



**1 A search for flavour changing neutral currents  
2 involving a top quark and a Z boson, using the  
3 data collected by the CMS collaboration at a  
4 centre-of-mass energy of 13 TeV**

**5 Isis Van Parijs**

**6 Proefschrift ingediend met het oog op het behalen van de academische graad  
7 Doctor in de Wetenschappen.**

Published in    Faculteit Wetenschappen & Bio-ingenieurswetenschappen  
                    Vrije Universiteit Brussel  
At                 10 November 2017.

**8**

Responsible Contact:    Isis Van Parijs  
                            Intitute for High Energy Physics  
Promotor:               Prof. Jorgen D'Hondt

9      First Referee:                            Prof. Dr. J. D'Hondt  
            Date of Hand-in:                        10 November 2017  
            Date of Defense:                        10 December 2017

# Contents

---

11	<b>1 Theoretical basis</b>	1
12	1.1 Elementary particles and forces . . . . .	1
13	1.2 Standard Model Lagrangian, connecting fields with particles . . . . .	3
14	1.3 Flavours in the SM . . . . .	5
15	1.4 The top quark in the SM . . . . .	6
16	1.5 Effective field theories . . . . .	11
17	1.6 Motivation for new physics . . . . .	14
18	1.7 An effective approach beyond the SM: FCNC involving a top quark . . . . .	15
19	1.8 Experimental constraints on top-FCNC . . . . .	18
20	<b>2 Experimental set-up</b>	23
21	2.1 The Large Hadron Collider . . . . .	23
22	2.2 The Compact Muon Solenoid . . . . .	27
23	2.2.1 CMS coordinate system . . . . .	27
24	2.2.2 Towards the heart of CMS . . . . .	28
25	2.2.3 Data acquisition . . . . .	37
26	2.2.4 Phase 1 upgrades . . . . .	37
27	2.2.5 CMS computing model . . . . .	39
28	<b>3 Analysis techniques</b>	41
29	3.1 Hadron collisions at high energies . . . . .	41
30	3.2 Event generation . . . . .	44
31	3.2.1 Fundamentals of simulating a proton collision . . . . .	44
32	3.2.2 Programs for event generation . . . . .	44
33	3.2.3 Generating FCNC top-Z interactions . . . . .	46
34	3.2.4 Generating SM background events . . . . .	48
35	3.3 Multivariate analysis techniques: Boosted Decision Trees . . . . .	50
36	3.4 Statistical methodology . . . . .	52
37	<b>4 Event reconstruction and identification</b>	55
38	4.1 Object Reconstruction . . . . .	55
39	4.1.1 Charged particle tracks . . . . .	57
40	4.1.2 Following the Muon's Footsteps . . . . .	57
41	4.1.3 The path of the Electron . . . . .	57

42	4.1.4 Primary Vertex Reconstruction . . . . .	58
43	4.1.5 Calorimeter clusters . . . . .	58
44	4.2 Particle flow identification . . . . .	59
45	4.3 Pileup mitigation and luminosity measurement . . . . .	59
46	4.4 Physics object reconstruction and identification . . . . .	60
47	4.4.1 Muons . . . . .	60
48	4.4.2 Electrons . . . . .	62
49	4.4.3 Jets . . . . .	63
50	4.4.4 Jets from b fragmentation . . . . .	66
51	4.4.5 Missing transverse energy . . . . .	69
52	4.5 Summary of corrections . . . . .	69
53	<b>5 Event selection and categorisation</b>	<b>71</b>
54	5.1 Baseline event selection and filters . . . . .	71
55	5.1.1 Event cleaning . . . . .	74
56	5.1.2 Estimation of the trigger efficiency . . . . .	76
57	5.1.3 Corrections . . . . .	77
58	5.1.4 Reconstruction of kinematic variables . . . . .	81
59	5.2 Analysis Strategy . . . . .	82
60	5.3 Data driven NPL background simulation . . . . .	83
61	5.4 Regions and channels . . . . .	85
62	5.4.1 WZCR . . . . .	86
63	5.4.2 TTCR and STCR . . . . .	87
64	5.4.3 TTSR and STSR . . . . .	87
65	<b>6 The search for FCNC involving a top quark and a Z boson</b>	<b>89</b>
66	6.1 Construction of template distributions . . . . .	89
67	6.1.1 Distributions of the BDT variables . . . . .	90
68	6.1.2 BDTs . . . . .	90
69	6.1.3 Transverse mass in WZCR . . . . .	90
70	6.2 Systematic uncertainties . . . . .	91
71	6.3 Limit setting procedure validation . . . . .	92
72	6.4 Result and discussion . . . . .	95
73	6.4.1 One dimensional limits . . . . .	106■
74	6.4.2 Two-dimensional limits . . . . .	109■
75	<b>7 Denouement of the top-Z FCNC hunt at 13 TeV</b>	<b>111■</b>
76	<b>Appendices</b>	<b>111■</b>
77	<b>A Trigger scale factors</b>	<b>113■</b>
78	<b>B Dilepton controlplots</b>	<b>125■</b>
79	<b>C Statistical independent regions</b>	<b>127■</b>
80	<b>Bibliography</b>	<b>129■</b>

81

# Theoretical basis

# 1

---

82 The Standard Model (SM) [1] is a name given in the 1970s to a theory describing the fun-  
 83 damental particles and their interactions. This quantum field theory describes the particles  
 84 and their interactions as fields and has successfully incorporated three of the four fundamental  
 85 forces in the universe. In [Section 1.1](#), the particle content of the SM is summarised, while  
 86 [Section 1.2](#) describes the SM Lagrangian and its symmetries. In [Section 1.3](#), the flavour content  
 87 of the SM is highlighted, and [Section 1.4](#) focusses on the top quark in the SM.

88 The successful theory of the SM has some shortcomings which are discussed in [Section 1.6](#)  
 89 and lead to searches for a more general theory. One of such is using an effective field theory  
 90 (EFT) approach [2] to search for new physics in a model independent way. In [Section 1.7](#) an  
 91 EFT model focussing on flavour changing neutral currents (FCNC) involving a top quark is  
 92 presented. Its current experimental constraints are given in [Section 1.8](#).

93 **1.1 Elementary particles and forces**

94 The interactions in nature can be described by four forces, the strong force, the electromagnetic  
 95 (EM) force, the weak force and the gravitational force. These interactions happen via particles  
 96 with an integer spin known as bosons. The strong interaction is mediated by eight gluons  $g$ ,  
 97 while the electromagnetic force is mediated by photons  $\gamma$ , and the weak force by Z and  $W^\pm$   
 98 bosons. In [Table 1.1](#), the forces and their characteristics are shown. The gravitational force is  
 99 the only force not included in the SM and can be neglected for energies lower than the Planck  
 scale ( $1.22 \times 10^{19}$  GeV).

**Table 1.1:** The four forces of nature and their characteristics.

	Range	Mediator
Strong force	$10^{-15}$ m	8 gluons
Electromagnetic force	$\infty$	photon
Weak force	$10^{-18}$ m	$W^\pm$ , Z bosons
Gravitational force	$\infty$	unknown

100

101 The fermions are the particles that make up the visible matter in the universe. They carry  
 102 half integer spin and can be subdivided into leptons and quarks, where leptons do not interact  
 103 strongly. Each fermion has a corresponding anti-fermion which has the same mass and is  
 104 oppositely charged. The electron  $e$  is the first elementary particle discovered [3] and belongs  
 105 to the first generation of leptons together with the electron neutrino  $\nu_e$ . The second generation  
 106 comprises the muon  $\mu$  and muon neutrino  $\nu_\mu$ , whereas the third generation consists of the tau  
 107  $\tau$  and tau neutrino  $\nu_\tau$ . The neutrinos are neutral particles, while the other leptons have charge  
 108  $\pm q_e$  with  $q_e$  representing the elementary charge of  $1.602 \times 10^{-19}$  C. The masses of charged  
 109 leptons differ by four orders of magnitude between the first and third generations. In the SM  
 110 the neutrinos are assumed to be massless, nonetheless it is experimentally established that  
 111 neutrinos do have a tiny non-zero mass [4, 5]. In Table 1.2, the leptons and their properties in  
 the SM are summarised.

**Table 1.2:** The properties of the leptons in the three generations of the SM [6], where  $q_e$  represents the elementary charge.

	Generation	Particle	Mass	Charge
First		$e^-$	0.511 MeV	$-q_e$
		$\nu_e$	$\approx 0$	0
Second		$\mu^-$	106 MeV	$-q_e$
		$\nu_\mu$	$\approx 0$	0
Third		$\tau^-$	1 777 MeV	$-q_e$
		$\nu_\tau$	$\approx 0$	0

112

113 The quarks can also be divided into three generations. Unlike the leptons, they carry  
 114 colour charge and can interact via the strong interaction. The top quark, discovered in 1995  
 115 at the Tevatron [7, 8], is the heaviest SM particle with a mass<sup>1</sup> measured to be  $173.34 \pm$   
 116  $0.27(\text{stat}) \pm 0.71(\text{syst})$  GeV [9]. The quarks and their properties are summarised in Table 1.3.  
 117 In nature, only colour neutral objects can exist. This has as consequence that quarks are bound  
 118 through gluons into mesons (quark+anti-quark) and baryons (three quarks). These mesons  
 119 and baryons are mostly short-lived and unstable particles that rapidly decay through  $W^\pm$  and  $Z$   
 120 bosons. The only known stable baryon is the proton, made up of two up quarks and one down  
 121 quark.

122 The scalar boson, commonly known as the Higgs boson, is the last piece of the SM discovered  
 123 in 2012 [10, 11]. It is responsible for the masses of the  $W^\pm$  and Zboson, and that of the fermions.

---

<sup>1</sup>In this thesis all masses and energies are expressed in natural units, where the speed of light and  $\hbar$  are taken to be equal to one.

**Table 1.3:** The properties of the quarks in the three generations of the SM [6], where  $q_e$  represents the elementary charge.

	Generation	Particle	Mass	Charge
First	up u	$2.2^{+0.6}_{-0.4}$ MeV	$\frac{2}{3} q_e$	
	down d	$4.7^{+0.5}_{-0.4}$ MeV	$\frac{-1}{3} q_e$	
Second	charm c	$1.28 \pm 0.03$ GeV	$\frac{2}{3} q_e$	
	strange s	$96^{+8}_{-4}$ MeV	$\frac{-1}{3} q_e$	
Third	top t	$173.1 \pm 0.6$ GeV	$\frac{2}{3} q_e$	
	bottom b	$4.18^{+0.04}_{-0.03}$ GeV	$\frac{-1}{3} q_e$	

## 124 1.2 Standard Model Lagrangian, connecting fields with particles

125 The SM is a quantum field theory and thus describes the dynamics and kinematics of particles  
 126 and forces by a Lagrangian  $\mathcal{L}$ . The theory is based on the  $SU(3)_C \times SU(2)_L \times U(1)_Y$  gauge  
 127 symmetry, where  $SU(2)_L \times U(1)_Y$  describes the electroweak interaction and  $SU(3)_C$  the strong  
 128 interaction. The indices refer to colour C, the left chiral nature of the  $SU(2)_L$  coupling L, and the  
 129 weak hypercharge Y. Its Lagrangian is constructed such that symmetries representing physics  
 130 conservation laws such as conservation of energy, momentum and angular momentum are  
 131 contained. The symmetries under local gauge transformations are sustained by demanding  
 132 gauge invariance<sup>2</sup>.

The  $U(1)_Y$  group has one generator Y with an associated gauge field  $B_\mu$ . The three gauge fields  $W_\mu^1$ ,  $W_\mu^2$ , and  $W_\mu^3$ , are associated to  $SU(2)_L$  with three generators that can be written as half of the Pauli matrices:

$$T_1 = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad T_2 = \frac{1}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \text{and} \quad T_3 = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (1.1)$$

The generators  $T^a$  satisfy the Lie algebra:

$$[T_a, T_b] = i\epsilon_{abc} T^c \text{ and } [T_a, Y] = 0, \quad (1.2)$$

133 where  $\epsilon_{abc}$  is an antisymmetric tensor. The gauge fields of  $SU(2)_L$  only couple to left-handed  
 134 fermions as required by the observed parity violating nature of the weak force. The  $SU(3)_C$   
 135 group represents quantum chromodynamics (QCD). It has eight generators corresponding to  
 136 eight gluon fields  $G_\mu^{1\dots 8}$ . Unlike  $SU(2)_L \times U(1)_Y$ ,  $SU(3)_C$  is not chiral.

Under  $SU(3)_C$  quarks are colour triplets while leptons are colour singlets. This implies that the quarks carry a colour index ranging between one and three, whereas leptons do not take part in strong interactions. Based on the chirality, the quarks and leptons are organized in doublets or singlets. Each generation  $i$  of fermions consists of left-handed doublets and right-handed

---

<sup>2</sup>Different field configurations of unobservable fields can result in identical quantities. Transformations between such configurations are called gauge transformation and the absence of change in the measurable quantities is a characteristic called gauge invariance.

singlets:

$$l_{L,i} = \begin{pmatrix} e_{L,i} \\ v_{L,i} \end{pmatrix}, e_{R,i}, q_{L,i} = \begin{pmatrix} u_{L,i} \\ d_{L,i} \end{pmatrix}, u_{R,i}, \text{ and } d_{R,i} \quad (1.3)$$

The SM Lagrangian can be decomposed as a sum of four terms

$$\mathcal{L}_{SM} = \mathcal{L}_{gauge} + \mathcal{L}_f + \mathcal{L}_{Yuk} + \mathcal{L}_\phi, \quad (1.4)$$

137 that are related to the gauge, fermion, Yukawa and scalar sectors. The gauge Lagrangian  
 138 regroups the gauge fields of all three symmetry groups, and the fermionic part consists of kinetic  
 139 energy terms for quarks and leptons. The interaction between fermions and the scalar doublet  
 140  $\phi$  gives rise to fermion masses and is described by the Yukawa Lagrangian. The scalar part of  
 141 the Lagrangian is composed of a kinematic and potential component related to the scalar boson.

For the electroweak theory, two coupling constants are introduced, namely  $g'$  for  $U(1)_Y$  and  $g$  for  $SU(2)_L$ . The physically observable gauge bosons of this theory are the photon field  $A_\mu$ , the Z boson field  $Z_\mu^0$ , and the W boson fields  $W_\mu^\pm$ . These are a superposition of the four gauge fields of  $SU(2)_L \times U(1)_Y$ :

$$\begin{aligned} A_\mu &= \sin \theta_W W_\mu^3 + \cos \theta_W B_\mu, \\ Z_\mu^0 &= \cos \theta_W W_\mu^3 - \sin \theta_W B_\mu, \text{ and} \\ W_\mu^\pm &= \sqrt{\frac{1}{2}} (W_\mu^1 \mp i W_\mu^2), \end{aligned} \quad (1.5)$$

142 where  $\theta_W$  represents the weak mixing angle defined as  $\tan \theta_W = \frac{g'}{g}$ .

143 The coupling constant representing the strength of the QCD interactions is denoted as  $g_s$ . In  
 144 QCD there is asymptotic freedom whereby the strong coupling constant becomes weaker as the  
 145 energy with which the interaction between strongly interacting particles is probed increases,  
 146 and stronger as the distance between the particles increases. A consequence of this is known as  
 147 colour confinement, the quarks and gluons can not exist on their own and are not observed  
 148 individually. They are bound in colour neutral states called hadrons, this process is known as  
 149 hadronisation.

## 150 Electroweak symmetry breaking

151 In  $\mathcal{L}_{gauge}$  and  $\mathcal{L}_f$  are no mass terms for fermions present because only singlets under  $SU(3)_C \times$   
 152  $SU(2)_L \times U(1)_Y$  can acquire a mass with an interaction of the type  $m^2 \phi^\dagger \phi$  without breaking  
 153 the gauge invariance. In order to accommodate mass terms for fermions and gauge fields,  
 154 electroweak symmetry breaking, leading to  $\mathcal{L}_\phi$  is introduced.

The scalar doublet is introduced in the SM as

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \varphi_1 + i\varphi_2 \\ \varphi_3 + i\varphi_4 \end{pmatrix}. \quad (1.6)$$

Its field potential is of the form

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2, \quad (1.7)$$

with  $\mu^2 < 0$  and  $\lambda$  a positive integer. This choice of parameters gives the potential a "Mexican hat" shape. It has an infinite set of minima (ground states) and by expanding the field around an arbitrary choice of ground state, the electroweak symmetry is broken (EW):

$$\phi = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix} + \hat{\phi}, \quad (1.8)$$

where  $v$  is the vacuum expectation value (vev), measured to be around 245 GeV and corresponds to  $\sqrt{\frac{-\mu}{\lambda}}$ . The scalar doublet's four degrees of freedom are reduced to three degrees of freedom that couple to the gauge fields and fix the  $W^+$ ,  $W^-$  and Zbosons. The remaining fourth degree of freedom has given rise to a physically observable particle, called the Brout-Englert-Higgs (BEH) boson. This spontaneous symmetry breaking leaves the gauge invariance intact and gives masses to the  $W^\pm$  and Zbosons as:

$$m_W = \frac{1}{2} v |g| \quad \text{and} \quad m_Z = \frac{1}{2} v \sqrt{g'^2 + g^2}. \quad (1.9)$$

<sup>155</sup> The Brout-Englert-Higgs field couples universally to fermions with a strength proportional to their masses, and to gauge bosons with a strength proportional to the square of their masses.

### 1.3 Flavours in the SM

Flavour changing charged currents are introduced in 1963 by Nicola Cabibbo [12]. Via interactions with a W boson the flavour of the quarks is changed. At the time of the postulation only up, down, and strange quarks were known and the charged weak current was described as a coupling between the up quark and  $d_{\text{weak}}$ , where  $d_{\text{weak}}$  is a linear combination of the down and strange quarks,  $d_{\text{weak}} = \cos\theta_c d + \sin\theta_c s$ . This linear combination is a direct consequence of the chosen rotation

$$\begin{pmatrix} d_{\text{weak}} \\ s_{\text{weak}} \end{pmatrix} = \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix} = \mathcal{R} \begin{pmatrix} d \\ s \end{pmatrix}, \quad (1.10)$$

where the rotation angle  $\theta_c$  is known as the Cabibbo angle. This provides a definition for the charged weak current between u and d quarks,

$$J_\mu = \bar{u} \gamma_\mu (1 + \gamma_5) d_{\text{weak}}. \quad (1.11)$$

A consequence of Cabibbo's approach is that the  $s_{\text{weak}}$  is left uncoupled, leading to Glashow, Iliopoulos and Maiani (GIM) [13–15] to require the existence of a fourth quark with charge  $\frac{2}{3}q_e$ . This quark, known as the charm quark, couples to  $s_{\text{weak}}$  and a new definition of the charged weak current is modified to

$$J_\mu = (\bar{u} \quad \bar{c}) \gamma_\mu (1 + \gamma_5) \mathcal{R} \begin{pmatrix} d \\ s \end{pmatrix} = \bar{U} \gamma_\mu (1 + \gamma_5) \mathcal{R} D. \quad (1.12)$$

The neutral weak current is defined as

$$J_3 = \bar{U} \gamma_\mu (1 + \gamma_5) [\mathcal{R}, \mathcal{R}^\dagger] D, \quad (1.13)$$

and is diagonal in flavour space. This has as consequence that no flavour changing neutral currents occur at tree-level interactions [1].

Kobayashi and Maskawa generalised the Cabibbo rotation matrix to accommodate a third generation of quarks. The result is a  $3 \times 3$  unitary matrix known as the CKM matrix, responsible for the mixing of weak interaction states of down-type quarks:

$$\begin{pmatrix} d_{\text{weak}} \\ s_{\text{weak}} \\ b_{\text{weak}} \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \mathcal{V}_{\text{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \quad (1.14)$$

where  $\mathcal{V}_{\text{CKM}}$  is unitary ( $\mathcal{V}_{\text{CKM}}^\dagger \mathcal{V}_{\text{CKM}} = \mathbb{1}$ ). A general  $3 \times 3$  unitary matrix depends on three real angles and six phases. For the CKM matrix, the freedom to redefine the phases of the quark eigenstates can remove five of the phases, leaving a single physical phase known as the Kobayashi-Maskawa phase. This phase is responsible for the charge parity violation in the SM [16]. Each element  $V_{ij}$  of  $\mathcal{V}_{\text{CKM}}$  represents the transition probability of a quark  $i$  going to a quark  $j$ , and is experimentally determined to be [6]

$$\mathcal{V}_{\text{CKM}} = \begin{pmatrix} 0.97425 \pm 0.00022 & 0.2253 \pm 0.0008 & (4.13 \pm 0.49) \times 10^{-3} \\ 0.225 \pm 0.008 & 0.986 \pm 0.016 & (41.1 \pm 1.3) \times 10^{-3} \\ (8.4 \pm 0.6) \times 10^{-3} & (40.0 \pm 2.7) \times 10^{-3} & 1.021 \pm 0.032 \end{pmatrix}. \quad (1.15)$$

From Equation 1.15 follows that top quarks predominantly decay via charged weak currents to bottom quarks, with a probability consistent with unity. In the SM, FCNC can only occur via higher loop interactions which are highly suppressed. The expected transition probabilities for a top quark decaying via a FCNC interaction in the SM are given in Table 1.4, where it is clear that the FCNC top quark interactions of the SM is still beyond the reach of the sensitivity of current experiments.

**Table 1.4:** The predicted branching fractions  $\mathcal{B}$  for FCNC decays involving the top quark in the SM [17].

Process	$\mathcal{B}$ in the SM	Process	$\mathcal{B}$ in the SM
$t \rightarrow uZ$	$8 \times 10^{-17}$	$t \rightarrow cZ$	$1 \times 10^{-14}$
$t \rightarrow u\gamma$	$4 \times 10^{-16}$	$t \rightarrow c\gamma$	$5 \times 10^{-14}$
$t \rightarrow ug$	$4 \times 10^{-14}$	$t \rightarrow cg$	$5 \times 10^{-12}$
$t \rightarrow uH$	$2 \times 10^{-17}$	$t \rightarrow cH$	$3 \times 10^{-15}$

165

## 1.4 The top quark in the SM

Discovered in 1995 by the CDF and D0 collaborations at the Tevatron with proton-antiproton data [18, 19], the top quark plays an important role in studying high energy physics. Its Yukawa

interaction is given by

$$\mathcal{L}_{\text{top-Yukawa}} = -\frac{\lambda_t \nu}{\sqrt{2}} \bar{t}_L t_R - \frac{\lambda_t}{\sqrt{2}} H \bar{t}_L t_R + \text{h.c.}, \quad (1.16)$$

yielding a Yukawa coupling of [6]

$$\lambda_t = \frac{\sqrt{2} m_t}{\nu} = 0.991 \pm 0.003. \quad (1.17)$$

167 This Yukawa coupling is very large compared to the other Yukawa couplings in the SM ( $\mathcal{O}(10^{-2})$ ),  
 168 leading to the belief that the top quark may have an important role in understanding the  
 169 mechanism of electroweak symmetry breaking. On top of this, the very short lifetime of the top  
 170 quark makes it an excellent candidate to study the properties of a bare quark. Its high mass,  
 171 almost 40 times higher than the mass of the closest fermion in mass, leads to a large coupling  
 172 with the Higgs boson and makes the top quark an interesting candidate to investigate how  
 173 particles acquire mass.

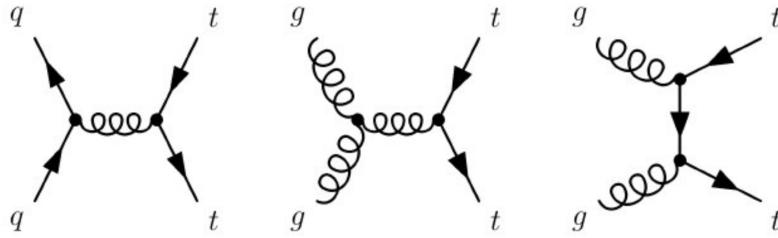
174 The CKM matrix element  $V_{tb}$ , given in [Equation 1.15](#), is experimentally found to be much  
 175 larger than  $V_{ts}$ ,  $V_{td}$ , and close to unity. The top quark decays through electroweak interactions  
 176 since the W boson mass is smaller than the top quark mass and the W boson can be on shell. A  
 177 consequence of this is that the top quark has a very short lifetime of only  $1/\Gamma_t \approx 5 \cdot 10^{-25}$  s [6]  
 178 leading to the fact that the formation of bound states involving top quarks are not allowed.  
 179 This lifetime is even shorter than the typical hadronisation timescale of  $1/\Lambda_{\text{QCD}} \approx 10^{-23}$  s,  
 180 prohibiting gluons to radiate from the top quark and keeping its spin coherent. Since the  
 181 electroweak interactions have a vector-axial vector (V-A) coupling structure<sup>3</sup>, the top quark  
 182 spin orientation can be derived from the angular distributions of its decay products. This makes  
 183 it possible to study the polarisation of top quarks from the angular distributions in various  
 184 processes.

185 The massiveness of the top quark leads to the fact that a large amount of energy is needed to  
 186 create one. This is only the case for high energy collisions such as those happening in the Earth's  
 187 upper atmosphere when cosmic rays collide with particles in air, or by particle accelerators.  
 188 The production of top quarks happens in two ways: single via the electroweak interaction or in  
 189 pairs via the strong interaction. At hadron colliders, the dominant production mechanism is top  
 190 quark production via gluon ( $gg \rightarrow t\bar{t}$ ) or quark fusion ( $q\bar{q} \rightarrow t\bar{t}$ ). In [Figure 1.1](#), the different top  
 191 quark pair production mechanisms are shown. The production channel of gluon fusion is the  
 192 main contributor to the top quark pair cross section at the LHC compared to quark fusion at the  
 193 Tevatron. The  $gg \rightarrow t\bar{t}$  process contributes 80-90% to the total top quark pair cross section in  
 194 the LHC centre-of-mass energy regime of 7-14 TeV [6]. In [Table 1.5](#) the predicted top quark  
 195 pair production cross sections are given for the LHC and the Tevatron, while in [Figure 1.2](#), a  
 196 summary plot of the LHC and Tevatron top quark pair cross section measurements as a function  
 197 of the centre-of-mass energy can be found.

198 The singly produced top quarks are produced via the electroweak interaction. These production  
 199 mechanisms are subdivided at leading order into three main channels based on the virtuality

---

<sup>3</sup>In the SM a vector - axial vector coupling structure  $(\gamma^\mu - \gamma^\mu \gamma^5)$  is predicted that only permits left-handed fermions or right-handed anti fermions to interact with a spin-1 particle.



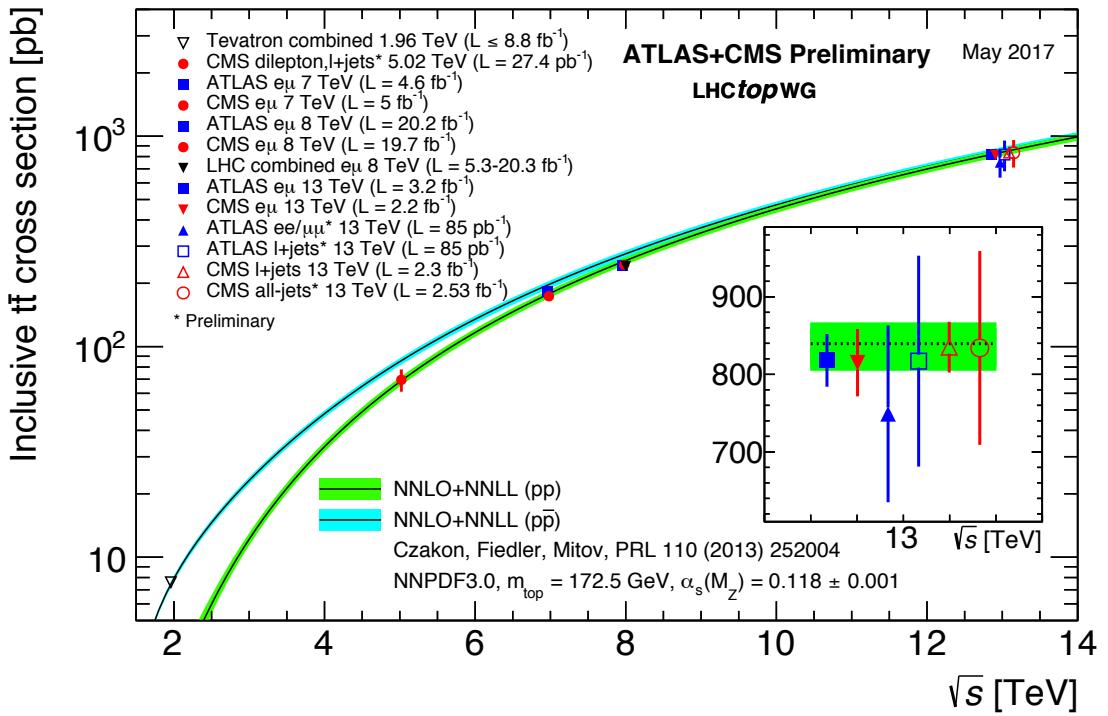
**Figure 1.1:** Leading order diagrams of the top quark pair production. Gluon fusion (right and middle) are the dominant processes at the LHC, while quark fusion (left) is the dominant one at the Tevatron.

**Table 1.5:** Predictions on the top quark pair production cross sections at next-to-next-to-leading order with next-to-next-to-leading log soft gluon resummation per centre-of-mass energy [6]. The first uncertainty is from scale dependence, while the second uncertainty originates from parton density functions.

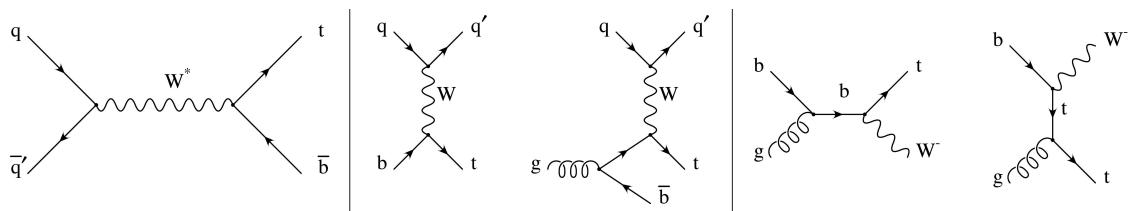
Experiment	Top quark mass	Centre-of-mass energy	Cross section (pb)
Tevatron	$m_t = 173.3$ GeV	$\sqrt{s} = 1.96$ TeV	$\sigma_{\bar{t}t} = 7.16^{+0.11+0.17}_{-0.20-0.12}$
LHC	$m_t = 173.2$ GeV	$\sqrt{s} = 7$ TeV	$\sigma_{\bar{t}t} = 173.6^{+4.5+8.9}_{-5.9-8.9}$
LHC	$m_t = 173.2$ GeV	$\sqrt{s} = 8$ TeV	$\sigma_{\bar{t}t} = 247.7^{+6.3+11.5}_{-8.5-11.5}$
LHC	$m_t = 173.2$ GeV	$\sqrt{s} = 13$ TeV	$\sigma_{\bar{t}t} = 816.0^{+19.4+34.4}_{-28.6-34.4}$

200 ( $Q^2 = -p_\mu p^\mu$ ) of the exchanged W boson. In Figure 1.3, the corresponding Feynman diagrams  
 201 are shown. The single top quark production cross sections, given in Table 1.6, are smaller than  
 202 the top quark pair production cross sections since the electroweak coupling strength is smaller  
 203 than the strong coupling strength. In addition, for the single top quark production, there is the  
 204 need of sea quarks ( $b, \bar{q}$ ) in the initial states for which the parton density functions increase  
 205 less steeply at low momentum fractions compared to the gluon parton density functions.

206 The production via the  $t$ -channel has a virtuality of the W boson  $Q^2 > 0$ , making it space-like.  
 207 It is produced via the scattering of the W boson of a bottom quark coming from a proton  
 208 or from gluon splitting ( $g \rightarrow b\bar{b}$ ). It has the highest single top quark cross section in proton  
 209 collisions and the top quark production is roughly twice as large than the antitop quark. This is  
 210 a consequence of the up-down valence quark composition of the proton. This feature makes the  
 211  $t$ -channel sensitive to the parton density functions of the proton. The  $s$ -channel is the production  
 212 mechanism with the smallest cross section. Here the W boson is time-like ( $Q^2 < 0$ ) which  
 213 requires the W boson to have a large virtuality to produce the heavier top quark. It is produced  
 214 from two quarks belonging to the same isodoublet (e.g.  $u\bar{d}$ ) and subsequently decays to  $t\bar{b}$ .  
 215 This process gets enhanced by many beyond the Standard Model scenarios via the addition of  
 216 new heavy particles such as  $W'$ . The  $tW$ -channel has a top quark produced in association with a  
 217 W boson produced on shell  $Q^2 = -m_W^2$ . This mode is negligible at the Tevatron, but of relevant  
 218 size at the LHC. The  $tW$ -channel is sensitive to new physics affecting the  $Wtb$  vertex. The single  
 219 top quark production cross section measurements by the CMS collaboration can be found in  
 220 Figure 1.4.



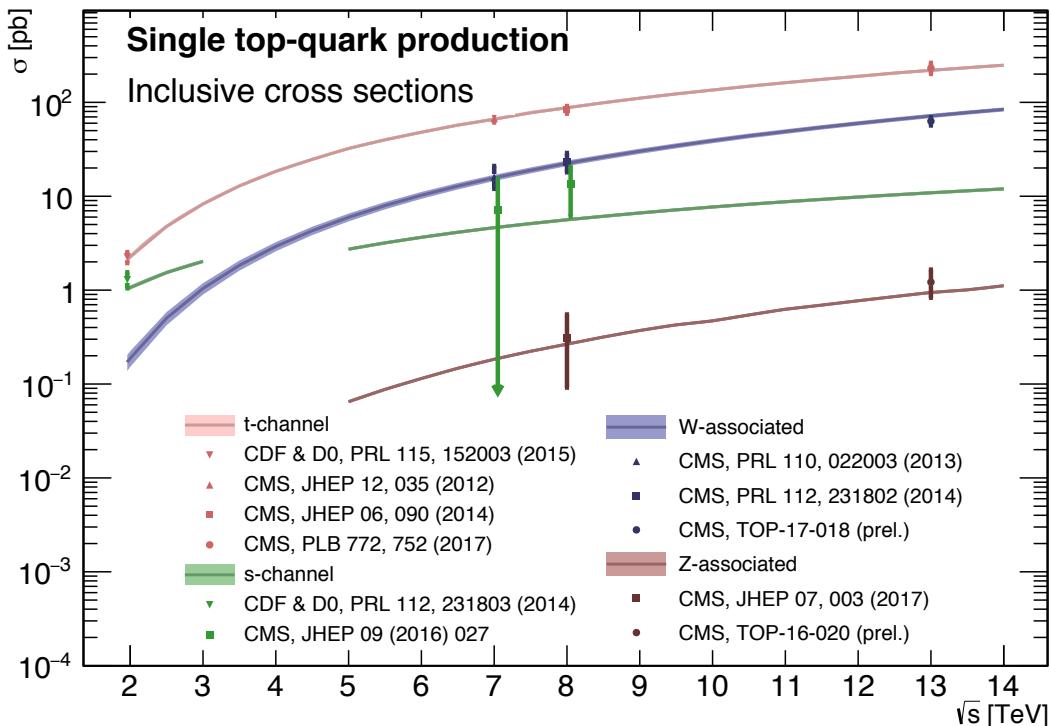
**Figure 1.2:** Summary of the LHC and the Tevatron measurements of the top quark pair production cross section as function of the centre-of-mass energy compared with the next-to-next-to-leading order QCD calculation. The theory bands are the uncertainties due to renormalization and factorisation scales, parton density functions and the strong coupling. The mass of the top quark is assumed to be 172.5 GeV. Measurements for the same centre-of-mass energy are slightly off-set for clarity. Figure taken from [20].



**Figure 1.3:** Leading order Feynman diagrams of the electroweak production of single top quarks in the  $s$ -channel (left),  $t$ -channel (middle), and for the  $tW$  associated production. Figure taken from [21].

**Table 1.6:** Predictions on the single top quark production cross sections at next-to-leading order per centre-of-mass energy [6]. The uncertainties from scale dependence and from parton density functions are combined in quadrature or given separately (scale + PDF). For the  $t$ -channel the relative proportions to  $t$  and  $\bar{t}$  are 65% and 35%. For the  $s$ -channel this is respectively 69% and 31%. The  $tW$ -channel has an equal proportion of top and antitop quarks. For Tevatron, the top quark mass is assumed to be 173.3 GeV, while the LHC predictions use  $m_t = 172.5$  GeV [6, 22].

Collider	Centre-of-mass energy	Cross section $\sigma_{t+\bar{t}}$ (pb)		
		$t$ -channel	$s$ -channel	$tW$ -channel
Tevatron	$\sqrt{s} = 1.96$ TeV	$2.06^{+0.13}_{-0.13}$	$1.03^{+0.05}_{-0.05}$	-
LHC	$\sqrt{s} = 7$ TeV	$63.89^{+2.91}_{-2.52}$	$4.29^{+0.19}_{-0.17}$	$15.74^{+0.40+1.10}_{-0.40-1.14}$
LHC	$\sqrt{s} = 8$ TeV	$84.69^{+3.76}_{-3.23}$	$5.24^{+0.22}_{-0.20}$	$22.37^{+0.60+1.40}_{-0.60-1.40}$
LHC	$\sqrt{s} = 13$ TeV	$216.99^{+9.04}_{-7.71}$	$10.32^{+0.40}_{-0.36}$	$71.7^{+1.80+3.40}_{-1.80-3.40}$



**Figure 1.4:** Summary of the measurements of the single top quark production cross section as function of the centre-of-mass energy. Figure taken from [23].

## 221 1.5 Effective field theories

222 Problems can be simplified if one looks at the relevant scale of the process that one want to  
 223 investigate, for example the chemical properties of an hydrogen atom can be described without  
 224 any knowledge of quark interactions inside the proton. In this case, the proton can be considered  
 225 the elementary object (indivisible) due to the fact that the binding energy of the constituents is  
 226 much bigger than the energy of the electron in orbit around the proton. Effective field theories  
 227 are based on this kind of separation of different energy scales in a system [24]. Effective field  
 228 theories can be used for theories where the perturbative expansion cannot be trusted, e.g. QCD  
 229 at low energy, or as bottom up approach to look for new physics in a model independent way.  
 230 The latter is the way effective field theory will be used throughout this thesis.

The main idea behind effective field theory is easily explained via the example of the Fermi theory. Fermi explained in 1933 [25] the  $\beta$ -decay as a product of currents:

$$\mathcal{L}_{\text{EFT}}^{\text{Fermi}} = -\frac{G_F}{\sqrt{2}} J^\mu J_\mu^\dagger, \quad (1.18)$$

where  $G_F$  is the Fermi coupling constant, measured to be  $G_F \approx 1.17 \times 10^{-5} \text{ GeV}^{-2}$ . The current  $J_\mu$  can written as the sum of an hadronic  $J_\mu^h$  and leptonic  $J_\mu^l$  current, where for simplicity only the leptonic current discussed.

$$J_\mu^l = \sum_i \bar{\nu}_i \gamma_\mu (1 - \gamma_5) l. \quad (1.19)$$

231 Historically, charged currents were flavour universal and the later discovered parity violation of  
 232 the weak interaction led to the V-A structure. After this, the  $SU(2)_L$  symmetry was postulated  
 233 and the existence of neutral currents was predicted. The effective Lagrangian used then (given  
 234 in Equation 1.18), could nowadays be build starting from  $SU(2)_L$  symmetries only.

The muon decay can be computed from two different starting points. The effective Fermi Lagrangian provides the decay width of the muon into an electron and two neutrinos

$$\Gamma(\mu \rightarrow e \bar{\nu}_e \nu_\mu) \approx \frac{1}{96\pi^3} \frac{m_\mu^2}{\Lambda_F^4}, \quad (1.20)$$

where  $\Lambda_F$  is the energy scale defined as

$$\frac{G_F}{\sqrt{2}} = \frac{1}{\Lambda_F^2}. \quad (1.21)$$

From muon decay measurements, the value of  $\Lambda_F$  is determined to be  $\Lambda_F \approx 348 \text{ GeV}$  [24]. From the SM Lagrangian, one could also calculate the muon decay. Considering that the momenta involved are small compared to the W boson mass, the propagator's denominator can be expanded as [1]

$$\frac{1}{p^2 - m_W^2} = -\frac{1}{m_W^2} - \frac{p^2}{m_W^4} + \dots \quad (1.22)$$

Looking at the first term, and identifying

$$\frac{g^2}{8m_W} = \frac{1^2}{\Lambda_F}, \quad (1.23)$$

one sees that this corresponds with [Equation 1.20](#), thus the effective Lagrangian in [Equation 1.18](#) is the first term of the expansion in  $\frac{1}{m_W^2}$  applied on the full Lagrangian.

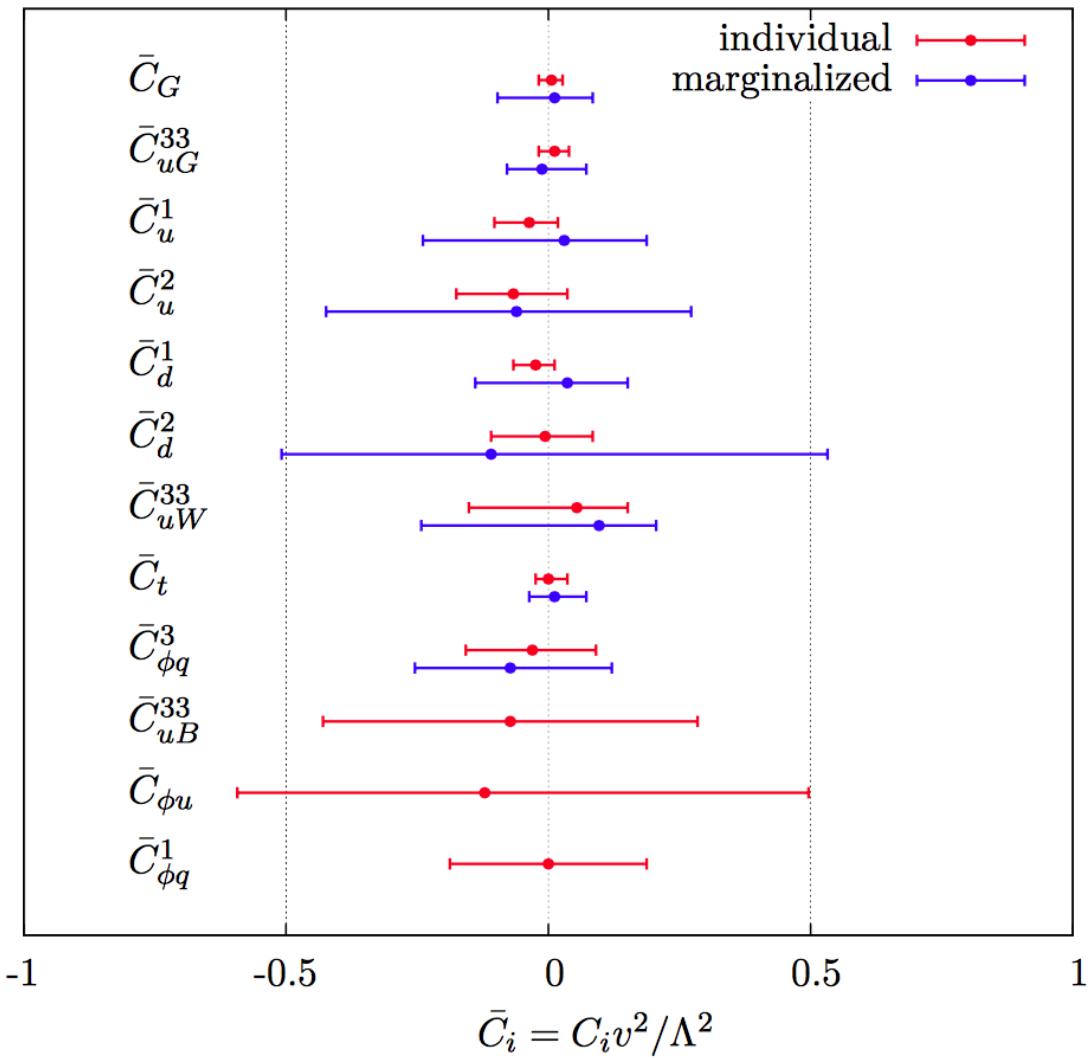
An effective theory is thus a Taylor expansion in the ratio of two scales and the only remnants of the full theory at low energies are the symmetries and the values of the coupling constants. If the expansion parameter is small, one can truncate the series leading to the Lagrangian containing a finite number of free coefficients, making predictions possible. The error on these predictions is then of the order as the truncated piece.

