged particle tracks, deposited energy in the calorimeters and hits in the muon chambers are grouped into blocks by the linking algorithm. Then these linked blocks are identified as different kinds of particles produced during proton-proton collisions. This algorithm gains a high efficiency and a low fake rate particle identification by improving the performance for jets, E_T^{miss} and the isolation of charged leptons compared to the generic object reconstruction. In the following, the reconstruction of main objects in this analysis, including electrons, muons, jets and E_T^{miss} , is described in more details.

4.3.4 Reconstruction of electrons

Electron reconstruction is based on the hits in the tracker system and energy deposits in the ECAL calorimeter. Electrons produced at the interaction point traverse about two radiation lengths of material in the silicon layers of the pixel and the silicon strip detectors before they reach the ECAL system. Due to their comparatively low masses and in the presence of a strong magnetic field, they lose significant amount of their energy by emitting bremsstrahlung photons. The path of electrons is curved in the tracker region hence the radiated photons reach the ECAL along the straight line tangent to the electron tracks leaving a spread of energy in the azimuthal direction ϕ . This energy loss affects the path of electron tracks and then has to be taken into account during electron reconstruction. As already mentioned in section 4.3.1, the object reconstruction is started by looking for related track seeds as initial point of track reconstruction.

Two complementary algorithms at the track seeding stage are used in the electron reconstruction: the ECAL-driven algorithm [95] suitable for isolated high p_T electrons and the tracker-driven algorithm [92] which is optimised for low p_T ones and also those inside jets. The energy of the electron is deposited into a cluster of ECAL crystals, while radiated photons are spread in the region with narrow width in the η coordinate and much larger in ϕ , all together form superclusters. The ECAL-driven algorithm starts by searching these superclusters which are constructed by summing energy deposits of the electrons. After that, the positions of these superclusters are propagated backwards towards the most inner layer of the pixel detector, used two hits in the tracker region and construct the seeds for possible electron tracks.

Unlike the ECAL-driven method, the tracker-driven seeding algorithm starts from the reconstructed tracks in the inner layers of tracking system which are compatible with the electron particle hypothesis. By extrapolating the selected electron tracks from tracker layers to the ECAL, the energy deposits of bremsstrahlung photons as clusters are estimated. Finally, if the ratio between the energy of the cluster and the track momentum is close to unity, then related track is used as seed for the electron reconstruction.

The energy loss distribution of a charged particle traversing a thin layer of material is Gaussian while in presence of bremsstrahlung photon emission, the shape is varied and well described by a Gaussian mixture. To take into account this effect, a non-linear generalization of the Kalman filter, Gaussian Sum Filter (GSF) [96] algorithm has been used. In the GSF method, the energy loss distribution is described by a weighted sum of Gaussian distributions rather than a single Gaussian one. After seed determination, the electron tracks are reconstructed using the hits

in the tracker and modelling the energy loss of electrons by fitting with the GSF algorithm. In this method, a straight line in the direction of the GSF track is extrapolated until the outermost tracker layer is reached. In the case of linking between this straight line with the ECAL cluster, the amount of cluster energy is taken into account as the total electron energy.

4.3.5 Reconstruction of muons

Muons as minimum ionizing particles traverse through the whole detector, lose only little of their energy in the calorimeters and producing hits in the different muon detectors. Due to this unique property, muon is one of the objects which reconstruct easier with high precision. Muons can be reconstructed both as tracks in the tracking system and hits in the four layers of the muon chambers [97]. Three different ways, based on independent reconstruction of tracks in the tracking system and the muon chambers, have been developed in CMS to identify and reconstruct muon candidates.

- **Standalone muons:** Muons which are reconstructed using only the information on the hits in the muon subdetectors. To reconstruct such muons, firstly hits within each DT and CSC chambers construct track segments. These track segments from the innermost chambers of muon system generate seeds which are starting point of muon track reconstruction. As already explained in section 4.3.1, the Kalman Filter technique is used here to propagate these tracks outwards to the layer surface of the next chamber. The track segments in the outer layer are compared with the extrapolated muon trajectory until the outermost chamber is reached. In order to estimate track parameters, the same Kalman Filter is applied in backward direction, working from outside to inner chambers.
- **Global muons:** Muons that are reconstructed by combined information of muon and tracking systems. This most common technique starts with the defined standalone muons, matches these tracks to the trajectories of muon candidates in the tracker and combines them in a fit procedure to construct a global muon track. To do the matching, a rectangular $\eta - \phi$ region around the stand-alone track is defined and select the tracker track which has the best compatibility with the standalone candidate in position and momentum. The energy loss and the multiple scattering in the calorimeter materials are also taken into account in this matching procedure. In the case of finding matched muon tracks, the global muon track is fitted based on the Kalman filter technique by using the hit information from both of tracker and standalone tracks. Details of the matching and fit approach can be found in [64].

- **Tracker muons:** Muons which are reconstructed using only tracker information. Instead of the outside-in method applied in the global muon reconstruction, the inside-out approach in the reverse flow is also used by matching tracker tracks to at least one segment of hits in the muon chambers. In this approach, reconstructed tracks in the tracking system with the transverse momentum of $p_T > 2.5 \text{ GeV}$ are extrapolate to the muon system with similar method used to reconstruct standalone muons. Finally, a tracker muon is identified if there is at least one finding segment in the muon chambers which matches with the extrapolated track without any fit procedure between the subdetectors.

Reconstructed muons with low p_T have a better efficiency in the tracker muon approach, while for high energy muon candidates with $p_T > 200 \text{ GeV}$, the global muon method leads to higher momentum resolution. Totally, by using some data samples, the muon identification and reconstruction efficiency has been obtained larger than 98% for muons with $p_T > 5 \text{ GeV}$ [97].

4.3.6 Reconstruction of jets

As was described in section 1.2.3, due to the QCD confinement effect, color partons can not be observed as free particles in collisions but they can hadronize into streams of colorless hadrons known as jets. These jets can be detected in the form of energy deposits in the calorimetry systems, so that total energy of them is determined as sum of deposited energy in the ECAL and the HCAL subdetectors. Since jets are collections of resulting particles from the hadronization, then to reconstruct them first of all it is necessary to identify all of their constituents. In this analysis, this is accomplished with good precision by the PF algorithm. Then to construct jets from combination of constituent particles, a jet clustering algorithm is applied. In these algorithms, jets can be considered as collections of physical objects like simulated partons, generated stable particles after hadronization and decay or detector objects like calorimeter entries.

In order to have compatible results with theory calculations, each jet cluster algorithm has to be passed some requirements such as infrared safety and collinear safety. The output of jet clustering algorithm is infrared safe, if the number of constructed jets is stable against additional soft radiation. The collinear safety is achieved if the algorithm output does not change under the splitting of one particle into two collinear particles with the same four momentum vector in total. There is a large variety of jet clustering algorithms which are subdivided into two main categories, the cone algorithms [98] and the sequential clustering algorithms like the k_T [99], the Cambridge/Aachen [100] and the anti- k_T [101]. In the CMS experiment as well as this analysis, the anti- k_T as a special case of the sequential clustering

algorithms is used.

Sequential clustering algorithms

Successive clustering algorithms work by defining a cluster of particles based on measured distances between any pair of entities and also beam distance for each object. To do that, distance between two objects i and j , d_{ij} , as well as distance between object i and the beam axis, d_{iB} are defined as:

$$\begin{aligned} d_{ij} &= \min \left\{ k_{T_i}^{2p}, k_{T_j}^{2p} \right\} \frac{\Delta_{ij}^2}{R^2}, \\ d_{iB} &= k_{T_i}^{2p}, \end{aligned} \tag{4.1}$$

where k_T is the transverse momentum of objects, $\Delta_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$ is the distance between two objects i and j in the $\eta - \phi$ plane and R is a radius parameter defining a characteristic size of the jet. The parameter p in this formula determines three different algorithms, when this parameter equals to $+1$, the k_T algorithm is reproduced, the $p = 0$ corresponds to the Cambridge/Aachen algorithm, and finally the case of $p = -1$ gives to the anti- k_T one.

At the first step, all the distances described above are calculated. If the minimal value of all found distances is between two objects, they are merged to a new object, while the original objects are removed from the list of jet's constituents. This procedure is repeated until the smallest distance in the event becomes a distance between one object and the beam. In this case, the corresponding object is considered as a reconstructed jets and is removed from the list of entities. When no more objects are remained, this process is finished. These algorithms comply the infrared and the collinear safety requirements. Also the anti- k_T algorithm as a special case, differ from other algorithms due to producing conical jets around entities with the largest transverse momentum.

In order to obtain four momenta of the jets in this analysis, the recombination scheme E is used. In the E scheme, four momentum of each jet is calculated by the combination of the four momenta of it's constituents. However, the measured four momenta of reconstructed jets are differ from the obtained from the generated jets on truth level. The main sources of this difference are the non linear calorimeter response to particles, the additional energy deposits caused by particles outside the jet cone or pileup events and the detector imperfections like it's limited energy resolution and small dead channels without any signal. To correct these effects, the Jet Energy Scale (JES) corrections [102] are applied after the jet reconstruction.

Jet energy corrections

In order to correct jets energy due to the non-linearity response of the calorimeters, a series of Jet energy corrections (JECs) have been developed in the CMS. The detector response is estimated by taking information of jet transverse momentum distributions at the generator and reconstruction levels. The CMS experiment uses factorized corrections into sequential steps, where each of them corrects for a specific set of effects. Each of these corrections is a scale factor of the jet four momentum which depends on jet related quantities such as p_T , η and flavour. A set of JECs used within this analysis is illustrated in Figure 4.2.

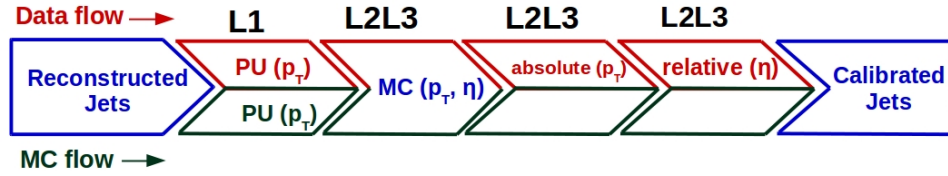


Figure 4.2: Illustration of all steps in the CMS jet energy correction scheme.

The main corrections can be categorized into following steps:

- **L1FastJet:** This correction is applied to remove additional measured energy which does not belong to the hard process and comes from pileup events. Such energy contributions increase the measured jet energies which have to be estimated and subtracted from the jet energy. In order to correct this effect, the FastJet [103] corrections are applied. To do that, a large number of soft ghost particles are injected into the event without changing properties of the hard jets. Each updated jets is reconstructed by using the anti- k_T algorithm. Then the average of p_T density per effective area in the event, ρ , can be determined from the median of p_T^i/A^i where i loops over all jets in the event. Now, the corrected energy of each jet can be calculated by subtracting the part of the underlying hadronic activity corresponding to the jet size.
- **L2Relative:** The inhomogeneous calorimeter layouts in the η direction lead to non-linear η -dependent response of the calorimeters. The aim of this correction is to ensure a homogeneous response of the calorimeter systems over the entire angular region. It is achieved by applying correction factor as a function of η on the jets so that energy response of jets at any η value perform the same relative to a jet in the homogeneous $|\eta| < 1.3$ region.
- **L3Absolute:** Like previous correction, the response of calorimeters is also non-linear with respect to the transverse momentum of the jets. The goal of

absolute JEC is to transform the measured jet energy into the energy corresponding to the true value of particle jet. This correction is p_T -dependent which designed to make a uniform calorimeter response with respect to the p_T value of the jet. These L2Relative and L3Absolute scale factors are derived from simulation events independent of selected criteria but applied on both of simulated and data events.

- **L2L3Residual:** As already mentioned, the last two corrections are purely determined from the simulation and applied on the data and simulation simultaneously. However since these corrections are completely based on simulation and the data does not behave exactly like the simulation, therefore additional data-driven corrections are needed to be applied to the data. These corrections are called L2L3Residual which correct for remaining differences between data and simulation in the η (L2) and p_T (L3) phase space.

In addition to described correction factors, some other optional corrections also exist. These corrections include making the jet response uniform with respect to the electromagnetic energy fraction, taking into account response dependency on the jet flavour and effects due to the underlying event, and also parton jet corrections which try to correct back to the parton level. However, in this analysis, those additional corrections have not been applied.

Jet energy resolution

In addition to corrections introduced in the previous part, the energy or transverse momentum of reconstructed jets have to be corrected based on the fact that the resolution of the jet transverse momentum is worse in data than in simulation. The ratio of the measured resolution in data to the resolution in simulation can be determined using the p_T balancing in QCD dijet events [102]. The simulated events are corrected with this η -dependent scale factors to make a better matching with the data ones. To do that, the energy and momenta of reconstructed PF jets are compared with the corresponding generator level jets, the smearing scale factors are extracted and applied on the energy and momenta of these PF jets.

4.3.7 b-tagging

Since each $t\bar{t}$ event has two b-quarks in the final state, identifying the b jets is a powerful tool to select signal events among multiple background processes. An efficient separation between b and non-b jets plays a crucial role in the event reconstruction for many analysis which contain $t\bar{t}$ in the final state, as well as this analysis. As written before, the outgoing b-quarks hadronize into B hadrons before decaying to

light quarks. These hadrons decay to lighter ones via the electroweak interaction, with the relatively large lifetimes $\tau_B \approx \mathcal{O}(\text{ps})$. Therefore they travel typically a few hundreds of micrometres from the primary vertex in the detector before decaying into lighter particles. This gives the possibility to tag the b jet with the help of displaced tracks which are not originating from the primary collision vertex but from a secondary displaced vertex. A schematic overview of a b jet with the related displaced tracks is shown in Figure 4.3.

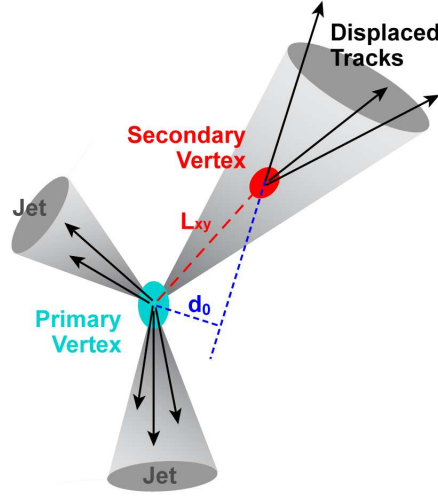


Figure 4.3: The decay of a bound state b-hadron at a secondary vertex after passing the flight distance of L_{xy} in the transverse plane, taken from [104]. The displaced tracks have a large impact parameter d_0 with respect to the primary vertex.

The displaced tracks are measured in the detector, extrapolated back and formed a secondary vertex. Finally the b jets are identified by using reconstructed secondary vertices and related track information. There is several b-tagging algorithms within the CMS software, to tag the b jets from other types of jets [105]. All of these methods trying to keep the b jet identification efficiency high while the mistag rate (the probability to tag a light parton as a b jet) should remain as low as possible. One group of them such as Track Counting High Efficiency (TCHE) and Track Counting High Purity (TCHP), exploits the impact parameter significance, d_0/σ_{d_0} , as a discriminating variable between b jets and light jets. Other group like Simple Secondary Vertex (SSV) and Combined Secondary Vertex (CSV) algorithms gives the possibility to tag the b jet by looking for the presence of a secondary vertex.

The b-tagging algorithm used in this analysis, is the CSV algorithm which combines information from the secondary vertex and track impact parameter to give the best possible discriminator. As a first step of this algorithm, the secondary vertex is reconstructed as a common intersection point of associated tracks to the

related jet. In the next step, some quantities like the flight distance significance in the transverse plane ($L_{xy}/\sigma_{L_{xy}}$), the mass of the secondary vertex and number of tracks associated to the vertex are combined in the MVA approach to form a single discriminating variable. While the SSV tagger can only tag jets with vertices, jets with no secondary vertex can still be constructed in the CSV algorithm by creating "pseudo vertex" based on tracks with $d_0/\sigma_{d_0} > 2.0$. Even in the case of not finding "pseudo vertex", the CSV algorithm works with the help of track-based variables.

CMS supports three working points for this algorithm as loose (L), medium (M) and tight (T) with different b-tagging efficiencies and mistag rates. Generally, the b-tagging efficiencies are about 85%, 65% and 50% with corresponding mistag rates of 10%, 1% and 0.1% for loose, medium and tight working points, respectively.

These efficiencies are different for data and simulation, then in order to correct the differences, the extracted b-tag efficiency scale factor has to be applied to simulated samples. This scale factor is derived as a ratio of b-tagging efficiencies based on the p_T for data and simulation events. The used method in this analysis to extract b-tag efficiencies and scale factor is explained in section 5.3.3 In Figure 4.4, the b-tagging efficiencies for data and MC as well as b-tag efficiency scale factor as a function of b-tag discriminator are shown.

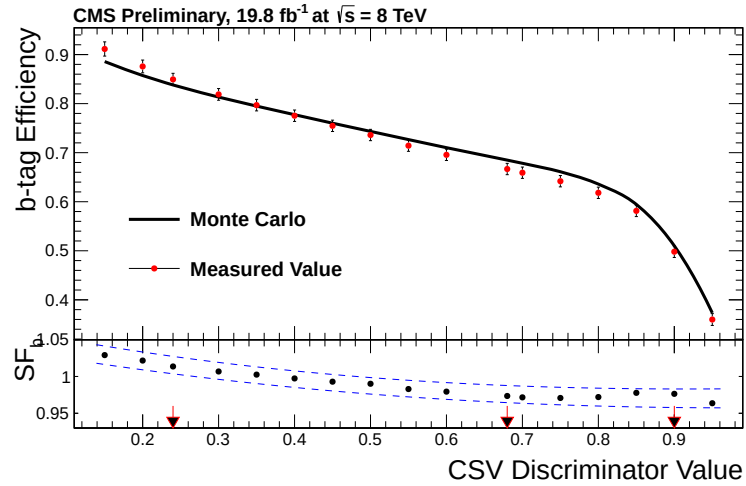


Figure 4.4: b-tag efficiencies in data and simulation as well as related scale factor as a function of the discriminator threshold for the CSV algorithm taken from [106]. The arrows indicate the standard operating working points.

4.3.8 Missing transverse energy

As mentioned earlier, in the case of dileptonic decay of $t\bar{t}$ system, each of the W boson decays into a lepton and undetectable corresponding neutrino. While these

neutrinos escape from the detector without any sensible interaction with matter, they can be detected indirectly for each event as a total missing transverse energy.

In collider machines like the LHC, flight directions of the colliding protons are parallel to the beam axis in the interaction point. As partons carrying only a fraction of the protons momenta, momentum conservation forces only the transverse part of momentum to be zero in each event, while the third component is not under any constraint. Due to momentum conservation in the transverse plane, any imbalance in the vector sum of all measured transverse energy is called the missing transverse energy, which used to infer the presence of neutrinos. Then the missing transverse energy is defined by taking the negative vectorial sum of all reconstructed particles in the event, given by:

$$\vec{E}_T^{raw} = - \sum_{i=1}^N (E_i \sin \theta_i \cos \phi_i \hat{x} + E_i \sin \theta_i \sin \phi_i \hat{y}), \quad (4.2)$$

where the sum runs over all of reconstructed objects and E_i is the deposit energy in the calorimeter tower for i th of particle. Once all of the objects are reconstructed by the PF algorithm, the resulting E_T^{miss} is also calculated as PF object.

In order to get a true value at the end, several corrections are applied during its computation [107]. For instance, in the above equation, uncorrected jets are used to calculate E_T^{miss} , while it seems necessary to propagate the effect of described JEC to the missing transverse energy. To avoid large JEC uncertainties for low p_T jets, this correction is only applied on jets with $p_T > 10$ GeV. The effect of this correction can be written as:

$$\vec{E}_T^{corr} = \vec{E}_T^{raw} - \sum_{i=1}^{N_{jets}} (p_{T,i}^{corr} - p_{T,i}^{raw}). \quad (4.3)$$

An other correction is related to the muons, since they deposit only a small amount of their energy in the calorimeters, the p_T values after subtracting little energy loss contribution, have to be added to the summation in equation 4.2. To make precise measurement on electron energy in the ECAL environment, electron contribution in the missing transverse energy computation, does not need to be corrected. The energy correction coming from jets and muons are referred to as Type-I correction.

In addition to the Type-I correction, there is also some other corrections from undetected particles escaping the detector close to the beamline, detector malfunctions, pileup and underlying events. Optional Type-II correction takes into account effects from pileup and underlying event but it is not recommended to use for the PF object. In this analysis, the missing transverse energy was reconstructed using the PF algorithm with the Type-I corrections.

Chapter 5

Event selection and background estimation

In the previous chapter, the $t\bar{t}$ object reconstruction was introduced with the identification criteria included for each object. In order to select events that contains $t\bar{t}$ final state, in addition to select interesting objects, several additional quality criteria are needed at the event reconstruction level. The goal of applying event selection criteria is to choose a sample of events which is enriched in $t\bar{t}$ and contains a reasonable statistics.

At the first step of event selection, samples of data and MC are needed. A brief description of datasets and MC signal and background simulation samples is given in section 5.1. In section 5.2, the optimising cuts recommended by the CMS Top physics group to select the signal events are introduced. These selection criteria are applied on simulation and data in order to maximize the signal-to-background ratio (S/B). As mentioned before, the event selection efficiencies and resolutions are different for data and simulated events. In order to make a correct interpretation, the effects of these efficiencies have to be taken into account (see section 5.3). Since there are many background processes with the same final state, the precise estimation of their contributions is an important topic in this analysis. The data driven estimation of main background source, Drell-Yan, is presented in section 5.4. A set of Data-MC comparison control plots for different kinematic observables are shown in section 5.5.

The subject of this analysis is the measurement of the W boson polarization in the top quark decay by using the cosine of the charged lepton emission angle θ_1^* in the W boson rest frame. To obtain θ_1^* and extract W boson helicity fractions, the full kinematic reconstruction of $t\bar{t}$ events is performed in section 5.6. In the presence of the neutrino and anti-neutrino at the final state, several ambiguities are found. The performance of $t\bar{t}$ reconstruction method as well as the comparison between the data and the MC prediction are presented in section 5.7. In this analysis, the fit re-weighting technique is used to extract the W boson helicity fractions from the $\cos \theta_1^*$

distribution, see section 5.8. In order to validate the fitting method for any possible biases in the results and also related uncertainties, several tests are performed which will be discussed in section 5.9. The whole procedure of this analysis with more details can be found in the related analysis note and the public PAS [43, 108].

5.1 Data and Monte Carlo samples

The used datasets for this analysis were recorded with the CMS experiment in proton-proton collisions at the center-of-mass energy of 8 TeV. The data taking is divided into some small time intervals, called luminosity blocks. These blocks are used if they are collected during periods in which the LHC circulating stable colliding beams. The total integrated luminosity for 2012 corresponds to 19.7 fb^{-1} . The full list of datasets with their corresponding integrated luminosities are shown in table 5.1.

Table 5.1: Summary of datasets with sample name and corresponding integrated luminosities.

Dataset	Dataset Name	$L_{\text{int}}(\text{pb}^{-1})$
Run2012A_ee	/DoubleElectron/Run2012A-13Jul2012-v1/AOD	810
Run2012A_ee_EcalRecovery	/DoubleElectron/Run2012A-recover-06Aug2012-v1/AOD	82
Run2012B_ee	/DoubleElectron/Run2012B-13Jul2012-v1/AOD	4404
Run2012C_ee	/DoubleElectron/Run2012C-13Jul2012-v1/AOD	6941
Run2012D_ee	/DoubleElectron/Run2012D-13Jul2012-v1/AOD	7273

The analysis presented here is based on the DoubleElectron primary datasets which contains the events passed through di-electron triggers at the HLT (see section 5.2.1).

