



**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
 Institute 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	1 Theoretical basis	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	2 Experimental set-up	23
21	2.1 The Large Hadron Collider	23
22	2.2 The Compact Muon Solenoid	27
23	2.2.1 CMS coordinate system	28
24	2.2.2 Towards the heart of CMS	29
25	2.2.3 Data acquisition	37
26	2.2.4 Phase 1 upgrades	37
27	2.2.5 CMS computing model	39
28	3 Analysis techniques	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	45
33	3.2.3 Generating FCNC top-Z interactions	46
34	3.2.4 Generating SM background events	49
35	3.3 Multivariate analysis techniques: Boosted Decision Trees	51
36	3.4 Statistical methodology	53
37	4 Event reconstruction and identification	57
38	4.1 Object Reconstruction	57
39	4.1.1 Charged particle tracks	59
40	4.1.2 Following the Muon's Footsteps	59
41	4.1.3 The path of the Electron	59

42	4.1.4 Primary Vertex Reconstruction	60
43	4.1.5 Calorimeter clusters	60
44	4.2 Particle flow identification	61
45	4.3 Pileup mitigation and luminosity measurement	61
46	4.4 Physics object reconstruction and identification	62
47	4.4.1 Muons	62
48	4.4.2 Electrons	65
49	4.4.3 Jets	65
50	4.4.4 Jets from b fragmentation	69
51	4.4.5 Missing transverse energy	71
52	4.5 Summary of corrections	72
53	5 Event selection and categorisation	73
54	5.1 Baseline event selection and filters	73
55	5.1.1 Event cleaning	76
56	5.1.2 Estimation of the trigger efficiency	78
57	5.1.3 Corrections	79
58	5.1.4 Reconstruction of kinematic variables	83
59	5.2 Analysis Strategy	84
60	5.3 Data driven NPL background simulation	85
61	5.4 Regions and channels	87
62	5.4.1 WZCR	88
63	5.4.2 TTCR and STCR	89
64	5.4.3 TTSR and STSR	89
65	6 The search for FCNC involving a top quark and a Z boson	91
66	6.1 Construction of template distributions	91
67	6.1.1 Distributions of the BDT variables	92
68	6.1.2 BDTs	92
69	6.1.3 Transverse mass in WZCR	92
70	6.2 Systematic uncertainties	93
71	6.3 Limit setting procedure validation	94
72	6.4 Result and discussion	95
73	6.4.1 One dimensional limits	95
74	6.4.2 Two-dimensional limits	97
75	7 Denouement of the top-Z FCNC hunt at 13 TeV	99
76	Appendices	99
77	A Trigger scale factors	101■
78	B Dilepton controlplots	113■
79	C Statistical independent regions	115■

Theoretical basis

1

81 The Standard Model (SM) [Peskin:257493] is a name given in the 1970s to a theory describing
 82 the fundamental particles and their interactions. This quantum field theory describes the
 83 particles and their interactions as fields and has successfully incorporated three of the four
 84 fundamental forces in the universe. In [Section 1.1](#), the particle content of the SM is summarised,
 85 while [Section 1.2](#) describes the SM Lagrangian and its symmetries. In [Section 1.3](#), the flavour
 86 content of the SM is highlighted, and [Section 1.4](#) focusses on the top quark in the SM.

87 The successful theory of the SM has some shortcomings which are discussed in [Section 1.6](#)
 88 and lead to searches for a more general theory. One of such is using an effective field theory
 89 (EFT) approach [Burgess:2007pt] to search for new physics in a model independent way. In
 90 [Section 1.7](#) an EFT model focussing on flavour changing neutral currents (FCNC) involving a
 91 top quark is presented. Its current experimental constraints are given in [Section 1.8](#).

92 1.1 Elementary particles and forces

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

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

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

100 The fermions are the particles that make up the visible matter in the universe. They carry half
 101 integer spin and can be subdivided into leptons and quarks, where leptons do not interact strongly.
 102 Each fermion has a corresponding anti-fermion which has the same mass and is oppositely
 103 charged. The electron e is the first elementary particle discovered [[electrondiscovery](#)] and
 104 belongs to the first generation of leptons together with the electron neutrino ν_e . The second
 105 generation comprises the muon μ and muon neutrino ν_μ , whereas the third generation consists
 106 of the tau τ and tau neutrino ν_τ . The neutrinos are neutral particles, while the other leptons
 107 have charge $\pm q_e$ with q_e representing the elementary charge of 1.602×10^{-19} C. The masses of
 108 charged leptons differ by four orders of magnitude between the first and third generations. In
 109 the SM the neutrinos are assumed to be massless, nonetheless it is experimentally established
 110 that neutrinos do have a tiny non-zero mass [[Fukuda:1998mi](#), [PhysRevLett.108.131801](#)]. 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 [PDG], where q_e represents the elementary charge.

Generation	Particle	Mass	Charge
First	e^-	0.511 MeV	$-q_e$
	ν_e	≈ 0	0
Second	μ^-	106 MeV	$-q_e$
	ν_μ	≈ 0	0
Third	τ^-	1 777 MeV	$-q_e$
	ν_τ	≈ 0	0

111

112 The quarks can also be divided into three generations. Unlike the leptons, they carry colour
 113 charge and can interact via the strong interaction. The top quark, discovered in 1995 at the
 114 Tevatron [[observationtopD0](#), [observationtopCDF](#)], is the heaviest SM particle with a mass¹
 115 measured to be $173.34 \pm 0.27(\text{stat}) \pm 0.71(\text{syst})$ GeV [[ATLAS:2014wva](#)]. The quarks and their
 116 properties are summarised in [Table 1.3](#). In nature, only colour neutral objects can exist. This
 117 has as consequence that quarks are bound through gluons into mesons (quark+anti-quark)
 118 and baryons (three quarks). These mesons and baryons are mostly short-lived and unstable
 119 particles that rapidly decay through W^\pm and Z bosons. The only known stable baryon is the
 120 proton, made up of two up quarks and one down quark.

121 The scalar boson, commonly known as the Higgs boson, is the last piece of the SM discovered
 122 in 2012 [[Chatrchyan:2012xdj](#), [Aad:2012tfa](#)]. It is responsible for the masses of the W^\pm and
 123 Z boson, and that of the fermions.

¹In 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 [PDG], 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 ± 0.03 GeV	$\frac{2}{3} q_e$	
	strange s	96^{+8}_{-4} MeV	$\frac{-1}{3} q_e$	
Third	top t	173.1 ± 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 invariance².

The $U(1)_Y$ group has one generator Y with an associated gauge field B_μ . The three gauge fields W_μ^1 , W_μ^2 , and W_μ^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 ϵ_{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

²Different 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:

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

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 ϕ 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_μ , the Z boson field Z_μ^0 , and the W boson fields W_μ^\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 θ_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}_ϕ 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 λ 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:

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

- 155 The Brout-Englert-Higgs field couples universally to fermions with a strength proportional to
 156 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 [PhysRevLett.10.531] ■
 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_{weak} , where d_{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 θ_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_{weak} is left uncoupled, leading to Glashow, Iliopoulos and Maiani (GIM) [PhysRevD.2.1285, BJORKEN1964255, Maiani:2013fpa] 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_{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 [Peskin:257493].

Kobayashi and Maskawa generalised the Cabibbo rotation matrix to accommodate a third generation of quarks. The result is a 3×3 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}_{CKM} is unitary ($\mathcal{V}_{\text{CKM}}^\dagger \mathcal{V}_{\text{CKM}} = \mathbb{1}$). A general 3×3 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 [CKM]. Each element V_{ij} of \mathcal{V}_{CKM} represents the transition probability of a quark i going to a quark j , and is experimentally determined to be [PDG]

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

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 ratios \mathcal{B} for FCNC decays involving the top quark in the SM [AguilarSaavedra:2004wm].

Process	\mathcal{B} in the SM	Process	\mathcal{B} in the SM
$t \rightarrow uZ$	8×10^{-17}	$t \rightarrow cZ$	1×10^{-14}
$t \rightarrow u\gamma$	4×10^{-16}	$t \rightarrow c\gamma$	5×10^{-14}
$t \rightarrow ug$	4×10^{-14}	$t \rightarrow cg$	5×10^{-12}
$t \rightarrow uH$	2×10^{-17}	$t \rightarrow cH$	3×10^{-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 [Abachi:1995iq, Abe:1995hr], 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 [PDG]

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

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 [PDG]
 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 structure³, 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 the
 194 LHC centre-of-mass energy regime of 7-14 TeV [PDG]. 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

³In 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.

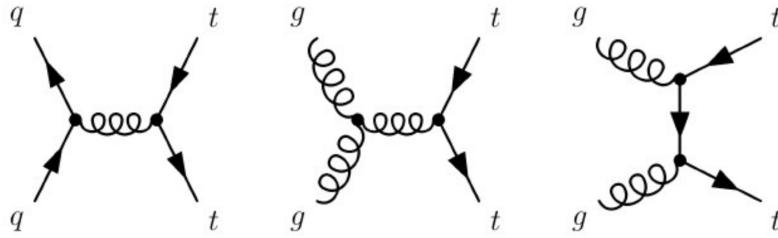


Figure 1.1: Leading order diagrams of the top quark pair production. Gluon fusion (right and middle) are the dominant processes at the LHC, while quark fusion (left) is the dominant one at the Tevatron.

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 [PDG]. 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.

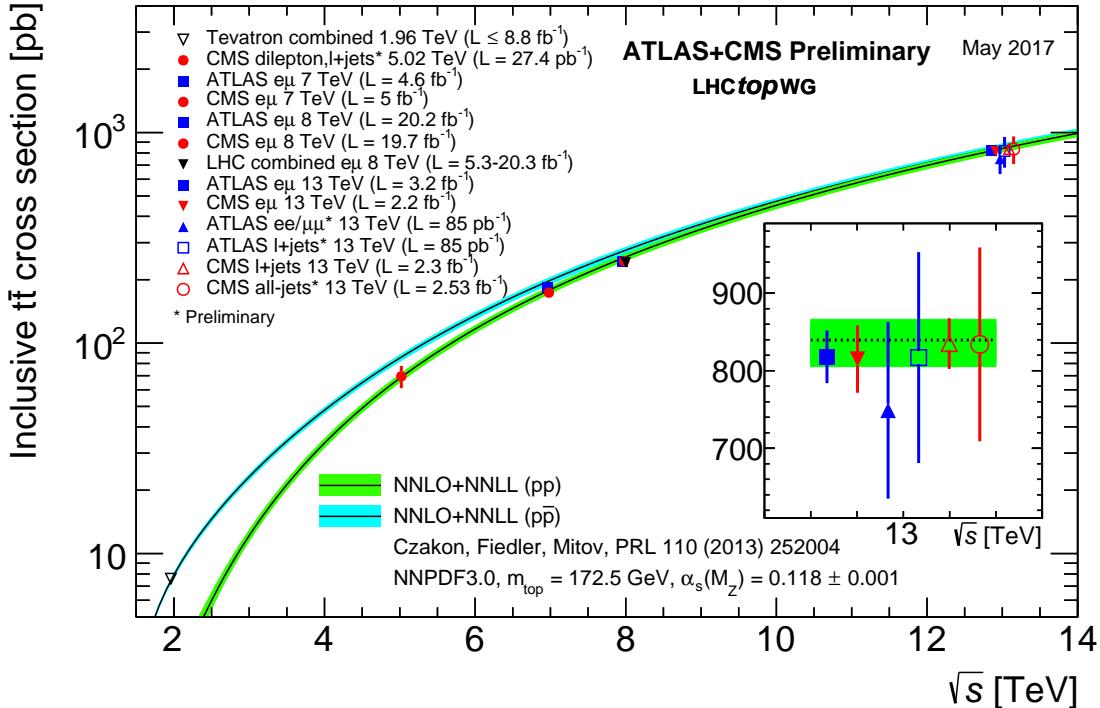


Figure 1.2: Summary of the LHC and the Tevatron measurements of the top quark pair production cross section as function of the centre-of-mass energy compared with the next-to-next-to-leading order QCD calculation. The theory bands are the uncertainties due to renormalization and factorisation scales, 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 [\[summarytwiki\]](#).

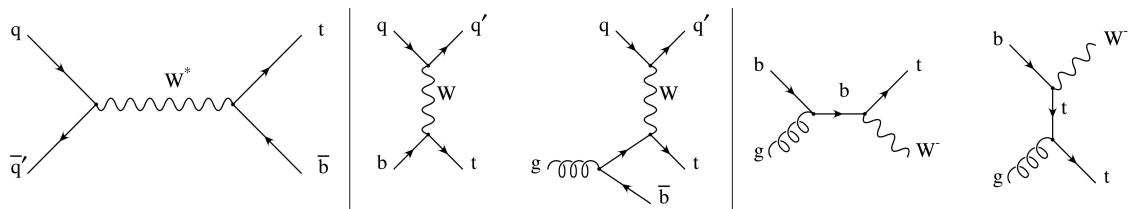


Figure 1.3: Leading order Feynman diagrams of the electroweak production of single top quarks in the s -channel (left), t -channel (middle), and for the tW associated production. Figure taken from [\[singletop\]](#).

Table 1.6: Predictions on the single top quark production cross sections at next-to-leading order per centre-of-mass energy [PDG]. 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 [PDG, stwiki].

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	$216.99^{+9.04}_{-7.71}$	$10.32^{+0.40}_{-0.36}$	$71.7^{+1.80+3.40}_{-1.80-3.40}$

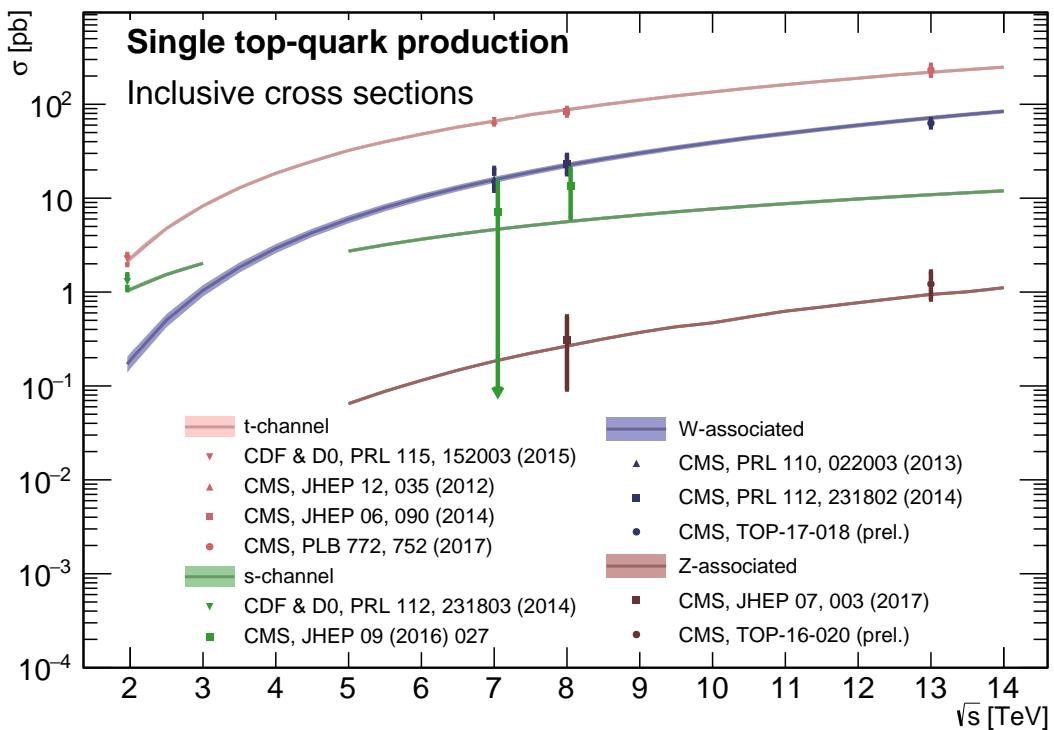


Figure 1.4: Summary of the measurements of the single top quark production cross section as function of the centre-of-mass energy. Figure taken from [summarywiki].

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 [thesisDeg]. Effective
 228 field theories can be used for theories where the perturbative expansion cannot be trusted, e.g.
 229 QCD at low energy, or as bottom up approach to look for new physics in a model independent
 230 way. 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 [Fermi2008] the β -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_μ can written as the sum of an hadronic J_μ^h and leptonic J_μ^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 Λ_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 Λ_F is determined to be $\Lambda_F \approx 348 \text{ GeV}$ [thesisDeg]. 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 [Peskin:257493]

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

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 Λ . 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 [[Grzadkowski:2010es](#)], 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. [[Grzadkowski:2010es](#)]. 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. [[durieuxEFT](#)]. These results are consistent with the SM expectation for which those operators vanish.

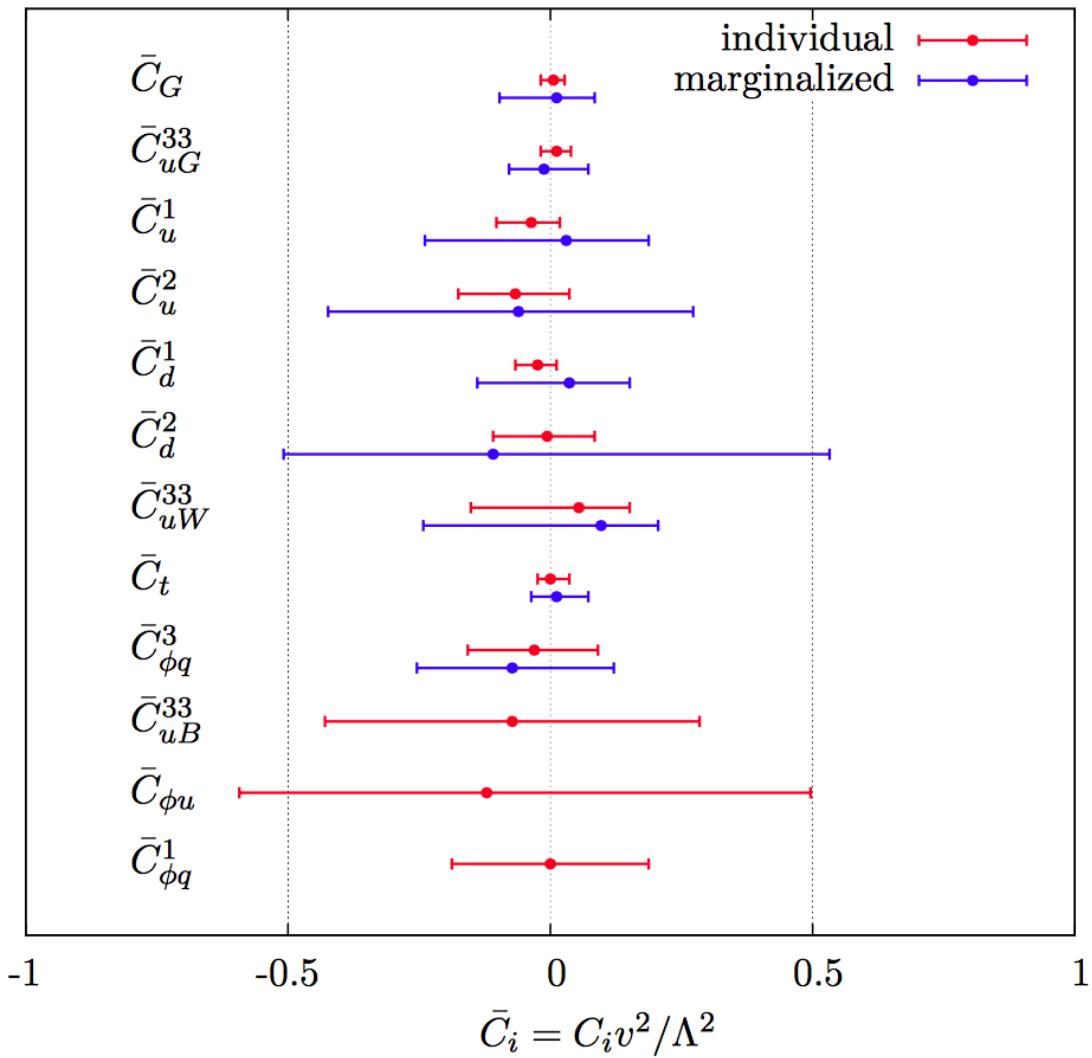


Figure 1.5: Global fit results of top quark effective field theory to experimental data including all constrained operators at dimension six. For the operators, the Warsaw basis of [Grzadkowski:2010es] is used. The bounds are set on the Wilson coefficients of various operators contributing to top quark production and decay in two cases (red) all other coefficients set to zero, or (blue) all other coefficients are marginalized over. Figure taken from [Buckley:2015lku].

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) [BSMWiley] is strongly motivated. These motivations
 258 are based on direct evidence from observation such as the existence of neutrino masses, the
 259 existence of dark matter and dark energy, or the matter-antimatter asymmetry, and also from
 260 theoretical problems such as the hierarchy problem, the coupling unification or the large numbers
 261 of free 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 [Fukuda:1998mi, PhysRevLett.108.131801]. These oscillations are only possible
 265 when neutrinos have masses. The flavour neutrinos (ν_e , ν_μ , ν_τ) are then linear expressions of
 266 the fields of at least three mass eigenstate neutrinos ν_1 , ν_2 , and ν_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 [Ade:2015xua], the density of cold non baryonic matter
 272 is determined. Cold dark matter is assumed to be only sensitive to the weak and gravitational
 273 force, leading to only one possible SM candidate: the neutrino. However, these are too light to
 274 account for the vast amount of dark matter and other models are needed. Dark energy is assumed
 275 to be responsible for the acceleration in the expansion of the universe [Peebles:2002gy].

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 [Sakharov]. These mechanisms are those of baryon and lepton number
 280 violation, that at a given moment in time there was a thermal imbalance for the interactions in
 281 the universe, and there is charge C and charge parity CP violation⁴.

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 [Burdman:2007ck] is related to the huge difference in energy be-
 288 tween the weak scale and the Planck scale. The vev of the Brout-Englert-Higgs field determines
 289 the weak scale that is approximately 246 GeV. The radiative corrections to the scalar boson
 290 squared mass m_H^2 , coming from its self couplings and couplings to fermions and gauge bosons,

⁴The 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 are quadratically proportional to the ultraviolet momentum cut-off Λ_{UV} . This cut-off is at least
 292 equal to the energy to which the SM is valid without the need of new physics. For the SM to be
 293 valid up to the Planck mass, the correction to m_H^2 becomes thirty orders of magnitude larger
 294 than m_H^2 . This implies that an extraordinary cancellation of terms should happen. This is also
 295 known as the 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 ϕ with a coupling λ_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 Λ_{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
 309 is a sensitive probe for new physics. Studying its properties is therefore an important topic
 310 of the experimental program at the LHC. Several extensions of the SM enhance the FCNC
 311 branching ratios and can be probed at the LHC [AguilarSaavedra:2004wm], from which
 312 some of them are shown in Table 1.7. Previous searches have been performed at the Tevatron
 313 by the CDF [PhysRevLett.101.192002] and D0 [Abazov:2010qk] collaborations, and
 314 at the LHC by the ATLAS [Aad:2015uza, Aad:2015gea, Aad:2015pj, Aad:2017mfd,
 315 ATLAS-CONF-2017-070] and CMS [Sirunyan:2017kkr, Chatrchyan:2013nwa, Khachatryan:2015att,
 316 Sirunyan:2017kkr, Chatrchyan:2016atv, CMS-PAS-TOP-17-003] collaborations.

317 The impact of BSM models can be written in a model independent way by means of an
 318 effective field theory valid up to an energy scale Λ . The leading effects are parametrized by a
 319 set of fully gauge symmetric operators that are added to the SM Lagrangian and can be reduced

Table 1.7: The predicted branching ratios \mathcal{B} for FCNC interactions involving the top quark in some BSM models [AguilarSaavedra:2004wm]: 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
$t \rightarrow uZ$	$\leq 1.1 \times 10^{-4}$	—	$\leq 2 \times 10^{-6}$	$t \rightarrow cZ$	$\leq 1.1 \times 10^{-4}$	$\leq 10^{-7}$	$\leq 2 \times 10^{-6}$
$t \rightarrow u\gamma$	$\leq 7.5 \times 10^{-9}$	—	$\leq 2 \times 10^{-6}$	$t \rightarrow c\gamma$	$\leq 7.5 \times 10^{-9}$	$\leq 10^{-6}$	$\leq 2 \times 10^{-6}$
$t \rightarrow ug$	$\leq 1.5 \times 10^{-7}$	—	$\leq 8 \times 10^{-5}$	$t \rightarrow cg$	$\leq 1.5 \times 10^{-7}$	$\leq 10^{-4}$	$\leq 8 \times 10^{-5}$
$t \rightarrow uH$	$\leq 4.1 \times 10^{-5}$	$\leq 5.5 \times 10^{-6}$	$\leq 10^{-5}$	$t \rightarrow cH$	$\leq 4.1 \times 10^{-5}$	$\leq 10^{-3}$	$\leq 10^{-5}$

320 to a minimal set of operators as seen in [Equation 1.24](#). For simplicity, the assumption is made
 321 that new physics effects are exclusively described by dimension-6 operators, thus neglecting
 322 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 Λ is large, contributions suppressed by powers of Λ 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 [[Grzadkowski:2010es](#)], parametrising the new physics effects relevant for the flavour changing neutral current interactions of the top quark as, all flavour indices understood,

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

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 Φ , 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 ϵ_{ij}^k . The gauge covariant derivatives are also standard defined as

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

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 [[Peskin:257493](#)]. 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 [[AguilarSaavedra:2004wm](#), [Beneke:2000hk](#)] 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

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

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

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

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

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

where the value of the FCNC couplings at scale Λ are represented by $\kappa_{tZq}, \kappa_{tgq}, \kappa_{t\gamma q}, \zeta_{tZq}$, and η_{Hqt} . These are assumed to be real and positive, with the unit of GeV^{-1} for κ_{txq}/Λ and no unit for ζ_{xqt} and η_{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 θ_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_μ , the gluon field $G_{\mu\nu}^{a=1..8}$, and the Z boson Z_μ^0 are defined as

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

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

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

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

330 1.8 Experimental constraints on top-FCNC

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

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

331 where $\Gamma_{t \rightarrow qX}$ represents the FCNC decay width⁵ for a coupling strength $\delta_{txq}^2 = 1$, and Γ_t the full
 332 decay width of the top quark. In the SM, supposing a top quark mass of 172.5 GeV, the full
 333 width becomes $\Gamma_t^{\text{SM}} = 1.32$ GeV [Gao:2012ja].

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

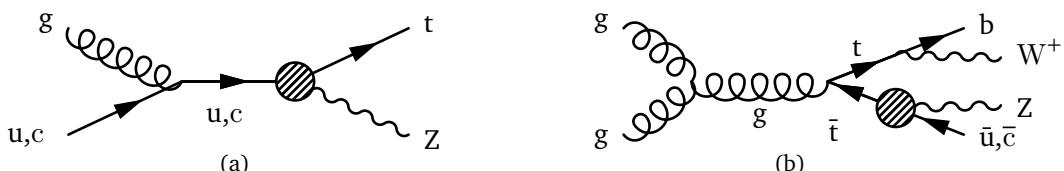


Figure 1.6: Feynman diagrams for the processes with a tZq FCNC interaction, where the FCNC interaction is indicated with the shaded dot. (a) Single top quark production through an FCNC interaction. (b) Top quark pair production with an FCNC induced decay.

⁵The 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.

339 The observation of top-FCNC interactions has yet to come and experiments have so far only
 340 been able to put upper bounds on the branching ratios. An overview of the best current limits
 341 is given in [Table 1.8](#). In [Figure 1.7](#) a comparison is shown between the current best limits set
 342 by ATLAS and CMS with respect to several BSM model benchmark predictions. From there
 343 one can see that FCNC searches involving a Z or H boson are close to excluding or confirming
 344 several BSM theories. In [Figure 1.9](#), the searches performed by CMS are summarised. For
 345 the tZq vertex, the best limit from CMS comes from Ref. [[Sirunyan:2017kkr](#)] where both
 346 single top quark and top quark pair are studied. The observed (expected) limits 95% CL at
 347 8 TeV for the FCNC tZq interaction by CMS are $\mathcal{B}(t \rightarrow uZ) < 2.2 \times 10^{-4}$ (2.7×10^{-4}) and
 348 $\mathcal{B}(t \rightarrow cZ) < 4.9 \times 10^{-4}$ (12×10^{-4}). In [Figure 1.8](#), the summary of the 95% confidence level
 349 observed limits on the branching ratios of the top quark decays to a charm or up quark and a
 350 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 ratios \mathcal{B} at 95% confidence level.