The SM can be seen as an effective theory applicable up to energies not exceeding a scale  $\Lambda$ . Therefore, remnants should still be valid and the theory above that scale should have a gauge group containing  $SU(3)_C \times SU(2)_L \times U(1)_Y$  and all the SM degrees of freedom, as well as reduce to the SM at lower energies. The general SM Lagrangian becomes then

$$\mathcal{L}_{SM+EFT} = \mathcal{L}_{SM}^{(4)} + \frac{1}{\Lambda} \sum_k C_k^{(5)} Q_k^{(5)} + \frac{1}{\Lambda^2} \sum_k C_k^{(6)} Q_k^{(6)} + \mathcal{O}\left(\frac{1}{\Lambda^3}\right), \quad (1.24)$$

where  $Q_k^{(n)}$  are dimension- $n$  operators (currents) and  $C_k^{(n)}$  the corresponding dimensionless coupling constants, so-called Wilson coefficients. The Wilson coefficients are determined by the underlying high energy theory.

In the Warsaw basis [\[26\]](#), a set of independent operators of dimension 5 and 6 are built out of the SM fields and are consistent with the SM gauge symmetries and is fully derived in Ref. [\[26\]](#). In general the various measurements show a good agreement with the SM predictions and by lack of deviations from the SM, limits on the anomalous couplings can be derived. The estimated coupling strengths per operator contributing to single top quark production obtained from various measurements at the LHC and Tevatron are shown in [Figure 1.5](#) for which the conventions are discussed in Ref. [\[27\]](#). These results are consistent with the SM expectation for which those operators vanish.



**Figure 1.5:** Global fit results of top quark effective field theory to experimental data including all constrained operators at dimension six. For the operators, the Warsaw basis of [26] is used. The bounds are set on the Wilson coefficients of various operators contributing to top quark production and decay in two cases (red) all other coefficients set to zero, or (blue) all other coefficients marginalized over. Figure taken from [28].

253 **1.6 Motivation for new physics**

254 Many high energy experiments confirm the success of the SM. In particular the scalar boson,  
 255 the cornerstone of the SM, has consecrated the theory. Unfortunately there are also strong  
 256 indications that the SM ought to be a lower energy expression of a more global theory. The  
 257 existence of physics beyond the SM (BSM) [29] is strongly motivated. These motivations are  
 258 based on direct evidence from observation such as the existence of neutrino masses, the existence  
 259 of dark matter and dark energy, or the matter-antimatter asymmetry, and also from theoretical  
 260 problems such as the hierarchy problem, the coupling unification or the large numbers of free  
 261 parameters in the SM.

262 In the SM, the neutrinos are assumed to be massless, while experiments with solar, atmospheric,  
 263 reactor and accelerator neutrinos have established that neutrinos can oscillate and change flavour  
 264 during flight [4, 5]. These oscillations are only possible when neutrinos have masses. The  
 265 flavour neutrinos ( $\nu_e$ ,  $\nu_\mu$ ,  $\nu_\tau$ ) are then linear expressions of the fields of at least three mass  
 266 eigenstate neutrinos  $\nu_1$ ,  $\nu_2$ , and  $\nu_3$ .

267 The ordinary or baryonic matter described by the SM describes only 5% of the mass (energy)  
 268 content of the universe. Astrophysical evidence indicated that dark matter is contributing  
 269 to approximately 27% and dark energy to 68% of the content of the universe. From the  
 270 measurements of the temperature and polarizations anisotropies of the cosmic microwave  
 271 background by the Planck experiment [30], the density of cold non baryonic matter is determined.  
 272 Cold dark matter is assumed to be only sensitive to the weak and gravitational force, leading  
 273 to only one possible SM candidate: the neutrino. However, these are too light to account for  
 274 the vast amount of dark matter and other models are needed. Dark energy is assumed to be  
 275 responsible for the acceleration in the expansion of the universe [31].

276 At the Big Bang, matter and antimatter are assumed to be produced in equal quantities.  
 277 However, it is clear that we are surrounded by matter. So where did all the antimatter go?  
 278 In 1967, Sakharov identified three mechanisms that are necessary to obtain a global matter  
 279 antimatter asymmetry [32]. These mechanisms are those of baryon and lepton number violation,  
 280 that at a given moment in time there was a thermal imbalance for the interactions in the universe,  
 281 and there is charge C and charge parity CP violation<sup>4</sup>.

282 The large number of free parameters in the SM comes from the nine fermion masses, three  
 283 CKM mixing angles and one CP violating phase, one EM coupling constant  $g'$ , one weak coupling  
 284 constant  $g$ , one strong coupling constant  $g_s$ , one QCD vacuum angle, one vacuum expectation  
 285 value, and one mass of the scalar boson. This large number of free parameters leads to the  
 286 expectation of a more elegant and profound theory beyond the SM.

287 The hierarchy problem [33] is related to the huge difference in energy between the weak  
 288 scale and the Planck scale. The vev of the Brout-Englert-Higgs field determines the weak scale  
 289 that is approximately 246 GeV. The radiative corrections to the scalar boson squared mass  $m_H^2$ ,  
 290 coming from its self couplings and couplings to fermions and gauge bosons, are quadratically

---

<sup>4</sup>The rate of a process  $i \rightarrow f$  can be different from the CP-conjugate process:  $\tilde{i} \rightarrow \tilde{f}$ . The SM includes sources of CP-violation through the residual phase of the CKM matrix. However, these could not account for the magnitude of the asymmetry observed.

291 proportional to the ultraviolet momentum cut-off  $\Lambda_{\text{UV}}$ . This cut-off is at least equal to the energy  
 292 to which the SM is valid without the need of new physics. For the SM to be valid up to the  
 293 Planck mass, the correction to  $m_H^2$  becomes thirty orders of magnitude larger than  $m_H^2$ . This  
 294 implies that an extraordinary cancellation of terms should happen. This is also known as the  
 295 naturalness problem of the H boson mass.

The correction to the squared mass of the scalar boson coming from a fermion  $f$ , coupling to the scalar field  $\phi$  with a coupling  $\lambda_f$  is given by

$$\Delta m_H^2 = -\frac{|\lambda_f|^2}{8\pi^2} \Lambda_{\text{UV}}^2, \quad (1.25)$$

while the correction to the mass from a scalar particle  $S$  with a mass  $m_S$ , coupling to the scalar field with a Lagrangian term  $-\lambda_S |\phi|^2 |S|^2$  is

$$\Delta m_H^2 = \frac{|\lambda_S|^2}{16\pi^2} \left( \Lambda_{\text{UV}}^2 - 2m_S^2 \ln \left( \frac{\Lambda_{\text{UV}}}{m_S} \right) + \dots \right). \quad (1.26)$$

296 As one can see the correction term to  $m_H^2$  is much larger than  $m_H^2$  itself. By introducing BSM  
 297 physics models that introduce new scalar particles at the TeV scale that couple to the scalar  
 298 boson one can cancel the  $\Lambda_{\text{UV}}^2$  divergence and avoid this fine-tuning.

299 The choice of the  $SU(3)_C \times SU(2)_L \times U(1)_Y$  symmetry group itself as well as the separate  
 300 treatment of the three forces included in the SM raises concern. The intensity of the forces  
 301 show a large disparity around the electroweak scale, but have comparable strengths at higher  
 302 energies. The electromagnetic and weak forces are unified in a electroweak interaction, but the  
 303 strong coupling constant does not encounter the other coupling constants at high energies. In  
 304 order to reach a grand unification, the running of couplings can be modified by the addition of  
 305 new particles in BSM models.

## 306 1.7 An effective approach beyond the SM: FCNC involving a top 307 quark

308 The closeness of the top quark mass to the electroweak scale led physicists to believe that it is a  
 309 sensitive probe for new physics. Studying its properties is therefore an important topic of the  
 310 experimental program at the LHC. Several extensions of the SM enhance the FCNC branching  
 311 fractions and can be probed at the LHC [17], from which some of them are shown in Table  
 312 1.7. Previous searches have been performed at the Tevatron by the CDF [34] and D0 [35]  
 313 collaborations, and at the LHC by the ATLAS [36–40] and CMS [41–45] collaborations.

314 The impact of BSM models can be written in a model independent way by means of an  
 315 effective field theory valid up to an energy scale  $\Lambda$ . The leading effects are parametrized by a  
 316 set of fully gauge symmetric operators that are added to the SM Lagrangian and can be reduced  
 317 to a minimal set of operators as seen in Equation 1.24. For simplicity, the assumption is made  
 318 that new physics effects are exclusively described by dimension-6 operators, thus neglecting

**Table 1.7:** The predicted branching fractions  $\mathcal{B}$  for FCNC interactions involving the top quark in some BSM models [17]: quark singlet (QS), generic two Higgs doublet model (2HDM) and the minimal supersymmetric extensions to the SM (MSSM);

Process	QS	2HDM	MSSM	Process	QS	2HDM	MSSM
$t \rightarrow uZ$	$\leq 1.1 \times 10^{-4}$	—	$\leq 2 \times 10^{-6}$	$t \rightarrow cZ$	$\leq 1.1 \times 10^{-4}$	$\leq 10^{-7}$	$\leq 2 \times 10^{-6}$
$t \rightarrow u\gamma$	$\leq 7.5 \times 10^{-9}$	—	$\leq 2 \times 10^{-6}$	$t \rightarrow c\gamma$	$\leq 7.5 \times 10^{-9}$	$\leq 10^{-6}$	$\leq 2 \times 10^{-6}$
$t \rightarrow ug$	$\leq 1.5 \times 10^{-7}$	—	$\leq 8 \times 10^{-5}$	$t \rightarrow cg$	$\leq 1.5 \times 10^{-7}$	$\leq 10^{-4}$	$\leq 8 \times 10^{-5}$
$t \rightarrow uH$	$\leq 4.1 \times 10^{-5}$	$\leq 5.5 \times 10^{-6}$	$\leq 10^{-5}$	$t \rightarrow cH$	$\leq 4.1 \times 10^{-5}$	$\leq 10^{-3}$	$\leq 10^{-5}$

319 neutrino physics. In the fully gauge symmetric case, the EFT Lagrangian is then given by

$$\mathcal{L}_{\text{SM+EFT}} = \mathcal{L}_{\text{SM}} + \sum_i \frac{\bar{c}_i}{\Lambda^2} O_i + \mathcal{O}\left(\frac{1}{\Lambda^3}\right), \quad (1.27)$$

where the Wilson coefficients  $\bar{c}_i$  depend on the considered theory and on the way that new physics couples to the SM particles. Taking into account that  $\Lambda$  is large, contributions suppressed by powers of  $\Lambda$  greater than two are neglected. Additionally, all four fermion operators are omitted for the rest of this thesis. The Warsaw basis is adopted for the independent effective operators [26], parametrising the new physics effects relevant for the flavour changing neutral current interactions of the top quark as, all flavour indices understood,

$$\begin{aligned} \mathcal{L}_{\text{EFT}}^t = & \frac{\bar{c}_{uG}}{\Lambda^2} \Phi^\dagger \cdot [\bar{Q}_L \sigma^{\mu\nu} \mathcal{T}_a u_R] G_{\mu\nu}^a + \frac{\bar{c}_{uB}}{\Lambda^2} \Phi^\dagger \cdot [\bar{Q}_L \sigma^{\mu\nu} u_R] B_{\mu\nu} + \frac{2\bar{c}_{uW}}{\Lambda^2} \Phi^\dagger T_i \cdot [\bar{Q}_L \sigma^{\mu\nu} u_R] W_{\mu\nu}^i \\ & + i \frac{\bar{c}_{hu}}{\Lambda^2} \left[ \Phi^\dagger \overleftrightarrow{D}_\mu \Phi \right] [\bar{u}_R \gamma^\mu u_R] + i \frac{\bar{c}_{hq}^{(1)}}{\Lambda^2} \left[ \Phi^\dagger \overleftrightarrow{D}_\mu \Phi \right] [\bar{Q}_L \gamma^\mu Q_L] \\ & + i \frac{4\bar{c}_{HQ}^{(3)}}{\Lambda^2} \left[ \Phi^\dagger T_i \overleftrightarrow{D}_\mu \Phi \right] [\bar{Q}_L \gamma^\mu T^i Q_L] + \frac{\bar{c}_{uh}}{\Lambda^2} \Phi^\dagger \Phi \Phi^\dagger \cdot [\bar{Q}_L u_R] + \text{h.c.}, \end{aligned} \quad (1.28)$$

where the left handed  $SU(2)_L$  doublet of the quark fields is denoted by  $Q_L$ , the up-type right handed fields by  $u_R$ , the down-type right handed fields by  $d_R$ , the  $SU(2)_L$  doublet of the Higgs field by  $\Phi$ , the field strength tensors as

$$\begin{aligned} B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu, \\ W_{\mu\nu}^k &= \partial_\mu W_\nu^k - \partial_\nu W_\mu^k - g \epsilon_{ij}^k W_\mu^i W_\nu^j, \\ G_{\mu\nu} &= \partial_\mu G_\nu^a - \partial_\nu G_\mu^a + g_s f_{bc}^a G_\mu^b G_\nu^c, \end{aligned} \quad (1.29)$$

denoting the structure constant of the  $SU(3)_C$  group as  $f_{bc}^a$  and the structure constant of the  $SU(2)_L$  group as  $\epsilon_{ij}^k$ . The gauge covariant derivatives are also standard defined as

$$D_\mu \Phi = \partial_\mu \Phi - \frac{1}{2} i g' B_\mu \Phi - i g T_k W_\mu^k \Phi \quad (1.30)$$

with the conventions of Section 1.2. The representation matrices  $T$  of  $SU(2)_L$  are defined in Equation 1.1, while the representation matrices  $\mathcal{T}$  of  $SU(3)_C$  are the Gell-Mann matrices [1].

The hermitian derivative operator is defined as

$$\Phi^\dagger \overleftrightarrow{D} \Phi = \Phi^\dagger D^\mu \Phi - D_\mu \Phi^\dagger \Phi. \quad (1.31)$$

After electroweak symmetry breaking, the operators induce [17, 46] both corrections to the SM couplings and new interactions at tree level such as FCNC interactions. The FCNC interactions of the top quark that are not present in the SM are given by

$$\mathcal{L}_{\text{EFT}}^t = \frac{\sqrt{2}}{2} \sum_{q=u,c} \left[ g' \frac{\kappa_{t\gamma q}}{\Lambda} A_{\mu\nu} \bar{t} \sigma^{\mu\nu} (f_{\gamma q}^L P_L + f_{\gamma q}^R P_R) q \right. \quad (1.32)$$

$$+ \frac{g}{2\cos\theta_W} \frac{\kappa_{tZq}}{\Lambda} Z_{\mu\nu} \bar{t} \sigma^{\mu\nu} (f_{Zq}^L P_L + f_{Zq}^R P_R) q \quad (1.33)$$

$$+ \frac{\sqrt{2}g}{4\cos\theta_W} \zeta_{tZq} \bar{t} \gamma^\mu (f_q^L P_L + f_q^R P_R) q Z_\mu \quad (1.34)$$

$$+ g_s \frac{\kappa_{tgq}}{\Lambda} Z_{\mu\nu} \bar{t} \sigma^{\mu\nu} (f_{gq}^L P_L + f_{gq}^R P_R) q G_\mu^a \quad (1.35)$$

$$+ \eta_{Hqt} \bar{t} (\hat{f}_q^L P_L + \hat{f}_q^R P_R) q H + \text{h.c.} \Big], \quad (1.36)$$

where the value of the FCNC couplings at scale  $\Lambda$  are represented by  $\kappa_{tZq}, \kappa_{tgq}, \kappa_{t\gamma q}, \zeta_{tZq}$ , and  $\eta_{Hqt}$ . These are assumed to be real and positive, with the unit of  $\text{GeV}^{-1}$  for  $\kappa_{txq}/\Lambda$  and no unit for  $\zeta_{xqt}$  and  $\eta_{xqt}$ . In the equation  $\sigma^{\mu\nu}$  equals to  $\frac{i}{2} [\gamma^\mu, \gamma^\nu]$ , and the left- and right-handed chirality projector operators are denoted by  $P_L$  and  $P_R$ . The electromagnetic coupling constant is denoted by  $g'$ , the strong interaction coupling is denoted as  $g_s$ , while the electroweak interaction is parametrised by the coupling constant  $g$  and the electroweak mixing angle  $\theta_W$ . The complex chiral parameters are normalized according to  $|f_{xq}^L|^2 + |f_{xq}^R|^2 = 1$ ,  $|\tilde{f}_q^L|^2 + |\tilde{f}_q^R|^2 = 1$ , and  $|\hat{f}_q^L|^2 + |\hat{f}_q^R|^2 = 1$ . In the expression for  $\mathcal{L}_{\text{EFT}}^t$ , the unitary gauge is adopted and the scalar field is expanded around its vacuum expectation value  $v$  with  $H$  being the SM scalar boson. The field strength tensors of the photon  $A_\mu$ , the gluon field  $G_\mu^{a..8}$ , and the Z boson  $Z_\mu^0$  are defined as

$$\begin{aligned} A_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu, \\ Z_{\mu\nu} &= \partial_\mu Z_\nu - \partial_\nu Z_\mu, \text{ and} \\ G_{\mu\nu} &= \partial_\mu G_\nu^a - \partial_\nu G_\mu^a + g_s f_{bc}^a G_\mu^b G_\nu^c. \end{aligned} \quad (1.37)$$

320 Note that there are two coupling constants arising in  $\mathcal{L}_{\text{EFT}}^t$ , which is a residue of electroweak  
 321 symmetry breaking. The massive Z boson will appear in both the  $Z_\mu^0$  field as well as the covariant  
 322 derivative, leading to an extra Z-vertex.

323 The relations between the Wilson coefficients in (1.28) and the coupling strengths of the  
 324 interactions in Equation 1.36 can be derived. The 14 effective operators are mapped onto 10  
 325 free parameters providing a more minimal parametrisation of the anomalous interactions of the  
 326 top quark.

$$\begin{aligned}
 \kappa_{tgq} f_{gq}^L &= \frac{\nu}{g_s \Lambda} [\bar{c}_{uG}]_{i3}^*, & \kappa_{tgq} f_{gq}^R &= \frac{\nu}{g_s \Lambda} [\bar{c}_{uG}]_{3i}, \\
 \kappa_{t\gamma q} f_{\gamma q}^L &= \frac{\nu}{g' \Lambda} [\cos \theta_W \bar{c}_{uB} - \sin \theta_W \bar{c}_{uW}]_{i3}^*, & \kappa_{t\gamma q} f_{\gamma q}^R &= \frac{\nu}{g' \Lambda} [\sin \theta_W \bar{c}_{uB} - \cos \theta_W \bar{c}_{uW}]_{3i}, \\
 \kappa_{tZq} f_{Zq}^L &= -\frac{2\cos \theta_W \nu}{g \Lambda} [\sin \theta_W \bar{c}_{uB} + \cos \theta_W \bar{c}_{uW}]_{i3}^*, & \kappa_{tZq} f_{Zq}^R &= -\frac{2\cos \theta_W \nu}{g \Lambda} [\cos \theta_W \bar{c}_{uB} + \sin \theta_W \bar{c}_{uW}]_{3i}, \\
 \zeta_{tZq} \tilde{f}_{Zq}^L &= -\frac{2\nu^2}{\Lambda^2} [(\bar{c}_{hq}^{(1)} - \bar{c}_{hq}^{(3)})_{i3} + (\bar{c}_{hq}^{(1)} - \bar{c}_{hq}^{(3)})_{3i}^*], & \zeta_{tZq} \tilde{f}_{Zq}^R &= -\frac{2\nu^2}{\Lambda^2} [(\bar{c}_{hu})_{i3} + (\bar{c}_{hu})_{3i}^*], \\
 \eta_{tHq} \hat{f}_{Hq}^L &= \frac{3\nu^2}{2\Lambda^2} [\bar{c}_{uh}]_{i3}^*, & \eta_{tHq} \hat{f}_{Hq}^R &= \frac{3\nu^2}{2\Lambda^2} [\bar{c}_{uh}]_{3i}.
 \end{aligned} \tag{1.38}$$

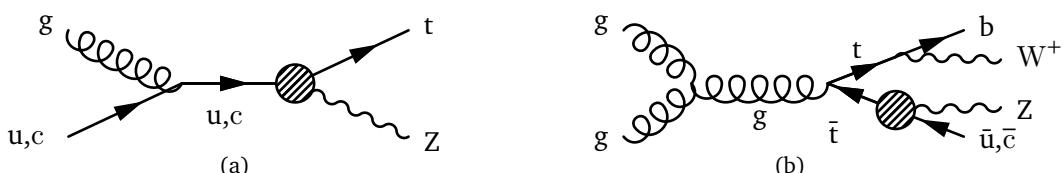
## 327 1.8 Experimental constraints on top-FCNC

Experiments commonly put limits on the branching fractions which allow an easier interpretation across different EFT models by use of the branching fraction

$$\mathcal{B}(t \rightarrow qX) = \frac{\delta_{txq}^2 \Gamma_{t \rightarrow qX}}{\Gamma_t}, \tag{1.39}$$

328 where  $\Gamma_{t \rightarrow qX}$  represents the FCNC decay width<sup>5</sup> for a coupling strength  $\delta_{txq}^2 = 1$ , and  $\Gamma_t$  the full  
 329 decay width of the top quark. In the SM, supposing a top quark mass of 172.5 GeV, the full  
 330 width becomes  $\Gamma_t^{\text{SM}} = 1.32$  GeV [47].

331 Searches for top-FCNC usually adopt a search strategy depending on the experimental set-up  
 332 and the FCNC interaction of interest, looking either for FCNC interactions in the production of a  
 333 single top quark or in its decay for top quark pair interactions. In Figure 1.6, these two cases  
 334 are shown for the tZq vertex.



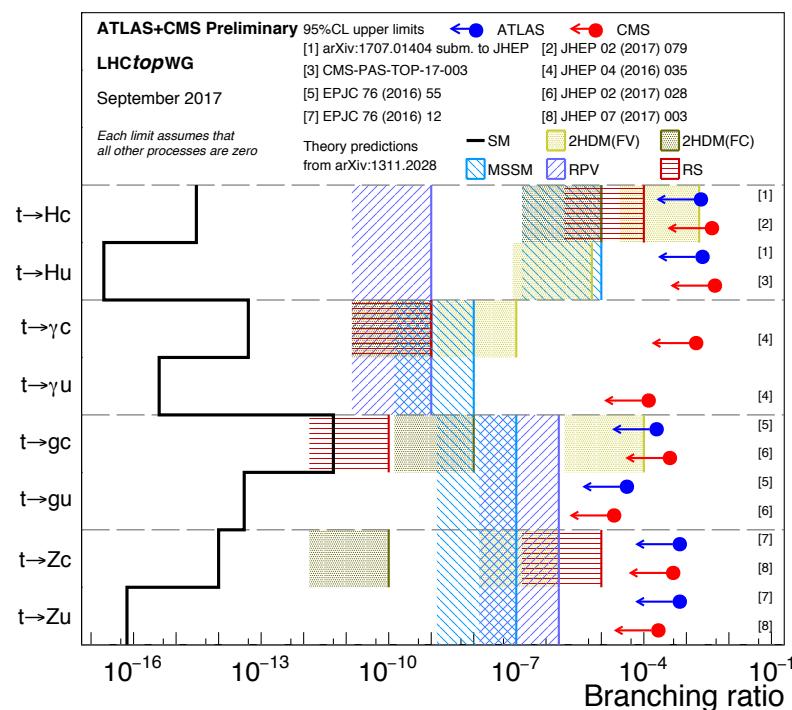
**Figure 1.6:** Feynman diagrams for the processes with a tZq FCNC interaction, where the FCNC interaction is indicated with the shaded dot. (a) Single top quark production through an FCNC interaction. (b) Top quark pair production with an FCNC induced decay.

<sup>5</sup>The decay width of a certain process represents the probability per unit time that a particle will decay. The total decay width, defined as the sum of all possible decay widths of a particle, is inversely proportional to its lifetime.

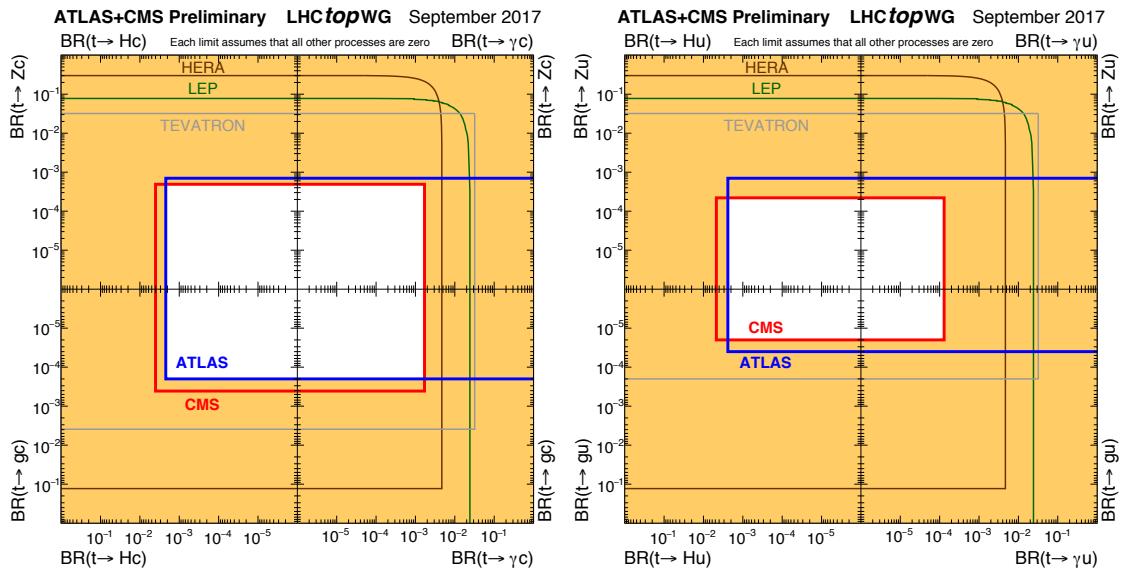
336 The observation of top-FCNC interactions has yet to come and experiments have so far only  
 337 been able to put upper bounds on the branching fractions. An overview of the best current limits  
 338 is given in [Table 1.8](#). In [Figure 1.7](#) a comparison is shown between the current best limits set by  
 339 ATLAS and CMS with respect to several BSM model benchmark predictions. From there one can  
 340 see that FCNC searches involving a Z or H boson are close to excluding or confirming several  
 341 BSM theories. In [Figure 1.9](#), the searches performed by CMS are summarised. For the tZq  
 342 vertex, the best limit from CMS comes from Ref. [41] where both single top quark and top quark  
 343 pair are studied. The observed (expected) limits 95% CL at 8 TeV for the FCNC tZq interaction  
 344 by CMS are  $\mathcal{B}(t \rightarrow uZ) < 2.2 \times 10^{-4}$ ( $2.7 \times 10^{-4}$ ) and  $\mathcal{B}(t \rightarrow cZ) < 4.9 \times 10^{-4}$ ( $12 \times 10^{-4}$ ). In  
 345 [Figure 1.8](#), the summary of the 95% confidence level observed limits on the branching fractions  
 346 of the top quark decays to a charm or up quark and a neutral boson is given, considering the  
 results from the HERA, the LEP, the Tevatron, and the LHC.

**Table 1.8:** Overview of the most stringent observed and expected experimental limits on top-FCNC branching fractions  $\mathcal{B}$  at 95% confidence level.

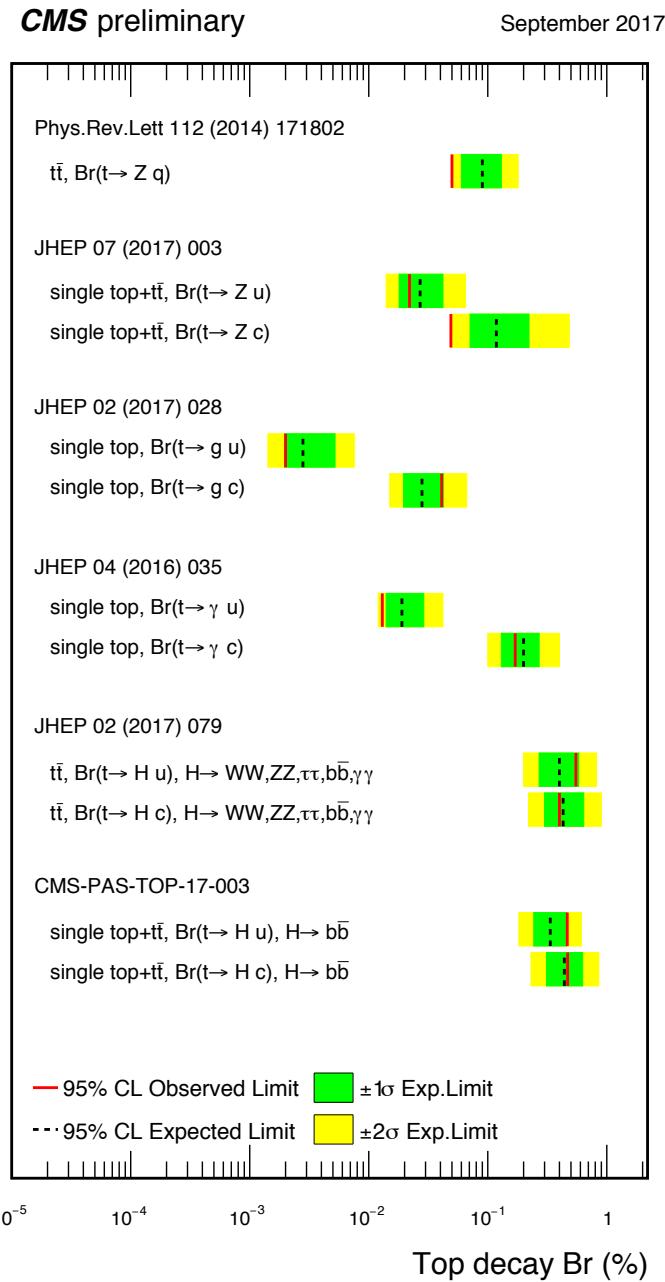
Process	Search mode	Observed $\mathcal{B}$	Expected $\mathcal{B}$	Experiment	
$t \rightarrow uZ$	top quark pair decay	$1.7 \times 10^{-4}$	$2.4 \times 10^{-4}$	ATLAS	[40]
$t \rightarrow u\gamma$	single top quark production	$1.3 \times 10^{-4}$	$1.9 \times 10^{-4}$	CMS	[43]
$t \rightarrow ug$	single top quark production	$4.0 \times 10^{-5}$	$3.5 \times 10^{-5}$	ATLAS	[37]
$t \rightarrow uH$	top quark pair decay	$2.4 \times 10^{-3}$	$1.7 \times 10^{-3}$	ATLAS	[39]
$t \rightarrow cZ$	top quark pair decay	$2.3 \times 10^{-4}$	$3.2 \times 10^{-4}$	ATLAS	[40]
$t \rightarrow c\gamma$	single top quark production	$2.0 \times 10^{-3}$	$1.7 \times 10^{-3}$	CMS	[43]
$t \rightarrow cg$	single top quark production	$2.0 \times 10^{-4}$	$1.8 \times 10^{-4}$	ATLAS	[37]
$t \rightarrow cH$	top quark pair decay	$2.2 \times 10^{-3}$	$1.6 \times 10^{-3}$	CMS	[39]



**Figure 1.7:** Current best limits set by CMS and ATLAS for top-FCNC interactions. Figure taken from [23]. (TO DO Remake with new atlas results)



**Figure 1.8:** Summary of the current 95% confidence level observed limits on the branching fractions of the top quark decays via flavour changing neutral currents to a charm (left) or up (right) quark and a neutral boson. The coloured lines represent the results from HERA (the most stringent limits between the ones obtained by the H1 and ZEUS collaborations, in brown), LEP (combined ALEPH, DELPHI, L3 and OPAL collaborations result, in green), TEVATRON (the most stringent limits between the ones obtained by the CDF and D0 collaborations, in grey). The yellow area represents the region excluded by the ATLAS and the CMS Collaborations. Figure taken from [20].



**Figure 1.9:** Summary of the FCNC branching fractions from CMS searches at 8 TeV. Figure taken from [23].

# Experimental set-up

# 2

349 A key objective of the Large Hadron Collider (LHC) was the search for the Brout-Englert-Higgs  
 350 boson. The Large Electron Positron (LEP) [48] and Tevatron [49] experiments established  
 351 that the mass of the scalar boson has to be larger than 114 GeV [50, 51], and smaller than  
 352 approximate 1 TeV due to unitarity and perturbativity constraints [52]. On top of this, the  
 353 search for new physics such as supersymmetry or the understanding of dark matter were part  
 354 of the motivation for building the LHC. Since the start of its operation, the LHC is pushing the  
 355 boundaries of the Standard Model, putting the most stringent limits on physics beyond the  
 356 Standard Model as well as precision measurements of the parameters of the Standard Model. A  
 357 milestone of the LHC is the discovery of the scalar boson in 2012 by the two largest experiments  
 358 at the LHC [10, 11].

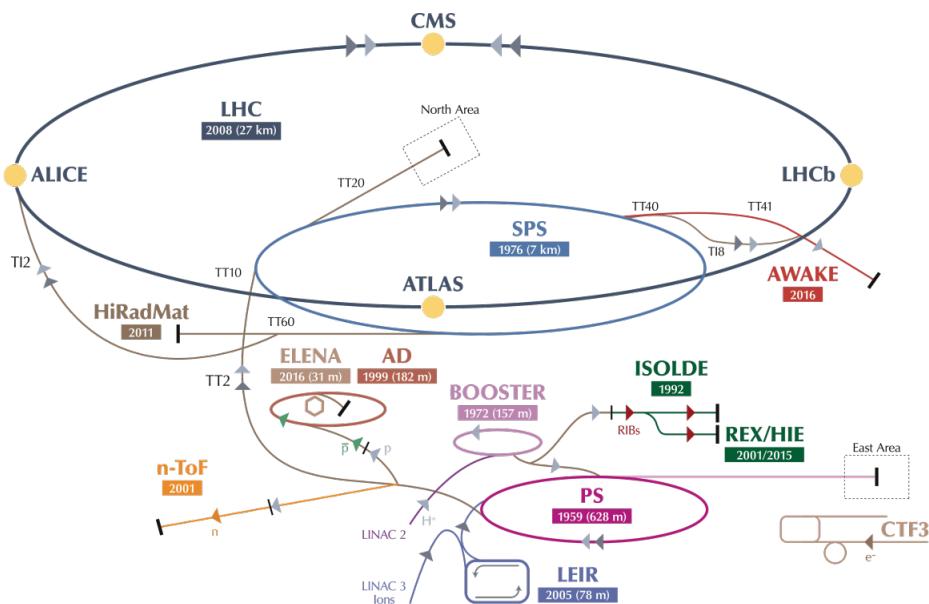
359 This chapter is dedicated to the experimental set-up of the LHC and the Compact Muon  
 360 Solenoid (CMS) experiment. Section 2.1 describes the LHC and its acceleration process for  
 361 protons to reach their design energies. The CMS experiment and its components are presented  
 362 in Section 2.2. The upgrades performed during the long shutdown in 2013 are discussed  
 363 in Section 2.2.4. The data acquisition of CMS is presented in Section 2.2.3, while the CMS  
 364 computing model is shown in Section 2.2.5.

## 365 2.1 The Large Hadron Collider

366 The LHC has started its era of cutting edge science on 10 September 2008 [53] after approval by  
 367 the European Organisation of Nuclear Research (CERN) in 1995 [54]. Installed in the previous  
 368 LEP tunnel, the LHC consists of a 26.7 km quasi ring, that is installed between 45 and 170 m  
 369 under the French-Swiss border amidst Cessy (France) and Meyrin (Switzerland). Built to study  
 370 rare physics phenomena at high energies, the LHC can accelerate mainly two types of particles,  
 371 protons and lead ions  $Pb^{45+}$ , and provides collisions at four interaction points, where the particle  
 372 bunches are crossing. Experiments for studying the collisions are installed at each interaction  
 373 point.

374 As can be seen in Figure 2.1, the LHC is the last element in a chain that creates, injects and  
 375 accelerates protons. The starting point is the ionisation of hydrogen, creating protons that are  
 376 injected in a linear accelerator (LINAC 2). Here, the protons obtain an energy of 50 MeV. They

377 continue to the Proton Synchrotron Booster (PSB or Booster), where the packs of protons are  
 378 accelerated to 1.4 GeV and each pack is split up in twelve bunches with 25 or 50 ns spacing.  
 379 The Proton Synchrotron (PS) then increases their energy to 25 GeV before the Super Proton  
 380 Synchrotron (SPS) increases the proton energy up to 450 GeV. Each accelerator ring expands in  
 381 radius in order to reduce the energy loss of the protons by synchrotron radiation<sup>1</sup>. Furthermore,  
 382 the magnets responsible for the bending of the proton trajectories have to be strong enough  
 383 to sustain the higher proton energy. Ultimately, the proton bunches are injected into opposite  
 384 directions into the LHC, where they are accelerated to 3.5 TeV (in 2010 and 2011), 4 TeV (in  
 2012 and 2013) or 6.5 TeV (in 2015 and 2016) [55].



**Figure 2.1:** Schematic representation of the accelerator complex at CERN [56]. The LHC (dark blue) is the last element in chain of accelerators. Protons are successively accelerated by LINAC 2, the Booster, the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS) before entering the LHC.

385

386 In Figure 2.2 the LHC programme is shown. the first data collisions, so-called Run 1 period,  
 387 lasted from 2008 until 16 February 2013 after which the CERN accelerator complex shut down  
 388 for two years of planned maintenance and consolidation during so-called long shutdown 1  
 389 (LS1). On 23 March 2015, the new data taking period known as Run 2 started. With a brief end  
 390 of the year extended technical stop (EYETS). The main activities carried out during the EYETS  
 391 were the maintenance of systems such as the cryogenics, the cooling, electrical systems, etc.; the  
 392 replacement of the magnet, as well as a de-cabling and cabling campaign on the SPS[57]. Run 2  
 393 will last until July 2018 when the long shutdown 2 (LS2) will begin for 2 years. The main goal  
 394 of this shutdown is the LHC injectors upgrade (LUI), but also maintenance and consolidation  
 395 will be performed. Furthermore, preparations for the High Luminosity LHC, which will start in

<sup>1</sup>This energy loss is proportional to the fourth power of the proton energy and inversely proportional to the bending radius.

396 2024, will be done. More information about phase 1 upgrades during LS1 and EYETS is given  
 397 in [Section 2.2.4](#).

398 Before the start of the LHC in 2010, the previous energy record was held by the Tevatron collider  
 399 at Fermilab, colliding protons with antiprotons at  $\sqrt{s} = 1.96$  TeV. When completely filled, the  
 LHC nominally contains 2220 bunches in Run 2, compared to 1380 in Run 1 (design: 2200).

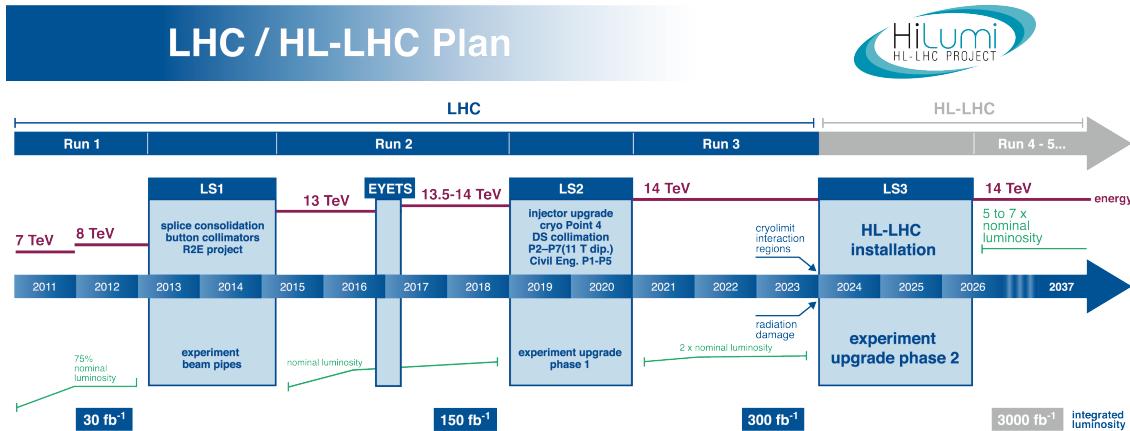


Figure 2.2: The HL-LHC timeline. Figure taken from [58].

400

401 Inside the LHC ring [59], the protons are accelerated by the means of radio frequency cavities,  
 402 while 1232 dipole magnets of approximately 15 m long, each weighing 35 t ensure the deflection  
 403 of the beams. The two proton beams circulate in opposite direction in separate pipes inside of  
 404 the magnet. Through the use of a strong electric current in the coils of the magnet, magnetic  
 405 fields are generated and cause the protons to bend in the required orbits. In order for the coil  
 406 to become superconducting and able to produce a strong magnetic field of 8.3 T, the magnet  
 407 structure is surrounded by a vessel. This vessel is filled with liquid Helium making it possible  
 408 to cool down the magnet to 1.9 K. In order to get more focussed and stabilised proton beams,  
 409 additional higher-order multipole and corrector magnets are placed along the LHC beam line.

410 The LHC is home to seven experiments, each located at an interaction point:

- 411 • A Toroidal LHC ApparatuS (ATLAS) [60] and the Compact Muon Solenoid (CMS) [61]  
 412 experiments are the two general purpose detectors at the LHC. They both have a hermetic,  
 413 cylindrical structure and were designed to search for new physics phenomena along with  
 414 precision measurements of the Standard Model. The existence of two distinct experiments  
 415 allows cross-confirmation of any discovery.
- 416 • A Large Ion Collider Experiment (ALICE) [62] and the LHC Beauty (LHCb) [63] experi-  
 417 ments are focusing on specific phenomena. ALICE studies strongly interacting matter  
 418 at extreme energy densities where a quark-gluon plasma forms in heavy ion collisions  
 (Pb-Pb or p-Pb). LHCb searches for differences between matter and antimatter with the  
 420 focus on b physics.

- The forward LHC (LHCf) [64] and the TOTal cross section, Elastic scattering and diffraction dissociation Measurement (TOTEM) [65] experiments are two smaller experiments that focus on head-on collisions. LHCf consists of two parts placed before and after ATLAS and studies particles created at very small angles. TOTEM is placed in the same cavern as CMS and measures the total proton-proton cross section and studies elastic and diffractive scattering.
- The Monopoles and Exotics Detector At the LHC (MoEDAL) [66] experiment is situated near LHCb and tries to find magnetic monopoles.