Monte Carlo simulation samples are essential tools in each high energy particle physics analysis. They are used to optimize selection procedure, to determine the signal acceptance and to evaluate the contribution from background processes. For specific physics analysis, the corresponding Monte Carlo simulation samples, based on the similar structure as the raw data, are produced. As described before in section 4.1, after event generation, the generated samples are processed with the GEANT4 detector simulation package and finally reconstructed and passed through the same analysis chain as the data. In order to get realistic simulated events comparable to the data, the output of the simulation chain is presented in an identical format to the real data. For this analysis, simulated samples containing $t\bar{t}$ signal, backgrounds and the ones for systematic variation studies, have been generated.

The standard top quark pair sample is created by the leading order (LO) MADGRAPH 5.1.48 [85] event generator scaled to the NLO, using PYTHIA 6.426 [86] for hadronization and showering processes. The $t\bar{t}$ simulated events are also interfaced

with Tauola [109] for the τ decay modelling. For the generation of the SM signal sample, top mass is fixed to $172.5 \text{ GeV}/c^2$ [110], also the W boson helicity fractions are set to the nominal values of the SM predictions. The parton distribution function (PDF) used here to describe parton structure is CTEQ6L1 [111]. The theoretical top pair production cross section normalized to the next-to-next-to-leading order (NNLO) is approximated to be $252.9 \pm 16.4 \text{ pb}$ [112].

To make a precision physics measurement, a profound understanding of the background contribution is necessary. The production of gauge bosons like W or Z bosons with additional jets are the dominant background processes for the $t\bar{t}$ events. Both samples are simulated using the MADGRAPH generator. The Z+jets (Drell-Yan) events are splitted in samples with low Z masses ($10 \text{ GeV} < M_{l+l-} < 50 \text{ GeV}$) and samples with higher Z masses ($M_{l+l-} > 50 \text{ GeV}$). Based on the decay products, Drell-Yan process can be categorized as ee , $\mu\mu$ and $e\mu$ through tau decays ($Z \rightarrow \tau^+\tau^- \rightarrow l^+\nu_l\bar{\nu}_\tau l^-\bar{\nu}_l\nu_\tau$).

Other backgrounds come from electroweak single top quark productions in the s-, t- and tW-channels are produced using POWHEG, interfaced to PYTHIA 6.426 for parton showering and hadronization. Additional sources of background arise from electroweak processes (WW, WZ and ZZ). While the WW sample was generated with MADGRAPH, the WZ and the ZZ were produced with PYTHIA 6.426. Similarly, the QCD multijet processes are also hadronized with PYTHIA 6.426. The probability to survive these events after applying tight event selections is very low, then they do not play an important role in our study. A detailed list of all used signal and background MC simulation samples with their corresponding cross sections are found in table 5.2.

The luminosity of simulated samples are re-weighted in order to make compatibility between MC and data. For systematic studies, additional MC samples are also generated. These samples are produced based on different settings for parameters such as factorisation scale, matching threshold, top quark masses and PDF values. Table 5.3 gives a summary of utilized MC samples for systematic studies.

5.2 Event selection

As mentioned before, the final state of a $t\bar{t}$ event in the di-electron channel is composed of two oppositely charged electrons, two jets from b-quarks and large missing E_T^{miss} due to presence of two undetected neutrinos. Then events with two high p_T isolated electrons, at least two b-jets and large missing transverse energy in the final state are considered as signal, including those with τ leptons decaying into electrons. The reconstruction and identification of individual objects are considered in different sections of previous chapter. In addition to the described object identification,

Table 5.2: Summary of signal and background simulated samples with corresponding cross section of each processes.

Dataset	Dataset Name	Cross section (pb)
$t\bar{t}$	/TTJets_FullLeptMGDecays_8TeV-madgraph-tauola/	26.5
	/TTJets_SemiLeptMGDecays_8TeV-madgraph-tauola/	111.1
	/TTJets_HadronicMGDecays_8TeV-madgraph/	115.3
Single t	/T_t-channel_TuneZ2star_8TeV-powheg-tauola/	59.5
	/Tbar_t-channel_TuneZ2star_8TeV-powheg-tauola/	32.1
	/T_s-channel_TuneZ2star_8TeV-powheg-tauola/	4.5
tW	/Tbar_s-channel_TuneZ2star_8TeV-powheg-tauola/	2.1
	/T_tW-channel-DR_TuneZ2star_8TeV-powheg-tauola/	11.2
$\bar{t}W$	/Tbar_tW-channel-DR_TuneZ2star_8TeV-powheg-tauola/	11.2
WW	/WWJetsTo2L2Nu_TuneZ2star_8TeV-madgraph-tauola/	5.8
WZ	/WZ_TuneZ2star_8TeV_pythia6-tauola/	22.4
ZZ	/ZZ_TuneZ2star_8TeV_pythia6-tauola/	9.0
W+jets	/WJetsToLNU_TuneZ2Star_8TeV-madgraph-tarball/	37509.0
$Z/\gamma^* \rightarrow l^+l^-$	/DYJetsToLL_M-10To50filter_8TeV-madgraph/	860.5
$Z/\gamma^* + \text{jets}$	/DYJetsToLL_M-50_TuneZ2Star_8TeV-madgraph-tarball/	3532.8
QCD.EMenriched2030	/QCD_Pt_20_30_EMEnriched_TuneZ2star_8TeV_pythia6/	29148.6
QCD.EMenriched3080	/QCD_Pt_30_80_EMEnriched_TuneZ2star_8TeV_pythia6/	4615893.0
QCD.EMenriched80170	/QCD_Pt_80_170_EMEnriched_TuneZ2star_8TeV_pythia6/	183294.9
QCD.MuEnrichedPt15	/QCD_Pt_20_MuEnrichedPt_15_TuneZ2star_8TeV_pythia6/	134680.0
QCD.BCtoE2030	/QCD_Pt_20_30_BCtoE_TuneZ2star_8TeV_pythia6/	167388.0
QCD.BCtoE3080	/QCD_Pt_30_80_BCtoE_TuneZ2star_8TeV_pythia6/	167040.0
QCD.BCtoE80170	/QCD_Pt_80_170_BCtoE_TuneZ2star_8TeV_pythia6/	12981.9

Table 5.3: Summary of Monte Carlo samples used for systematic studies.

Dataset	Dataset Name	Cross section (pb)
$t\bar{t}$	/TTJets_matchingup_TuneZ2star_8TeV-madgraph-tauola/	252.9
	/TTJets_matchingdown_TuneZ2star_8TeV-madgraph-tauola/	252.9
	/TTJets_scaleup_TuneZ2star_8TeV-madgraph-tauola/	252.9
	/TTJets_scaledown_TuneZ2star_8TeV-madgraph-tauola/	252.9
	/TTJets_mass161_5_TuneZ2star_8TeV-madgraph-tauola/	252.9
	/TTJets_mass163_5_TuneZ2star_8TeV-madgraph-tauola/	252.9
	/TTJets_mass166_5_TuneZ2star_8TeV-madgraph-tauola/	252.9
	/TTJets_mass169_5_TuneZ2star_8TeV-madgraph-tauola/	252.9
	/TTJets_mass175_5_TuneZ2star_8TeV-madgraph-tauola/	252.9
	/TTJets_mass178_5_TuneZ2star_8TeV-madgraph-tauola/	252.9
	/TTJets_mass181_5_TuneZ2star_8TeV-madgraph-tauola/	252.9
	/TTJets_mass184_5_TuneZ2star_8TeV-madgraph-tauola/	252.9
	/TTTo2L2Nu2B_8TeV-powheg-pythia6/	252.9
	/TT_CT10_TuneZ2star_8TeV-powheg-tauola/	252.9
	/TT_CT10_TuneZ2star_8TeV-powheg-tauola/	252.9
$Z/\gamma^* + \text{jets}$	/TT_8TeV-mcatnlo/	252.9
	/DYJetsToLL_M-50_matchingup_8TeV-madgraph-tauola/	3532.8
	/DYJetsToLL_M-50_matchingdown_8TeV-madgraph/	3532.8
	/DYJetsToLL_M-50_scaleup_8TeV-madgraph-tauola/	3532.8
W+jets	/DYJetsToLL_M-50_scaledown_8TeV-madgraph-tauola/	3532.8
	/WJetsToLNU_matchingup_8TeV-madgraph-tauola/	37509.0
	/WJetsToLNU_matchingdown_8TeV-madgraph-tauola/	37509.0
	/WJetsToLNU_scaleup_8TeV-madgraph-tauola/	37509.0
	/WJetsToLNU_scaledown_8TeV-madgraph-tauola/	37509.0

in order to reduce as many background events and keep as many signal events as possible, a set of requirements are also applied. These additional requirements for objects and event selections are based on Top Physics Analysis Group (TopPAG) recommendations for 2012 analysis [113]. In this section, the selection criteria for each objects as well as $t\bar{t}$ are briefly described.

5.2.1 Trigger selection

Before any event selection, a specific trigger is used to select events with two electrons in the final state. The list of the triggers in data and simulated MC samples used for di-electron channel is shown in table 5.4.

Table 5.4: Triggers used in data and simulation for di-electron channel.

Sample	Channel	Trigger path
Data	Di-Electron	HLT_Ele17_CaloIdT_CaloIsoVL_TrkIdVL
	Di-Electron	TrkIsoVL_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_v*
MC	Di-Electron	HLT_Ele17_CaloIdT_CaloIsoVL_TrkIdVL
	Di-Electron	TrkIsoVL_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_v17

For di-electron channel, selected electrons must be matched to the mentioned trigger paths with p_T threshold of 17 GeV/c and 8 GeV/c for the leading and sub-leading electrons, respectively.

5.2.2 Event cleaning

Once an event passes the trigger selection, the triggered events must have at least one good primary vertex. In order to ensure the events originated from the collision and to reject the non-collision scraping events, only mentioned primary vertices with at least 10 high purity tracks are considered. These high purity tracks are selected by requiring $\chi^2/N_{\text{dof}} > 4$. In addition to above condition, the selected primary vertices must satisfy the following criteria:

- The maximum longitudinal distance from the PV to the nominal interaction point is 24 cm, $|z| < 24$ cm.
- The radial distance from the PV to the nominal interaction point is less than 2 cm, $\rho < 2$ cm.

In the case of no leading vertex satisfying these criteria, the event is dropped from the analysis. The good vertices are sorted according to a decreasing sum of the p_T^2 , where p_T is the transverse momentum of each track associated to studied vertex. If more than one PV is found, the one with the highest $\sum p_T^2$ is chosen.

As discussed in section 4.1.4, by increasing the luminosity of the LHC machine, multiple proton-proton collisions can occur, known as pileup effect. Pileup effects result in additional energy deposits that affect the jet reconstruction and E_T^{miss} . The contribution of charged hadrons from pileup events which do not come from the primary vertex is subtracted event-by-event while neutral hadron contributions are considered in the jet energy correction [102]. The number of reconstructed primary vertices is affected by the pileup at the LHC with different values for data and MC samples. The events of simulated samples are re-weighted such that the average number of primary vertices was the same in data and MC. In figure 5.1, the number of primary vertices in data and MC are compared for the case of before and after applying pileup re-weighting procedure.

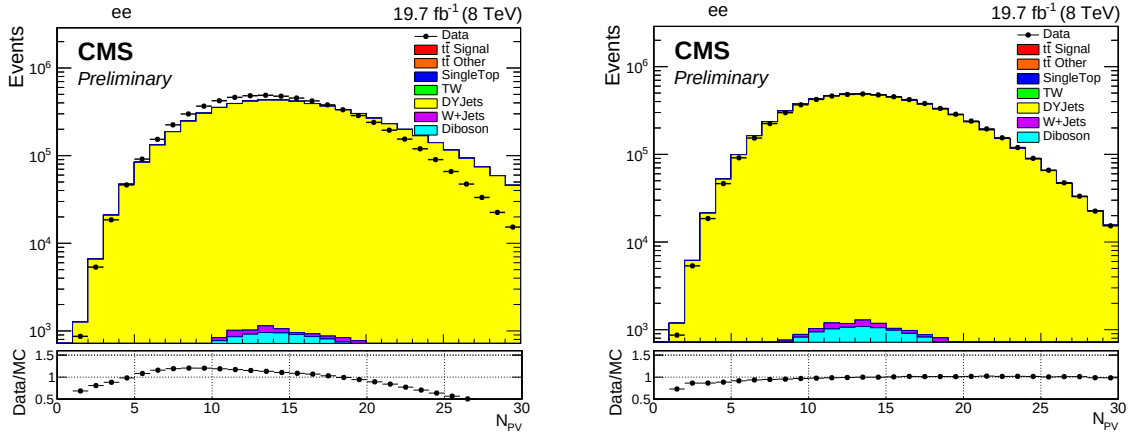


Figure 5.1: Data and MC primary vertices distributions in the di-electron channel before (a) and after (b) pileup re-weighting. These plots are created after two oppositely charged electrons requirement.

5.2.3 Electron selection

As described in section 4.3.4, the electron object is defined as a reconstructed energy deposits in the electromagnetic calorimeter associated to reconstructed tracks in the tracker subdetectors. Starting from the trigger requirement, the reconstructed offline candidate electrons have to pass additional selection cuts. In this analysis the PF electron candidates are reconstructed by using the GSF algorithm [113]. Only electrons with at least one hit in more than one layer of the tracker system are considered here. Reconstructed electrons are required to have a transverse momentum $p_T > 20 \text{ GeV}/c$ and a pseudorapidity from the central part of the detector $|\eta| < 2.5$. Moreover, the electron candidates with an ECAL supercluster in the region of $1.4442 < |\eta| < 1.5660$, corresponding to the gap between the barrel and end-cap parts of the electromagnetic calorimeter, are not taken into account.

For electrons, it is required that the transverse impact parameter of their tracks with respect to the primary vertex is less than 0.04 cm. Furthermore, a multivariate electron identification algorithm with selection cut of Multi-Variate Analysis $MVA > 0.5$ is required as an optimal compromise between efficiency of electron identification and rejection of fake candidate electrons.

In order to provide good separation between isolated electrons and jets, the electron candidates have to pass isolation criteria, which means little energy flow is found in the vicinity of the studied electron. This property is given by ρ corrected relative isolation quantity, defined as the ratio of the sum of the charged, neutral and photon isolation within a cone of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.3$ around the electron track over the p_T of electron candidate:

$$REI = \frac{(I^{CH} + \max(I^{NH} + I^\gamma - \rho_{Iso} \times A_{eff}(\gamma + NH), 0))}{p_T(e)}, \quad (5.1)$$

where I^{CH} and I^{NH} are sum p_T of the charged and neutral hadrons, respectively. In this equation, I^γ and A_{eff} denote sum E_T of PF photons and values of effective areas calculated on data derived from recommendation in reference [114]. The ρ corrected relative isolation for electron candidate in this analysis is required to be less than 0.15.

Beyond electron requirements, additional electrons can also appear from the conversions of photons which have to be rejected [115]. To do that, an electron without any hit in the innermost tracker layer is rejected. The direction of electron from photon conversion is parallel to the positron track at the decay point. In the case of existence opposite curvature track with minimal distance less than 0.02 ($|\Delta \cos \theta| < 0.02$ and $d_{r-\phi} < 0.02$), electron is considered to stem from a photon conversion and is therefore rejected. Finally, in order to distinguish an electron from a combination of a muon (which provides the track) and hadrons in the ECAL, we require $\Delta R > 0.1$ between the electron and any global muon [116]. All electron and positron candidates passing above requirements are considered in this analysis.

5.2.4 Jet selection and b-tag requirement

In section 4.3.6, jet clustering algorithms to reconstruct jet objects have been introduced. In this analysis, the selected jets based on PF objects are reconstructed with the anti- k_T algorithm within a cone size of $R = 0.5$. The defined L1FastJet, L2Relative, L3Absolute and L2L3Residual jet energy correction steps are also applied to the raw jet momenta (see section 4.3.6). Most of the jets decaying from $t\bar{t}$ dilepton channels are generally central with high p_T . The event selection requires the presence of at least two jets with transverse momentum of $p_T > 30 \text{ GeV}/c$ and pseudorapidity of $|\eta| < 2.4$.

In order to avoid contributions from misidentified photons, electrons, pions or instrumental noise, some common identification requirements are applied to the selected jets [117]. These additional selection criteria to identify loose jets contain: neutral hadronic energy fraction < 0.99 , charged hadronic energy fraction > 0 , neutral electromagnetic energy fraction < 0.99 , charged electromagnetic energy fraction < 0.99 and charged particle multiplicity > 0 . Furthermore, in order to avoid double counting, the jets overlapping with accepted electrons are removed if the distance between the jet and each of two selected electrons is $\Delta R < 0.5$.

From the signal event topology, we request minimal number of two jets passing selection criteria which at least two of them are tagged as b-jets. Since the top quark decays into a W boson and a b quark, a b-tagging requirement is used to improve the background rejection. Moreover, by using the b-tagging, the fraction of properly reconstructed jets can be increased significantly. Within this analysis, the Combined Secondary Vertex algorithm with the loose (CSVL) working point as detailed in section 4.3.7 is used for b-jet identification. The output discriminator value of $b_{\text{disc}} > 0.244$ corresponding to loose working point is applied to distinguish between the b jet and other jets coming from light-flavor quarks or gluons. The chosen working point gives b-tagging efficiency around 85%, while the probability of a non-b-quark jet being mis-tagged is about 10% [118].

5.2.5 Missing transverse energy cut

The event missing transverse energy, E_T^{miss} is an observable in the transverse plane which represents presence of two undetectable neutrinos. As defined before in section 4.3.8, the E_T^{miss} is constructed as the magnitude of the transverse momentum imbalance, which is defined as a negative vectorial sum of the transverse momenta of all the reconstructed particles (equation 4.2). In this analysis, the E_T^{miss} is reconstructed using PF algorithm with Type I corrections, where the jet energy corrections from PF jets are also propagated to the MET calculation.

Among the different backgrounds, two most important sources of backgrounds are DY and QCD multijet. With no neutrinos in the final state, the contributions of DY and QCD+jets are significantly reduced by requiring $E_T^{\text{miss}} > 40 \text{ GeV}$.

5.2.6 $t\bar{t}$ event selection

In this thesis, the final state of the di-electron $t\bar{t}$ events is characterized by the presence of two oppositely charged electrons, at least two b-tagged jets and missing transverse energy. In the case there are more than two electrons in events, the electron-positron pair which maximize the sum of the transverse momenta, is considered as the best electron-positron candidate. Also when more than two b-tagged

jets are exist, the two with the highest p_T are selected.

Following the selection requirements of reconstructed objects, the selection of event consists of a sequential cuts as follow:

- The suitable double electron trigger path is fired depending on the used data taking periods.
- At least one well-defined primary vertex candidate reconstructed from at least 10 high purity tracks is required.
- Only events with two oppositely charged and isolated electrons are considered in this study.
- In order to avoid the low mass resonances background region, the di-electron invariant mass is required to satisfy $M_{e^-e^+} > 20 \text{ GeV}/c^2$.
- Additionally, to suppress backgrounds from Z+jets events in the di-electron channel, events in the Z mass region of $76 < M_{e^-e^+} < 106 \text{ GeV}/c^2$ are rejected.
- Events with at least two b-tagged jets, passing the jets selection criteria are considered in this study.
- Selected events in the e^-e^+ mode must satisfy MET cut by $E_T^{\text{miss}} > 40 \text{ GeV}$.
- Events are required to pass the $t\bar{t}$ reconstruction criteria, which will be described in section 5.6.

5.3 Corrections to Monte Carlo events

The MC generators are not simulated with all effects that appear in data. To account for the observed differences between data and simulation, some data-to-MC scale factors are applied on the simulated events. After introducing the pileup reweighting in section 5.2.2, the observed data still do not describe properly with the MC simulation. As a consequence, in order to improve the data-MC agreement, several scale factors from various sources of efficiencies like trigger, electron identification and isolation, as well as for the b-jet identification are applied to the MC events.

These scale factors are obtained as a function of the transverse momentum and pseudorapidity of the related objects. Introduced efficiencies are explained in the following, while other corrections related to the jet object, like jet energy resolution and jet energy corrections are studied in sections 4.3.6. After applying all corrections, it is found that the transverse momentum distribution of the top

quark in the data is softer than the one in the simulation [119, 120]. To solve this inconsistency, a new scale factor will be introduced in section 5.6.4, so that the related distribution in the simulation is re-weighted to match the data.

5.3.1 Trigger efficiency

The trigger efficiency is measured using triggers that are only weakly correlated to the di-electron triggers and counts the number of such events passing and failing the di-electron trigger selection, as described in [121]. The advantage of this method is that the applied electron definition is identical to the one used for the event selection. In this section, the principles of this method to estimate the di-electron trigger selection efficiency as well as trigger efficiency scale factor are summarized.

First, we determine a series of cross triggers which are weakly correlated with the di-electron trigger. Second, the number of events passing the cross triggers and the $t\bar{t}$ di-electron event selection are counted, as N_{Xtrig} . Similarly, the number of events $N_{Xtrig \& e^-e^+trig}$ which pass the cross trigger, the $t\bar{t}$ di-electron selection and the di-electron trigger selection are numerated. Third, the di-electron trigger efficiency in the data is given by:

$$\epsilon_{trigger} = \frac{N_{Xtrig \& e^-e^+trig}}{N_{Xtrig}}. \quad (5.2)$$

Since the MET triggers are found to be weakly correlated with the di-electron trigger, the missing transverse energy based datasets were used as the cross triggers. This method produces a large enough number of di-electron events so that the statistical uncertainty keep below 1%.

The trigger efficiencies are calculated for different period of datasets separately, combined by weighting factors based on the corresponding luminosities. The trigger efficiency scale factor is derived by dividing efficiency in data by MC. The global trigger scale factors, as well as their uncertainties, is given in reference [121] for the different selection steps. From these results, it can be concluded that the trigger SFs are compatible for different steps within the quoted uncertainty. In this analysis, the scale factor measured for events after the requirement of two oppositely charged isolated electrons is chosen to apply to MC simulation events. The trigger efficiency can be parametrized for any binning of kinematic variables. A dependency of trigger efficiencies for data and MC as well as resulting scale factor on the transverse momentum and pseudorapidity of the leading electron are shown in figure 5.2.

In general, within the uncertainties, the derived scale factor is stable around $97\% \pm 1\%$ with respect to the electron transverse momentum. Only a small dependency on the electron pseudorapidity is observed, therefore the scale factors are determined and applied to the simulations in bins of η . Figure 5.3 shows the trigger

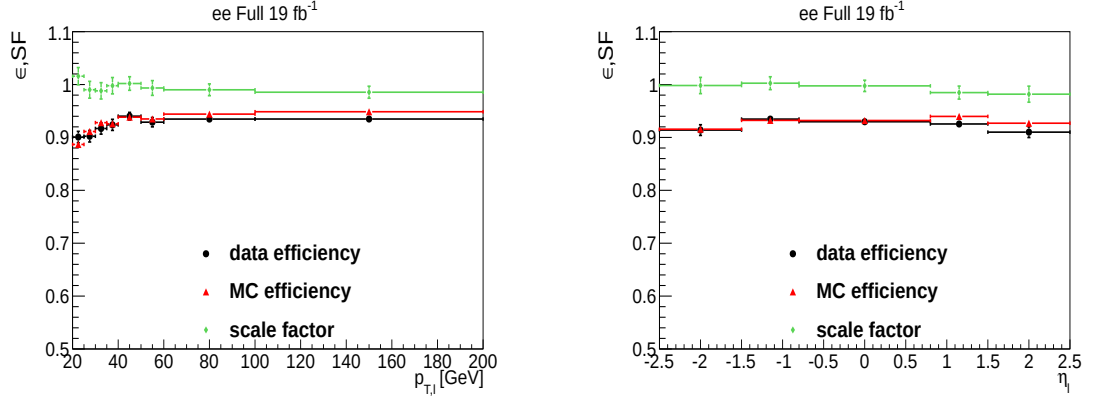


Figure 5.2: Di-electron trigger efficiency as a function of electron p_T (left) and η (right), for data (black circles) and MC (red triangles) and correction factor (green diamonds), taken from [121].

scale factor as two dimension of η for the first and second leading electron candidates.