Process	Search mode	Observed \mathcal{B}	Expected \mathcal{B}	Experiment	
$t \rightarrow uZ$	top quark pair decay	1.7×10^{-4}	2.4×10^{-4}	ATLAS	[ATLAS-CONF-2017-070]
$t \rightarrow u\gamma$	single top quark production	1.3×10^{-4}	1.9×10^{-4}	CMS	[Khachatryan:2015att]
$t \rightarrow ug$	single top quark production	4.0×10^{-5}	3.5×10^{-5}	ATLAS	[Aad:2015gea]
$t \rightarrow uH$	top quark pair decay	2.4×10^{-3}	1.7×10^{-3}	ATLAS	[Aaboud:2017mfd]
$t \rightarrow cZ$	top quark pair decay	2.3×10^{-4}	3.2×10^{-4}	ATLAS	[ATLAS-CONF-2017-070]
$t \rightarrow c\gamma$	single top quark production	2.0×10^{-3}	1.7×10^{-3}	CMS	[Khachatryan:2015att]
$t \rightarrow cg$	single top quark production	2.0×10^{-4}	1.8×10^{-4}	ATLAS	[Aad:2015gea]
$t \rightarrow cH$	top quark pair decay	2.2×10^{-3}	1.6×10^{-3}	CMS	[Aaboud:2017mfd]

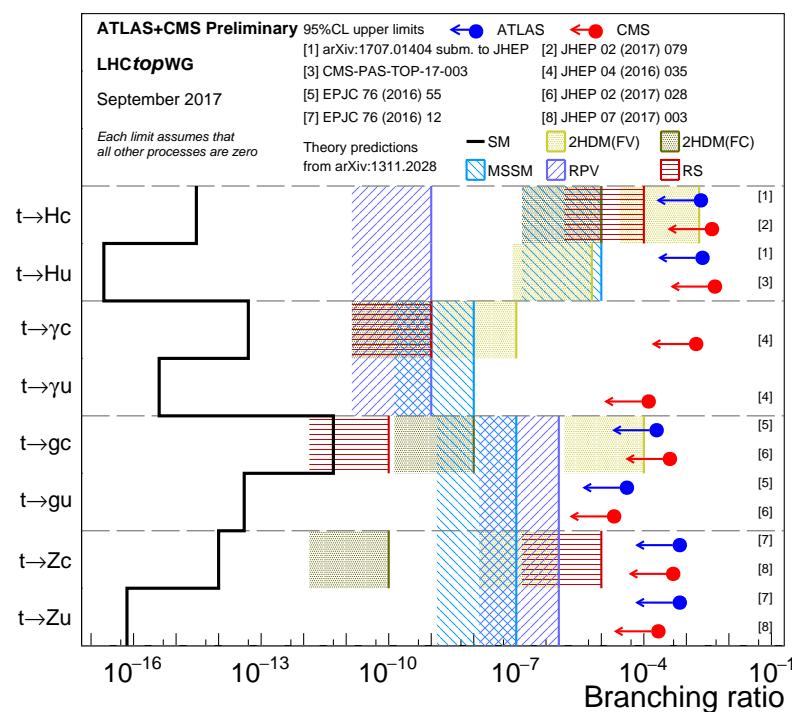


Figure 1.7: Current best limits set by CMS and ATLAS for top-FCNC interactions. Figure taken from [summarywiki]. (TO DO Remake with new atlas results)

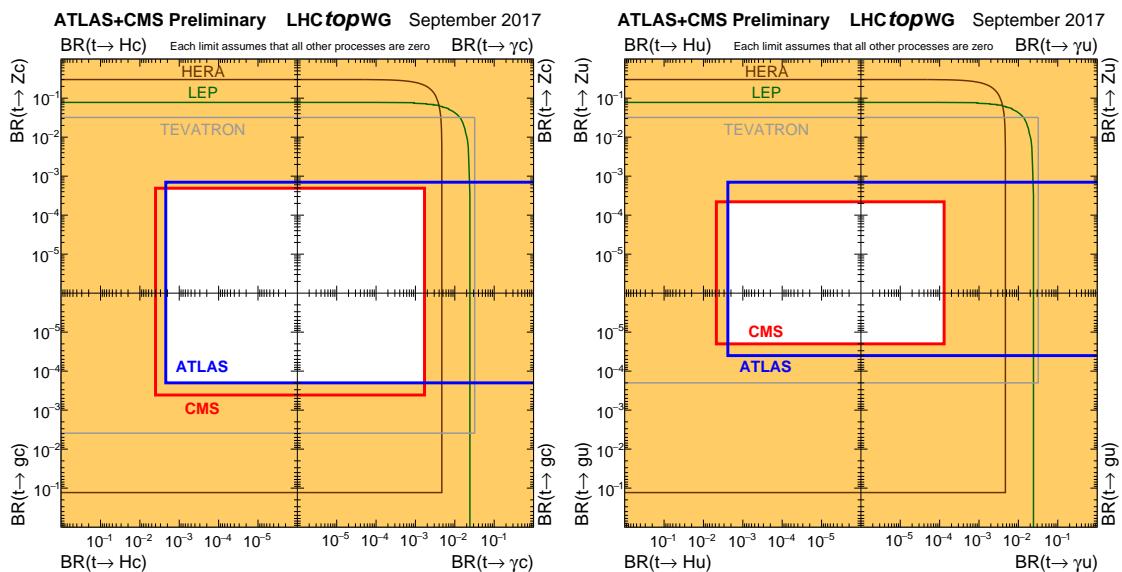


Figure 1.8: Summary of the current 95% confidence level observed limits on the branching ratios of the top quark decays via flavour changing neutral currents to a charm (left) or up (right) quark and a neutral boson. The coloured lines represent the results from HERA (the most stringent limits between 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 [[summarytwiki](#)].

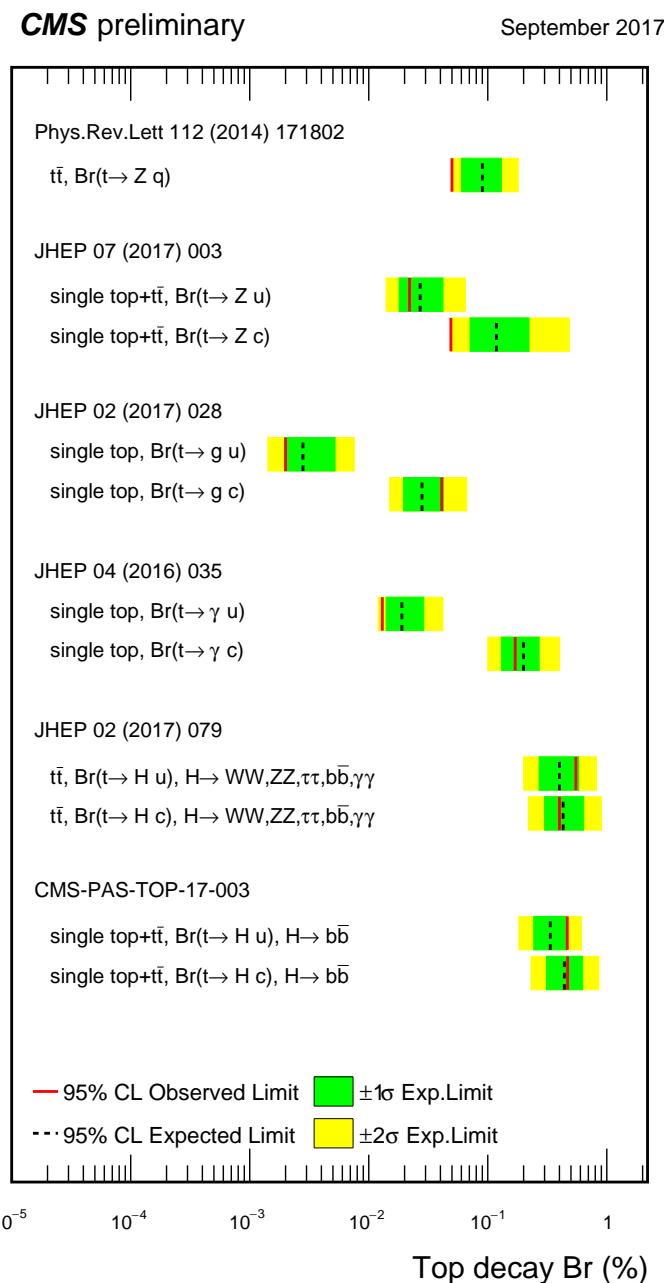


Figure 1.9: Summary of the FCNC branching ratios from CMS searches at 8 TeV. Figure taken from [\[summarywiki\]](#).

Experimental set-up

2

353 A key objective of the Large Hadron Collider (LHC) was the search for the Brout-Englert-Higgs bo-
 354 son. The Large Electron Positron (LEP) [Myers:226776] and Tevatron [1748-0221-6-08-T08001]■
 355 experiments established that the mass of the scalar boson has to be larger than 114 GeV [Barate:2003sz]■
 356 Herner:2016woc], and smaller than approximate 1 TeV due to unitarity and perturbativity
 357 constraints [Djouadi:2005gi]. On top of this, the search for new physics such as supersymmetry
 358 or the understanding of dark matter were part of the motivation for building the LHC. Since the
 359 start of its operation, the LHC is pushing the boundaries of the Standard Model, putting the most
 360 stringent limits on physics beyond the Standard Model as well as precision measurements of the
 361 parameters of the Standard Model. A milestone of the LHC is the discovery of the scalar boson
 362 in 2012 by the two largest experiments at the LHC [Chatrchyan:2012xdj, Aad:2012tfa].

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

369 2.1 The Large Hadron Collider

370 The LHC has started its era of cutting edge science on 10 September 2008 [LHC:2008] after ap-
 371 proval by the European Organisation of Nuclear Research (CERN) in 1995 [Pettersson:291782].
 372 Installed in the previous LEP tunnel, the LHC consists of a 26.7 km quasi ring, that is installed
 373 between 45 and 170 m under the French-Swiss border amidst Cessy (France) and Meyrin
 374 (Switzerland). Built to study rare physics phenomena at high energies, the LHC can accelerate
 375 mainly two types of particles, protons and lead ions Pb^{45+} , and provides collisions at four inter-
 376 action points, where the particle bunches are crossing. Experiments for studying the collisions
 377 are installed at each interaction point.

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

381 continue to the Proton Synchrotron Booster (PSB or Booster), where the packs of protons are
 382 accelerated to 1.4 GeV and each pack is split up in twelve bunches with 25 or 50 ns spacing.
 383 The Proton Synchrotron (PS) then increases their energy to 25 GeV before the Super Proton
 384 Synchrotron (SPS) increases the proton energy up to 450 GeV. Each accelerator ring expands in
 385 radius in order to reduce the energy loss of the protons by synchrotron radiation¹. Furthermore,
 386 the magnets responsible for the bending of the proton trajectories have to be strong enough
 387 to sustain the higher proton energy. Ultimately, the proton bunches are injected into opposite
 388 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) [Wenninger:2254678].

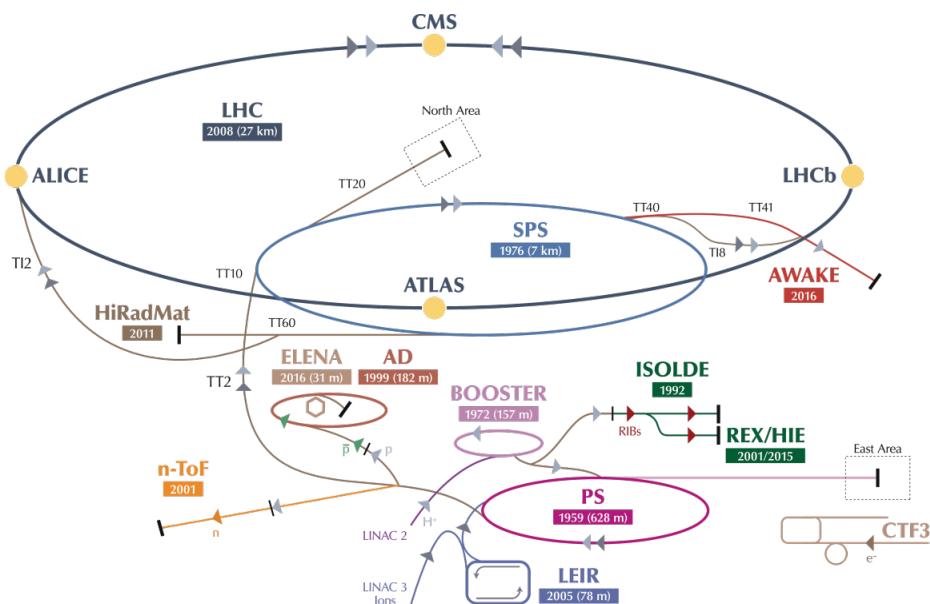


Figure 2.1: Schematic representation of the accelerator complex at CERN [DeMelis:2197559]. The LHC (dark blue) is the last element in chain of accelerators. Protons are successively accelerated by LINAC 2, the Booster, the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS) before entering the LHC.

389

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

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

399 will be performed. Furthermore, preparations for the High Luminosity LHC, which will start in
 400 2024, will be done. More information about phase 1 upgrades during LS1 and EYETS is given
 401 in [Section 2.2.4](#).

402 Before the start of the LHC in 2010, the previous energy record was held by the Tevatron collider
 403 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).

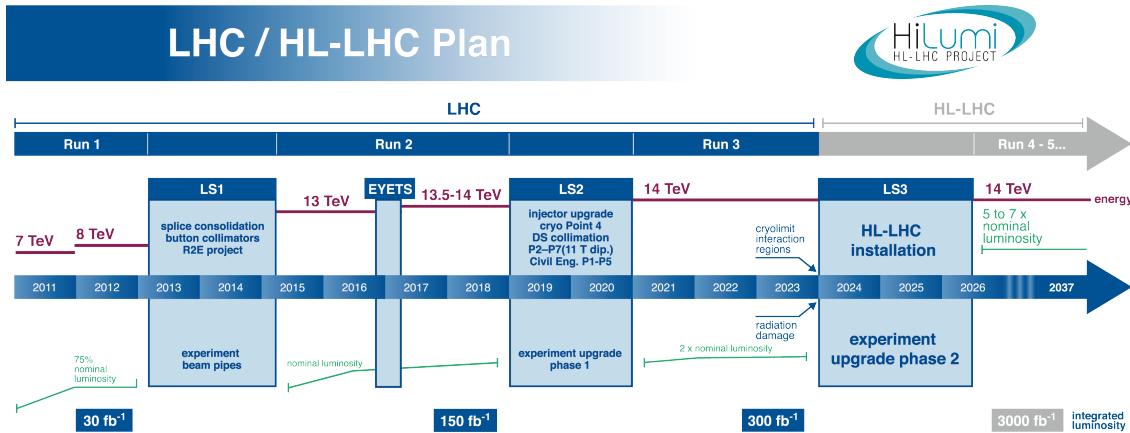


Figure 2.2: The HL-LHC timeline. Figure taken from [\[Antonella:1975962\]](#).

404

405 Inside the LHC ring [\[Bruning:782076\]](#), the protons are accelerated by the means of radio
 406 frequency cavities, while 1232 dipole magnets of approximately 15 m long, each weighing 35 t
 407 ensure the deflection of the beams. The two proton beams circulate in opposite direction in
 408 separate pipes inside of the magnet. Through the use of a strong electric current in the coils
 409 of the magnet, magnetic fields are generated and cause the protons to bend in the required
 410 orbits. In order for the coil to become superconducting and able to produce a strong magnetic
 411 field of 8.3 T, the magnet structure is surrounded by a vessel. This vessel is filled with liquid
 412 Helium making it possible to cool down the magnet to 1.9 K. In order to get more focussed and
 413 stabilised proton beams, additional higher-order multipole and corrector magnets are placed
 414 along the LHC beam line.

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

- 416 • A Toroidal LHC ApparatuS (ATLAS) [\[Aad:2008zzm\]](#) and the Compact Muon Solenoid
 417 (CMS) [\[Chatrchyan:2008aa\]](#) experiments are the two general purpose detectors at the
 418 LHC. They both have a hermetic, cylindrical structure and were designed to search for
 419 new physics phenomena along with precision measurements of the Standard Model. The
 420 existence of two distinct experiments allows cross-confirmation of any discovery.
- 421 • A Large Ion Collider Experiment (ALICE) [\[Aamodt:2008zz\]](#) and the LHC Beauty (LHCb) [\[Alves:2008zz\]](#)
 422 experiments are focusing on specific phenomena. ALICE studies strongly interacting mat-
 423 ter 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 focus on b physics.

- The forward LHC (LHCf) [Bongi:2010zz] and the TOTal cross section, Elastic scattering and diffraction dissociation Measurement (TOTEM) [Anelli:2008zza] 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) [Acharya:2014nyr] 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 [Gillies:1997001] 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 σ_{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_{rev} the revolution frequency, γ_r the relativistic gamma factor, ϵ_n the normalized transverse beam emittance - a quality for the confinement of the beam, β^* the beta function at the collision point - a measurement for the width of the beam, θ_c the angle between two beams at the interaction point, σ_z the mean length of one bunch, and σ^* 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 [Harriet:2012301]. 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 $13-17 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$ 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 fb^{-1} during 2016 data taking.

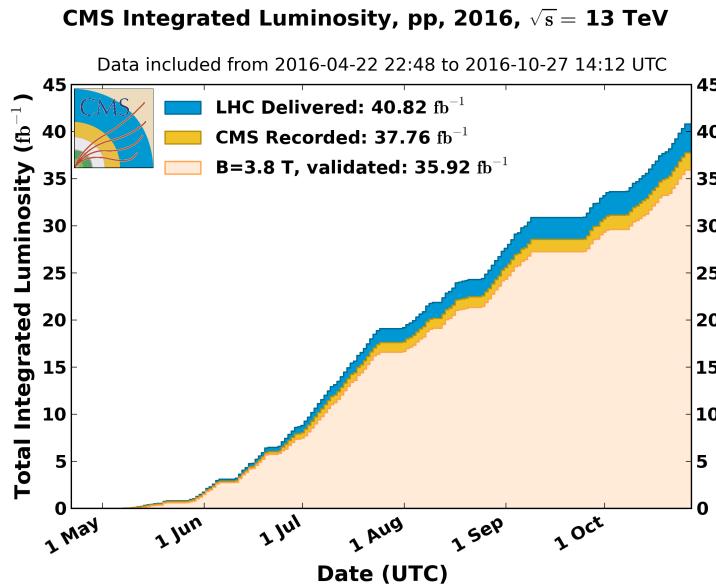


Figure 2.3: Cumulative off-line luminosity measured versus day delivered by the LHC (blue), and recorded by CMS (orange), and certified as good physics analysis during stable beams (light orange) during stable beams and for proton collisions at 13 TeV centre-of-mass energy in 2016. [LumiWiki].

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

457 2.2 The Compact Muon Solenoid

458 At one of the collision points of the LHC, the CMS detector [CMS, Bayatian:2006zz, Bayatian:922757]²
 459 is placed. Weighing 14 000 t, this cylindrical detector is about 28.7 m long and 15 m in diameter.
 460 It has an onion like structure of several specialised detectors and contains a superconducting
 461 solenoid with a magnetic field of 3.8 T. Living in a hadronic environment, multi-jet processes
 462 produced by the strong interaction are the main source of background for rare physics pro-
 463 cesses. Therefore, good identification, momentum resolution, and charge determination of
 464 muons, electrons and photons are one of the main goals of the CMS detector. Additionally, a
 465 good charged particle momentum resolution and reconstruction efficiency in the inner tracker
 466 provides identification for jets coming from b quarks or tau particles. Also the electromagnetic
 467 resolution for an efficient photon and lepton isolation as well as a good hadronic calorimeter
 468 for the missing transverse energy² were kept into account while designing CMS. In Figure 2.4,
 469 an overview of the CMS detector is shown.

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

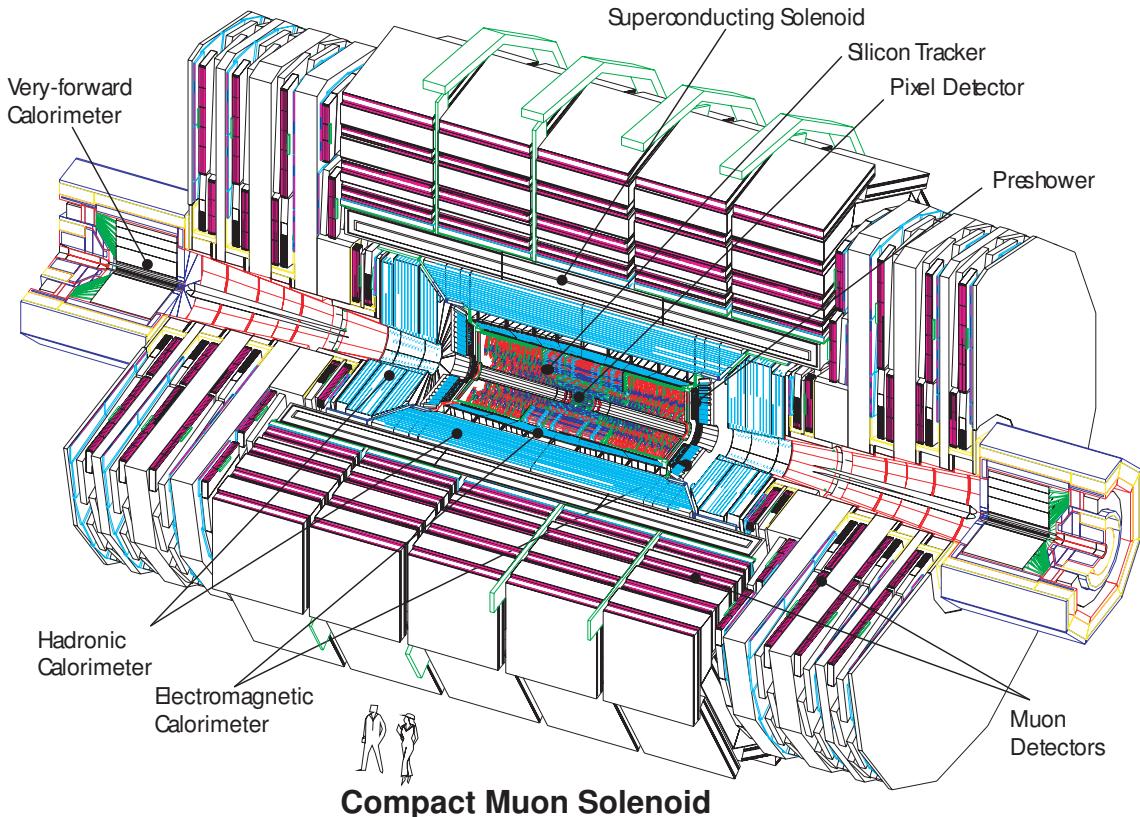


Figure 2.4: Mechanical layout of the CMS detector. Figure taken from [CMSdraw].

470 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 angle³ $\phi \in [-\pi, \pi]$, the pseudo-rapidity⁴ η :

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

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

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

- 471 where the difference of rapidities of two particles is invariant under a Lorentz boost in the
472 z-direction.

³The 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 .

⁴The pseudo rapidity is expressed by the polar angle θ between the direction of \vec{p} and the beam.

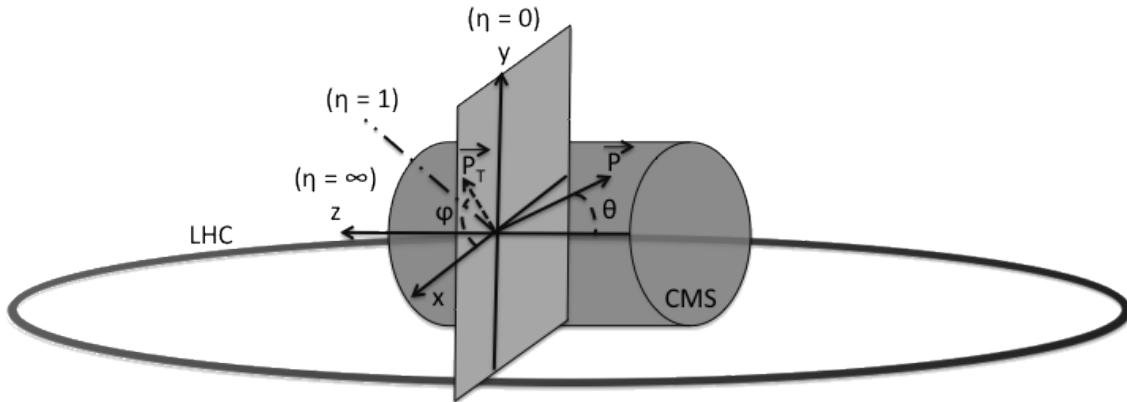


Figure 2.5: Representation of the coordinate system used by CMS. The point of origin is put at the collision point. The x-axis points towards the centre of the LHC ring such that the z-axis lies tangent to the beam axis.

473 2.2.2 Towards the heart of CMS

474 The CMS detector can be divided into two parts. A central barrel is placed around the beam
 475 pipe ($|\eta| < 1.4$), and two plugs (endcaps) ensure the hermeticity of the detector. In [Figure 2.4](#)
 476 and [Figure 2.6](#) the onion like structure of the CMS detector is visible. The choice of a solenoid
 477 of 12.9 m long and 5.9 m diameter gives the advantage of bending the particle trajectories in the
 478 transverse plane. The hadronic calorimeter ([Section 2.2.2.3](#)), the electromagnetic calorimeter
 479 ([Section 2.2.2.4](#)) and the tracker ([Section 2.2.2.5](#)) are within the solenoid ([Section 2.2.2.2](#)),
 480 while the muon chambers ([Section 2.2.2.1](#)) are placed outside the solenoid. The data used for
 481 the search presented in this thesis is collected after the long shutdown 1. After discussing each
 482 part of CMS in their Run 1 configuration, [Section 2.2.4](#) elaborates on their different upgrades
 483 for the data collected in Run 2.

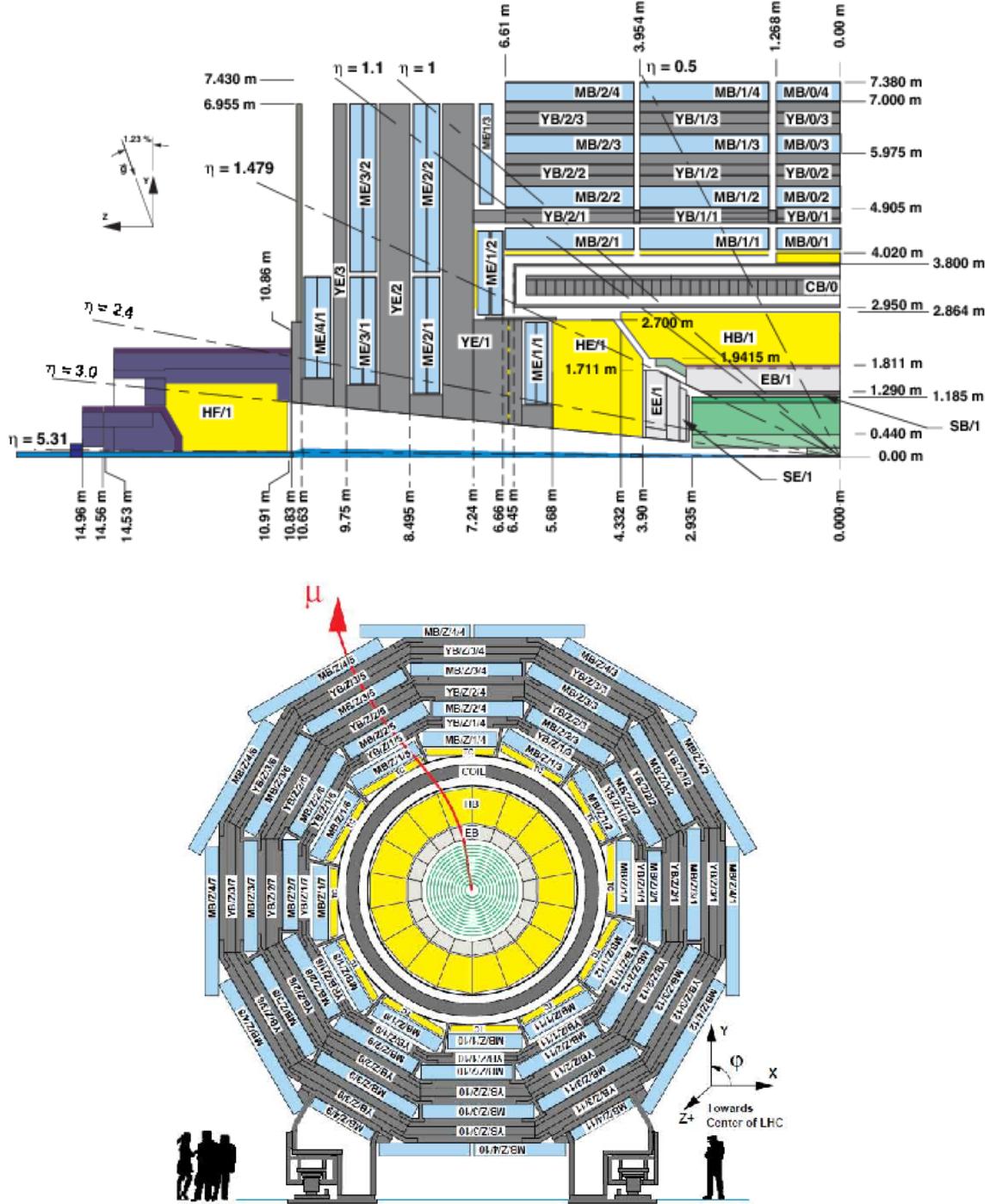


Figure 2.6: Schematic view of the CMS detector in the Run 1 configuration. The longitudinal view of one quarter of the detector is given at the top, while the transversal view is shown at the bottom. The muon system barrel elements are denoted as MBZ/N/S, where Z= -2...+2 is the barrel wheel number, N= 1...4 the station number and S= 1...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 [Chatrchyan:1223944].

484 **2.2.2.1 Muon system**

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

490 The muon system is divided into three parts, shown in [Figure 2.7](#). The muon rate and neutron
 491 induced backgrounds are small and the magnetic field is very low for the barrel, thus CMS can
 492 use drift tube (DT) chambers. For the endcaps however, the muon and background flux is much
 493 higher and there is a need to use cathode strip chambers (CSC) which are able to provide a
 494 faster response, higher granularity and have a better resistance against radiation. In order to
 495 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.

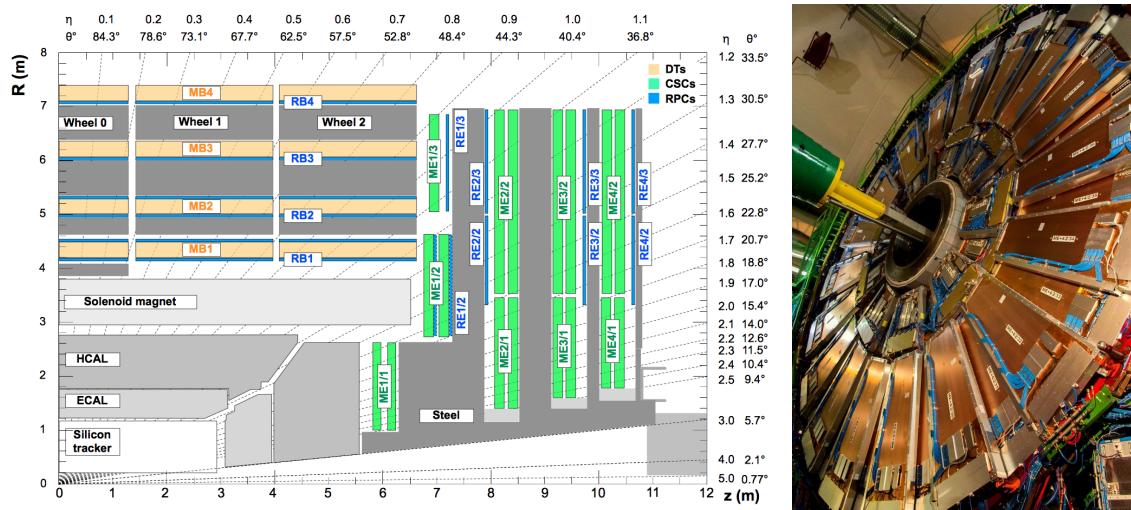


Figure 2.7: (Left) Schematic view of one quarter of the CMS muon system in the Run 1 configuration. The cathode strip chambers (CSC) are shown in green, the drift tubes (DT) are shown in yellow, while the resistive plate chambers (RPC) are shown in blue. Figure taken from [[Chatrchyan:1223944](#)]. (Right) Cathode strip chambers (ME+4/2 chambers on YE+3). Photo taken from [[muon](#)].