For the enhancement of the exploration of rare events and thus enhancing the number of collisions, high beam energies as well as high beam intensities are required. The luminosity [67] is a measurement of the number of collisions that can be produced in a detector per square meter and per second and is the key role player in this enhancement. The LHC collisions create a number of events per second given by

$$N_{\text{event}} = L\sigma_{\text{event}}, \quad (2.1)$$

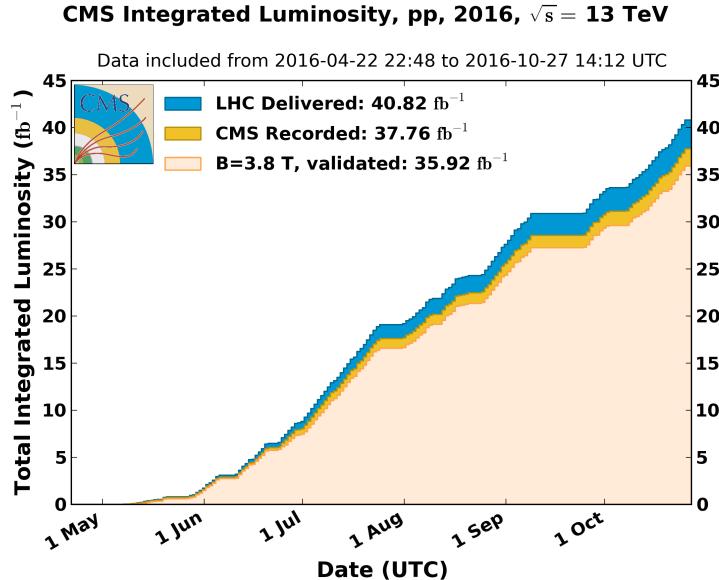
where  $\sigma_{\text{event}}$  is the cross section of the process of interest and  $L$  the machine instantaneous luminosity. This luminosity depends only on the beam parameters and is for a Gaussian beam expressed as

$$L = \frac{1}{4\pi} N_b n_b f_{\text{rev}} \frac{N_b}{\epsilon_n} \left( 1 + \left( \frac{\theta_c \sigma_z}{2\sigma^*} \right)^2 \right)^{-\frac{1}{2}} \frac{\gamma_r}{\beta^*}. \quad (2.2)$$

The number of particles per bunch is expressed by  $N_b$ , while  $n_b$  is the number of bunches per beam,  $f_{\text{rev}}$  the revolution frequency,  $\gamma_r$  the relativistic gamma factor,  $\epsilon_n$  the normalized transverse beam emittance - a quality for the confinement of the beam,  $\beta^*$  the beta function at the collision point - a measurement for the width of the beam,  $\theta_c$  the angle between two beams at the interaction point,  $\sigma_z$  the mean length of one bunch, and  $\sigma^*$  the mean height of one bunch. In Equation 2.2, the blue part represents the stream of particles, the red part the brilliance, and the green part the geometric reduction factor due to the crossing angle at the interaction point.

The peak design luminosity for the LHC is  $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ , which leads to about 1 billion proton interactions per second. In 2016, the LHC was around 10% above this design luminosity [68]. The luminosity is not a constant in time since it diminishes due to collisions between the beams, and the interaction of the protons and the particle gas that is trapped in the centre of the vacuum tubes due to the magnetic field. The diffusion of the beam degrades the emittance and therefore also the luminosity. For this reason, the mean lifetime of a beam inside the LHC is around 15 h. The integrated luminosity - the luminosity provided in a certain time range - recorded by CMS and ATLAS over the year 2016 is given in Figure 2.3. In Run 2, the peak luminosity is  $13\text{-}17 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$  compared to  $7.7 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$  in Run 1. The recorded luminosity is validated for physics analysis keeping  $35.9 \text{ fb}^{-1}$  during 2016 data taking.

Multiple proton-proton interactions can occur during one bunch crossing, referred to as pileup. On average, the number of pileup events is proportional to the luminosity times the total inelastic proton-proton cross section. In 2016, an average of about 27 of pileup interactions has been observed in 13 TeV proton collisions at the interaction point of CMS. For 2012, this number was about 21 pileup interactions for 8 TeV collisions.



**Figure 2.3:** Cumulative off-line luminosity measured versus day delivered by the LHC (blue), and recorded by CMS (orange), and certified as good physics analysis during stable beams (light orange) during stable beams and for proton collisions at 13 TeV centre-of-mass energy in 2016. [69].

## 451 2.2 The Compact Muon Solenoid

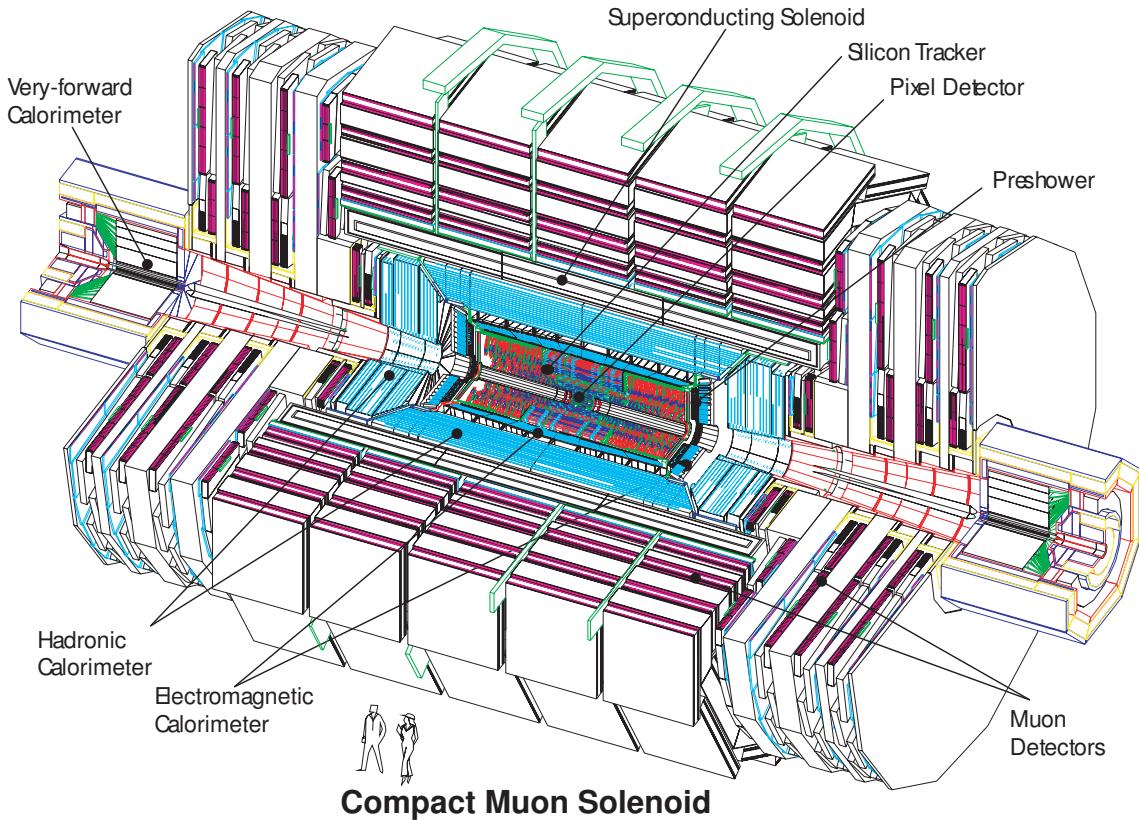
452 At one of the collision points of the LHC, the CMS detector [70–72] is placed. Weighing 14 000 t,  
453 this cylindrical detector is about 28.7 m long and 15 m in diameter. It has an onion like structure  
454 of several specialised detectors and contains a superconducting solenoid with a magnetic field of  
455 3.8 T. Living in a hadronic environment, multi-jet processes produced by the strong interaction  
456 are the main source of background for rare physics processes. Therefore, good identification,  
457 momentum resolution, and charge determination of muons, electrons and photons are one of  
458 the main goals of the CMS detector. Additionally, a good charged particle momentum resolution  
459 and reconstruction efficiency in the inner tracker provides identification for jets coming from b  
460 quarks or tau particles. Also the electromagnetic resolution for an efficient photon and lepton  
461 isolation as well as a good hadronic calorimeter for the missing transverse energy<sup>2</sup> were kept  
462 into account while designing CMS. In Figure 2.4, an overview of the CMS detector is shown.

### 463 2.2.1 CMS coordinate system

The coordinate system used by CMS can be found in Figure 2.5. The origin of the right handed orthogonal coordinate system is chosen to be the point of collisions. The x-axis points towards the centre of the LHC ring such that the y-axis points towards the sky, and the z-axis lies tangent to the beam axis. Since the experiment has a cylindrical shape, customary coordinates are used to describe the momentum  $\vec{p}$ : the distance  $p = |\vec{p}|$ , the azimuthal angle<sup>3</sup>  $\phi \in [-\pi, \pi]$ , the

<sup>2</sup>The missing transverse energy comes from an imbalance in the transverse plane. This will be discussed in Chapter 4.

<sup>3</sup>The azimuthal angle is the angle between the x-axis and the projection in the transverse plane of the momentum  $\vec{p}$ , denoted as  $\vec{p}_T$ .



**Figure 2.4:** Mechanical layout of the CMS detector. Figure taken from [73].

pseudo-rapidity<sup>4</sup>  $\eta$  :

$$\eta = -\ln \left( \tan \left( \frac{\theta}{2} \right) \right). \quad (2.3)$$

For the energies considered at the LHC, where  $E \gg m$ , the pseudo-rapidity is a good approximation of the rapidity  $y$

$$y = \frac{1}{2} \ln \left( \frac{E + p_z}{E - p_z} \right), \quad (2.4)$$

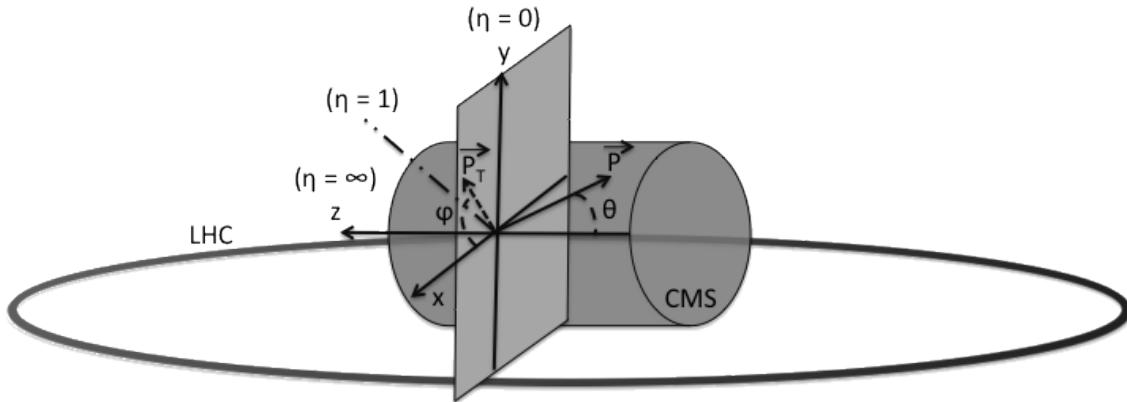
<sup>464</sup> where the difference of rapidities of two particles is invariant under a Lorentz boost in the  
<sup>465</sup> z-direction.

### 466 2.2.2 Towards the heart of CMS

<sup>467</sup> The CMS detector can be divided into two parts. A central barrel is placed around the beam  
<sup>468</sup> pipe ( $|\eta| < 1.4$ ), and two plugs (endcaps) ensure the hermeticity of the detector. In [Figure 2.4](#)  
<sup>469</sup> and [Figure 2.6](#) the onion like structure of the CMS detector is visible. The choice of a solenoid  
<sup>470</sup> of 12.9 m long and 5.9 m diameter gives the advantage of bending the particle trajectories in the

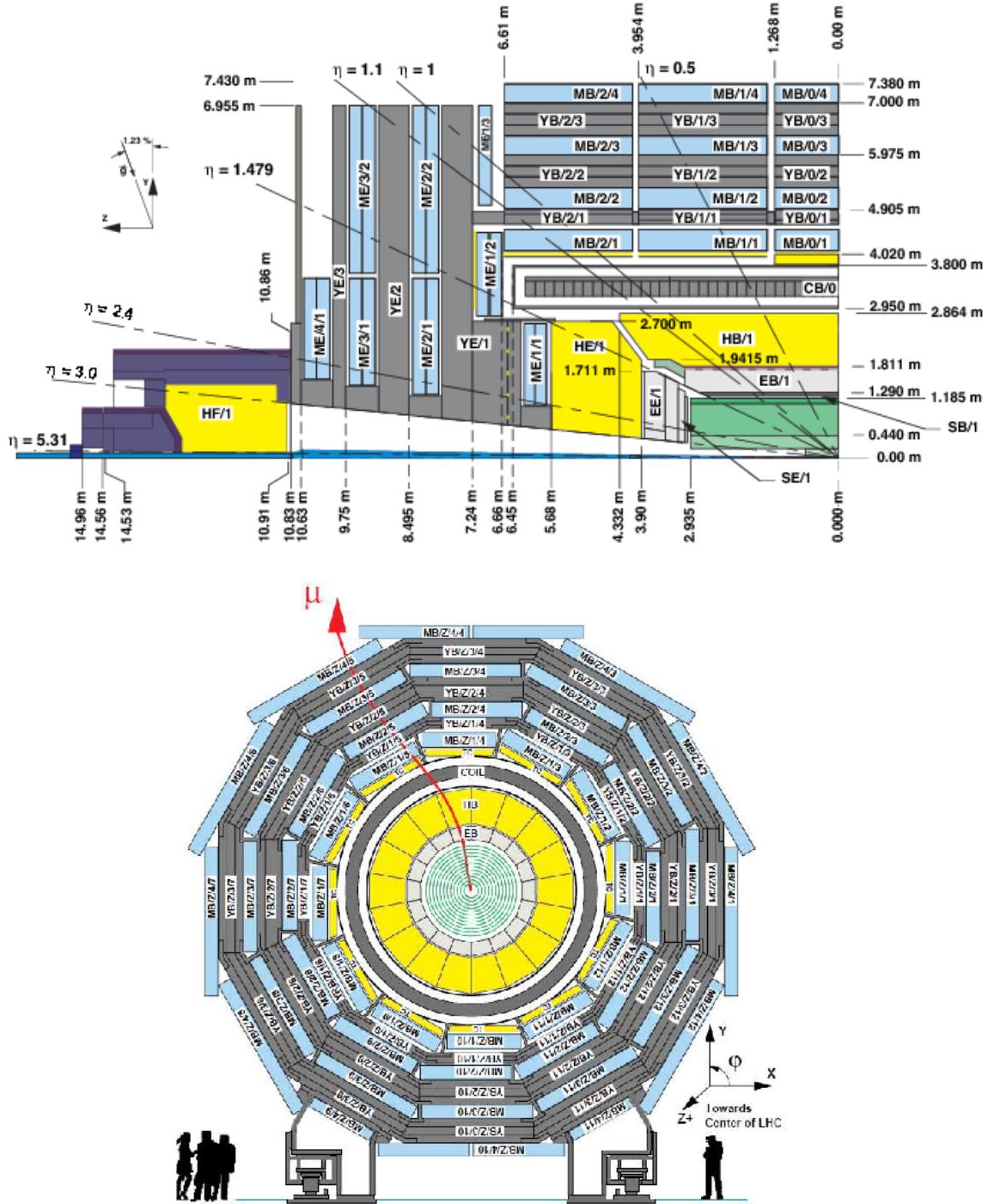
---

<sup>4</sup>The pseudo rapidity is expressed by the polar angle  $\theta$  between the direction of  $\vec{p}$  and the beam.



**Figure 2.5:** Representation of the coordinate system used by CMS. The point of origin is put at the collision point. The x-axis points towards the centre of the LHC ring such that the z-axis lies tangent to the beam axis.

471 transverse plane. The hadronic calorimeter (Section 2.2.2.3), the electromagnetic calorimeter  
 472 (Section 2.2.2.4) and the tracker (Section 2.2.2.5) are within the solenoid (Section 2.2.2.2),  
 473 while the muon chambers (Section 2.2.2.1) are placed outside the solenoid. The data used for  
 474 the search presented in this thesis is collected after the long shutdown 1. After discussing each  
 475 part of CMS in their Run 1 configuration, Section 2.2.4 elaborates on their different upgrades  
 476 for the data collected in Run 2.

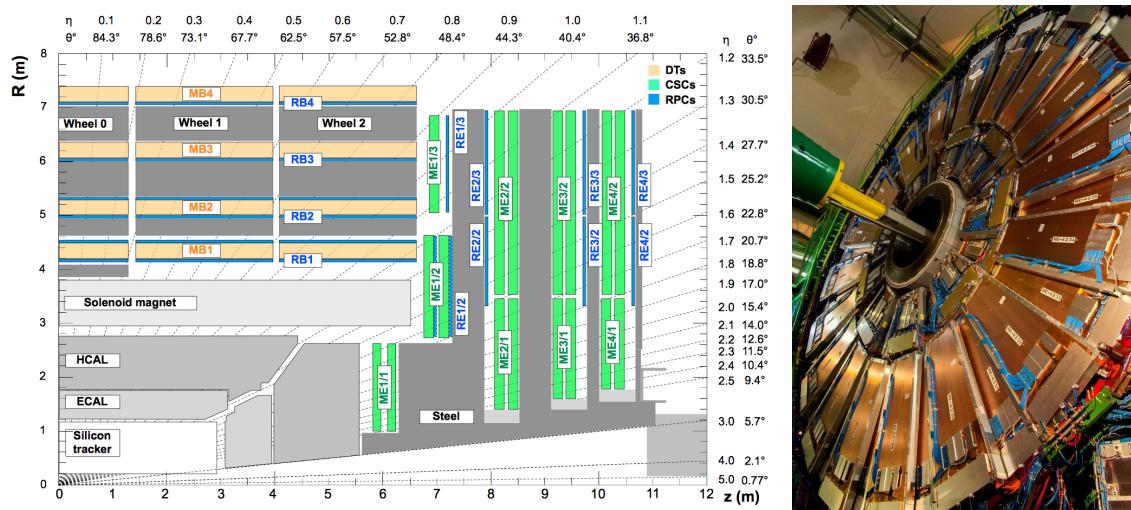


**Figure 2.6:** Schematic view of the CMS detector in the Run 1 configuration. The longitudinal view of one quarter of the detector is given at the top, while the transversal view is shown at the bottom. The muon system barrel elements are denoted as  $MBZ/N/S$ , where  $Z = -2 \dots +2$  is the barrel wheel number,  $N = 1 \dots 4$  the station number and  $S = 1 \dots 12$  the sector number. Similarly, the steel return yokes are denoted as  $YBZ/N/S$ . The solenoid is denoted as  $CB0$ , while the hadronic calorimeter is denoted as  $HE$  (endcap)/  $HB$  (barrel)/  $HF$  (forward) and the electromagnetic calorimeter as  $EE$  (endcap)/  $EB$  (barrel). The green part represents the tracking system (tracker + pixel). Figure taken from [74].

477 **2.2.2.1 Muon system**

478 The outermost part of CMS consists of the muon system. The magnet return yoke is interleaved  
 479 with gaseous detector chambers for muon identification and momentum measurement. The  
 480 barrel contains muon stations arranged in five separate iron wheels, while in the endcap four  
 481 muon stations are mounted onto three independent iron discs on each side. Each barrel wheel  
 482 has 12 sectors in the azimuthal angle.

483 The muon system is divided into three parts, shown in [Figure 2.7](#). The muon rate and neutron  
 484 induced backgrounds are small and the magnetic field is very low for the barrel, thus CMS can  
 485 use drift tube (DT) chambers. For the endcaps however, the muon and background flux is much  
 486 higher and there is a need to use cathode strip chambers (CSC) which are able to provide a  
 487 faster response, higher granularity and have a better resistance against radiation. In order to  
 488 form a redundant trigger system, resistive plate chambers (RPC) are added. This makes a total  
 of 250 DT, 540 CSC and 610 RPC chambers. In [Figure 2.6](#) the arrangement is shown.



**Figure 2.7:** (Left) Schematic view of one quarter of the CMS muon system in the Run 1 configuration. The cathode strip chambers (CSC) are shown in green, the drift tubes (DT) are shown in yellow, while the resistive plate chambers (RPC) are shown in blue. Figure taken from [74]. (Right) Cathode strip chambers (ME+4/2 chambers on YE+3). Photo taken from [75].

489

490 Providing a measurement for  $|\eta| < 1.2$ , the DT chambers in the barrel are on average  $2 \times 2.5\text{m}^2$   
 491 in size and consist of 12 layers of DT cells<sup>5</sup> arranged in three groups of four. The  $r\phi$  coordinate is  
 492 provided by the two outside groups, while the middle group measures the  $z$  coordinate. For the  
 493 outer muon station, the DT chambers contain only 8 layers of DT cells, providing a muon position  
 494 in the  $r\phi$  plane. There are four CSC stations in each endcap, providing muon measurements for  
 495  $0.9 < |\eta| < 2.4$  (Run 1 configuration). These CSCs are multi-wired proportional chambers that  
 496 consist of 6 anode wire planes crossed by 7 copper strips cathode panels in a gas volume. The  
 497  $r$  coordinate is provided by the copper strips, while the  $\phi$  coordinate comes from the anode  
 498 wires, giving a two dimensional position measurement. There are six layers of RPCs in the

<sup>5</sup>The DT cells are 4 cm wide gas tubes with positively charged stretched wires inside.

499 barrel muon system and one layer into each of the first three stations of the endcap. They are  
 500 made from two high resistive plastic plates with an applied voltage and separated by a gas  
 501 volume. Read-out strips mounted on top of the plastic plates detect the signal generated by a  
 502 muon passing through the gas volume. The RPCs provide a fast response with a time resolution  
 503 of 1 ns and cover a range of  $|\eta| < 1.8$  for the Run 1 configuration.

504 The muon system provides triggering on muons, identifying muons and improves the momen-  
 505 tum measurement and charge determination of high  $p_T$  muons. On top of the muon system,  
 506 a fraction of the muon energy is deposited in the electromagnetic calorimeter, the hadronic  
 507 calorimeter, and outer calorimeter. The high magnetic field enables an efficient first level trigger  
 508 and allows a good momentum resolution of  $\Delta p/p \approx 1\%$  for a  $p_T$  of 100 GeV and  $\approx 10\%$  for a  
 509  $p_T$  of 1 TeV. There is an efficient muon measurement up to  $|\eta| < 2.4$ .

### 510 2.2.2.2 Solenoid

511 Making use of the knowledge of previous experiments like ALEPH and DELPHI at LEP and H1  
 512 at HERA, CMS chose for a large super conducting solenoid with a length of 12.9 m and a  
 513 inner bore of 5.9 m [72]. With 2168 turns, a current of 18.5 kA resulting in a magnetic field of  
 514 3.8 T, and a total energy of 2.7 GJ, a large bending power can be obtained for a modestly-sized  
 515 solenoid. In order to ensure a good momentum resolution in the forward regions, a favourable  
 516 length/radius was necessary. In [Figure 2.8](#), a photo of the CMS solenoid is shown.

517 The solenoid uses a high-purity aluminium stabilised conductor with indirect cooling from  
 518 liquid helium, together with full epoxy impregnation. A four-layer winding is implemented that  
 519 can withstand an outward pressure of 64 atm. The NbTi cable is co-extruded by pure aluminium  
 520 that acts as a thermal stabilizer and has an aluminium alloy for mechanical reinforcement. The  
 521 return of the magnetic field is done by five wheels, noted by YB in [Figure 2.6](#).

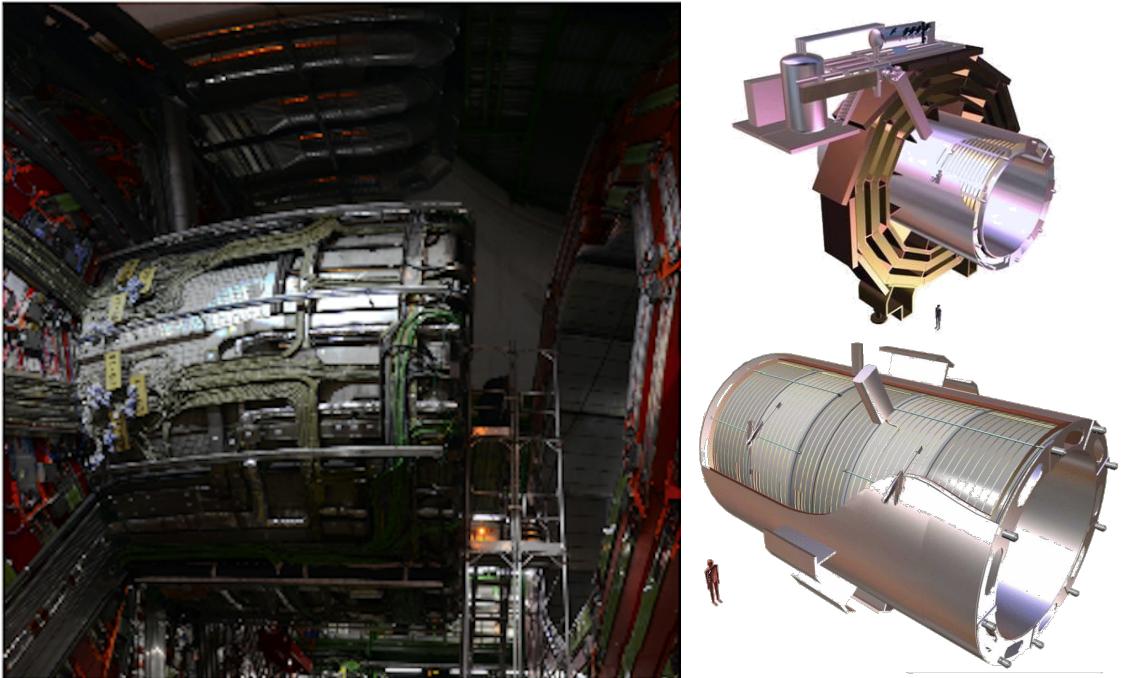
### 522 2.2.2.3 Hadronic calorimeter

523 The hadronic calorimeter (HCAL) is dedicated to precisely measure the energy of charged and  
 524 neutral hadrons via a succession of absorbers and scintillators. This makes it crucial for physics  
 525 analyses with hadronic jets or missing transverse energy. The HCAL barrel extends between  
 526  $1.77 < r < 2.95$  m, where  $r$  is the radius in the transverse plane with respect to the beam. Due  
 527 to space limitations, the HCAL needs to be as small as possible and is made from materials  
 528 with short interaction lengths<sup>6</sup>. On top of this, the HCAL should be as hermetic as possible and  
 529 extend to large absolute pseudo rapidities such that it can provide a good measurement of the  
 530 missing transverse energy.

531 The quality of the energy measurements is dependent on the fraction of the hadronic shower  
 532 that can be detected. Therefore, the HCAL barrel (HB) inside the solenoid is reinforced by an  
 533 outer hadronic calorimeter between the solenoid and muon detectors (HO, see [Figure 2.9](#)), using  
 534 the solenoid as extra absorber. This increases the thickness to 12 interaction lengths. The HB

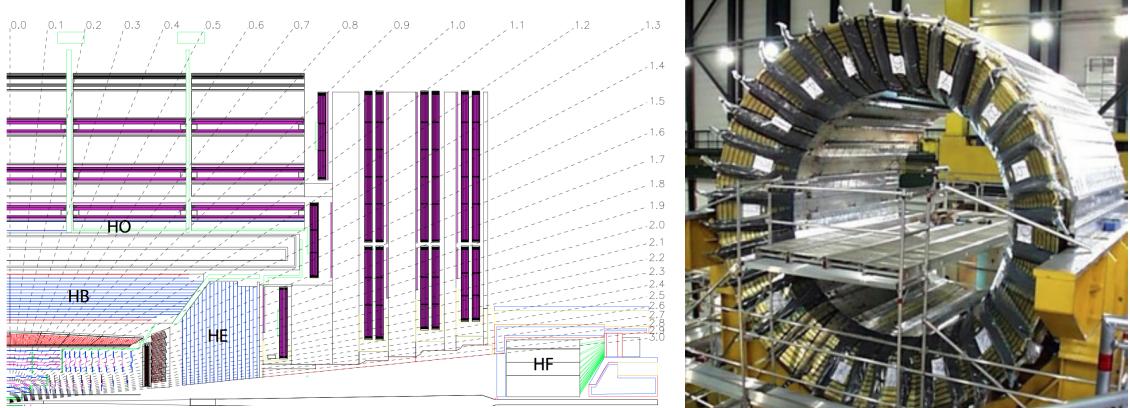
---

<sup>6</sup>Here the interaction length is the nuclear interaction length and this is the length needed for absorbing 36.7% of the relativistic charged particles. For the electromagnetic calorimeter this is defined in radiation length  $X_0$ . The radiation length is the mean distance over which a high energy electron loses all but  $1/e$  of its energy by bremsstrahlung.



**Figure 2.8:** (Left) CMS solenoid during the long shutdown in 2013. (Right) An impression of the solenoid magnet taken from [76].

535 and HO provide measurements for  $|\eta| < 1.3$ , while an endcap on each side (HE,  $1.3 < |\eta| < 3$ )  
 536 and a forward calorimeter (HF,  $3.0 < |\eta| < 5.2$ ) extend the pseudo rapidity range.



**Figure 2.9:** (Left) Longitudinal view of the CMS detector showing the locations of the HB, HE, HO, and HF calorimeters. Figure taken from [61]. (Right) CMS barrel calorimeter. Photo taken from [77].

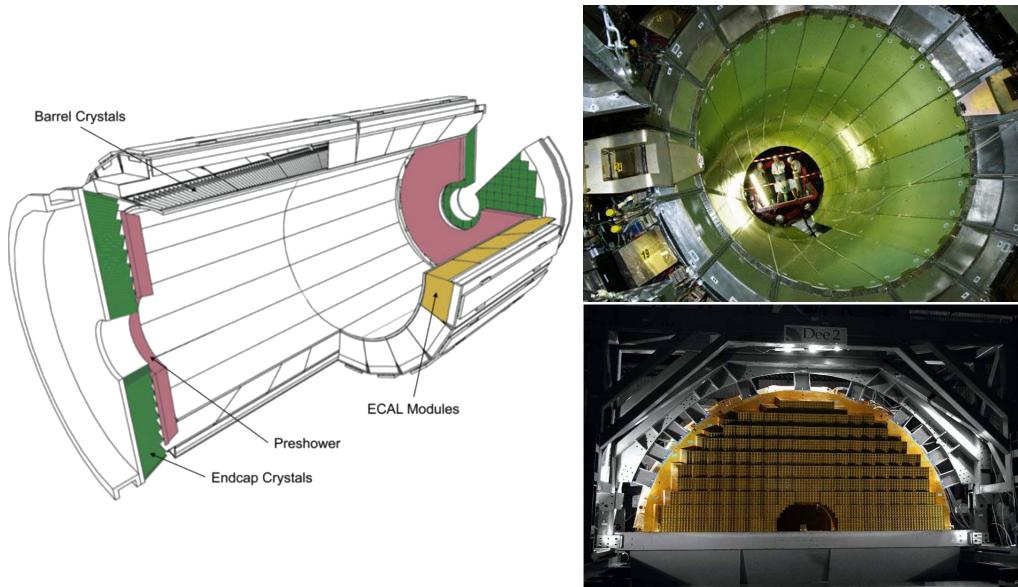
537 The HB is made of 16 absorber plates where most of them are built from brass and others are  
 538 made from stainless steel and is about five to ten interaction lengths thick. It is divided in  $\eta \times \phi$   
 539 towers and contains 2592 read-out channels. The HO complements the HB and extends the  
 540 reach up to twelve interaction lengths. This subsystem contains 2160 read-out channels. The  
 541 HE is also composed of brass absorber plates and has a thickness corresponding to approximately

ten interaction lengths, with 2592 read-out channels. The HF experiences intense particle fluxes with an expected energy of 760 GeV deposited on average in a proton interaction at a centre-of-mass of 14 TeV, compared to 100 GeV in the rest of the detector. Therefore, these are Cherenkov light detectors made of radiation hard quartz fibres. The main causes of such large energy events are high energy muons and charged particles from late showering hadrons. During Run 1, it became clear that the glass windows of the photon multiplier tubes (PMTs) had to be replaced, which was done during LS1 [78]. The HF represents 1728 read-out channels.

The HCAL and electromagnetic calorimeter combined, can measure the hadron energy with a resolution  $\Delta E/E \approx 100\% \sqrt{E[\text{GeV}]} + 5\%$ .

#### 2.2.2.4 Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) is designed to measure the energy of photons and electrons and covers  $|\eta| < 3$ . It is an hermetic, homogeneous detector and consists of 75 848 lead tungstate ( $\text{PbWO}_4$ ) crystals. These crystals have a fast response time - 80% of the light is emitted within 25 ns- and are radiation hard. The electromagnetic showers produced by passing electrons or photons, ionize the crystal atoms which emit a blue-green scintillation light, that is collected by silicon avalanche photodiodes (APDs) in the barrel and vacuum phototriodes (VPTs) in the endcaps. The crystals and the APD response are sensitive to temperature changes and require a stable temperature.



**Figure 2.10:** (Left) Schematic cross section of the electromagnetic calorimeter taken from [61]. (Right top) The ECAL barrel during construction [79]. (Right bottom) One half of an EE [80].

There are three parts: a central barrel (EB), an endcap region (EE) and a preshower (ES) (Figure 2.10). The EB has an inner radius of 129 cm and corresponds to a pseudo rapidity of  $0 < |\eta| < 1.479$ . At a distance of 314 cm from the vertex and covering a pseudo rapidity of  $1.479 < |\eta| < 3.0$ , are the EE. They consist of semi-circular aluminium plates from which structural units of  $5 \times 5$  crystals (super crystals) are supported. The ES is placed in front of

565 the crystal calorimeter over the endcap pseudo rapidity range with two planes of silicon strip  
 566 detectors as active elements.

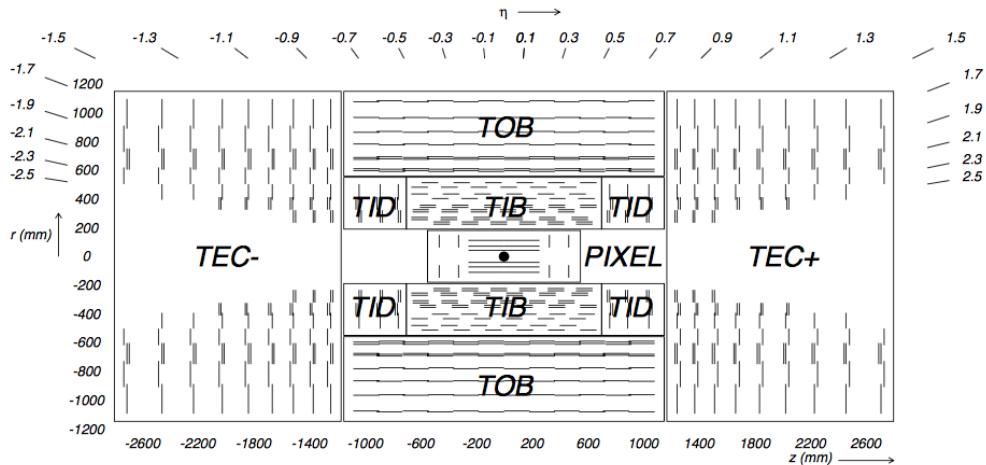
The electromagnetic shower will typically involve more than one channel. More than 90% of the energy of a 35 GeV electron or photon is contained in a  $5 \times 5$  matrix of crystals. Therefore, a clustering algorithm is performed in order to associate the energy deposits to the particles impinging the calorimeter. The achieved precision [81] for the barrel is  $2 \times 10^{-3}$  rad in  $\phi$  and  $10^{-3}$  in  $\eta$ . For the endcaps this is  $5 \times 10^{-3}$  rad in  $\phi$  and  $2 \times 10^{-3}$  in  $\eta$ . The energy is reconstructed by a supercluster algorithm, taking into account energy radiated via bremsstrahlung or conversion [61]. The energy resolution is given by

$$\frac{\sigma(E)}{E} = \frac{2.8\%}{\sqrt{E}} \oplus \frac{0.128}{E(\text{GeV})} \oplus 0.3\%, \quad (2.5)$$

567 in the absence of a magnetic field, where the contributions come from the stochastic, noise and  
 568 constant terms respectively. The dominating term is the constant term ( $E_{\text{shower}} \approx 100$  GeV)  
 569 and thus the performance is highly dependent on the quality of calibration and monitoring.

#### 570 2.2.2.5 Inner tracking system and operations

571 The tracking system (tracker) [82] is the detecting unit closest to the point of interaction.  
 572 Responsible for the reconstruction of trajectories from charged particles with  $|\eta| < 2.5$  that  
 573 are bent by the magnetic field, it provides a measurement of the momentum. The tracker is  
 574 also responsible for the determination of the interaction point or vertex. It should be able to  
 575 provide high granularity as well as fast read-out, and be able to endure high radiation. For  
 576 these reasons, the CMS collaboration chose silicon detector technology.



**Figure 2.11:** Schematic cross section through the CMS tracker. Each line represents a detector module. Double lines indicate back-to-back modules that deliver stereo hits. Figure taken from [61].

577 The tracking system consists of a cylinder of 5.8 m long and 2.5 m in diameter. It is immersed  
 578 in a co-axial magnetic field of 3.8 T provided by the solenoid. As shown in Figure 2.11, the  
 579 tracker is built up from a large silicon strip tracker with a small silicon pixel tracker inside. The

580 inner pixel region ( $4.4 < r < 10.2$  cm), gets the highest flux of particles. Therefore, pixel silicon  
 581 sensors of  $100 \times 150 \mu\text{m}^2$  are used. It consists of three cylindrical barrels that are complemented  
 582 by two discs of pixel modules at each side. The silicon strip tracker ( $20 < r < 116$  cm) has three  
 583 subdivisions. The Tracker Inner Barrel and Discs (TIB, TID, see [Figure 2.13](#)) are composed  
 584 of four barrel layers accompanied by three discs at each end. The outer part of the tracker -  
 585 Tracker Outer Barrel (TOB) - consists of 6 barrel layers. In the outer discs, there are nine discs  
 586 of silicon sensors, referred to as Tracker End Caps (TEC).



**Figure 2.12:** The pixel barrel being re-installed after the Long Shutdown in 2015, around the beam pipe at CMS [83].



**Figure 2.13:** First half of the inner tracker barrel, consisting of three layers of silicon modules [84].

587 The pixel detector, shown in [Figure 2.12](#), has 1440 modules that cover an area of about  $1 \text{ m}^2$   
 588 and have 66 million pixels. It provides a three-dimensional position measurement of the hits  
 589 arising from the interaction from charged particles with the sensors. In transverse coordinate  
 590 ( $r\phi$ ), the hit position resolution is about  $10 \mu\text{m}$ , while  $20\text{-}40 \mu\text{m}$  is obtained in the longitudinal  
 591 coordinate ( $z$ ). The sensor plane position provides the third coordinate. The TIB/TID, shown in  
 592 [Figure 2.13](#), delivers up to four  $r\phi$ -measurements using  $320 \mu\text{m}$  thick silicon micro-strip sensors.  
 593 These sensors are placed with their strips parallel to the beam axis in the barrel and radial  
 594 in the discs. In the TIB, the first two layers have a strip pitch of  $80 \mu\text{m}$ , while the remaining  
 595 two have a strip pitch of  $120 \mu\text{m}$ . This leads to a respective single point resolution of  $23 \mu\text{m}$   
 596 and  $35 \mu\text{m}$ . For the TID, the pitch varies between  $100 \mu\text{m}$  and  $141 \mu\text{m}$ . The TOB provides six  
 597  $r\phi$ -measurements with a single point resolution of  $53 \mu\text{m}$  in the first four layers, and  $35 \mu\text{m}$  in  
 598 the last two layers. It consists of  $500 \mu\text{m}$  thick microstrip sensors with strip pitches of  $183 \mu\text{m}$   
 599 (first 4 layers) or  $122 \mu\text{m}$  (last two layers). The TEC provides up to 9  $\phi$ -measurements via 9  
 600 discs consisting of up to 7 rings of silicon microstrip sensors of  $97 \mu\text{m}$  to  $184 \mu\text{m}$  average pitch.

601 A second co-ordinate measurement ( $z$  in the barrel,  $r$  on the discs) is provided through the  
 602 use of a second micro strip detector module mounted back-to-back with a stereo angle of 100  
 603 mrad. This is done on the modules in the first two layers and rings of the TIB, TID, and TOB, as  
 604 well as rings 1,2, and 5 of the TECs (blue line in [Figure 2.11](#)). The resolution in the  $z$  direction  
 605 is approximately  $230 \mu\text{m}$  in the TIB and  $530 \mu\text{m}$  in the TOB, and is varying with pitch in the  
 606 TID and TEC. To allow overlay and avoid gaps in acceptance, each module is shifted slightly  
 607 in  $r$  or  $z$  with respect to its neighbouring modules within a layer. With this detector lay-out,  
 608 at least nine points per charged particle trajectory can be measured in an  $|\eta|$  range up to 2.4,  
 609 where at least four of them being two dimensional. The CMS silicon tracker provides 9.3 million  
 610 read-out channels and covers an active area of about  $198 \text{ m}^2$ .

611 **2.2.3 Data acquisition**

612 At a design luminosity of  $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ , the proton interaction rate exceeds 1 GHz. Given the  
 613 large size of an event (about 1 MB), the high crossing rate, and that typically tens of collisions  
 614 happen at the same time, it is impossible for the CMS experiment to store all the data generated.  
 615 In order to deal with the large amount of data, a two level trigger system has been put in place.  
 616 The first level (Level-1) is a custom hardware system, while a second high level trigger (HLT) is  
 617 software based, running on a large farm of computers.

618 **CMS Level-1 Trigger**

619 The Level-1 Trigger has to be a flexible, maintainable system, capable of adapting to the  
 620 evolving physics programme of CMS [85]. Its output rate is restricted to 100 kHz imposed  
 621 by the CMS read-out electronics. It is implemented by custom hardware and selects events  
 622 containing candidate objects - e.g. ionization deposits consistent with a muon, or energy clusters  
 623 corresponding to an electron / photon / tau lepton / missing transverse energy / jet. Collisions  
 624 with large momenta can be selected with the use of the scalar sum of the transverse momenta  
 625 of the jets.

626 By buffering the raw data from the CMS subdetectors in front-end drivers, the Level-1 Trigger  
 627 has a pipeline memory of 3.2  $\mu\text{s}$  to decide whether to keep an event or reject it. The trigger  
 628 primitives (TP) from the calorimeters and muon detectors are processed in several steps and  
 629 combined into a global trigger. This information is then combined with the input from the other  
 630 subsystems for the HLT. The separate inputs are synchronized to each other and the LHC orbit  
 631 clock and sent to the global trigger module. Here, Level-1 Trigger algorithms are performed  
 632 within 1  $\mu\text{s}$  to decide whether to keep the event.

633 **CMS HLT Trigger**

634 The HLT is an array of commercially available computers with a programmable menu that has  
 635 an output rate of on average 400 Hz for off-line event storage. The data processing is based on  
 636 an HLT path. This is a set of algorithmic steps to reconstruct objects to define selection criteria.  
 637 Here, the information of all subdetectors can be used to perform algorithms on higher level  
 638 reconstructed objects.

639 **2.2.4 Phase 1 upgrades**

640 Before the start of taking collision data for 13 TeV operations on 3 June 2015, CMS had a  
 641 long shutdown (LS1) [86]. During this shutdown, the section of the beryllium beam pipe  
 642 within CMS was replaced by a narrower one. This operation required the pixel detector to be  
 643 removed and reinserted into CMS. During Run 2, higher particle fluxes with respect to Run  
 644 1 are expected. To avoid longterm damage caused by the intense particle flux at the heart of  
 645 CMS, the tracker has been made ready to operate at much lower temperature than during Run  
 646 1. The electromagnetic calorimeter preshower system was damaged during Run 1, therefore the  
 647 preshower discs were removed, repaired and reinstalled successfully inside CMS in 2014. To  
 648 help the discrimination between interesting low momentum muons coming from collisions and  
 649 muons caused by backgrounds, a fourth triggering and measurement station for muons was

650 added in each of the endcaps. Several new detectors were installed into CMS for measuring the  
 651 collision rate within the detector and to monitor beam related backgrounds.

652 During the LS1, the muon system underwent major upgrades [87, 88]. In the fourth station  
 653 of each endcap, the outermost rings of CSC and RPC chambers were completed, providing an  
 654 angular coverage of  $1.2 < |\eta| < 1.8$  for Run 2, increasing the system redundancy, and allowing  
 655 tighter cuts on the trigger quality. In order to reduce the environmental noise, outer yoke discs  
 656 have been placed on both sides for the endcaps. At the innermost rings of the first station,  
 657 the CSCs have been upgraded by refurbishing the read-out electronics to make use of the full  
 658 detector granularity instead of groups of three as was the case for Run 1. In Figure 2.7 (right),  
 659 the refurbishing of the CSCs is shown.