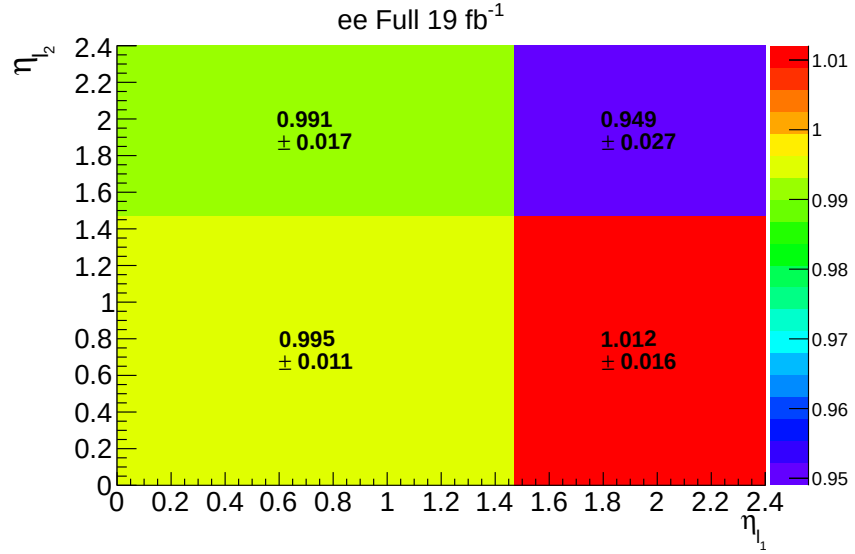


Figure 5.3: Scale factor for the di-electron trigger as a function of leading electron η (x-axis) and second leading electron η (y-axis), taken from [121].

5.3.2 Electron identification and isolation efficiencies

The efficiencies for electron identification and isolation are measured separately using a tag and probe method, following a simple cut and count approach [121]. In order to make a precise efficiency measurement, a sample of high purity leptons from Z processes with clean decay mode is chosen. Oppositely charged di-electron candidates from decay of the Z bosons, known as tag and probe electrons, are considered here to estimate the efficiency. The invariant mass of the tag and probe

electron candidates fall in the range of $76 < M_{e^-e^+} < 106 \text{ GeV}/c^2$. In this analysis, the electron selection, isolation and identification criteria are applied to identify tagged electrons. Tagged electrons are matched to the electrons passing through the high level trigger path. For the probe electron, the applied selection criteria depend on the selected efficiency. In fact, the electron identification and isolation efficiencies are calculated by measuring the ratio of di-electron candidates with and without applied selection cut on the related probe electrons.

The identification efficiency is estimated as the number of probe electrons passing the complete selection criteria except for the isolation cuts over the total number of probe electrons without passing any requirements. The isolation efficiency is defined as the ratio of electrons passing identification and isolation cuts to the total number of electrons passing only identification requirements. The electron identification and isolation scale factors are also obtained as ratios of related efficiencies in data to Drell-Yan:

$$SF_{Id,Iso} = \frac{\epsilon_{Id,Iso}^{data}}{\epsilon_{Id,Iso}^{DY}}. \quad (5.3)$$

The electron identification and isolation efficiencies as well as related scale factors are presented in terms of the electron transverse momentum p_T and pseudorapidity η in figure 5.4 and 5.5, respectively.

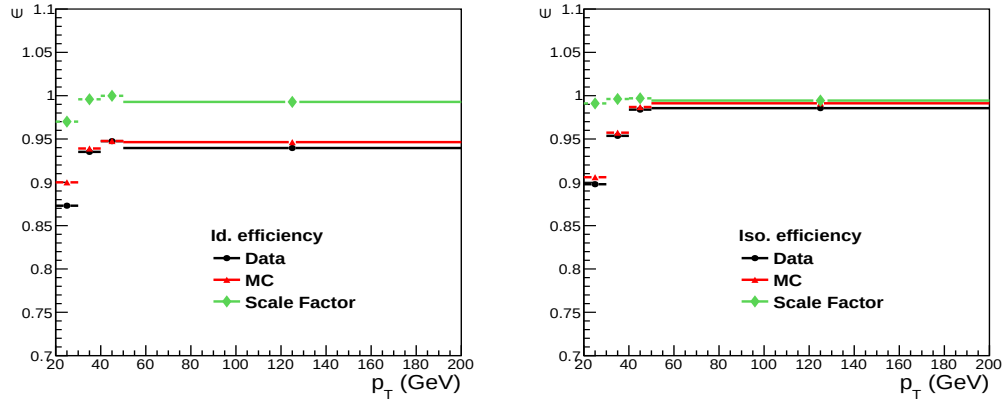


Figure 5.4: Identification (left) and isolation (right) efficiencies for a loose electron in terms of transverse momentum. Black circles correspond to efficiency from data, red triangles to efficiency from the DY MC and green diamonds to the scale factor, taken from [121].

By multiplying these two scale factors, the global electron scale factor is obtained to be around $98\% \pm 1\%$. In addition to the mentioned statistical uncertainty, an additional 1% flat systematic uncertainty will be also considered, to account for possible differences between Z and $t\bar{t}$ event kinematics. Figure 5.6 shows the global

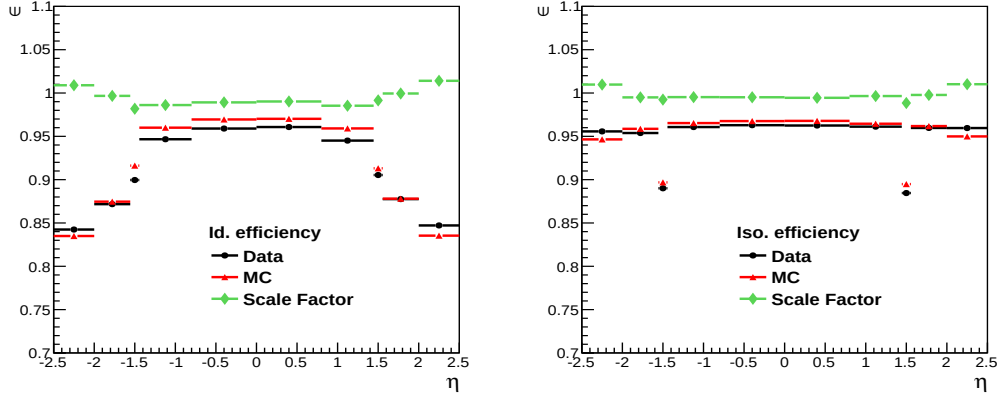


Figure 5.5: Identification (left) and isolation (right) efficiencies for a loose electron in terms of pseudorapidity. Black circles correspond to efficiency from data, red triangles to efficiency from the DY MC and green diamonds to the scale factor, taken from [121].

electron scale factors as a two dimensional scan in $p_T - \eta$ plane.

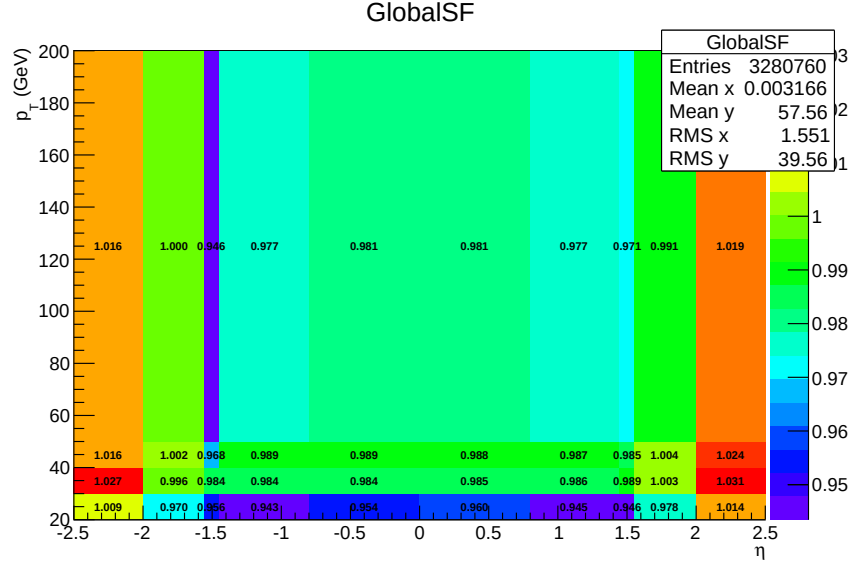


Figure 5.6: Electron global scale factor as a two dimensional function of (η, p_T) , taken from [121].

5.3.3 B-tagging efficiency and scale factor

The b-tagging algorithm in section 4.3.7 performs differently on recorded data and simulated events. The b-tagging efficiency procedure follows the recommendation of the BTV POG group [122]. The b-jet and mis-tag efficiencies are estimated from the MC prediction. The mis-tag means that a c-jet or light-flavour is identified

as a b-jet. To estimate b-tagging efficiency per event, the efficiencies per jets are calculated separately for each b, c and light flavours as a function of p_T and $|\eta|$ of the jet from the $t\bar{t}$ sample with no b-jet requirement. The efficiency for each jet flavour is defined by counting the number of jets with studied flavour, identified as a b-jet passing CSVL tagger cut divided by the total number of jets with corresponding flavour:

$$\epsilon_f(i, j) = \frac{N_f^{b\text{-tagged}}(i, j)}{N_f^{\text{Total}}(i, j)}, \quad (5.4)$$

where subscript f is jet flavour obtained from the MC truth jet flavour information. The (i, j) denotes bin index in the (p_T , η) phase space. The resulting efficiencies for b, c and light-flavour jets derived from the $t\bar{t}$ signal sample are shown in figure 5.7.

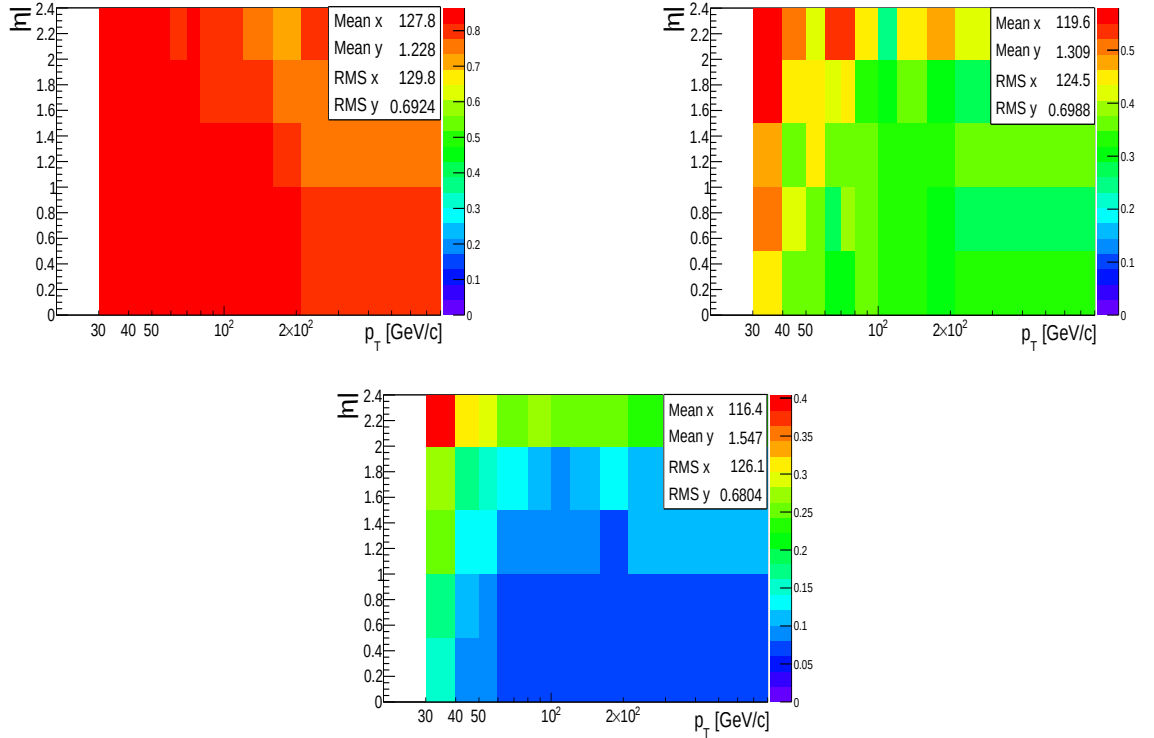


Figure 5.7: The $p_T - \eta$ dependent efficiencies for b (top-left), c (top-right) and light flavour jets (bottom) in the di-electron channel.

In order to calculate b-tagging efficiencies per event, the corrected b-jet identification efficiencies are also defined by applying the provided SFs for each individual jet flavours as $\epsilon_f^{\text{corr}} = \text{SF}_f \times \epsilon_f$. In our study, individual jet flavour scale factors are provided by BTV group as a function of the jet p_T from 20 to 800 GeV/c and jet $|\eta| < 2.4$ using QCD enriched dataset [123]. In this analysis, the method 1-a described in reference [122] is used to extract the b-tagging efficiencies per event. In this approach, the b-tagging efficiencies per jet are translated to event probabilities

for MC and data, written as:

$$\begin{aligned} P(MC) &= \prod_{i=\text{tagged}} \epsilon_i \prod_{j=\text{non-tagged}} (1 - \epsilon_j), \\ P(Data) &= \prod_{i=\text{tagged}} \epsilon_i^{\text{corr}} \prod_{j=\text{non-tagged}} (1 - \epsilon_j^{\text{corr}}). \end{aligned} \quad (5.5)$$

The final form of the b-tagging scale factor per event is extracted as ratio of event probabilities:

$$SF_{b\text{-tag}} = \frac{P(Data)}{P(MC)}. \quad (5.6)$$

From this formula, the scale factors are parametrised as a function of p_T and η of the constituent jets for the kinematic range of $p_T = 20 - 800 \text{ GeV}/c$ and $|\eta| < 2.4$. In the measurements, it is estimated that the SF for the b-jet identification efficiency to be around 0.98.

5.4 Drell-Yan background estimation

One source of background contribution comes from W+jets and QCD multi-jet processes that contain misidentified non-isolated electrons, known as the fake electron background. The contribution from these sources are well suppressed after requesting two isolated electrons and b-tagged jets. As mentioned before, the main background source in the di-electron channel is the Drell-Yan ($Z/\gamma^* \rightarrow e^-e^+$) which produced with associated jets and large E_T^{miss} due to resolutions effects.

By applying event selection criteria, discussed in section 5.2.6, especially requiring MET cut and excluding Z boson mass window, the contribution of Drell-Yan background is reduced significantly. However, a small fraction of such events with large E_T^{miss} and di-electron invariant mass away from the Z boson mass window still remain. Due to large uncertainties on the non-Gaussian E_T^{miss} tails and on the Z boson multi-jet cross section, the proper modelling of remaining events is difficult. Therefore, the contribution of Drell-Yan is estimated using the data driven method described in the following, more details can be found in [124].

First, two event regions corresponding to the inside and outside of the Z boson mass window are defined as control and signal regions, respectively. Then a scale factor from the control region is extracted, so that it makes the number of entries in data and MC simulation to be equal in this region. This scale factor can be used to normalize the simulation prediction in the signal region. Finally the estimated number of Drell-Yan events in the signal region can be written as:

$$\begin{aligned}
N_{DY}^{SR} &= SF^{CR} \times N_{DY,MC}^{SR}, \\
&= \frac{N_{Data}^{CR} - N_{Non-DY}^{CR}}{N_{DY,MC}^{CR}} \times N_{DY,MC}^{SR},
\end{aligned} \tag{5.7}$$

while N_{DY}^{SR} denotes the number of estimated events in the signal region, $N_{DY,MC}^{CR}$ represents the number of simulated ones in the control regions. The variable N_{Non-DY}^{CR} is the number of simulated events except ones from Drell-Yan contribution in the control region. The parameter N_{Data}^{CR} gives the number of observed events inside Z boson mass window.

Non Drell-Yan background processes contributing to the Z boson mass window with a continuous distribution are estimated from data. This is achieved by measuring the number of observed events in the control region for the $e\mu$ channel:

$$N_{Non-DY}^{CR} = \frac{1}{2} k_{e^-e^+}^{CR} \times N_{Data,e\mu}^{CR}. \tag{5.8}$$

The constant k is the combinatoric factor to take into account the difference in efficiency to reconstruct and to select an electron compared to a muon. This correction factor can be estimated from the number of di-electron and di-muon events in the control region passing standard lepton and jet selection regardless of MET requirement which can be expressed as:

$$k_{e^-e^+}^{CR} = \sqrt{\frac{N_{Data,e^-e^+}^{CR}}{N_{Data,\mu^-\mu^+}^{CR}}}. \tag{5.9}$$

By substituting equation 5.8 in 5.7, the number of estimated Drell-Yan events in signal region is written as:

$$\begin{aligned}
N_{DY}^{SR} &= \frac{(N_{Data,e^-e^+}^{CR} - \frac{1}{2} k_{e^-e^+}^{CR} \times N_{Data,e\mu}^{CR})}{N_{DY,MC}^{CR}} \times N_{DY,MC}^{SR}, \\
&= R_{CR}^{SR} \times (N_{Data,e^-e^+}^{CR} - \frac{1}{2} k_{e^-e^+}^{CR} \times N_{Data,e\mu}^{CR}),
\end{aligned} \tag{5.10}$$

where $R_{CR}^{SR} = N_{DY,MC}^{SR}/N_{DY,MC}^{CR}$ is the ratio of DY events in signal and control regions. The number of observed and expected Drell-Yan events in the signal region and the corresponding normalization scale factors at different levels of the event selection are summarized in table 5.5.

The quoted uncertainties of the reported scale factors in table 5.5 are estimated by taking into account the statistical fluctuation of the data in the control region. The scale factor of Drell-Yan events for di-electron channel with two b-tagged jets is found to be 1.24 ± 0.23 .

Table 5.5: Number of observed and expected Drell-Yan events in the signal region and related scale factors of the di-electron channel at different event selection steps.

e^-e^+ channel	MC DY	Estimated DY	SF
$\geq 2\text{Jets} + E_T^{\text{miss}}$	2525.8 ± 77.8	3259.8 ± 108.9	1.29 ± 0.06
$\geq 2\text{Jets} + E_T^{\text{miss}} + \geq 2b - \text{tag}$	161.2 ± 20.5	200.5 ± 28.6	1.24 ± 0.23

5.5 Data-MC comparison

In order to check how well the simulation represents the distributions from data, kinematic distributions of the selected physics objects in the simulated di-electron $t\bar{t}$ events are compared to the corresponding ones in data. The distributions obtained from the MC are normalized according to their cross section values and the data luminosity to the number of observed events in the data. To make a real data-MC comparison, pileup re-weighting, jet smearing, electron identification and isolation, trigger efficiency and b-tagging correction scale factors are applied to the MC simulation events.

The following control plots are drawn after requiring at least two b-tagged jets before reconstructing the final $t\bar{t}$. In these plots, the $t\bar{t}$ di-electron event topologies, including those come from τ decays, are referred to as $t\bar{t}$ signal, while other $t\bar{t}$ events contain semi-electron or full hadronic final state are considered as $t\bar{t}$ other. The error bands on the following distributions account only for the statistical uncertainty. The kinematic properties of selected electrons, p_T and η , are shown in figure 5.8.

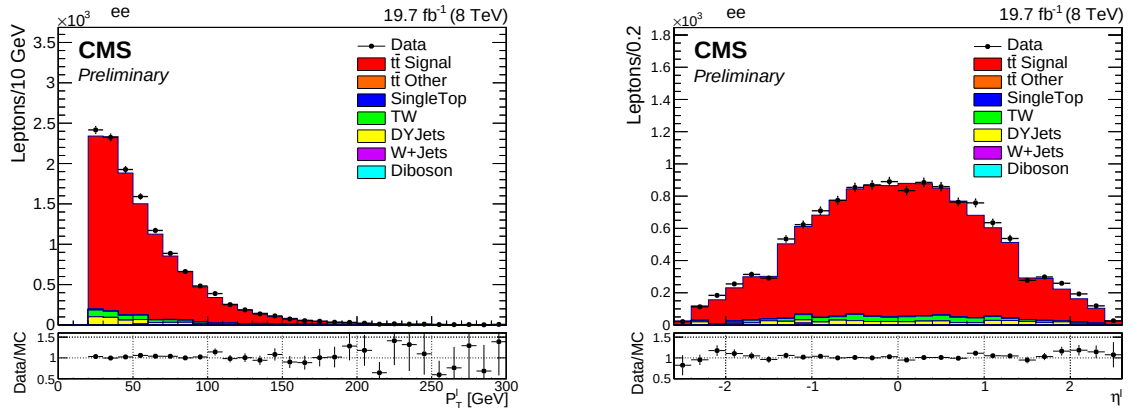


Figure 5.8: The electron p_T (left) and η (right) distributions in the di-electron channel after requesting two b-tagged jets. At this step, all of described correction scale factors are applied on the MC simulation events.

In figure 5.9, the expected number of selected jets and the b-tagged jet multiplicity are displayed.

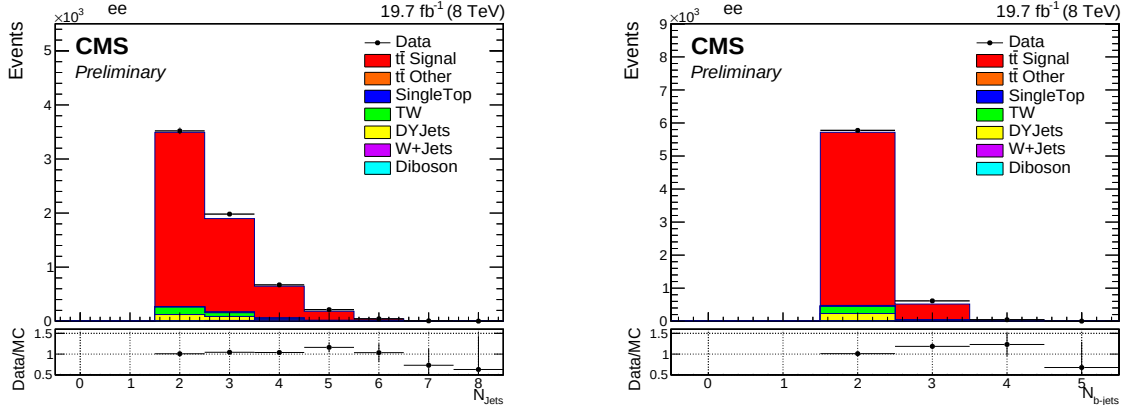


Figure 5.9: The jet multiplicity (left) and b-tagged jet multiplicity (right) distributions in the di-electron channel after requesting two b-tagged jets. At this step, all of described correction scale factors are applied on the MC simulation events.

The comparison of the data and the MC for the transverse momentum of the first and the second leading jets are presented in figure 5.10.