496

497 Providing a measurement for $|\eta| < 1.2$, the DT chambers in the barrel are on average $2 \times 2.5\text{m}^2$
 498 in size and consist of 12 layers of DT cells⁵ arranged in three groups of four. The $r\phi$ coordinate is
 499 provided by the two outside groups, while the middle group measures the z coordinate. For the
 500 outer muon station, the DT chambers contain only 8 layers of DT cells, providing a muon position
 501 in the $r\phi$ plane. There are four CSC stations in each endcap, providing muon measurements for
 502 $0.9 < |\eta| < 2.4$ (Run 1 configuration). These CSCs are multi-wired proportional chambers that
 503 consist of 6 anode wire planes crossed by 7 copper strips cathode panels in a gas volume. The
 504 r coordinate is provided by the copper strips, while the ϕ coordinate comes from the anode
 505 wires, giving a two dimensional position measurement. There are six layers of RPCs in the

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

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

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

517 2.2.2.2 Solenoid

518 Making use of the knowledge of previous experiments like ALEPH and DELPHI at LEP and H1
 519 at HERA, CMS chose for a large super conducting solenoid with a length of 12.9 m and a
 520 inner bore of 5.9 m [Bayatian:922757]. With 2168 turns, a current of 18.5 kA resulting in a
 521 magnetic field of 3.8 T, and a total energy of 2.7 GJ, a large bending power can be obtained
 522 for a modestly-sized solenoid. In order to ensure a good momentum resolution in the forward
 523 regions, a favourable length/radius was necessary. In [Figure 2.8](#), a photo of the CMS solenoid is
 524 shown.

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

530 2.2.2.3 Hadronic calorimeter

531 The hadronic calorimeter (HCAL) is dedicated to precisely measure the energy of charged and
 532 neutral hadrons via a succession of absorbers and scintillators. This makes it crucial for physics
 533 analyses with hadronic jets or missing transverse energy. The HCAL barrel extends between
 534 $1.77 < r < 2.95$ m, where r is the radius in the transverse plane with respect to the beam. Due
 535 to space limitations, the HCAL needs to be as small as possible and is made from materials
 536 with short interaction lengths⁶. On top of this, the HCAL should be as hermetic as possible and
 537 extend to large absolute pseudo rapidities such that it can provide a good measurement of the
 538 missing transverse energy.

539 The quality of the energy measurements is dependent on the fraction of the hadronic shower
 540 that can be detected. Therefore, the HCAL barrel (HB) inside the solenoid is reinforced by an
 541 outer hadronic calorimeter between the solenoid and muon detectors (HO, see [Figure 2.9](#)), using

⁶Here 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.

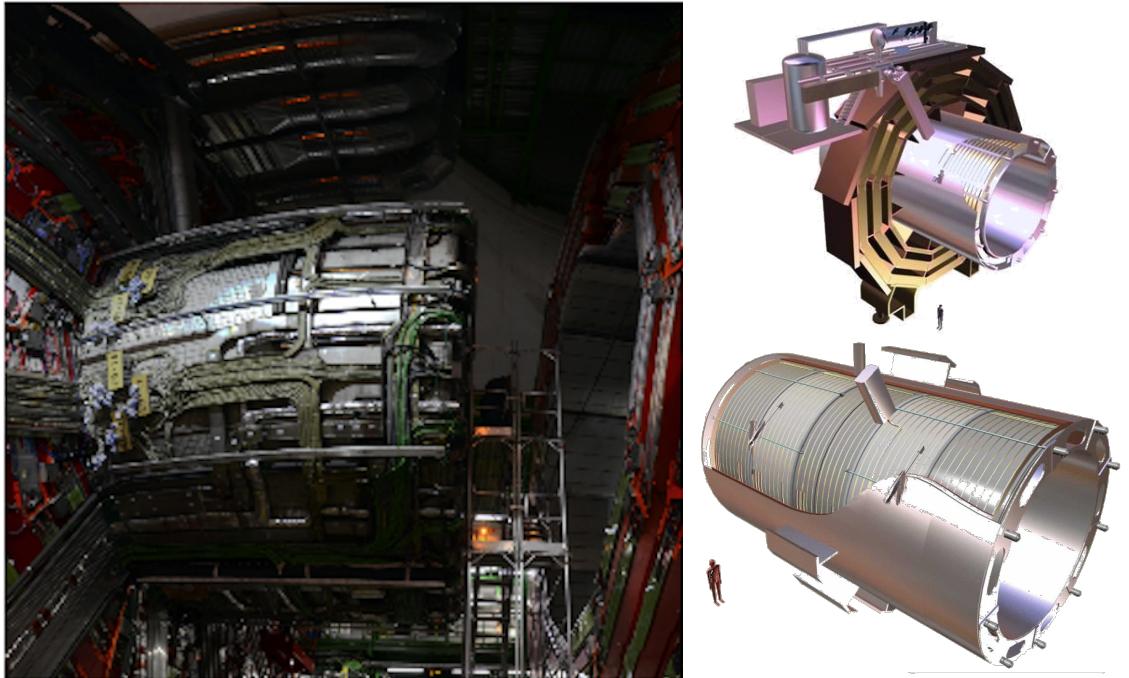


Figure 2.8: (Left) CMS solenoid during the long shutdown in 2013. (Right) An impression of the solenoid magnet taken from [[solenoid](#)].

542 the solenoid as extra absorber. This increases the thickness to 12 interaction lengths. The HB
 543 and HO provide measurements for $|\eta| < 1.3$, while an endcap on each side (HE, $1.3 < |\eta| < 3$)
 544 and a forward calorimeter (HF, $3.0 < |\eta| < 5.2$) extend the pseudo rapidity range.

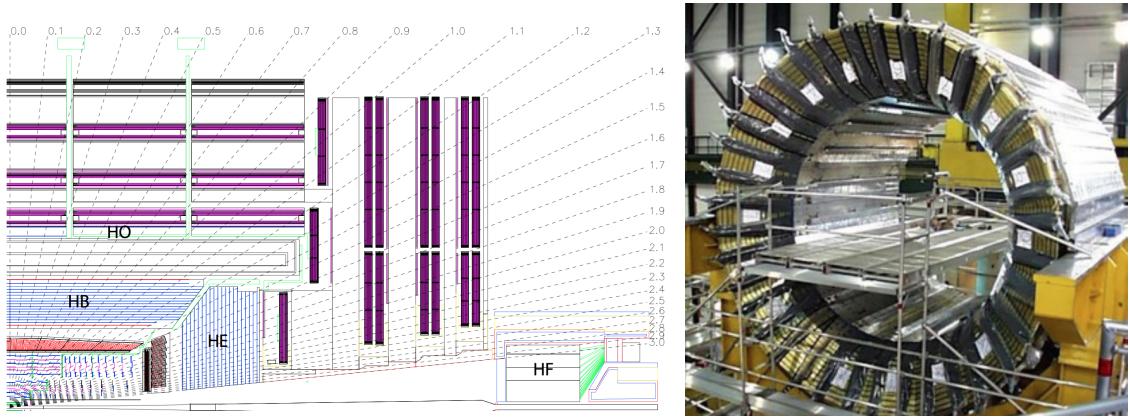


Figure 2.9: (Left) Longitudinal view of the CMS detector showing the locations of the HB, HE, HO, and HF calorimeters. Figure taken from [[Chatrchyan:2008aa](#)]. (Right) CMS barrel calorimeter. Photo taken from [[HCAL](#)].

545 The HB is made of 16 absorber plates where most of them are built from brass and others are
 546 made from stainless steel and is about five to ten interaction lengths thick. It is divided in $\eta \times \phi$
 547 towers and contains 2592 read-out channels. The HO complements the HB and extends the

548 reach up to twelve interaction lengths. This subsystem contains 2160 read-out channels. The
 549 HE is also composed of brass absorber plates and has a thickness corresponding to approximately
 550 ten interaction lengths, with 2592 read-out channels.. The HF experiences intense particle
 551 fluxes with an expected energy of 760 GeV deposited on average in a proton interaction at a
 552 centre-of-mass of 14 TeV, compared to 100 GeV in the rest of the detector. Therefore, these
 553 are Cherenkov light detectors made of radiation hard quartz fibres. The main causes of such
 554 large energy events are high energy muons and charged particles from late showering hadrons.
 555 During Run 1, it became clear that the glass windows of the photon multiplier tubes (PMTs) had
 556 to be replaced which was done during LS1 [Tiras:2016ghv]. The HF represents 1728 read-out
 557 channels.

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

560 2.2.2.4 Electromagnetic calorimeter

561 The electromagnetic calorimeter (ECAL) is designed to measure the energy of photons and
 562 electrons and covers $|\eta| < 3$. It is an hermetic, homogeneous detector and consists of 75 848
 563 lead tungstate (PbWO_4) crystals. These crystals have a fast response time - 80% of the light is
 564 emitted within 25 ns- and are radiation hard. The electromagnetic showers produced by passing
 565 electrons or photons ionize the crystal atoms which emit a blue-green scintillation light, that
 566 is collected by silicon avalanche photodiodes (APDs) in the barrel and vacuum phototriodes
 567 (VPTs) in the endcaps. The crystals and the APD response is sensitive to temperature changes
 568 and require a stable temperature.

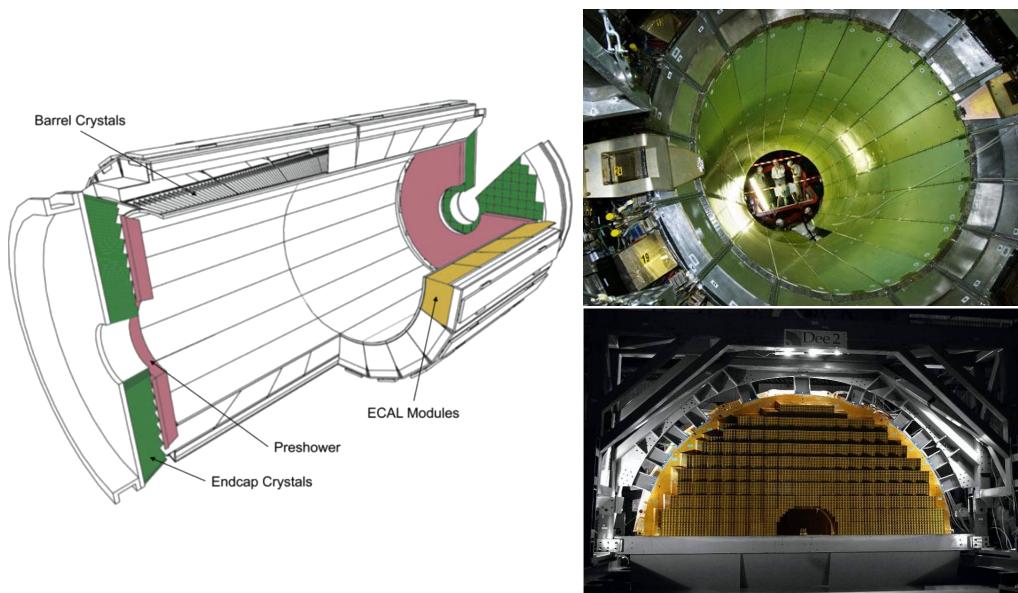


Figure 2.10: (Left) Schematic cross section of the electromagnetic calorimeter taken from [Chatrchyan:2008aa]. (Right top) The ECAL barrel during construction [ECAL]. (Right bottom) One half of an EE [EE].

569 There are three parts: a central barrel (EB), an endcap region (EE) and a preshower (ES)

570 (Figure 2.10). The EB has an inner radius of 129 cm and corresponds to a pseudo rapidity
 571 of $0 < |\eta| < 1.479$. At a distance of 314 cm from the vertex and covering a pseudo rapidity
 572 of $1.479 < |\eta| < 3.0$, are the EE. They consist of semi-circular aluminium plates from which
 573 structural units of 5×5 crystals (super crystals) are supported. The ES is placed in front of
 574 the crystal calorimeter over the endcap pseudo rapidity range with two planes of silicon strip
 575 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 5×5 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 [1748-0221-12-01-C01069] for the barrel is 2×10^{-3} rad in ϕ and 10^{-3} in η . For the endcaps this is 5×10^{-3} rad in ϕ and 2×10^{-3} in η . The energy is reconstructed by a supercluster algorithm, taking into account energy radiated via bremsstrahlung or conversion [Chatrchyan:2008aa]. The energy resolution is given by

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

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

579 2.2.2.5 Inner tracking system and operations

580 The tracking system (tracker) [Chatrchyan:1704291] is the detecting unit closest to the point
 581 of interaction. Responsible for the reconstruction of trajectories from charged particles with $|\eta|$
 582 < 2.5 that are bent by the magnetic field, it provides a measurement of the momentum. The
 583 tracker is also responsible for the determination of the interaction point or vertex. It should be
 584 able to provide high granularity as well as fast read-out, and be able to endure high radiation.
 585 For these reasons, the CMS collaboration choose silicon detector technology.

586 The tracking system consists of a cylinder of 5.8 m long and 2.5 m in diameter. It is immersed
 587 in a co-axial magnetic field of 3.8 T provided by the solenoid. As shown in Figure 2.11, the
 588 tracker is built up from a large silicon strip tracker with a small silicon pixel tracker inside. The
 589 inner pixel region ($4.4 < r < 10.2$ cm), gets the highest flux of particles. Therefore, pixel silicon
 590 sensors of $100 \times 150 \mu\text{m}^2$ are used. It consists of three cylindrical barrels that are complemented
 591 by two discs of pixel modules at each side. The silicon strip tracker ($20 < r < 116$ cm) has three
 592 subdivisions. The Tracker Inner Barrel and Discs (TIB, TID, see Figure 2.13) are composed
 593 of four barrel layers accompanied by three discs at each end. The outer part of the tracker -
 594 Tracker Outer Barrel (TOB) - consists of 6 barrel layers. In the outer discs, there are nine discs
 595 of silicon sensors, referred to as Tracker End Caps (TEC).

596 The pixel detector, shown in Figure 2.12, has 1440 modules that cover an area of about 1 m^2
 597 and have 66 million pixels. It provides a three-dimensional position measurement of the hits
 598 arising from the interaction from charged particles with the sensors. In transverse coordinate
 599 ($r\phi$), the hit position resolution is about $10 \mu\text{m}$, while $20\text{-}40 \mu\text{m}$ is obtained in the longitudinal
 600 coordinate (z). The sensor plane position provides the third coordinate. The TIB/TID, shown in
 601 Figure 2.13, delivers up to four $r\phi$ -measurements using $320 \mu\text{m}$ thick silicon micro-strip sensors.

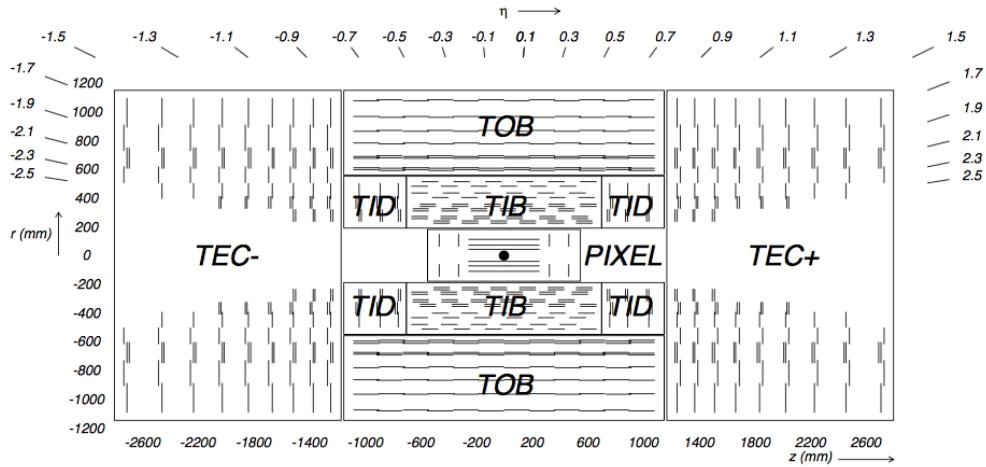


Figure 2.11: Schematic cross section through the CMS tracker. Each line represents a detector module. Double lines indicate back-to-back modules that deliver stereo hits. Figure taken from [Chatrchyan:2008aaa].



Figure 2.12: The pixel barrel being re-installed after the Long Shutdown in 2015, around the beam pipe at CMS [Christine:2024986].

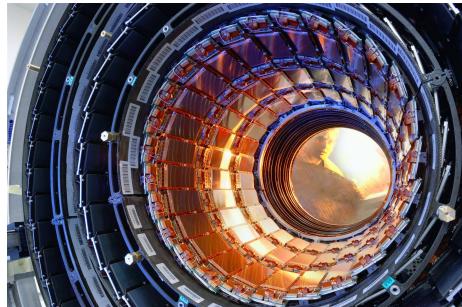


Figure 2.13: First half of the inner tracker barrel, consisting of three layers of silicon modules [beautiful:1998635].

These sensors are placed with their strips parallel to the beam axis in the barrel and radial in the discs. In the TIB, the first two layers have a strip pitch of $80\mu\text{m}$, while the remaining two have a strip pitch of $120\mu\text{m}$. This leads to a respective single point resolution of $23\mu\text{m}$ 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 $r\phi$ -measurements with a single point resolutions of $53\mu\text{m}$ in the first four layers, and $35\mu\text{m}$ in the last two layers. It consists of $500\mu\text{m}$ thick microstrip sensors with strip pitches of $183\mu\text{m}$ (first 4 layers) or $122\mu\text{m}$ (last two layers). The TEC provides up to 9 ϕ -measurements via 9 discs consisting of up to 7 rings of silicon microstrip sensors of $97\mu\text{m}$ to $184\mu\text{m}$ average pitch.

A second co-ordinate measurement (z in the barrel, r on the discs) is provided through the use of a second micro strip detector module mounted back-to-back with a stereo angle of 100 mrad . This is done on the modules in the first two layers and rings of the TIB, TID, and TOB, as well as rings 1, 2, and 5 of the TECs (blue line in Figure 2.11). The resolution in the z direction is approximately $230\mu\text{m}$ in the TIB and $530\mu\text{m}$ in the TOB, and is varying with pitch in the TID

and TEC. To allow overlay and avoid gaps in acceptance, each module is shifted slightly in r or z with respect to its neighbouring modules within a layer. With this detector lay out, at least nine points per charged particle trajectory can be measured in an $|\eta|$ range up to 2.4, where at least four of them being two dimensional. The CMS silicon tracker provides 9.3 million readout channels and covers an active area of about 198 m^2 .

2.2.3 Data acquisition

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

CMS Level-1 Trigger

The Level-1 Trigger has to be a flexible, maintainable system, capable of adapting to the evolving physics programme of CMS [Khachatryan:2016bia]. Its output rate is restricted to 100 kHz imposed by the CMS readout electronics. It is implemented by custom hardware and selects events containing candidate objects - e.g. ionization deposits consistent with a muon, or energy clusters corresponding to an electron / photon / tau lepton / missing transverse energy / jet. Collisions with large momenta can be selected by using scalar sum of the transverse momenta of the jets.

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

CMS HLT Trigger

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

2.2.4 Phase 1 upgrades

Before the start of taking collision data for 13 TeV operations on 3 June 2015, CMS had a long shutdown (LS1) [Pralavorio:2024977]. During this shutdown, the section of the beryllium beam pipe within CMS was replaced by a narrower one. This operation required the pixel detector to be removed and reinserted into CMS. During Run 2, higher particle fluxes with

respect to Run 1 are expected. To avoid longterm damage caused by the intense particle flux at the heart of CMS, the tracker is been made ready to operate at much lower temperature than during Run 1. The electromagnetic calorimeter preshower system was damaged during Run 1, therefore the preshower discs were removed, repaired and reinstalled successfully inside CMS in 2014. To help the discrimination between interesting low momentum muons coming from collisions and muons caused by backgrounds, a fourth triggering and measurement station for muons was added in each of the endcaps. Several new detectors were installed into CMS for measuring the collision rate within the detector and to monitor beam related backgrounds.

During the LS1, the muon system underwent major upgrades [Guiducci:1966038, Battilana:2239185]. In the fourth station of each endcap, the outermost rings of CSC and RPC chambers were completed, providing an angular coverage of $1.2 < |\eta| < 1.8$ for Run 2, increasing the system redundancy, and allowing tighter cuts on the trigger quality. In order to reduce the environmental noise, outer yoke discs have been placed on both sides for the endcaps. At the innermost rings of the first station, the CSCs have been upgraded by refurbishing the readout electronics to make use of the full detector granularity instead of groups of three as was the case for Run 1. In Figure 2.7 (right), the refurbishing of the CSCs is shown.

Since the HF experiences intense particle fluxes, it became clear during Run 1 that the glass windows of the PMTs need replacing. For the ECAL in Run 1, the energy reconstruction happened via a weighted sum of the digitized samples [Chatrchyan:2013dga]. For Run 2 however, the reconstruction had to be made more resistant for out of time pileup and a multi-fit approach has been set into place. In this approach, the pulse shape is modelled as a sum of one in-time pulse plus the out of time pulses [1748-0221-12-01-C01069]. The energy resolution is better than 2% in the central barrel region and 2-5 % elsewhere.

During the first data taking period of the LHC (2010 to 2013), the tracker operated at $+4^{\circ}\text{C}$. 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:

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

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

During the LS1, the CMS cooling plant was refurbished [running:1998606] and the fluorocarbon cooling system overhauled. To help to suppress the humidity inside the tracker, new methods for vapour sealing and insulation were applied. Furthermore, several hundred high-precision sensors are used to monitor the humidity and temperature. In order to get as dry air as possible, a new dry-gas plant provides eight times more dry gas (air or nitrogen) than during the first run, and allows regulation of the flow. As a final addition, the cooling bundles outside the tracker are equipped with heater wires and temperature sensors in order to maintain safe operations above the cavern dew point. For the data taking in 2015-2016, the tracker operated at -15°C .

687 In Run 2, with the increase in centre of mass energy and a higher luminosity, a larger number
688 of simultaneous inelastic collisions per crossing is expected with respect to Run 1. For this,
689 the CMS Level-1 has been upgraded [[1748-0221-12-03-C03021](#)]. All hardware, software,
690 databases and the timing control system have been replaced for Run 2, where the main changes
691 are that the muon system now uses the redundancy of the muon detector system earlier to make
692 a high resolution muon trigger. Other upgrades are that the calorimeter system isn't bound any
693 more for streaming data and the global trigger has more Level-1 Trigger algorithms.

694 After the first half of Run 2, the innermost part of detection in CMS (pixel detector) was
695 replaced, enhancing the particle tracking capabilities of CMS. The data used in the framework
696 of this thesis however is from before this upgrade. More information on the Pixel upgrade can
697 be found in Refs. [[CMS:2012sda](#), [Perrey:1644757](#)].

698 **2.2.5 CMS computing model**

699 The selected data is stored, processed and dispersed via the Worldwide Large Hadron Collider
700 Computing Grid (WLCG) [[Grandi:814248](#), [Eck:840543](#)]. This has a tiered structure that
701 functions as a single, coherent system.

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

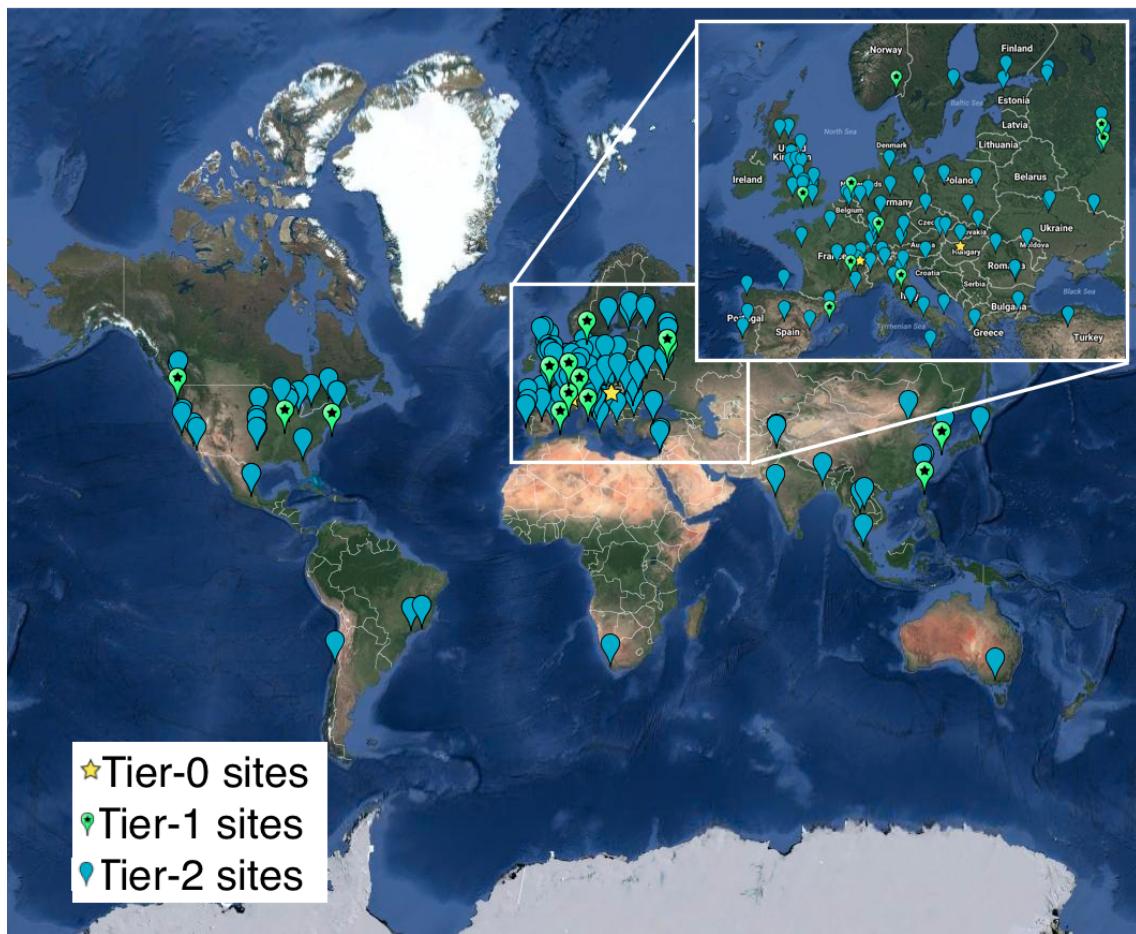


Figure 2.14: Worldwide LHC Computing Grid in 2017 [WLCG].

3

714

Analysis techniques

715 In order to study the collisions coming from high energy experiments, many tools have been
716 developed. In [Section 3.1](#), the physics of hadron collision at high energies are presented. These
717 insights are used to generate events via Monte Carlo event generators, explained in [Section](#)
718 [3.2](#). Machine learning helps to differentiate between signal- and background like events. In
719 [Section 3.3](#), the multivariate technique of boosted decision trees is explained. This yields
720 powerful discriminants for separating signal and background events and provides distributions
721 for template-based maximum likelihood fits. The fitting method used in the search presented
722 in this thesis is discussed in [Section 3.4](#).

723 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 $\hat{\sigma}_{ij \rightarrow X}$ [[Collins:1989gx](#)]

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

724 where i and j are the partons resolved from protons A and B. The parton density functions
725 (PDF) are denoted as $f_i(x_j, Q^2)$, and Q^2 is the factorisation scale more commonly denoted as
726 μ_F . This factorisation scale represents the energy at which the hadronic interaction can be
727 expressed as a product of the partonic cross section and the process independent PDF. In [Figure](#)
728 [3.1](#), the kinematic regions in x and μ_F are shown for fixed target and collider experiments.