660 Since the HF experiences intense particle fluxes, it became clear during Run 1 that the glass  
 661 windows of the PMTs need replacing. For the ECAL in Run 1, the energy reconstruction happened  
 662 via a weighted sum of the digitized samples [89]. For Run 2 however, the reconstruction had to  
 663 be made more resistant for out-of-time pileup and a multi-fit approach has been set into place.  
 664 In this approach, the pulse shape is modelled as a sum of one in-time pulse plus the out-of-time  
 665 pulses [81]. The energy resolution is better than 2% in the central barrel region and 2-5 %  
 666 elsewhere.

During the first data taking period of the LHC (2010 to 2013), the tracker operated at +4°C. With the higher LHC beam intensities from 2015 onwards, the tracker needs to be operated at much lower temperatures. The reason for this is that with intense irradiation, the doping concentration changes, the leakage current increases proportional to the fluence and the charge collection efficiency decreases due to charge trapping. Mostly the leakage current ( $I$ ) is affected by the temperature change:

$$I \propto T^2 e^{-\frac{E_g}{2kT}}, \quad (2.6)$$

667 where  $T$  is the operating temperature,  $E_g$  the band gap and  $k$  the Boltzmann constant. There is  
 668 approximately a factor 15 between the leakage currents at room temperatures and at -10 °C.

669 During the LS1, the CMS cooling plant was refurbished [90] and the fluorocarbon cooling  
 670 system overhauled. To help suppressing the humidity inside the tracker, new methods for vapour  
 671 sealing and insulation were applied. Furthermore, several hundred high-precision sensors are  
 672 used to monitor the humidity and temperature. In order to get as dry air as possible, a new  
 673 dry-gas plant provides eight times more dry gas (air or nitrogen) than during the first run, and  
 674 allows regulation of the flow. As a final addition, the cooling bundles outside the tracker are  
 675 equipped with heater wires and temperature sensors in order to maintain safe operations above  
 676 the cavern dew point. For the data taking in 2015-2016, the tracker operated at -15 °C.

677 In Run 2, with the increase in centre-of-mass energy and a higher luminosity, a larger number  
 678 of simultaneous inelastic collisions per crossing is expected with respect to Run 1. For this, the  
 679 CMS Level-1 has been upgraded [91]. All hardware, software, databases and the timing control  
 680 system have been replaced for Run 2, where the main changes are that the muon system now  
 681 uses the redundancy of the muon detector system earlier to make a high resolution muon trigger.

682 Other upgrades are also performed, including providing the global trigger with more Level-1  
683 Trigger algorithms.

684 After the first half of Run 2, the innermost part of detection in CMS (pixel detector) was  
685 replaced, enhancing the particle tracking capabilities of CMS. The data used in the framework  
686 of this thesis however is from before this upgrade. More information on the Pixel upgrade can  
687 be found in Refs. [92, 93].

### 688 2.2.5 CMS computing model

689 The selected data is stored, processed and dispersed via the Worldwide Large Hadron Collider  
690 Computing Grid (WLCG) [94, 95]. This has a tiered structure that functions as a single, coherent  
691 system.

692 At CERN and the Wigner Research Centre for physics, a single Tier-0 is located. The raw data  
693 collected by the experiments is archived here, and a first reconstruction of the data is done.  
694 This data is then already in a file format usable for physics analysis. Furthermore, it is able to  
695 reprocess data when new calibrations become available. The Tier-0 site distributes this data  
696 to a total of 14 Tier-1 centres. They carry out data reprocessing and store real data as well as  
697 simulated data. The Tier-1 further distributes the data to over 150 Tier-2 centres. These make  
698 the data accessible for physics analysis and are also being used for the production of simulated  
699 data. The data is made accessible for physicists around the world. For CMS, the Tier-0 site at  
700 CERN reconstructs the full collision events and the backup of the data is sent to seven Tier-1  
701 computer centres in France, Germany, Italy, Spain, Taiwan, UK, and the US. At the Tier-1 sites  
702 the events are again reconstructed using refined calibration constants. The patterns are created  
703 and the more complex events are sent to forty Tier-2 centres for specific analysis tasks.

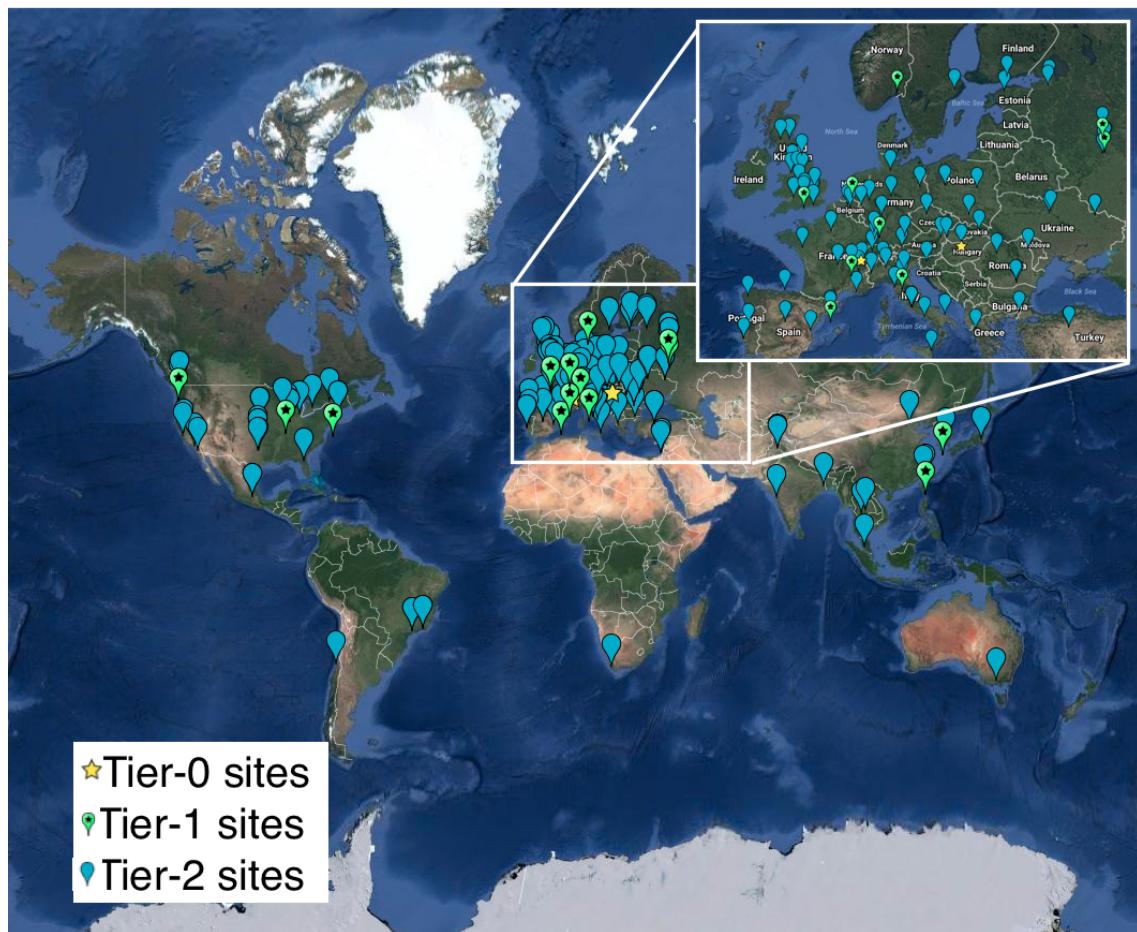


Figure 2.14: Worldwide LHC Computing Grid in 2017 [96].

# Analysis techniques

# 3

705 In order to study the collisions coming from high energy experiments, many tools have been  
 706 developed. In [Section 3.1](#), the physics of hadron collision at high energies are presented. These  
 707 insights are used to generate events via Monte Carlo event generators, explained in [Section](#)  
 708 [3.2](#). Machine learning helps to differentiate between signal- and background like events. In  
 709 [Section 3.3](#), the multivariate technique of boosted decision trees is explained. This yields  
 710 powerful discriminants for separating signal and background events and provides distributions  
 711 for template-based maximum likelihood fits. The fitting method used in the search presented  
 712 in this thesis is discussed in [Section 3.4](#).

## 713 3.1 Hadron collisions at high energies

All partons can be approximated as free when there is sufficiently high momentum transfer in hadron collisions. This makes it possible to treat a hadron-hadron scattering as a single parton-parton interaction. The momentum of the parton can then be expressed as a fraction of the hadron momentum

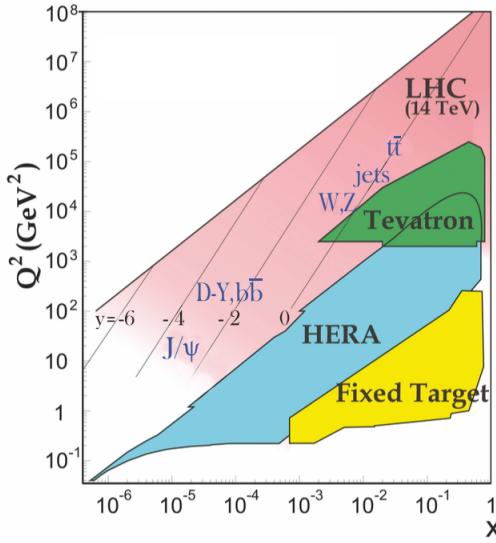
$$\vec{p}_{\text{parton}} = x \vec{p}_{\text{hadron}}, \quad (3.1)$$

where  $x$  is referred to as the Björken scaling variable. The interaction  $p_A p_B \rightarrow X$  can then be factorised in terms of partonic cross sections  $\hat{\sigma}_{ij \rightarrow X}$  [97]

$$\sigma_{p_A p_B \rightarrow X} = \sum_{ij} \iint dx_1 dx_2 f_i^A(x_1, Q^2) f_j^B(x_2, Q^2) d\hat{\sigma}_{ij \rightarrow X}, \quad (3.2)$$

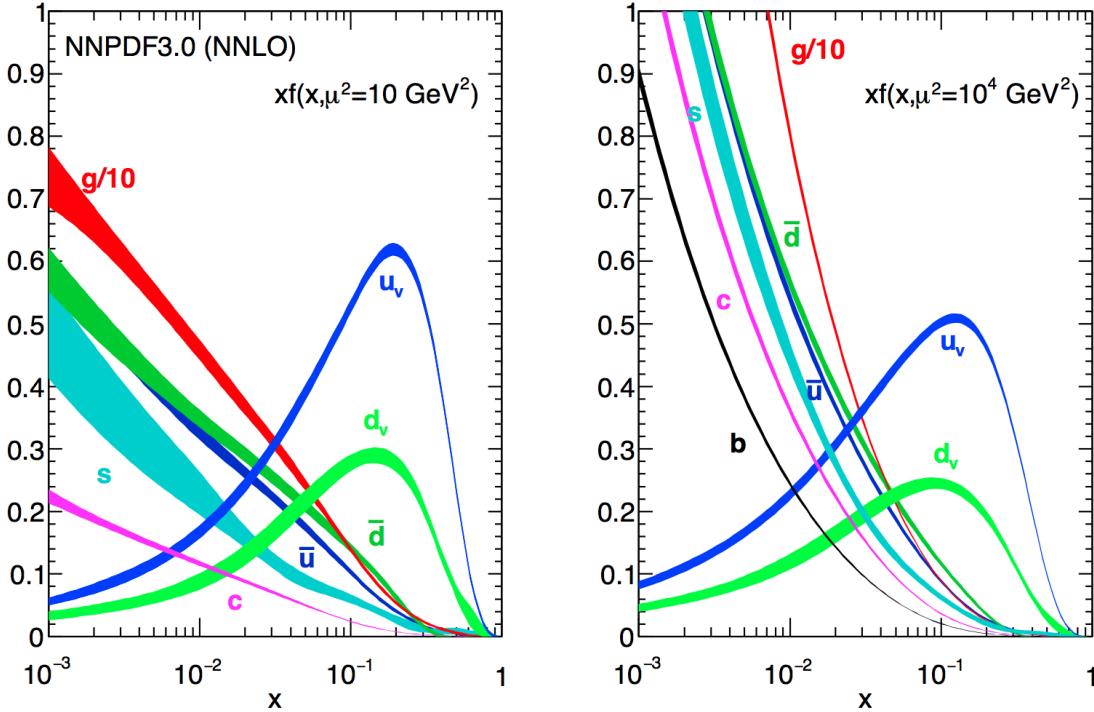
714 where  $i$  and  $j$  are the partons resolved from protons A and B. The parton density functions  
 715 (PDF) are denoted as  $f_i(x_j, Q^2)$ , and  $Q^2$  is the factorisation scale more commonly denoted as  
 716  $\mu_F$ . This factorisation scale represents the energy at which the hadronic interaction can be  
 717 expressed as a product of the partonic cross section and the process independent PDF. In [Figure](#)  
 718 [3.1](#), the kinematic regions in  $x$  and  $\mu_F$  are shown for fixed target and collider experiments.

719 The parton density functions (PDF) [98–100] represent the momentum distribution of the  
 720 proton amongst its partons at an energy scale  $\mu_F$ . These functions are obtained from global fits  
 721 to data since they can not be determined from first principles. From measurements on deep



**Figure 3.1:** Kinematic regions in momentum fraction  $x$  and factorisation scale  $Q^2$  probed by fixed-target and collider experiments. Some of the final states accessible at the LHC are indicated in the appropriate regions, where  $y$  is the rapidity. In this figure, the incoming partons have  $x_{1,2} = (M/14 \text{ TeV})e^{\pm y}$  with  $Q = M$  where  $M$  is the mass of the state shown in blue in the figure. For example, exclusive  $J/\psi$  and  $\Upsilon$  production at high  $|y|$  at the LHC may probe the gluon PDF down to  $x \sim 10^{-5}$ . Figure taken from [6].

722 inelastic scattering using lepton-proton collision by the HERA collider [101], supplemented  
 723 with proton-antiproton collisions from the Tevatron [102], and proton collision data from the  
 724 ATLAS, CMS and LHCb collaborations at the LHC (Run 1) [103] the PDFs are determined and  
 725 included in global PDF sets known as the PDF4LHC recommendation [100]. Their measurement  
 726 at scale  $\mu_F$  is extrapolated to higher energies by use of the DGLAP equations [104]. Once  
 727 these PDFs are known, the cross section of a certain process can be calculated and used as  
 728 input for the Monte Carlo generators used to make the simulated data samples at the LHC. In  
 729 the framework of this thesis, the NLO PDF4LHC15\_100 set is used. This set is an envelope of  
 730 three sets: CT14, MMHT2014 and NNPDF3.0 [100]. As illustration, the dependency of the PDFs  
 731 on the momentum fraction  $x$  is shown for the NNPDF3.0 set on hadronic scale  $\mu_F^2 = 10 \text{ GeV}^2$   
 732 and LHC scale  $\mu_F^2 = 10^4 \text{ GeV}^2$  in Figure 3.2. The gluon density dominated for most values of  
 733 the momentum fraction, implying that it is easier to probe gluons than the quarks. When the  
 734 Björken scale is to one, the parton densities of the valence quarks of the proton, up and down  
 735 quarks, dominate over the gluon density. The sea quarks originating from gluon splitting, the  
 736 charm, anti-up, and anti-down quarks, have lower densities in general for the proton. The  
 737 resolution scale  $Q^2$  is typically taken to be the energy scale of the collision. For the top quark pair  
 738 production a scale of  $Q^2 = (350 \text{ GeV})^2$  is chosen, meaning that the centre-of-mass energy of the  
 739 hard interaction is about twice the top quark mass. The uncertainty on the parton distributions  
 740 is evaluated using the Hessian technique [105], where a matrix with a dimension identical to  
 741 the number of free parameters needs to be diagonalised. In the case of PDF4LHC15\_100 set, this  
 742 translates into 100 orthonormal eigenvectors and 200 variations of the PDF parameters in the  
 743 plus and minus direction.



**Figure 3.2:** The momentum fraction  $x$  times the parton distribution functions  $f(x)$ , where  $f = u_v, d_v, \bar{u}, \bar{d}, s, c$ , or  $g$  as a function of the momentum fraction obtained in the NNLO NNPDF3.0 global analysis at factorisation scales  $\mu^2 = 10 \text{ GeV}^2$  (left) and  $\mu^2 = 10^4 \text{ GeV}^2$  (right), with  $\alpha_s(M_Z^2) = 0.118$ . The gluon PDF has been scaled down by a factor of 0.1. Figure taken from [6].

Quantum fluctuations can cause divergences at high energies. This is solved by introducing a renormalization scale  $\mu_R$  to redefine physical quantities, making the theory still able to describe the experimental regime. A consequence of this method is that the coupling constants will run as a function of  $\mu_R$ . Beyond the renormalization scale, the high energy effects such as the loop corrections to propagators (self energy) are absorbed in the physical quantities through a renormalization of the fields. In particular, the running behaviour of the strong coupling constant<sup>1</sup>  $\alpha_S$  is found to be

$$\alpha_S = \frac{\alpha_S(\mu_0^2)}{1 + \alpha_S(\mu_0^2) \frac{33-2n_f}{12\pi} \ln\left(\frac{|\mu_R^2|}{\mu_0^2}\right)}, \quad (3.3)$$

744 with  $n_f$  the number of quarks and  $\mu_0$  the reference scale at which the coupling is known. The  
 745 current world average of the strong coupling constant at the Z boson mass is  $\alpha_S(\mu_R = m_Z) =$   
 746  $0.1181 \pm 0.0011$  [6]. From Equation 3.3 one can see easily that the coupling strength decreases  
 747 with increasing renormalization scale, this is known as asymptotic freedom. Additionally,  
 748 following the behaviour of  $\alpha_S(\mu_R^2)$ , a limit  $\Lambda_{\text{QCD}} \approx 200 \text{ MeV}$  is found for which  $\alpha_S$  becomes  
 749 larger than one. Under this limit, the perturbative calculations of observables can no longer be  
 750 done.

<sup>1</sup>The strong coupling constant is defined as  $\alpha_S = \frac{g_S^2}{4\pi}$ .

Cross sections can be written in terms of interacting vertices contributing to the matrix element (ME) originating from elements of a perturbative series [106], allowing them to be expanded as a power series of the coupling constant  $\alpha$

$$\sigma = \sigma_{\text{LO}} \left( 1 + \left( \frac{\alpha}{2\pi} \right) \sigma_1 + \left( \frac{\alpha}{2\pi} \right)^2 \sigma_2 + \dots \right). \quad (3.4)$$

751 Leading order (LO) accuracy contains the minimal amount of vertices in the process, then  
 752 depending on where the series is cut-off one speaks of next-to-leading order (NLO), or next-to-  
 753 next-to-leading order (NNLO) accuracy in  $\alpha$ . Predictions including higher order corrections  
 754 tend to be less affected by theoretical uncertainties originating from a variation of the chosen  
 755 renormalization and factorisation scales.

## 756 3.2 Event generation

757 In order to compare reconstructed data with theoretical predictions, collision events are gener-  
 758 ated and passed through a simulation of the CMS detector and an emulation of its read-out. For  
 759 the detector simulation, a so-called Full Simulation package [107, 108] based on the Geant4  
 760 toolkit [109] is employed. This allows detailed simulations of the interactions of the particles  
 761 with the detector material.

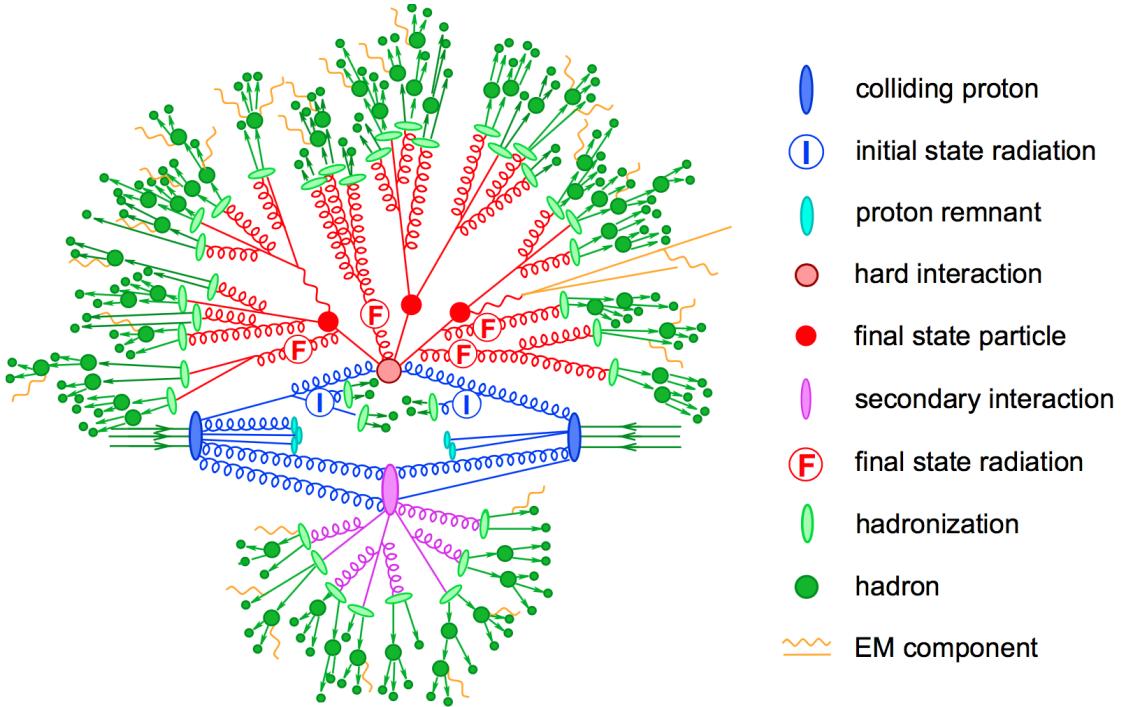
### 762 3.2.1 Fundamentals of simulating a proton collision

763 The generation of  $\text{pp} \rightarrow \text{X}$  events is subdivided into sequential steps [110–112], as shown in  
 764 Figure 3.3.

765 The interaction of two incoming protons is often soft and elastic leading to events that are not  
 766 interesting in the framework of this thesis. More intriguing are the hard interactions between  
 767 two partons from the incoming protons. The event generation starts from the matrix elements of  
 768 a hard scattering process of interest. The corresponding cross section integral is sampled using  
 769 Monte Carlo techniques and the resulting sample of events reflects the probability distribution  
 770 of a process over its final state phase space. A parton shower (PS) program is then used to  
 771 simulate the hadronisation of final state partons, coming from the sample of events of the hard  
 772 interaction, into hadrons which then decay further. On top of this, radiation of soft gluons or  
 773 quarks from initial or final state partons is simulated. These are respectively referred to as  
 774 initial state radiation (ISR) or final state radiation (FSR). The contributions from soft secondary  
 775 interactions, the so-called underlying event (UE), and colour reconnection effects are also taken  
 776 into account. A brief overview of the programs used for the event generation of the signal and  
 777 main background processes used in the search presented in this thesis, is given in Section 3.2.2.

### 778 3.2.2 Programs for event generation

779 The FEYNRULES package [113] allows for the calculation of the Feynman rules in momentum  
 780 space for any quantum field theory model. By use of a Lagrangian, the set of Feynman rules asso-  
 781 ciated with this Lagrangian is calculated. Via the Universal FeynRules Output (UFO) [114]  
 782 the results are then passed to matrix element generators.



**Figure 3.3:** Sketch of a hadron collision as simulated by a Monte-Carlo event generator. The red blob in the centre represents the hard collision, surrounded by a tree-like structure representing Bremsstrahlung as simulated by parton showers. The purple blob indicates a secondary hard scattering event. Parton-to-hadron transitions are represented by light green blobs, dark green blobs indicate hadron decays, while yellow lines signal soft photon radiation. Figure taken from [112].

783     The MadGraph program [115] is used to interpret the physics model and calculate the cor-  
 784     responding Feynman diagrams and matrix elements. After this, MadEvent [116] is used to  
 785     calculate the corresponding partons. These generated parton configurations are then merged  
 786     with Pythia [117–119] parton showers using the MLM merging scheme [120].

787     The MadGraph5\_aMC@NLO program [121] combines the LO MadGraph [115] and the aMC@NLO  
 788     program into a common framework. This combination supports the generation of samples at  
 789     LO or NLO together with a dedicated matching to parton showers using the MC@NLO [122]  
 790     or FxFx [123] schemes respectively. The FxFx scheme produces a certain fraction of events  
 791     with negative weights originating from the subtraction of amplitudes that contain additional  
 792     emissions from the NLO matrix element to prevent double-counting.

793     The POWHEG box (versions 1 and 2) [124–129] contains predefined implementations of various  
 794     processes at NLO. It applies the POWHEG method for ME- to PS- matching, where the hardest  
 795     radiation generated from the ME has priority over subsequent PS emission to remove the overlap  
 796     with the PS simulation.

797     The JHU generator (version 7.02) [130–133] is used to generate the parton level information  
 798     including full spin and polarization correlations. It is commonly used for studying the spin and  
 799     parity properties of new resonances such as  $ab \rightarrow X \rightarrow VV$ , where  $V = Z, W, \gamma$ .

800     The generation of events from processes involving the production and decay of resonances  
 801 creates a computational heavy load, especially at NLO. The narrow width approximation assumes  
 802 that the resonant particle is on-shell. This factorizes the production and decay amplitude,  
 803 allowing to perform the simulation of the production and decay of heavy resonances like top  
 804 quarks or Higgs bosons to be performed in separate steps. The MadSpin program [134] extends  
 805 this approach and accounts for off-shell effects through a partial reweighting of the events.  
 806 Additionally, spin correlation effects between production and decay products are taken into  
 807 account.

808     The Pythia program (versions 6 and 8) [117–119] generates events of various processes  
 809 at LO. However, more commonly it is only used for its PS simulation and is then used after  
 810 other LO and NLO event generators to perform subsequent parton showering, hadronisation,  
 811 and simulation of the underlying event. In this thesis the underlying event tunes [135] are the  
 812 CUETP8M2T4, CUETP8M1 and CUETP8M2.

813     The detector response is simulated via the Geant4 [109] program. This program tracks the  
 814 particles through the detector material via a detailed description of the detector and generates  
 815 several hits throughout several sensitive layers. In addition, the response of the detector  
 816 electronics to these hits are simulated.

### 817 3.2.3 Generating FCNC top-Z interactions

818     The FCNC processes are generated by interfacing the Lagrangian in Equation 1.36 with  
 819 MadGraph5\_aMC@NLO by means of the FeynRules package and its Universal FeynRules  
 NOTE: RH<sup>820</sup>  
 and LH<sup>821</sup>  
 gave the<sup>822</sup>  
 same re-<sup>823</sup>  
 sulting vari-<sup>824</sup>  
 ables and<sup>825</sup>  
 RH is easier<sup>826</sup>  
 to simulate<sup>827</sup>  
 since those<sup>828</sup>  
 are singlets<sup>829</sup>  
 under SU<sup>827</sup>  
 (no doublet<sup>828</sup>  
 with b)<sup>829</sup>  
 output format. The complex chiral parameters are arbitrary chosen to be  $f_{Xq}^L = 0$  and  $f_{Xq}^R = 1$ .  
 The processes are generated with the MadGraph5\_aMC@NLO (version 2.2.2) and showered with  
 Pythia (version 8.22). The signal consists of two components: events describing the top quark  
 pair production followed by an FCNC decay of one top quark ( $t \rightarrow Zq$ ), and events with the FCNC  
 single top quark production ( $Zq \rightarrow t$ ) for which the top quark decays according to theSM. The  
 leading order generation of the single top quark FCNC process  $tZ + 0,1$  jet including a merging  
 technique can not be done since  $tZ + 1$  jet also contains contributions from top quark pair FCNC  
 where one quark is decaying in  $tZ$ . Therefore, single top quark and top quark pair processes  
 are generated independently, where the single top quark process is generated without the extra  
 hard jet, and the top quark pair FCNC process is generated with up to two extra jets.

830     The signal rates are estimated by use of the MadGraph5\_aMC@NLO program for estimating the  
 831 partial widths. The anomalous couplings are left free to float for this estimation, and only one  
 832 coupling is allowed to be non-vanishing at a time. The results are presented in Table 3.1.

833     The anomalous single top quark cross sections are calculated by convolution of the hard  
 834 scattering matrix elements with the LO order set of NN23L01 [136] partons densities. The NLO  
 835 effects are modelled by multiplying each LO cross section by a global  $k$ -factor. The LO single top  
 836 quark production cross section and the global  $k$ -factors for the top-Z production are shown in  
 837 Table 3.2. The hard scattering events are then matched to parton showers to Pythia to account  
 838 for the simulation of the QCD environment relevant for hadronic collisions.

**Table 3.1:** Leading order partial widths related to the anomalous decay modes of the top quark, where the new physics scale  $\Lambda$  is given in GeV.

Anomalous coupling	vertex	Partial decay width (GeV)
$\kappa_{tZq}/\Lambda$	tZu	$1.64 \times 10^4 \times (\kappa_{tZu}/\Lambda)^2$
	tZc	$1.64 \times 10^4 \times (\kappa_{tZc}/\Lambda)^2$
$\zeta_{tZq}$	tZu	$1.69 \times 10^{-1} \times (\zeta_{tZu})^2$
	tZc	$1.68 \times 10^{-1} \times (\zeta_{tZc})^2$

**Table 3.2:** Leading order single top quark production cross section at a centre-of-mass of 13 TeV for  $pp \rightarrow tZ$  or  $t\bar{Z}$ , where the new physics scale is given in GeV. The NLO  $k$ -factors [137] are given in the last column.

Anomalous coupling	vertex	Cross section (pb) $pp \rightarrow t + pp \rightarrow \bar{t}$	$\sigma_{pp \rightarrow \bar{t}}/\sigma_{pp \rightarrow t}$	NLO $k$ -factor
$\kappa_{tZq}/\Lambda$	tZu	$1.92 \times 10^7 \times (\kappa_{tZu}/\Lambda)^2$	0.12	1.40
	tZc	$2.65 \times 10^6 \times (\kappa_{tZc}/\Lambda)^2$	0.50	1.40
$\zeta_{tZq}$	tZu	$8.24 \times 10 \times (\zeta_{tZu})^2$	0.14	1.40
	tZc	$1.29 \times 10 \times (\zeta_{tZc})^2$	0.50	1.40

The top quark pair cross sections are derived from the SM  $t\bar{t}$  cross section, calculated with MadGraph5\_aMC@NLO at NLO at a centre-of-mass of 13 TeV ( $\sigma_{t\bar{t}}^{\text{SM,NLO}} = 6.741 \times 10^2 \text{ pb}$ ), and considering the decay  $t\bar{t} \rightarrow (bW^\pm)(X_{qt})$ . The branching fraction  $\mathcal{B}(t \rightarrow bW^\pm)$  is assumed to be equal to one and the FCNC branching fraction is calculated as

$$\mathcal{B}(t \rightarrow qX) = \frac{\Gamma_{t \rightarrow qX}}{\Gamma_t^{\text{SM}} + \Gamma_t^{\text{FCNC}}} \approx \frac{\Gamma_{t \rightarrow qX}}{\Gamma_t^{\text{SM}}}, \quad (3.5)$$

where  $\Gamma_{t \rightarrow qX}$  is given in Table 3.1,  $\Gamma_t^{\text{SM}} = 1.32 \text{ GeV}$  [47], and the assumption  $\Gamma_t^{\text{FCNC}} \ll \Gamma_t^{\text{SM}}$  is made. In Table 3.3 the resulting NLO cross sections for the top-Z FCNC interactions are given.

**Table 3.3:** Next to leading order top quark pair cross section for the top-Z FCNC interactions  $t\bar{t} \rightarrow (bW^\pm)(X_{qt})$  with a full leptonic decay at a centre-of-mass of 13 TeV, where  $\sigma_{pp \rightarrow t\bar{t}}^{\text{SM,NLO}} = 6.741 \times 10^2 \text{ pb}$ ,  $\mathcal{B}(Z \rightarrow \ell\bar{\ell}) = 3.36 \times 3 \times 10^{-2}$ , and  $\mathcal{B}(W \rightarrow \ell\nu) = 10.80 \times 3 \times 10^{-2}$ .

Anomalous coupling	vertex	Process	Cross section (pb)
$\kappa_{tZq}/\Lambda$	tZu	$t\bar{t} \rightarrow (b\ell^+\nu)(\bar{u}\ell^+\ell^-)$	$2.727 \times 10^5 \times (\kappa_{tZu}/\Lambda)^2$
		$t\bar{t} \rightarrow (\bar{b}\ell^-\bar{\nu})(u\ell^+\ell^-)$	$2.727 \times 10^5 \times (\kappa_{tZu}/\Lambda)^2$
	tZc	$t\bar{t} \rightarrow (b\ell^+\nu)(\bar{c}\ell^+\ell^-)$	$2.726 \times 10^5 \times (\kappa_{tZc}/\Lambda)^2$
		$t\bar{t} \rightarrow (\bar{b}\ell^-\bar{\nu})(c\ell^+\ell^-)$	$2.726 \times 10^5 \times (\kappa_{tZc}/\Lambda)^2$
$\zeta_{tZq}$	tZu	$t\bar{t} \rightarrow (b\ell^+\nu)(\bar{u}\ell^+\ell^-)$	$2.807 \times (\zeta_{tZu})^2$
		$t\bar{t} \rightarrow (\bar{b}\ell^-\bar{\nu})(u\ell^+\ell^-)$	$2.807 \times (\zeta_{tZu})^2$
	tZc	$t\bar{t} \rightarrow (b\ell^+\nu)(\bar{c}\ell^+\ell^-)$	$2.807 \times (\zeta_{tZc})^2$
		$t\bar{t} \rightarrow (\bar{b}\ell^-\bar{\nu})(c\ell^+\ell^-)$	$2.807 \times (\zeta_{tZc})^2$

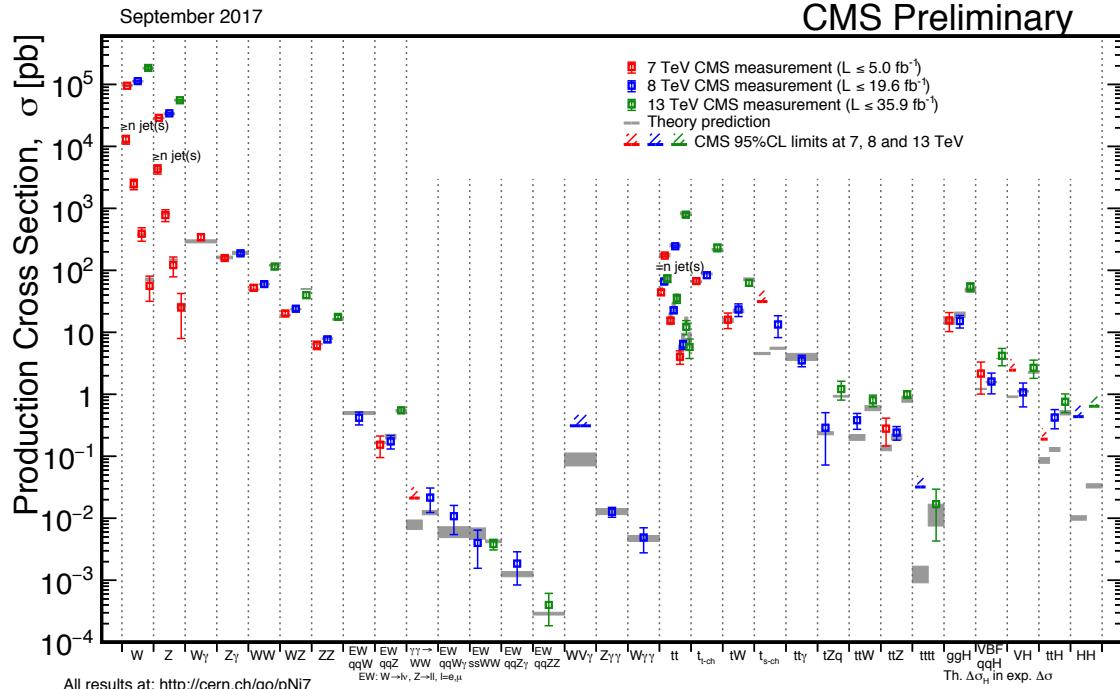
### 3.2.4 Generating SM background events

The SM tZq sample is generated using the MadGraph5\_aMC@NLO generator (version 2.2.2) [138] at leading order accuracy. The t $\bar{t}Z$  and triboson samples were generated using the MadGraph5\_aMC@NLO generator (version 2.2.2), interfaced through the dedicated MC@NLO matching scheme [122]. The WZ+jets and t $\bar{t}W$  samples are produced with up to one additional parton at next-to-leading order accuracy using MadGraph5\_aMC@NLO (version 2.2.2) and using FxFx approach [139] for matching and merging. The samples of t $\bar{t}H$ , WW, ZZ, and single top quark production channels are generated with the POWHEG box (versions 1 and 2) [124–129]. The JHU generator [130–133] is used for the tqH sample, while the tWZ sample is generated using MadGraph5\_aMC@NLO at leading order [115]. All events are interfaced to Pythia version 8.22 [119] to simulate parton shower, hadronisation, and underlying event. Additionally, MadSpin is used for the tZq, WZ+jets, t $\bar{t}Z$ , t $\bar{t}W$ , tWZ, and triboson samples.

The complete list of SM samples is given in Table 3.4, along with their cross sections at a centre-of-mass of 13 TeV. The cross sections without a reference are coming from the generator with which the sample has been made, for some of them the uncertainties are provided by the Generator Group [140]. For each MC sample, the integrated luminosity that the sample represents is estimated as the number of simulated events divided by the cross section of the generated process. This luminosity is then matched to integrated luminosity of  $35.9 \text{ fb}^{-1}$  represented by the data used for analysis. For processes generated with MadGraph5\_aMC@NLO, the effective number of simulated events is used, taking into account positive and negative event weights. In Figure 3.4, a summary is given of the SM cross section measurements performed by the CMS collaboration. These cross sections are all in agreement with their SM predictions.

**Table 3.4:** SM MC samples used in this analysis with their corresponding cross section at a centre-of-mass of 13 TeV and MadGraph5\_aMC@NLO correction C when applicable. The generators used for each sample are indicated and the simulation of the parton shower, hadronisation, and underlying event is done by Pythia version 8.22 [119] for all samples.

Process	Generator	Cross section (pb)	C	Ref.
$WZ \rightarrow 3\ell\nu$	MadGraph5_aMC@NLO+MadSpin	5.26	1.61	[140]
$tZq$ with $Z \rightarrow \ell^+\ell^-$	MadGraph5_aMC@NLO+MadSpin	0.0758	3.77	[140]
$tqH$ with $H \rightarrow ZZ \rightarrow \ell^+\ell^-\ell^+\ell^-$	JHU	$8.80 \cdot 10^{-6}$	-	[140]
$t\bar{t}W + \text{jets}$ with $W \rightarrow \ell\nu$	MadGraph5_aMC@NLO+MadSpin	$0.2043 \pm 0.0020$	1.94	[140]
$t\bar{t}Z \rightarrow 2\ell + 2\nu + \text{other}$ , with $m_{\ell\ell} > 10$ GeV	MadGraph5_aMC@NLO+MadSpin	$0.2529 \pm 0.0004$	2.15	[140]
$t\bar{t}H, \text{no } b\bar{b} \text{ decays}$	POWHEG	0.2151	-	[140]
$t\bar{t}H, b\bar{b} \text{ decays}$	POWHEG	0.2934	-	[140]
$WW \rightarrow 2\ell 2\nu$	POWHEG	12.178	-	[141]
$ZZ \rightarrow 4\ell$	POWHEG	0.3366	-	[140]
$WZZ$	MadGraph5_aMC@NLO+ MadSpin	0.05565	1.14	[140]
$ZZZ$	MadGraph5_aMC@NLO	0.01398	1.17	[140]
single top quark $tWZ$ , with $Z_\mu \rightarrow \ell^+\ell^-$	MadGraph5_aMC@NLO(LO)+MadSpin	0.001123	-	[140]
single top quark t-channel $\bar{t}$	POWHEG+MadSpin	$44.33^{+1.76}_{-1.49}$	-	[140]
single top quark t-channel $t$	POWHEG+MadSpin	$26.38^{+1.32}_{-1.18}$	-	[140]
single top quark $t\bar{W}$	POWHEG	$35.85 \pm 0.90 \text{ (scale)} \pm 1.70 \text{ (PDF)}$	-	[140]
single top quark $tW$	POWHEG	$35.85 \pm 0.90 \text{ (scale)} \pm 1.70 \text{ (PDF)}$	-	[140]
$t\bar{t}$	POWHEG	$831.76^{+19.77+35.06}_{-29.20-35.06}$	-	[140]
$Z/\gamma^* + \text{jets}$ , with $m_{\ell\ell} > 50$ GeV	MadGraph5_aMC@NLO	$3 \times (1921.8 \pm 0.6 \pm 33.2)$	1.49	[140]
$Z/\gamma^* + \text{jets}$ , with $10 \text{ GeV} < m_{\ell\ell} < 50 \text{ GeV}$	MadGraph	18610	-	[140]



**Figure 3.4:** Summary of the SM cross section measurements performed by the CMS collaboration. Figure taken from [23].

### 863 3.3 Multivariate analysis techniques: Boosted Decision Trees

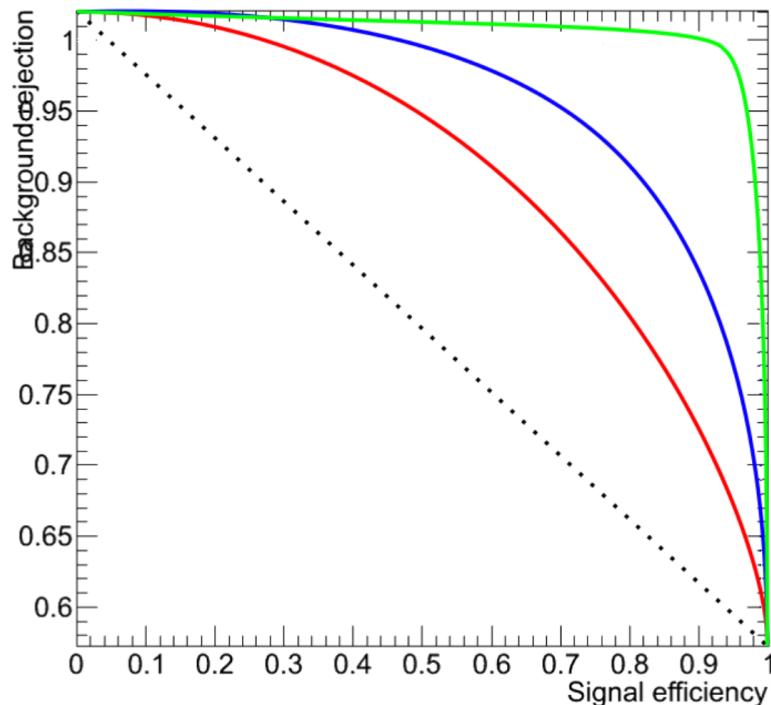
864 The need for processing large quantities of data and discriminating between events with largely  
 865 similar experimental signatures makes multivariate analysis (MVA) a largely used method in the  
 866 physics community. Multivariate classification methods based on machine learning techniques  
 867 are a fundamental ingredient to most analyses. The advantage of using a MVA classifier is  
 868 that it can achieve a better discrimination power with respect to a simpler analysis based on  
 869 individual selection criteria or poorly discriminating variables. A risk of using MVA classifiers  
 870 is overtraining. This happens when there are too many model parameters of an algorithm  
 871 adjusted to too few data points. This leads to an increase in the classification performance over  
 872 the objectively achievable one.