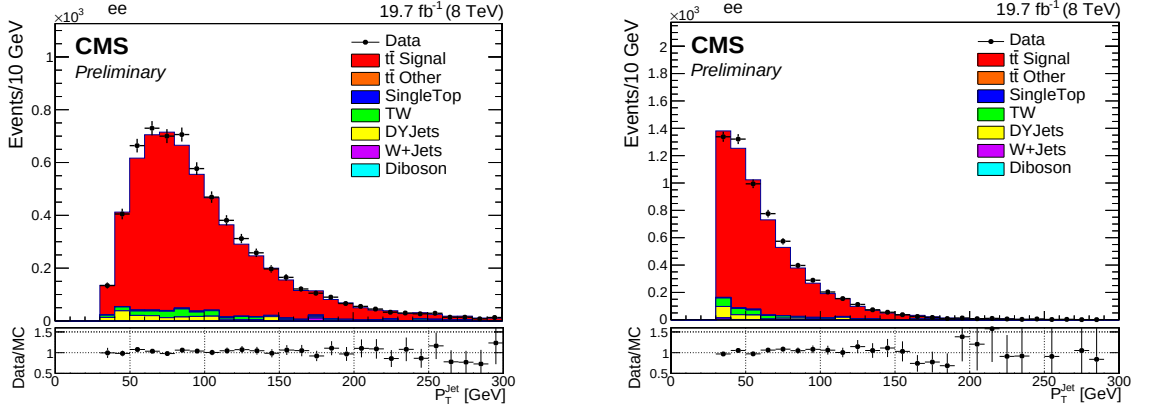


Figure 5.10: The p_T distributions for first (left) and second (right) leading jet in the di-electron channel after requesting two b-tagged jets. At this step, all of described correction scale factors are applied on the MC simulation events.

Figure 5.11, provides the di-electron invariant mass and the E_T^{miss} distributions.

One can clearly see from these distributions that a reasonable agreement between data and MC within the statistical fluctuations is obtained. The same comparison plots in combination of dileptonic channels can be found in appendix B. The consistency between data and simulation allows to study the full kinematic $t\bar{t}$ reconstruction which is the subject of the next section.

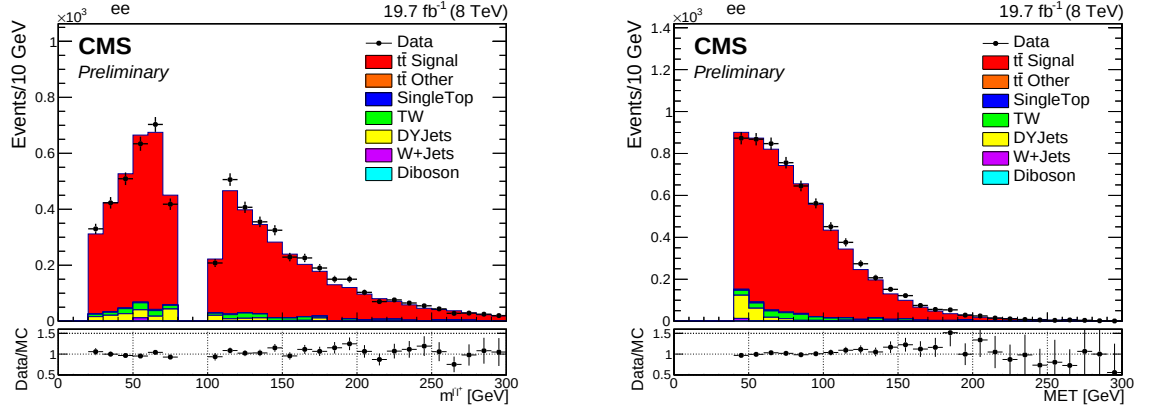


Figure 5.11: The distributions of oppositely charged di-electron invariant mass (left) and missing transverse energy (right) in the di-electron channel after requesting two b-tagged jets. At this step, all of described correction scale factors are applied on the MC simulation events.

5.6 Full reconstruction of the $t\bar{t}$ events

As discussed in the beginning of this chapter, the main parameter of this analysis is the $\cos \theta_1^*$. In order to define the $\cos \theta_1^*$ parameter, the four-momentum vector of the W boson and the top quark are needed. Since these particles are not observed directly, they are reconstructed from their decay products. To fully reconstruct $t\bar{t}$ events, the four-momentum vector of electrons, jets and neutrinos as decay products are measured.

During the full reconstruction of $t\bar{t}$ events, several ambiguities can occur, so that solving these uncertainties is the hardest issue in our study. Although electrons and b-jets are reconstructed precisely, the presence of two undetected neutrinos is inferred by missing transverse energy, leading to an under-constrained system. Another challenge is the lack of jet charge information which leads to two assignment possibilities for reconstructed jets as the b and $\bar{b}$ quarks. In this section, the method to reconstruct $t\bar{t}$ via solving these ambiguities will be discussed.

By assuming that only two undetectable neutrinos construct the missing transverse energy, following two equations from the projections of MET in the transverse plane are obtained:

$$\begin{aligned} (E_T^{miss})_x &= p_{x,\nu} + p_{x,\bar{\nu}}, \\ (E_T^{miss})_y &= p_{y,\nu} + p_{y,\bar{\nu}}. \end{aligned} \quad (5.11)$$

Additionally, the momenta of the two neutrinos are constrained by the four-momentum vector conservations in the W boson and the top quark decay vertices:

$$\begin{aligned}
m_{W^+}^2 &= (E_{e^+} + E_\nu)^2 - (p_{x,e^+} + p_{x,\nu})^2 - (p_{y,e^+} + p_{y,\nu})^2 - (p_{z,e^+} + p_{z,\nu})^2, \\
m_{W^-}^2 &= (E_{e^-} + E_{\bar{\nu}})^2 - (p_{x,e^-} + p_{x,\bar{\nu}})^2 - (p_{y,e^-} + p_{y,\bar{\nu}})^2 - (p_{z,e^-} + p_{z,\bar{\nu}})^2,
\end{aligned} \tag{5.12}$$

$$\begin{aligned}
m_t^2 &= (E_{e^+} + E_\nu + E_b)^2 - (p_{x,e^+} + p_{x,\nu} + p_{x,b})^2 \\
&\quad - (p_{y,e^+} + p_{y,\nu} + p_{y,b})^2 - (p_{z,e^+} + p_{z,\nu} + p_{z,b})^2, \\
m_{\bar{t}}^2 &= (E_{e^-} + E_{\bar{\nu}} + E_{\bar{b}})^2 - (p_{x,e^-} + p_{x,\bar{\nu}} + p_{x,\bar{b}})^2 \\
&\quad - (p_{y,e^-} + p_{y,\bar{\nu}} + p_{y,\bar{b}})^2 - (p_{z,e^-} + p_{z,\bar{\nu}} + p_{z,\bar{b}})^2.
\end{aligned} \tag{5.13}$$

To solve above six non-linear equations, the masses of electrons and corresponding neutrinos are neglected and the masses of the W boson, b and top quarks are set to the nominal values used in the simulation. These non-linear equations can be transformed into a fourth-order polynomial equation in terms of one neutrino momentum component, written as [125]:

$$h_0 p_{\nu_x}^4 + h_1 p_{\nu_x}^3 + h_2 p_{\nu_x}^2 + h_3 p_{\nu_x} + h_4 = 0, \tag{5.14}$$

where the coefficients depend on the mass and the four-momentum vector of particles.

This system of equations has been solved analytically with a maximum four-fold ambiguity in reference [125]. These six constraints on kinematic variables of particles are implemented in a package suggested by the CMS physics group [126] to find the neutrino and anti-neutrino momenta. In addition to the four-fold ambiguity, further complications arise due to the lack of jet charge information. Thus, for any events there are up to eight solutions for the neutrino and anti-neutrino momenta. In order to select the best $t\bar{t}$ candidate as a solver of the under constrained system, several different approaches are available which will be discussed in the next section.

5.6.1 Minimum effective mass of $t\bar{t}$ approach

The $t\bar{t}$ event reconstruction in this method is conceptually straightforward and has been used before in reference [127]. The four-momentum vector of input particles contain the reconstructed electrons and jets as well missing transverse energy are fed into this method without any smearing. Instead of using constant values for the masses of the W bosons and the top quarks, these parameters are varied randomly based on the relativistic Breit-Wigner distributions with the nominal simulated mean and width quantities. By using non constant values for input mass parameters, the number of events with one solution would be increased.

With the described conditions and fixing b-jet assignment, the solver can find up

to four solutions for neutrinos momenta. The same procedure is repeated once again for the other possible b-jet pairing and finally, up to eight solutions are obtained for each input values of the W boson and the top quark masses. Among all possible solutions, the neutrino solution is chosen according to a minimal invariant mass of $t\bar{t}$ system as the best candidate. The events are thrown from the rest of analysis, if no solution is given from the solver package. More details on the minimum effective mass method as well as output results can be found in [127].

5.6.2 Analytical Matrix Weighting Technique (AMWT)

In this analysis, to find a probable solution among possible ambiguities, the AMWT method is used. This method has been already used within the CMS in other analysis with the $t\bar{t}$ dileptonic events [110, 128]. In this technique, in order to achieve good resolution, each event is reconstructed assuming fixed values for the W bosons and the top quarks equal to 80.41 GeV/c² and 172.5 GeV/c², respectively.

For a sizeable fraction of events, there is no solution found for values of input particle momenta and missing transverse energy. In order to improve the efficiency of this technique, the input variables of analytical solver contain electrons, b-jets, and MET are smeared. For that, each event is reconstructed 300 times, each time smearing factors are selected randomly according to the detector resolutions for the relevant parameters. The smearing factors of p_T , η and ϕ for reconstructed particles are determined from the MC simulation matched to the corresponding kinematic inputs at the truth level. The distributions of smearing factors for input jet and MET observables in the MC signal sample are shown in figures 5.12 and 5.13, respectively.

As described before, for a given event smearing iteration, the resulting up to eight solutions for the $t\bar{t}$ kinematics can be found. In order to choose a preferred b-jet assignment and the best top pair candidate, a proper weight based on the probability of observing such a configuration is assigned for each of these solutions [129]:

$$w(\vec{X}|m_t) = \left\{ \sum_{\text{Initial partons}} F(x_1)F(x_2) \right\} p(E_{e^+}^*|m_t)p(E_{e^-}^*|m_t), \quad (5.15)$$

where the summation of parton distribution function $F(x_{1,2})$ is performed over the possible leading order initial parton states ($u\bar{u}$, $\bar{u}u$, $d\bar{d}$, $\bar{d}d$, $s\bar{s}$, $\bar{s}s$, gg), and evaluated at the scale of $Q^2 = m_t^2$. In this equation, x_1 , x_2 indicate Bjorken values for the initial parton states. The variable $p(E_{e^\pm}^*|m_t)$ denotes the probability density of observing a charged electron with energy $E_{e^\pm}^*$ in the rest frame of the top quark with mass m_t ,

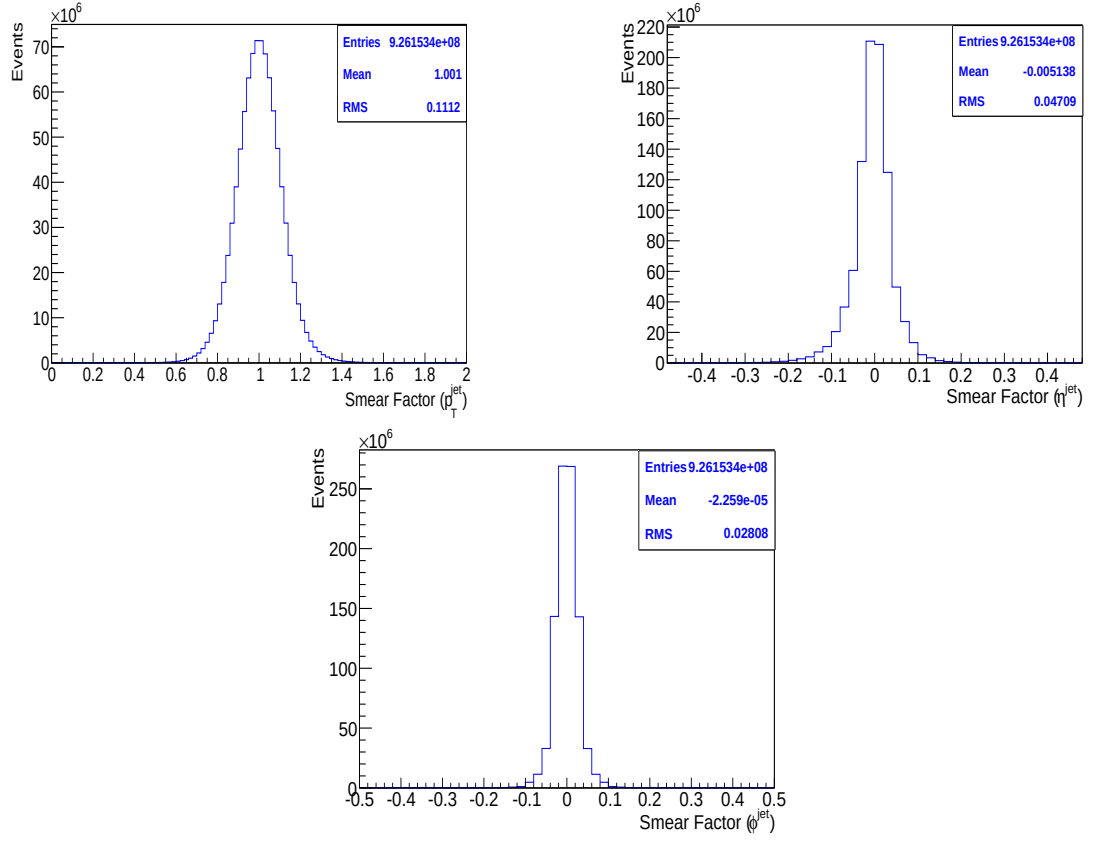


Figure 5.12: The extracted smear factors from detector resolution information for jet p_T (top-left), jet η (top-right) and jet ϕ (bottom).

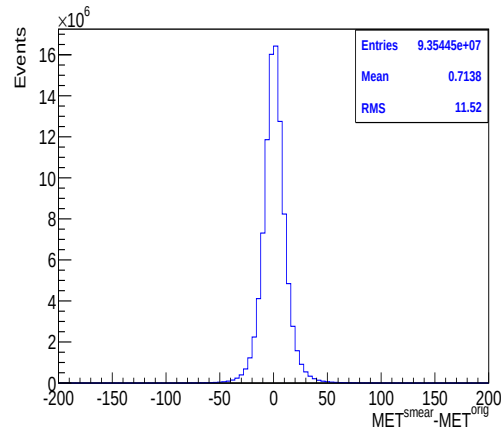


Figure 5.13: The propagated effects of smeared objects into the MET.

given by [129].

$$p(E_{e^\pm}^*|m_t) = \frac{4m_t E_{e^\pm}^* (m_t^2 - m_b^2) - 2m_t E_{e^\pm}^*}{(m_t^2 - m_b^2)^2 + m_W^2 (m_t^2 - m_b^2) - 2m_W^4}. \quad (5.16)$$

To achieve the best correct electron-jet combination, sum of weights for each of b-jet assignment configurations is calculated and finally one with the highest sum of weights is chosen as a correct assignment. By fixing the preferred b-jet assignment, the best top pair candidate is selected as a solution with the highest weight value among up to four possible solutions. This procedure is repeated 300 times for both of data and MC events and finally, the top quark three-momentum vector is extracted by averaging over all found best top pair candidates. Afterwards, the four-momentum vector of the top quark is obtained from the corresponding three-momentum vector with the fixed value for the top quark mass. To determine kinematics of other particles, similar averaged method with the corresponding mass values is used.

The advantage of the AMWT method compared to the minimum effective mass approach is that it suppresses possible biases in kinematic variables by integrating over smearing iterations. The bias study is performed by comparing kinematic observables at the reconstruction level to the truth level for both of $t\bar{t}$ reconstruction methods. The results of one and two dimensional bias study for one of the important parameter of the analysis, $M_{t\bar{t}}$, are shown in figures 5.14 and 5.15.

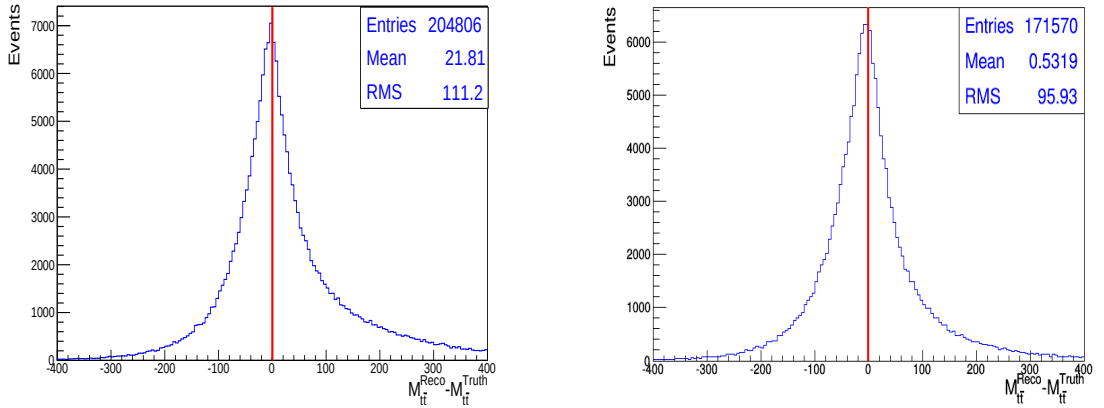


Figure 5.14: Bias study for the invariant mass of $t\bar{t}$ system in one dimensional for minimum effective mass (left) and AMWT (right) methods.

Despite applying smearing procedure to improve solver efficiency, cases with no solution are also found approximately 10% and 13% of events in MC and data samples, respectively. The events with no solution are not considered in the rest of analysis. The solver efficiency is defined as the ratio of events with solution to total events before $t\bar{t}$ kinematic reconstruction step. These efficiencies in the

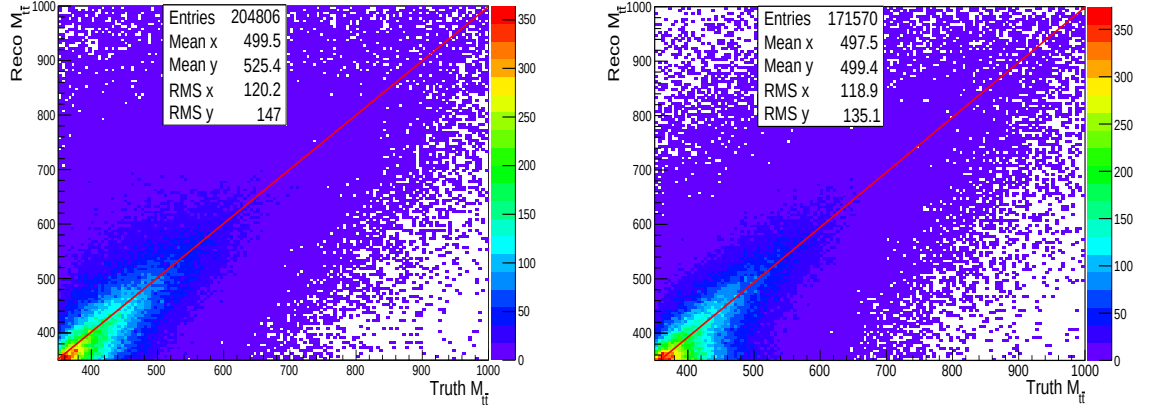


Figure 5.15: Bias study for the invariant mass of $t\bar{t}$ system in two dimensional for minimum effective mass (left) and AMWT (right) methods.

MC simulation are compared to the data for different kinematic observables. To compensate for the possible differences, solver scale factor is defined as the ratio of solver efficiencies in data and MC. In figure 5.16, the solver efficiencies for both data and simulated events as well as solver scale factors are shown as a function of electron transverse momentum, pseudorapidity and jet multiplicity.

The data and MC solver efficiencies, and related scale factor independent of the kinematic observables are listed in table 5.6.

Table 5.6: Data and MC efficiencies and also related scale factor for di-electron channel.

Channel	Data Solver Eff	MC Solver Eff	Solver SF
Di-Electron	0.86	0.89	0.97

5.6.3 Kinematics method

Another method to fully reconstruct the $t\bar{t}$ events is known as kinematic top quark reconstruction algorithm. This new method has been described with more details in [120]. Similar to the AMWT method, while electrons, jets and MET are smeared based on the information from detector resolution, the masses of the W bosons and the top quarks are fixed to the nominal values. Instead of using the weight function in the AMWT method to disambiguate the electron-jet combination, the b-jet assignment with the largest sum of weights according to the expected invariant mass of electron-jet pairs at the truth level is taken as a correct one. The M_{lb} distribution estimated from the signal MC at particle level is shown in figure 5.17.

In this method, by fixing the correct b-jet assignment, the best $t\bar{t}$ candidate

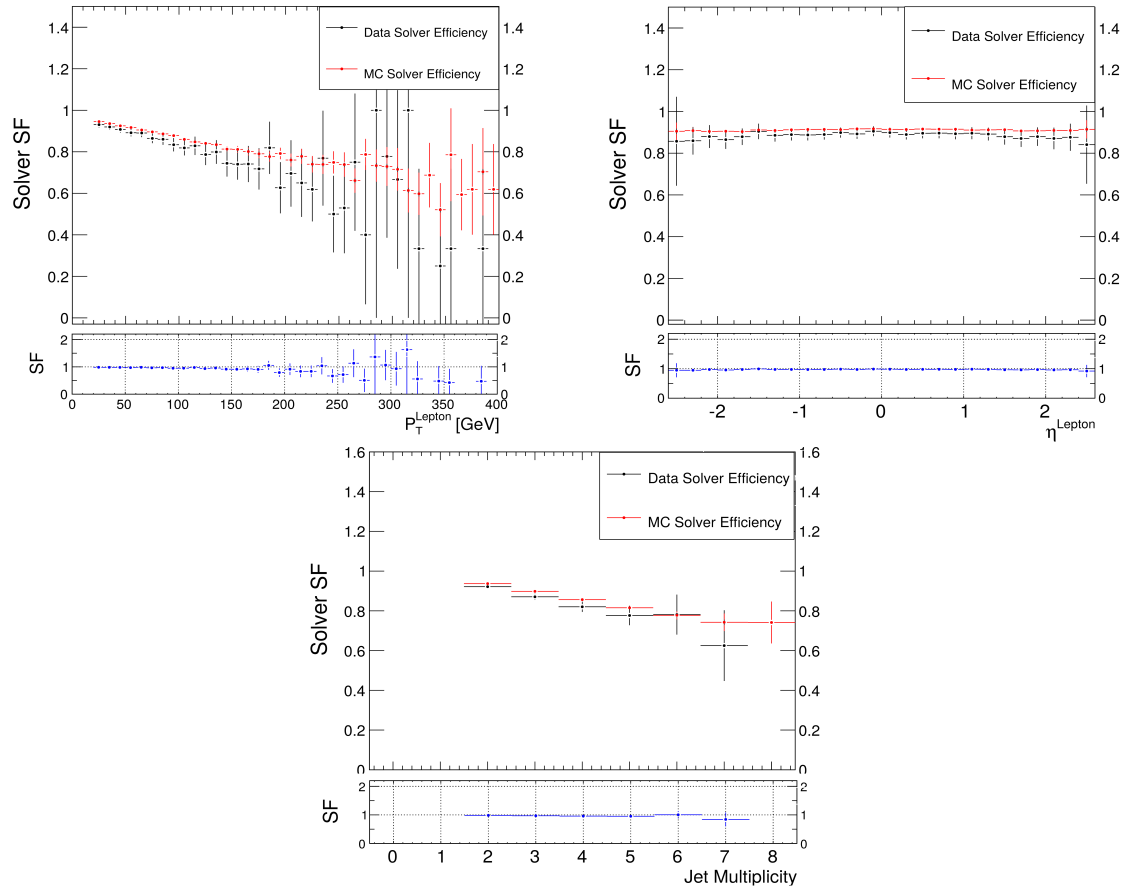


Figure 5.16: Solver efficiencies of data and MC as well as extracted scale factors for electron p_T (top-left), electron η (top-right) and jet multiplicity (bottom).