729 The parton density functions (PDF) [[Placakyte:2011az](#), [Ball2015](#), [Butterworth:2015oua](#)]
730 represent the momentum distribution of the proton amongst its partons at an energy scale
731 μ_F . These functions are obtained from global fits to data since they can not be determined

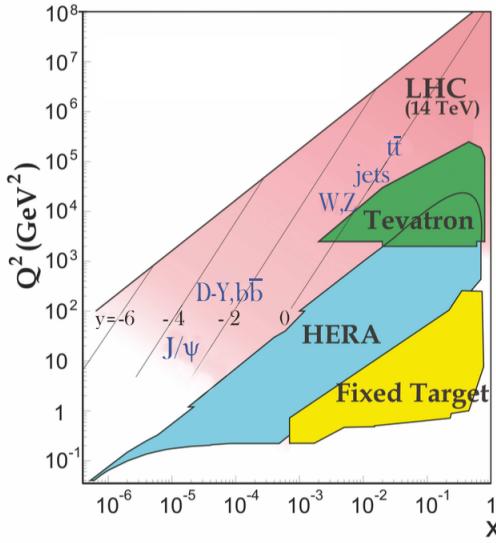


Figure 3.1: Kinematic regions in momentum fraction x and factorisation scale Q^2 probed by fixed-target and collider experiments. Some of the final states accessible at the LHC are indicated in the appropriate regions, where y is the rapidity. In this figure, the incoming partons have $x_{1,2} = (M/14\text{TeV})e^{\pm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

from first principles. From measurements on deep inelastic scattering using lepton-proton collision by the HERA collider [Abramowicz:1998ii], supplemented with proton-antiproton collisions from the Tevatron [Holmes:2011ey], and proton collision data from the ATLAS, CMS and LHCb collaborations at the LHC (Run 1) [Rojo:2015acz] the PDFs are determined and included in global PDF sets known as the PDF4LHC recommendation [Butterworth:2015oua]. Their measurement at scale μ_F is extrapolated to higher energies by use of the DGLAP equations [Martin:2008cn]. Once these PDFs are known, the cross section of a certain process can be calculated and used as input for the Monte Carlo generators used to make the simulated data samples at the LHC. In the framework of this thesis, the NLO PDF4LHC15_100 set is used. This set is an envelope of three sets: CT14, MMHT2014 and NNPDF3.0 [Butterworth:2015oua]. As illustration, the dependency of the PDFs on the momentum fraction x is shown for the NNPDF3.0 set on hadronic scale $\mu_F^2 = 10\text{GeV}^2$ and LHC scale $\mu_F^2 = 10^4\text{GeV}^2$ in Figure 3.2. The gluon density dominated for most values of the momentum fraction, implying that it is easier to probe gluons than the quarks. When the Björken scale is to one, the parton densities of the valence quarks of the proton, up and down quarks, dominate over the gluon density. The sea quarks originating from gluon splitting, the charm, anti-up, and anti-down quarks, have lower densities in general for the proton. The resolution scale Q^2 is typically taken to be the energy scale of the collision. For the top quark pair production a scale of $Q^2 = (350\text{ GeV})^2$ is chosen, meaning that the centre-of-mass energy of the hard interaction is about twice the top quark mass. The uncertainty on the parton distributions is evaluated using the Hessian technique [Pumplin:2001ct], where a matrix with a dimension identical to the number of free parameters needs to be diagonalised. In the case of PDF4LHC15_100 set, this translates into

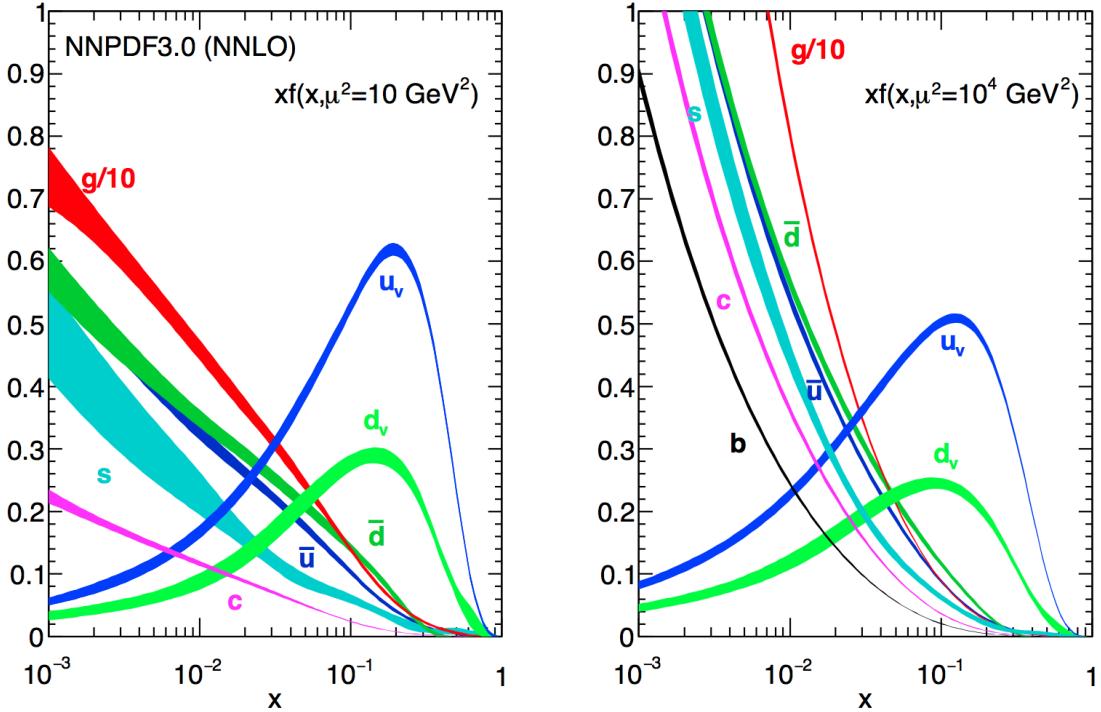


Figure 3.2: The momentum fraction x times the parton distribution functions $f(x)$, where $f = u_v, d_v, \bar{u}, \bar{d}, s, c$, or g as a function of the momentum fraction obtained in the NNLO NNPDF3.0 global analysis at factorisation scales $\mu^2 = 10 \text{ GeV}^2$ (left) and $\mu^2 = 10^4 \text{ GeV}^2$ (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 [PDG].

754 100 orthonormal eigenvectors and 200 variations of the PDF parameters in the plus and minus
 755 direction.

Quantum fluctuations can cause divergences at high energies. This is solved by introducing a renormalization scale μ_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 μ_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 constant¹ α_S is found to be

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

756 with n_f the number of quarks and μ_0 the reference scale at which the coupling is known.
 757 The current world average of the strong coupling constant at the Z boson mass is $\alpha_S(\mu_R = m_Z) = 0.1181 \pm 0.0011$ [PDG]. From Equation 3.3 one can see easily that the coupling
 758

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

759 strength decreases with increasing renormalization scale, this is known as asymptotic freedom.
 760 Additionally, following the behaviour of $\alpha_S(\mu_R^2)$, a limit $\Lambda_{\text{QCD}} \approx 200 \text{ MeV}$ is found for which α_S
 761 becomes larger than one. Under this limit, the perturbative calculations of observables can no
 762 longer be done.

Cross sections can be written in terms of interacting vertices contributing to the matrix element (ME) originating from elements of a perturbative series [Mandl:1236742], allowing them to be expanded as a power series of the coupling constant α

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

763 Leading order (LO) accuracy contains the minimal amount of vertices in the process, then
 764 depending on where the series is cut off one speaks of next-to-leading order (NLO), or next-to-
 765 next-to-leading order (NNLO) accuracy in α . Predictions including higher order corrections
 766 tend to be less affected by theoretical uncertainties originating from a variation of the chosen
 767 renormalization and factorisation scales.

768 3.2 Event generation

769 In order to compare reconstructed data with theoretical predictions, collision events are gen-
 770 erated and passed through a simulation of the CMS detector and an emulation of its readout.
 771 For the detector simulation, a so-called Full Simulation package [1742-6596-396-2-022003,
 772 1742-6596-664-7-072022] based on the Geant4 toolkit [AGOSTINELLI2003250] is employed.
 773 This allows detailed simulations of the interactions of the particles with the detector material.

774 3.2.1 Fundamentals of simulating a proton collision

775 The generation of $\text{pp} \rightarrow \text{X}$ events is subdivided into sequential steps [Seymour:2013ega,
 776 Sjostrand:2009ad, Hoche:2014rga], as shown in Figure 3.3.

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

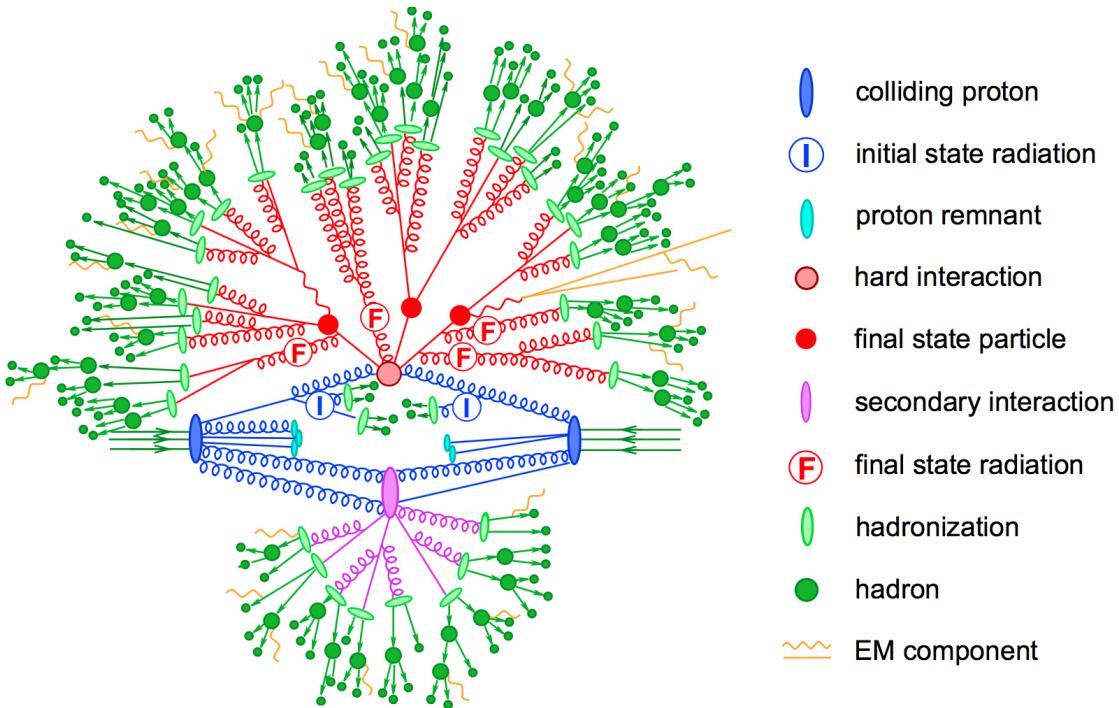


Figure 3.3: Sketch of a hadron collision as simulated by a Monte-Carlo event generator. The red blob in the centre represents the hard collision, surrounded by a tree-like structure representing Bremsstrahlung as simulated by parton showers. The purple blob indicates a secondary hard scattering event. 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 [Hoche:2014rga].

3.2.2 Programs for event generation

The FEYNRULES package [Alloul:2013bka] allows for the calculation of the Feynman rules in momentum space for any quantum field theory model. By use of a Lagrangian, the set of Feynman rules associated with this Lagrangian is calculated. Via the Universal FeynRules Output (UFO) [Degrande:2011ua] the results are then passed to matrix element generators.

The MadGraph program [Alwall:2011uj] is used to interpret the physics model and calculate the corresponding Feynman diagrams and matrix elements. After this, MadEvent [Mangano:2006rw] is used to calculate the corresponding partons. These generated parton configurations are then merged with Pythia [Sjostrand:2015l59, Sjostrand:2006za, Sjostrand:2014zea] parton showers using the MLM merging scheme [Alwall:2007fs].

The MadGraph5_aMC@NLO program [Alwall:2014hca] combines the LO MadGraph [Alwall:2011uj] and the aMC@NLO program into a common framework. This combination supports the generation of samples at LO or NLO together with a dedicated matching to parton showers using the MC@NLO [Frixione:2002ik] or FxFx [Frederix:2012ps] schemes respectively. The FxFx scheme produces a certain fraction of events with negative weights originating from the subtraction of amplitudes that contain additional emissions from the NLO matrix element to prevent double-counting.

808 The POWHEG box (versions 1 and 2) [Alioli2010, 1126-6708-2009-09-111, 1126-6708-2007-11-070] contains predefined implementations of various processes at NLO. It applies the POWHEG method for ME- to PS- matching, where the hardest radiation generated from the ME has priority over subsequent PS emission to remove the overlap with the PS simulation.

813 The JHU generator (version 7.02) [Gritsan:2016hjl, Anderson:2013afp, Bolognesi:2012mm, Gao:2010qx] is used to generate the parton level information including full spin and polarization correlations. It is commonly used for studying the spin and parity properties of new resonances such as $ab \rightarrow X \rightarrow VV$, where $V = Z, W, \gamma$.

817 The generation of events from processes involving the production and decay of resonances creates a computational heavy load, especially at NLO. The narrow width approximation assumes 818 that the resonant particle is on-shell. This factorizes the production and decay amplitude, allowing 819 to perform the simulation of the production and decay of heavy resonances like top quarks 820 or Higgs bosons to be performed in separate steps. The MadSpin program [Artoisenet:2012st] 821 extends this approach and accounts for off-shell effects through a partial reweighting of the 822 events. Additionally, spin correlation effects between production and decay products are taken 823 into account.

825 The Pythia program (versions 6 and 8) [Sjostrand2015159, Sjostrand:2006za, Sjostrand:2014zea] generates events of various processes at LO. However more commonly it is only used for its PS simulation and is then used after other LO and NLO event generators to perform subsequent 827 parton showering, hadronisation, and simulation of the underlying event. In this thesis the underlying event tunes [Khachatryan2016] are the CUETP8M2T4, CUETP8M1 and CUETP8M2.

830 The detector response is simulated via the Geant4 [AGOSTINELLI2003250] program. This 831 program tracks the particles through the detector material via a detailed description of the 832 detector and generates several hits throughout several sensitive layers. In addition, the response 833 of the detector electronics to these hits are simulated.

834 3.2.3 Generating FCNC top-Z interactions

835 The FCNC processes are generated by interfacing the Lagrangian in [Equation 1.36](#) with 836 MadGraph5_aMC@NLO by means of the FeynRules package and its Universal FeynRules Output format. The complex chiral parameters are arbitrary chosen to be $f_{Xq}^L = 0$ and $f_{Xq}^R = 1$.
 NOTE: RH³⁷
 and LH³⁸
 gave the same re-³⁹
 sulting vari-⁴⁰
 ables and⁴¹
 RH is easier⁴²
 to simulate⁴³
 since those⁴⁴
 are singlets⁴⁵
 under SU⁴⁶
 (no doublet
 with b)⁴⁷

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 the SM. 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.

847 The signal rates are estimated by use of the MadGraph5_aMC@NLO program for estimating the
 848 partial widths. The anomalous couplings are left free to float for this estimation, and only one
 coupling is allowed to be non-vanishing at a time. The results are presented in [Table 3.1](#).

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

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

849

850 The anomalous single top quark cross sections are calculated by convolution of the hard
 851 scattering matrix elements with the LO order set of NN23L01 [[Ball:2012cx](#)] partons densities.
 852 The NLO effects are modelled by multiplying each LO cross section by a global k -factor. The LO
 853 single top quark production cross section and the global k -factors for the top-Z production are
 854 shown in [Table 3.2](#). The hard scattering events are then matched to parton showers to Pythia
 to account for the simulation of the QCD environment relevant for hadronic collisions.

Table 3.2: Leading order single top quark production cross section at a centre-of-mass of 13 TeV for $pp \rightarrow tZ$ or $\bar{t}Z$, where the new physics scale is given in GeV. The NLO k -factors [[Zhang:2011gh](#)] 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
κ_{tZq}/Λ	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
ζ_{tZq}	tZu	$8.24 \times 10^{-1} \times (\zeta_{tZu})^2$	0.14	1.40
	tZc	$1.29 \times 10^{-1} \times (\zeta_{tZc})^2$	0.50	1.40

855

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 $t\bar{t} \rightarrow (bW^\pm)(X_{\text{qt}})$. The branching ratio $\mathcal{B}(t \rightarrow bW^\pm)$ is assumed to be equal to one and the FCNC branching ratio 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](#), $\Gamma_t^{\text{SM}} = 1.32 \text{ GeV}$ [[Gao:2012ja](#)], 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 $t\bar{t} \rightarrow (bW^\pm)(X_{\text{qt}})$ 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)
κ_{tZq}/Λ	tZ_u	$t\bar{t} \rightarrow (b\ell^+\nu)(\bar{u}\ell^+\ell^-)$	$2.727 \times 10^5 \times (\kappa_{tZu}/\Lambda)^2$
		$t\bar{t} \rightarrow (\bar{b}\ell^-\bar{\nu})(u\ell^+\ell^-)$	$2.727 \times 10^5 \times (\kappa_{tZu}/\Lambda)^2$
	tZ_c	$t\bar{t} \rightarrow (b\ell^+\nu)(\bar{c}\ell^+\ell^-)$	$2.726 \times 10^5 \times (\kappa_{tZc}/\Lambda)^2$
		$t\bar{t} \rightarrow (\bar{b}\ell^-\bar{\nu})(c\ell^+\ell^-)$	$2.726 \times 10^5 \times (\kappa_{tZc}/\Lambda)^2$
ζ_{tZq}	tZ_u	$t\bar{t} \rightarrow (b\ell^+\nu)(\bar{u}\ell^+\ell^-)$	$2.807 \times (\zeta_{tZu})^2$
		$t\bar{t} \rightarrow (\bar{b}\ell^-\bar{\nu})(u\ell^+\ell^-)$	$2.807 \times (\zeta_{tZu})^2$
	tZ_c	$t\bar{t} \rightarrow (b\ell^+\nu)(\bar{c}\ell^+\ell^-)$	$2.807 \times (\zeta_{tZc})^2$
		$t\bar{t} \rightarrow (\bar{b}\ell^-\bar{\nu})(c\ell^+\ell^-)$	$2.807 \times (\zeta_{tZc})^2$

859 **3.2.4 Generating SM background events**

860 The SM tZq sample is generated using the MadGraph5_aMC@NLO generator (version 2.2.2) [Alwall:2016jdl] at leading order accuracy. The t \bar{t} Z and triboson samples were generated using the
 861 MadGraph5_aMC@NLO generator (version 2.2.2), interfaced through the dedicated MC@NLO
 862 matching scheme [Frixione:2002ik]. The WZ+jets and t \bar{t} W, samples are produced with up to
 863 one additional parton at next-to-leading order accuracy using MadGraph5_aMC@NLO (version
 864 2.2.2) and using FxFx approach [Frederix:2014hta] for matching and merging. Other minor
 865 background are simulated using different generators. The samples of t \bar{t} H, WW, ZZ, and single
 866 top quark production channels are generated with the POWHEG box (versions 1 and 2) [Alioli2010,
 867 1126-6708-2009-09-111, 1126-6708-2007-11-070, Alioli:2010xd, Frixione:2007vw, Nason:2004rx]
 868 The JHU generator [Gritsan:2016hjl, Anderson:2013afp, Bolognesi:2012mm, Gao:2010qx]
 869 is used for the tqH sample, while the tWZ sample is generated using MadGraph [Alwall:2011uj]
 870 interfaced with the MLM matching scheme [Alwall:2007fs]. All events are interfaced to
 871 Pythia version 8.22 [Sjostrand:2014zea] to simulate parton shower, hadronisation, and un-
 872 derlying event. Additionally, MadSpin is used for the tZq, WZ+jets, t \bar{t} Z, t \bar{t} W, tWZ, and triboson
 873 samples

875 The complete list of SM samples is given in Table 3.4, along with their cross sections at a
 876 centre-of-mass of 13 TeV. The cross sections without a reference are coming from the generator
 877 with which the sample has been made, for some of them the uncertainties are provided by
 878 the Generator Group [generator]. For each MC sample, the integrated luminosity that the
 879 sample represents is estimated as the number of simulated events divided by the cross section
 880 of the generated process. This luminosity is then matched to integrated luminosity of 35.9 fb^{-1}
 881 represented by the data used for analysis. For processes generated with MadGraph5_aMC@NLO,
 882 the effective number of simulated events is used, taking into account positive and negative event
 883 weights. In Figure 3.4, a summary is given of the SM cross section measurements performed by
 884 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 [Sjostrand:2014zea] for all samples.

Process	Generator	Cross section (pb)	C	Ref.
$WZ \rightarrow 3\ell\nu$	MadGraph5_aMC@NLO+MadSpin	5.26	1.61	[generator]
tZq with $Z \rightarrow \ell^+\ell^-$	MadGraph5_aMC@NLO+MadSpin	0.0758	3.77	[generator]
tqH with $H \rightarrow ZZ \rightarrow \ell^+\ell^-\ell^+\ell^-$	JHU	$8.80 \cdot 10^{-6}$	-	[generator]
$t\bar{t}W + \text{jets}$ with $W \rightarrow \ell\nu$	MadGraph5_aMC@NLO+MadSpin	0.2043 ± 0.0020	1.94	[generator]
$t\bar{t}Z \rightarrow 2\ell + 2\nu + \text{other}$, with $m_{\ell\ell} > 10$ GeV	MadGraph5_aMC@NLO+MadSpin	0.2529 ± 0.0004	2.15	[generator]
$t\bar{t}H, \text{no } b\bar{b} \text{ decays}$	POWHEG	0.2151	-	[generator]
$t\bar{t}H, b\bar{b} \text{ decays}$	POWHEG	0.2934	-	[generator]
$WW \rightarrow 2\ell 2\nu$	POWHEG	12.178	-	[Gehrmann:2014fva]
$ZZ \rightarrow 4\ell$	POWHEG	0.3366	-	[generator]
WZZ	MadGraph5_aMC@NLO+ MadSpin	0.05565	1.14	[generator]
ZZZ	MadGraph5_aMC@NLO	0.01398	1.17	[generator]
single top quark tWZ , with $Z_\mu \rightarrow \ell^+\ell^-$	MadGraph5_aMC@NLO(LO)+MadSpin	0.001123	-	[generator]
single top quark t-channel \bar{t}	POWHEG+MadSpin	$44.33^{+1.76}_{-1.49}$	-	[generator]
single top quark t-channel t	POWHEG+MadSpin	$26.38^{+1.32}_{-1.18}$	-	[generator]
single top quark $\bar{t}W$	POWHEG	$35.85 \pm 0.90 \text{ (scale)} \pm 1.70 \text{ (PDF)}$	-	[generator]
single top quark tW	POWHEG	$35.85 \pm 0.90 \text{ (scale)} \pm 1.70 \text{ (PDF)}$	-	[generator]
$t\bar{t}$	POWHEG	$831.76^{+19.77+35.06}_{-29.20-35.06}$	-	[generator]
$Z/\gamma^* + \text{jets}$, with $m_{\ell\ell} > 50$ GeV	MadGraph5_aMC@NLO	$3 \times (1921.8 \pm 0.6 \pm 33.2)$	1.49	[generator]
$Z/\gamma^* + \text{jets}$, with $10 \text{ GeV} < m_{\ell\ell} < 50 \text{ GeV}$	MadGraph	18610	-	[generator]

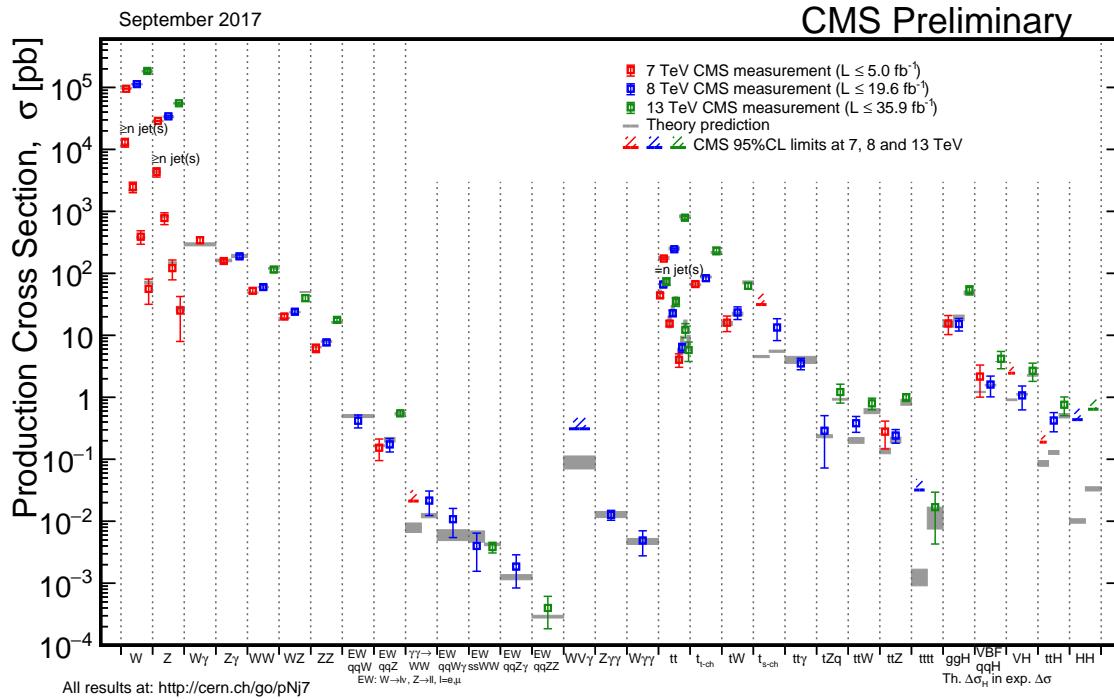


Figure 3.4: Summary of the SM cross section measurements performed by the CMS collaboration. Figure taken from [\[summarywiki\]](#)

885 3.3 Multivariate analysis techniques: Boosted Decision Trees

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

895 There are many software tools that exist for MVA. In this thesis the Tool for Multivariate
 896 Analysis (TMVA) [2007physics3039H] is used. This software is an open source project in-
 897 cluded into ROOT [Brun:1997pa]. By training on events for which the classification is known,
 898 a mapping function is determined that describes a classification or an approximation of the
 899 underlying behaviour defining the target value (regression). In this thesis boosted decision
 900 trees (BDT) are employed for the classification of events as implemented in the TMVA frame-
 901 work [2007physics3039H]. This multivariate technique is based on a set of decision trees where
 902 each tree yields a binary output depending on the fact that an event is signal- or background-like.
 903 This has as advantage that several discriminating variables can be combined into a powerful

904 one-dimensional discriminant D.

905 The decision tree is constructed by training on a dataset for which the outcome is already
 906 provided, such as simulation datasets with signal and background processes (supervised learn-
 907 ing). Different trees can be combined into a forest where the final output is determined by the
 908 majority vote of all trees, so-called boosting. This stabilises the decision trees against statistical
 909 fluctuations and makes it possible to keep the decision trees very shallow, making the method
 910 more robust against overtraining. Examples of such boosting algorithms are Adaptive Boosting
 911 (AdaBoost) and Gradient Boosting [2014arXiv1403.1452M]. In this thesis Gradient boost is
 912 used with a learning rate of 0.2-0.3 and the depth of the tree is set to three. Additionally, the
 913 Gradient boost is used in combination with bagging, so-called stochastic gradient boosting. Bag-
 914 ging smears the statistical fluctuations in the training data and therefore stabilises the response
 915 of the classifier and increases the performance by eliminating overtraining. More information
 916 about stochastic gradient boosting can be found in Ref. [Behnke:2013:DAH:2564838].

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

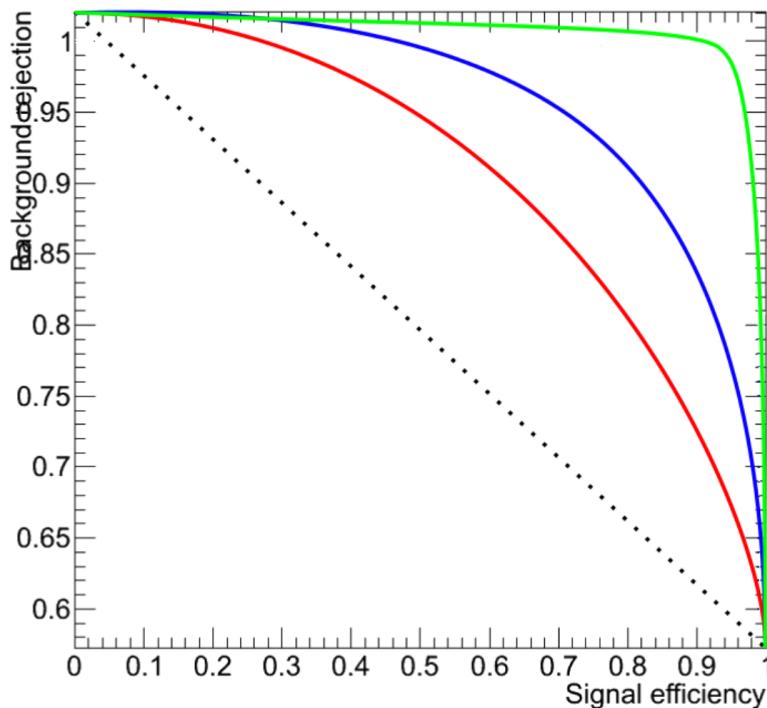


Figure 3.5: Example of ROC curves. In this example, the green method is better than the blue one,
 which is better than the red one. The dashed line represents a case where there is no separation.
 Figure taken from [ROC].

923 **3.4 Statistical methodology**

924 The search performed in the framework of this thesis requires the simultaneous analysis of
 925 data from different decay channels. The statistical methodology used for this search is de-
 926 veloped by the ATLAS and CMS collaborations in the context of the LHC Higgs Combination
 927 group [[Chatrchyan:2012tx](#), [Cowan:2010js](#), [CMS-PAS-HIG-12-020](#), [CMS-NOTE-2011-005](#)].
 928 The Higgs Combined Tool [[HiggsCombine](#)] is a RooStats [[Moneta:2010pm](#)] framework
 929 which runs different statistical methods. In this section, only the statistical tools necessary for
 930 the performed search are described [[CLs](#)].

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

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

In this thesis, the modified frequentist approach [[JUNK1999435](#), [0954-3899-28-10-313](#)] for confidence levels that adopts the classical frequentist method to allow nuisance parameters, is used. It constructs a likelihood $\mathcal{L}(\text{data} | \mu, \theta)$ as

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

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

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

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

947 where $\pi_\theta(\theta)$ is the hyper prior for the (imaginary) measurements. The the pdfs used by the
 948 Higgs Combine Tool are described in Ref. [[CMS-NOTE-2011-005](#)].

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 μ 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_μ^{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_μ 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 μ is achieved by adjusting μ 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 μ 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

964 $\pm 1\sigma$ (68%) and $\pm 2\sigma$ (95%) bands on the expected limit are defined by the crossings of the
965 16% and 84%, and 2.5% and 97.5% quantiles.

966 In order to significantly reduce computing time, the Asymptotic CL method is used. This
967 method avoids an ensemble of toy Monte Carlo samples and instead replaces it by one represen-
968 tative dataset, called Asimov dataset. This dataset is constructed such that all observed quantities
969 are set equal to their MLE values ($\hat{\theta}_{\text{Asimov}} = \theta_0$). More information about this procedure can be
970 found in Refs. [Cowan:2010js].

Event reconstruction and identification

4

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

979 4.1 Object Reconstruction

980 In [Figure 4.1](#), the particle interaction in a transverse slice of the CMS detector is shown. When
981 a particle enters the detector, it first enters the tracker where charged particle trajectories,
982 so-called tracks, and origins, so-called vertices, are reconstructed from signals or hits in the
983 sensitive layers. The magnetic field bends the charged particles making it able to measure the
984 electric charges and momenta of charged particles. The electron and photons are absorbed in
985 the ECAL and the corresponding electromagnetic showers are detected as clusters of energy
986 in adjacent cells. From this, the energy and the direction of the particles can be determined.
987 The charged and neutral hadrons can also initiate a hadronic shower in the ECAL that is fully
988 absorbed in the HCAL. The clusters from these showers are also used to estimate the energy
989 and direction. Muons and neutrino's pass through the calorimeters without little to no energy
990 loss and the neutrinos even escape the CMS detector undetected while muons produce hits in
991 the muon detectors.