873 There are many software tools that exist for MVA. In this thesis, the Tool for Multivariate  
 874 Analysis (TMVA) [142] is used. This software is an open source project included into  
 875 ROOT [143]. By training on events for which the classification is known, a mapping function  
 876 is determined that describes a classification or an approximation of the underlying behaviour  
 877 defining the target value (regression). Boosted decision trees (BDT) are employed for the clas-  
 878 sification of events as implemented in the TMVA framework [142]. This multivariate technique  
 879 is based on a set of decision trees where each tree yields a binary output depending on the fact  
 880 that an event is signal- or background-like. This has as advantage that several discriminating  
 881 variables can be combined into a powerful one-dimensional discriminant D.

882 The decision tree is constructed by training on a dataset for which the outcome is already

883 provided, such as simulation datasets with signal and background processes (supervised learning). Different trees can be combined into a forest where the final output is determined by  
 884 the majority vote of all trees, so-called boosting. This stabilises the decision trees against  
 885 statistical fluctuations and makes it possible to keep the decision trees very shallow, making the  
 886 method more robust against overtraining. Examples of such boosting algorithms are Adaptive  
 887 Boosting (AdaBoost) and Gradient Boosting [144]. For the search presented in the following  
 888 chapters, Gradient boost is used with a learning rate of 0.2-0.3 and the depth of the tree is  
 889 set to three. Additionally, the Gradient boost is used in combination with bagging, so-called  
 890 stochastic gradient boosting. Bagging smears the statistical fluctuations in the training data and  
 891 therefore stabilises the response of the classifier and increases the performance by eliminating  
 892 overtraining. More information about stochastic gradient boosting can be found in Ref. [145].

894 The discriminating power of a BDT is assessed by analysing the receiver operating characteristic  
 895 (ROC) curve. This curve represents the background rejection over the signal efficiency of the  
 896 remaining sample. The area under the curve (AUC) is compared to random guessing in order  
 897 to identify the best classifier. When the multivariate discriminator has no discriminating power,  
 898 the resulting AUC will be 0%, while 50% means fully separated event classes. In Figure 3.5  
 examples of ROC curves are shown.



**Figure 3.5:** Example of ROC curves. In this example, the green method is better than the blue one, which is better than the red one. The dashed line represents a case where there is no separation. Figure taken from [146].

### 3.4 Statistical methodology

The search performed in the framework of this thesis requires the simultaneous analysis of data from different decay channels. The statistical methodology used for this search is developed by the ATLAS and CMS collaborations in the context of the LHC Higgs Combination group [147–150]. The Higgs Combined Tool [151] is a RooStats [152] framework which runs different statistical methods. In this section, only the statistical tools necessary for the performed search are described [153].

The event yields of signal and background processes are denoted as  $s$  and  $b$  respectively. These represent event counts in multiple bins or unbinned probability density functions. By use of simulation, predictions on both signal and background yields are made. The multiple uncertainties on these predictions are accounted for by introducing nuisance parameters  $\theta$  such that  $s = s(\theta)$  and  $b = b(\theta)$ .

The Bayesian and modified classical frequentist statistical approaches are used in high energy physics to characterise the absence of a signal. The level of incompatibility of data with a signal hypothesis is quantified in terms of confidence levels (CL). The convention is to require a 95% CL for excluding a signal. In general, limits are not set on the signal cross section directly, but are set on the signal strength modifier  $\mu$ . The signal strength modifier is defined such that it equally changes all the cross sections of all production mechanisms of the signal by the same scale.

In this thesis, the modified frequentist approach [154, 155] for confidence levels that adopts the classical frequentist method to allow nuisance parameters, is used. It constructs a likelihood  $\mathcal{L}(\text{data}|\mu, \theta)$  is as

$$\mathcal{L}(\text{data}|\mu, \theta) = \text{Poisson}(\text{data}|\mu s(\theta) + b(\theta)) \text{pdf}(\tilde{\theta}|\theta). \quad (3.6)$$

The probability density function  $\text{pdf}(\tilde{\theta}|\theta)$  describes all sources of uncertainty. In this thesis, all sources of uncertainties are assumed to be either 100% correlated or uncorrelated. Partially correlated uncertainties are broken down to subcomponents that fit those requirements, allowing to include all constraints in the likelihoods in a clean factorised form.

A systematic uncertainty pdf  $\rho(\theta|\tilde{\theta})$  for the nuisance  $\theta$  with nominal value  $\tilde{\theta}$  is used. It reflects the degree of belief of what the true value of the  $\theta$  is. In this thesis, the approach from the Higgs Combined Tool is used where the pdfs  $\rho(\theta|\tilde{\theta})$  are re-interpreted as posteriors of real or imaginary measurements  $\tilde{\theta}$

$$\rho(\theta|\tilde{\theta}) \sim \text{pdf}(\tilde{\theta}|\theta) \pi_\theta(\theta), \quad (3.7)$$

where  $\pi_\theta(\theta)$  is the hyper prior for the (imaginary) measurements. The pdfs used by the Higgs Combine Tool are described in Ref. [150].

The data in Equation 3.6 represents either the actual observation or pseudo-data to construct sampling distributions. For a binned likelihood, the Poisson probabilities to observe  $n_i$  events in bin  $i$  is given as

$$\text{Poisson}(\text{data}|\mu s(\theta) + b(\theta)) = \prod_i \frac{(\mu s_i(\theta) + b_i(\theta))^{n_i}}{n_i!} e^{-\mu s_i(\theta) - b_i(\theta)}. \quad (3.8)$$

At the LHC, the test statistic is defined as

$$q_\mu = -2 \ln \frac{\mathcal{L}(\text{data} | \mu, \hat{\theta}_\mu)}{\mathcal{L}(\text{data} | \hat{\mu}, \hat{\theta}_\mu)}, \quad (3.9)$$

where the likelihood is maximised in the numerator (maximum likelihood estimator, MLE) for a given  $\mu$  and (pseudo) data at  $\hat{\theta}_\mu$ , while  $\hat{\mu}$  combined with  $\hat{\theta}$  defines the point for which the likelihood reaches its global maximum. The estimated signal strength modifier  $\hat{\mu}$  can not become negative since a signal rate is positive defined by physics. Furthermore, an upper constraint on the MLE  $\hat{\mu} \leq \mu$  is imposed.

The signal is excluded at  $1 - \alpha$  confidence level when

$$\text{CLs} = \frac{P(q_\mu \geq q_\mu^{\text{obs}} | \mu s + b)}{P(q_\mu \geq q_\mu^{\text{obs}} | b)} \leq \alpha, \quad (3.10)$$

with  $P(q_\mu \geq q_\mu^{\text{obs}} | \mu s + b)$  the probability to observe a value of the test statistic at least as large as the one observed in data  $q_\mu^{\text{obs}}$ , under the signal plus background ( $s + b$ ) hypothesis, and  $P(q_\mu \geq q_\mu^{\text{obs}} | b)$  for the background only ( $b$ ) hypothesis. These probabilities are defined as

$$p_\mu = P(q_\mu \geq q_\mu^{\text{obs}} | \mu s + b) = \int_{q_\mu^{\text{obs}}}^{\infty} f(q_\mu | \mu, \hat{\theta}_\mu^{\text{obs}}) dq_\mu, \\ 1 - p_b = P(q_\mu \geq q_\mu^{\text{obs}} | b) = \int_{q_{\mu=0}^{\text{obs}}}^{\infty} f(q_\mu | \mu = 0, \hat{\theta}_{\mu=0}^{\text{obs}}) dq_\mu, \quad (3.11)$$

where  $p_\mu$  and  $p_b$  are the p-values associated to the two hypothesis, and  $f(q_\mu | \mu, \hat{\theta}_\mu^{\text{obs}})$  and  $f(q_\mu | \mu = 0, \hat{\theta}_{\mu=0}^{\text{obs}})$  are the probability density functions of the signal plus background and background only hypothesis constructed from toy Monte Carlo pseudo data. These are generated with nuisance parameters fixed to  $\hat{\theta}_{\mu=0}^{\text{obs}}$  (background only) and  $\hat{\theta}_\mu^{\text{obs}}$  (signal plus background). The 95% CL level upper limit on  $\mu$  is achieved by adjusting  $\mu$  until  $\text{CL} = 0.05$ , this is the so-called observe limit. The expected median upper limit and the  $\pm 1\sigma$  and  $\pm 2\sigma$  bands for a hypothesis is generated by a large set of pseudo data and calculate the CLs and the value of  $\mu$  at 95% CL for each of them. A cumulative probability distribution can be build by starting the integration from the side corresponding to low event yields. The median expected value, so-called expected limit at 95% CL, is where the cumulative distribution function crosses the 50% quantile. The  $\pm 1\sigma$  (68%) and  $\pm 2\sigma$  (95%) bands on the expected limit are defined by the crossings of the 16% and 84%, and 2.5% and 97.5% quantiles.

In order to significantly reduce computing time, the Asymptotic CL method is used. This method avoids an ensemble of toy Monte Carlo samples and instead replaces it by one representative dataset, called Asimov dataset. This dataset is constructed such that all observed quantities are set equal to their MLE values ( $\hat{\theta}_{\text{Asimov}} = \theta_0$ ). More information about this procedure can be found in Refs. [148].



# 4

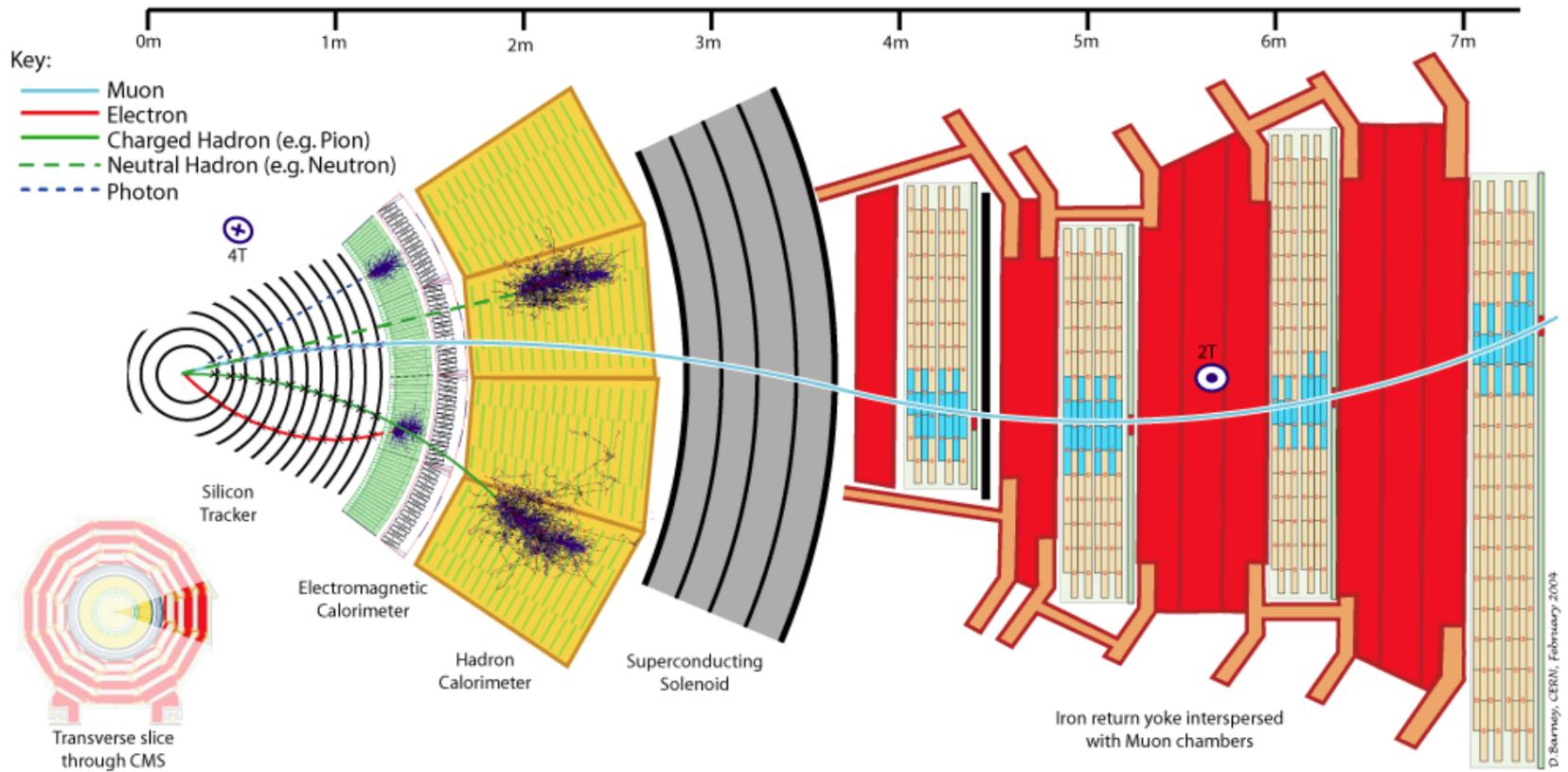
## Event reconstruction and identification

947

948 The simulated data after the detector simulation described in [Section 3.2](#), has the exact same  
949 format as the real collision data recorded at the CMS experiment. Therefore the same software  
950 can be used for the reconstruction of both simulation and real data. In [Section 4.1](#), the object  
951 reconstruction is explained. After reconstructing the objects, they are connected to physics  
952 objects, which need to be identified ([Section 4.2](#)) and corrected for pileup ([Section 4.3](#)). The  
953 objects used for physics analysis have extra requirements as shown in [Section 4.4](#). A summary  
954 of all the corrections applied to data and simulation is given in [Section 4.5](#).

### 955 4.1 Object Reconstruction

956 In [Figure 4.1](#), the particle interaction in a transverse slice of the CMS detector is shown. When  
957 a particle enters the detector, it first enters the tracker where charged particle trajectories,  
958 so-called tracks, and origins, so-called vertices, are reconstructed from signals or hits in the  
959 sensitive layers. The magnetic field bends the charged particles making it able to measure the  
960 electric charges and momenta of charged particles. The electrons and photons are absorbed in  
961 the ECAL and the corresponding electromagnetic showers are detected as clusters of energy  
962 in adjacent cells. From this, the energy and the direction of the particles can be determined.  
963 The charged and neutral hadrons can also initiate a hadronic shower in the ECAL that is fully  
964 absorbed in the HCAL. The clusters from these showers are also used to estimate the energy  
965 and direction. Muons and neutrinos pass through the calorimeters without little to no energy  
966 loss and the neutrinos even escape the CMS detector undetected while muons produce hits in  
967 the muon detectors.



**Figure 4.1:** Cross-section of the CMS detector with all parts of the detector labelled. This sketch shows the specific particle interactions from a beam interaction region to the muon detector. The muon and charged pion are positively charged, the electron is negatively charged. Figure taken from [156].

968 The particle flow (PF) [156] reconstruction algorithm correlates the tracks and clusters  
 969 from all detector layers with the identification of each final state particle, and combines the  
 970 corresponding measurements to reconstruct the properties. The muon is identified by a track  
 971 in the inner tracker, connected to a track in the muon detector as described in Section 4.1.2.  
 972 The electrons are identified by a track and an ECAL cluster, not connected to an HCAL cluster  
 973 as described in Section 4.1.3. The ECAL and HCAL clusters without a track link identify the  
 974 photons and neutral hadrons, while the addition of the tracker determines the energy and  
 975 direction of a charged hadron.

#### 976 4.1.1 Charged particle tracks

977 An iterative tracking algorithm is responsible for the reconstruction of the tracks made by  
 978 charged particles in the inner tracking system. Each iteration consists of four steps [72]: the  
 979 track-seed generation, the pattern recognition algorithm, removal of track-hit ambiguities and  
 980 a final track fit. The pattern recognitions done by use the Kalman filter method [157, 158]  
 981 which takes into account the magnetic field and multiple scattering effects. All hits that are  
 982 unambiguously associated to the final track are removed from the list of available hits. In order to  
 983 associate the remaining hits, the procedure is repeated with looser track reconstruction criteria.  
 984 The use of the iterative track reconstruction procedure has a high track finding efficiency, where  
 985 the fake track reconstruction rate is negligible.

**NOTE:** Ik kan hier stoppen en 4.1.1, 4.1.2, 4.1.3.4.1.4 volledig schrappen (dus enkel primary vertex houden)

#### 986 4.1.2 Following the Muon's Footsteps

987 The muon reconstruction [159] has three subdivisions: local reconstruction, regional reconstruc-  
 988 tion and global reconstruction. The local reconstruction is performed on individual detector  
 989 elements such as strip and pixel hits in the inner tracking system, and muon hits and/or segments  
 990 in the muon chambers. Independent tracks are reconstructed in the inner tracker - called tracker  
 991 tracks - and in the muon system, called standalone muon tracks. Based on these tracks, two  
 992 reconstructions are considered: Global Muon reconstruction and Tracker Muon reconstruction.  
 993 The first is an outside-in approach starting from a standalone muon track while the second uses  
 994 an inside-out approach starting from tracker tracks. For low transverse momenta ( $p_T \lesssim 5$  GeV),  
 995 the tracker muon reconstruction is more efficient than the global muon approach. This is due  
 996 to the fact that tracker muons only require a single muon segment in muon system, while the  
 997 global muon approach requires typically segments in at least two muon stations. These tracker  
 998 muons are used for identifying muons from the hadronisation of b or c quarks. The global muon  
 999 approach typically improves the tracker reconstruction for  $p_T \gtrsim 200$  GeV.

#### 1000 4.1.3 The path of the Electron

1001 Standard tracking algorithms are based on Kalman filtering which assume that the energy loss  
 1002 is Gaussian distributed. Since the electron tracks are increasingly curved in the magnetic field  
 1003 as a function of its flight distance, these standard tracking algorithms are not suitable to fit the  
 1004 electron tracks and different filtering algorithm. The Gaussian sum filter (GSF) [160] is used  
 1005 instead.

1006 In CMS, the electrons are reconstructed in two ways. The older ECAL based tracking is  
 1007 developed to identify high energetic isolated electrons. This tracking algorithm starts from

1008 ECAL clusters with a transverse energy above 4 GeV and extrapolates from these cluster the  
 1009 position of the hits in the tracker. Another, tracker based algorithm uses all the tracks with  
 1010 a  $p_T$  higher than 2 GeV found with iterative tracking as seeds. The electron seeds from the  
 1011 ECAL- and tracker-based procedures are merged into a unique collection and are then refitted  
 1012 by using the summed Gaussian distributions as uncertainty per hit in the track fit. The electron  
 1013 efficiency is measured in 8 TeV proton collision data to be better than 93% for electrons with an  
 1014 ECAL supercluster energy of  $E_T > 20$  GeV [161]. For electrons with an  $E_T > 25$  GeV in 13 TeV  
 1015 proton collision data, the efficiency is about 96%[162].

#### 1016 4.1.4 Primary Vertex Reconstruction

1017 The primary vertex (PV) reconstruction is able to measure the location of all proton interaction  
 1018 vertices in each event consisting of the signal vertex and all vertices from pileup events. First,  
 1019 tracks are selected to be consistent with being produced promptly in the primary interaction [82].  
 1020 Then the tracks are grouped according to the  $z$  coordinate of their closest approach to the beam  
 1021 line [163] and a vertex fitting algorithm [164] is performed. The primary vertex is found as the  
 1022 vertex corresponding to the highest sum of squared track transverse momenta and is taken to  
 1023 be the main interaction point. The resolution on the primary vertex is about 14  $\mu\text{m}$  in  $r\phi$  and  
 1024 about 19  $\mu\text{m}$  in the  $z$  direction for primary vertices with the sum of the track  $p_T > 100$  GeV  
 1025 for 2016 data taking.

#### 1026 4.1.5 Calorimeter clusters

1027 The energy and direction of stable neutral particles such as photons and neutral hadron are recon-  
 1028 structed using a cluster algorithm. This algorithm also separates neutral particles from charged  
 1029 hadron energy deposits, and reconstructs and identifies electrons and their bremsstrahlung  
 1030 photons. Furthermore, the cluster algorithm is contributing to the energy measurements of  
 1031 charged hadrons that don't have accurate tracks parameters, e.g. for low quality and high  
 1032 transverse momentum tracks. The clustering is performed separately in each subdetector:  
 1033 ECAL barrel and endcaps, HCAL barrel and endcaps, and the two preshower layers. The HF has  
 1034 no clustering algorithm since the electromagnetic or hadronic components give rise to an HF  
 1035 EM or HF HAD cluster.

1036 The clustering algorithm consist of different steps. First seeds are identified when cells have  
 1037 an energy larger than the seeding threshold and larger than their neighbouring cells. Then  
 1038 topological clusters are made by accumulating cells that share at least a corner with a cell  
 1039 already in the cluster and an energy above a cell threshold set to twice the noise level. The  
 1040 third step is an expectation maximization algorithm that reconstructs the cluster [156] and  
 1041 assumes that the energy deposits are Gaussian distributed. The calorimeter clusters are used  
 1042 for reconstructing photons and neutral hadrons. The clusters that are not in the vicinity of the  
 1043 extrapolated charged tracks are identified as neutral hadrons or photons. If the energy deposits  
 1044 are in vicinity of charged tracks, such is the case for charged hadrons, the neutral particle energy  
 1045 deposit is measured as an excess over the charged particle deposit.

1046 **4.2 Particle flow identification**

1047 The several PF elements from the various CMS subdetectors are connected through a link  
1048 algorithm. This algorithm tests any pair of elements in an event, only considering nearest  
1049 neighbours in the  $\eta\phi$ -plane. The quality of the link is determined via the distance between the  
1050 two elements and PF blocks of elements are formed from elements with a direct link or indirect  
1051 link through common elements. The identification and reconstruction follows a particular order  
1052 in each PF block. After each identification and reconstruction the corresponding PF elements  
1053 (tracks and clusters) are removed from the PF block.

1054 The muons are the first to be identified and reconstructed. These are reconstructed if their  
1055 momenta are compatible with corresponding track only momenta. Then the electron and its  
1056 corresponding brehmstrahung photons, are identified and reconstructed by using of the GSF  
1057 tracking. At the same time, the energetic and isolated photons are identified as well. The  
1058 remaining elements in the PF block are subjected to a cross identification of charged hadrons,  
1059 neutral hadrons, and photons that arise from parton fragmentation, hadronisation, and decays  
1060 in jets. The charged hadron candidate is made from the remaining candidates that have a  
1061 charged particle track associated with them. Then the charged particle energy fraction is  
1062 subtracted from the calibrated energy of the linked calorimeter clusters and the remaining  
1063 energy is assigned to the neutral energy. Depending on the excess of neutral energy in the ECAL  
1064 and HCAL clusters, a photon or a neutral hadron is assigned respectively. The pseudorapidity  
1065 range of the inner tracker limits the information on the particles charge to  $|\eta| < 2.4$ . Outside  
1066 this range a simplified identification is done for hadronic and electromagnetic candidates only.

1067 **4.3 Pileup mitigation and luminosity measurement**

For the 8 TeV dataset, an average of about 21 pileup interactions happen per bunch cross section. For the dataset taken at 13 TeV, the number of pileup interactions increases to about 27 interactions per bunch crossing. These interactions are spread around the beam axis around the centre of the CMS coordinate system and follow a normal distribution with a standard deviation of about 5 cm [156]. The number of pileup interactions is estimated from the number of interaction vertices reconstructed from charged particle tracks, or from the instantaneous luminosity of the given bunch crossing with dedicated detectors and the inelastic proton-proton crossing. The luminosity of the CMS interaction point is estimated from measuring certain process rates with luminometers such as the pixel detector, HF calorimeter, and the pixel luminosity telescope [165]. The instantaneous luminosity from recorded process rate  $R$  is then determined as

$$Ldt = \frac{Rdt}{\sigma_{fid}}, \quad (4.1)$$

1068 where  $\sigma_{fid} = \sigma \times A$  corresponds to the fiducial cross section recorded in the luminometer  
1069 acceptance  $A$  which is determined using van der Meer scans [166]. The overall uncertainty on  
1070 the luminosity measurement is estimated to be 2.5%.

1071 The luminosity is used to infer the number of pileup interactions in data, which can be used  
1072 to correct the predefined pileup interactions in simulation. Then an event weight can be derived

1073 from the ratio of the distributions of pileup interactions in data and simulation. For 13 TeV  
 1074 collisions, the inelastic cross section is measured to be  $71.3 \pm 3.5$  mb [167]. However a better  
 1075 agreement in data and simulation for the pileup sensitive variables, such as the number of  
 1076 primary vertices, is found with a lower cross section of 69.2 mb with an uncertainty of 4.6%.

## 1077 4.4 Physics object reconstruction and identification

1078 The particle flow objects are used for building physics objects that are used for analysis. Analyses  
 1079 use jets, muons, electrons, photons, taus and missing transverse momentum  $\vec{p}_T$  with extra,  
 1080 analysis dependent requirements. In the following section, only the physics objects used  
 1081 throughout this thesis are discussed.

### 1082 4.4.1 Muons

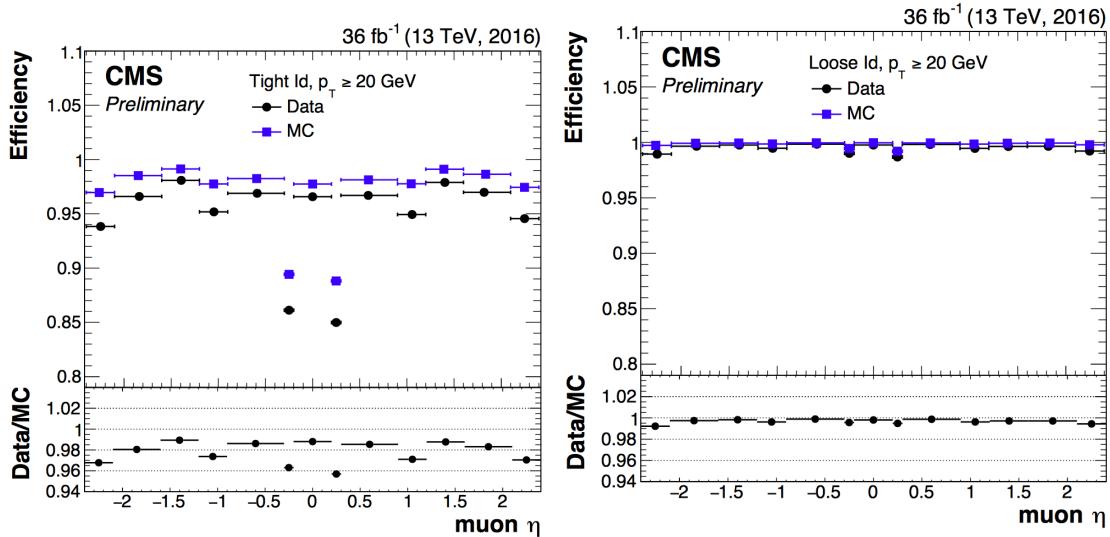
1083 The muon candidates used for analysis in this thesis correspond to the tight and loose working  
 1084 point. Detailed reports on the performance can be found in [168].

1085 The tight working point rejects objects wrongly reconstructed as muons from hadron showers  
 1086 that reach the muon system (punch-throughs), by requiring that the global muon fit includes  
 1087 at least one valid hit in the muon chambers for which at least two muon segments in two  
 1088 muon stations are present. Furthermore, the muon tracks should have a global fit yielding a  
 1089 goodness-of-fit of  $\chi^2/\text{ndof} < 10$ . Requiring at least one pixel hit in the muon track suppresses  
 1090 the decay of muons in flight. Also a minimum of five hits in the tracker is required. Cosmic  
 1091 muons and muons originating from pileup interactions are rejected by constricting the distance  
 1092 of the muon with respect to the primary vertex by putting limits on  $d_{x,y} < 2$  mm and  $d_z < 5$   
 1093 mm. Also muons according to the loose muon working point will be used in the thesis. These  
 1094 are either global muons or tracker muons reconstructed from the particle flow muon object. In  
 1095 Table 4.1, the muon requirements for the muons used throughout this thesis are summarised.  
 1096 In Figure 4.2, the muon efficiencies for data and simulation are presented. These efficiencies  
 1097 are estimated from tag-and-probe methods [168]. Overall, the efficiency is about 95-100%,  
 1098 with two drops due to the crack between the wheels of the DT system. The differences between  
 1099 data and simulation are corrected by applying  $p_T$ - and  $\eta$ -dependent scale factors ( $\epsilon_{\text{data}}/\epsilon_{\text{MC}}$ )  
 1100 to simulated events.

In addition to the identification criteria, the muons are required to be spatially isolated from electromagnetic and hadronic activity. The relative lepton isolation is defined as estimating the total transverse energy of the particles emitted around the direction of the lepton by defining a cone of radius  $\Delta R$  in  $\eta\phi$  plane around the lepton direction. Then a summed energy is calculated from the charged hadrons (CH), neutral hadrons (NH), and photons ( $\gamma$ ), excluding the lepton itself. This sum is then corrected to remove the energy coming from pileup interactions. The relative isolation for muons  $\mathcal{I}_{\mu}$  is defined as [156]:

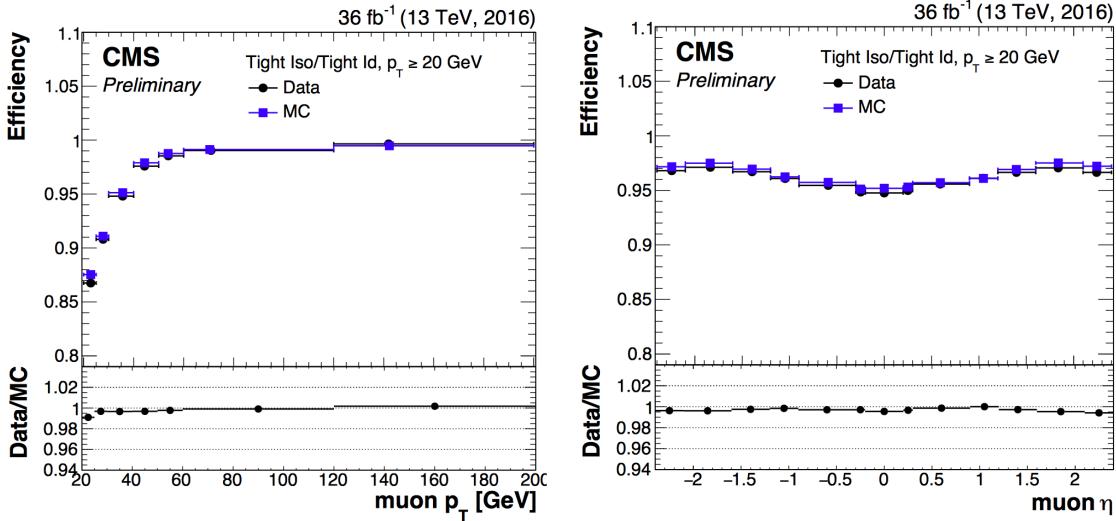
$$\mathcal{I}_{\mu} = \frac{\sum p_T(\text{CH}) + \max(0., \sum E_T(\text{NH}), \sum E_T(\gamma) - 0.5 \times \sum E_T(\text{CH}))}{p_T(\mu)}, \quad (4.2)$$

1101 where a cone of  $\Delta R = 0.4$  is adopted and the pileup mitigation is based on the  $\Delta\beta$  correction.  
 1102 The  $\Delta\beta$  correction estimates the pileup energy as half of the contribution coming from charged



**Figure 4.2:** Comparison of the muon tight ID (left) and loose ID (right) efficiencies in data and simulation as a function of the pseudorapidity of the muon using the full 2016 dataset. Figure taken from [168].

hadrons. For tight ID muons, this relative isolation should  $\mathcal{I}_\mu < 0.15$ , while for loose muons this should be  $\mathcal{I}_\mu < 0.25$ . In Figure 4.3, the isolation efficiencies as a function of the pseudo rapidities using the tag and probe method are shown for the tight muon working point. The efficiencies are 85-100% and have a decline for low- $p_T$  muons. The differences between data and simulation are accounted for by applying  $\eta$ - and  $p_T$ -dependent scale factors on the simulation.



**Figure 4.3:** Comparison of the muon tight isolation requirement with the muon tight ID efficiencies in data and simulation as a function of the transverse momentum (left) or pseudorapidity (right) of the muon using the full 2016 dataset. Figure taken from [168].

**Table 4.1:** Muon requirements for the tight and loose working points, used throughout this thesis.

Properties	Loose Muons	Tight Muons
Global muon or Tracker Muon	One or the other	Both
Particle Flow muon	Y	Y
$\chi^2/\text{ndof}$ of global muon track fit	N/A	< 10
Nb. of hit muon chambers	N/A	> 0
Nb. of muon stations contained in the segment	N/A	> 1
Size of the transverse impact parameter of the track wrt. the PV	N/A	$d_{xy} < 2 \text{ mm}$
Longitudinal distance wrt. the PV	N/A	$d_z < 5 \text{ mm}$
Nb. of pixel hits	N/A	> 0
Nb. of tracker layers with hits	N/A	> 5
Relative Isolation	<0.25	<0.15

#### 1109 4.4.2 Electrons

1110 The electron candidates used, correspond to the tight and veto working points. The study of  
 1111 the electron reconstruction and identification performance can be found in [162].

1112 Starting from an electron PF candidate with a GSF track that is outside the barrel-endcap  
 1113 transition region ( $1.4443 < |\eta| < 1.5660$ ), several requirements are set. The electron track  
 1114 should not have more than one (two or three) missing hit(s) in the innermost layer for the tight  
 1115 (veto) working point. This dismisses electrons from photon conversions. Additionally, a photon  
 1116 conversion veto is applied by testing if a pair of electron tracks is originating from a common  
 1117 displaced vertex. Furthermore, refined cuts are applied on the shower shape variables such as  
 1118 the difference in  $\eta$  or  $\phi$  between the energy weighted supercluster position in the ECAL and  
 1119 the track direction at the innermost tracker position ( $\Delta\eta_{\text{in}}$ ,  $\Delta\phi_{\text{in}}$ ), and the ECAL crystal based  
 1120 shower covariance in the  $\eta$  direction ( $\sigma_{\eta\eta}$ ). These cuts also include energy related variables  
 1121 such as the absolute difference between the inverse electron energy measured in the ECAL  
 1122 and the inverse momentum measured in the tracker ( $|1/E - 1/p|$ ), and the ratio of the energy  
 1123 measured in the HCAL and ECAL (H/E). Unlike the muon case, the identification criteria also  
 1124 contain requirements on the isolation of the electrons.

Similar to the muons, the electron relative isolation is determined from the sum of the particles in a cone around the electron itself. The cone radius used for electrons is  $\Delta R = 0.3$  and a  $\rho$  correction for pileup mitigation is applied. For this correction, the expected pileup energy inside the isolation cone is estimated from the median density energy per area of pileup contamination ( $\rho$ ), computed event by event, and the effective area ( $A_{\text{eff.}}$ ) [156]. This effective area is estimated from simulation and denotes the expected amount of neutral energy from

pileup interactions per  $\rho$  within the isolation cone as a function of the pseudo rapidity of the associated ECAL superclusters. Table 4.2 shows the values used for 13 TeV data. The relative electron isolation  $\mathcal{I}_e$  is calculated as

$$\mathcal{I}_e = \frac{\sum p_T(\text{CH}) + \max(0., \sum E_T(\text{NH}), \sum E_T(\gamma) - \rho \times A_{\text{eff}})}{p_T(e)}. \quad (4.3)$$

**Table 4.2:** The effective areas  $A_{\text{eff}}$  used for the electron relative isolation [169].

$\eta$ region	$A_{\text{eff}}$
$0 <  \eta  < 0.1752$	0.1703
$1.0 <  \eta  < 0.1479$	0.1715
$1.479 <  \eta  < 2.0$	0.1213
$2.0 <  \eta  < 2.2$	0.1230
$2.2 <  \eta  < 2.3$	0.1635
$2.3 <  \eta  < 2.4$	0.1937
$2.4 <  \eta  < 2.5$	0.2393

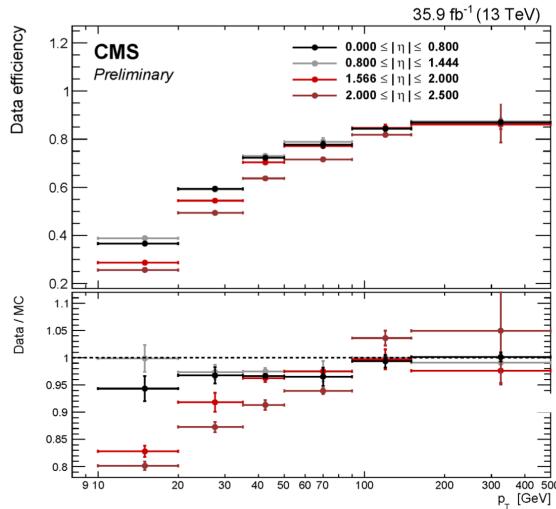
1125

1126 The efficiency of electron identification is estimated from  $Z \rightarrow e^- e^+$  events via the tag-and-  
 1127 probe method and is shown in Figure 4.4 for the tight working point. The efficiencies reach  
 1128 95 – 100%. The difference between data and simulation is corrected for by dedicated  $p_T$ - and  
 1129  $\eta$  dependent scale factors as well.

1130 **4.4.3 Jets**

1131 Jets are reconstructed from all reconstructed particles without the charged hadrons associated to  
 1132 pileup vertices. The clustering is done via the anti –  $k_T$  algorithm [170] with a radius parameter  
 1133 for the cone size of the resulting jet of  $R = 0.4$ . More information about the jet algorithm  
 1134 performance can be found in Ref. [171].

1135 The jets used for the analysis in this thesis, use the loose identification working point sum-  
 1136 marised in Table 4.4. The requirements on the jet constituents are based on the assumption  
 1137 that a proper jet originating from the hadronisation of a quark or gluon consists of multiple PF  
 1138 particles and types. Therefore, the jet should consist of more than one constituent, and the  
 1139 neutral hadron fraction and neutral EM energy fractions should be less than 99%. For the jets  
 1140 within the tracker acceptance ( $|\eta| < 2.4$ ), at least one constituent has to be a charged hadron  
 1141 resulting in a charged hadron energy fraction above 0%. Additionally the charged EM energy  
 1142 fraction should be less than 99%. On top of these requirements, objects that are labelled as jets  
 1143 and found in vicinity of any isolated lepton,  $\Delta R < 0.3$ , are removed from the jet collection in  
 1144 that event to avoid duplications of objects.



**Figure 4.4:** Electron identification efficiency as function of the electron transverse momentum from the full 2016 dataset. Figure taken from [162].

**Table 4.3:** Electron requirements used in this analysis. The requirements are set in the barrel ( $|\eta_{supercluster}| \leq 1.479$ ) and the endcaps ( $|\eta_{supercluster}| > 1.479$ ).

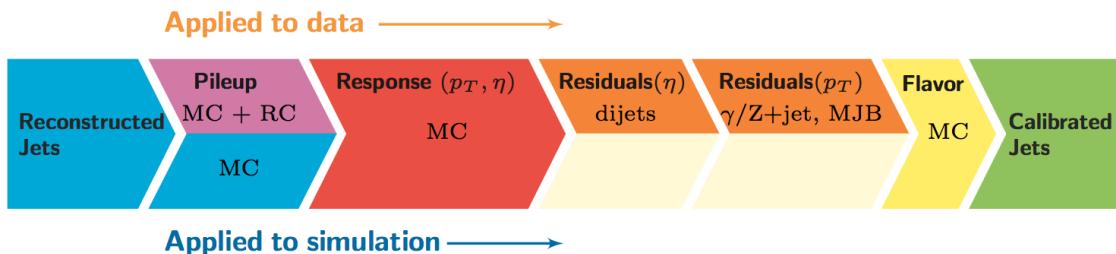
Properties	$ \eta_{supercluster}  \leq 1.479$		$ \eta_{supercluster}  > 1.479$	
	Veto electron	Tight electron	Veto electron	Tight electron
$\sigma_{\eta\eta}$	< 0.0115	< 0.00998	< 0.037	< 0.0292
$ \Delta\eta_{in} $	< 0.00749	< 0.00308	< 0.00895	< 0.00605
$ \Delta\phi_{in} $	< 0.228	< 0.0816	< 0.213	< 0.0394
H/E	< 0.356	< 0.0414	< 0.211	< 0.0641
relative isolation	< 0.175	< 0.0588	< 0.159	< 0.0571
$ 1/E - 1/p $	$< 0.299 \text{ GeV}^{-1}$	$< 0.0129 \text{ GeV}^{-1}$	$< 0.15 \text{ GeV}^{-1}$	$< 0.0129 \text{ GeV}^{-1}$
expected missing inner hits	$\leq 2$	$\leq 1$	$\leq 3$	$\leq 1$
pass conversion veto	Y	Y	Y	Y

**Table 4.4:** Jet criteria used throughout the thesis. The last three requirements are only for jets within the tracker acceptance.

Properties	Loose Jet ID
Neutral hadron fraction	< 0.99
Neutral EM fraction	< 0.99
Number of constituents	> 1
Charged hadron fraction	> 0
Charged multiplicity	> 0
Charged EM fraction	< 0.99

1145 The energy of the reconstructed jets deviates from the energies of the corresponding jets  
 1146 clustered from the hadronisation products of true partons from simulations due to non-linear  
 1147 subdetector response and efficiencies. The jet energy corrections (JEC) calibrate the jets  
 1148 in order to have the correct energy scale and resolution. Jet energy scale corrections (JES)  
 1149 are determined as a function of pseudorapidity and the transverse momentum from data and  
 1150 simulated events by combining several channels and methods. This is extensively described  
 1151 in [172]. These corrections account for the effects of pileup, the uniformity of the detector  
 1152 response, and residual data-simulation jet energy scale differences. Furthermore, the jet energy  
 1153 resolution (JER) is measured in data and simulation as function of pileup, jet size and jet flavour.  
 1154 The performance of the jet energy corrections for the 13 TeV dataset can be found in [173].

1155 The JEC are factorised and subsequently correct for the off-set energy due to pileup, the  
 1156 detector response to hadrons, and residual differences between data and simulation as a function  
 1157 of the jet pseudorapidity and transverse momentum. The consecutive steps of JEC are shown  
 in Figure 4.5. The off-set corrections remove the dependency of the jet energy response of



**Figure 4.5:** The sequence of the JEC for data and simulations. The corrections marked with MC are derived from simulation studies, while RC stands for random cone, and MJB for the analysis of multi-jet events. Figure taken from [172].

1158 additional pileup activity. It is based on the jet area method, which uses the effective area  
 1159 of the jets multiplied by the average density in the event, to calculate the off-set energy to  
 1160 be subtracted of the jets. The correction factors are derived by comparing the jet response

with and without pileup events overlaid. The residual differences between data and detector simulation are determined using the random cone method (RC). For this method, many jets are reconstructed in each event by clustering particles through placing random cones. This provides a mapping of the  $\eta\phi$ -space and the average  $p_T$  of those jets gives the average energy off-set due to pileup [172]. The next level of corrections have as goal to have a uniform energy response independent of the transverse momentum or pseudorapidity of the jet. These corrections are determined from simulated events by matching the reconstructed to true particle jets and comparing their momenta. The residual corrections between data and simulation are determined by comparing the transverse momentum balance in various types of events (multi-jet, Z + jets, and  $\gamma$  + jets), using a reference jet in the barrel region. The jet flavour corrections are optional and not used for this thesis. More information on the jet flavour corrections can be found in [172]. For jets with a transverse momentum above 30 GeV, the uncertainties from the various corrections are 3-5% for the 13 TeV dataset [173].

After applying JEC, the transverse momentum resolution of the jet is extracted from data and simulated events. There are two methods used to rescale the reconstructed four momentum based on whether or not the simulated jet can be matched to a true jet in simulation. The factors are defined as

$$c_{\text{matched}} = 1 + \frac{p_T^{\text{reco.}} - p_T^{\text{true}}}{p_T^{\text{reco.}}} (s_{\text{JER}} - 1), \\ c_{\text{unmatched}} = 1 + N(0, \sigma_{\text{JER}}) \sqrt{\max(s_{\text{JER}}^2 - 1, 0)}, \quad (4.4)$$

where  $N(0, \sigma_{\text{JER}})$  denotes a sample value from a normal distribution centred at zero with standard deviation  $\sigma_{\text{JER}}$ , and  $s_{\text{JER}}$  the  $\eta$ -dependent resolution scale factors. These scale factors are derived from data from di-jet or  $\gamma$  + jets events and analysing the  $p_T$  balance. The resolution scale factors (data/simulation) are found to be 1.1-1.2, except for the transition regions around  $|\eta| = 3$  and  $|\eta| = 1.4$  [173] and given in Table 4.5.