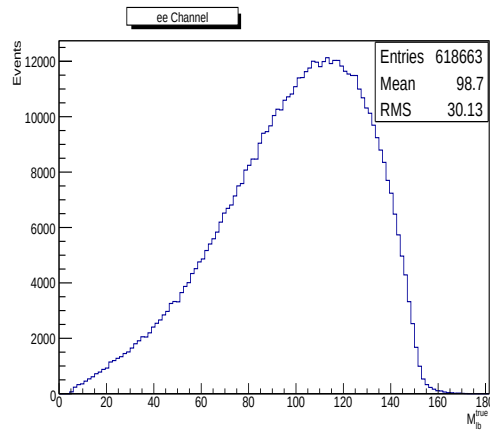


Figure 5.17: Distribution of M_{lb} at the truth level.

among up to four ones is defined as the solution with the lowest value of top pair invariant mass. To determine three and four-momentum vector of the top quark and other kinematic observables, same approach to the AMWT technique is used here. The same solver efficiencies and bias measurements are also done for this algorithm, but in the rest of thesis, the results of the AMWT method are reported.

5.6.4 Top quark p_T re-weighting

Despite applying described scale factors in section 5.3, CMS has observed that the distribution of the top quark transverse momentum is softer in data than in MC simulation [130]. This inconsistency can be occur especially for the simulated samples which are produced by the MADGRAPH-PYTHIA generator. The full explanation about source of this discrepancy can be found in reference [131]. To correct this effect, MC simulation events are re-weighted using the p_T values of top and anti-top quarks at the parton truth level. The form of re-weighting function is given as following:

$$SF_{p_T \text{ re-weighting}} = \sqrt{SF(p_T^{\text{top}}) \times SF(p_T^{\text{anti-top}})}, SF(p_T) = \exp(0.156 - 0.00137p_T). \quad (5.17)$$

This additional scale factor is applied in the simulation events for the top quark p_T and other distributions. In figure 5.18, the effect of the top quark p_T re-weighting is demonstrated by comparing the data-MC top quark p_T distributions.

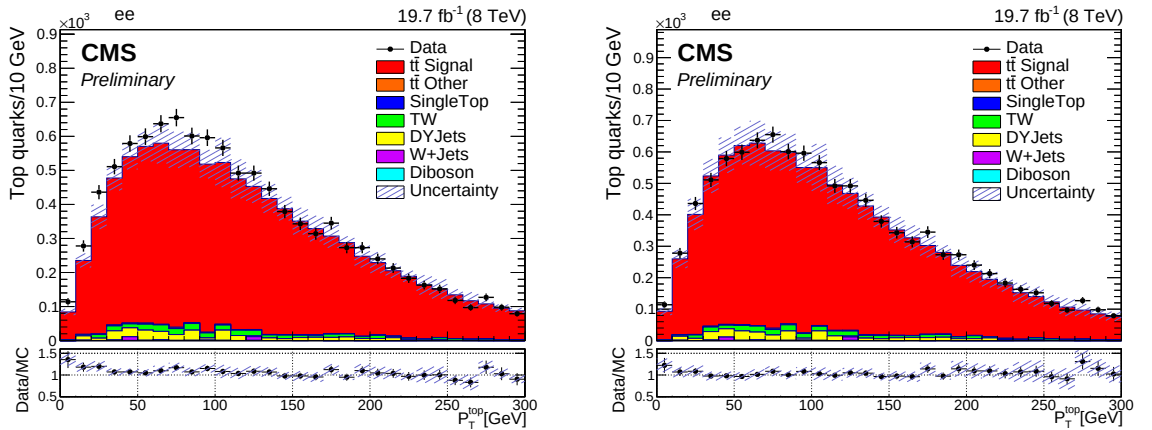


Figure 5.18: Transverse momentum distributions of the top quark before (left) and after (right) applying top p_T re-weighting in the di-electron channel. All of described correction scale factors are applied on the MC simulation events.

As can be seen, the agreement for the top quark p_T variable is improved after applying the top quark p_T re-weighting scale factor.

5.7 Event yield and control plots

The quality of the AMWT method is checked with the data-MC comparison results. The number of observed compared to the expected simulated events after each consecutive selection steps are given in table 5.7.

Table 5.7: Number of events for each MC samples, data and total MC in different steps of the di-electron channel. The statistical uncertainty on MC samples are also included in this table.

e^-e^+ channel	≥ 2 electrons	≥ 2 jets	MET	≥ 2 b-tag	$t\bar{t}$
Diboson	4456.9 ± 21.4	352.9 ± 4.8	155.8 ± 3.7	10.3 ± 0.9	5.4 ± 0.6
Wjets	1102.9 ± 107.3	76.9 ± 28.8	44.5 ± 22.4	10.6 ± 10.6	10.3 ± 10.3
DY	542787 ± 1160.5	18605.4 ± 215.9	3821.8 ± 108.7	222.5 ± 26.8	150.8 ± 21.6
tW	1887.4 ± 30.4	653.0 ± 18.0	518.8 ± 16.1	193.4 ± 9.7	114.0 ± 7.3
SingleTop	22.8 ± 2.8	4.4 ± 1.3	3.5 ± 1.1	1.7 ± 0.8	0.2 ± 0.2
$t\bar{t}$ other	98.4 ± 2.7	83.2 ± 2.5	58.8 ± 2.1	23.7 ± 1.3	17.9 ± 1.1
$t\bar{t}$ signal	20004.7 ± 30.9	13598.1 ± 25.4	13598.1 ± 25.4	5761.4 ± 16.5	4997.6 ± 15.1
Sum MC	570361 ± 1166.5	33373.9 ± 220.1	15338.8 ± 114.5	6223.7 ± 34.7	5296.4 ± 29.3
Data	596122 ± 772.1	35896 ± 189.5	15311 ± 123.7	6437 ± 80.2	5490 ± 74.1

The MC simulation predicts that the contribution of $t\bar{t}$ signal events is dominated. The MC yields are normalized to the integrated luminosity of 19.7 fb^{-1} using the cross section values from table 5.2. In this table, to make a better description of the data by the MC, each MC events are weighted with the correction scale factors. The comparison between data and MC simulation for the kinematic variables of W bosons and top quarks are shown in figures 5.19 and 5.20. In the following figures, the hatched areas correspond to the systematic uncertainty on the MC simulation which will be explained in the next chapter.

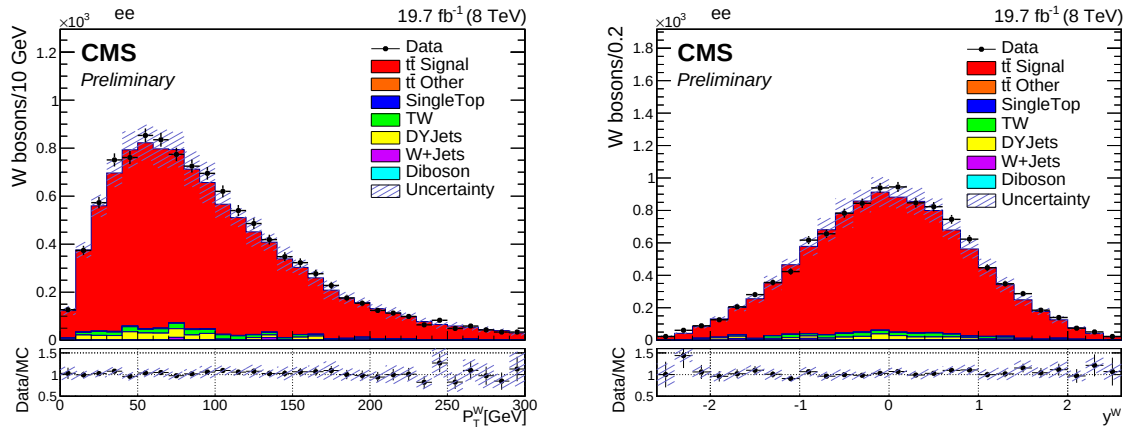


Figure 5.19: The transverse momentum (left) and rapidity (right) distributions of the W boson in the di-electron channel after full kinematic reconstruction of $t\bar{t}$ events. All described correction scale factors are applied on the MC simulation events. The hatched areas show the systematic uncertainties on the MC events.

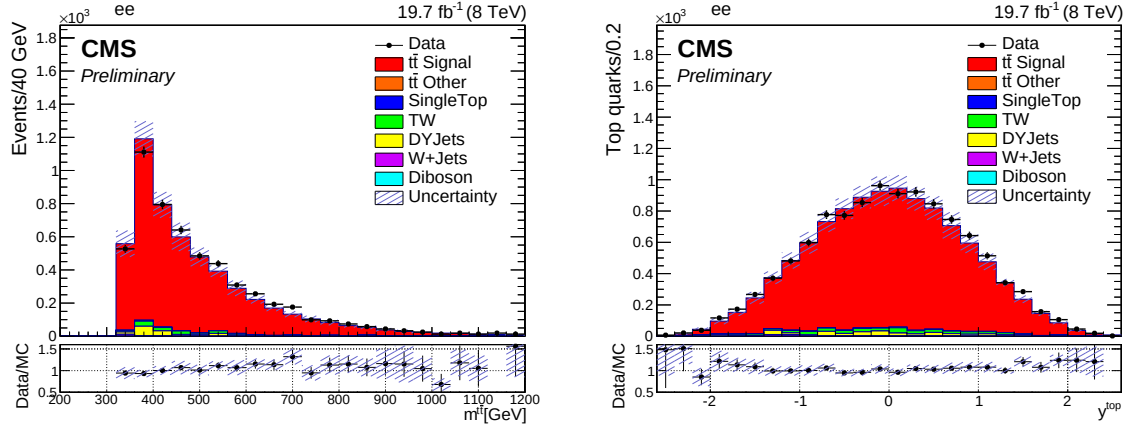


Figure 5.20: The distributions of $t\bar{t}$ invariant mass (left) and top rapidity (right) in the di-electron channel after full kinematic reconstruction of $t\bar{t}$ events. All described correction scale factors are applied on the MC simulation events. The hatched areas show the systematic uncertainties on the MC events.

In figure 5.21, the distributions of $\cos\theta_1^*$ for the di-electron and combination of all dileptonic channels are presented.

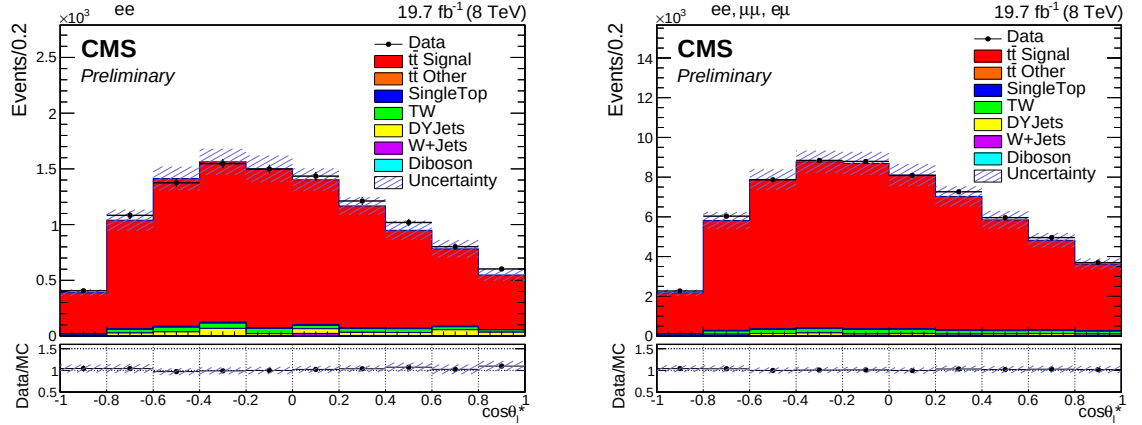


Figure 5.21: The distributions of $\cos\theta_1^*$ in the di-electron (left) and the combination of all channels (right) after full kinematic reconstruction of $t\bar{t}$ events. All described correction scale factors are applied on the MC simulation events. The hatched areas show the systematic uncertainties on the MC events.

More distributions comparing data to MC for combination of all dileptonic channels can be found in appendix B. In general, after applying correction scale factors, the simulation gives a good description of the shape of the data.

5.8 Fit re-weighting method

The main goal of this analysis is to provide a precise measurement of the W boson helicity fractions. This measurement at the reconstruction level is affected by the top kinematic reconstruction, detector acceptance cuts and the detector resolution. To take into account these effects accurately, the measurement of the W boson helicity fractions are done at the generation or the truth level. The MC signal events at the truth level are generated with the fixed SM values for the W boson helicity fractions. In order to measure observed fractions with any possible values, transition from truth level with the fixed values to the cases with the arbitrary polarization fractions are necessary.

In this analysis, to account simultaneously both resolution effects and dependency on the polarization fractions without any biased effects by the SM expectations, a fitting method according to re-weighting technique is used [41, 127, 132]. In this method, each MC signal events at the reconstruction level is re-weighted by using the truth distribution of $\cos\theta_1^*$. Since the other non signal samples are not expected to provide sensitive effects, this re-weighting procedure has not been applied on them. The new obtained $\cos\theta_1^*$ distribution for the MC signal sample after being added to the expected backgrounds is fitted to the observed one by using likelihood function.

In this section, the conceptual basic of the re-weighting method is explained with more details. Due to detector effects, the observed distribution for each observables is different from the one obtained in the truth level. The migration of events from the true bins to the observed one is performed by introducing a response function, calculated using MC information of the signal process. The mathematical expression explains the relation between distributions of true quantity c in the parton level with the measured quantity e written as [133]:

$$g(e) = \int P(e, c)f(c)dc, \quad (5.18)$$

where, the $P(e, c)$ is response function representing the migration probability from true value (c) to the measured one (e). If defined distributions $g(e)$ and $f(c)$ represent by histograms, the response function is moved to response matrix and also integral is substituted by summation as:

$$E_i = \sum_{j=1}^m P_{ij}C_j, \quad i = 1, 2, \dots, n. \quad (5.19)$$

In this equation, E is a measured histogram with n elements and C stands for the true distribution with m elements. The element P_{ij} in the $n \times m$ response matrix refers to the probability of the true value c in histogram bin j to the observe value

e in histogram bin i.

This procedure is extended to the $\cos \theta_1^*$ variable by replacing distribution functions to $d\Gamma/d\cos \theta_1^* = \rho(\cos \theta_1^*)$ as following:

$$\rho(\cos \theta_{e^\pm, reco}^* | \vec{F}^{Free}) \propto \int d\cos \theta_{e^\pm, gen}^* \rho(\cos \theta_{e^\pm, gen}^* | \vec{F}^{Free}) \mathcal{R}(\cos \theta_{e^\pm, gen}^*, \cos \theta_{e^\pm, reco}^*), \quad (5.20)$$

where $\mathcal{R}(\cos \theta_{e^\pm, gen}^*, \cos \theta_{e^\pm, reco}^*)$ is the response matrix that transfers generated observable $\cos \theta_{e^\pm, gen}^*$ to the reconstructed one $\cos \theta_{e^\pm, reco}^*$. As mentioned before, this matrix includes all of the detector resolution, acceptance, inefficiencies as well as intrinsic biases from the top pair kinematic reconstruction. The function $\rho(\cos \theta_{e^\pm, gen}^* | \vec{F}^{Free})$ with free values for the W boson helicity fractions, can be expressed versus the same parameter with the fixed SM W boson polarization values:

$$\rho(\cos \theta_{e^\pm, gen}^* | \vec{F}^{Free}) = \frac{\rho^{Free}(\cos \theta_{e^\pm, gen}^*)}{\rho^{SM}(\cos \theta_{e^\pm, gen}^*)} \times \rho^{SM}(\cos \theta_{e^\pm, gen}^* | \vec{F}), \quad (5.21)$$

where $\rho^{SM}(\cos \theta_{e^\pm, gen}^*)$ is phase space density function for $\cos \theta_1^*$ variable at the truth level, satisfying normalization constraint $\int d\cos \theta_{e^\pm, gen}^* \rho^{SM}(\cos \theta_{e^\pm, gen}^*) = 1$. By substituting above relation into equation 5.20, the following final form is obtained:

$$\rho(\cos \theta_{e^\pm, reco}^* | \vec{F}^{Free}) \propto \int d\cos \theta_{e^\pm, gen}^* \frac{\rho^{Free}(\cos \theta_{e^\pm, gen}^*)}{\rho^{SM}(\cos \theta_{e^\pm, gen}^*)} \rho(\cos \theta_{e^\pm, gen}^* | \vec{F}^{SM}) \times \mathcal{R}(\cos \theta_{e^\pm, gen}^*, \cos \theta_{e^\pm, reco}^*). \quad (5.22)$$

This equation can also be written in terms of binned histogram for $t\bar{t}$ signal sample:

$$N_i^{reco}(\vec{F}^{Free}) \propto \sum_{j=1}^m W_j(\cos \theta_{e^-e^+, gen}^*; \vec{F}) N_j^{gen} \mathcal{R}_{ij}(\cos \theta_{e^\pm, gen}^*, \cos \theta_{e^\pm, reco}^*). \quad (5.23)$$

Here, index i and j denote to ith and jth bins of the $\cos \theta_1^*$ distributions at the reconstruction and the truth levels, respectively. Since each event of $t\bar{t}$ sample has two tWb vertices, the weight $W_j(\cos \theta_{e^-e^+, gen}^*; \vec{F})$ parameter can be expressed as:

$$W_j(\cos \theta_{e^-e^+, gen}^*; \vec{F}) = W_j(\cos \theta_{e^-, gen}^*; \vec{F}) \times W_j(\cos \theta_{e^+, gen}^*; \vec{F}), \quad (5.24)$$

$$W_j(\cos \theta_{e^\pm, gen}^*; \vec{F}) = \frac{\rho^{Free}(\cos \theta_{e^\pm, gen}^*)}{\rho^{SM}(\cos \theta_{e^\pm, gen}^*)}.$$

By using equation 1.60, weight function can be written as:

$$W_j(\cos \theta_{e^\pm, gen}^*; \vec{F}) = \frac{\frac{3}{8}(1 - \cos \theta_{e^\pm, gen}^*)^2 F_L^{Free} + \frac{3}{8}(1 + \cos \theta_{e^\pm, gen}^*)^2 F_R^{Free} + \frac{3}{4} \sin^2 \theta_{e^\pm, gen}^* F_0^{Free}}{\frac{3}{8}(1 - \cos \theta_{e^\pm, gen}^*)^2 F_L^{SM} + \frac{3}{8}(1 + \cos \theta_{e^\pm, gen}^*)^2 F_R^{SM} + \frac{3}{4} \sin^2 \theta_{e^\pm, gen}^* F_0^{SM}} \Big|_{bin j}. \quad (5.25)$$

In the case of event-by-event re-weighting, the i-j element of response matrix ($\mathcal{R}_{ij}$) equals to one or zero, so that equation 5.23 is simplified to this form:

$$N_i^{reco}(\vec{F}^{Free}) \propto \sum_{Events, bin i} W(\cos \theta_{e^- e^+, gen}^*; \vec{F}), \quad (5.26)$$

where the summation is over events in the i th bin of reconstructed $\cos \theta_1^*$ distribution. In order to extract helicity fractions, the final fit is performed by implementing event-by-event re-weighting algorithm in a minimization program. Assuming Poisson statistics for the $\cos \theta_1^*$ distribution, a binned likelihood function is introduced as:

$$\mathcal{L}(\vec{F}^{Free}) = \prod_{i \in bins} \frac{N_{MC}(i; \vec{F})^{N_{data}(i)}}{N_{data}(i)!} \times e^{-N_{MC}(i; \vec{F})}, \quad (5.27)$$

where $N_{data}(i)$ and $N_{MC}(i; \vec{F})$ are the number of observed and expected events in i th bin of the $\cos \theta_1^*$ distribution, respectively. By using equation 5.26, the number of MC events is written as a combination of re-weighted $t\bar{t}$ signal events and backgrounds:

$$N_{MC}(i; \vec{F}) = N_{Bkg}(i) + N_{t\bar{t}}(i; \vec{F}), \quad (5.28)$$

$$N_{t\bar{t}}(i; \vec{F}) = \mathcal{F}_{t\bar{t}} \left[\sum_{t\bar{t} \text{ events, bin } i} W(\cos \theta_{gen}^*; \vec{F}) \right], \quad (5.29)$$

$$N_{Bkg}(i) = N_{single \text{ top}}(i) + N_{DY}(i) + N_{di-boson}(i) + N_{W+jets}(i) + N_{t\bar{t} \text{ other}}(i). \quad (5.30)$$

As discussed before, due to low statistics, top quark pair events in $t\bar{t}$ other sample and single top quark events are not re-weighted. In above equation, $\mathcal{F}_{t\bar{t}}$ is a free normalization parameter for the $t\bar{t}$ component which is not sensitive to the extracted W boson helicity fractions. This variable covers most of possible biases due to detector inefficiencies, luminosity estimation, tagging efficiencies, acceptance of cuts and theory uncertainties. The uncertainties related to the background normalization are considered as a separate source of systematics.

After constructing likelihood function, the final measured helicity fractions are obtained by maximizing the defined likelihood or minimizing the $-2 \ln(\mathcal{L}(\vec{F}^{Free}))$ function. This can be done by performing a 3-D fit on the likelihood function using

MINUIT package [134]. As can be seen from equations 5.25 and 5.29, the helicity fractions F_0 , F_L and F_R and the normalization constant $\mathcal{F}_{t\bar{t}}$ are the free parameters of the fit. The unitary constraint $F_0 + F_L + F_R = 1$ can be used to determine right handed fraction versus other two polarizations. Therefore, in this analysis we implement F_0 , F_L and $\mathcal{F}_{t\bar{t}}$ as free parameters into described 3-D fit. The full results of helicity fractions as well as related statistic and systematic uncertainties will be presented in the next chapter.

5.9 Validation of the fitting method

In order to check biases in the extracted fractions and related uncertainties, some consistency tests are needed to be considered. To verify the validity of the applied method, different kinds of checks are performed. In these validation tests, ensembles of pseudo experiments obtained from MC samples with different settings are generated. In this analysis, two linearity and error estimation tests are investigated to prove that the output results are unbiased and to check whether the method gives proper statistical uncertainties. These two validation checks will be studied with more details in the following sections.

5.9.1 Linearity check of method

The main idea of this test is to check whether the fit re-weighting method could return the same output values as input ones, while the input helicity fractions are being fixed. The results of the linearity check is interpreted as a two dimensional phase space of input-output helicity fractions. Firstly, 300 pseudo-experiments using 1/3 fraction of events in the simulated samples to create pseudo-data are generated. To build these pseudo-data collections, events are selected randomly from simulated samples. In each pseudo-experiments, the number of pseudo-data events in the i th bin of $\cos\theta_1^*$ histogram can be written as:

$$N_{pseudo-data}(i; \vec{F}) = \sum_{1/3 \text{ of } t\bar{t} \text{ sample, bin } i} \frac{\rho^{Free}(\cos\theta_{e^\pm, gen}^*)}{\rho^{SM}(\cos\theta_{e^\pm, gen}^*)} + N_{BKG}^{(1/3)}(i; \vec{F}), \quad (5.31)$$

where the sum and background parts run over 1/3 of randomly selected events in the MC simulation samples. For each of these 300 pseudo-experiments, while the right handed fraction is fixed to the SM value at the truth level (0.0017), the F_0 value is varied linearly in 300 steps in $[0, 1 - F_R]$ interval and the F_L fraction is also changed according to the unitary constraint. At the next step, each of obtained pseudo-data set is normalized to the integrated luminosity of 19.7 fb^{-1} . Finally,

the pseudo-data is fitted to the entire simulated samples using same re-weighting method described in the previous section. This procedure is repeated 300 times for each of 300 pseudo-data. For each configurations with the fixed input helicity fractions, 300 output values are obtained.