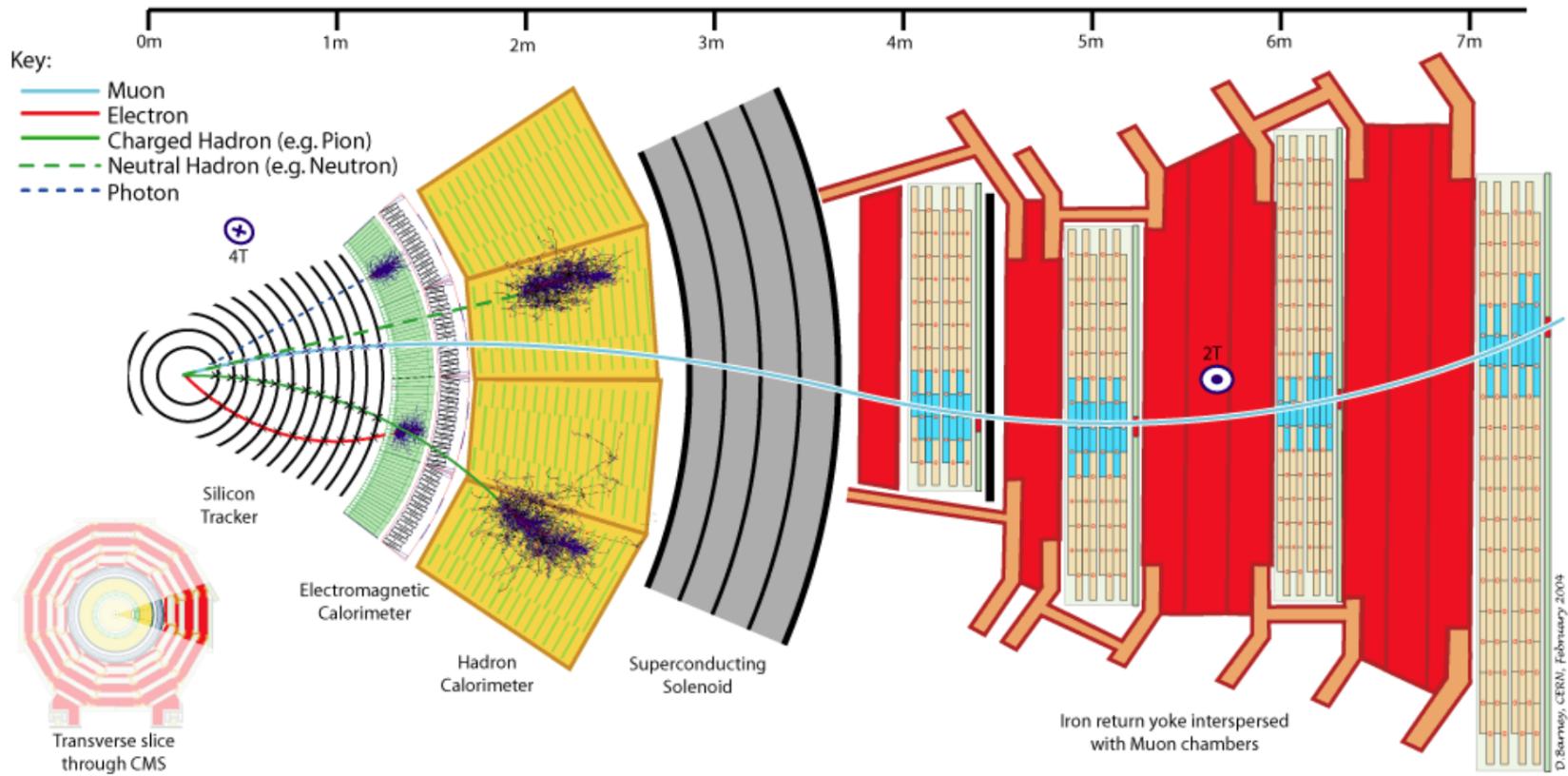


Figure 4.1: Cross-section of the CMS detector with all parts of the detector labelled. This sketch shows the specific particle interactions from a beam interaction reign to the muon detector. The muon and charged pion are positively charged, the electron is negatively charged. Figure taken from [CMS-PRF-14-001].

992 The particle flow (PF) [CMS-PRF-14-001] reconstruction algorithm correlates the tracks and
 993 clusters from all detector layers with the identification of each final state particle, and combines
 994 the corresponding measurements to reconstruct the properties. The muon is identified by a
 995 track in the inner tracker connected to a track in the muon detector as described in [Section](#)
 996 [4.1.2](#). The electrons are identified by a track and an ECAL cluster, not connected to an HCAL
 997 cluster as described in [Section 4.1.3](#). The ECAL and HCAL clusters without a track link identify
 998 the photons and neutral hadrons, while the addition of the tracker determines the energy and
 999 direction of a charged hadron.

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)

1000 **4.1.1 Charged particle tracks**

1001 An iterative tracking algorithm is responsible for the reconstruction of the tracks made by charged
 1002 particles in the inner tracking system. Each iteration consists of four steps [[Bayatian:922757](#)]:
 1003 the track-seed generation, the pattern recognition algorithm, removal of track-hit ambiguities
 1004 and a final track fit. The pattern recognitions done by use a Kalman filter method [[FRUHWIRTH1987444](#),
 1005 [Billoir:1989mh](#)] which into account the magnetic field and multiple scattering effects. All
 1006 hits that are unambiguously associated to the final track are removed from the list of available
 1007 hits. In order to associate the remaining hits, the procedure is repeated with looser track
 1008 reconstruction criteria. The use of the iterative track reconstruction procedure has a high track
 1009 finding efficiency, where the fake track reconstruction rate is negligible.

1010 **4.1.2 Following the Muon's Footsteps**

1011 The muon reconstruction [[Chatrchyan:2012xi](#)] has three subdivisions: local reconstruction,
 1012 regional reconstruction and global reconstruction. The local reconstruction is performed on
 1013 individual detector elements such as strip and pixel hits in the inner tracking system, and muon
 1014 hits and/or segments in the muon chambers. Independent tracks are reconstructed in the inner
 1015 tracker - called tracker tracks - and in the muon system, called standalone muon tracks. Based
 1016 on these tracks, two reconstructions are considered: Global Muon reconstruction and Tracker
 1017 Muon reconstruction. The first is an outside in approach starting from a standalone muon track
 1018 while the second uses an inside-out approach starting from tracker tracks. For low transverse
 1019 momenta ($p_T \lesssim 5$ GeV), the tracker muon reconstruction is more efficient than the global muon
 1020 approach. This is due to the fact that tracker muons only require a single muon segment in
 1021 muon system, while the global muon approach requires typically segments in at least two muon
 1022 stations. These tracker muons are used for identifying muons from the hadronisation of b or
 1023 c quarks. The global muon approach typically improves the tracker reconstruction for $p_T \gtrsim$
 1024 200 GeV.

1025 **4.1.3 The path of the Electron**

1026 Standard tracking algorithms are based on Kalman filtering which assume that the energy
 1027 loss is Gaussian distributed. Since the electron tracks are increasingly curved in the mag-
 1028 netic field as a function of its flight distance, these standard tracking algorithms are not
 1029 suitable to fit the electron tracks and different filtering algorithm, the Gaussian sum filter
 1030 (GSF) [[0954-3899-31-9-N01](#)] is used instead.

1031 In CMS, the electrons are reconstructed in two ways. The older ECAL based tracking is
 1032 developed to identify high energetic isolated electrons. This tracking algorithm starts from
 1033 ECAL clusters with a transverse energy above 4 GeV and extrapolates from these cluster the
 1034 position of the hits in the tracker. Another, tracker based algorithm uses all the tracks with
 1035 a p_T higher than 2 GeV found with iterative tracking as seeds. The electron seeds from the
 1036 ECAL- and tracker-based procedures are merged into a unique collection and are then refitted
 1037 by using the summed Gaussian distributions as uncertainty per hit in the track fit. The electron
 1038 efficiency is measured in 8 TeV proton collision data to be better than 93% for electrons with an
 1039 ECAL supercluster energy of $E_T > 20$ GeV [1748-0221-10-06-P06005]. For electrons with an
 1040 $E_T > 25$ GeV in 13 TeV proton collision data, the efficiency is about 96%[CMS-DP-2017-004].

1041 4.1.4 Primary Vertex Reconstruction

1042 The primary vertex (PV) reconstruction is able to measure the location of all proton interac-
 1043 tion vertices in each event consisting of the signal vertex an all vertices from pileup events.
 1044 First tracks are selected to be consistent with being produced promptly in the primary interac-
 1045 tion [Chatrchyan:1704291]. Then the tracks are grouped according to the z coordinate of their
 1046 closest approach to the beam line [726788] and a vertex fitting algorithm [Waltenberger:1166320]■
 1047 is performed. The primary vertex is found as the vertex corresponding to the highest sum of
 1048 squared track transverse momenta and is taken to be the main interaction point. The resolution
 1049 on the primary vertex is about 14 μm in $r\phi$ and about 19 μm in the z direction for primary
 1050 vertices with the sum of the track $p_T > 100$ GeV for 2016 data taking.

1051 4.1.5 Calorimeter clusters

1052 The energy and direction of stable neutral particles such as photons and neutral hadron are recon-
 1053 structed using a cluster algorithm. This algorithm also separates neutral particles from charged
 1054 hadron energy deposits, and reconstructs and identifies electrons and their bremsstrahlung
 1055 photons. Furthermore, the cluster algorithm is contributing to the energy measurements of
 1056 charged hadrons that don't have accurate tracks parameters, e.g. for low quality and high
 1057 transverse momentum tracks. The clustering is performed separately in each subdetector:
 1058 ECAL barrel and endcaps, HCAL barrel and endcaps, and the two preshower layers. The HF has
 1059 no clustering algorithm since the electromagnetic or hadronic components give rise to an HF
 1060 EM or HF HAD cluster.

1061 The clustering algorithm consist of different steps. First seeds are identified when cells have
 1062 an energy larger than the seeding threshold and larger than their neighbouring cells. Then
 1063 topological clusters are made by accumulating cells that share at least a corner with a cell
 1064 already in the cluster and an energy above a cell threshold set to twice the noise level. The third
 1065 step is an expectation maximization algorithm that reconstructs the cluster [CMS-PRF-14-001]
 1066 and assumes that the energy deposits are Gaussian distributed. The calorimeter clusters are
 1067 used for reconstructing photons and neutral hadrons. The clusters that are not in the vicinity of
 1068 the extrapolated charged tracks identified as neutral hadrons or photons. If the energy deposits
 1069 are in vicinity of charged tracks, such is the case for charged hadrons, the neutral particle energy
 1070 deposit is measured as an excess over the charged particle deposit.

1071 4.2 Particle flow identification

1072 The several PF elements from the various CMS subdetectors are connected through a link
 1073 algorithm. This algorithm tests any pair of elements in an event, only considering nearest
 1074 neighbours in the $\eta\phi$ -plane. The quality of the link is determined via the distance between the
 1075 two elements and PF blocks of elements are formed from elements with a direct link or indirect
 1076 link through common elements. The identification and reconstruction follows a particular order
 1077 in each PF block. After each identification and reconstruction the corresponding PF elements
 1078 (tracks and clusters) are removed from the PF block.

1079 The muons are the first to be identified and reconstructed. These are reconstructed if their
 1080 momenta are compatible with corresponding track only momenta. Then the electron and its
 1081 corresponding brehmstrahung photons, are identified and reconstructed by using of the GSF
 1082 tracking. At the same time, the energetic and isolated photons are identified as well. The
 1083 remaining elements in the PF block are subjected to a cross identification of charged hadrons,
 1084 neutral hadrons, and photons that arise from parton fragmentation, hadronisation, and decays
 1085 in jets. The charged hadron candidate is made from the remaining candidates that have a
 1086 charged particle track associated with them. Then the charged particle energy fraction is
 1087 subtracted from the calibrated energy of the linked calorimeter clusters and the remaining
 1088 energy is assigned to the neutral energy. Depending on the excess of neutral energy in the ECAL
 1089 and HCAL clusters, a photon or a neutral hadron is assigned respectively. The pseudorapidity
 1090 range of the inner tracker limits the information on the particles charge to $|\eta| < 2.4$. Outside
 1091 this range a simplified identification is done for hadronic and electromagnetic candidates only.

1092 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 [CMS-PRF-14-001]. 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 [Kornmayer:2039978]. The instantaneous luminosity from recorded process rate R is then determined as

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

1093 where $\sigma_{fid} = \sigma \times A$ corresponds to the fiducial cross section recorded in the luminometer
 1094 acceptance A which is determined using van der Meer scans [CMS-PAS-LUM-17-001]. The
 1095 overall uncertainty on the luminosity measurement is estimated to be 2.5%.

1096 The luminosity is used to infer to number of pileup interactions in data, which can be used to
 1097 corrected the predefined pileup interactions in simulation. Then an event weight can be derived

from the ratio of the distributions of pileup interactions in data and simulation. For 13 TeV collisions, the inelastic cross section is measured to be 71.3 ± 3.5 mb [CMS-PAS-FSQ-15-005]. However a better agreement in data and simulation for the pileup sensitive variables, such as the number of primary vertices, is found with a lower cross section of 69.2 mb with an uncertainty of 4.6%.

4.4 Physics object reconstruction and identification

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

4.4.1 Muons

The muon candidates used for analysis in this thesis correspond to the tight and loose working point. Detailed reports on the performance can be found in [CMS-DP-2017-007].

The tight working point rejects objects wrongly reconstructed as muons from hadron showers that reach the muon system (punch-throughs), by requiring that the global muon fit includes at least one valid hit in the muon chambers for which at least two muon segments in two muon stations are present. Furthermore, the muon tracks should have a global fit yielding a goodness-of-fit of $\chi^2/\text{ndof} < 10$. Requiring at least one pixel hit in the muon track suppresses the decay of muons in flight. Also a minimum of five hits in the tracker is required. Cosmic muons and muons originating from pileup interactions are rejected by constricting the distance of the muon with respect to the primary vertex by putting limits on $d_{x,y} < 2$ mm and $d_z < 5$ mm. Also muons according to the loose muon working point will be used in the thesis. These are either global muons or tracker muons reconstructed from the particle flow muon object. In Table 4.1, the muon requirements for the muons used throughout this thesis are summarised. In Figure 4.2, the muon efficiencies for data and simulation is presented. These efficiencies are estimated from tag-and-probe methods [CMS-DP-2017-007]. Overall, the efficiency is about 95-100%, with two drops due to the crack between the wheels of the DT system. The differences between data and simulation are corrected by applying p_T - and η -dependent scale factors ($\epsilon_{\text{data}}/\epsilon_{\text{MC}}$) 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 ΔR in $\eta\phi$ plane around the lepton direction. Then a summed energy is calculated from the charged hadrons (CH), neutral hadrons (NH), photons (γ), excluding the lepton itself. This sum is then corrected to remove the energy coming from pileup interactions. The relative isolation for muons \mathcal{I}_{μ} is defined as [CMS-PRF-14-001]:

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

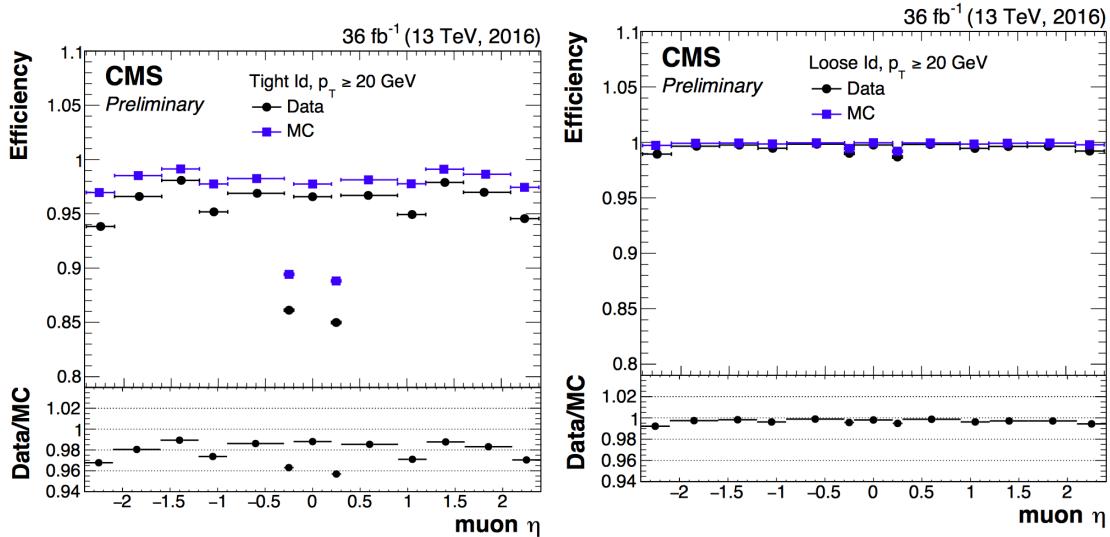


Figure 4.2: Comparison of the muon tight ID (left) and loose ID (right) efficiencies in data and simulation as a function of the pseudorapidity of the muon using the full 2016 dataset. Figure taken from [CMS-DP-2017-007].

1127 where a cone of $\Delta R = 0.4$ is adopted and the pileup mitigation is based on the $\Delta\beta$ correction.
 1128 The $\Delta\beta$ correction estimates the pileup energy as half of the contribution coming from charged
 1129 hadrons. For tight ID muons, this relative isolation should $\mathcal{I}_\mu < 0.15$, while for loose muons
 1130 this should be $\mathcal{I}_\mu < 0.25$. In Figure 4.3, the isolation efficiencies as a function of the pseudo
 1131 rapidities using the tag and probe method are shown for the tight muon working point. The
 1132 efficiencies are 85–100% and have a decline for low- p_T muons. The differences between data and
 1133 simulation are accounted for by applying η - and p_T -dependent scale factors on the simulation.

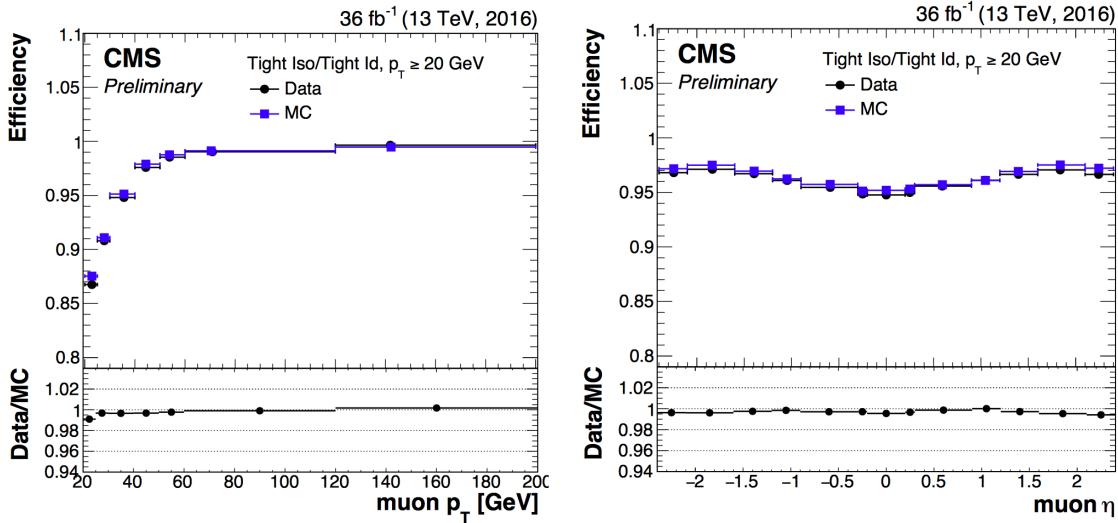


Figure 4.3: Comparison of the muon tight isolation requirement with the muon tight ID efficiencies in data and simulation as a function of the transverse momentum (left) or pseudorapidity (right) of the muon using the full 2016 dataset. Figure taken from [CMS-DP-2017-007].

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
$\chi^2/ndof$ 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. to 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

1134 **4.4.2 Electrons**

1135 The electrons candidates used correspond to the tight and veto working points. The study of the
 1136 electron reconstruction and identification performance can be found in [CMS-DP-2017-004].

1137 Starting from an electron PF candidate with a GSF track that is outside the barrel-endcap
 1138 transition region ($1.4443 < |\eta| < 1.5660$), several requirements are set. The electron track
 1139 should not have more than one (two or three) missing hit in the innermost layer for the tight (veto)
 1140 working point. This dismisses electrons from photon conversions are dismissed. Additionally, a
 1141 photon conversion veto is applied by testing if a pair of electron tracks is originating from a
 1142 common displaced vertex. Furthermore, refined cuts are applied on the shower shape variables
 1143 such as the difference in η or ϕ between the energy weighted supercluster position in the
 1144 ECAL and the track direction in at the innermost tracker position ($\Delta\eta_{\text{in}}$, $\Delta\phi_{\text{in}}$), and the ECAL
 1145 crystal based shower covariance in the η direction ($\sigma_{\eta\eta}$). These cuts also include energy related
 1146 variables such as the absolute difference between the inverse electron energy measured in the
 1147 ECAL and the inverse momentum measured in the tracker ($|1/E - 1/p|$), and the ratio of the
 1148 energy measured in the HCAL and ECAL (H/E). Unlike the muon case, the identification criteria
 1149 also 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 ρ 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 (ρ), computed event by event, and the effective area (A_{eff}) [CMS-PRF-14-001]. This effective area is estimated from simulation and denotes the expected amount of neutral energy from pileup interactions per ρ 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

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

1150

1151 The efficiency of electron identification is estimated from $Z \rightarrow e^- e^+$ events via the tag-and-
 1152 probe method and is shown in Figure 4.4 for the tight working point. The efficiencies reach
 1153 95 – 100%. The difference between data and simulation are corrected by dedicated p_T - and η
 1154 dependent scale factors as well.

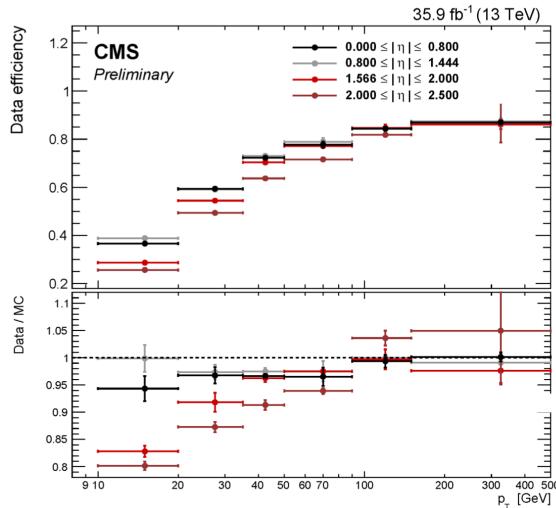
1155 **4.4.3 Jets**

1156 Jets are reconstructed from all reconstructed particles without the charged hadrons associated
 1157 to pileup vertices. The clustering is done via the anti – k_T algorithm [Cacciari:2008gp] with a
 1158 radius parameter for the cone size of the resulting jet of $R = 0.4$. More information about the
 1159 jet algorithm performance can be found in Ref. [CMS-PAS-JME-16-003].

1160 The jets used for the analysis discussed in this thesis uses the loose identification working
 1161 point summarised in Table 4.4. The requirements on the jet constituents are based on the

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

η 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

**Figure 4.4:** Electron identification efficiency as function of the electron transverse momentum from the full 2016 dataset. Figure taken from [CMS-DP-2017-004].

assumption that a proper jet originating from the hadronisation of a quark or gluon consists of multiple PF particles and types. Therefore, the jet should consist of more than one constituent and the neutral hadron fraction and neutral EM energy fractions should be less than 99%. For the jets within the tracker acceptance ($|\eta| < 2.4$), at least one constituent has to be a charged hadron resulting in a charged hadron energy fraction above 0%. Additionally the charged EM energy fraction should be less than 99%. On top of these requirements, objects that are labelled as jets and found in vicinity of any isolated lepton, $\Delta R < 0.3$, are removed from the jet collection in that event to avoid duplications of objects.

The energy of the reconstructed jets deviates from the energies of the corresponding jets clustered from the hadronisation products of true partons from simulations due to non linear subdetector response and efficiencies. The jet energy corrections (JEC) calibrate the jets

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

Properties	$ \eta_{\text{supercluster}} \leq 1.479$		$ \eta_{\text{supercluster}} > 1.479$	
	Veto electron	Tight electron	Veto electron	Tight electron
$\sigma_{\eta\eta}$	< 0.0115	< 0.00998	< 0.037	< 0.0292
$ \Delta\eta_{\text{in}} $	< 0.00749	< 0.00308	< 0.00895	< 0.00605
$ \Delta\phi_{\text{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 GeV^{-1}	< 0.0129 GeV^{-1}	< 0.15 GeV^{-1}	< 0.0129 GeV^{-1}
expected missing inner hits	≤ 2	≤ 1	≤ 3	≤ 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

in order to have the correct energy scale and resolution. Jet energy scale corrections (JES) are determined as a function of pseudorapidity and the transverse momentum from data and simulated events by combining several channels and methods. This is extensively described in [1748-0221-12-02-P02014]. These corrections account for the effects of pileup, the uniformity of the detector response, and residual data-simulation jet energy scale differences. Furthermore, the jet energy resolution (JER) is measured in data and simulation as function of pileup, jet size and jet flavour. The performance of the jet energy corrections for the 13 TeV dataset can be found in [CMS-DP-2016-020].

The JEC are factorised and subsequently correct for the offset energy due to pileup, the detector response to hadrons, and residual differences between data and simulation as a function of

1183 the jet pseudorapidity and transverse momentum. The consecutive steps of JEC are shown
in Figure 4.5. The off set corrections remove the dependence of the jet energy response of

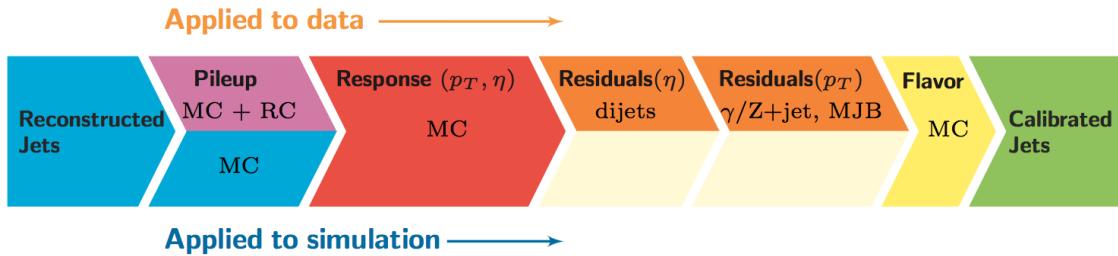


Figure 4.5: The sequence of the JEC for data and simulations The corrections marked with MC are derived from simulation studies, while RC stands for random cone, and MJB for the analysis of multi-jet events. Figure taken from [1748-0221-12-02-P02014].

1184
1185 additional pileup activity. It is based on the jet area method, which uses the effective area
1186 of the jets multiplied by the average density in the event to calculate the offset energy to
1187 be subtracted of the jets. The correction factors are derived by comparing the jet response
1188 with and without pileup events overlaid. The residual differences between data and detector
1189 simulation are determined using the random cone method (RC). For this method, many jets are
1190 reconstructed in each event by clustering particles through placing random cones. This provides
1191 a mapping of the $\eta\phi$ -space and the average p_T of those jets gives the average energy offset due
1192 to pileup [1748-0221-12-02-P02014]. The next level of corrections have as goal to have an
1193 uniform energy response independent of the transverse momentum or pseudorapidity of the
1194 jet. These corrections are determined from simulated events by matching the reconstructed
1195 to true particle jets and comparing their momenta. The residual corrections between data
1196 and simulation are determined by comparing the transverse momentum balance in various
1197 types of events (multi-jet, Z + jets, and γ + jets), using a reference jet in the barrel region. The
1198 jet flavour corrections are optional and not used for this thesis. More information on the jet
1199 flavour corrections can be found in [1748-0221-12-02-P02014]. For jets with a transverse
1200 momentum above 30 GeV, the uncertainties from the various corrections are 3-5% for the 13
1201 TeV dataset [CMS-DP-2016-020].

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 chosen based on whether or not the simulated jet can be matched to a true jet in simulation. The factors are defined as

$$c_{\text{matched}} = 1 + \frac{p_T^{\text{reco.}} - p_T^{\text{true}}}{p_T^{\text{reco.}}} (s_{\text{JER}} - 1),$$

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

1202 where $N(0, \sigma_{\text{JER}})$ denotes a sample value from a normal distribution centred at zero with
1203 as standard deviation the relative resolution in simulation σ_{JER} , and s_{JER} the η -dependent
1204 resolution scale factors. These scale factors are derived from data from di-jet or γ + jets events
1205 and analysing the p_T balance. The resolution scale factors (data/simulation) are found to be

1206 1.1-1.2 except for the transition regions around $|\eta| = 3$ and $|\eta| = 1.4$ [CMS-DP-2016-020] and given in Table 4.5.

Table 4.5: Jet energy scale factors in bins of η 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

1207

1208 4.4.4 Jets from b fragmentation

1209 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.
 1210 There are several algorithms developed within CMS to perform b-tagging [1748-0221-8-04-P04013]
 1211 **CMS-PAS-BTV-15-001**] on jets that fall within the pseudorapidity acceptance of the trackers.
 1212 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
 1213 hard momentum spectra. Additionally, their semi-leptonic decays can be exploited as well. To
 1214 use b-jet identification in an analysis, one needs to know its efficiency and misidentification
 1215 probability. In general these are function of the pseudorapidity and transverse momentum of
 1216 the considered jet. Their performances are directly measured from data by use of b-jet enriched
 1217 jet samples (multi-jet or top-quark decays).

1218 This thesis uses b-jets identified by the Combined Secondary Vertex version 2 (CSVv2)
 1219 algorithm [1748-0221-8-04-P04013]. This algorithm combines secondary vertices together
 1220 with track based lifetime information by use of a multivariate technique. The secondary vertex
 1221 is reconstructed from displaced tracks within a jet, as illustrated in Figure 4.6. The final state b
 1222 quark is encapsulated in a B meson (e.g. B^\pm , B_0 , B_S) after the hadronisation. This B meson has
 1223 relatively long lifetime and can travel a measurable distance from the primary vertex before
 1224 decaying. After reconstruction, the secondary vertices are required to be in accordance with the
 1225 B meson hypothesis bases on the amount of shared tracks with the primary vertex, the invariant
 1226 vertex mass to reject kaon decays, and the direction of the tracks compared to the jet axis.

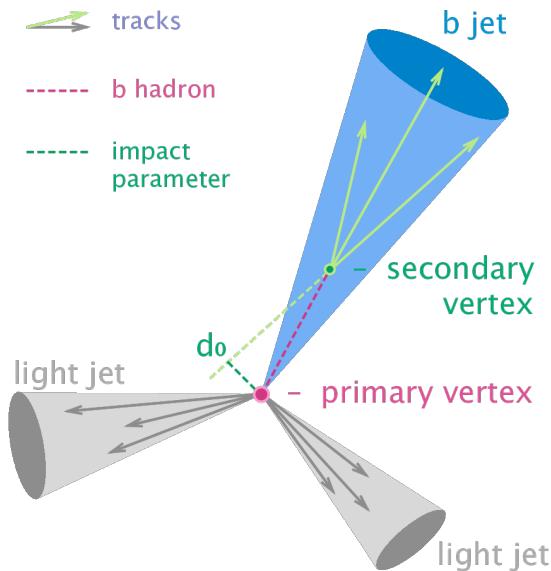


Figure 4.6: Sketch showing the common principle of the identification of b-jets. Figure taken from [btagjet]

1229 The b-tagging algorithm performances are evaluated taking into account two cases: discrimination
 1230 of b-tagged jets originating from charm quarks, and discrimination of b-tagged jets
 1231 against jets coming from gluons or light (u, d, s) quarks. In Figure 4.7, the misidentification
 1232 probabilities for different b-tagging algorithms within CMS are shown. Based on the misidentification
 1233 probabilities for a certain threshold on the CSVv2 discriminator, different working points
 1234 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?