#### 4.4.4 Jets from b fragmentation

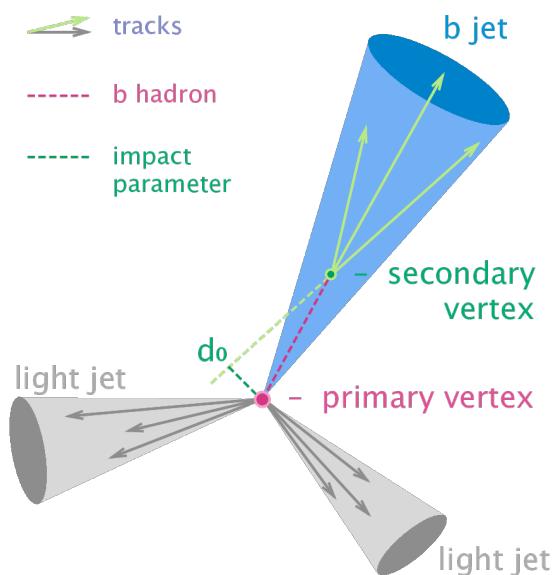
Jets originating from the hadronisation of bottom quarks can be discriminated from jets from gluons and light-flavour quarks as well as charm quark fragmentation through the use of b-tagging. There are several algorithms developed within CMS to perform b-tagging [174, 175] on jets that fall within the pseudorapidity acceptance of the trackers. These algorithms exploit the properties of the b quark to identify the jets formed by its fragmentation. These hadrons have relative large masses, long lifetimes and daughter particles with hard momentum spectra. Additionally, their semi-leptonic decays can be exploited as well. To use b-jet identification in an analysis, one needs to know its efficiency and misidentification probability. In general, these are function of the pseudorapidity and transverse momentum of the considered jet. Their performances are directly measured from data by use of b-jet enriched jet samples (multi-jet or top-quark decays).

This thesis uses b-jets identified by the Combined Secondary Vertex version 2 (CSVv2) algorithm [174]. This algorithm combines secondary vertices together with track based lifetime

**Table 4.5:** Jet energy scale factors in bins of  $\eta$  with uncertainty

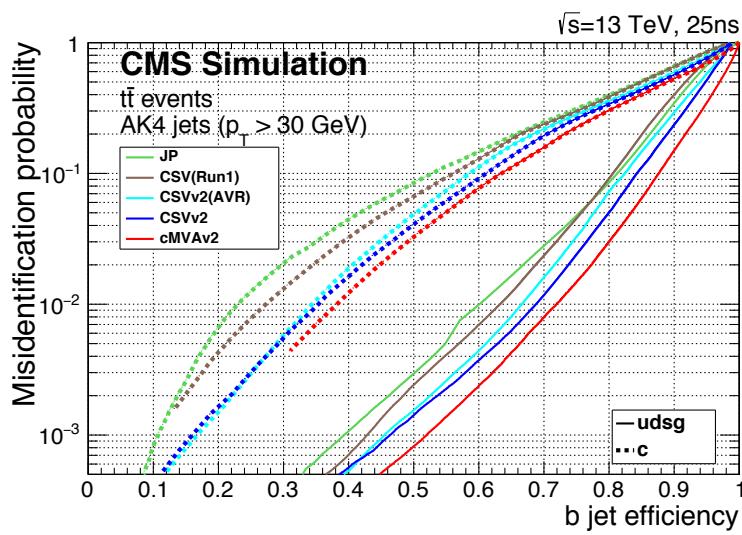
$ \eta $	SF	Uncertainty ( $\pm$ )
0-0.5	1.109	0.008
0.5-0.8	1.138	0.013
0.8-1.1	1.114	0.013
1.1-1.3	1.123	0.024
1.3-1.7	1.084	0.011
1.7-1.9	1.082	0.035
1.9-2.1	1.140	0.047
2.1-2.3	1.067	0.053
2.3-2.5	1.177	0.041

information by use of a multivariate technique. The secondary vertex is reconstructed from displaced tracks within a jet, as illustrated in Figure 4.6. The final state b quark is encapsulated in a B meson (e.g.  $B^\pm$ ,  $B_0$ ,  $B_S$ ) after the hadronisation. This B meson has relatively long lifetime and can travel a measurable distance from the primary vertex before decaying. After reconstruction, the secondary vertices are required to be in accordance with the B meson hypothesis based on the amount of shared tracks with the primary vertex, the invariant vertex mass to reject kaon decays, and the direction of the tracks compared to the jet axis.

**Figure 4.6:** Sketch showing the common principle of the identification of b-jets. Figure taken from [176].

1202 The b-tagging algorithm performances are evaluated taking into account two cases: discrimination  
 1203 of b-tagged jets originating from charm quarks, and discrimination of b-tagged jets  
 1204 against jets coming from gluons or light ( $u, d, s$ ) quarks. In Figure 4.7, the misidentification  
 1205 probabilities for different b-tagging algorithms within CMS are shown. Based on the misidentification  
 1206 probabilities for a certain threshold on the CSVv2 discriminator, different working points  
 1207 are defined. These are shown in Table 4.6. The analysis presented in this thesis uses the loose  
 working point which has an average efficiency of 81% and a misidentification rate of 10%.

**NOTE:** Find reason why I am not using cMVA, omdat CSVv2 op ttbar events is gemeten en cMVA op multijet?



**Figure 4.7:** Misidentification probabilities of various b-tagging algorithms in simulation. Figure taken from [175].

**Table 4.6:** Working points used for tagging jets as coming from b quarks for the CSVv2 discriminant.

WP	CSVv2 discr cut	b-tag eff.	misid. prob.
Loose (L)	$> 0.5426$	$\approx 81\%$	$\approx 10\%$
Medium (M)	$> 0.8484$	$\approx 66\%$	$\approx 1\%$
Tight (T)	$> 0.9535$	$\approx 46\%$	$\approx 0.1\%$

1208

1209 The efficiency of tagging a jet as coming from a bottom quark in simulation typically deviates  
 1210 somewhat from data. Efficiency scale factors  $e_b^{\text{data}}/e_b^{\text{MC}}$  are derived from data to account for  
 1211 those differences. These scale factors are  $\eta$ -,  $p_T$ -, and flavour dependent, where the flavour  
 1212 of the jet is determined from matched generated hadrons. For cut based analyses these scale  
 1213 factors are applied to the b-tagging efficiencies and mistag rates according to the chosen working  
 1214 point [175]. For shape based analysis however, such as the one presented in this thesis, the  
 1215 scale factors are applied on the distribution of the b-tagging discriminator. This is the so-called  
 1216 IterativeFit method [177].

1217 The uncertainties related to the IterativeFit method cover possible shape discrepancies between  
 1218 data and simulation. The uncertainty coming from the purity of the sample is subdivided into  
 1219 two uncorrelated uncertainties based on the purity of the light flavoured (lf) and heavy flavoured  
 1220 (hf) jet contributions in the sample. Furthermore, the jet energy scale results in jets migrating  
 1221 from one  $p_T$  bin to another, having an influence on bin dependent scale factors. The statistical  
 1222 fluctuations of the limited amount of entries in each bin are also accounted for and have an  
 1223 influence on the scale factor uncertainties. These have four uncorrelated sources: two for heavy  
 1224 flavour and two for light flavour jets. Since the uncertainty on the scale factors for the jets  
 1225 originating from a charm quark (cf) is determined from the uncertainty on the b scale factors  
 1226 resulting in two independent uncertainties [177].

#### 1227 4.4.5 Missing transverse energy

The missing transverse momentum  $\vec{p}_T$  and energy  $E_T^{\text{miss}}$  resulting from particles that do not interact with the detector material, are calculated by balancing the vectorial sum of the transverse momenta of all particles:

$$E_T = |\vec{p}_T|, \quad (4.5)$$

$$\vec{p}_T = - \sum_{i=1}^{N_{\text{particles}}} \vec{p}_{T,i}.$$

1228 The missing transverse energy is influenced by the minimum thresholds in calorimeters, the  
 1229 inefficiencies in the tracker, and the non-linear response of the calorimeter to hadronic particles.  
 1230 The bias is reduced by correcting the transverse momentum of the jets to particle jet  $p_T$  via the  
 1231 JEC and propagating it to the missing transverse momentum. The performance of the missing  
 1232 transverse energy reconstruction can be found in [178].

### 1233 4.5 Summary of corrections

1234 Throughout the chapter several corrections are introduced to improve the agreement between  
 1235 data and simulation. These corrections are sources of systematic uncertainties for the anal-  
 1236 ysis presented in this thesis. Therefore a summary of the corrections and their associated  
 1237 uncertainties is provided.

1238 **Lepton scale factors** The systematic uncertainty on the lepton scale factors consists of three  
 1239 sources: identification, isolation and tracking. The applied scale factors are varied  
 1240 independently within one standard deviation of their measured uncertainties to account  
 1241 for their systematic impact on the measurements.

1242 **Jet energy corrections** The momenta of the reconstructed jets are corrected to match the  
 1243 expected true energy derived from the hadronisation products of partons in simulation.  
 1244 Furthermore, residual corrections and smearing is applied to match the overall energy  
 1245 scale and resolution for simulation and data. These corrections are also propagated to  
 1246 the missing transverse energy. The systematic uncertainties due to these scale factors are  
 1247 estimated by varying them within their uncertainties and repeating the measurements  
 1248 with recalibrated jets and missing transverse energy.

1249 **CSVv2 discriminant shape reweighting** There are three sources of uncertainty contributing  
1250 to the measurement of the scale factors: statistical uncertainties, jet energy scale and  
1251 the purity of the sample. The jet energy scale uncertainty is 100% correlated to the jet  
1252 energy uncertainties and is evaluated simultaneously. The uncertainty coming from the  
1253 purity of the sample is subdivided into two uncorrelated uncertainties based on the purity  
1254 of the light flavoured (lf) and heavy flavoured (hf) jet contributions in the sample. A  
1255 one sigma shift in each of the two purity contributions corresponds to a higher/lower  
1256 contribution in the purity of the considered flavours. The statistical uncertainties have  
1257 four uncorrelated sources, two for heavy flavour and two for light flavour jets. One of  
1258 the uncertainties correspond to the shift consistent with the statistical uncertainties of  
1259 the sample, while the other is propagated in a way that the upper and lower ends of the  
1260 distribution are affected with respect to the centre of the distribution. The uncertainty  
1261 on the charm jet scale factors (cf) is obtained from the uncertainty on the heavy flavour  
1262 scale factors, doubling it in size and constructing two nuisance parameters to control the  
1263 charm flavour scale factors and treating them as independent uncertainties.

1264 **Pileup** Varying the minimum bias cross section, used to calculate the pileup distribution by  
1265  $\pm 4.6\%$ , results in a systematic shift in the pileup distribution. The uncertainty is estimated  
1266 by recalculating the pileup weights to the distributions associated to the minimum bias  
1267 cross sections.

1268 **Luminosity** The luminosity is measured with a global uncertainty of 2.5%, affecting the ex-  
1269 pected number of events.

# Event selection and categorisation

5

1270

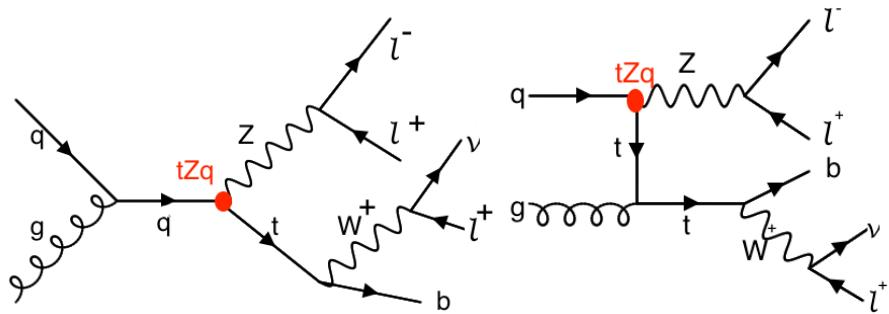
1271 A basic event selection is made for selecting signal like events and is discussed in [Section 5.1](#).  
1272 Also the effect corrections applied to simulation and data, summarised in [Section 4.5](#) is shown.  
1273 The analysis strategy is presented in [Section 5.2](#), defining the signal and background regions  
1274 to constrain the huge SM background compared to the expected signal. [Section 5.4](#) discusses  
1275 each region that is entering the analysis. On top of the use of background estimation from  
1276 control regions, backgrounds that have prompt leptons contaminated by real leptons either from  
1277 decays of tau leptons or from hadronized mesons or baryons (collectively commonly referred as  
1278 “non-prompt leptons”) as well as by hadrons or jets misidentified as leptons<sup>1</sup> are evaluated with  
1279 a data-driven method discussed in [Section 5.3](#).

## 1280 5.1 Baseline event selection and filters

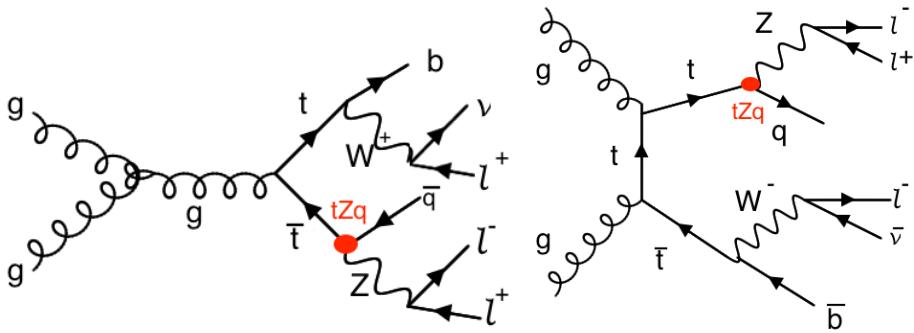
1281 In this analysis a search is performed in a final state made up of a Z boson and a top quark,  
1282 associated or not with a jet, in the leptonic decay of the Z boson and the top quark. The leading  
1283 order Feynman diagrams can be seen in [Figure 5.1](#) and [Figure 5.2](#). The signal considers both the  
1284 single top quark FCNC ( $tZ$  in the final state) and the top quark pair FCNC ( $tZq$  with  $q = c, u$  in  
1285 the final state) events. Their final state signatures consist of three leptons, considering electrons  
1286 or muons, and a jet originating from a b quark. For FCNC  $tZq$ , there is an additional up or charm  
1287 jet. Leptons from tau decays are not vetoed and are entering the analysis via their leptonic  
1288 decays. Four different lepton channels based on lepton flavour are considered: 3e, 2e1 $\mu$ , 1e2 $\mu$ ,  
1289 and 3 $\mu$ .

1290 The CMS collaboration recorded in the course of 2016 proton at 13 TeV collisions data with  
1291 a total recorded integrated luminosity of  $35.9 \text{ fb}^{-1}$ . The baseline event selection has a goal to  
1292 substantially reject SM background events, whilst maintaining a high signal efficiency. The  
1293 CMS trigger system, described in [Section 2.2.3](#), filters out the main of the collision events from  
1294 uninteresting processes and dedicated trigger paths are defined to single out the events with our  
1295 required detector signature. The trigger paths are chosen based on on-line triggering objects  
1296 with at least one muon (M), at least one electron (E), at least two muons (MM), at least two

<sup>1</sup>These two classes of contamination will be referred to as not prompt-lepton (NPL) samples.



**Figure 5.1:** Single top quark Feynman diagrams at leading order. The vertex labelled  $tZq$  is the sought-for FCNC interaction.



**Figure 5.2:** Top quark pair Feynman diagram at leading order. The vertex labelled  $tZq$  is the sought-for FCNC interaction.

1297 electrons (EE), at least one muon and an electron (ME), at least three muons (MMM), at least  
 1298 three electrons (EEE), at least two muons and one electron (MME), or at least two electrons  
 1299 and one muon (EEM). For the MC simulation a simple *or* of all triggers is taken and the event  
 1300 is considered when it passes one of the trigger paths. For data however, double counting of the  
 1301 same event has to be taken into account and a procedure to avoid double counting has been  
 1302 put into place. It consists of vetoing in a given dataset the events that are already selected in  
 1303 another, as given in Table 5.1.

1304 For the single lepton triggers, at least one electron (muon) with a transverse momentum  
 1305  $p_T$  higher than 32 (24) GeV is required. The dilepton triggers require an electron (muon)  
 1306 with  $p_T > 23$  GeV and a muon (electron) with  $p_T > 8$  GeV, or an electron (muon) with  
 1307  $p_T > 23$  (17) GeV and an electron (muon) with  $p_T > 12$  (8) GeV. Events collected by the  
 1308 trilepton triggers require an electron (muon) with  $p_T > 16$  (12) GeV, a second electron (muon)  
 1309 of  $p_T > 12$  (10) GeV, and a third electron (muon) with  $p_T > 8$  (5) GeV. The mixed trilepton  
 1310 trigger events require two electrons (muons) with  $p_T > 12$  (9) GeV and a third muon (electron)  
 1311 with  $p_T > 8$  (9) GeV. The HLT trigger paths used in data and simulation are summarised in  
 1312 Table 5.2.

1313 In order to ensure a full trigger efficiency, the off-line  $p_T$  thresholds are set higher than the on-

**Table 5.1:** Trigger logic used to select data events in order to avoid double counting.

Dataset	Trigger Logic
1e1μ	EM    EEM    MME
2μ	(MM    MMM) && !( EM    EEM    MME)
2e	(EE    EEE) && !(MM    MMM) && !( EM    EEM    MME)
single μ	M && !(EE    EEE) && !(MM    MMM) && !( EM    EEM    MME)
single e	E && !M && !(EE    EEE) && !(MM    MMM) && !( EM    EEM    MME)

line trigger thresholds. Selected electrons (muons) are required to have  $p_T > 35$  (30) GeV and  $|\eta| < 2.1(2.4)$ . The electrons and muons corresponding to a tight working point, as discussed in Section 4.4.1 (Table 4.1) and Section 4.4.2 (Table 4.3), are used for analysis. Only events with exactly three leptons are being considered. Events with extra leptons according to looser working points are vetoed. The trigger efficiency estimation is described in Section 5.1.2 and is approximately 100%. To ensure that all reconstructed particles considered for the analysis are corresponding to a proton interaction and to remove signals from beam halo particles as well as detector noise,, several filters are used. These are described in Section 5.1.1.

In addition to three leptons, jets and missing transverse energy are expected from the signal signature. The jets are reconstructed using the anti –  $k_T$  algorithm with a distance parameter of 0.4 using the particle flow particles that are not identified as isolated leptons as input. The jet momentum is determined as the vectorial sum of the particles contained in the jet. Additional selection criteria are applied to each event to remove spurious jet-like features originating from isolated noise patterns in certain hadron calorimeter regions. The jet requirements are discussed Section 4.4.3. The jets are calibrated in simulation and in data separately, accounting for deposits from pileup and the non-linear detector response. Calibrated jets with  $p_T > 30$  GeV and  $|\eta| < 2.4$  are selected for the analysis. A selected jet may still overlap with the selected leptons leading to a double-counting of reconstructed objects. To prevent such cases, jets that are found within a cone of  $R = 0.3$  around any of the signal electrons (muons) are removed from the collection of selected jets. The jets originating from b quarks are tagged using the CSVv2 algorithm, which combines the information of displaced tracks with information of secondary vertices associated with the jet using a multivariate technique. The jets are tagged as a jet coming from a b quark (b-tagged) if the CSVv2 discriminator is above a certain threshold. This analysis uses a threshold that results in an average b-tagging efficiency of 81% and a misidentification rate of 10%. More information about b-tagging can be found in Section 4.4.4.

The missing transverse momentum vector  $\vec{p}_T$  is defined as the projection on the plane perpendicular to the beams of the negative vector sum of the momenta of all reconstructed particles in an event. Its magnitude is denoted by  $E_T^{\text{miss}}$  as shown in Section 4.4.5. Its longitudinal component is calculated by limiting the lepton + neutrino to the W boson mass. In case two solutions arise, the mass closest to the known top quark mass is used.

**Table 5.2:** HLT trigger paths used to select data and simulation events.

Trigger path name	Trigger type
HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL_v	ME
HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL_DZ_v	ME
HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_v	ME
HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v	ME
HLT_DiMu9_Ele9_CaloIdL_TrackIdL_v	MME
HLT_Mu8_DiEle12_CaloIdL_TrackIdL_v	EEM
HLT_IsoMu24_v	M
HLT_IsoTkMu24_v	M
HLT_Ele32_eta2p1_WPTight_Gsf_v	E
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_v	MM
HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_v	MM
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v	MM
HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v	MM
HLT_TripleMu_12_10_5_v	MMM
HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v	EE
HLT_Ele16_Ele12_Ele8_CaloIdL_TrackIdL_v	EEE

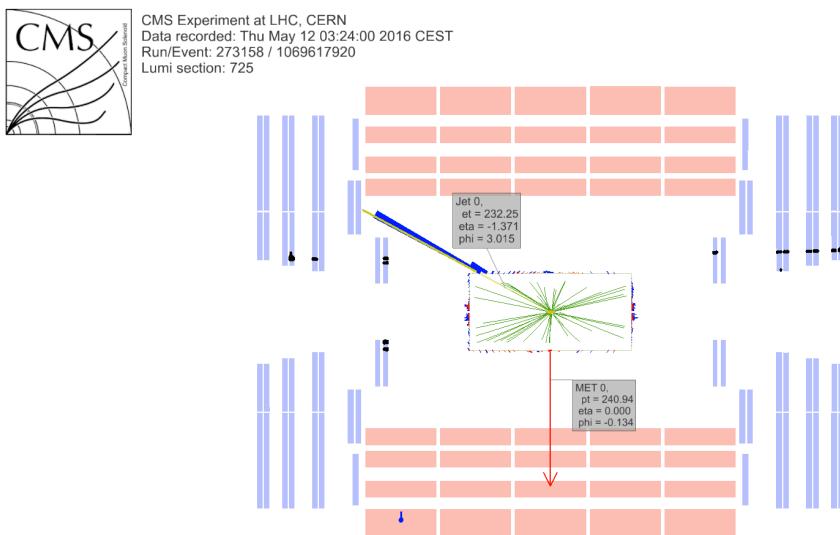
### <sup>1344</sup> 5.1.1 Event cleaning

<sup>1345</sup> Some events arising from instrumental noise and beam backgrounds might end up in the  
<sup>1346</sup> data [178, 179]. Spurious deposits may appear in the ECAL from non collision origins such as  
<sup>1347</sup> beam halo particles, or from particles hitting the sensors in the ECAL photodetectors. Conjointly,  
<sup>1348</sup> dead ECAL cells can cause artificial missing transverse energy. Also the HCAL can cause spurious  
<sup>1349</sup> energy from particle interactions with the light guides and the photomultiplier tubes of the  
<sup>1350</sup> HF, as well as noisy hybrid photo diodes. In CMS, different algorithms, so-called filters, are  
<sup>1351</sup> developed to identify and suppress these events.

<sup>1352</sup> The ECAL electronics noise and spurious signals from particle interactions with photo detectors  
<sup>1353</sup> are mostly removed via topological and timings based selection using ECAL information only.  
<sup>1354</sup> The remaining effects such as anomalously high energy crystals and the lack of information  
<sup>1355</sup> for channels due to inefficiencies in the read out are removed through dedicated events filters.  
<sup>1356</sup> Five ECAL endcap supercrystal have been identified for giving anomalously high energies due  
<sup>1357</sup> to high amplitude pulses in several channels at once, and are masked. Furthermore, the crystal  
<sup>1358</sup> read out from a small amount of ECAL towers is not available. However, their trigger primitive

information is still available making it possible to estimate the magnitude of unmeasured energy and when the value is too large, the event is filtered out.

The machine induced particles, via beam-gas / beam-pipe/... interactions, that are flying with the beam affect the physics analysis. They leave a calorimeter deposit along a line at constant  $\phi$  in the calorimeter, and interactions in the CSCs will often line up with this deposit. This can be seen in Figure 5.3. Therefore, events containing such beam halo particles are removed from the selection with the CSC Beam Halo Filter. This filter uses information related to the geometric quantities, energy deposits, and timing signatures. For 2016, the filter rejects 85% in a halo-enriched sample, whereas the mistag probability determined from simulation if found to be less than 0.01%.



**Figure 5.3:** Event display of a beam halo event with collinear hits in the CSC (black), missing transverse energy of 250 GeV and a jet of 232 GeV. The hadronic deposit is spread in  $\eta$ , but narrow in  $\phi$ . Figure taken from [178].

Furthermore, there is anomalous high missing transverse energy coming from low quality muons that lead to high- $p_T$  tracks, but are considered not good by the particle flow algorithm. These low quality tracks will be mislabelled as charged hadrons and will therefore be used in the calculation of the missing transverse energy. By investigation the purity of the reconstructed tracks and the relative transverse momentum error of the muons, these events can be filtered out.

Supplementary to previous filters, only events with where the first primary vertex is a well reconstructed primary vertex are selected. The reconstructed primary vertex should have at least five degrees of freedom, the longitudinal distance from the beam spot is maximally 24 cm ( $d_z < 24$  cm), and the transversal distance from the beam spot is maximally 2 cm ( $d_{xy} < 2$  cm).

1379 **5.1.2 Estimation of the trigger efficiency**

1380 The trigger efficiency in data is estimated using a data sample collected using unprescaled  $E_T^{\text{miss}}$   
 1381 triggers. These allow events with a missing transverse energy higher than 110 GeV(120 GeV)  
 1382 and that the scalar sum of the transverse momenta of the reconstructed PF jets  $H_T^{\text{trig.}}$  is at least  
 1383 300 GeV (120 GeV), or a calorimeter (PF)  $E_T^{\text{miss}}$  higher than 200 GeV(300 GeV). For an HB-HE  
 1384 cleaned event the PF missing transverse energy threshold is lowered to 170 GeV. These trigger  
 1385 paths are summarised in [Table 5.3](#) and chosen to be completely uncorrelated with the lepton  
 triggers given in [Table 5.2](#).

**Table 5.3:** Unprescaled  $E_T^{\text{miss}}$  HLT trigger paths for estimating the trigger efficiency. An *or* of the triggers  
 is used to select events.

Trigger path	Requirement
HLT_PFHT300_PFMET110_v*	PF $E_T^{\text{miss}} > 110$ GeV, PF $H_T^{\text{trig.}} > 300$ GeV
HLT_MET200_v*	calorimeter $E_T^{\text{miss}} > 200$ GeV
HLT_PFMET300_v*	PF $E_T^{\text{miss}} > 300$ GeV
HLT_PFMET120_PFHT120_IDTight_v*	PF $E_T^{\text{miss}} > 120$ GeV, PF $H_T^{\text{trig.,tightWP}} > 120$ GeV
HLT_PFMET170_HBHECleaned_v*	PF $E_T^{\text{miss}} > 170$ GeV, cleaned for HB/HE anomalous signals

1386

The trigger efficiency is studied on the main background, namely WZ+jets, with all corrections applied. For this study, the events passing a three lepton cut and at least one jet are being used. The corresponding efficiencies are then calculated as

$$\epsilon_{\text{data}} = \frac{\text{Nb. of events passing lepton and MET triggers}}{\text{Nb. of events passing MET triggers}} \quad (5.1)$$

$$\epsilon_{\text{MC}} = \frac{\text{Nb. of events passing lepton triggers}}{\text{Nb. of total events}} \quad (5.2)$$

The resulting efficiencies and scale factors can be found in [Table 5.4](#), where the scale factors are defined as

$$SF = \frac{\epsilon_{\text{data}}}{\epsilon_{\text{MC}}} \quad (5.3)$$

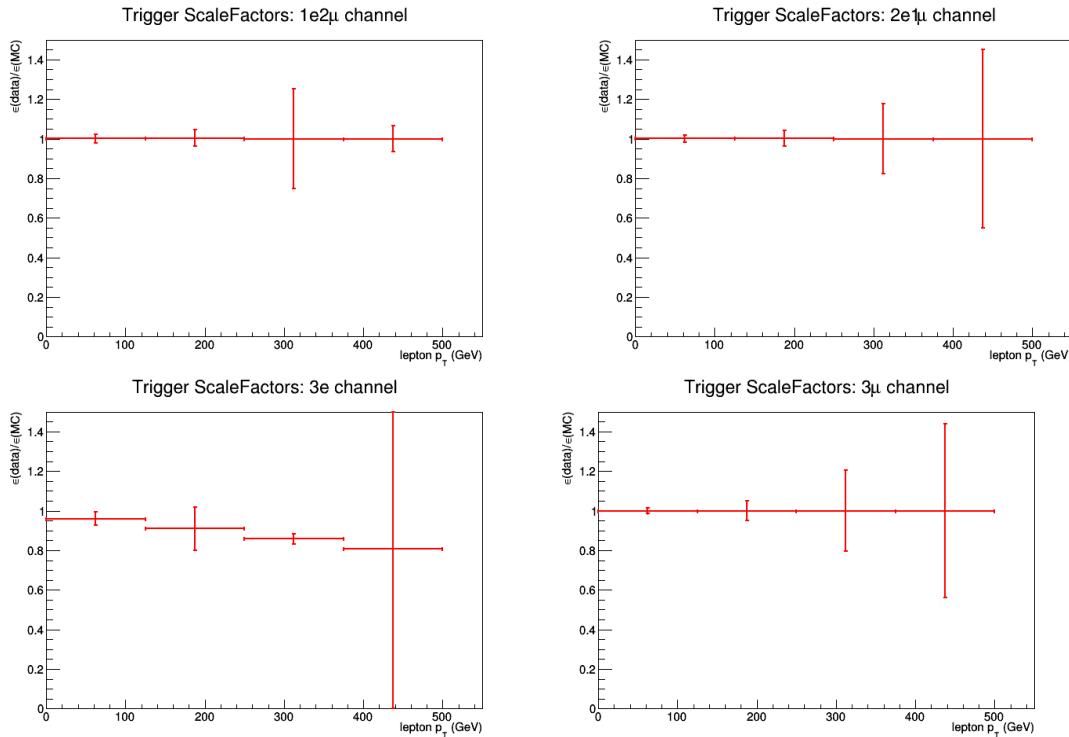
1387 More detailed scale factors and efficiencies can be found in [Appendix A](#).

1388 The trigger efficiencies are also measured in function of the  $p_T$  of the leptons for which the  
 1389 distributions can be found in [Appendix A](#). The resulting scale factors can be found in Figure  
 1390 [Figure 5.4](#). The trigger efficiencies are measured to be nearly 100% for both simulation and  
 1391 data. The results are dominated by statistics and assigning a large uncertainty to the trigger  
 1392 efficiency based on the dataset collected by  $E_T^{\text{miss}}$  triggers would be over conservative. A one  
 1393 percent uncertainty on the trigger selection for the 2e1 $\mu$  and 3 $\mu$  final states, and 5% for the  
 1394 3e and 1e2 $\mu$  final states is assigned instead. No scale factors will be applied on simulation as  
 1395 they are close to unity. Control plots are made in the dilepton region to validate all corrections  
 applied to simulation and can be found in [Appendix B](#).

**NOTE:** 1396  
 Make sure  
 this is the  
 case

**Table 5.4:** Trigger scale factors for each channel, after three lepton + jets cut, in the Z mass window, by counting number of events.

	all	3 $\mu$	3e	2e1 $\mu$	1e2 $\mu$
	1.0000	1.0000	0.9541	1.0006	1.0004%



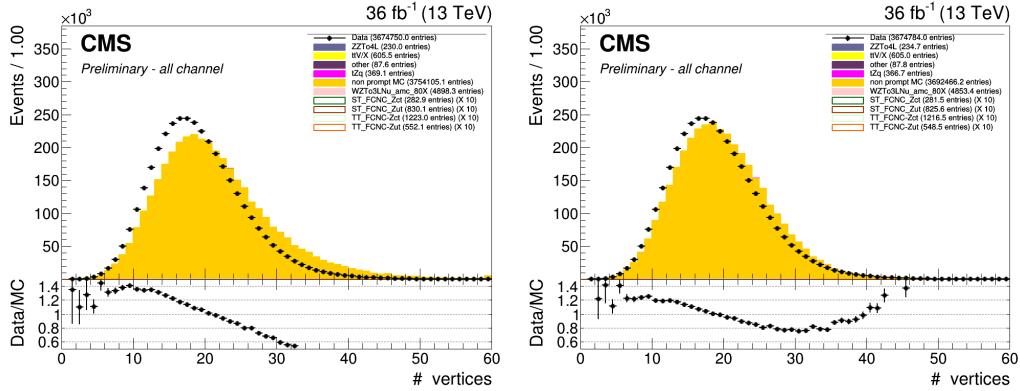
**Figure 5.4:** The trigger scale factors measured as a function of lepton  $p_T$ , using the dataset collected by  $E_T^{\text{miss}}$  triggers and WZ simulation, after a 3 lepton and jets selection, in the Z mass window. All corrections to simulation are applied. Left, upper: 1e2 $\mu$  channel. Right, upper: 2e1 $\mu$  channel. Left, lower: 3e channel. Right, lower: 3 $\mu$  channel

### 1397 5.1.3 Corrections

1398 Mismatches between data and simulation are corrected for via the use of scale factors. These are  
 1399 elaborately discussed in [Section 4.4](#). In this section a short overview of the applied corrections  
 1400 is given and their effect on a dilepton selection is shown.

### 1401 Pileup reweighting

1402 In data, the number of interactions per bunch crossing (pileup) is calculated with a minimum  
 1403 bias cross section of 69.2 mb. The number of simulated pileup events is then reweighted to  
 1404 match the expected number of pileup events in data. Pileup reweighting manifests itself as an  
 1405 altered shape of the number of reconstructed primary vertices as can be seen in [Figure 5.5](#).



**Figure 5.5:** The number of primary vertices before (left) and after (right) pileup reweighting. After a dilepton plus jets selection, in the Z mass window.

1406 Note that Figure 5.5 indicates that even after pileup reweighting, the primary vertex multi-  
1407 plicity is not well described by simulation. This is a known effect, and using a minimum bias  
1408 cross section with a slightly lower value is found to better describe the data. However, the b  
1409 tagging scale factors are only provided for the nominal inelastic cross section, and thus this  
1410 value is used.

### 1411 Lepton scale factors

The efficiency to select leptons is different in simulation ( $\epsilon_{\text{MC}}$ ) compared to the data ( $\epsilon_{\text{data}}$ ). This is corrected for by applying lepton scale factors (SF) to the simulation that are defined as

$$SF = \frac{\epsilon_{\text{data}}}{\epsilon_{\text{MC}}} \quad (5.4)$$

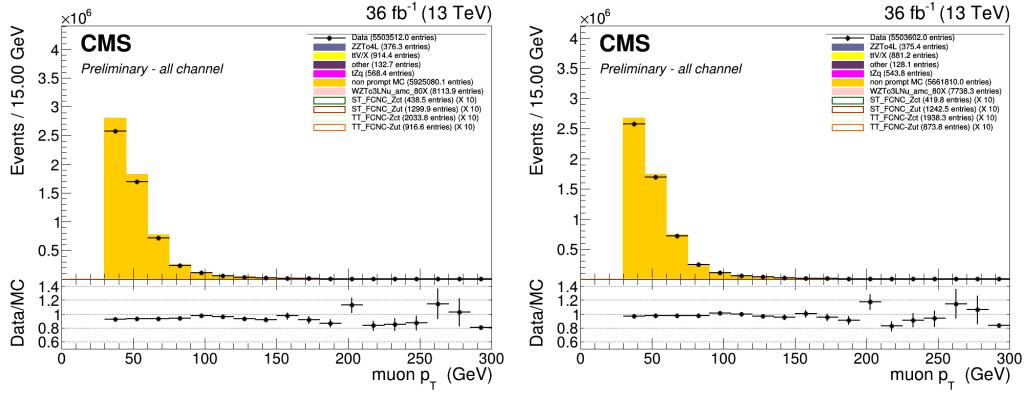
These scale factors are measured for the identification, isolation, tracking and trigger efficiencies of the objects as a function of  $p_{\text{T}}$  and  $\eta$  (see Section 4.4.1 and Section 4.4.2). Multiplying these scale factors for each lepton provides an overall efficiency:

$$SF_{\text{global}}^{\mu} = \prod_{\text{i}}^{\#\mu} SF_{\text{ID}}^{\mu}(p_{\text{T}}, \eta) SF_{\text{Iso.}}^{\mu}(p_{\text{T}}, \eta) SF_{\text{Trig.}}^{\mu}(p_{\text{T}}, \eta) SF_{\text{track}}^{\mu}(p_{\text{T}}, \eta), \quad (5.5)$$

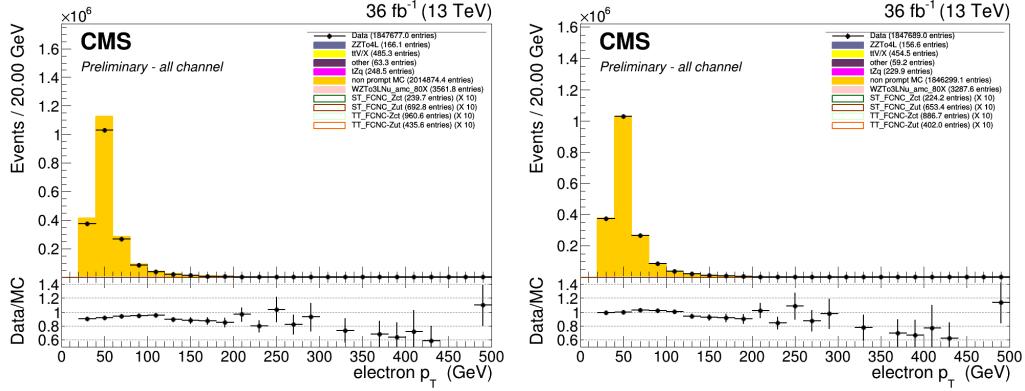
$$SF_{\text{global}}^e = \prod_{\text{i}}^{\#e} SF_{\text{ID}}^e(p_{\text{T}}, \eta) SF_{\text{Iso.}}^e(p_{\text{T}}, \eta) SF_{\text{Trig.}}^e(p_{\text{T}}, \eta) SF_{\text{track}}^e(p_{\text{T}}, \eta). \quad (5.6)$$

1412 The effect of the scale factors can be found in Figure 5.7 for the electron scaling and Figure 5.6  
1413 for the muons. The trigger efficiencies are estimated in Section 5.1.2.

1414 Additionally, corrections are made for the energy resolution of the leptons. For the electrons,  
1415 energy smearing and regression is applied [180]. The energy regression uses the detector  
1416 information to correct the electron energy in order to have the best energy resolution by  
1417 correcting local energy containment, material effects, etc.. The energy scale and smearing

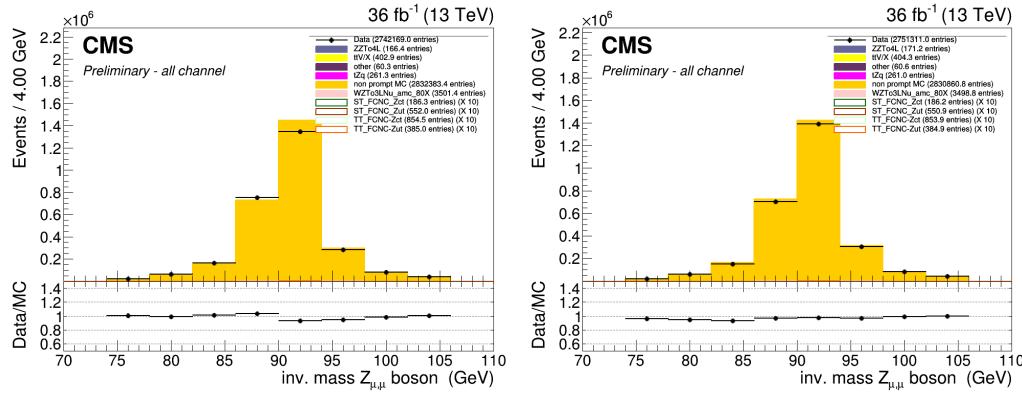


**Figure 5.6:** The  $p_T$  of the muons before (left) and after (right) muon scale factors. After a dilepton plus jets selection, in the Z mass window. Both after the Rochester correction.



**Figure 5.7:** The  $p_T$  of the electrons before (left) and after (right) electron scale factors. After a dilepton plus jets selection, in the Z mass window. Both after energy scale corrections and smearing.

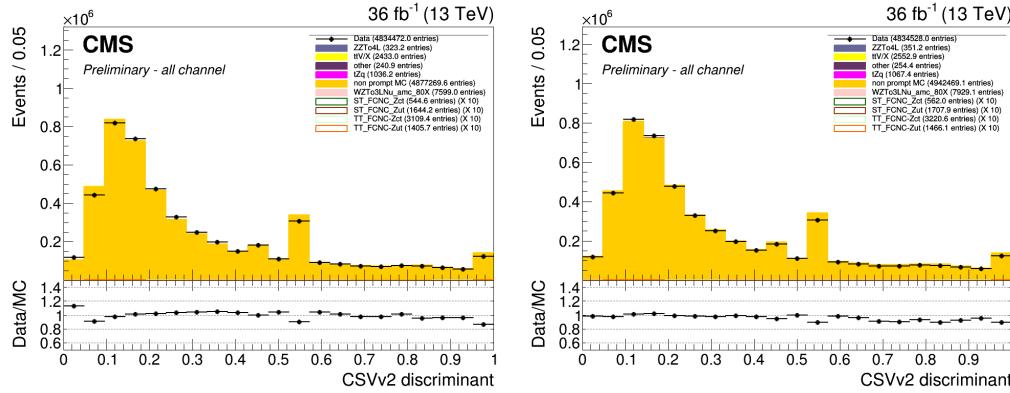
1418 is done in order to bring the data energy scale to simulation level. It smears the simulation  
 1419 energies to have identical energy resolution in simulation and data. For the muons, the  $p_T$  is  
 1420 corrected using the Rochester method [181, 182]. This correction removes the bias of the muon  
 1421  $p_T$  from any detector misalignment or any possible error of the magnetic field. The effect of  
 1422 the Rochester correction can be found in Figure 5.8.



**Figure 5.8:** The mass of the Z boson consisting of the muons before (left) and after (right) the Rochester correction. After a dilepton plus jets selection, in the Z mass window.

#### 1423 CSVv2 shape correction

1424 In order to make the shape of the CSVv2 b-tagging discriminant in simulation agree with data,  
 1425 jet-by-jet based scale factors are applied. These scale factors are a function of the  $p_T$ ,  $\eta$  and  
 1426 CSVv2 value of the jet as discussed in Section 4.4.4. The effect of these scale factors can be  
 1427 found in Figure 5.9.

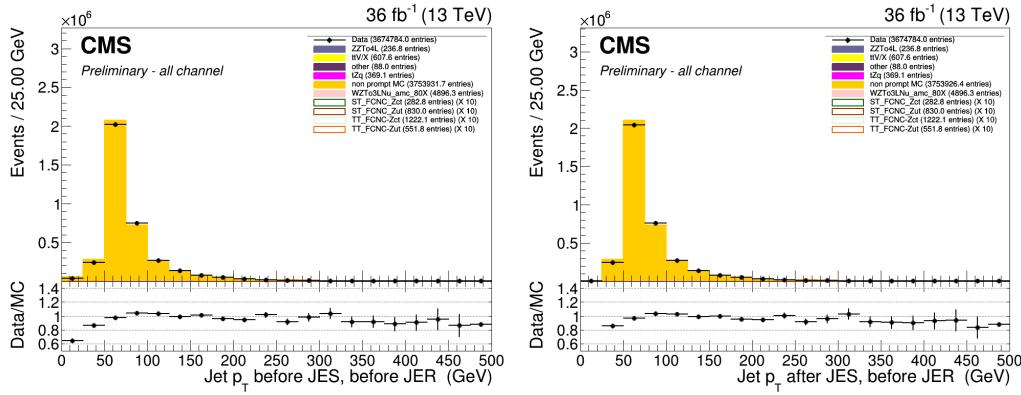


**Figure 5.9:** The CSVv2 discriminant of the jets before (left) and after (right) b-tag scale factors. After a dilepton plus jets selection, in the Z mass window.