For a given fixed F_R value, the dependency of reconstructed F_0 and F_L mean values over 300 pseudo-experiments on the corresponding input fractions are shown in figure 5.22.

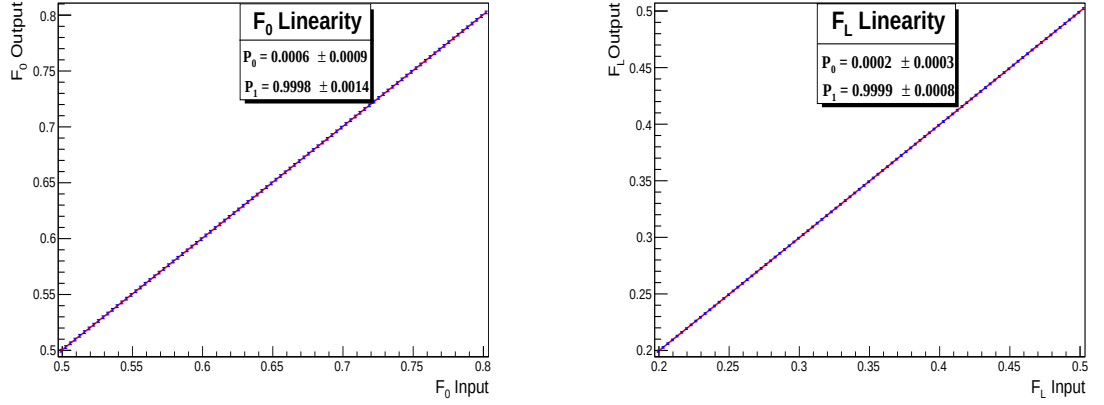


Figure 5.22: Linearity checks for F_0 (left) and F_L helicity fractions in the di-electron channel. The linear polynomial functions are fitted on these graphs.

By fitting the distributions with a straight line, the slopes perfectly consistent with unity and offset values are compatible with zero for both F_0 and F_L helicity fractions. For the right handed fraction, the mean of 300 extracted output values for each of pseudo-experiment are displayed versus the sequential number of pseudo-experiment in figure 5.23.

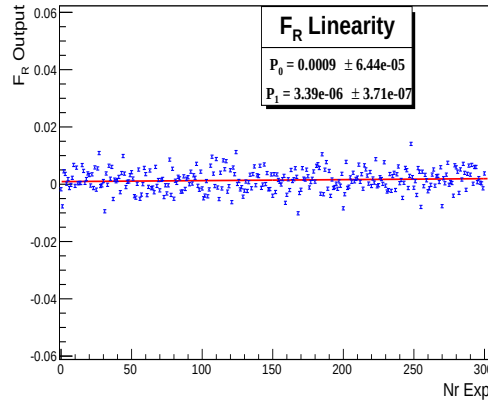


Figure 5.23: Linearity check for the right handed helicity fraction with the linear polynomial fitted function.

The result of helicity fraction is then fitted with a linear polynomial function, where the fit parameters are compatible with the expectation values. From the linearity check results, it is concluded that the fit re-weighting method does not introduce any bias, even for the case of input values away from the SM prediction.

5.9.2 Statistical properties of the estimators

To study statistical properties of the extracted helicity estimators, the pull and the residual test are introduced in this section. In these tests, 1000 sets of pseudo-data are created from the entire MC signal and background samples, where the helicity fractions are fixed to the SM expectations. For each of pseudo-experiments, the number of selected events are normalized to an integrated luminosity of 19.7 fb^{-1} , result in compatibility of the statistical uncertainties in the simulated MC and dataset. The events in these pseudo-experiments are selected randomly in order to ensure presence only marginal correlation between the individual fit results.

The constructed pseudo-data sets are fitted into the entire simulated samples with the same re-weighting method used in section 5.8. For these tests, 1000 values for all of extracted helicity fractions are obtained. In order to study statistical biases on the fitting procedure, the pull and residual variables are defined as:

$$\begin{aligned} \text{residual} &= F^{\text{est}} - F^{\text{exp}}, \\ \text{pull} &= \frac{F^{\text{est}} - F^{\text{exp}}}{\sigma^{\text{est}}}, \end{aligned} \tag{5.32}$$

where σ^{est} is the statistical uncertainty on the estimated W boson helicity fractions. In the case of correct estimation, it is expected to have a Gaussian shape with zero mean value for the residual distribution. Also, a Gaussian function centred at zero with unity width can describe the expected distribution of pull results. Figures 5.24, 5.25 and 5.26 display the residual and pull distributions for different helicity fractions as well as fitted Gaussian functions.

The mean values within uncertainties for all residual and pull distributions are equal to zero which means the fit method does not introduce any bias in the measurements. Also, the width of pull distributions are compatible with unity which proves that the uncertainties are estimated properly.

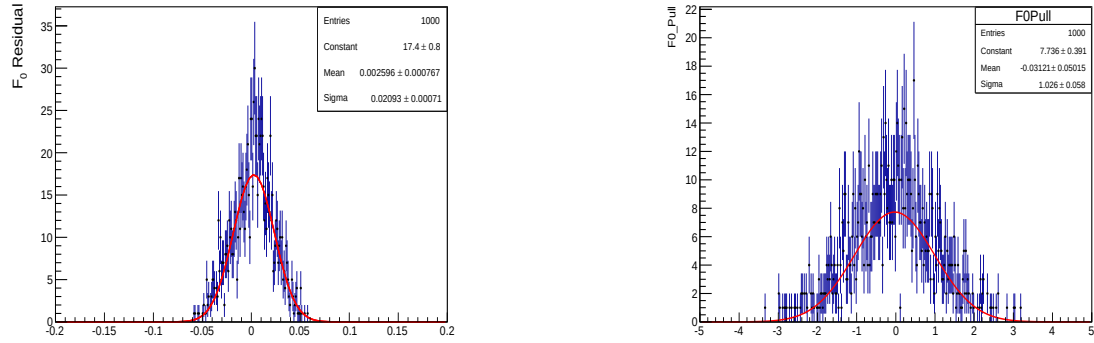


Figure 5.24: The residual (left) and pull (right) distributions for the longitudinal helicity fraction. Each distribution is fitted with a Gaussian function.

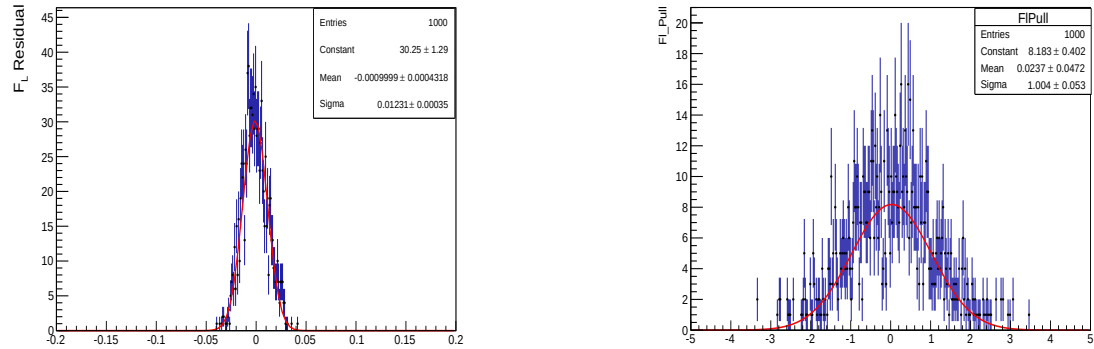


Figure 5.25: The residual (left) and pull (right) distributions for the left handed helicity fraction. Each distribution is fitted with a Gaussian function.

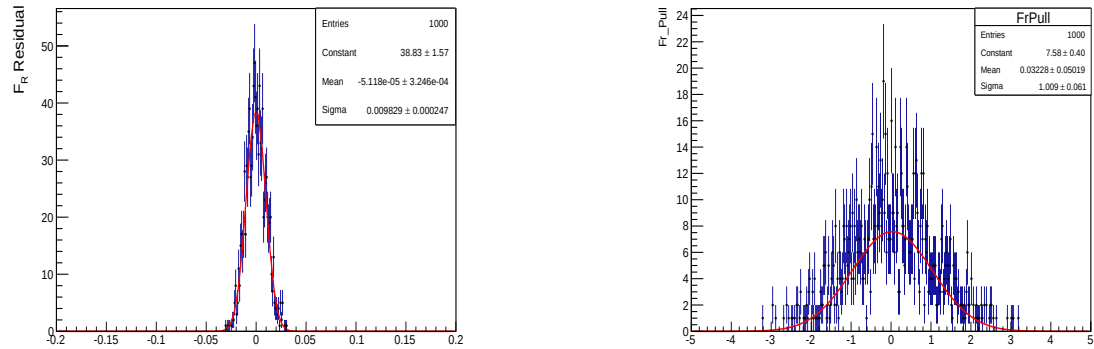


Figure 5.26: The residual (left) and pull (right) distributions for the right handed helicity fraction. Each distribution is fitted with a Gaussian function.

Chapter 6

Systematic uncertainties and results

In this chapter, after introducing main systematic sources, the final results of helicity fractions include the statistical and systematic uncertainties will be presented. The statistical fluctuation of data coming from the finite set of observed events can bias output results. On the other hand, systematic uncertainties are originated from the nature of the measurement. To make a reliable results at the end, the source of all uncertainties must be accounted properly. Several sources of systematic uncertainty are considered in this analysis which can be categorized into theoretical miss-modelling of the signal and background processes, the experimental setup and the reconstruction method.

For each systematic effect, the corresponding simulated pseudo-data sets by varying related variable are created. The effect of these variations are considered as one standard deviation in either upward or downward directions with respect to the central values. The up and down variations of each systematic source are reported as the related mean values of these varied pseudo-data distributions after subtracting from the nominal values. Finally by averaging over the up and down variations with respect to the nominal values, the effect of systematic uncertainty is obtained individually for each of systematic sources. The different systematic estimations are assumed to be uncorrelated with each other. Then, the total systematic uncertainty is computed by the quadratic sum of individual systematic values. In the following, a short description of the systematic sources as well as their estimations in the di-electron decay channel of $t\bar{t}$ system are reported.

6.1 Electron identification and trigger efficiencies

The methods to determine di-electron trigger scale factor as well as electron identification and isolation efficiencies were described with details in sections 5.3.1 and 5.3.2, respectively. The limited efficiency of the di-electron trigger leads to uncertainty on the trigger efficiency. The uncertainty on the trigger efficiency determined in different pseudorapidity bins of two electrons, around 1% (see figure 5.3).

In order to identify and to reconstruct electrons perfectly, uncertainties on the electron identification and isolation are also taken into account. These uncertainties are estimated in terms of the transverse momentum and pseudorapidity of electrons using the same method described in section 5.3.2. The resulting global efficiencies have a flat systematic trend with the nearly 1% variation in all of p_T and η bins. The overall systematic uncertainties on F_0 and F_L helicity fractions due to these efficiencies are less than 0.001.

6.2 B-tagging scale factors

The b-tagging efficiency and mis-tag rate as well as b-tagging scale factor were derived from the b-tagging at loose working point in section 5.3.3. They are provided as a function of the jet p_T from 20 to 800 GeV/c and $|\eta| < 2.4$. The individual scale factors for b, c and light flavours are varied separately within one standard deviation [123]. These variations lead to a new total b-tagging uncertainty using the same procedure described in section 5.3.3. The uncertainty due to b-tagging and non b-tagging efficiencies are evaluated separately. They are then added in quadrature to obtain the total b-tagging systematic uncertainty.

6.3 Backgrounds normalization

To account the uncertainties on backgrounds cross section, the tW and DY production cross sections are varied by 30% [135]. For the diboson and single top quark processes (t- and s-channels), an extreme variation of 50% is taken into account. Due to lower statistics of diboson and single top MC events, the corresponding uncertainties on the helicity measurements are negligible.

The total systematic uncertainty due to background normalizations is obtained by adding the uncertainties coming from each described background samples in quadrature.

6.4 Jet energy resolution and jet energy scale

The jet energy resolution, as described in section 4.3.6, is determined by matching the reconstructed jets in the calorimeter and the tracking system to the jet at the truth level. To estimate the impact of this uncertainty on the measurement of the W boson helicity fractions, each jet is smeared within the η -dependent JER uncertainties. For each case of up and down variations, the jet smearing effect is propagated into E_T^{miss} as new MET. Full top pair reconstruction is done for each nominal, scaled up and down scenarios and finally the related variations with respect to the nominal value are considered as the JER uncertainty.

The uncertainty on the jet energy scale is composed of different components [136] originating from the detector simulation, the calorimeter response and pileup effect. For each sources, the contribution is estimated by varying the energy of reconstructed jets within one standard deviation in bins of transverse momentum and pseudorapidity [102]. By assuming uncorrelated individual effects, the up and down jet energy scale uncertainties are evaluated by the quadratic sum of different components.

6.5 Pileup effect

As discussed in section 5.2.2, the MC events are re-weighted to make a good description in the number of reconstructed pileup events in the data. The effect of pileup re-weighting is checked by comparing the number of vertices in data and MC before and after applying pileup effect, shown in figure 5.1. In order to evaluate the systematic uncertainty on the pileup re-weighting, the same procedure as reference [137] is followed.

It's contribution is estimated by producing two additional samples with a variation of $\pm 5\%$ in the cross section of the minimum bias events. Based on the official recommendation, the selection of this variation covers all uncertainties due to physics and pileup modelling.

6.6 Integrated luminosity

The uncertainty on the integrated luminosity for the 2012 data taking is estimated to be 2.6% [138]. The effect of the uncertainty is investigated by varying the normalization factors in the signal and background samples simultaneously.

6.7 Top quark p_T re-weighting

The effect of the top p_T scale factor on the shape of the top p_T distribution was shown in figure 5.18. The full explanation of procedure was discussed in section 5.6.4. For this new scale factor, there is several source of uncertainties such as different selection of fitting function and also possible correlation between data points. According to [130], the systematic uncertainty of the top p_T re-weighting is evaluated by taking the differences on the measured helicity fractions when the weight is applied twice with respect to the case without applying any weight.

6.8 The model dependent uncertainties

6.8.1 Top quark mass

The nominal $t\bar{t}$ MC signal sample used in this analysis is generated with a top quark mass of $172.5 \text{ GeV}/c^2$. In order to evaluate systematic uncertainty due to the top quark mass variation, MC samples with different top quark mass values from $169 \text{ GeV}/c^2$ to $175 \text{ GeV}/c^2$ in steps of $2.5 \text{ GeV}/c^2$ are generated.

The whole procedure to extract helicity fractions in previous chapter is repeated on the new MC signal samples. The errors on the helicity fractions are estimated by interpolating the obtained mass uncertainties to the measured top quark mass uncertainty window $1 \text{ GeV}/c^2$ [139].

6.8.2 Factorization and renormalization scales

Similar to the top quark mass, the W boson helicity fractions are also affected by the different values for factorization and renormalization scales. In order to estimate the impact of these scales on the measurement presented in this analysis, the Q scale is varied by a factor of 0.5 and 2 with respect to the nominal default value. To do that, while new $t\bar{t}$ signal and DY+Jets background samples are used with the varied factorization scales, the other background simulated ones keep unchanged and same fitting procedure is repeated for both up and down variations.

6.8.3 Jet-parton matching scale

In addition to the factorization scale uncertainty, the uncertainty on the matching threshold for the matching between matrix element and parton shower in the generation of $t\bar{t}$ signal events is also taken into account. In order to evaluate jet-parton matching uncertainty, new $t\bar{t}$ simulated samples with the increased and decreased values of matching threshold were generated. Similar to the systematic case of

factorization scale, the nominal simulated samples are used for the background contributions and finally up and down helicity fractions are extracted with the same fit re-weighting method.

From the systematic results on the factorization and jet-parton matching scales, it can be concluded that the magnitude of these uncertainties is relatively large with respect to the contributions of other systematic sources.

6.8.4 Hadronization model

The MADGRAPH generator, interfaced with PYTHIA has been used as the baseline signal sample in this analysis. The impact of different hadronization modelling of the $t\bar{t}$ signal on the measured helicity fractions is studied using two new MC@NLO+HERWIG and POWHEG+PYTHIA simulated samples. The differences in the measured fractions between two processes are taken as the systematic uncertainty on the hadronization modelling.

6.8.5 Parton distribution function uncertainty

The uncertainty due to the parton distribution functions (PDFs), is evaluated using a re-weighting procedure with the help of 44 CTEQ66 error PDF sets [140] and the Lhapdf package [141]. In order to avoid sample generations for each of PDF sets, a re-weighting method is used. For each PDF sets, the simulated events are re-weighted and the final fit results are determined from the mean values of several produced pseudo-experiments. The results of re-weighted samples are compared with the nominal $t\bar{t}$ signal one, and the difference is taken as the systematic uncertainty. The total systematic uncertainty due to errors on the PDF is defined as the square root of the quadratically summed single contributions.

The contribution of this uncertainty is evaluated in the differential top pair cross section analysis [135]. In di-electron channel, the value of error propagation on the cross section is estimated to be 0.5%. For analysis presented in this thesis, the uncertainty due to the PDF is estimated by applying corresponding up and down varied top pair cross section values.

6.9 Limited statistics of signal sample

For some of systematic sources, the statistics of generated signal events are very low which leads to the additional source of systematic due to the lack of statistics for MC samples. The uncertainties from the lack of statistics for signal sample are estimated by dividing signal sample into N subsamples and performing the same re-weighting

procedure to extract the helicity fractions [41]. This procedure is repeated for many times, distribution of output helicity fractions is obtained at the end.

It is expected, the mean values of distributions correspond to the extracted fractions using the full sample. The systematic error is obtained by dividing the spread of distributions to the root mean squared number of subsamples $\sigma/\sqrt{N}$, and subtracting from the nominal values.

6.10 Systematics conclusion

The individual systematic uncertainty of each helicity fraction is estimated as the average value of up and down variations with respect to the nominal. For each helicity fractions, the single contributions are assumed to be uncorrelated and the total systematic uncertainty is therefore calculated as the square root of the quadratic sum of all the single contributions. By adding all single uncertainties in quadrature, we obtain a total systematic uncertainties of ± 0.65 and ± 0.48 for the longitudinal and left handed fractions, respectively. A summary of all various sources of systematic uncertainties in di-electron channel is found in table 6.1.

Table 6.1: Summary of systematic uncertainties originating from various sources.

Source of systematics	ΔF_L	ΔF_0
Lepton ID and trigger	<0.001	<0.001
b-tagging	0.001	0.002
Background normalisations	0.003	0.006
Jet energy resolution	0.003	0.003
Jet energy scale	0.006	0.015
Top p_T reweighting	0.005	0.010
Factorization/renormalization scales (signal)	0.035	0.052
Factorization/renormalization scales (DY)	0.006	0.010
Hadronization model	0.008	0.010
Jet-parton matching	0.030	0.030
Top mass ($\pm 1\text{GeV}/c^2$)	0.003	0.005
Pileup	0.001	0.002
PDF	<0.001	<0.001
Integrated luminosity	0.001	<0.001
Limited simulated signal statistics	0.005	0.008
Total uncertainty	0.048	0.065

As can be seen in table 6.1, the dominant estimated systematic uncertainty in this measurements are the factorization and normalization scales as well as the jet-parton matching.

By using the unitary constraint, the central value of the right handed helicity fraction is also obtained. The statistical and systematic uncertainties on the F_0 and the F_L polarizations are propagated to the measurement of the F_R helicity fraction. By taking into account the correlation between F_0 and F_L measurements for both of statistical and systematic uncertainties, the errors on the right handed fraction is written as:

$$\begin{aligned}\delta F_R^{stat} &= \sqrt{(\delta F_L^{stat})^2 + (\delta F_0^{stat})^2 + 2\rho^{stat}\delta F_L^{stat}\delta F_0^{stat}}, \\ \delta F_R^{syst} &= \sqrt{\sum_{i=1}^N [(\delta F_L^{syst;i})^2 + (\delta F_0^{syst;i})^2 + 2\rho_{ij}^{syst}\delta F_L^{syst;i}\delta F_0^{syst;j}]},\end{aligned}\quad (6.1)$$

where the summation is over single systematic contributions and N is expressed as the total number of systematic sources. The ρ^{stat} variable indicates the statistical correlation coefficient between the F_0 and the F_L polarizations which is found to be -0.92 for the di-electron channel. The parameter ρ_{ij}^{syst} denotes the ij matrix element of systematic correlation between these two helicity fractions for i th and j th systematic sources with the following definitions:

Table 6.2: The value of systematic correlation matrix elements for different variation conditions.

Definition	Condition
$\rho_{ij}^{syst} = 1$	$\delta^{up}F_0^{syst} \times \delta^{up}F_L^{syst} > 0, \delta^{down}F_0^{syst} \times \delta^{down}F_L^{syst} > 0$
$\rho_{ij}^{syst} = -1$	$\delta^{up}F_0^{syst} \times \delta^{up}F_L^{syst} > 0, \delta^{down}F_0^{syst} \times \delta^{down}F_L^{syst} < 0 $ $\delta^{up}F_0^{syst} \times \delta^{up}F_L^{syst} < 0, \delta^{down}F_0^{syst} \times \delta^{down}F_L^{syst} > 0$
$\rho_{ij}^{syst} = 0$	otherwise

In this analysis, the different sources of systematic uncertainties are assumed to be one of anti-correlated, uncorrelated or fully correlated cases with the corresponding -1, 0 or +1 values for the elements of systematic correlation matrix.

Figure 6.1 shows graphical comparisons between the longitudinal and the left handed helicity fractions including their uncertainties to the SM expectation value in the $F_0 - F_L$ plane for the di-electron and the combination channels.

These ellipses indicate the measurements within 95% confidence level for the full experimental uncertainty. The red and green dashed ellipses show the 2σ statistical and systematic uncertainties, and the black dashed line indicates the corresponding 2σ total error. As can be seen from figure 6.1, the measurements are consistent with the SM prediction.

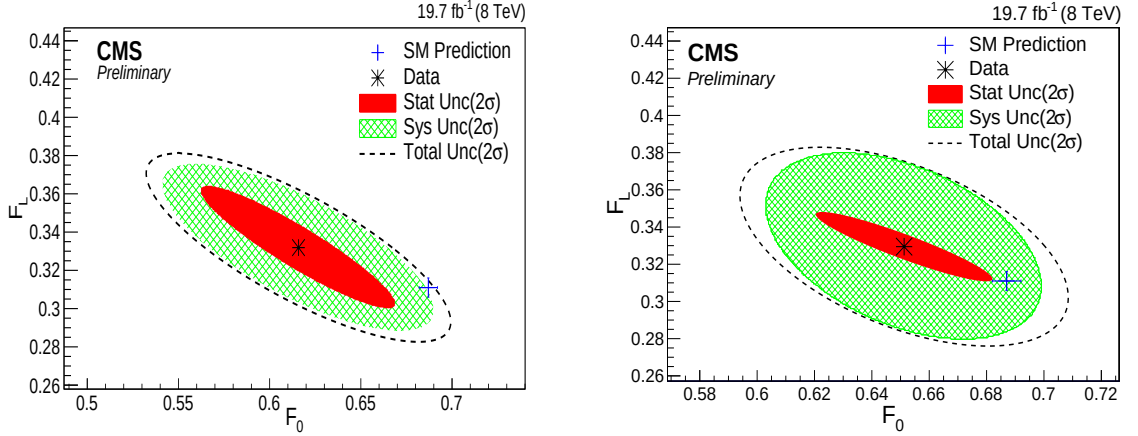


Figure 6.1: The 95% region in the (F_0-F_L) plane obtained from the fit to data for the di-electron (left) and the combination (right) channels. The measured and theoretical values of the W boson helicity fractions are shown as well.