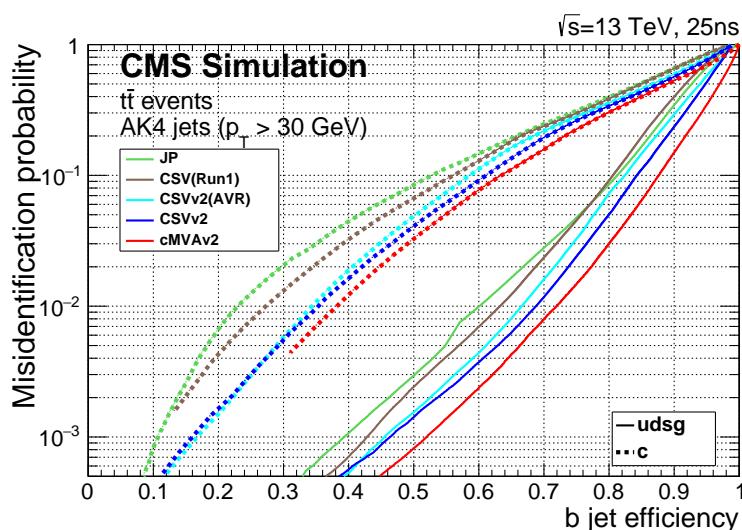


Figure 4.7: Misidentification probabilities of various b-tagging algorithm in simulation. Figure taken from [CMS-PAS-BTV-15-001].

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	$\approx 46\%$	$\approx 0.1\%$

1235

1236 The efficiency of tagging a jet as coming from a bottom quark in simulation typically deviates
 1237 somewhat from data. Efficiency scale factors $\epsilon_b^{\text{data}}/\epsilon_b^{\text{MC}}$ are derived from data to account for
 1238 those differences. These scale factors are η -, p_T -, and flavour dependent, where the flavour
 1239 of the jet is determined from matched generated hadrons. For cut based analyses these scale
 1240 factors are applied to the b-tagging efficiencies and mistag rates according to the chosen working
 1241 point [CMS-PAS-BTV-15-001]. For shape based analysis however, such as the one presented in
 1242 this thesis, the scale factors are applied on the distribution of the b-tagging discriminator. This
 1243 is the so-called IterativeFit method [CMS-PAS-BTV-16-001].

1244 The uncertainties related to the IterativeFit method cover possible shape discrepancies between
 1245 data and simulation. The uncertainty coming from the purity of the sample is subdivided into
 1246 two uncorrelated uncertainties based on the purity of the light flavoured (lf) and heavy flavoured
 1247 (hf) jet contributions in the sample. Furthermore, the jet energy scale results in jets migrating
 1248 from one p_T bin to an other, having an influence on bin dependent scale factors. The statistical
 1249 fluctuations of the limited amount of entries in each bin are also accounted for and have an
 1250 influence on the scale factor uncertainties. These have four uncorrelated sources: two for heavy
 1251 flavour and two for light flavour jets. Since the uncertainty on the scale factors for the jets
 1252 originating from a charm quark (cf) is determined from the uncertainty on the b scale factors
 1253 resulting in two independent uncertainties [CMS-PAS-BTV-16-001].

1254 4.4.5 Missing transverse energy

The missing transverse momentum \vec{p}_T and energy E_T^{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:

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

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

1255 The missing transverse energy is influenced by the minimum thresholds in calorimeters, the
 1256 inefficiencies in the tracker, and the non-linear response of the calorimeter to hadronic particles.
 1257 The bias is reduced by correcting the transverse momentum of the jets to particle jet p_T via the
 1258 JEC and propagating it to the missing transverse momentum. The performance of the missing
 1259 transverse energy reconstruction can be found in [CMS-PAS-JME-16-004].

1260 4.5 Summary of corrections

1261 Throughout the chapter several corrections are introduced to improve the agreement between
 1262 data and simulation. These corrections are sources of systematic uncertainties for the anal-
 1263 ysis presented in this thesis. Therefore a summary of the corrections and their associated
 1264 uncertainties is provided.

1265 **Lepton scale factors** The systematic uncertainty on the lepton scale factors consist of three
 1266 sources: identification, isolation and tracking. The applied scale factors are varied
 1267 independently within one standard deviation of their measured uncertainties to account
 1268 for their systematic impact on the measurements.

1269 **Jet energy corrections** The momenta of the reconstructed jets are corrected to match the expected true energy derived from the hadronisation products of partons in simulation.
 1270 Furthermore, residual corrections and smearing is applied to match the overall energy scale and resolution for simulation and data. These corrections are also propagated to the missing transverse energy. The systematic uncertainties due to these scale factors are estimated by varying them within their uncertainties and repeating the measurements with recalibrated jets and missing transverse energy.

1276 **CSVv2 discriminant shape reweighting** There are three sources of uncertainty contributing to the measurement of the scale factors: statistical uncertainties, jet energy scale and the purity of the sample. The jet energy scale uncertainty is 100% correlated to the jet energy uncertainties and is evaluated simultaneously. The uncertainty coming from the purity of the sample is subdivided into two uncorrelated uncertainties based on the purity of the light flavoured (lf) and heavy flavoured (hf) jet contributions in the sample. A one sigma shift in each of the two purity contributions corresponds to a higher/lower contribution in the purity of the considered flavours. The statistical uncertainties has four uncorrelated sources, two for heavy flavour and two for light flavour jets. One of the uncertainties correspond to the shift consistent with the statistical uncertainties of the sample, while the other is propagated in a way that the upper and lower ends of the distribution are affected with respect to the centre of the distribution. The uncertainty on the charm jet scale factors (cf) is obtained from the uncertainty on the heavy flavour scale factors, doubling it in size and constructing two nuisance parameters to control the charm flavour scale factors and treating them as independent uncertainties.

1291 **Pileup** Varying the minimum bias cross section, used to calculate the pileup distribution by $\pm 4.6\%$, results in a systematic shift in the pileup distribution. The uncertainty is estimated by recalculating the pileup weights to the distributions associated to the minimum bias cross sections.

1295 **Luminosity** The luminosity is measured with a global uncertainty of 2.5%, affecting the ex-
 1296 pected number of events.

Event selection and categorisation

5

1297

1298 A basic event selection is made for selecting signal like events and is discussed in [Section 5.1](#).
1299 Also the effect corrections applied to simulation and data, summarised in [Section 4.5](#) is shown.
1300 The analysis strategy is presented in [Section 5.2](#), defining the signal and background regions
1301 to constrain the huge SM background compared to the expected signal. [Section 5.4](#) discusses
1302 each region that is entering the analysis. On top of the use of background estimation from
1303 control regions, backgrounds that have prompt leptons contaminated by real leptons either from
1304 decays of tau leptons or from hadronized mesons or baryons (collectively commonly referred as
1305 “non-prompt leptons”) as well as by hadrons or jets misidentified as leptons¹ are evaluated with
1306 a data-driven method discussed in [Section 5.3](#).

1307 5.1 Baseline event selection and filters

1308 In this analysis a search is performed in a final state made up of a Z boson and a top quark,
1309 associated or not with a jet, in the leptonic decay of the Z boson and the top quark. The leading
1310 order Feynman diagrams can be seen in [Figure 5.1](#) and [Figure 5.2](#). The signal considers both the
1311 single top quark FCNC (tZ in the final state) and the top quark pair FCNC (tZq with $q = c, u$ in
1312 the final state) events. Their final state signatures consist of three leptons, considering electrons
1313 or muons, and a jet originating from a b quark. For FCNC tZq , there is an additional up or charm
1314 jet. Leptons from tau decays are not vetoed and are entering the analysis via their leptonic
1315 decays. Four different lepton channels based on lepton flavour are considered: 3e, 2e1 μ , 1e2 μ ,
1316 and 3 μ .

1317 The CMS collaboration recorded in the course of 2016 proton at 13 TeV collisions data with
1318 a total recorded integrated luminosity of 35.9 fb^{-1} . The baseline event selection has a goal to
1319 substantially reject SM background events, whilst maintaining a high signal efficiency. The
1320 CMS trigger system, described in [Section 2.2.3](#), filters out the main of the collision events from
1321 uninteresting processes and dedicated trigger paths are defined to single out the events with our
1322 required detector signature. The trigger paths are chosen based on on-line triggering objects
1323 with at least one muon (M), at least one electron (E), at least two muons (MM), at least two

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

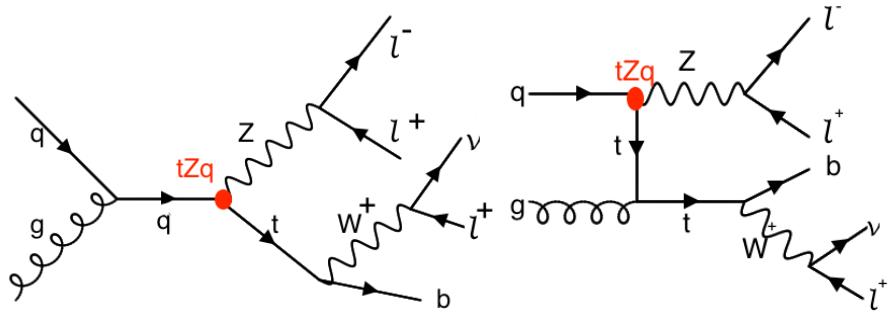


Figure 5.1: Single top quark Feynman diagrams at leading order. The vertex labelled tZq is the sought-for FCNC interaction.

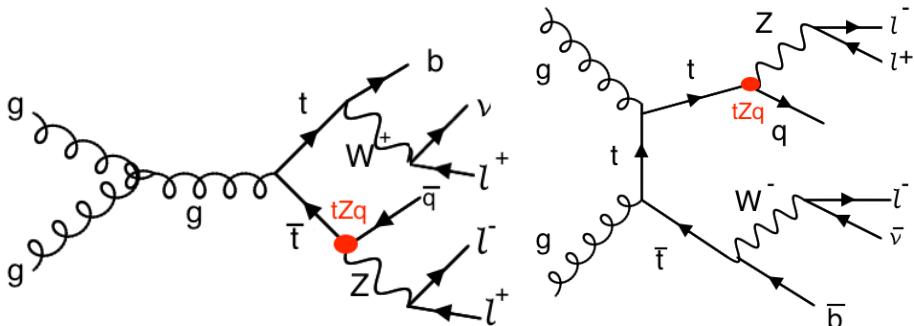


Figure 5.2: Top quark pair Feynman diagram at leading order. The vertex labelled tZq is the sought-for FCNC interaction.

1324 electrons (EE), at least one muon and an electron (ME), at least three muons (MMM), at least
 1325 three electrons (EEE), at least two muons and one electron (MME), or at least two electrons
 1326 and one muon (EEM). For the MC simulation a simple *or* of all triggers is taken and the event
 1327 is considered when it passes one of the trigger paths. For data however, double counting of the
 1328 same event has to be taken into account and a procedure to avoid double counting has been
 1329 put into place. It consists of vetoing in a given dataset the events that are already selected in
 1330 another, as given in Table 5.1.

1331 For the single lepton triggers, at least one electron (muon) with a transverse momentum
 1332 p_T higher than 32 (24) GeV is required. The dilepton triggers require an electron (muon)
 1333 with $p_T > 23$ GeV and a muon (electron) with $p_T > 8$ GeV, or an electron (muon) with
 1334 $p_T > 23$ (17) GeV and an electron (muon) with $p_T > 12$ (8) GeV. Events collected by the
 1335 trilepton triggers require an electron (muon) with $p_T > 16$ (12) GeV, a second electron (muon)
 1336 of $p_T > 12$ (10) GeV, and a third electron (muon) with $p_T > 8$ (5) GeV. The mixed trilepton
 1337 trigger events require two electrons (muons) with $p_T > 12$ (9) GeV and a third muon (electron)
 1338 with $p_T > 8$ (9) GeV. The HLT trigger paths used in data and simulation are summarised in
 1339 Table 5.2.

1340 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^{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

1371 5.1.1 Event cleaning

1372 Some events arising from instrumental noise and beam backgrounds might end up in the
 1373 data [**Filters, CMS-PAS-JME-16-004**]. Spurious deposits may appear in the ECAL from non
 1374 collision origins such as beam halo particles, or from particles hitting the sensors in the ECAL
 1375 photodetectors. Conjointly, dead ECAL cells can cause artificial missing transverse energy.
 1376 Also the HCAL can cause spurious energy from particle interactions with the light guides and
 1377 the photomultiplier tubes of the HF, as well as noisy hybrid photo diodes. In CMS, different
 1378 algorithms, so-called filters, are developed to identify and suppress these events.

1379 The ECAL electronics noise and spurious signals from particle interactions with photo detectors
 1380 are mostly removed via topological and timings based selection using ECAL information only.
 1381 The remaining effects such as anomalously high energy crystals and the lack of information
 1382 for channels due to inefficiencies in the read out are removed through dedicated events filters.
 1383 Five ECAL endcap supercrystal have been identified for giving anomalously high energies due
 1384 to high amplitude pulses in several channels at once, and are masked. Furthermore, the crystal
 1385 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 ϕ 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%.

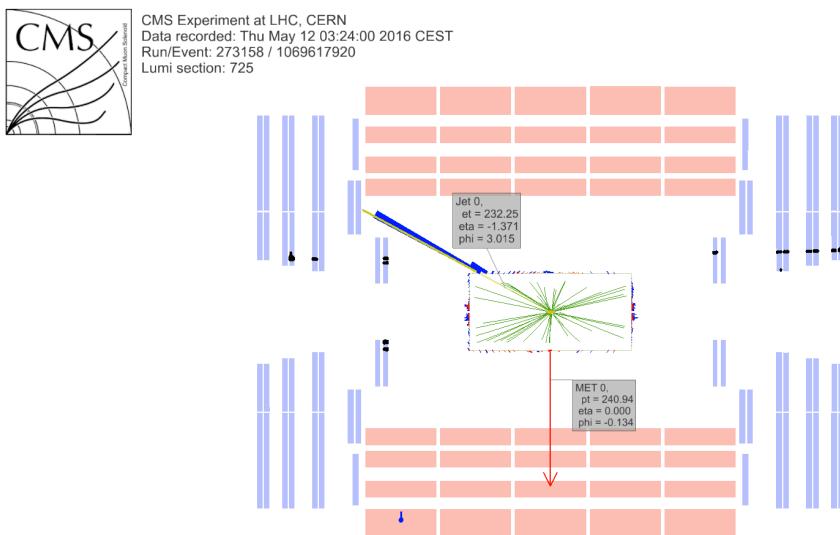


Figure 5.3: Event display of a beam halo event with collinear hits in the CSC (black), missing transverse energy of 250 GeV and a jet of 232 GeV. The hadronic deposit is spread in η , but narrow in ϕ . Figure taken from [CMS-PAS-JME-16-004].

1395

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).

1406 **5.1.2 Estimation of the trigger efficiency**

1407 The trigger efficiency in data is estimated using a data sample collected using unprescaled E_T^{miss}
 1408 triggers. These allow events with a missing transverse energy higher than 110 GeV(120 GeV)
 1409 and that the scalar sum of the transverse momenta of the reconstructed PF jets $H_T^{\text{trig.}}$ is at least
 1410 300 GeV (120 GeV), or a calorimeter (PF) E_T^{miss} higher than 200 GeV(300 GeV). For an HB-HE
 1411 cleaned event the PF missing transverse energy threshold is lowered to 170 GeV. These trigger
 1412 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^{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

1413

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

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

1414 More detailed scale factors and efficiencies can be found in [Appendix A](#).

1415 The trigger efficiencies are also measured in function of the p_T of the leptons for which the
 1416 distributions can be found in [Appendix A](#). The resulting scale factors can be found in Figure
 1417 [Figure 5.4](#). The trigger efficiencies are measured to be nearly 100% for both simulation and
 1418 data. The results are dominated by statistics and assigning a large uncertainty to the trigger
 1419 efficiency based on the dataset collected by E_T^{miss} triggers would be over conservative. A one
 1420 percent uncertainty on the trigger selection for the 2e1 μ and 3 μ final states, and 5% for the
 1421 3e and 1e2 μ final states is assigned instead. No scale factors will be applied on simulation as
 1422 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: 1423
 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 μ	3e	2e1 μ	1e2 μ
	1.0000	1.0000	0.9541	1.0006	1.0004%

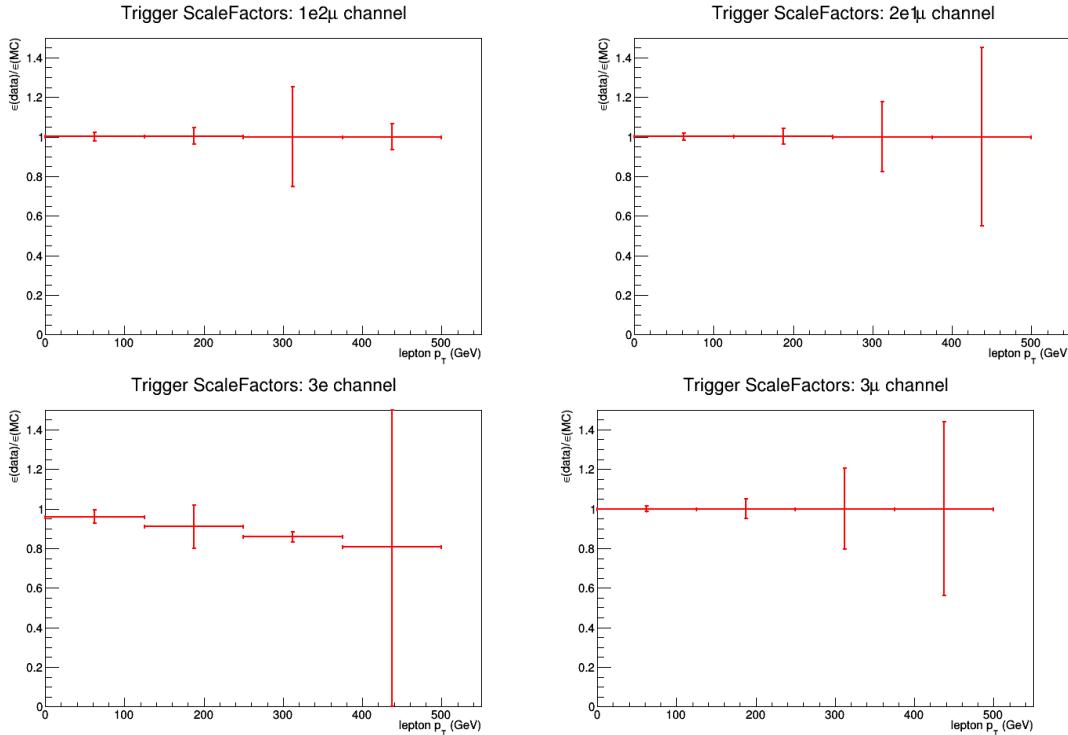


Figure 5.4: The trigger scale factors measured as a function of lepton p_T , using the dataset collected by E_T^{miss} triggers and WZ simulation, after a 3 lepton and jets selection, in the Z mass window. All corrections to simulation are applied. Left, upper: 1e2 μ channel. Right, upper: 2e1 μ channel.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

5.1.3 Corrections

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

Pileup reweighting

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

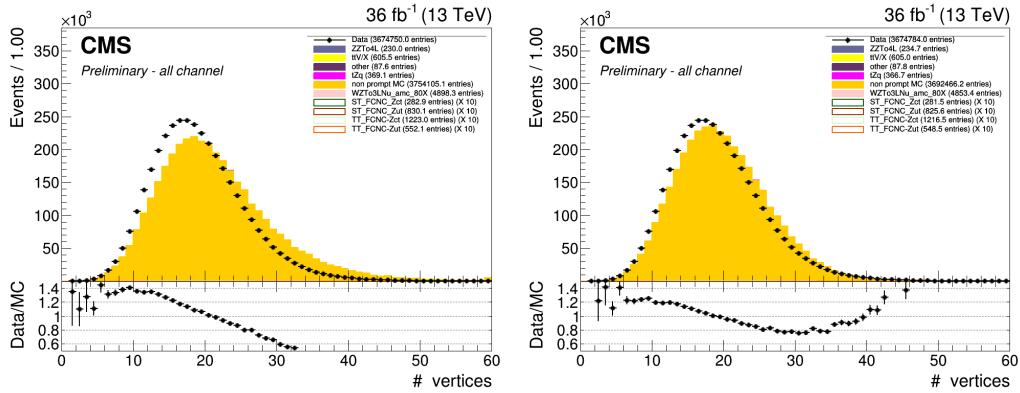


Figure 5.5: The number of primary vertices before (left) and after (right) pileup reweighting. After a dilepton plus jets selection, in the Z mass window.

1433 Note that Figure 5.5 indicates that even after pileup reweighting, the primary vertex multi-
1434 plicity is not well described by simulation. This is a known effect, and using a minimum bias
1435 cross section with a slightly lower value is found to better describe the data. However, the b
1436 tagging scale factors are only provided for the nominal inelastic cross section, and thus this
1437 value is used.

1438 Lepton scale factors

The efficiency to select leptons is different in simulation (ϵ_{MC}) compared to the data (ϵ_{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_{T} and η (see Section 4.4.1 and Section 4.4.2). Multiplying these scale factors for each lepton provides an overall efficiency:

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

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

1439 The effect of the scale factors can be found in Figure 5.7 for the electron scaling and Figure 5.6
1440 for the muons. The trigger efficiencies are estimated in Section 5.1.2.

1441 Additionally, corrections are made for the energy resolution of the leptons. For the electrons,
1442 energy smearing and regression is applied [smearing]. The energy regression uses the detector
1443 information to correct the electron energy in order to have the best energy resolution by
1444 correcting local energy containment, material effects, etc.. The energy scale and smearing

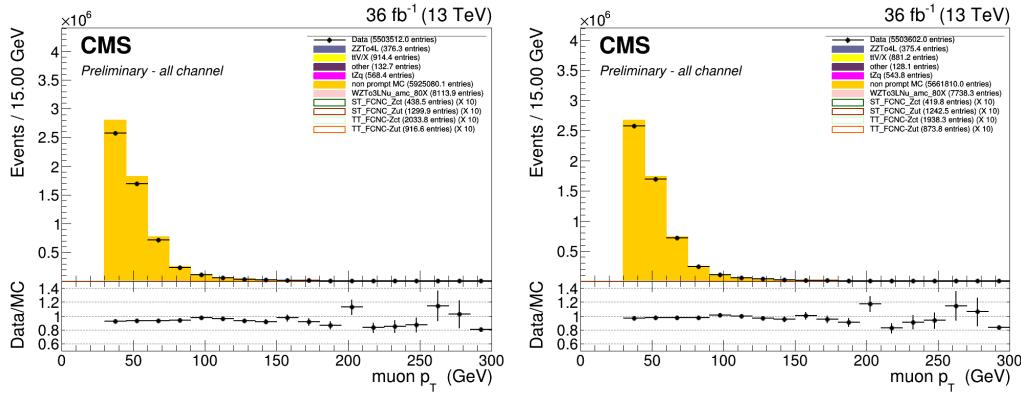


Figure 5.6: The p_T of the muons before (left) and after (right) muon scale factors. After a dilepton plus jets selection, in the Z mass window. Both after the Rochester correction.

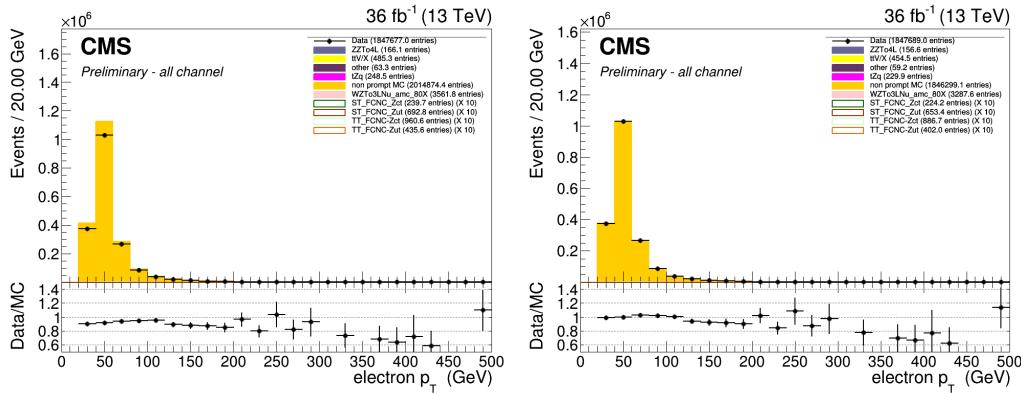


Figure 5.7: The p_T of the electrons before (left) and after (right) electron scale factors. After a dilepton plus jets selection, in the Z mass window. Both after energy scale corrections and smearing.

is done in order to bring the data energy scale to simulation level. It smears the simulation energies to have identical energy resolution in simulation and data. For the muons, the p_T is corrected using the Rochester method [roch, roch2]. This correction removes the bias of the muon p_T from any detector misalignment or any possible error of the magnetic field. The effect of the Rochester correction can be found in Figure 5.8.

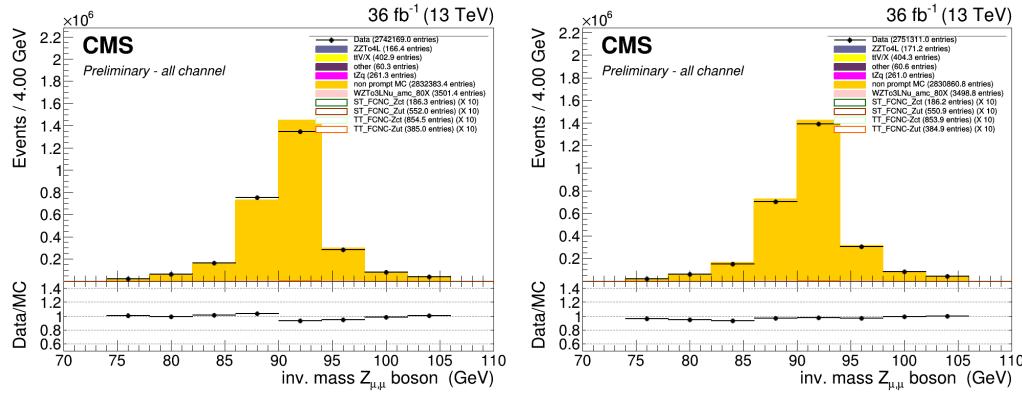


Figure 5.8: The mass of the Z boson consisting of the muons before (left) and after (right) the Rochester correction. After a dilepton plus jets selection, in the Z mass window.

1450 CSVv2 shape correction

1451 In order to make the shape of the CSVv2 b-tagging discriminant in simulation agree with data,
 1452 jet-by-jet based scale factors are applied. These scale factors are a function of the p_T , η and
 1453 CSVv2 value of the jet as discussed in Section 4.4.4. The effect of these scale factors can be
 1454 found in Figure 5.9.

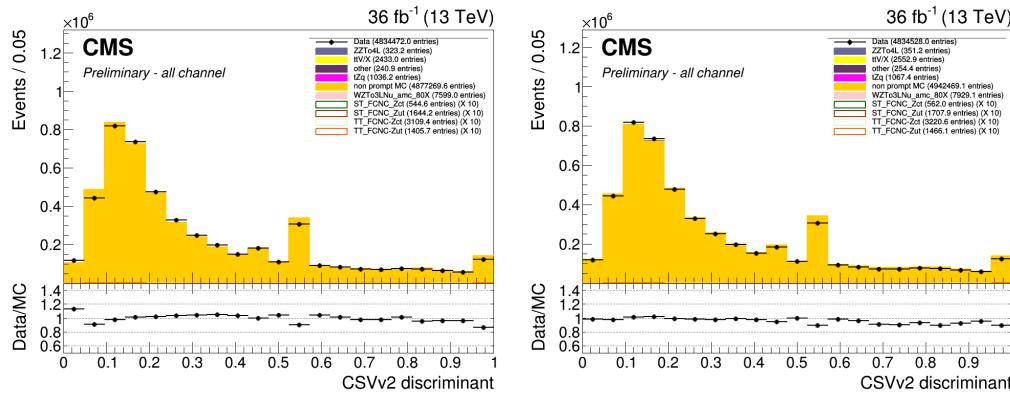


Figure 5.9: The CSVv2 discriminant of the jets before (left) and after (right) b-tag scale factors. After a dilepton plus jets selection, in the Z mass window.

1455 Jet energy

1456 The jet energy in data and simulation is corrected by the measured energy response of the
 1457 detector. This provides p_T - and η dependent scale factors and are directly taken from the
 1458 frontier condition database as discussed in Section 4.4.3. The effect of the jet energy corrections
 1459 can be found in Figure 5.10 and Figure 5.11.

NOTE: Figure 5.9
jet plot

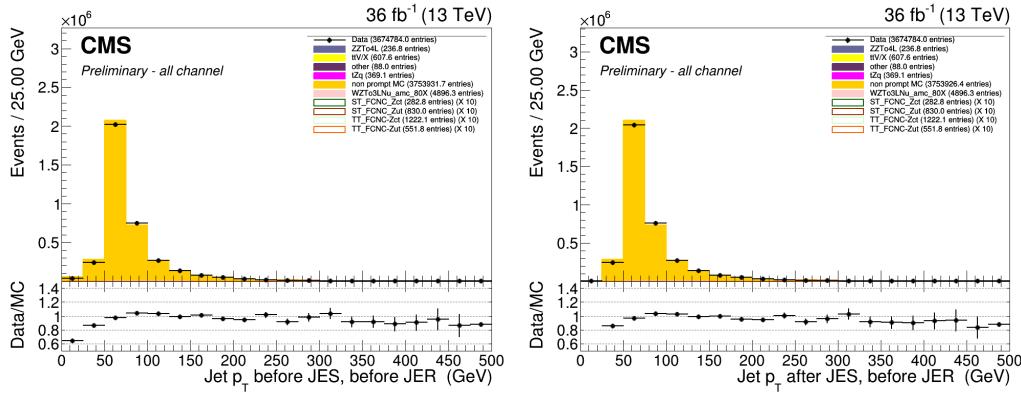


Figure 5.10: The p_T of the jets before (left) and after (right) jet energy scale corrections. After a dilepton plus jets selection, in the Z mass window.