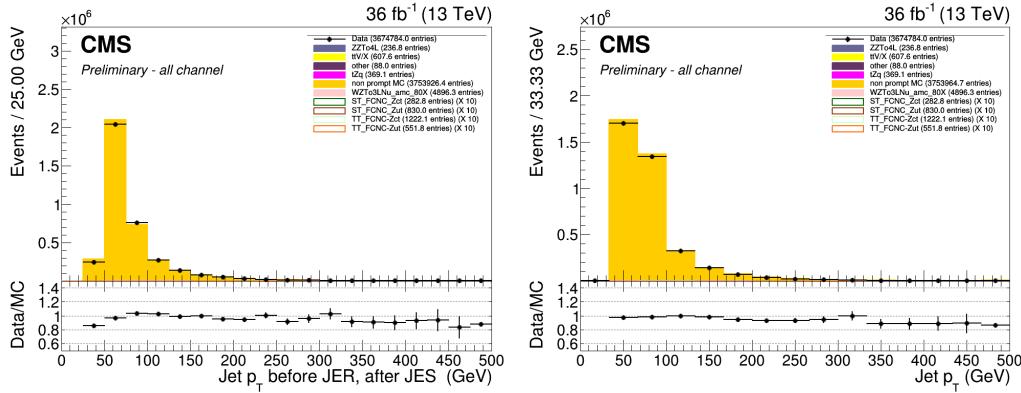
#### 1428 Jet energy

1429 The jet energy in data and simulation is corrected by the measured energy response of the  
 1430 detector. This provides  $p_T$ - and  $\eta$  dependent scale factors and are directly taken from the  
 1431 frontier condition database as discussed in Section 4.4.3. The effect of the jet energy corrections  
 1432 can be found in Figure 5.10 and Figure 5.11.

**NOTE:** Fix32  
jer plot



**Figure 5.10:** The  $p_T$  of the jets before (left) and after (right) jet energy scale corrections. After a dilepton plus jets selection, in the Z mass window.



**Figure 5.11:** The  $p_T$  of the jets before (left) and after (right) jet energy resolution smearing. After a dilepton plus jets selection, in the Z mass window.

## 1433 Missing transverse energy

1434 The energy scale and resolution corrections applied to the jets are propagated back to the  
1435 missing transverse energy (smeared Type I correction) as discussed in [Section 4.4.5](#). This  
1436 rebalances the transverse net momentum of the event and improves the missing transverse  
1437 energy resolution itself.

### 1438 5.1.4 Reconstruction of kinematic variables

1439 The variables used for the training are related to the reconstructed leptons, jets, Z boson and  
 1440 top quark candidates. The Z boson is reconstructed as the sum of the four vectors of the  
 1441 two same flavour leptons of opposite sign giving the closest value to the Z mass. The third  
 1442 remaining lepton is assigned as the lepton coming from the W boson decay. The reconstruction  
 1443 of the SM top quark candidate is more difficult and done by summing the third lepton, the  
 1444 SM b-jet and the neutrino ( $E_T^{\text{miss}}$ ). The SM b jet is assigned to the jet with the highest CSVv2  
 1445 discriminant. The longitudinal momentum of the neutrino is calculated by putting a constraint

1446 on the lepton+neutrino system with the W mass. In case two solutions are found for the  $p_z^\nu$   
 1447 component, the one that gives the reconstructed mass (lepton + neutrino + b jet) to the known  
 1448 top quark mass is used. The FCNC top quark is reconstructed by summing the reconstructed Z  
 1449 boson and the jet giving the mass closest to the top mass, from the jet collection from which the  
 1450 SM b jet is removed.

1451 The reconstructed objects are validated using simulation by matching the reconstructed  
 1452 objects to their generated counterpart by minimizing  $\Delta R$ . The efficiencies derived for the  
 1453 simulated signal samples and the SM tZq background process can be found in [Table 5.5](#) and  
 1454 [Table 5.6](#).

**Table 5.5:** Efficiencies of assigning the correct leptons in the analysis.

	FCNC tZq	FCNC tZ	SM tZq
W lepton	99%	98%	97%
leptons from the Z boson	99%	98%	97%
all leptons in the decay	99%	98%	97%

**Table 5.6:** Efficiencies of assigning the correct jets in the analysis.

	FCNC tZq	FCNC tZ	SM tZq
SM b jet	99%	98%	80%
c jet	71%	N/A	50%
u jet	83%	N/A	54%

## 1455 5.2 Analysis Strategy

1456 The analysis strategy uses five statistically independent regions to extract limits using a likelihood  
 1457 fit of various observables. Two signal regions, the tZ (STSR) and tZq (TTSR) signal regions, are  
 1458 constructed using the jet multiplicity, focussed on each signal signature (see Tab. 5.9). In order  
 1459 to constrain the rate of WZ+jet events as well as that of NPL backgrounds, three control regions  
 1460 are defined. The WZ control region (WZCR) focusses on NPLs originating from  $Z/\gamma^*$  + jets  
 1461 and simultaneously constrains the WZ+jets background rate. The NPL backgrounds coming  
 1462 from a  $t\bar{t}$  process are constrained by two control regions, TTCSR and STCSR, one for each signal  
 1463 region (TTSR and STSR). In the STSR and TTSR multivariate discriminants based on Boosted  
 1464 Decision Trees (BDT) (see [Section 3.3](#)) are used to respectively discriminate FCNC tZ and FCNC  
 1465 tZq from backgrounds. In the WZCR a discriminating variable between the two backgrounds,  
 1466 WZ+jets and NPLs, is used. In TTCSR and STCSR the dominating process is the  $t\bar{t}$  process, and  
 1467 its rate is estimated by subtracting all other background predictions from data. A simultaneous  
 1468 global fit using the Higgs Combined Tool ([Section 3.4](#)) is performed taking into account each  
 1469 region (STSR, TTSR, WZCR, TTCSR and STCSR) for the four different leptonic channels.

<sup>1470</sup> **5.3 Data driven NPL background simulation**

<sup>1471</sup> One of the most important background consist of events with not prompt leptons. These are  
<sup>1472</sup> mostly instrumental background and are therefore very difficult to model. The NPL background  
<sup>1473</sup> is estimated from data for both its shape and its normalisation.

<sup>1474</sup> The NPL background originates from hadronic objects wrongly reconstructed as leptons, real  
<sup>1475</sup> leptons coming from the semi-leptonic decay of a b or c hadron, or real leptons coming from  
<sup>1476</sup> the conversion of photons that pass the identification and isolation requirements. The dominant  
<sup>1477</sup> source of these NPLs depend on the flavour of the lepton and therefore the events with a not  
<sup>1478</sup> prompt muon ( $\text{NP}\mu$ ) are treated differently than those with a not prompt electron ( $\text{NPe}$ ). For  
<sup>1479</sup> muons, the dominant source is the semi-leptonic decay of heavy flavour hadrons, while for  
<sup>1480</sup> electrons, the dominant sources are hadrons and photon conversions.

<sup>1481</sup> The backgrounds causing NPL contributions are mostly  $Z/\gamma^* + \text{jets}$ (Drell–Yan) and  $t\bar{t} + \text{jets}$   
<sup>1482</sup> dilepton processes, and in a smaller amount  $WW$  processes. All of these backgrounds contain  
<sup>1483</sup> two real leptons and one NPL. Due to the fact that the probability for a lepton to be a NPL is  
<sup>1484</sup> small, backgrounds containing two or more NPL s are neglected in thus search. The assumption  
<sup>1485</sup> is made that for  $Z/\gamma^* + \text{jets}$  the two leptons compatible with a  $Z$  boson decay are the real leptons,  
<sup>1486</sup> and the additional lepton is coming from a NPL source, while for  $t\bar{t} + \text{jets}$  the NPL is associated  
<sup>1487</sup> to the  $Z$  boson. This has been validated using Monte Carlo simulation.

<sup>1488</sup> The NPL sample is constructed from data by requiring exactly three leptons, from which two  
<sup>1489</sup> are considered real, isolated leptons and the third lepton is identified as a NPL. This NPL is  
<sup>1490</sup> created by loosening its identification and inverting its isolation criteria. The full requirements  
<sup>1491</sup> on the not prompt leptons are given in [Table 5.7](#) and [Table 5.8](#). For NPe, a large fraction  
<sup>1492</sup> is coming from misidentified photons. These are removed by applying a tighter cut on the  
<sup>1493</sup>  $1/E - 1/p$  variable, and by limiting the isolation values to be smaller than one.

**Table 5.7:** Non prompt electron requirements used in this analysis. The requirements are set in the barrel ( $|\eta_{supercluster}| \leq 1.479$ ) and the end caps ( $|\eta_{supercluster}| > 1.479$ ).

Properties	$ \eta_{supercluster}  \leq 1.479$	$ \eta_{supercluster}  > 1.479$
$\sigma_{\eta\eta}$	< 0.011	< 0.0314
$ \Delta\eta_{in} $	< 0.00477	< 0.00868
$ \Delta\phi_{in} $	< 0.222	< 0.212
H/E	< 0.298	< 0.101
relative isolation	$\geq 0.0588 \ \&\& < 1$	$\geq 0.0571 \ \&\& < 1$
$ 1/E - 1/p $	$< 0.0129 \text{ GeV}^{-1}$	$< 0.0129 \text{ GeV}^{-1}$
expected missing inner hits	$\leq 1$	$\leq 1$
pass conversion veto	Y	Y
$p_T$	$> 35 \text{ GeV}$	$> 35 \text{ GeV}$

**Table 5.8:** Non prompt muon requirements used in the analysis.

Properties	modified Loose Muon WP
Global muon or Tracker Muon	Both
Particle Flow muon	Y
$\chi^2/ndof$ of global muon track fit	N/A
Nb. of hit muon chambers	N/A
Nb. of muon stations contained in the segment	N/A
Size of the transverse impact parameter of the track wrt. PV	N/A
Longitudinal distance wrt. PV	N/A
Nb. of pixel hits	N/A
Nb. of tracker layers with hits	N/A
Relative Isolation	$\geq 0.15$
$p_T$	$> 30 \text{ GeV}$

<sup>1494</sup> **5.4 Regions and channels**

The regions are defined as in Table 5.9 after a common selection of exactly three leptons containing one opposite sign, same flavour pair that is assigned to the Z boson, at least one jet and at the most three jets, the transverse mass of the W boson to be maximal 300 GeV. The transverse mass  $m_T(W)$  is reconstructed using

$$m_T(W) = \sqrt{(p_T(l_W) + p_T(v_W))^2 - (p_x(l_W) + p_x(v_W))^2 - (p_y(l_W) + p_y(v_W))^2} \quad (5.7)$$

**Table 5.9:** The statistically independent regions used in the analysis.

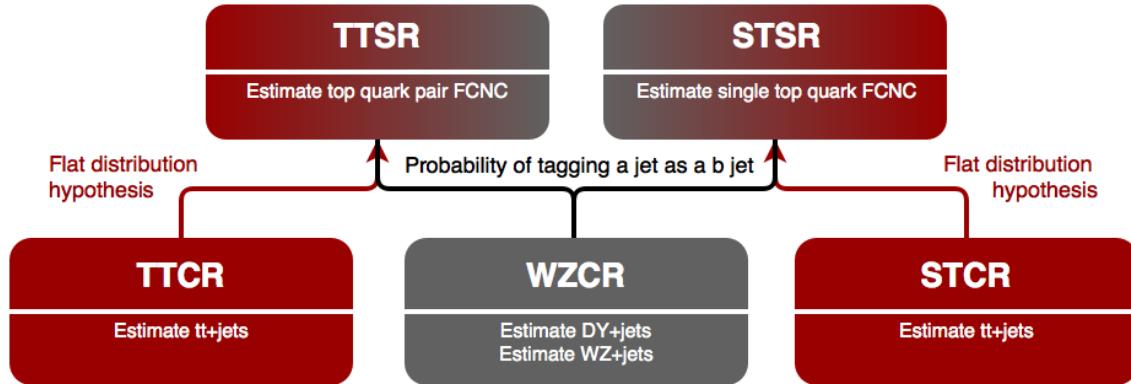
	WZCR	STS R	TTSR	STCR	TTCR
Number of jets	$\geq 1$	1	$\geq 2$	1	$\geq 2$
Number of b jets	0	1	$\geq 1$	1	$\geq 1$
$ m_Z^{\text{reco}} - m_Z  < 7.5 \text{ GeV}$	Yes	Yes	Yes	No	No
$ m_Z^{\text{reco}} - m_Z  < 30 \text{ GeV}$	Yes	Yes	Yes	Yes	Yes
Number of leptons	3	3	3	3	3

<sup>1495</sup> Additional leptons with a looser identification are vetoed in order to reduce the contamination  
<sup>1496</sup> of backgrounds with four or more leptons in the final state, e.g. ZZ, t̄Z, and t̄H. The most  
<sup>1497</sup> important backgrounds are the ones that contain three prompt leptons in the final state. These  
<sup>1498</sup> are mainly WZ+jets, t̄Z and SM tZq. For these backgrounds, the three lepton topology is  
<sup>1499</sup> identical to the FCNC signal: two opposite sign leptons of the same flavour decaying from the Z  
<sup>1500</sup> boson, and a third additional, high  $p_T$  lepton coming from the W boson decay.

<sup>1501</sup> For the FCNC tZ final state, one b jet coming from the SM top decay is expected. For the  
<sup>1502</sup> FCNC tZq, an additional light jet is expected. In the t̄Z final state, two b jets are present in the  
<sup>1503</sup> final state. However, due to inefficiencies of the b-tagging algorithm, one of the two b jets may  
<sup>1504</sup> be identified as a light quark jet, giving the same final state as the FCNC tZq final state. For  
<sup>1505</sup> the WZ+jets final states, one of the b jets produced by gluon splitting can be b-tagged or light  
<sup>1506</sup> flavour jets coming from the WZ+jets production can be mis-tagged as b jets. The SM tZq final  
<sup>1507</sup> state expects the same signal as FCNC tZq.

<sup>1508</sup> The NPL events give a significant background contribution. This background is coming mainly  
<sup>1509</sup> from Z/ $\gamma^*$  + jets and t̄t + jets processes (in a less significant way, also WW and tWZ contributes),  
<sup>1510</sup> which have very high cross sections and causes a large number of NPL background events  
<sup>1511</sup> compared to signal.

<sup>1512</sup> In order to reduce the large uncertainties in backgrounds, five independent regions are used  
<sup>1513</sup> as defined in Table 5.9. In Figure 5.12, the strategy and usage of each region is illustrated.



**Figure 5.12:** The strategy used for this search. The WZCR region is used to estimate the WZ+jets background process as well as the NPL background coming from the  $Z/\gamma^* + \text{jets}$  process. The TTCR and STCR regions are used to estimate the contributions of the NPL background coming from the  $t\bar{t} + \text{jets}$  process.

#### 1514 5.4.1 WZCR

1515 The WZCR is constructed by vetoing events with jets tagged as being a b-jet, making it statistically  
 1516 independent from the signal regions where at least one b-tagged jet is required. In this region,  
 1517 a fit is performed on the transverse mass of the W boson, in order to estimate the NPL yield  
 1518 coming from  $Z/\gamma^* + \text{jets}$  and the WZ+jets backgrounds.

A transfer factor is used to account for going from a region without b-tagged jets to a region with exactly or at least one b-tagged jet. For this the probability of tagging at least one jet with the CSVv2 algorithm at the loose working point is used to calculate the expected number of events,  $N_b$ , after b-tagging:

$$N_b = \frac{\sum_{\text{events}} \mathcal{P}_b}{\text{total nb of events}}, \quad (5.8)$$

where  $\mathcal{P}_b$  is the probability that an event survives the b-tagging requirement

$$\begin{aligned} \mathcal{P}_b &= 1 - P(\text{event doesn't survive b tag}) \\ &= 1 - \left( \prod_b P(\text{b not tagged}) \prod_c P(\text{c not tagged}) \prod_{uds} P(\text{light not tagged}) \right) \end{aligned} \quad (5.9)$$

1519 with the products going over all b-, c-, and light jets. The jet flavour is determined by means of  
 1520 matching the reconstructed jet to the generated quark based on the distance in the  $\eta\phi$  plane.  
 1521 In order to estimate the probability for exactly one b-tagged jet, the expected number of events  
 1522 is corrected by the fraction of events with exactly one jet in the WZCR. The resulting transfer  
 1523 factors are given in [Appendix C](#). The yield of WZ+jets events in the signal region estimated  
 1524 using the above described transfer factor and the yield calculated with simulated events are in  
 1525 agreement.

1526 **5.4.2 TTCR and STCR**

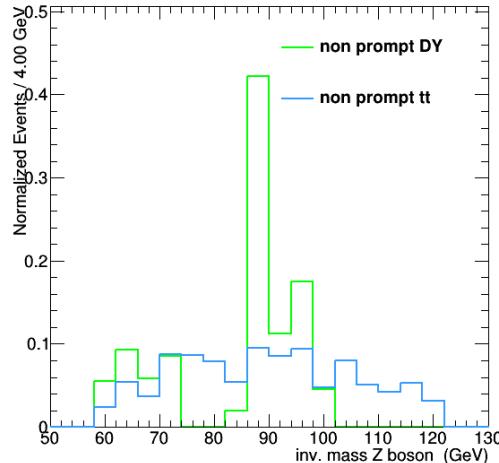
The TTCR and STCR have are constructed with the same selection criteria as TTSR and STSR but are outside the Z mass window (sidebands):

$$7.5 \text{ GeV} < |m_Z^{\text{reco}} - m_Z| < 30 \text{ GeV}. \quad (5.10)$$

where  $m_Z^{\text{reco}}$  is the reconstructed mass of the Z boson in the event, and  $m_Z$  the mass of the Z boson. These regions are dominated by  $t\bar{t}$ +jets (see Appendix C) and are used to estimate the NPL coming from  $t\bar{t}$ +jets in the STSR and TTSR. Since there aren't enough events entering these regions, no shapes are used in the fit. The distribution of the mass of the Z boson is flat for  $t\bar{t}$ +jets events, as shown in Fig. Figure 5.13, and thus the number of expected events,  $N_s$ , in the signal regions estimated from the number of expected events,  $N_c$ , in the control region is obtained as

$$N_s = \frac{15}{60 - 15} N_c. \quad (5.11)$$

1527 The resulting transfer factors are given in Appendix C. The expected yield in the signal region  
1528 estimated from the TTCR (STCR) is in agreement with the yield calculated from simulated events.



**Figure 5.13:** The normalized distribution for  $Z/\gamma^* + \text{jets}$  and  $t\bar{t} + \text{jets}$  events before dividing the events in to regions, after  $|m_Z^{\text{reco}} - m_Z| < 30 \text{ GeV}$ . All leptonic channels combined.

1529

1530 **5.4.3 TTSR and STSR**

1531 The TTSR is defined to target top quark pair FCNC ( $tZq$ ), while the STSR focusses on single top  
1532 quark FCNC ( $tZ$ ). They have NPL contributions coming from  $Z/\gamma^* + \text{jets}$  and  $t\bar{t} + \text{jets}$  events. In  
1533 this region, the data driven NPL template is split into two templates based on the presence of the  
1534 NPL in the Z boson. The NPL associated with W boson is assigned to  $Z/\gamma^* + \text{jets}$  and estimated  
1535 in the WZCR, while the NPL associated with Z boson is assigned to  $t\bar{t} + \text{jets}$  and estimated in  
1536 the TTCR and STCR.



# The search for FCNC involving a top quark and a Z boson

6

1537

## 1538 6.1 Construction of template distributions

1539 There were no selection criteria found to make a clear rejection of the background events without  
1540 sacrificing a significant amount of signal. For this reason, a multivariate approach using Boosted  
1541 Decision Trees that combines several discriminating variables in the TMVA framework is used.  
1542 For the training, the BDTs are trained against all backgrounds, where the NPL background is not  
1543 taken into account. The BDT settings avoid over-training and maintain a good discriminating  
1544 power against all backgrounds. The background and signal yields follow the relative fractions  
1545 predicted by the simulation.

1546 The variables used to construct the BDTs include the angles, distances, masses and transverse  
1547 momenta:

- 1548 1. pseudo rapidity of the SM top: TT+ST tZu, ST tZc
- 1549 2. invariant mass of the W lepton and the SM b jet: TT+ST tZu, TT+ST tZc
- 1550 3.  $\Delta\Phi$  between the W lepton and the SM b jet: TT+ST tZu, ST tZc
- 1551 4. minimal  $\Delta R$  between the W lepton and jets: TT tZu
- 1552 5. invariant mass of the Z boson: TT tZu, TT tZc
- 1553 6.  $\Delta\Phi$  between the W lepton and the Z boson: TT+ST tZu, TT tZc
- 1554 7.  $\Delta R$  between the W lepton and the SM b jet: TT+ST tZu, TT tZc
- 1555 8. number of CSVv2 medium WP b jets: TT tZu, TT tZc
- 1556 9. invariant mass of the FCNC top: TT tZu, TT tZc
- 1557 10.  $\Delta R$  between the Z boson and the FCNC light jet: TT tZu, TT tZc
- 1558 11.  $\Delta R$  between the FCNC light jet and the SM b jet: TT tZu, TT tZc

1559 12. charge of the W lepton times the absolute pseudo rapidity of the W lepton: ST tZu

1560 13. b discriminant of the highest  $p_T$  jet: ST tZu, ST tZc

1561 14. total Ht of the leptons: ST tZu

1562 15. the  $p_T$  of the W lepton times its charge: ST tZu

1563 16. total invariant mass of the leptons: ST tZc

1564 17.  $\Delta R$  between the W lepton and Z boson: TT+ST tZc

1565 18. total invariant mass of the event: TT tZc

1566 19. b discriminant of the FCNC light jet: TT tZc

1567 20.  $\Delta R$  between the SM b jet and Z boson: TT tZc

1568 The pre fit distributions of the variables used for creating the multivariate discriminator are  
 1569 given in [Section 6.1.1](#). The resulting multivariate discriminator are shown in [Section 6.1.2](#).

### 1570 6.1.1 Distributions of the BDT variables

**NOTE:** include distributions of the vars

### 1572 6.1.2 BDTs

**NOTE:** better to do normalised?

**Figure 6.1:** Distributions of the discriminating variable before the fit, all different leptonic channels together. Upper left: TTSR tZu, upper right: TTSR tZc; lower left: STSR tZu, lower right: STSR tZc.

1573

**Figure 6.2:** Distributions of the discriminating variable before the fit, 3 $\mu$  channel. Upper left: TTSR tZu, upper right: TTSR tZc; lower left: STSR tZu, lower right: STSR tZc.

**Figure 6.3:** Distributions of the discriminating variable before the fit, 1e2 $\mu$  channel. Upper left: TTSR tZu, upper right: TTSR tZc; lower left: STSR tZu, lower right: STSR tZc.

**Figure 6.4:** Distributions of the discriminating variable before the fit, 2e1 $\mu$  channel. Upper left: TTSR tZu, upper right: TTSR tZc; lower left: STSR tZu, lower right: STSR tZc.

### 1574 6.1.3 Transverse mass in WZCR

1575 The WZCR is used to estimate the contribution from WZ+jets and NPL background. In this  
 1576 region, a fit is performed on the transverse mass distribution of the W boson. The pre-fit  
 1577 templates are given in [Figure 6.6](#).

**Figure 6.5:** Distributions of the discriminating variable before the fit, 3e channel. Upper left: TTSR tZ<sub>u</sub>, upper right: TTSR tZ<sub>c</sub>; lower left: STSR tZ<sub>u</sub>, lower right: STSR tZ<sub>c</sub>.

**Figure 6.6:** The transverse mass of the W boson in the WZCR, before the fit. All different leptonic channels together. Left: scaled to the data, right: normalized.

## 1578 6.2 Systematic uncertainties

1579 The systematic uncertainties entering the analysis are coming from different sources. The  
 1580 experimental uncertainties arise from the reconstruction of the objects and are discussed in  
 1581 [Section ??](#). These influence the number of events passing the selection, so-called normalisation  
 1582 uncertainties, or the relative occupancies of the distributions, so-called shape uncertainties. The  
 1583 normalisation uncertainties coming from reconstruction include the uncertainty of 2.5% on  
 1584 the measured integrated luminosity and the efficiency of the trigger logic used for the analysis  
 1585 which has a 1% (5%) uncertainty on the 3μ and 1e2μ(2e1μ and 3e) channels. The pileup  
 1586 distribution is calculated via the minimum bias cross section which has a 4.6% uncertainty.  
 1587 This uncertainty results in a systematic shift in the pileup distribution and its shape effect is  
 1588 estimated by recalculating the pileup distribution for each variation of the minimum bias cross  
 1589 section. The effect of the systematic upwards and downwards shift on the pileup distribution is  
 1590 demonstrated in [Section ??](#). The shape uncertainties also include the uncertainties coming from  
 1591 the applied lepton scale factors. Their systematic uncertainty originates from three sources:  
 1592 identification, isolation and tracking. The effect of systematic upwards or downwards shift  
 1593 on shapes is demonstrated in [Section ??](#) and [Section ??](#). The uncertainties arising from jet  
 1594 energy corrections require a recalculation of all jet related kinematic observables and its effect  
 1595 is propagated to the missing transverse energy. The resulting effect of the systematic upwards  
 1596 or downwards shift on shapes is demonstrated in [Section ??](#) and [Section ??](#). The reweighting of  
 1597 the CSVv2 discriminant is also a source of uncertainty. There are three sources of uncertainty  
 1598 contributing to the measurement of the b-tag related scale factors: statistical uncertainties, jet  
 1599 energy scale and the purity of the sample. These result in eight uncorrelated contributions for  
 1600 which the effects on the shapes are shown in [Section ??](#).

1601 Since the NPL sample is artificially made from data by inverting the isolation of the third  
 1602 lepton. Its effect has to be estimated. The shape uncertainty one the NPL processes is obtained  
 1603 by varying the isolation inversion with respect to tight working point to the loose working point  
 1604 for electrons and muons at the same time. This found to have negligible effect. The uncertainty  
 1605 on the normalisation of the overall NPL yield is taken as 50% in accordance the the SM tZq  
 1606 search.

**NOTE:** Add source

1607 The uncertainty on the expected yield of the simulated backgrounds is taken to be 30%  
 1608 of the yield such that it covers all uncertainties at next to leading order accuracy. Theory  
 1609 uncertainties originating from the modelling of the main backgrounds are estimated to account  
 1610 for the effect on the shape of the distributions from the choice of parton density funnctions, and  
 1611 renormalization ( $\mu_R$ ) and factorization ( $\mu_F$ ) scales. The effect of the renormalization ( $\mu_R$ ) and  
 1612 factorization ( $\mu_F$ ) scales is estimated by varying each independently and correlated up and  
 1613 down by a factor of two, where the anti-correlated variations are dropped. The envelope of  
 1614 these variations is used as an uncertainty. The uncertainties coming from the parton density

functions used for simulation are estimated using the PDF4LHC recipe [183], which combines the MMHT14, CT14, and NNPDF3.0 PDF sets [183]. The theory uncertainties are considered for the main backgrounds coming from simulation: WZ+jets, ZZ+jets, t̄Z, and tZq. This is found to have a negligible effect.

The way the uncertainties are treated as nuisance parameters is summarized in Table 6.1.

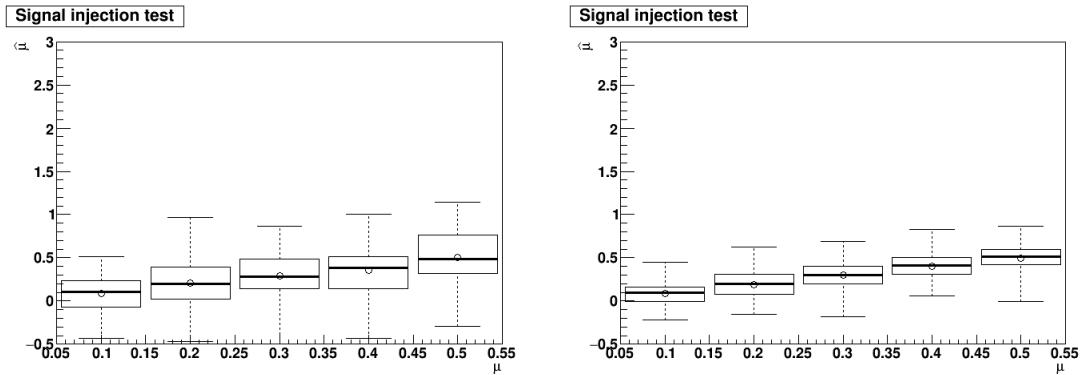
**Table 6.1:** Uncertainties used in this analysis. The column labelled type represents how the uncertainty is treated for the fit.

Source	Systematic input	Type
nonprompt muon norm.	50%	normalisation
nonprompt electron norm.	50%	normalisation
background t̄Z norm.	30%	normalisation
background WZ norm.	30%	normalisation
background tZq norm.	30%	normalisation
background ZZ norm.	30%	normalisation
background other MC norm.	30%	normalisation
trigger	1% (5%)	normalisation
lepton identification	$\pm\sigma(p_T, \eta)$	shape
JES	$\pm\sigma(p_T, \eta)$	shape
JER	$\pm\sigma(p_T, \eta)$	shape
b-tagging	$\pm\sigma(p_T, \eta)$	shape
pileup	$\pm\sigma$ of min. bias cross section	shape
PDF	PDF4LHC recipe	shape (WZ,tZq, ttZ, ZZ)
luminosity	2.5%	normalisation
renorm. and fact. scales	varying indep. and corr.	shape

### 6.3 Limit setting procedure validation

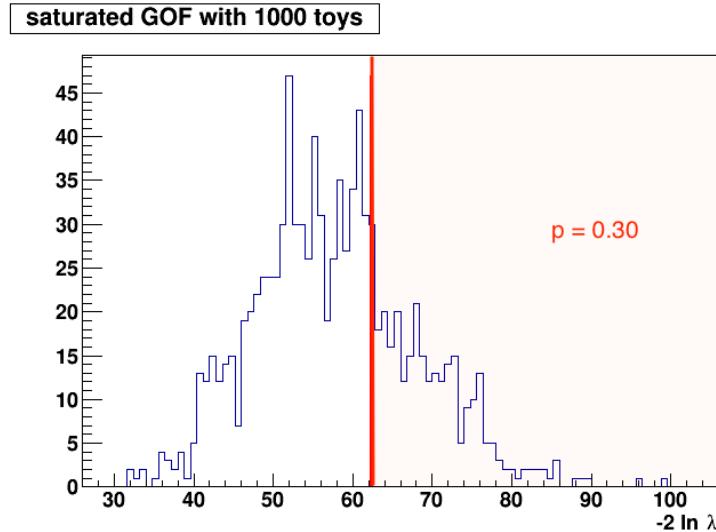
The analysis strategy has been established using a blinded strategy. Through the use of a pseudo dataset, the limit setting procedure has been validated. Signal injection tests for which the signal strength from a pseudo dataset with a pre-set signal strength is estimated are performed and shown in Figure 6.7.

Another validation has been done by performing a Maximum Likelihood fit in the WZCR only, considering all lepton channels. A simultaneous fit of the signal strength of the NPe, NPμ and

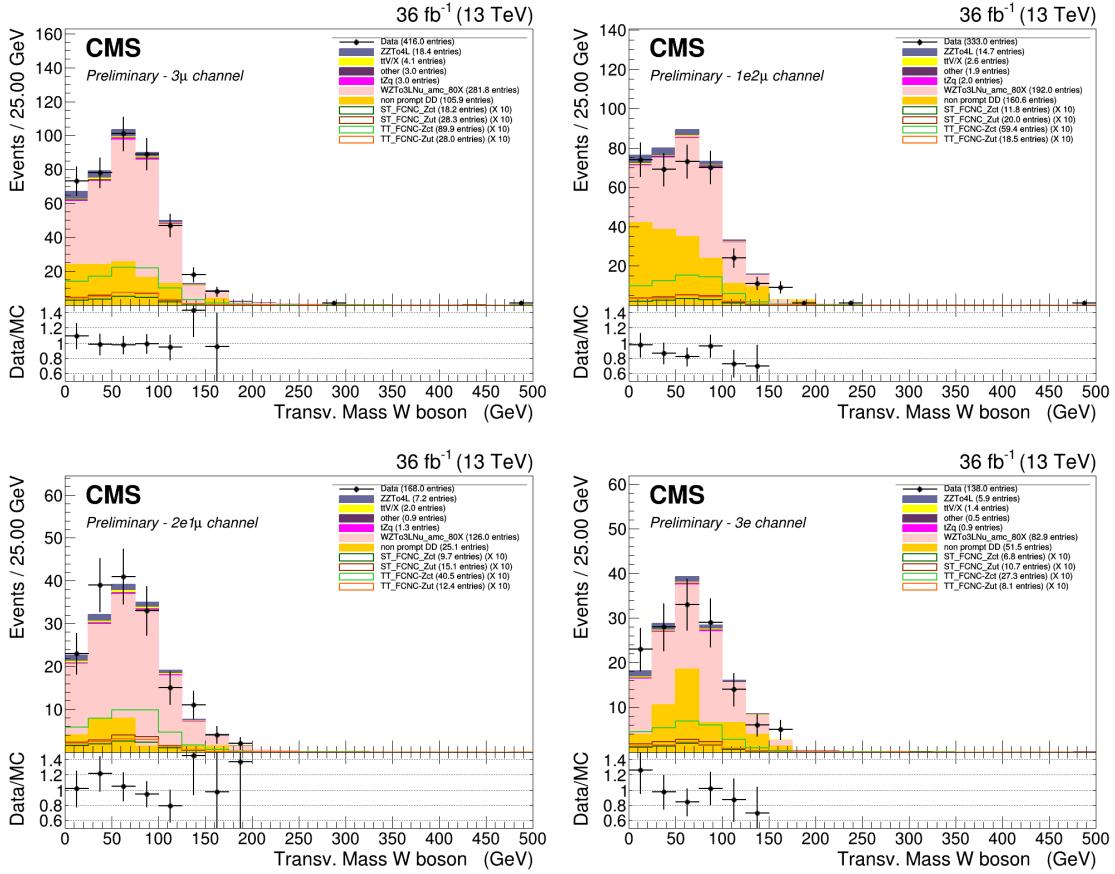


**Figure 6.7:** The obtained signal strength with the Maximum Likelihood method is in agreement with the signal strength used to generate the Asimov data set for the tZu (left) and tZc (right) couplings.

1627 the WZ+jets backgrounds is done by using the multi-dimensional fit in Higgs Combine Tool.  
 1628 The resulting signal strengths can then be applied on the distribution of the transverse mass of  
 1629 the W boson to verify data/MC agreement, as can be seen in Fig. 6.9. Furthermore, a goodness  
 1630 of fit test is performed, resulting in a p-value of 0.3 (see Fig. 6.8 ).



**Figure 6.8:** Goodness of fit testing in the WZCR with 1000 toys. The likelihood ration is denoted as  $\lambda$ .

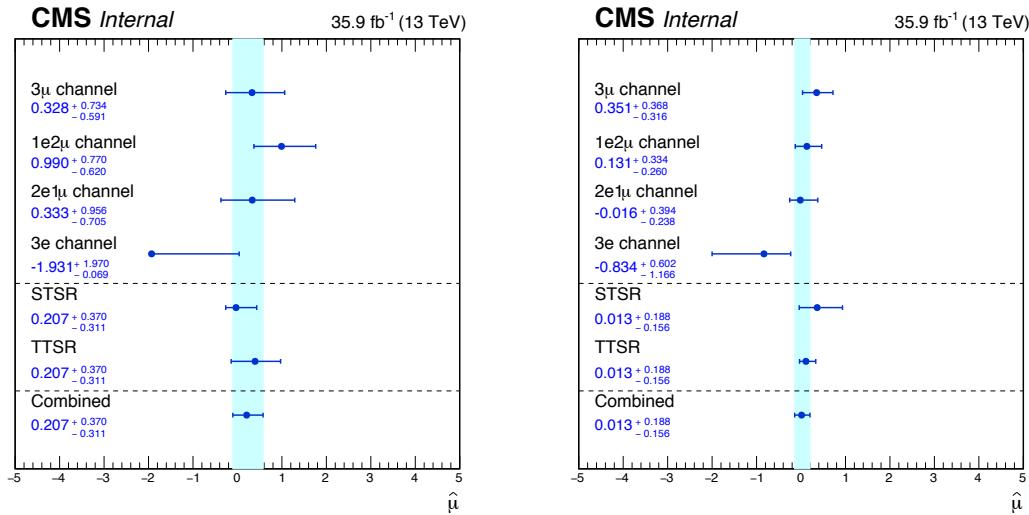


**Figure 6.9:** The transverse mass of the W boson in the WZCR for the 3 $\mu$  channel (left, upper), 1e2 $\mu$  channel (right, upper), 2e1 $\mu$  channel (left, lower), and 3 electrons channel (right, lower).

1631 Using the normalisations obtained by fitting the WZCR only, one can look at the data/MC  
 1632 agreement in the WZCR for all the variables used to create the BDT variables. These show good  
 1633 agreement as illustrated in Figure ??.

## 6.4 Results and discussion

The limit setting procedure explained in [Section 3.4](#) is applied and results are obtained for each lepton channel separately as well as the combination. For both the tZu and tZc coupling, the maximum likelihood estimator of their signal strengths  $\hat{\mu}$  is compatible with zero. This is shown in [Figure 6.10](#). One can see that the 3e leptonic channel has the largest uncertainty. This is due to the fact that this channel is the most influences by the lack of statistics for this search.



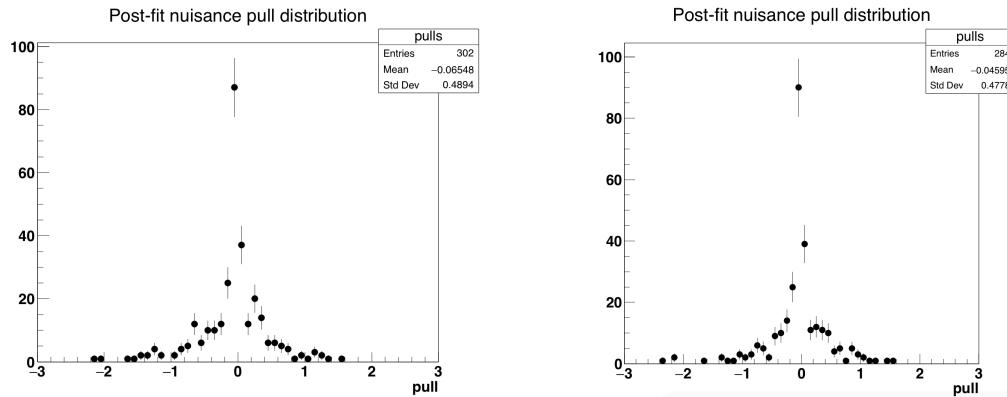
**Figure 6.10:** The maximum likelihood estimators for the signal strengths for the tZu vertex (right) and tZc vertex (left) per lepton channel as well as the combination in the STSR, TTSR, and all regions combined.

1639

The maximum likelihood estimators for the nuisance parameters  $\hat{\theta}$  are shown in [Figure 6.12](#) and [Figure 6.13](#). Their values obtained from the signal plus background or background only fits are in agreement. The transfer factors have an initial value different than one and have small uncertainties. The normalisation uncertainties on the yields of the simulated backgrounds get constrained by the fit. A well known verification of the error coverage for the fit is the pull distribution. For a random variable  $x$  following a Gaussian distribution of mean  $m$  and width  $w$ , the pull  $g$  is defined as

$$g = \frac{x - m}{w}, \quad (6.1)$$

and follows a standard Gaussian with mean zero and unit width. This property can be applied to our case of parameter estimation due to the central limit theorem [184]. The pull distributions are shown in [Figure 6.11](#). These are peaking at zero with tails going to  $\pm 2.5$ .