6.11 Final results

The results of helicity fractions with corresponding statistical and systematic uncertainties in the di-electron and the combination of all dilepton channels are summarized in table 6.3.

Table 6.3: The measured W helicity fractions with corresponding uncertainties in the di-electron and the combination of all dilepton channels.

Channel	Result
e^-e^+	$F_0 = 0.617 \pm 0.037(\text{stat}) \pm 0.065(\text{syst})$
	$F_L = 0.330 \pm 0.022(\text{stat}) \pm 0.048(\text{syst})$
	$F_R = 0.053 \pm 0.019(\text{stat}) \pm 0.047(\text{syst})$
All channels	$F_0 = 0.653 \pm 0.016(\text{stat}) \pm 0.024(\text{syst})$
	$F_L = 0.329 \pm 0.009(\text{stat}) \pm 0.025(\text{syst})$
	$F_R = 0.018 \pm 0.008(\text{stat}) \pm 0.026(\text{syst})$

The measured W boson helicity fractions are consistent with the SM expectation at the NNLO within 2σ error bars. The statistical and systematic uncertainties on the right handed helicity fraction are estimated with the method explained in the previous section. Due to negative sign of statistical correlation for both channels, a lower statistical uncertainty on the F_R polarization with respect to the other two ones is obtained.

The results of measured values for all of dilepton channels containing di-electron, di-muon and electron-muon channels as well as combination of them are presented in figure 6.2.

To make an overall comparison, the other measurements performed by the CMS experiment in the $t\bar{t}$ semileptonic channel [40, 41] and the single top quark channels

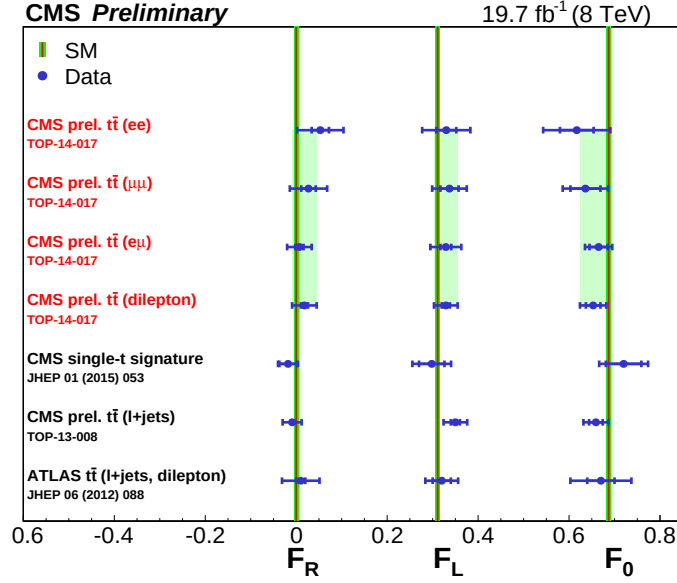


Figure 6.2: Measured W boson helicity fractions in the $t\bar{t}$ dilepton channels compared with results in different final states from the CMS experiment [40–42] and from the ATLAS experiment [39].

[42] are shown in this figure. In addition, the measured values of the helicity fractions from ATLAS [39] are also presented.

6.12 Conclusions

The W boson helicity fractions in the di-electron channel of the top pair events have been measured using a total integrated luminosity of 19.7 fb⁻¹ in proton-proton collisions. The observed events are collected by the CMS experiment during 2012 at the center-of-mass energy of 8 TeV. The sensitive observable used to measure the helicity fractions is the cosine of the decay angle θ_1^* . This variable is defined as the angle between the charged lepton momentum in the W boson rest frame and that of the W boson in the top quark rest frame. Since the measurement of $\cos \theta_1^*$ is based on the full top pair reconstruction, a $t\bar{t}$ kinematic reconstruction method has been developed.

This analysis focuses on the di-electron final state which is characterized by the presence of two oppositely charged isolated electrons, two b-tagged jets and missing transverse energy. To fully reconstruct the top pair system, several ambiguities in the reconstruction of two undetected neutrinos as well as the b-jet assignment are occurred. In order to select the best candidates for the W bosons and the top quarks, the AMWT method is used. By applying pileup re-weighting, electron identification and isolation, trigger efficiency, jet energy resolution and scale, b-tagging scale factor, and top p_T re-weighting good agreement between data and MC

is observed.

After fully kinematic reconstruction of the $t\bar{t}$ events, the $\cos\theta_1^*$ distribution is constructed. The resulting distribution is then subjected to the fit re-weighting approach. The W boson helicity fractions are extracted using binned likelihood method. Within the quoted statistical and systematic uncertainties, the different results indicate good agreement with the SM predictions at the 95% confidence level. Furthermore, the results are compared to other analyses in the CMS and ATLAS experiments, and it is found to be consistent. Precise measurements of the W boson helicity fractions will be provided soon with the center-of-mass energy of 13 TeV.

Bibliography

- [1] F. Mandl and G. Shaw, *Quantum Field Theory*. Wiley, 1984.
- [2] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*. Westview Press Reading (Mass.), 1995.
- [3] S. Weinberg, *The Quantum Theory of Fields. Vol. 1: Foundations*. Cambridge University Press, 1995.
- [4] A. Einstein, *Grundlage der Allgemeinen Relativit atstheorie*, *Annalen Phys.* **49** (1916) 769–822.
- [5] J. Beringer *et. al.*, *Particle Data Group. Review of particle physics*, *Phys. Rev. D* **86** (2012) 010001.
- [6] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (1964) 321–323.
- [7] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964) 508–509.
- [8] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, *Phys. Rev. Lett.* **13** (1964) 585–587.
- [9] **CMS** Collaboration, S. Chatrchyan *et. al.*, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys. Lett. B* **716** (2012) 30 [1207.7235v2].
- [10] **ATLAS** Collaboration, G. Aad *et. al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett. B* **716** (2012) 1 [1207.7214v2].
- [11] R. P. Feynman, *Mathematical Formulation of the Quantum Theory of Electro-magnetic Interaction*, *Phys. Rev.* **80** (1950) 440–457.
- [12] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, *Phys. Rev. Lett.* **10** (1963) 531–533.

- [13] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, *Prog. Theor. Phys.* **49** (1973) 652–657.
- [14] S. Glashow, *Partial Symmetries of Weak Interactions*, *Nucl. Phys.* **22** (1961) 579–588.
- [15] A. Salam and J. C. Ward, *Electromagnetic and weak interactions*, *Phys. Lett.* **13** (1964) 168–171.
- [16] S. Weinberg, *A Model of Leptons*, *Phys. Rev. Lett.* **19** (1967) 1264–1266.
- [17] W. James Stirling, R. Keith Ellis and B. R. Webber, *QCD and Collider Physics*. Cambridge University Press, 1996.
- [18] F. Abe *et. al.*, *Observation of top quark production in $\bar{p}p$ collisions*, *Phys. Rev. Lett.* **74** (1995) 2626–2631 [[hep-ex/9503002](#)].
- [19] S. Abachi *et. al.*, *Observation of the top quark*, *Phys. Rev. Lett.* **74** (1995) 2632–2637 [[hep-ex/9503003](#)].
- [20] **ATLAS, CDF, CMS and DØ** Collaboration, *First combination of Tevatron and LHC measurements of the top-quark mass*, [1403.4427v1](#).
- [21] T. Aaltonen *et. al.*, *Observation of Electroweak Single Top Quark Production*, *Phys. Rev. Lett.* **103** (2009) 092002.
- [22] V. M. Abazov *et. al.*, *Observation of Single Top-Quark Production*, *Phys. Rev. Lett.* **103** (2009) 092001.
- [23] J. C. Collins, D. E. Soper and G. Sterman, *Factorization for one-loop corrections in the Drell-Yan process*, *Nucl. Phys. B* **223** (1983) 381–421.
- [24] J. C. Collins and D. E. Soper, *Back-to-back jets in QCD*, *Nucl. Phys. B* **193** (1981) 381–443.
- [25] M. Czakon, P. Fiedler and A. Mitov, *Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$* , *Phys. Rev. Lett.* **110** (2013) 252004.
- [26] M. Fischer, S. Groote, J. G. Krner and M. C. Mauser, *Longitudinal, transverse-plus and transverse-minus W bosons in unpolarized top quark decays at $O(\alpha_S)$* , *Phys. Rev. D* **63** (2001) 031501.
- [27] A. Czarnecki, J. G. Korner and J. H. Piclum, *Helicity fractions of W bosons from top quark decays at NNLO in QCD*, *Phys. Rev. D* **81** (2010) 111503 [[1005.2625v1](#)].

- [28] J. Cao, R. J. Oakes, F. Wang and J. M. Yang, *Supersymmetric effects in top quark decay into polarized W boson*, *Phys. Rev. D* **68** (2003) 054019 [[hep-ph/0306278v2](#)].
- [29] W. Buchmuller and D. Wyler, *Effective Lagrangian Analysis Of New Interactions And Flavor Conservation*, *Nucl. Phys. B* **268** (1986) 621–653.
- [30] S. Weinberg, *Baryon- and Lepton-Nonconserving Processes*, *Phys. Rev. Lett.* **43** (1979) 1566.
- [31] E. Barberio *et. al.*, *Averages of b-hadron properties at the end of 2006*, [0704.3575v1](#).
- [32] J. A. Aguilar-Saavedra, J. Carvalho, N. Castro, F. Veloso and A. Onofre, *Probing anomalous Wtb couplings in top pair decays*, *Eur. Phys. J. C* **50** (2007) 519–533 [[hep-ph/0605190v2](#)].
- [33] G. L. Kane, G. A. Ladinsky and C. P. Yuan, *sing the top quark for testing standard-model polarization and CP predictions*, *Phys. Rev. D* **45** (1992) 124.
- [34] Z. Bern *et. al.*, *Left-Handed W Bosons at the LHC*, *Phys. Rev. D* **84** (2011) 034008 [[1103.5445v2](#)].
- [35] **CDF** Collaboration, T. Aaltonen *et. al.*, *W boson polarization measurement in the $t\bar{t}$ dilepton channel using the CDF II Detector*, *Phys. Lett. B* **722** (2013) 48–54 [[1205.0354v4](#)].
- [36] **CDF** Collaboration, T. Aaltonen *et. al.*, *Measurement of W-Boson Polarization in Top-quark Decay using the Full CDF Run II Data Set*, *Phys. Rev. D* **87** (2013) 03110 [[1211.4523v2](#)].
- [37] **DØ** Collaboration, V. M. Abazov *et. al.*, *Measurement of the W boson helicity in top quark decays using 5.4 fb⁻¹ of ppbar collision data*, *Phys. Rev. D* **83** (2011) 032009 [[1011.6549v2](#)].
- [38] **CDF and DØ** Collaboration, T. Aaltonen *et. al.*, *Combination of CDF and DØ measurements of the W boson helicity in top quark decays*, *Phys. Rev. D* **85** (2012) 071106 [[1202.5272v1](#)].
- [39] **ATLAS** Collaboration, G. Aad *et. al.*, *Measurement of the W boson polarization in top quark decays with the ATLAS detector*, *JHEP* **1206** (2012) 088 [[1205.2484v1](#)].

- [40] **CMS** Collaboration, S. Chatrchyan *et. al.*, *Measurement of the W-boson helicity in top-quark decays from $t\bar{t}$ production in lepton+jets events in pp collisions at $\sqrt{s} = 7\text{ TeV}$* , *JHEP* **10** (2013) 167 [1308.3879v2].
- [41] **CMS** Collaboration, *Measurement of the W-boson helicity in top decays from $t\bar{t}$ production in lepton+jets events at the LHC at $\sqrt{s} = 8\text{ TeV}$* , Tech. Rep. CMS-PAS-TOP-13-008, CERN, Geneva, 2013.
- [42] **CMS** Collaboration, S. Chatrchyan *et. al.*, *Measurement of the W boson helicity in events with a single reconstructed top quark in pp collisions at $\sqrt{s} = 8\text{ TeV}$* , *JHEP* **01** (2015) 053 [1410.1154v2].
- [43] **CMS** Collaboration, *Measurement of the W boson helicity using $t\bar{t}$ events in the dilepton final state at $\sqrt{s} = 8\text{ TeV}$* , Tech. Rep. CMS-PAS-TOP-14-017, CERN, Geneva, 2015.
- [44] L. Evans and P. Bryant, *LHC Machine*, *JINST* **03** (2015) S08001.
- [45] **CMS** Collaboration, G. L. Bayatian *et. al.*, *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*. Technical Design Report CMS. CERN, Geneva, 2006.
- [46] **CMS** Collaboration, G. L. Bayatian *et. al.*, *CMS Physics Technical Design Report, Volume II: Physics Performance*, *J. Phys. G: Nucl. Part. Phys.* **34** (2007) 995.
- [47] <https://cmsddt.cern.ch/SDT/doxygen>.
- [48] **ATLAS** Collaboration, G. Aad *et. al.*, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003.
- [49] **CMS** Collaboration, S. Chatrchyan *et. al.*, *The CMS experiment at the CERN LHC*, *JINST* **3** (2008) S08004.
- [50] **ALICE** Collaboration, K. Aamodt *et. al.*, *The ALICE experiment at the CERN LHC*, *JINST* **3** (2008) S08002.
- [51] **LHCb** Collaboration, A. A. Alves Jr *et. al.*, *The LHCb Detector at the LHC*, *JINST* **3** (2008) S08005.
- [52] **LHCf** Collaboration, O. Adriani *et. al.*, *The LHCf detector at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08006.
- [53] **TOTEM** Collaboration, G. Anelli *et. al.*, *The TOTEM Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08007.

- [54] **CMS** Collaboration, *The CMS Tracker System Project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997.
- [55] **CMS** Collaboration, *The CMS Tracker: Addendum to the Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 2000.
- [56] H. Behnamian, *Sensor RD for the CMS Outer Tracker Upgrade for the HL-LHC*, *JINST* **9** (2014) C04033.
- [57] **CMS** Collaboration, *The CMS electromagnetic calorimeter project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997.
- [58] **CMS** Collaboration, *The CMS hadron calorimeter project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997.
- [59] **CMS** Collaboration, *The CMS muon project : Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997.
- [60] **CMS** Collaboration, *CMS TriDAS project : Technical Design Report, Volume 1: The Trigger Systems*. Technical Design Report CMS. CERN, Geneva, 2000.
- [61] **CMS** Collaboration, *CMS The TriDAS Project : Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger*. Technical Design Report CMS. CERN, Geneva, 2000.
- [62] **CMS** Collaboration, W. Adam *et. al.*, *The CMS high level trigger*, *EPJC* **46** (2006) 605–667.
- [63] C. Eck *et. al.*, *LHC computing Grid : Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 2005.
- [64] F. Gianotti *et. al.*, *Physics potential and experimental challenges of the LHC luminosity upgrade*, *Eur. Phys. J. C* **39** (2005) 293.
- [65] S. Muller, *The Beam Condition Monitor and the Radiation Environment of the CMS Detector at the LHC*. PhD thesis, Karlsruhe Institute of Technology, 2011.
- [66] A. Dierlamm, *Characterisation of silicon sensor materials and designs for the CMS Tracker Upgrade*, Tech. Rep. AIDA-CONF-2014-006, CERN, Geneva, Sep, 2012.

- [67] A. Dierlamm, *Silicon sensor developments for the CMS Tracker upgrade*, *JINST* **7** (2012) C01110.
- [68] M. Moll, *Development of radiation tolerant semiconductor detectors for the Super-LHC*, *Nucl. Instr. Meth. A.* **546** (2005) 99.
- [69] W. Ferguson *et. al.*, *The CBC microstrip readout chip for CMS at the High Luminosity LHC*, *JINST* **7** (2012) C08006.
- [70] G. Casse *et. al.*, *Measurements of charge collection efficiency with microstrip detectors made on various substrates after irradiations with neutrons and protons with different energies*, *PoS (VERTEX 2008)* (2008) 036.
- [71] G. Casse *et. al.*, *Evidence of enhanced signal response at high bias voltages in planar silicon detectors irradiated up to 2.2×10^{16} n_{eq}/cm²*, *Nucl. Instr. Meth. A.* **636** (2011) S56.
- [72] A. Bhardwaj *et. al.*, *Development of Radiation Damage Models for Irradiated Silicon Sensors Using TCAD Tools*, Tech. Rep. CMS-CR-2014-120, CERN, Geneva, Jun, 2014.
- [73] T. Sjostrand, *Monte Carlo Tools*, 0911.5286v1.
- [74] M. H. Seymour and M. Marx, *Monte Carlo Event Generators*, 1304.6677v1.
- [75] <https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookCMSSWFramework>.
- [76] M. Dobbs and J. B. Hansen, *The HepMC C++ Monte Carlo Event Record for High Energy Physics*, *Comput. Phys. Commun.* **134** (2001) 41–46.
- [77] V. N. Gribov and L. N. Lipatov, *Deep inelastic ep scattering in perturbation theory*, *Sov. J. Nucl. Phys.* **15** (1972) 438–450.
- [78] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, *Nucl. Phys. B* **126** (1977) 298.
- [79] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^- Annihilation by Perturbation Theory in Quantum Chromodynamics*, *Sov. Phys. JETP.* **46** (1977) 641–653.
- [80] B. Andersson *et. al.*, *The Lund Fragmentation Process for a Multi-gluon String According to the Area Law*, *Eur. Phys. J.* **21** (2001) 631–647 [[hep-ph/0106185v1](#)].

- [81] B. R. Webber, *A QCD Model for Jet Fragmentation Including Soft Gluon Interference*, *Nucl. Phys. B* **238** (1984) 492.
- [82] D. Amati and G. Veneziano, *Preconfinement as a Property of Perturbative QCD*, *Phys. Lett. B* **83** (1979) 87.
- [83] P. Skands, *QCD for Collider Physics*, 1104.2863v2.
- [84] **GEANT4** Collaboration, S. Agostinelli *et. al.*, *GEANT4: a simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250303.
- [85] J. Alwall *et. al.*, *MadGraph 5 : Going Beyond*, *JHEP* **06** (2011) 128 [1106.0522v1].
- [86] T. Sjostrand, P. Skands and S. Mrenna, *PYTHIA 6.4 Physics and Manual*, *JHEP* **05** (2006) 026 [hep-ph/0603175v2].
- [87] G. Corcella *et. al.*, *HERWIG 6.5: an event generator for Hadron Emission Reactions With Interfering Gluons (including supersymmetric processes)*, *JHEP* **01** (2001) 010 [hep-ph/0011363v3].
- [88] M. Mangano *et. al.*, *Matching matrix elements and shower evolution for the top-quark production in hadronic collisions*, *JHEP* **01** (2007) 013 [hep-ph/0611129v1].
- [89] S. Catani, F. Krauss, R. Kuhn and B. R. Webber, *QCD Matrix Elements + Parton Showers*, *JHEP* **0111** (2001) 063 [hep-ph/0109231v1].
- [90] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, *JHEP* **11** (2007) 070 [0709.2092v1].
- [91] S. Frixione *et. al.*, *The MC@NLO 4.0 Event Generator*, 1010.0819v1.
- [92] **CMS** Collaboration, *Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector*, Tech. Rep. CMS-PAS-PFT-10-001, 2010.
- [93] T. Speer *et. al.*, *Track Reconstruction in the CMS Tracker*, *Nucl. Instrum. Meth. A* **559** (2006) 143–147.
- [94] R. Fruhwirth, *Application of Kalman Filtering to Track and Vertex Fitting*, *Nucl. Instrum. Meth. A* **262** (1987) 444–450.
- [95] S. Baffioni *et. al.*, *Electron reconstruction in CMS*, *Eur. Phys. J. C* **49** (2006) 1099–1116.

- [96] W. Adam, R. Fruhwirth, A. Strandlie and T. Todorov, *Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC*, *J. Phys. G: Nucl. Part. Phys.* **31** (2005) N9.
- [97] S. Chatrchyan *et. al.*, *Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7\text{ TeV}$* , *JINST* **7** (2012) P10002 [1206.4071v2].
- [98] G. P. Salam and G. Soyez, *A practical Seedless Infrared-Safe Cone jet algorithm*, *JHEP* **5** (2007) 86 [0704.0292v2].
- [99] S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, *Longitudinally Invariant K_t Clustering Algorithms for Hadron Hadron Collisions*, *Nucl. Phys. B* **406** (1993) 187224.
- [100] Y. L. Dokshitzer, G. D. Leder, S. Moretti and B. R. Webber, *Better Jet Clustering Algorithms*, *JHEP* **08** (1997) 001 [hep-ph/9707323v2].
- [101] M. Cacciari, G. P. Salam and G. Soyez, *The Anti- K_t Jet Clustering Algorithm*, *JHEP* **04** (2008) 063 [0802.1189v2].
- [102] CMS Collaboration, S. Chatrchyan *et. al.*, *Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS*, *JINST* **6** (2011) 11002 [1107.4277v1].
- [103] M. Cacciari, G. P. Salam and G. Soyez, *FastJet User Manual*, *Eur. Phys. J.* **C72** (2012) 1896 [1111.6097v1].
- [104] A. B. Galtieri, F. Margaroli and I. Volobouev, *Precision measurements of the top quark mass from the Tevatron in the pre-LHC era*, 1109.2163v2.
- [105] N. Dhingra, *Precision measurements of the top quark mass from the Tevatron in the pre-LHC era*, 1409.0251v1.
- [106] CMS Collaboration, *Performance of b tagging at $\sqrt{s} = 8\text{ TeV}$ in multijet, $t\bar{t}$ and boosted topology events*, Tech. Rep. CMS-PAS-BTV-13-001, CERN, Geneva, 2013.
- [107] CMS Collaboration, *Performance of the Missing Transverse Energy Reconstruction in the 8 TeV Data*, Tech. Rep. CMS-PAS-JME-12-002, CERN, Geneva, 2013.
- [108] H. Behnamian, M. Khakzad, M. Mohammadi Najafabadi and M. Naseri, *Measurement of the W boson helicity using $t\bar{t}$ events in the dilepton final state at $\sqrt{s} = 8\text{ TeV}$* , Tech. Rep. CMS-AN-14-052, 2015.