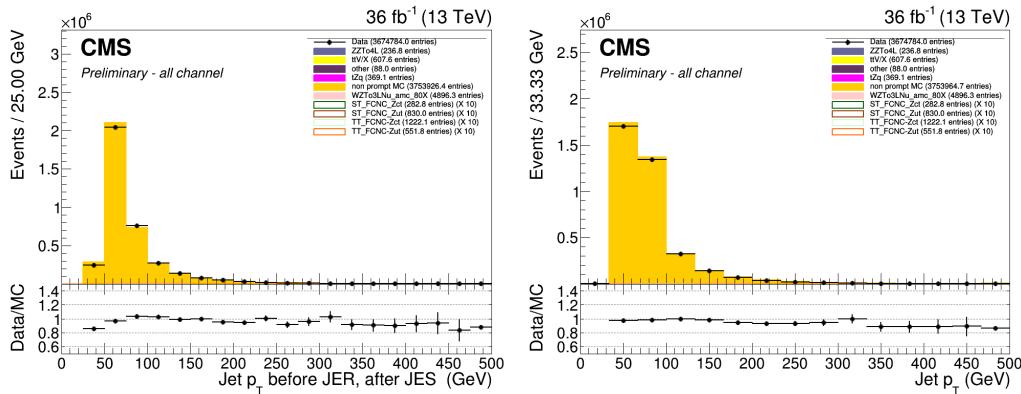


Figure 5.11: The p_T of the jets before (left) and after (right) jet energy resolution smearing. After a dilepton plus jets selection, in the Z mass window.

1460 Missing transverse energy

1461 The energy scale and resolution corrections applied to the jets are propagated back to the
1462 missing transverse energy (smeared Type I correction) as discussed in [Section 4.4.5](#). This
1463 rebalances the transverse net momentum of the event and improves the missing transverse
1464 energy resolution itself.

1465 5.1.4 Reconstruction of kinematic variables

1466 The variables used for the training are related to the reconstructed leptons, jets, Z boson and
 1467 top quark candidates. The Z boson is reconstructed as the sum of the four vectors of the
 1468 two same flavour leptons of opposite sign giving the closest value to the Z mass. The third
 1469 remaining lepton is assigned as the lepton coming from the W boson decay. The reconstruction
 1470 of the SM top quark candidate is more difficult and done by summing the third lepton, the
 1471 SM b-jet and the neutrino (E_T^{miss}). The SM b jet is assigned to the jet with the highest CSVv2
 1472 discriminant. The longitudinal momentum of the neutrino is calculated by putting a constraint

1473 on the lepton+neutrino system with the W mass. In case two solutions are found for the p_z^ν
 1474 component, the one that gives the reconstructed mass (lepton + neutrino + b jet) to the known
 1475 top quark mass is used. The FCNC top quark is reconstructed by summing the reconstructed Z
 1476 boson and the jet giving the mass closest to the top mass, from the jet collection from which the
 1477 SM b jet is removed.

1478 The reconstructed objects are validated using simulation by matching the reconstructed
 1479 objects to their generated counterpart by minimizing ΔR . The efficiencies derived for the
 1480 simulated signal samples and the SM tZq background process can be found in [Table 5.5](#) and
 1481 [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%

1482 5.2 Analysis Strategy

1483 The analysis strategy uses five statistically independent regions to extract limits using a likelihood
 1484 fit of various observables. Two signal regions, the tZ (STSR) and tZq (TTSR) signal regions, are
 1485 constructed using the jet multiplicity, focussed on each signal signature (see Tab. 5.9). In order
 1486 to constrain the rate of WZ+jet events as well as that of NPL backgrounds, three control regions
 1487 are defined. The WZ control region (WZCR) focusses on NPLs originating from Z/γ^* + jets
 1488 and simultaneously constrains the WZ+jets background rate. The NPL backgrounds coming
 1489 from a $t\bar{t}$ process are constrained by two control regions, TTCSR and STCSR, one for each signal
 1490 region (TTSR and STSR). In the STSR and TTSR multivariate discriminants based on Boosted
 1491 Decision Trees (BDT) (see [Section 3.3](#)) are used to respectively discriminate FCNC tZ and FCNC
 1492 tZq from backgrounds. In the WZCR a discriminating variable between the two backgrounds,
 1493 WZ+jets and NPLs, is used. In TTCSR and STCSR the dominating process is the $t\bar{t}$ process, and
 1494 its rate is estimated by subtracting all other background predictions from data. A simultaneous
 1495 global fit using the Higgs Combined Tool ([Section 3.4](#)) is performed taking into account each
 1496 region (STSR, TTSR, WZCR, TTCSR and STCSR) for the four different leptonic channels.

5.3 Data driven NPL background simulation

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

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

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

The NPL sample is constructed from data by requiring exactly three leptons, from which two are considered real, isolated leptons and the third lepton is identified as a NPL. This NPL is created by loosening its identification and inverting its isolation criteria. The full requirements on the not prompt leptons are given in [Table 5.7](#) and [Table 5.8](#). For NPe, a large fraction is coming from misidentified photons. These are removed by applying a tighter cut on the $1/E - 1/p$ 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	≤ 1	≤ 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	≥ 0.15
p_T	$> 30 \text{ GeV}$

1521 **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	≥ 1	1	≥ 2	1	≥ 2
Number of b jets	0	1	≥ 1	1	≥ 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

1522 Additional leptons with a looser identification are vetoed in order to reduce the contamination
 1523 of backgrounds with four or more leptons in the final state, e.g. ZZ, t̄Z, and t̄H. The most
 1524 important backgrounds are the ones that contain three prompt leptons in the final state. These
 1525 are mainly WZ+jets, t̄Z and SM tZq. For these backgrounds, the three lepton topology is
 1526 identical to the FCNC signal: two opposite sign leptons of the same flavour decaying from the Z
 1527 boson, and a third additional, high p_T lepton coming from the W boson decay.

1528 For the FCNC tZ final state, one b jet coming from the SM top decay is expected. For the
 1529 FCNC tZq, an additional light jet is expected. In the t̄Z final state, two b jets are present in the
 1530 final state. However, due to inefficiencies of the b-tagging algorithm, one of the two b jets may
 1531 be identified as a light quark jet, giving the same final state as the FCNC tZq final state. For
 1532 the WZ+jets final states, one of the b jets produced by gluon splitting can be b-tagged or light
 1533 flavour jets coming from the WZ+jets production can be mis-tagged as b jets. The SM tZq final
 1534 state expects the same signal as FCNC tZq.

1535 The NPL events give a significant background contribution. This background is coming mainly
 1536 from Z/ γ^* + jets and t̄t + jets processes (in a less significant way, also WW and tWZ contributes),
 1537 which have very high cross sections and causes a large number of NPL background events
 1538 compared to signal.

1539 In order to reduce the large uncertainties in backgrounds, five independent regions are used
 1540 as defined in Table 5.9. In Figure 5.12, the strategy and usage of each region is illustrated.

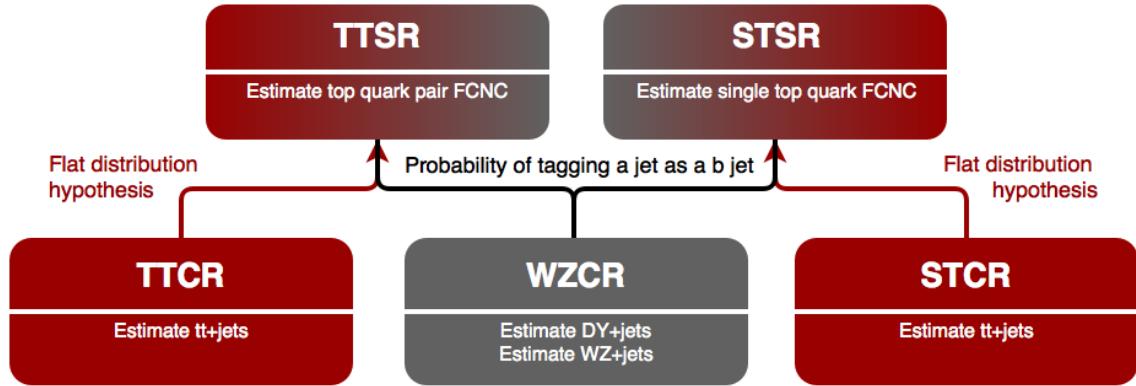


Figure 5.12: The strategy used for this search. The WZCR region is used to estimate the WZ+jets background process as well as the NPL background coming from the $Z/\gamma^* + \text{jets}$ process. The TTCSR and STCSR regions are used to estimate the contributions of the NPL background coming from the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

1541 5.4.1 WZCR

1542 The WZCR is constructed by vetoing events with jets tagged as being a b-jet, making it statistically
 1543 independent from the signal regions where at least one b-tagged jet is required. In this region,
 1544 a fit is performed on the transverse mass of the W boson, in order to estimate the NPL yield
 1545 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

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

1546 with the products going over all b-, c-, and light jets. The jet flavour is determined by means of
 1547 matching the reconstructed jet to the generated quark based on the distance in the $\eta\phi$ plane.
 1548 In order to estimate the probability for exactly one b-tagged jet, the expected number of events
 1549 is corrected by the fraction of events with exactly one jet in the WZCR. The resulting transfer
 1550 factors are given in [Appendix C](#). The yield of WZ+jets events in the signal region estimated
 1551 using the above described transfer factor and the yield calculated with simulated events are in
 1552 agreement.

1553 **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^{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

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

1554 The resulting transfer factors are given in Appendix C. The expected yield in the signal region
 1555 estimated from the TTCR (STCR) is in agreement with the yield calculated from simulated events.

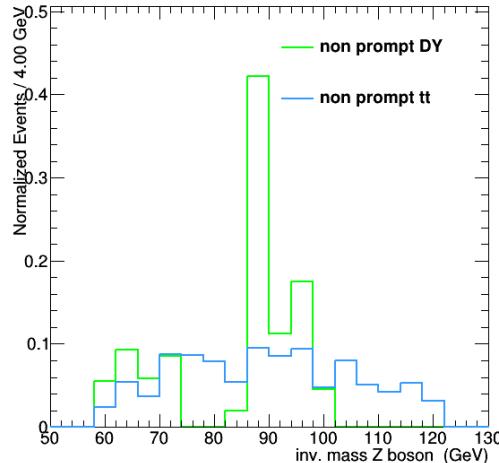


Figure 5.13: The normalized distribution for $Z/\gamma^* + \text{jets}$ and $t\bar{t} + \text{jets}$ events before dividing the events in to regions, after $|m_Z^{\text{reco}} - m_Z| < 30 \text{ GeV}$. All leptonic channels combined.

1556

1557 **5.4.3 TTSR and STSR**

1558 The TTSR is defined to target top quark pair FCNC (tZq), while the STSR focusses on single top
 1559 quark FCNC (tZ). They have NPL contributions coming from $Z/\gamma^* + \text{jets}$ and $t\bar{t} + \text{jets}$ events. In
 1560 this region, the data driven NPL template is split into two templates based on the presence of the
 1561 NPL in the Z boson. The NPL associated with W boson is assigned to $Z/\gamma^* + \text{jets}$ and estimated
 1562 in the WZCR, while the NPL associated with Z boson is assigned to $t\bar{t} + \text{jets}$ and estimated in
 1563 the TTCR and STCR.

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

6

1564

1565 6.1 Construction of template distributions

1566 There were no selection criteria found to make a clear rejection of the background events without
1567 sacrificing a significant amount of signal. For this reason, a multivariate approach using Boosted
1568 Decision Trees that combines several discriminating variables in the TMVA framework is used.
1569 For the training, the BDTs are trained against all backgrounds, where the NPL background is not
1570 taken into account. The BDT settings avoid over-training and maintain a good discriminating
1571 power against all backgrounds. The background and signal yields follow the relative fractions
1572 predicted by the simulation.

1573 The variables used to construct the BDTs include the angles, distances, masses and transverse
1574 momenta:

- 1575 1. pseudo rapidity of the SM top: TT+ST tZu, ST tZc
- 1576 2. invariant mass of the W lepton and the SM b jet: TT+ST tZu, TT+ST tZc
- 1577 3. $\Delta\Phi$ between the W lepton and the SM b jet: TT+ST tZu, ST tZc
- 1578 4. minimal ΔR between the W lepton and jets: TT tZu
- 1579 5. invariant mass of the Z boson: TT tZu, TT tZc
- 1580 6. $\Delta\Phi$ between the W lepton and the Z boson: TT+ST tZu, TT tZc
- 1581 7. ΔR between the W lepton and the SM b jet: TT+ST tZu, TT tZc
- 1582 8. number of CSVv2 medium WP b jets: TT tZu, TT tZc
- 1583 9. invariant mass of the FCNC top: TT tZu, TT tZc
- 1584 10. ΔR between the Z boson and the FCNC light jet: TT tZu, TT tZc
- 1585 11. ΔR between the FCNC light jet and the SM b jet: TT tZu, TT tZc

1586 12. charge of the W lepton times the absolute pseudo rapidity of the W lepton: ST tZu

1587 13. b discriminant of the highest p_T jet: ST tZu, ST tZc

1588 14. total Ht of the leptons: ST tZu

1589 15. the p_T of the W lepton times its charge: ST tZu

1590 16. total invariant mass of the leptons: ST tZc

1591 17. ΔR between the W lepton and Z boson: TT+ST tZc

1592 18. total invariant mass of the event: TT tZc

1593 19. b discriminant of the FCNC light jet: TT tZc

1594 20. ΔR between the SM b jet and Z boson: TT tZc

1595 The pre fit distributions of the variables used for creating the multivariate discriminator are
 1596 given in [Section 6.1.1](#). The resulting multivariate discriminator are shown in [Section 6.1.2](#).

1597 6.1.1 Distributions of the BDT variables

NOTE: include distributions of the vars

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.

1600

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

1601 6.1.3 Transverse mass in WZCR

1602 The WZCR is used to estimate the contribution from WZ+jets and NPL background. In this
 1603 region, a fit is performed on the transverse mass distribution of the W boson. The pre-fit
 1604 templates are given in [Figure 6.6](#).

Figure 6.5: Distributions of the discriminating variable before the fit, 3e channel. Upper left: TTSR tZ_u, upper right: TTSR tZ_c; lower left: STSR tZ_u, lower right: STSR tZ_c.

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.

1605 6.2 Systematic uncertainties

1606 The systematic uncertainties entering the analysis are coming from different sources. The
 1607 experimental uncertainties arise from the reconstruction of the objects and are discussed in
 1608 [Section ??](#). These influence the number of events passing the selection, so-called normalisation
 1609 uncertainties, or the relative occupancies of the distributions, so-called shape uncertainties. The
 1610 normalisation uncertainties coming from reconstruction include the uncertainty of 2.5% on
 1611 the measured integrated luminosity and the efficiency of the trigger logic used for the analysis
 1612 which has a 1% (5%) uncertainty on the 3μ and 1e2μ(2e1μ and 3e) channels. The pileup
 1613 distribution is calculated via the minimum bias cross section which has a 4.6% uncertainty.
 1614 This uncertainty results in a systematic shift in the pileup distribution and its shape effect is
 1615 estimated by recalculating the pileup distribution for each variation of the minimum bias cross
 1616 section. The effect of the systematic upwards and downwards shift on the pileup distribution is
 1617 demonstrated in [Section ??](#). The shape uncertainties also include the uncertainties coming from
 1618 the applied lepton scale factors. Their systematic uncertainty originates from three sources:
 1619 identification, isolation and tracking. The effect of systematic upwards or downwards shift
 1620 on shapes is demonstrated in [Section ??](#) and [Section ??](#). The uncertainties arising from jet
 1621 energy corrections require a recalculation of all jet related kinematic observables and its effect
 1622 is propagated to the missing transverse energy. The resulting effect of the systematic upwards
 1623 or downwards shift on shapes is demonstrated in [Section ??](#) and [Section ??](#). The reweighting of
 1624 the CSVv2 discriminant is also a source of uncertainty. There are three sources of uncertainty
 1625 contributing to the measurement of the b-tag related scale factors: statistical uncertainties, jet
 1626 energy scale and the purity of the sample. These result in eight uncorrelated contributions for
 1627 which the effects on the shapes are shown in [Section ??](#).

1628 Since the NPL sample is artificially made from data by inverting the isolation of the third
 1629 lepton. Its effect has to be estimated. The shape uncertainty one the NPL processes is obtained
 1630 by varying the isolation inversion with respect to tight working point to the loose working point
 1631 for electrons and muons at the same time. This found to have negligible effect. The uncertainty
 1632 on the normalisation of the overall NPL yield is taken as 50% in accordance the the SM tZq
 1633 search.

NOTE: Add source

1634 The uncertainty on the expected yield of the simulated backgrounds is taken to be 30%
 1635 of the yield such that it covers all uncertainties at next to leading order accuracy. Theory
 1636 uncertainties originating from the modelling of the main backgrounds are estimated to account
 1637 for the effect on the shape of the distributions from the choice of parton density funnctions, and
 1638 renormalization (μ_R) and factorization (μ_F) scales. The effect of the renormalization (μ_R) and
 1639 factorization (μ_F) scales is estimated by varying each independently and correlated up and
 1640 down by a factor of two, where the anti-correlated variations are dropped. The envelope of these
 1641 variations is used as an uncertainty. The uncertainties coming from the parton density functions

1642 used for simulation are estimated using the PDF4LHC recipe [[Ball:2017nwa](#)], which combines
 1643 the MMHT14, CT14, and NNPDF3.0 PDF sets [[Ball:2017nwa](#)]. The theory uncertainties are
 1644 considered for the main backgrounds coming from simulation: WZ+jets, ZZ+jets, tt>Z, and tZq.
 1645 This is found to have a negligible effect.

1646 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 t <bar>t>Z norm</bar>	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

1647 6.3 Limit setting procedure validation

1648 The analysis strategy has been established using a blinded strategy. Through the use of a pseudo
 1649 dataset, the limit setting procedure has been validated. Signal injection tests for which the
 1650 signal strength from a pseudo dataset with a pre-set signal strength is estimated are performed
 and shown in [Figure 6.7](#).

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.

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, NP μ and the WZ+jets backgrounds is done by using the multi-dimensional fit in Higgs Combine Tool. The resulting signal strengths with their maximum and minimum value in the 68% CL interval according to a one-dimensional chi-square are:

$$\mu_{FakeMu} = 1.825 - 0.807 / + 0.757 \quad (6.1)$$

$$\mu_{FakeEl} = 1.321 - 0.501 / + 0.323 \quad (6.2)$$

$$\mu_{WZ} = 1.235 - 0.301 / + 0.401 \quad (6.3)$$

When applying these signal strengths as scale factors, the data agrees with simulation as can be seen in Fig. 6.9. Furthermore, a goodness of fit test using the saturated model in combine, resulting in a p-value of 0.3 (see Fig. 6.8).

Figure 6.8: Goodness of fit testing in the WZCR using the saturated model with 1000 toys.

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

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

6.4 Result 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. The maximum likelihood estimators for the nuisance parameters $\hat{\theta}$ are shown in . Their values obtained from the signal plus background or background only fits are in agreement. Furthermore, the pulls on each of the nuisance parameters are shown in . Here one can see that the nuisance parameters related to the NPL normalisations are shifted with respect to their initial values. This is to be expected since their initial normalisation is arbitrary. 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 . The post fit event yields are given in .

NOTE:
MAKE THIS
PLOT

NOTE:
MAKE THIS
PLOT

NOTE:
MAKE THIS
PLOT

NOTE:
MAKE
PLOTS

NOTE:
MAKE TA-
BLE

6.4.1 One dimensional limits

The limit setting procedure used in this search returns limits on the signal strength modifier which can be translated to signal cross sections. These limits are translated to a limit on the branching fraction using Equation 1.39. Additionally, the limit on the couplings are extracted using the fact that the cross sections are quadratically dependent on the couplings. In Figure ??, the resulting limits at 95% CL on the branching fraction and couplings related to the tZu vertex is shown. This

observed (expected) limit amounts to $\mathcal{B} < 0.024\%$ (0.015%) when $\kappa_{tZu}/\Lambda \neq 0$ and $\kappa_{tZc}/\Lambda = 0$. The expected limit surpasses the CMS search at a centre-of-mass of 8 TeV expected limits of 0.027% [Sirunyan:2017kkr]. The observed limit of 0.024% for the tZu interaction doesn't surpass the CMS 8 TeV observed limit of 0.022% [Sirunyan:2017kkr]. The ATLAS collaboration has set limits 95% CL at a centre-of-mass of 13 TeV [ATLAS-CONF-2017-070] with observed (expected) limits of 0.017% (0.024%) for the tZu coupling. The expected limit presented in this analysis surpasses the expected limit for tZu.

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

Table 6.2: Expected limits on the branching ratios at 95% CL for the tZu coupling [Sirunyan:2017kkr, ATLAS-CONF-2017-070].

	expected	+2 σ	+1 σ	-1 σ	-2 σ	observed
3 μ	0.032%	0.074%	0.050%	0.021%	0.014%	0.053%
1e2 μ	0.032%	0.074%	0.050%	0.021%	0.014%	0.064%
2e1 μ	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 fb^{-1})	0.027%	-%	0.42%	0.018%	-%	0.022%
13 TeV ATLAS (36 fb^{-1})	0.024%	-%	0.35%	0.017%	-%	0.017%

1687

In Figure ??, the resulting limits at 95% CL on the branching fraction and couplings related to the tZc vertex is shown. or this coupling, the observed (expected) limit at 95% CL is $\mathcal{B} < 0.045\%$ (0.037%) when $\kappa_{tZc}/\Lambda \neq 0$ and $\kappa_{tZu}/\Lambda = 0$. The expected limit surpasses the CMS search at a centre-of-mass of 8 TeV expected limits of 0.118% [Sirunyan:2017kkr]. This also the case for the observed limit of 0.045% for the tZc interaction, which surpasses the CMS 8 TeV observed limit of 0.049% [Sirunyan:2017kkr]. The observed (expected) limits set by the ATLAS collaboration at a centre-of-mass of 13 TeV [ATLAS-CONF-2017-070] are 0.023% (0.032%) for the tZc coupling. The expected limit presented in this analysis is in accordance with this limit.

The limits for each leptonic channel separate as well as their combined limits are shown for the tZc vertex in Figure ?? and Table 6.3. Also here, the leptonic channels are in agreement

¹⁶⁹⁹ with eachother, and the presence of a muon helps the sensitivity. For the tZu vertex, the TTSR
¹⁷⁰⁰ is the most sensitive region and by combining single top quark and top quark pair signals, one
¹⁷⁰¹ also gains in sensitivity.

Table 6.3: Expected limits on the branching ratios at 95% CL for the tZc coupling [Sirunyan:2017kkr, ATLAS-CONF-2017-070].

	expected	+2 σ	+1 σ	-1 σ	-2 σ	observed
3 μ	0.070%	0.15%	0.10%	0.046%	0.032%	0.12%
1e2 μ	0.079%	0.18%	0.12%	0.052%	0.036%	0.096%
2e1 μ	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 fb^{-1})	0.118%	-%	0.222%	0.071%	-%	0.049%
13 TeV ATLAS (36 fb^{-1})	0.032%	-%	0.046%	0.022%	-%	0.023%

¹⁷⁰² 6.4.2 Two-dimensional limits

¹⁷⁰³ The expected limits at 95% CL on the BR of 0.015% for tZu and 0.037% for tZc surpass the
¹⁷⁰⁴ expected limits of 0.027% for tZu and 0.118% for tZc from the 8 TeV analysis. The observed
¹⁷⁰⁵ limit at 95% CL on the BR for tZu is 0.024% and isn't improving the current best limit of
¹⁷⁰⁶ 0.022%. For tZc the observed limit 95% CL is 0.045% and is improving the current best limit of
¹⁷⁰⁷ 0.049%. The latest result on 13 TeV from ATLAS [ATLAS-CONF-2017-070] set limits at 95%
¹⁷⁰⁸ CL of observed (expected) limits of 0.017% (0.024%) for tZu and 0.023% (0.032%) for tZc.

Denouement of the top-Z FCNC hunt at 13 TeV

1709

7

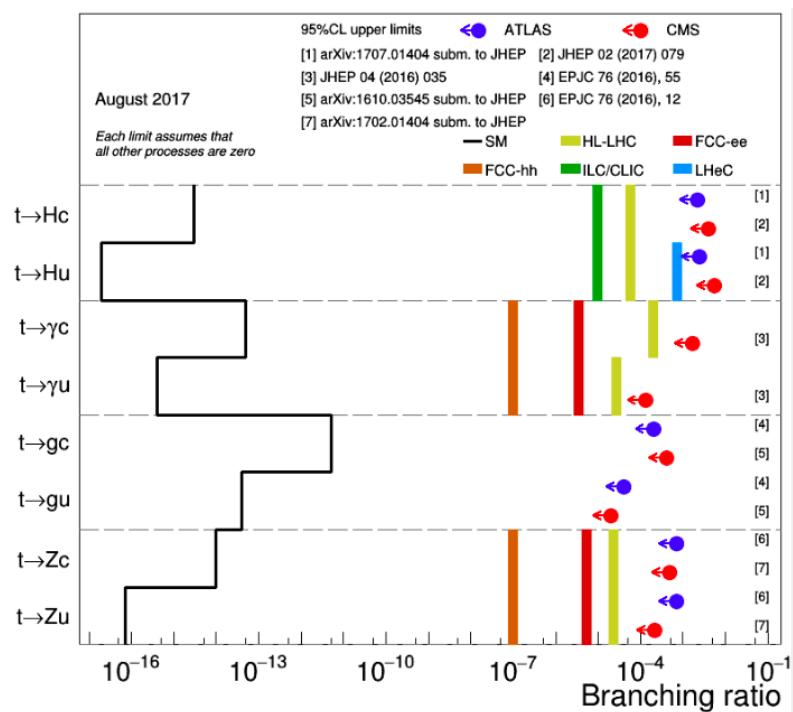


Figure 7.1:

1710

Trigger scale factors



1711 The trigger scale factors measured as a function of lepton p_T , using the dataset collected by
 1712 E_T^{miss} triggers and WZ simulation, after a 3 lepton and jets selection, in the Z mass window. All
 1713 corrections to simulation are applied.

Table A.1: Trigger efficiencies on data events selected with E_T^{miss} triggers and WZ simulation for all leptonic channels together. The unweighed number of events is quoted. When there are no events passing the cuts, it is indicated with N/A. The uncertainties are statistical uncertainties.

ALL CHANNEL	data		WZ simulations	
	Efficiency	unc.	Efficiency	unc.
3 lep, at least one jet	117/118 = 99.15 %	12.94%	18047/18055 = 99.96%	1.05%
STSR	6/6 = 100.00%	57.74%	1541/1541 = 100.00%	3.60%
TTSR	26/27 = 96.30%	26.46%	1791/1792 = 99.94%	3.34%
WZCR	69/69 = 100.00 %	17.03%	14405/14412=99.95%	1.18%

Table A.2: Trigger efficiencies on data events selected with E_T^{miss} triggers and WZ simulation for 3 μ leptonic channel together. The unweighed number of events is quoted. When there are no events passing the cuts, it is indicated with N/A. The uncertainties are statistical uncertainties.

3 μ CHANNEL	data		WZ simulations	
	Efficiency	unc.	Efficiency	unc.
3 lep, at least one jet	40/40 = 100.00 %	22.36 %	7814/7814 = 100.00%	1.60%
STSR	N/A	N/A	687/687 = 100%	5.40%
TTSR	13/13 = 100.00%	39.22%	763/763 = 100.00%	5.12%
WZCR	22/22 = 100.00 %	30.15%	6238/6238=100.00%	1.79%

Table A.3: Trigger efficiencies on data events selected with E_T^{miss} triggers and WZ simulation for 3e leptonic channel together. The unweighted number of events is quoted. When there are no events passing the cuts, it is indicated with N/A. The uncertainties are statistical uncertainties.

3e CHANNEL	data		WZ simulations	
	Efficiency	unc.	Efficiency	unc.
3 lep, at least one jet	20/21 = 95.24%	29.76%	2211/2215 = 99.82 %	3.00%
STS R	4/4 = 100.00%	70.71%	176/176 = 100.00%	10.66%
TTS R	2/3 = 66.67%	60.86%	242/242 = 100.00%	9.09%
WZCR	14/14 = 100.00 %	37.80%	1744/1748=99.77%	3.38%

Table A.4: Trigger efficiencies on data events selected with E_T^{miss} triggers and WZ simulation for 2e1 μ leptonic channel together. The unweighted number of events is quoted. When there are no events passing the cuts, it is indicated with N/A. The uncertainties are statistical uncertainties.

2e1 μ CHANNEL	data		WZ simulations	
	Efficiency	unc.	Efficiency	unc.
3 lep, at least one jet	32/32 = 100.00 %	25.00 %	3116/3118 = 99.94%	2.53%
STS R	1/1 = 100.00%	141.42%	255/255 = 100%	8.86%
TTS R	9/9 = 100.00%	47.14%	291/291 = 100.00%	8.29%
WZCR	14/14 = 100.00 %	37.80%	2529/2531=99.92%	2.81%

Table A.5: Trigger efficiencies on data events selected with E_T^{miss} triggers and WZ simulation for 1e2 μ leptonic channel together. The unweighted number of events is quoted. When there are no events passing the cuts, it is indicated with N/A. The uncertainties are statistical uncertainties.