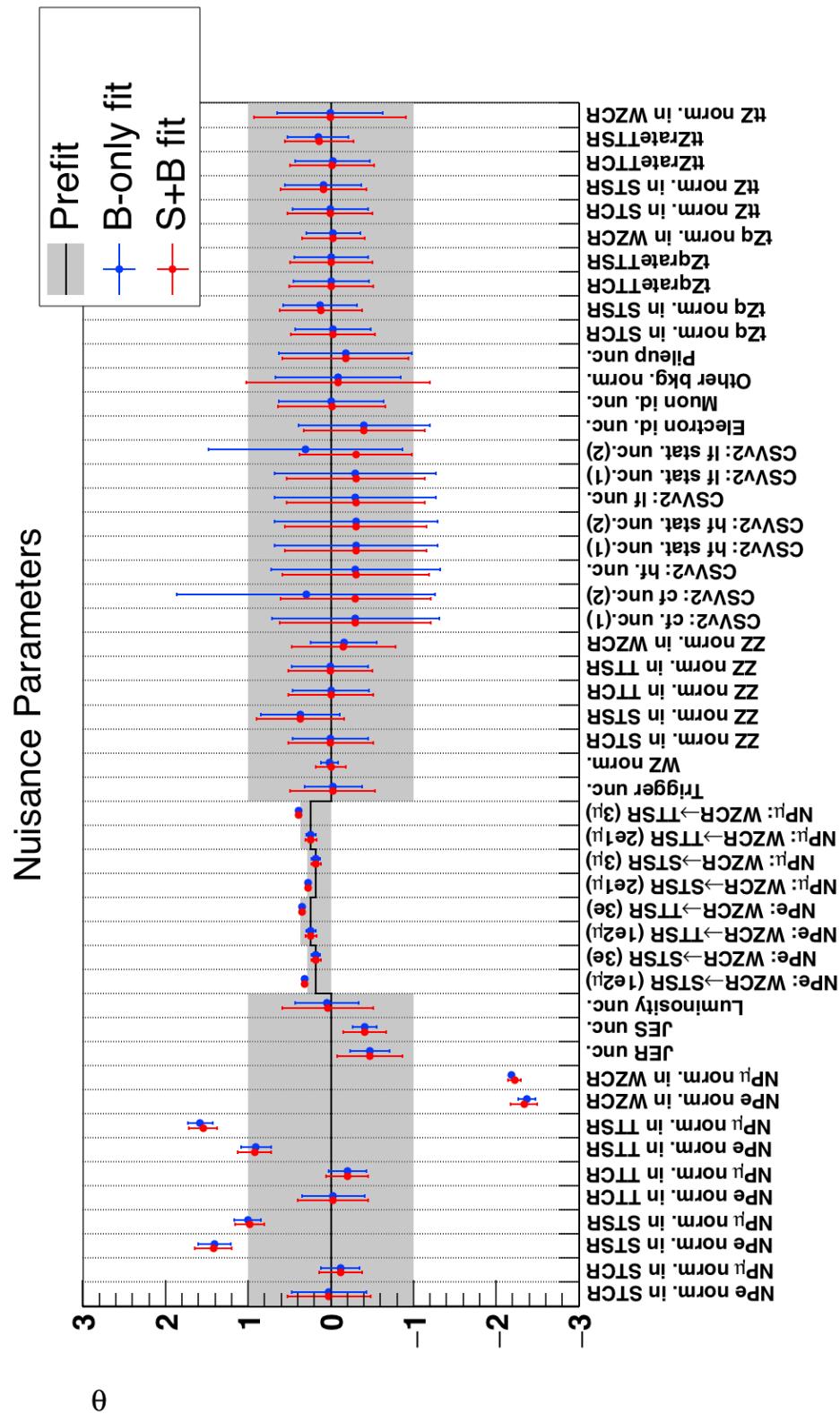


**Figure 6.11:** Post-fit distribution of the pulls for the tZu (left) and tZc (right) vertex.

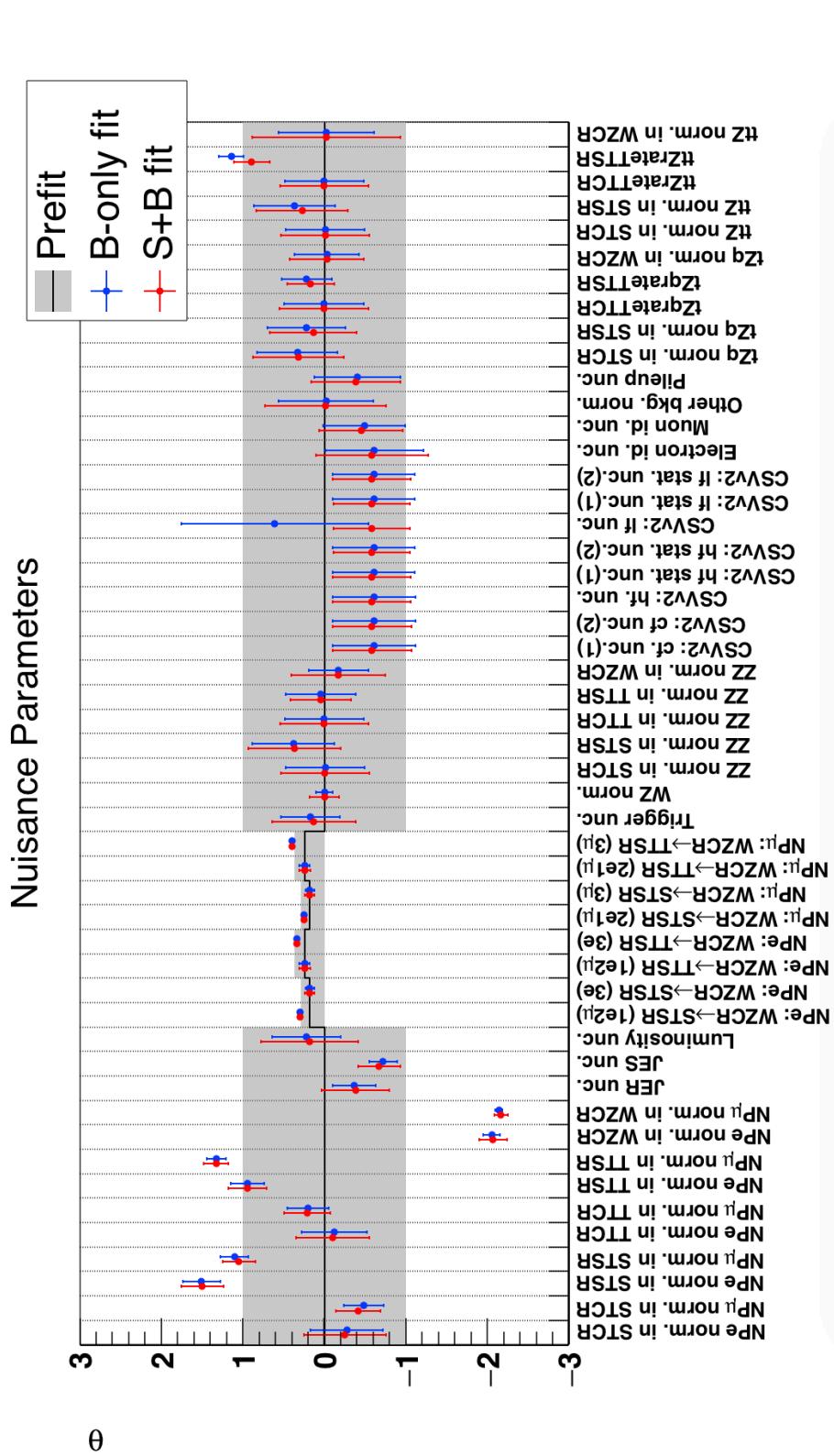
1643 The nuisance parameters causing the tails in the pull distributions are the ones related to the  
 1644 NPL background as can be seen in Figure 6.14, Figure 6.15, Figure 6.16, and Figure 6.17. Here  
 1645 one can see that the nuisance parameters related to the NPL normalisations are shifted with  
 1646 respect to their initial values. This is to be expected since their initial normalisation is arbitrary.  
 1647 Furthermore, the effect of the uncertainties on the maximum likelihood estimate of the signal  
 strength is shown.

**NOTE:** Zeer  
welke sys  
de grootste  
invl hebbe

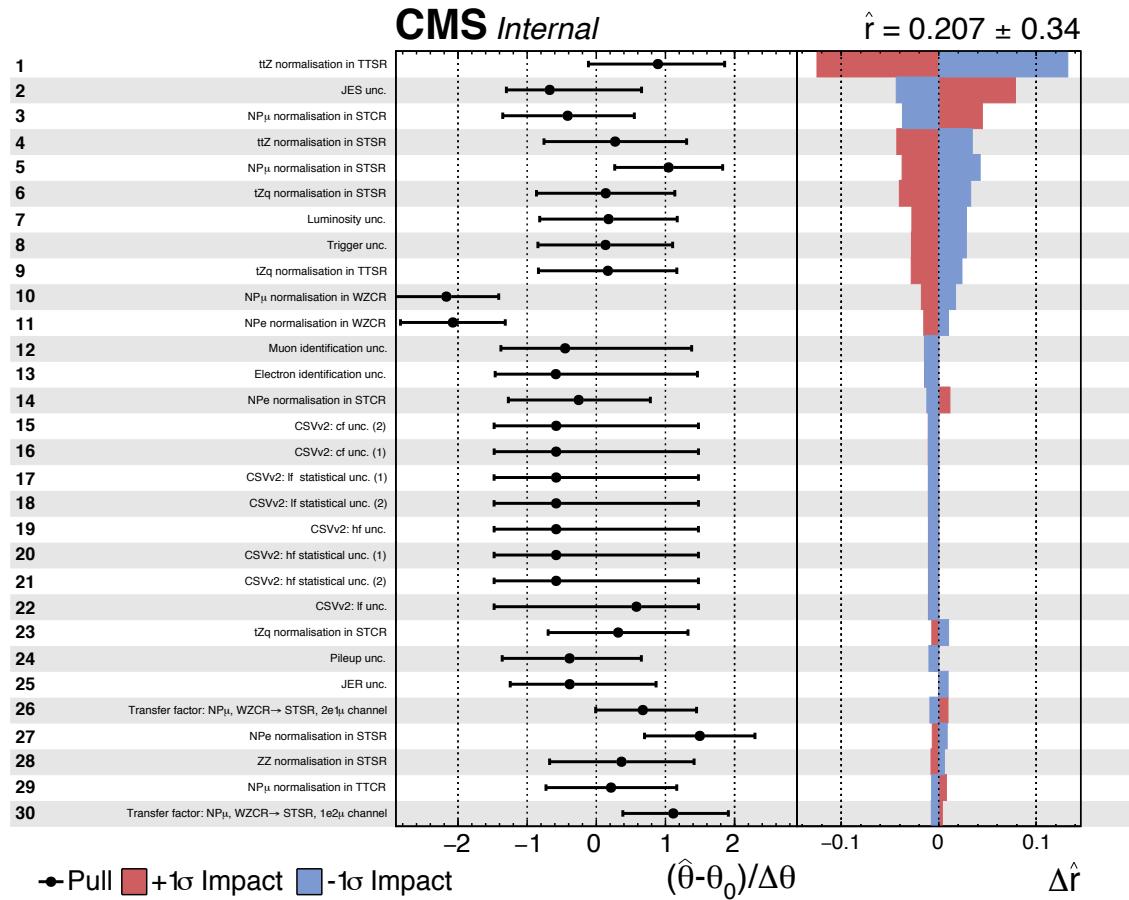
1648  
 1649  
 1650  
 1651  
 1652 The distributions of multivariate discriminating variables as well as the distribution of the  
 transverse mass of the W boson are recreated with the maximum likelihood estimations of the  
 nuisance parameters. The resulting distributions are shown in Figure 6.18, Figure 6.19, Figure  
 6.20, Figure 6.21, and Figure 6.22.



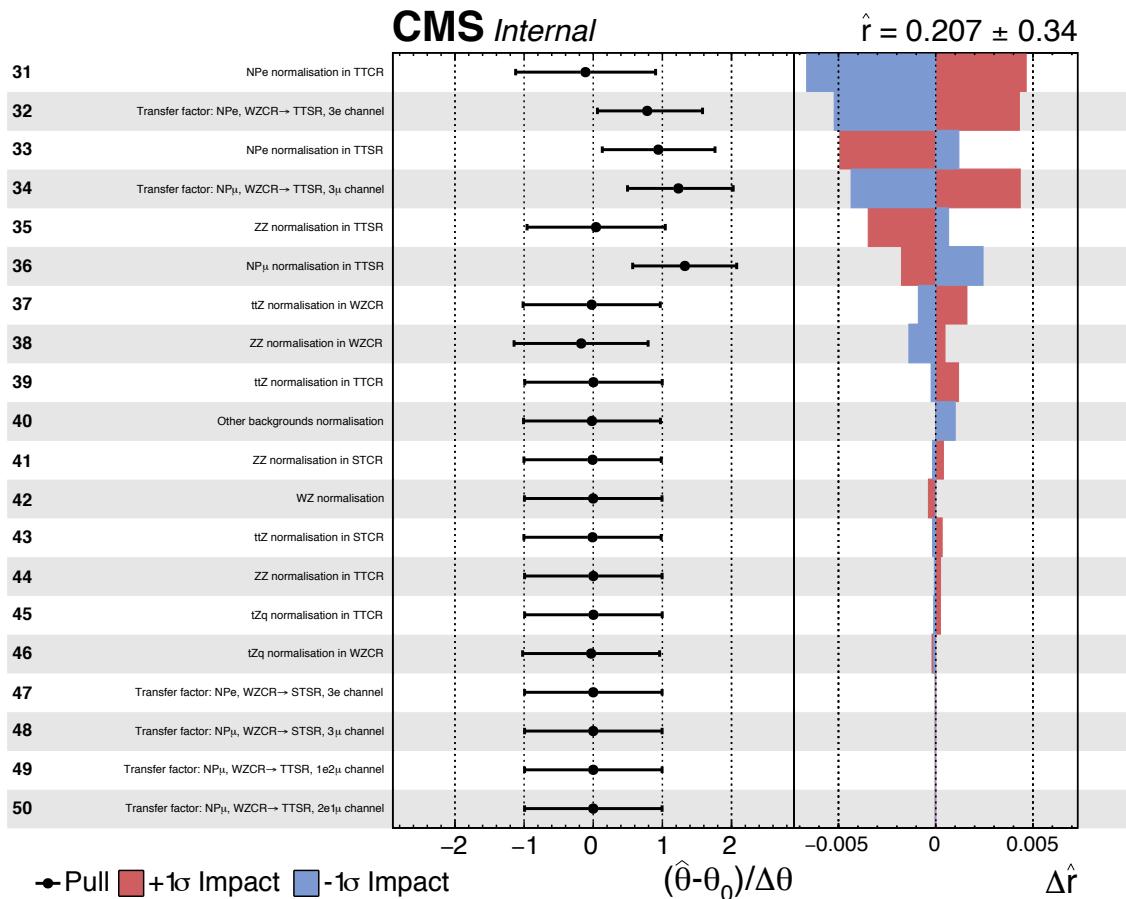
**Figure 6.12:** Maximum likelihood estimators for the tZc vertex.



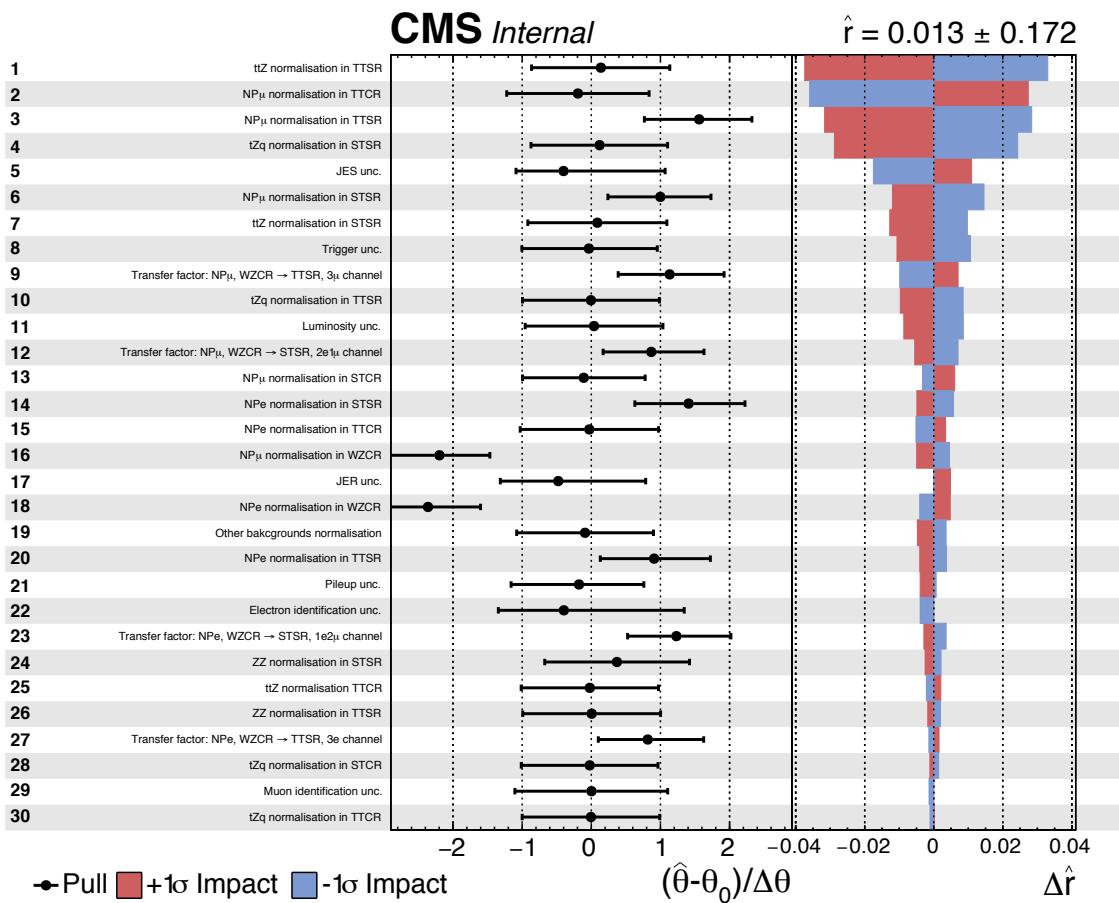
**Figure 6.13:** Maximum likelihood estimators for the tZu vertex.



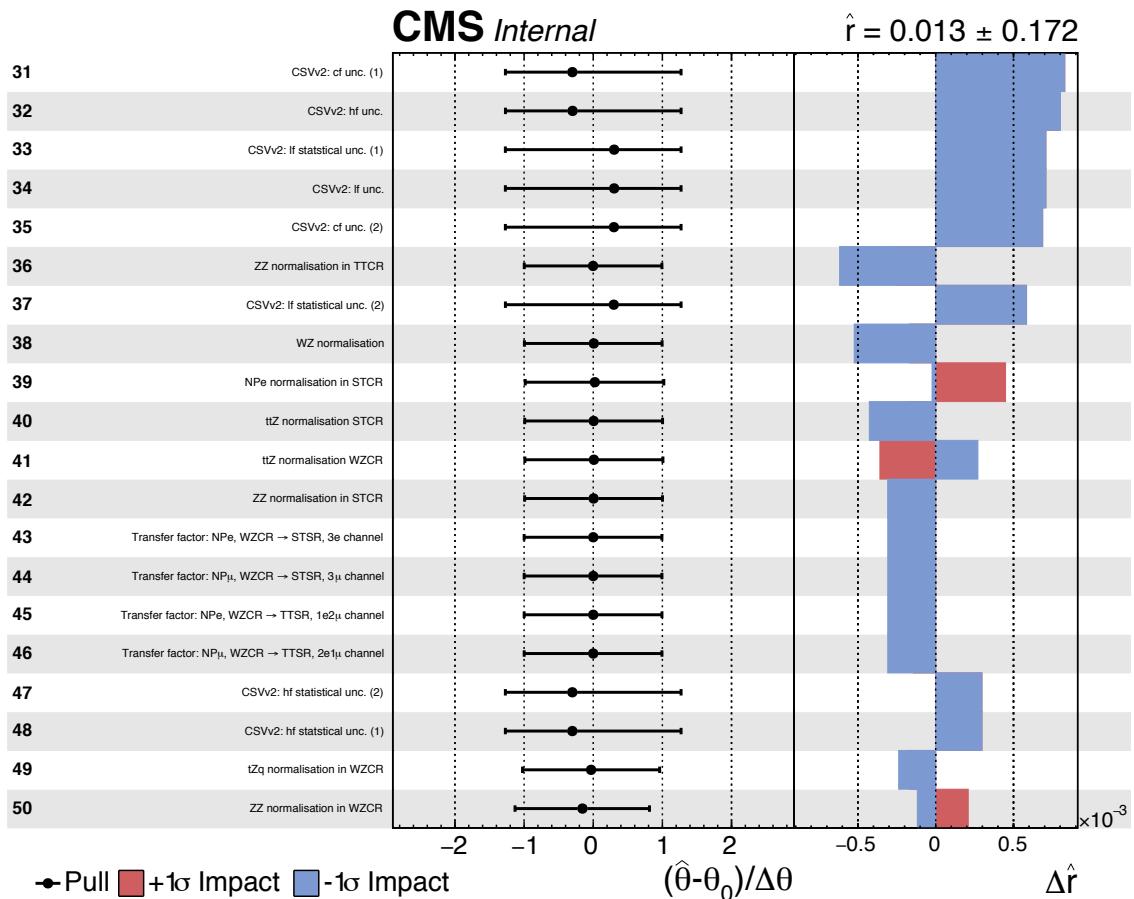
**Figure 6.14:** The pulls of each nuisance parameter and the influence of their uncertainty on the maximum likelihood estimation of the signal strength  $\hat{r}$  for the tZu vertex, part one.



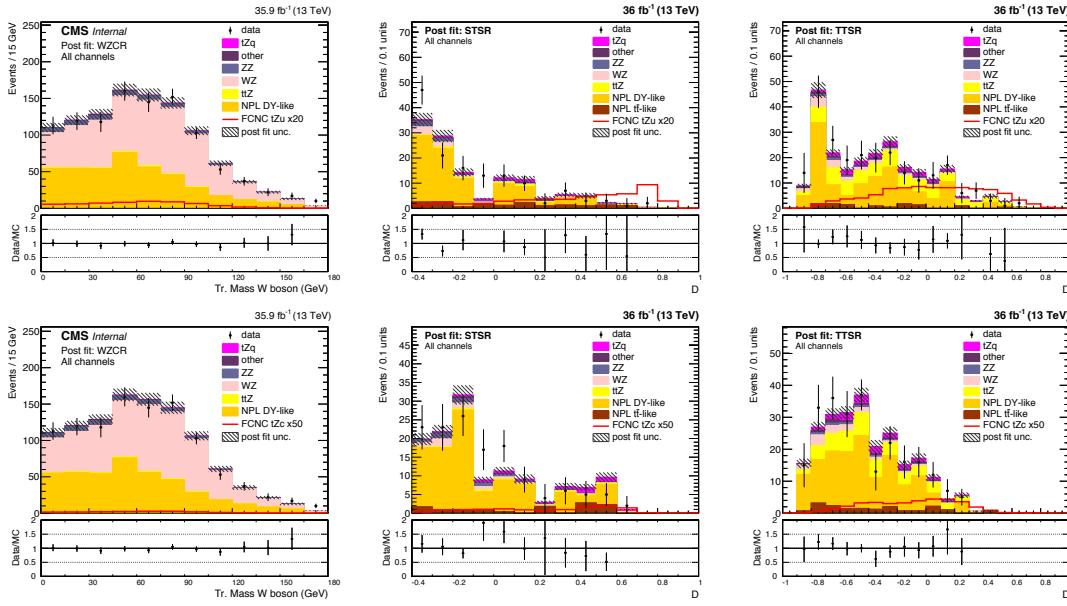
**Figure 6.15:** The pulls of each nuisance parameter and the influence of their uncertainty on the maximum likelihood estimation of the signal strength  $\hat{r}$  for the tZu vertex, part two.



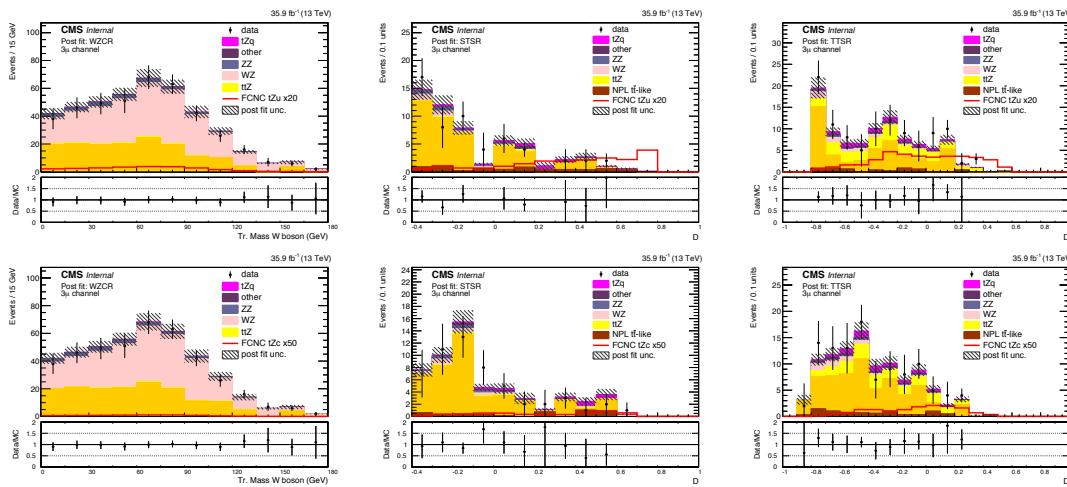
**Figure 6.16:** The pulls of each nuisance parameter and the influence of their uncertainty on the maximum likelihood estimation of the signal strength  $\hat{r}$  for the tZc vertex, part one.



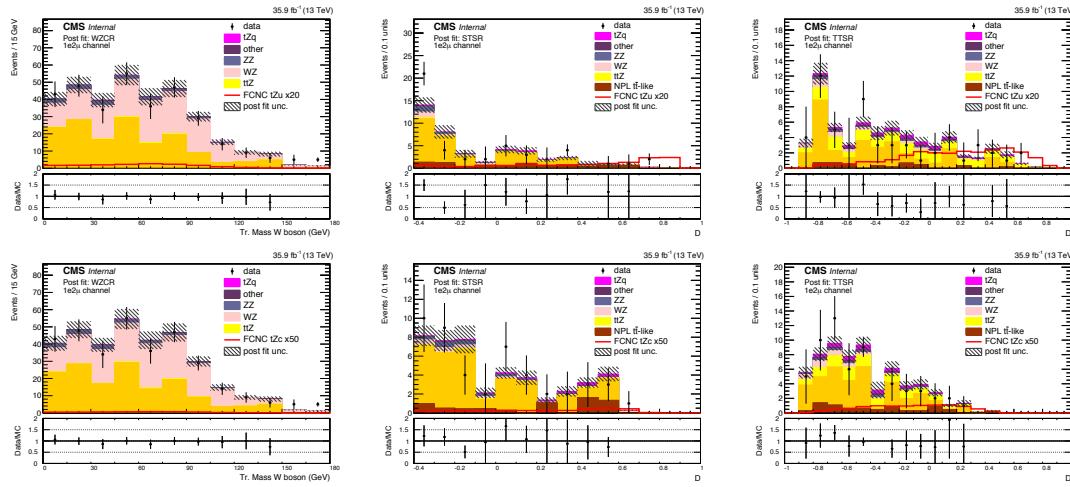
**Figure 6.17:** The pulls of each nuisance parameter and the influence of their uncertainty on the maximum likelihood estimation of the signal strength  $\hat{r}$  for the tZc vertexg, part two.



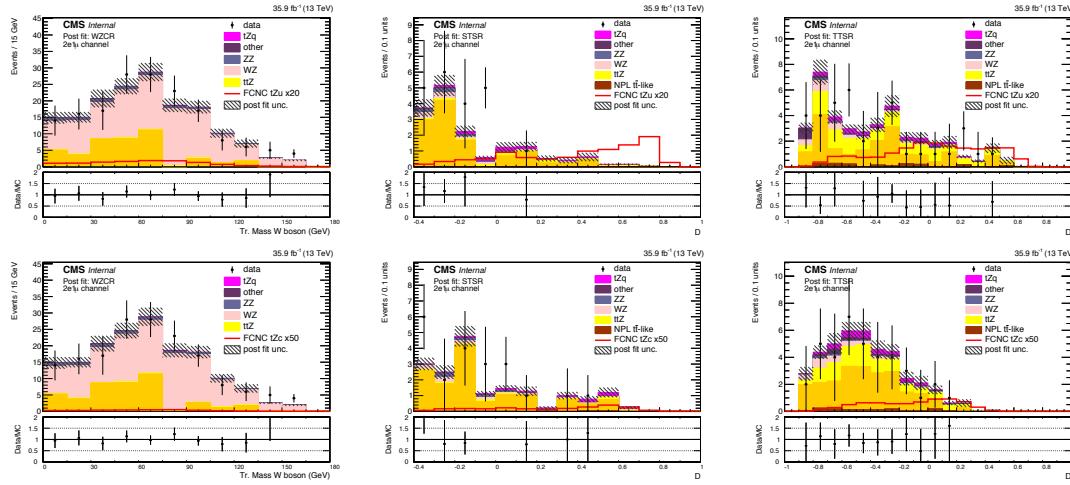
**Figure 6.18:** Post fit distributions for the all leptonic channels of the transverse mass of the W boson in the WZCR (left), the multivariate discriminating variable in the STSR (middle), and the multivariate discriminating variable in the TTSR (right) for the tZu (top) and tZc (bottom) couplings.



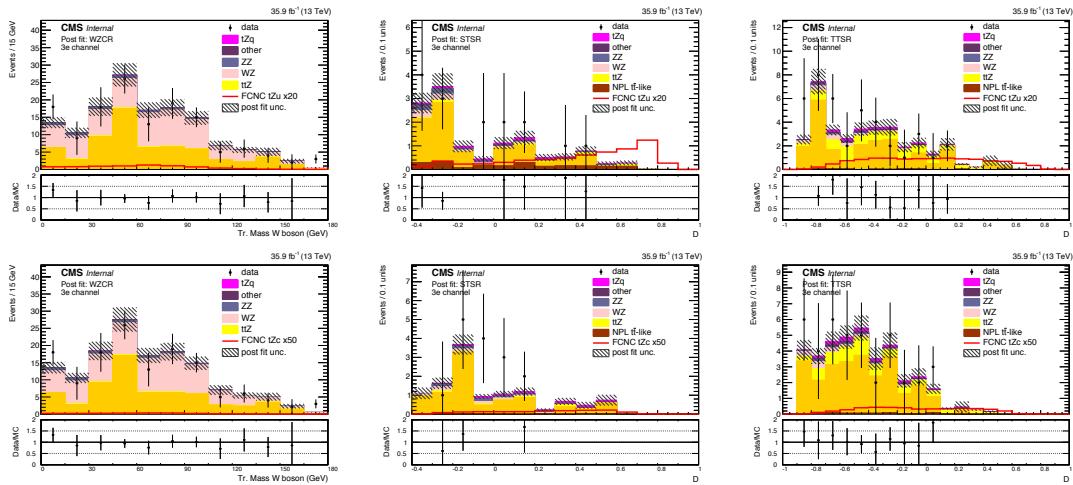
**Figure 6.19:** Post fit distributions for the  $3\mu$  leptonic channel of the transverse mass of the W boson in the WZCR (left), the multivariate discriminating variable in the STSR (middle), and the multivariate discriminating variable in the TTSR (right) for the tZu (top) and tZc (bottom) couplings.



**Figure 6.20:** Post fit distributions for the  $1e2\mu$  leptonic channel of the transverse mass of the W boson in the WZCR (left), the multivariate discriminating variable in the STSR (middle), and the multivariate discriminating variable in the TTSR (right) for the tZu (top) and tZc (bottom) couplings.



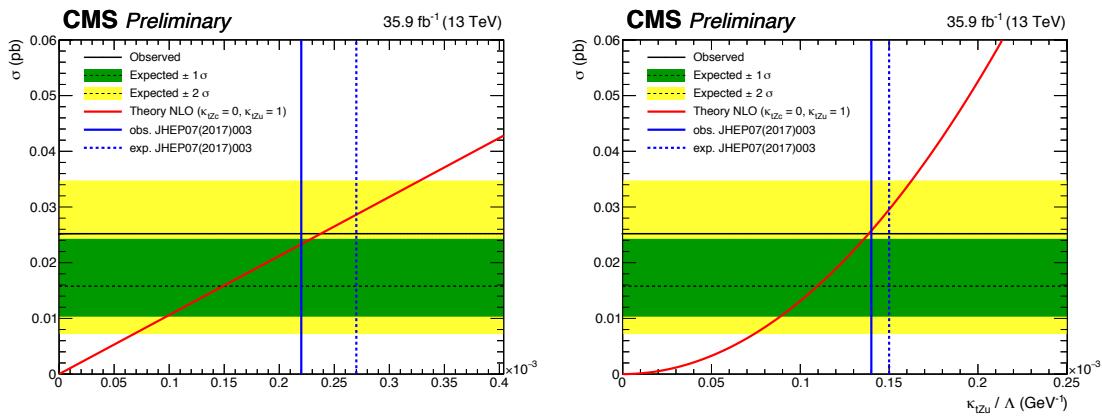
**Figure 6.21:** Post fit distributions for the  $2e1\mu$  leptonic channel of the transverse mass of the W boson in the WZCR (left), the multivariate discriminating variable in the STSR (middle), and the multivariate discriminating variable in the TTSR (right) for the tZu (top) and tZc (bottom) couplings.



**Figure 6.22:** Post fit distributions for the 3e leptonic channel of the transverse mass of the W boson in the WZCR (left), the multivariate discriminating variable in the STSR (middle), and the multivariate discriminating variable in the TTSR (right) for the tZu (top) and tZc (bottom) couplings.

1653 **6.4.1 One dimensional limits**

1654 The limit setting procedure used in this search returns limits on the signal strength modifier  
 1655 which can be translated to signal cross sections. These limits are translated to a limit on the  
 1656 branching fraction using [Equation 1.39](#). Additionally, the limit on the couplings are extracted  
 1657 using the fact that the cross sections are quadratically dependent on the couplings. In [Figure](#)  
 1658 [6.23](#), the resulting limits at 95% CL on the branching fraction and couplings related to the  
 1659 tZu vertex is shown. This observed (expected) limit amounts to  $\mathcal{B} < 0.024\%$  ( $0.015\%$ ) when  
 1660  $\kappa_{tZu}/\Lambda \neq 0$  and  $\kappa_{tZc}/\Lambda = 0$ . The expected limit surpasses the CMS search at a centre-of-mass of  
 1661 8 TeV expected limits of 0.027% [41]. The observed limit of 0.024% for the tZu interaction  
 1662 doesn't surpass the CMS 8 TeV observed limit of 0.022% [41]. The ATLAS collaboration has set  
 1663 limits 95% CL at a centre-of-mass of 13 TeV [40] with observed (expected) limits of 0.017%  
 1664 ( $0.024\%$ ) for the tZu coupling. The expected limit presented in this analysis surpasses the  
 expected limit for tZu.



**Figure 6.23:** Exclusion limits at 95% CL on the FCNC branching fractions (left) and couplings (right) as a function of the cross section of the FCNC process, considering only the tZu vertex. The CMS observed (expected) limit at 95% CL at a centre-of-mass of 8 TeV [41] is given with a blue line (dashed line).

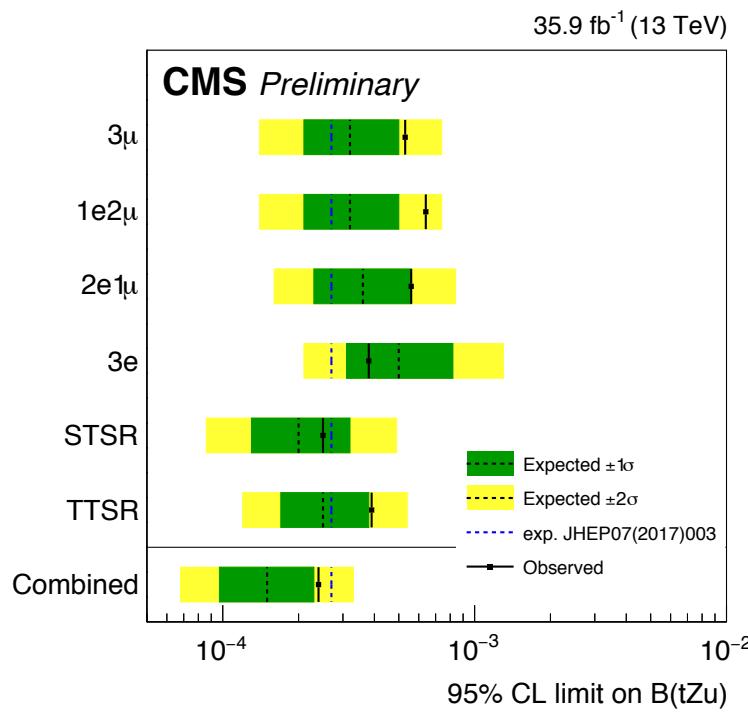
1665

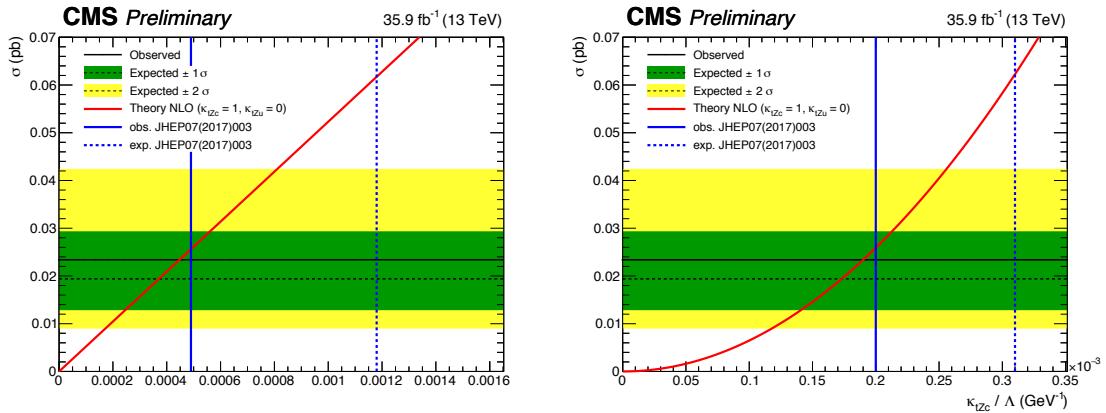
1666 In [Figure 6.24](#) and [Table 6.2](#), the limits for each leptonic channel separate as well as their  
 1667 combined limits are shown for the tZu vertex. The leptonic channels are in agreement with  
 1668 each other, where the presence of a muon helps pushing the limit further. The STSR is the  
 1669 most sensitive region because of the higher presence of the targeted single top quark signal.  
 1670 Further, one can see that by combining signle top quark and top quark pair signals, one gains  
 1671 in sensitivity.

1672 In [Figure 6.25](#), the resulting limits at 95% CL on the branching fraction and couplings  
 1673 related to the tZc vertex is shown. or this coupling, the observed (expected) limit at 95%  
 1674 CL is  $\mathcal{B} < 0.045\%$  ( $0.037\%$ ) when  $\kappa_{tZc}/\Lambda \neq 0$  and  $\kappa_{tZu}/\Lambda = 0$ . The expected limit surpasses  
 1675 the CMS search at a centre-of-mass of 8 TeV expected limits of 0.118% [41]. This also the  
 1676 case for the observed limit of 0.045% for the tZc interaction, which surpasses the CMS 8 TeV  
 1677 observed limit of 0.049% [41]. The observed (expected) limits set by the ATLAS collaboration  
 1678 at a centre-of-mass of 13 TeV [40] are 0.023% ( $0.032\%$ ) for the tZc coupling. The expected  
 1679 limit presented in this analysis is in accordance with this limit.

**Table 6.2:** Expected limits on the branching fractions at 95% CL for the tZu coupling [40, 41].

	expected	+2 $\sigma$	+1 $\sigma$	-1 $\sigma$	-2 $\sigma$	observed
3 $\mu$	0.032%	0.074%	0.050%	0.021%	0.014%	0.053%
1e2 $\mu$	0.032%	0.074%	0.050%	0.021%	0.014%	0.064%
2e1 $\mu$	0.036%	0.084%	0.056%	0.023%	0.016%	0.056%
3e	0.050%	0.13%	0.082%	0.031%	0.021%	0.038%
STSR only	0.020%	0.049%	0.032%	0.013%	0.0086%	0.025%
TTSR only	0.025%	0.054%	0.038%	0.025%	0.017%	0.039%
combined	0.015%	0.033%	0.023%	0.0097%	0.0068%	0.024%
8 TeV CMS ( $19.7 \text{ fb}^{-1}$ )	0.027%	-%	0.42%	0.018%	-%	0.022%
13 TeV ATLAS ( $36 \text{ fb}^{-1}$ )	0.024%	-%	0.35%	0.017%	-%	0.017%

**Figure 6.24:** Exclusion limits at 95% CL for each leptonic channel and signal region on the FCNC tZu branching fractions considering one non-vanishing coupling at a time. The CMS observed (expected) limit at 95% CL at a centre-of-mass of 8 TeV [41] is given with a blue line (dashed line).

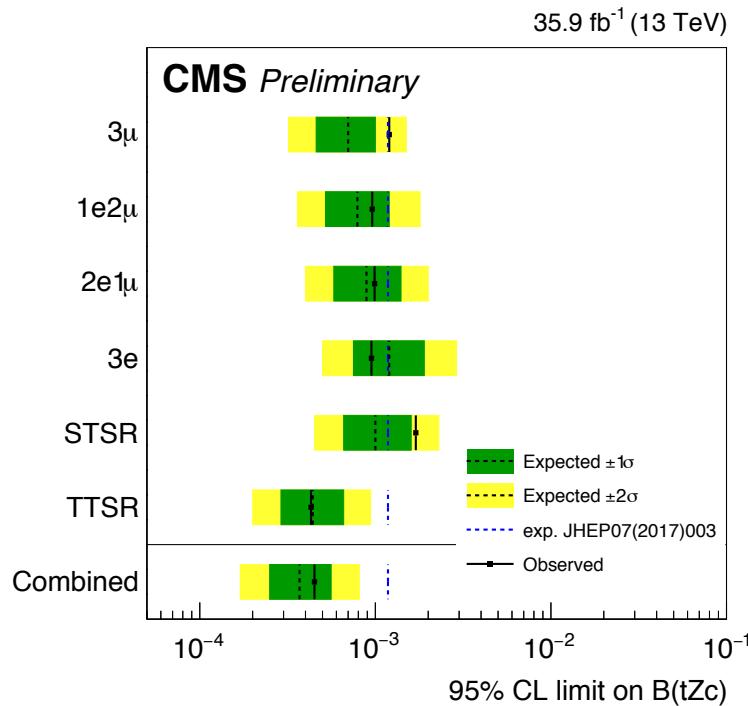


**Figure 6.25:** Exclusion limits at 95% CL on the FCNC branching fractions (left) and couplings (right) as a function of the cross section of the FCNC process, considering only the tZc vertex. The CMS observed (expected) limit at 95% CL at a centre-of-mass of 8 TeV [41] is given with a blue line (dashed line).

1680 The limits for each leptonic channel separate as well as their combined limits are shown for  
 1681 the tZc vertex in Figure 6.26 and Table 6.3. Also here, the leptonic channels are in agreement  
 1682 with each other, and the presence of a muon helps the sensitivity. For the tZu vertex, the TTSR  
 1683 is the most sensitive region and by combining single top quark and top quark pair signals, one  
 1684 also gains in sensitivity.

**Table 6.3:** Expected limits on the branching fractions at 95% CL for the tZc coupling [40, 41].

	expected	+2σ	+1σ	-1σ	-2σ	observed
3μ	0.070%	0.15%	0.10%	0.046%	0.032%	0.12%
1e2μ	0.079%	0.18%	0.12%	0.052%	0.036%	0.096%
2e1μ	0.089%	0.20%	0.14%	0.058%	0.040%	0.099%
3e	0.12%	0.29%	0.19%	0.075%	0.050%	0.095%
STSR only	0.10%	0.23%	0.16%	0.066%	0.045%	0.17%
TTSR only	0.044%	0.094%	0.066%	0.029%	0.020%	0.043%
combined	0.037%	0.081%	0.056%	0.025%	0.017%	0.045%
8 TeV CMS ( $19.7 \text{ fb}^{-1}$ )	0.118%	-%	0.222%	0.071%	-%	0.049%
13 TeV ATLAS ( $36 \text{ fb}^{-1}$ )	0.032%	-%	0.046%	0.022%	-%	0.023%



**Figure 6.26:** Exclusion limits at 95% CL for each leptonic channel and signal region on the FCNC  $tZc$  branching fractions considering one non-vanishing coupling at a time. The CMS observed (expected) limit at 95% CL at a centre-of-mass of 8 TeV [41] is given with a blue line (dashed line).

#### 1685 6.4.2 Two-dimensional limits

One can interpolate the one dimensional limits  $\lambda_{tZq}^{1D}$  to a scenario where both couplings are non-vanishing. The interpolation is taken from Ref. , where an experimental extrapolation formulae

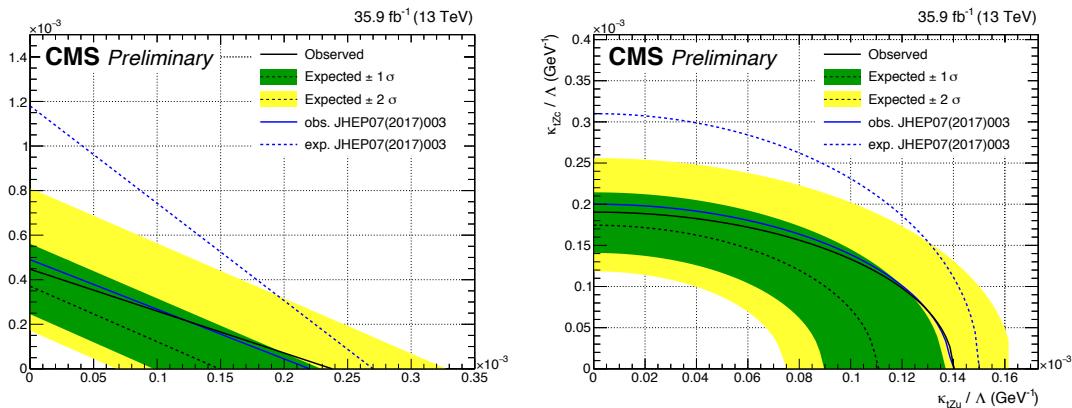
$$\kappa_{tZc}/\Lambda = \lambda_{tZc}^{1D} \sqrt{1 - \frac{\kappa_{tZu}/\Lambda}{\lambda_{tZu}^{1D}}}, \quad (6.2)$$

is found from 100 benchmark scenarios. These scenarios are constructed from existing signal samples as

$$\text{Signal yield} = (\kappa_{tZu}/\Lambda)^2(\text{ST Zut yield} + \text{TT Zut yield}) + (\kappa_{tZc}/\Lambda)^2(\text{ST Zct yield} + \text{TT Zct yield}). \quad (6.3)$$

1686 For each scenario a dedicated BDT training is performed and the expected limit at 95% CL are  
1687 calculated. The resulting two-dimensional limits are shown in Figure 6.27.

**NOTE:** Add source



**Figure 6.27:** Two-dimensional limits on the branching fractions (left) and couplings (right) for FCNC interactions involving a tZq vertex. The CMS observed (expected) limit at 95% CL at a centre-of-mass of 8 TeV [41] is given with a blue line (dashed line).