- [109] N. Davidson *et. al.*, *Universal Interface of TAUOLA Technical and Physics Documentation*, *Comput. Phys. Commun.* **183** (2010) 821 [1002.0543v1].
- [110] **CMS** Collaboration, *Measurement of the Top Quark Mass in Dilepton $t\bar{t}$ Decays at $\sqrt{s} = 8$ TeV*, Tech. Rep. CMS-TOP-2014-010, CERN, Geneva, 2014.
- [111] J. Pumplin *et. al.*, *New generation of parton distributions with uncertainties from global QCD analysis*, *JHEP* **07** (2002) 012 [hep-ph/0201195v3].
- [112] N. Kidonakis, *NNLL threshold resummation for top-pair and single-top production*, *Phys. Part. Nucl.* **45** (2014) 4 [1210.7813v2].
- [113] <https://twiki.cern.ch/twiki/bin/view/CMS/TopEventSelection>.
- [114] <https://twiki.cern.ch/twiki/bin/view/CMS/EgammaEARhoCorrection>.
- [115] <https://twiki.cern.ch/twiki/bin/view/CMS/ConversionBackgroundRejection>.
- [116] **CMS** Collaboration, S. Chatrchyan *et. al.*, *Measurement of $t\bar{t}$ production cross section in the dilepton channel in pp collisions at $\sqrt{s} = 8$ TeV*, *JHEP* **02** (2014) 024 [1312.7582v2].
- [117] <https://twiki.cern.ch/twiki/bin/view/CMS/JetID>.
- [118] S. Chatrchyan *et. al.*, *Results on b -tagging identification in 8 TeV pp collisions*, Tech. Rep. CMS-DP-2013-005, CERN, Geneva, 2013.
- [119] **CMS** Collaboration, S. Chatrchyan *et. al.*, *Measurement of differential top-quark pair production cross sections in pp collisions at $\sqrt{s} = 7$ TeV*, *Eur. Phys. J. C* **73** (2013) 2339 [1211.2220v3].
- [120] **CMS** Collaboration, V. Khachatryan *et. al.*, *Measurement of the differential cross section for top quark pair production in pp collisions at $\sqrt{s} = 8$ TeV*, 1505.04480v1.
- [121] C. Diez Pardos and J. Kieseler, *Dilepton trigger and lepton identification efficiencies for the top quark pair production cross section measurements at 8 TeV in the dilepton decay channel*, Tech. Rep. CMS-AN-2012-389, CERN, Geneva, 2012.
- [122] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BTagSFMethods>.
- [123] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation53XReReco>.

- [124] W. Andrews *et. al.*, *A Method to Measure the Contribution of $DY \rightarrow l + \bar{l}$ to a di-lepton + MET Selection*, Tech. Rep. CMS-AN-2009-023, CERN, Geneva, 2009.
- [125] L. Sonnenschein, *Analytical solution of $t\bar{t}$ dilepton equations*, *Phys. Rev. D* **73** (2006) 054015 [[hep-ph/0603011v3](#)].
- [126] <https://github.com/gerbaudo/ttbar-kinsol-comp/tree/master/dileptonSolver>.
- [127] CMS Collaboration, *Measurement of the W helicity in top pair production with dileptons at 7 TeV*, Tech. Rep. CMS-PAS-TOP-12-015, CERN, Geneva, 2013.
- [128] CMS Collaboration, S. Chatrchyan *et. al.*, *Measurements of the $t\bar{t}$ charge asymmetry using the dilepton decay channel in pp collisions at $\sqrt{s} = 7$ TeV*, *JHEP* **04** (2014) 191 [[1402.3803v2](#)].
- [129] R. H. Dalitz and G. R. Goldstein, *Decay and polarization properties of the top quark*, *Phys. Rev. D* **45** (1992) 1531.
- [130] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/TopPtReweighting>.
- [131] N. Kidonakis, *NNLL threshold resummation for top-pair and single-top production*, *Phys. Part. Nucl.* **45** (2014) 714 [[1210.7813v2](#)].
- [132] CMS Collaboration, S. Chatrchyan *et. al.*, *Measurement of the W boson helicity in events with a single reconstructed top quark in pp collisions at $\sqrt{s} = 8$ TeV*, *JHEP* **01** (2015) 053 [[1410.1154v2](#)].
- [133] V. Blobel, *An Unfolding Method for High Energy Physics Experiments*, [hep-ex/0208022v1](#).
- [134] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, *Comput. Phys. Commun* **10** (1975) 343.
- [135] CMS Collaboration, *Measurement of the differential top-quark pair production cross section in the dilepton channel in pp collisions at $\sqrt{s} = 8$ TeV*, Tech. Rep. CMS-TOP-2012-028, CERN, Geneva, 2012.
- [136] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/JECUncertaintySources>.
- [137] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/PileupInformation>.

- [138] **CMS** Collaboration, *CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update*, Tech. Rep. CMS-PAS-LUM-13-001, CERN, Geneva, 2013.
- [139] **CMS** Collaboration, *Combination of the CMS top-quark mass measurements from Run 1 of the LHC*, Tech. Rep. CMS-PAS-TOP-14-015, CERN, Geneva, 2014.
- [140] P. M. Nadolsky *et. al.*, *Implications of CTEQ global analysis for collider observables*, *Phys. Rev. D* **78** (2008) 013004 [0802.0007v3].
- [141] M. R. Whalley, D. Bourilkov and R. C. Group, *The Les Houches Accord PDFs (LHAPDF) and Lhaglu*, hep-ph/0508110v1.

Appendix A

Amplitude calculation of top quark decay

The Lorentz-invariant amplitude of top quark decay can be written as:

$$i\mathcal{M} = \frac{ig}{\sqrt{2}} \bar{b}(q) \gamma^\mu \left(\frac{1 - \gamma^5}{2} \right) t(p) \epsilon_\mu^*(k). \quad (\text{A.1})$$

The spin-averaged square of the matrix element is given by:

$$\langle |\mathcal{M}|^2 \rangle = \frac{1}{2} \frac{1}{3} \sum_{\text{quark spins}} \sum_{W \text{ polarizations}} |\mathcal{M}|^2, \quad (\text{A.2})$$

where the first sum is over spin states of the top and the b-quarks while the second sum is corresponding to all possible W boson helicity configurations. By substituting equation A.1 in above formula, we have:

$$\begin{aligned} \langle |\mathcal{M}|^2 \rangle &= \frac{g^2}{2} \sum_{\text{spins}} [\bar{b}(q) \gamma^\mu \left(\frac{1 - \gamma^5}{2} \right) t(p)] [\bar{b}(q) \gamma^\nu \left(\frac{1 - \gamma^5}{2} \right) t(p)]^\dagger \times \\ &\quad \frac{1}{3} \sum_{W \text{ polarizations}} \epsilon_\mu^*(k) \epsilon_\nu(k), \end{aligned} \quad (\text{A.3})$$

gives

$$\begin{aligned} \langle |\mathcal{M}|^2 \rangle &= \frac{g^2}{2} \sum_{\text{spins}} [\bar{b}(q) \gamma^\mu \left(\frac{1 - \gamma^5}{2} \right) t(p)] [\bar{t}(p) \left(\frac{1 + \gamma^5}{2} \right) \gamma^\nu b(q)] \times \frac{1}{3} \sum_{W \text{ polarizations}} \epsilon_\mu^*(k) \epsilon_\nu(k), \\ &= \frac{g^2}{2} \sum_{r,s} [\bar{b}_i^r(q) \gamma_{ik}^\mu \left(\frac{1 - \gamma^5}{2} \right)_{kj} t_j^s(p)] [\bar{t}_l^s(p) \left(\frac{1 + \gamma^5}{2} \right)_{lm} \gamma_{mn}^\nu b_n^r(q)] \times \\ &\quad \frac{1}{3} \sum_{W \text{ polarizations}} \epsilon_\mu^*(k) \epsilon_\nu(k), \end{aligned} \quad (\text{A.4})$$

Since all the quantities in equation A.4 are just numbers, with the indices keeping track of the matrix multiplication, this can be written as:

$$\begin{aligned} \langle |\mathcal{M}|^2 \rangle &= \frac{g^2}{2} [\gamma_{ik}^\mu (\frac{1-\gamma^5}{2})_{kj} (\frac{1+\gamma^5}{2})_{lm} \gamma_{mn}^\nu] [\sum_{r=1}^2 b_n^r(q) \bar{b}_i^r(q)] [\sum_{s=1}^2 t_j^s(p) \bar{t}_l^s(p)] \times \\ &\quad \frac{1}{3} \sum_{W \text{ polarizations}} \epsilon_\mu^*(k) \epsilon_\nu(k), \end{aligned} \quad (\text{A.5})$$

Using the completeness relation for spinors, the top quark decay amplitude becomes:

$$\begin{aligned} \langle |\mathcal{M}|^2 \rangle &= \frac{g^2}{2} [\gamma_{ik}^\mu (\frac{1-\gamma^5}{2})_{kj} (\not{p} + m_t)_{jl} (\frac{1+\gamma^5}{2})_{lm} \gamma_{mn}^\nu (\not{q} + m_b)_{ni}] \times \\ &\quad \frac{1}{3} \sum_{W \text{ polarizations}} \epsilon_\mu^*(k) \epsilon_\nu(k), \\ &= \frac{g^2}{2} \text{Tr} [\gamma^\mu (\frac{1-\gamma^5}{2}) (\not{p} + m_t) (\frac{1+\gamma^5}{2}) \gamma^\nu (\not{q} + m_b)] \times \\ &\quad \frac{1}{3} \sum_{W \text{ polarizations}} \epsilon_\mu^*(k) \epsilon_\nu(k). \end{aligned} \quad (\text{A.6})$$

This can be evaluated by writing $\not{p} = \gamma^\alpha p_\alpha$ and $\not{q} = \gamma^\beta q_\beta$. The calculation of the spin-averaged square of the matrix element is reduced to some of trace calculations involve combinations of γ -matrices. By using trace theorems, this can be written as:

$$\langle |\mathcal{M}|^2 \rangle = g^2 [p^\mu q^\nu + p^\nu q^\mu - g^{\mu\nu} p \cdot q - i \epsilon^{\mu\nu\alpha\beta} p_\alpha q_\beta] \times \frac{1}{3} \sum_{W \text{ polarizations}} \epsilon_\mu^*(k) \epsilon_\nu(k). \quad (\text{A.7})$$

Appendix B

Data-MC comparison plots for combination channel

The full analysis procedure, described in chapters 5 and 6 has been done for other dilepton channels like di-muon and electron-muon. The number of observed compared to the expected simulated events after each consecutive selection steps are given in table B.1.

Table B.1: Number of events for each MC samples, Data and total MC in different steps for combination of all dilepton channels. The statistical uncertainty on MC samples are also included in this table.

Combination channel	≥ 2 leptons	≥ 2 jets	MET	≥ 2 b-tag	$t\bar{t}$
Diboson	21470.5 ± 48.2	1299.0 ± 10.0	850.1 ± 8.8	49.9 ± 2.2	29.8 ± 1.6
Wjets	3161.3 ± 180.0	266.1 ± 52.5	220.7 ± 47.8	19.1 ± 13.6	10.3 ± 10.4
DY	1315520 ± 1811.4	44376.8 ± 334.3	10137.7 ± 174.5	605.6 ± 44.2	420.4 ± 36.3
tW	9497.5 ± 67.7	3111.2 ± 38.9	2815.9 ± 37.0	1054.4 ± 22.5	677.4 ± 17.9
SingleTop	159.5 ± 7.4	26.7 ± 3.0	23.2 ± 2.8	6.4 ± 1.5	4.0 ± 1.1
$t\bar{t}$ other	568.7 ± 6.6	457.1 ± 5.9	402.2 ± 5.5	126.8 ± 3.1	102.6 ± 2.8
$t\bar{t}$ signal	99205.6 ± 68.6	67723.3 ± 56.5	61281.2 ± 54.0	33350.9 ± 39.5	29840.9 ± 37.1
Sum MC	1449580 ± 1823.6	117260 ± 345.5	75731.1 ± 192.7	35213.1 ± 65.0	31085.5 ± 55.9
Data	1530445	122382	75544	36266	31881

Several control plots of kinematic observables are presented in the following. The kinematic properties of selected leptons, p_T and η , are shown in figure B.1. In figure B.2, the expected number of selected jets and the b-tagged jet multiplicity are displayed. The comparison of the data and the MC for the transverse momentum of the first and the second leading jets are presented in figure B.3. Figure B.4, provides the dilepton invariant mass and the E_T^{miss} distributions.

In the following figures, the hatched areas correspond to the systematic uncertainty on the MC simulation, explained in the chapter 6. Figures B.5 and B.6, show the kinematic variables of W bosons and top quarks, respectively. In figure B.7, the distribution of $t\bar{t}$ invariant mass for the combination of dilepton channels is presented.

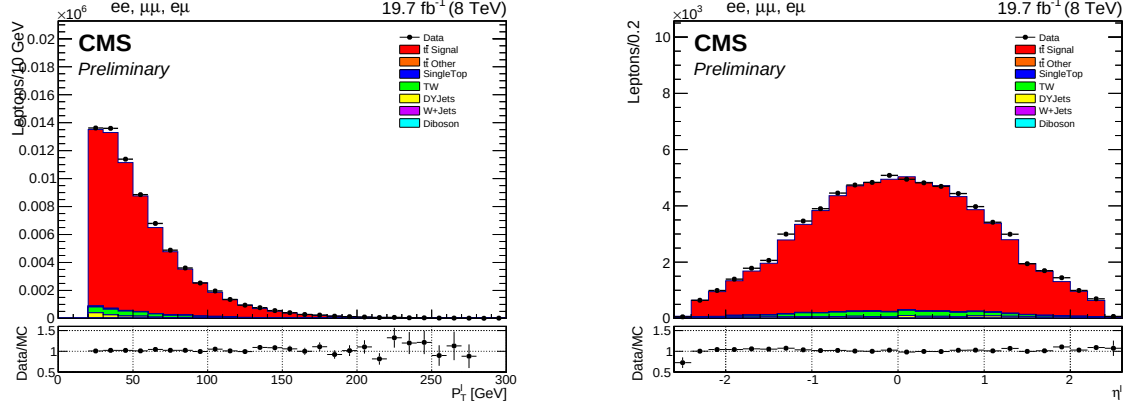


Figure B.1: The lepton p_T (left) and η (right) distributions in combination of all dilepton channels after requesting two b-tagged jets. At this step, all of described correction scale factors are applied on the MC simulation events.

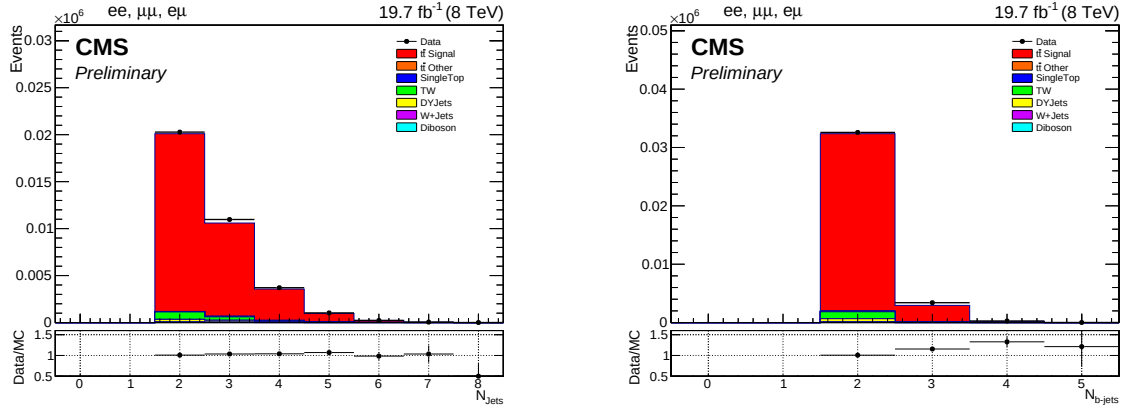


Figure B.2: The jet multiplicity (left) and b-tagged jet multiplicity (right) distributions in combination of all dilepton channels after requesting two b-tagged jets. At this step, all of described correction scale factors are applied on the MC simulation events.

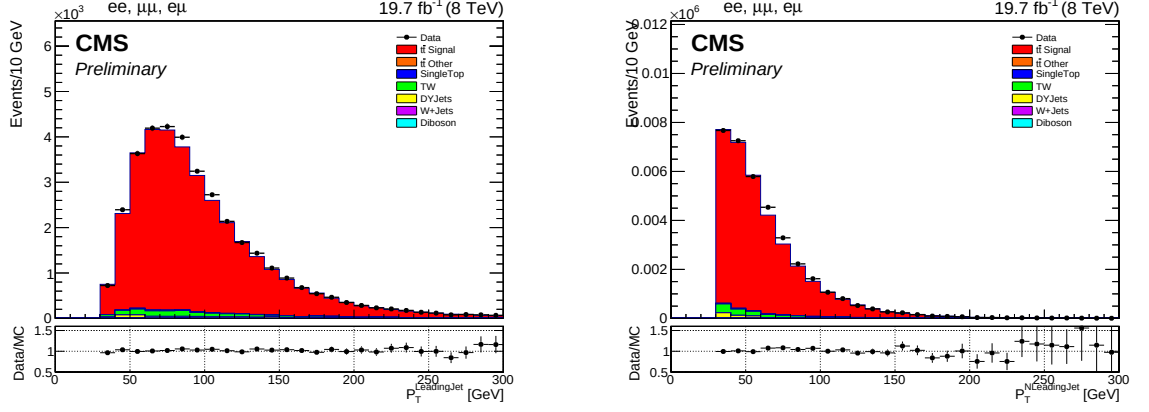


Figure B.3: The p_T distributions for first (left) and second (right) leading jet in combination of all dilepton channels after requesting two b-tagged jets. At this step, all of described correction scale factors are applied on the MC simulation events.

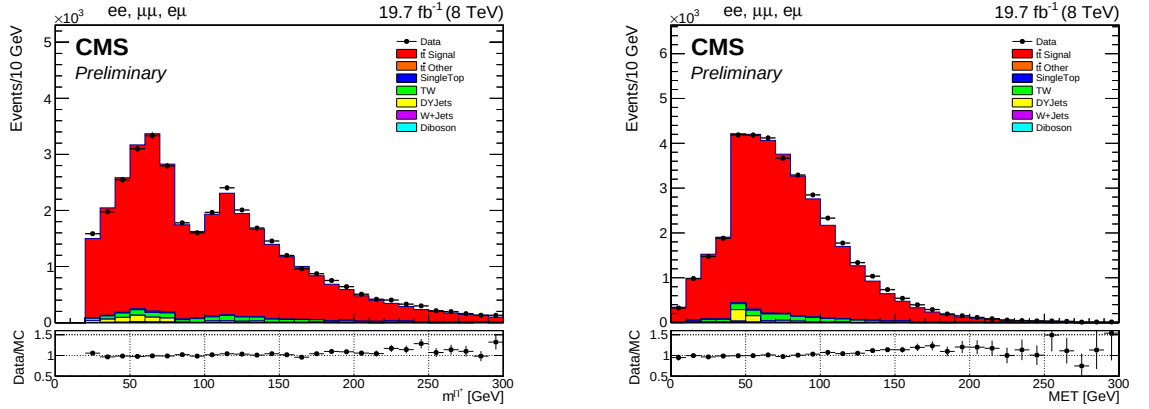


Figure B.4: The distributions of oppositely charged dilepton invariant mass (left) and missing transverse energy (right) in combination of all dilepton channels after requesting two b-tagged jets. At this step, all of described correction scale factors are applied on the MC simulation events.

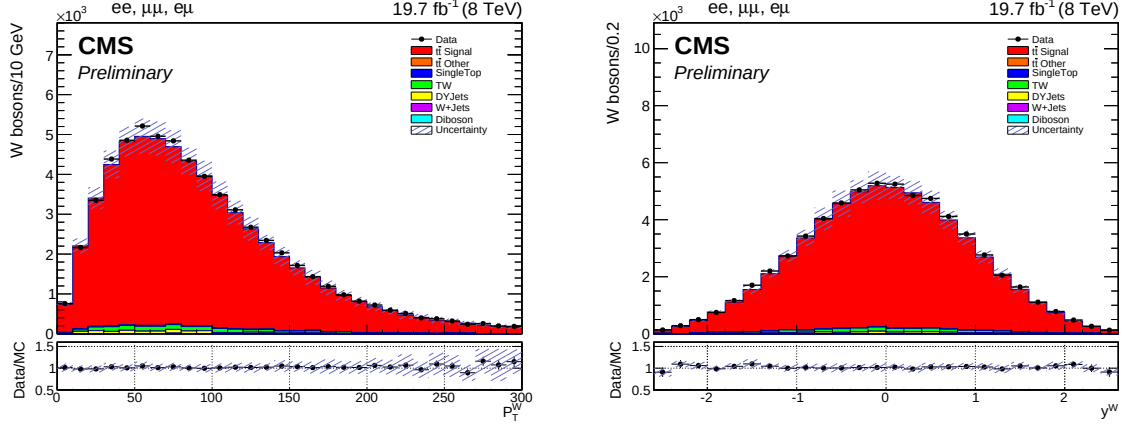


Figure B.5: The transverse momentum (left) and rapidity (right) distributions of W boson in combination of all dilepton channels after full kinematic reconstruction of $t\bar{t}$ events. All described correction scale factors are applied on the MC simulation events. The hatched areas show the systematic uncertainties on the MC events.

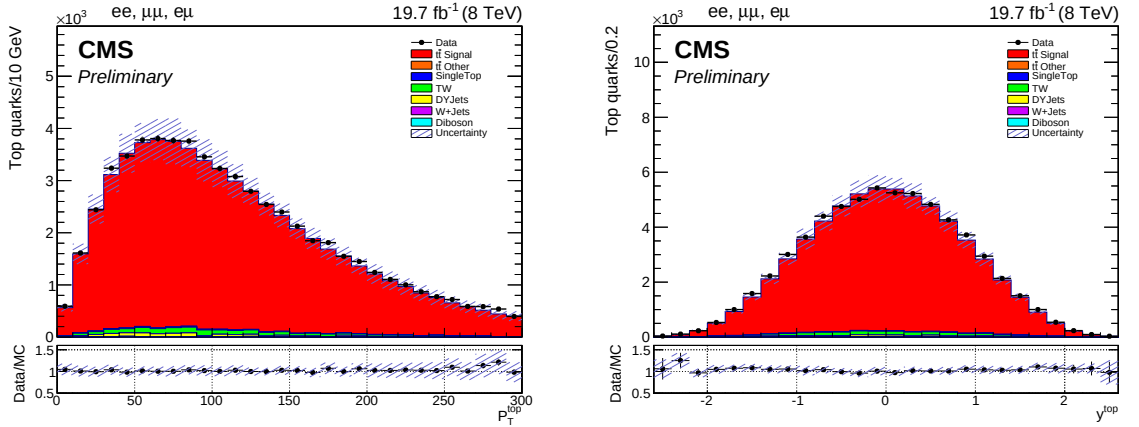


Figure B.6: The distributions of top p_T (left) and top rapidity (right) in combination of all dilepton channels after full kinematic reconstruction of $t\bar{t}$ events. All described correction scale factors are applied on the MC simulation events. The hatched areas show the systematic uncertainties on the MC events.

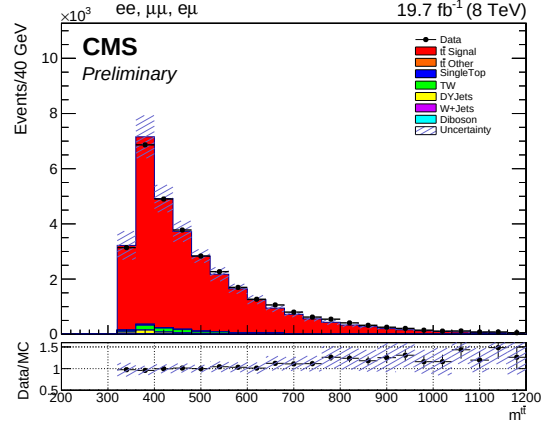


Figure B.7: The distributions of $t\bar{t}$ invariant mass for combination of all channels after full kinematic reconstruction of $t\bar{t}$ events. All described correction scale factors are applied on the MC simulation events. The hatched areas show the systematic uncertainties on the MC events.

One can clearly see from these distributions that the reasonable agreement between data and MC within the statistical and systematic uncertainties is obtained.