1e2 μ CHANNEL	data		WZ simulations	
	Efficiency	unc.	Efficiency	unc.
3 lep, at least one jet	25/25 = 100.00%	28.28%	4906/4908 = 99.96 %	2.02%
STS R	1/1 = 100.00%	141.42%	423/423 = 100.00%	6.88%
TTS R	2/2 = 100.00%	100.00%	495/496 = 99.80%	6.34%
WZCR	19/19 = 100.00 %	32.44%	3894/3895 =99.97%	2.27%

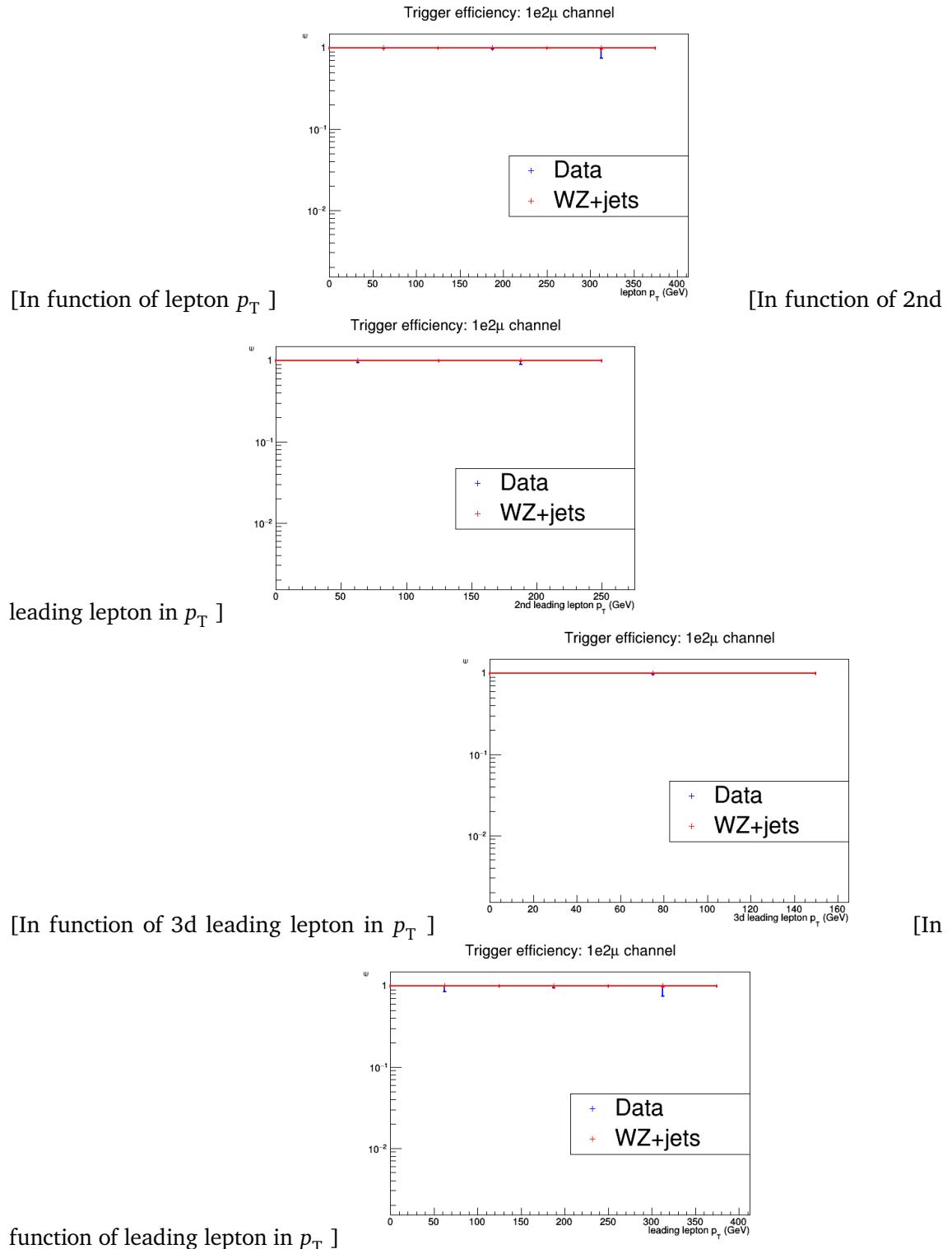


Figure A.1: The trigger efficiencies measured as a function of lepton p_T , using the dataset collected by E_T^{miss} triggers and WZ simulation, after a 3 lepton and jets selection, in the Z mass window. All corrections to simulation are applied. 1e2μ channel.

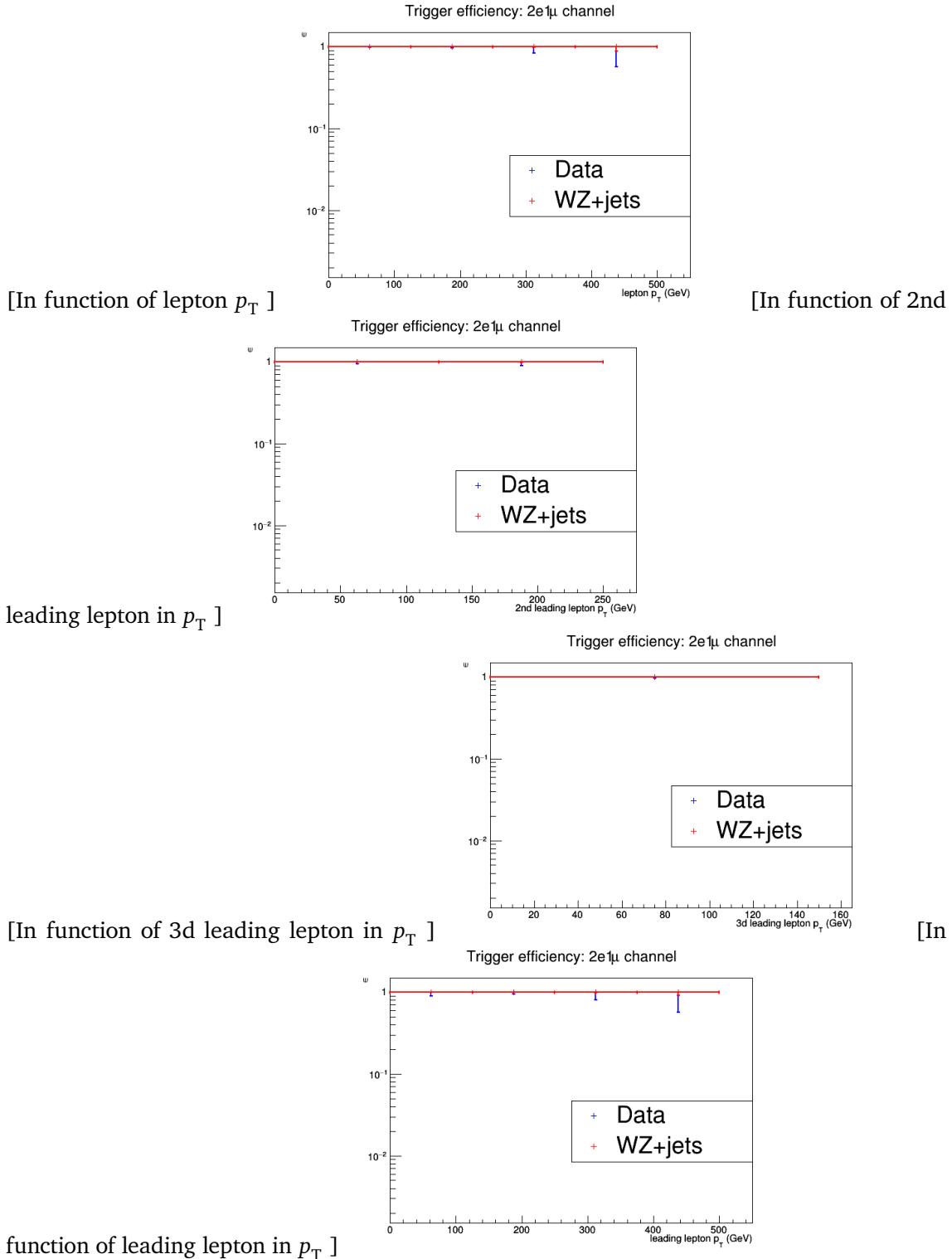


Figure A.2: The trigger efficiencies measured as a function of lepton p_T , using the dataset collected by E_T^{miss} triggers and WZ simulation, after a 3 lepton and jets selection, in the Z mass window. All corrections to simulation are applied. 2e1μ channel.

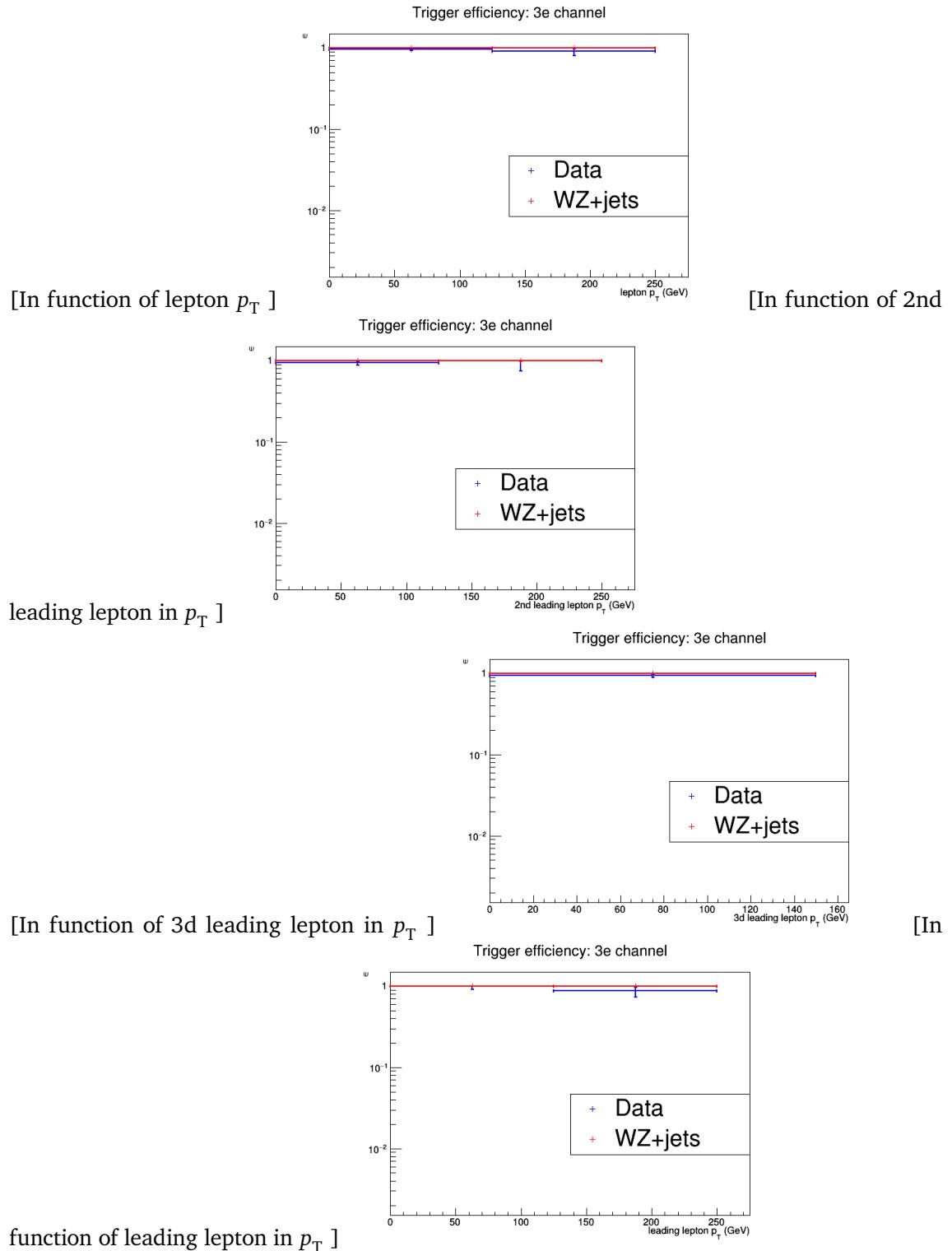


Figure A.3: The trigger efficiencies measured as a function of lepton p_T , using the dataset collected by E_T^{miss} triggers and WZ simulation, after a 3 lepton and jets selection, in the Z mass window. All corrections to simulation are applied. 3e channel.

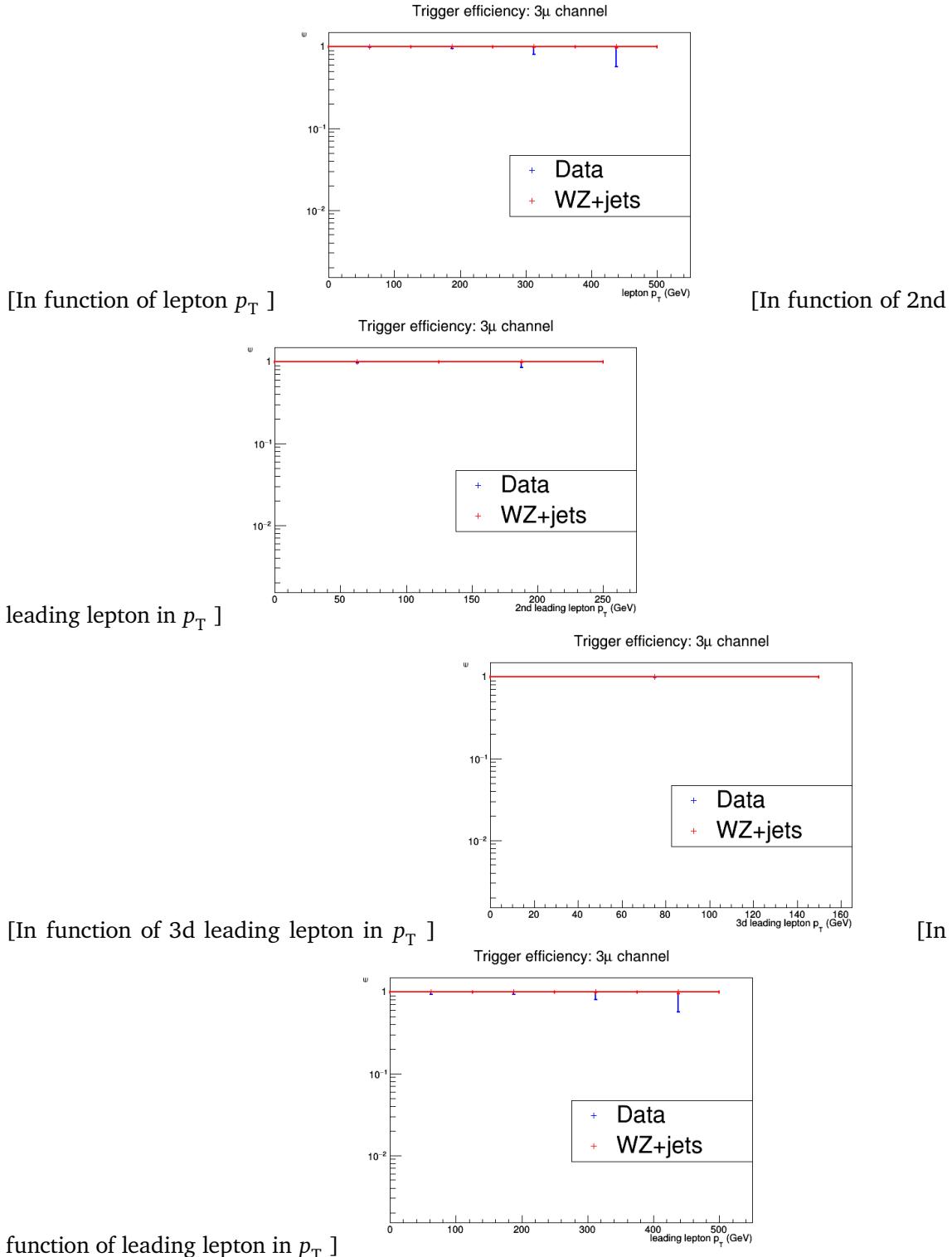


Figure A.4: The trigger efficiencies measured as a function of lepton p_T , using the dataset collected by E_T^{miss} triggers and WZ simulation, after a 3 lepton and jets selection, in the Z mass window. All corrections to simulation are applied. 3μ channel.

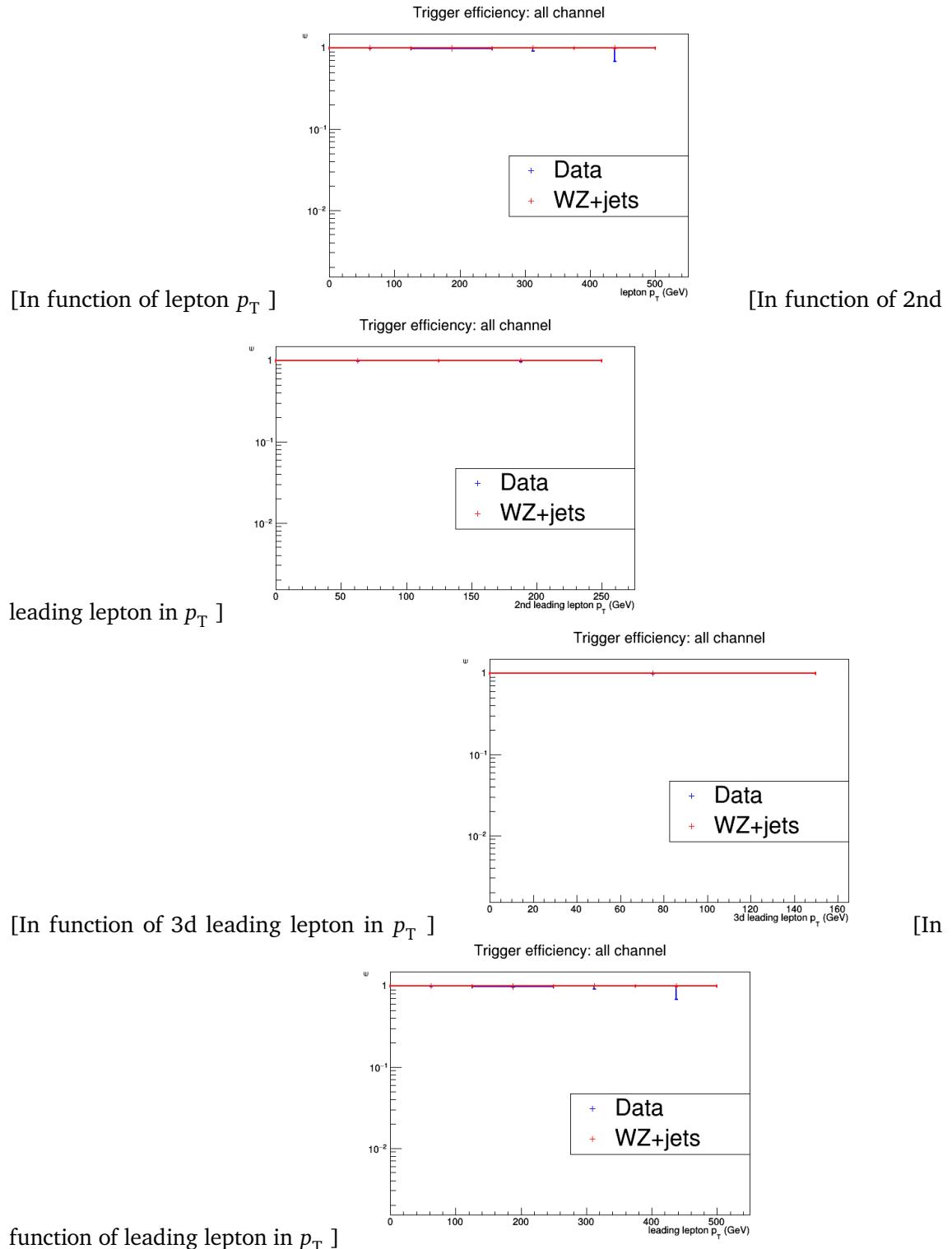


Figure A.5: The trigger efficiencies measured as a function of lepton p_T , using the dataset collected by E_T^{miss} triggers and WZ simulation, after a 3 lepton and jets selection, in the Z mass window. All corrections to simulation are applied. All channel.

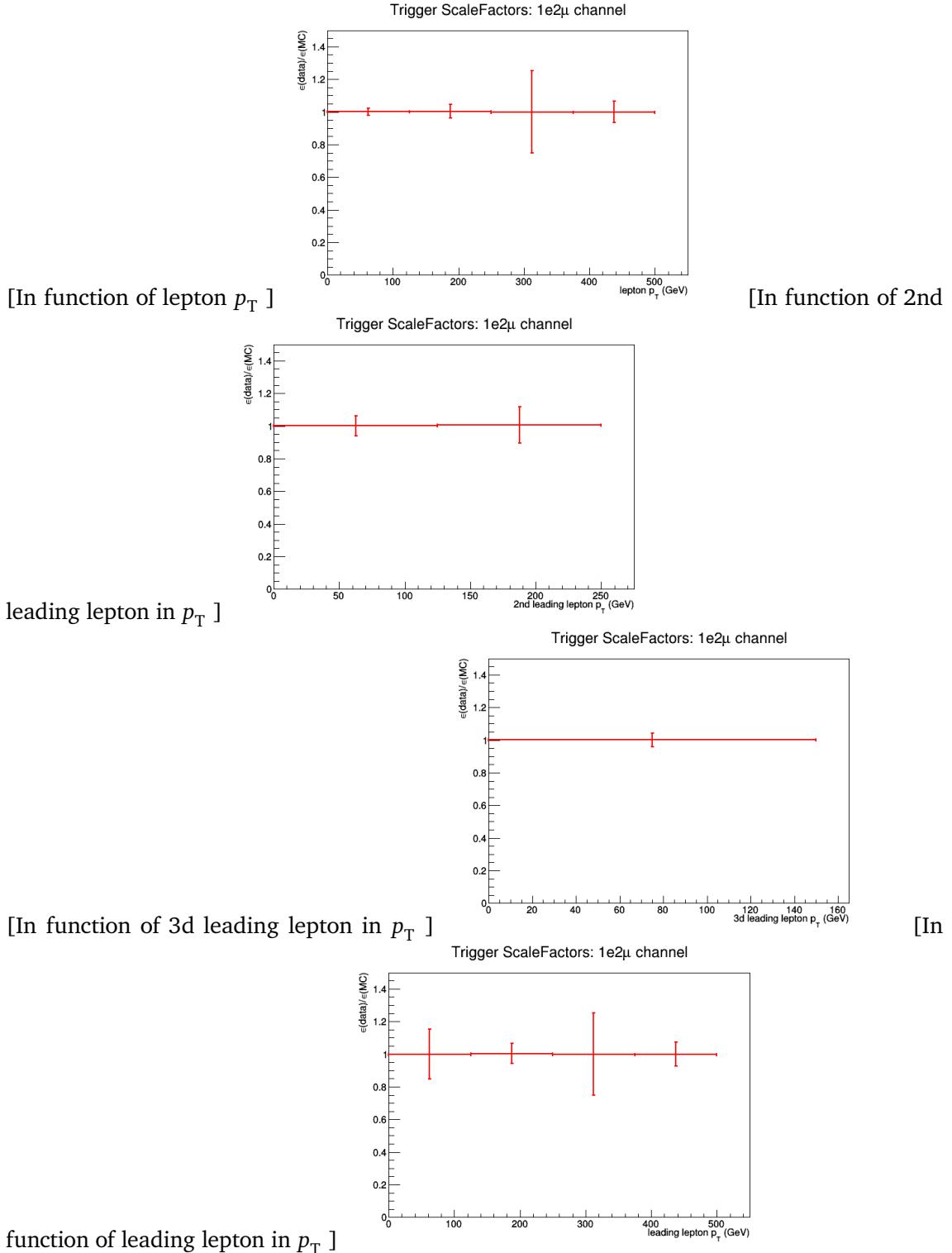


Figure A.6: The trigger scale factors measured as a function of lepton p_T , using the dataset collected by E_T^{miss} triggers and WZ simulation, after a 3 lepton and jets selection, in the Z mass window. All corrections to simulation are applied. 1e2μ channel.

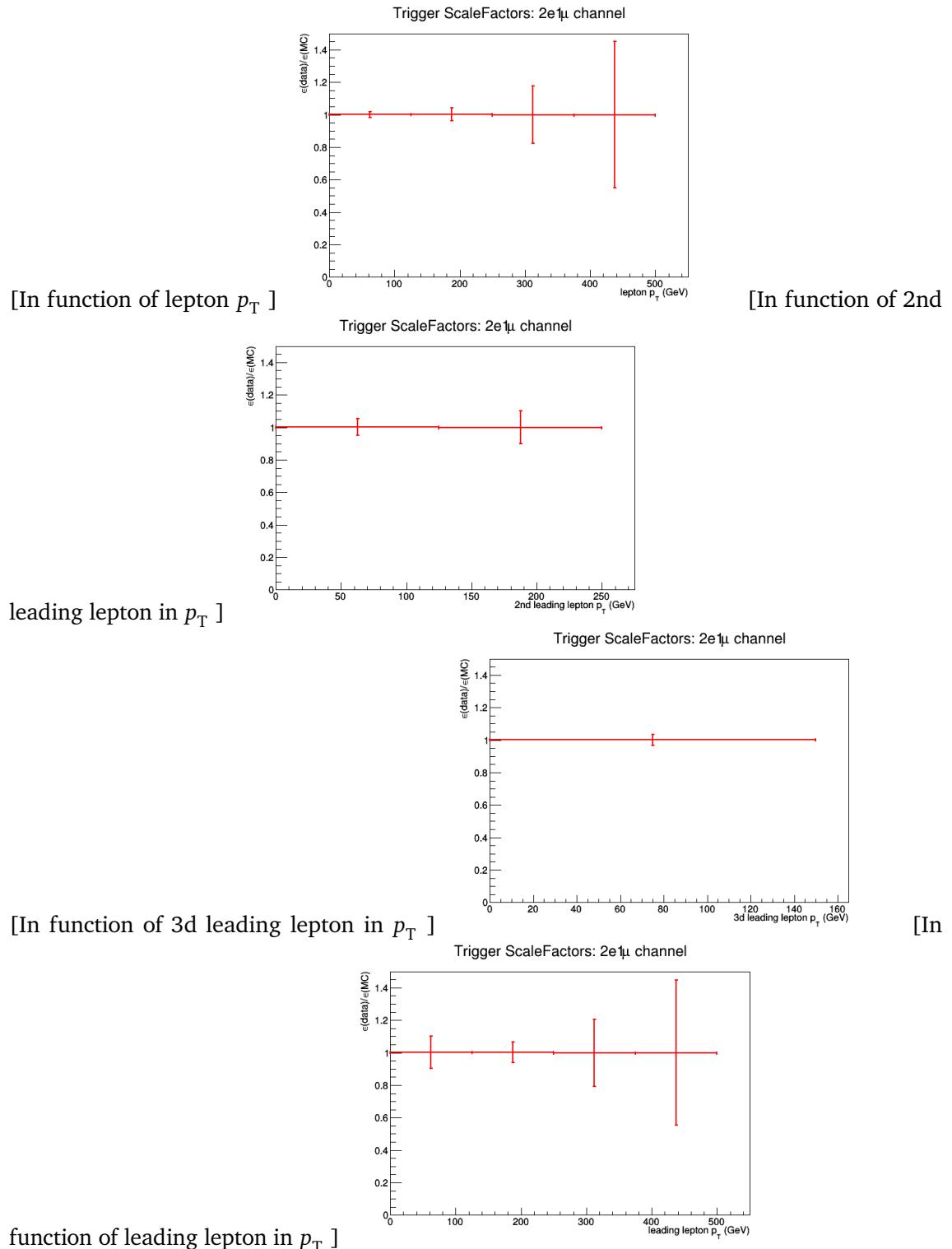


Figure A.7: The trigger scale factors measured as a function of lepton p_T , using the dataset collected by E_T^{miss} triggers and WZ simulation, after a 3 lepton and jets selection, in the Z mass window. All corrections to simulation are applied. 2e1μ channel.

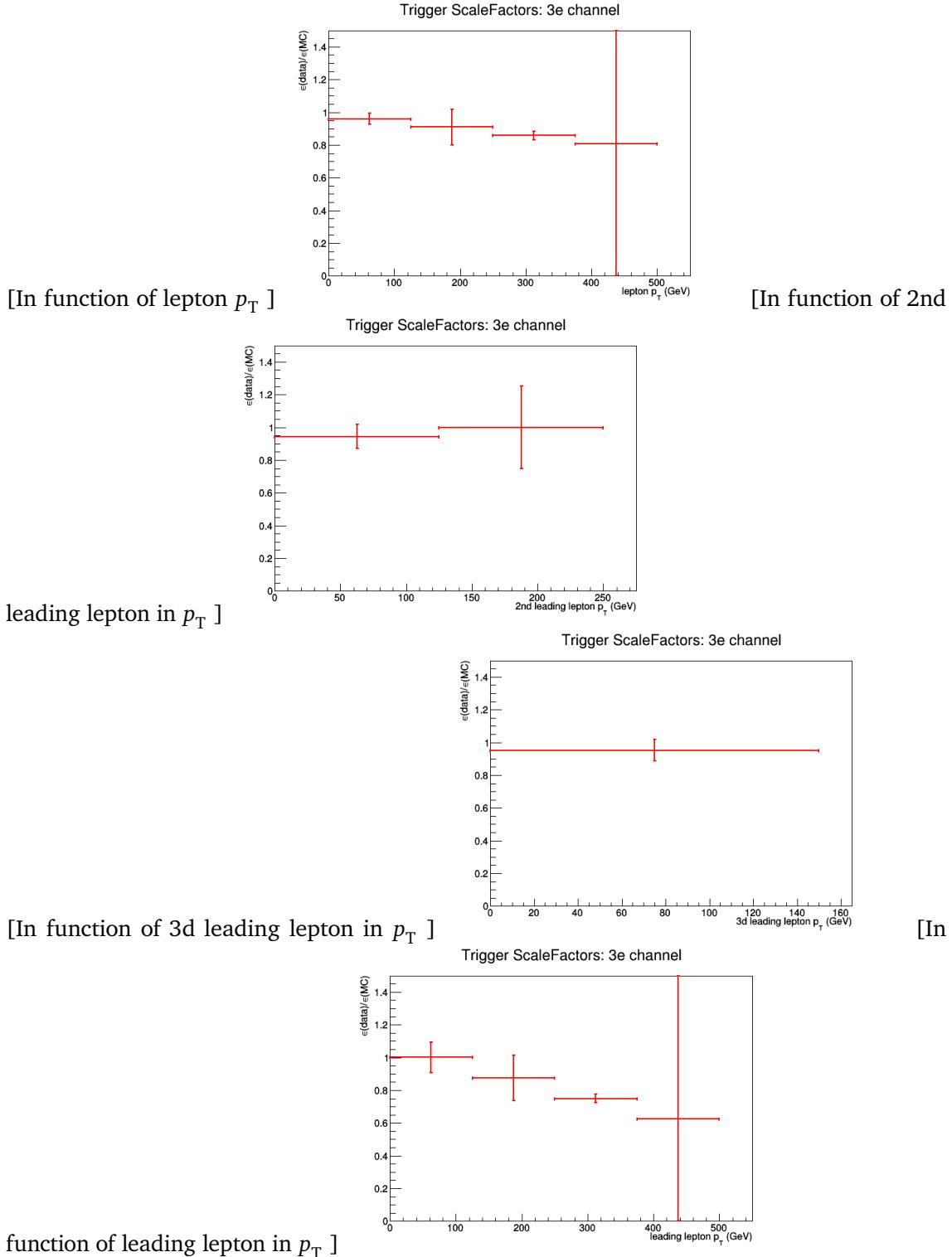


Figure A.8: The trigger scale factors measured as a function of lepton p_T , using the dataset collected by E_T^{miss} triggers and WZ simulation, after a 3 lepton and jets selection, in the Z mass window. All corrections to simulation are applied. 3e channel.

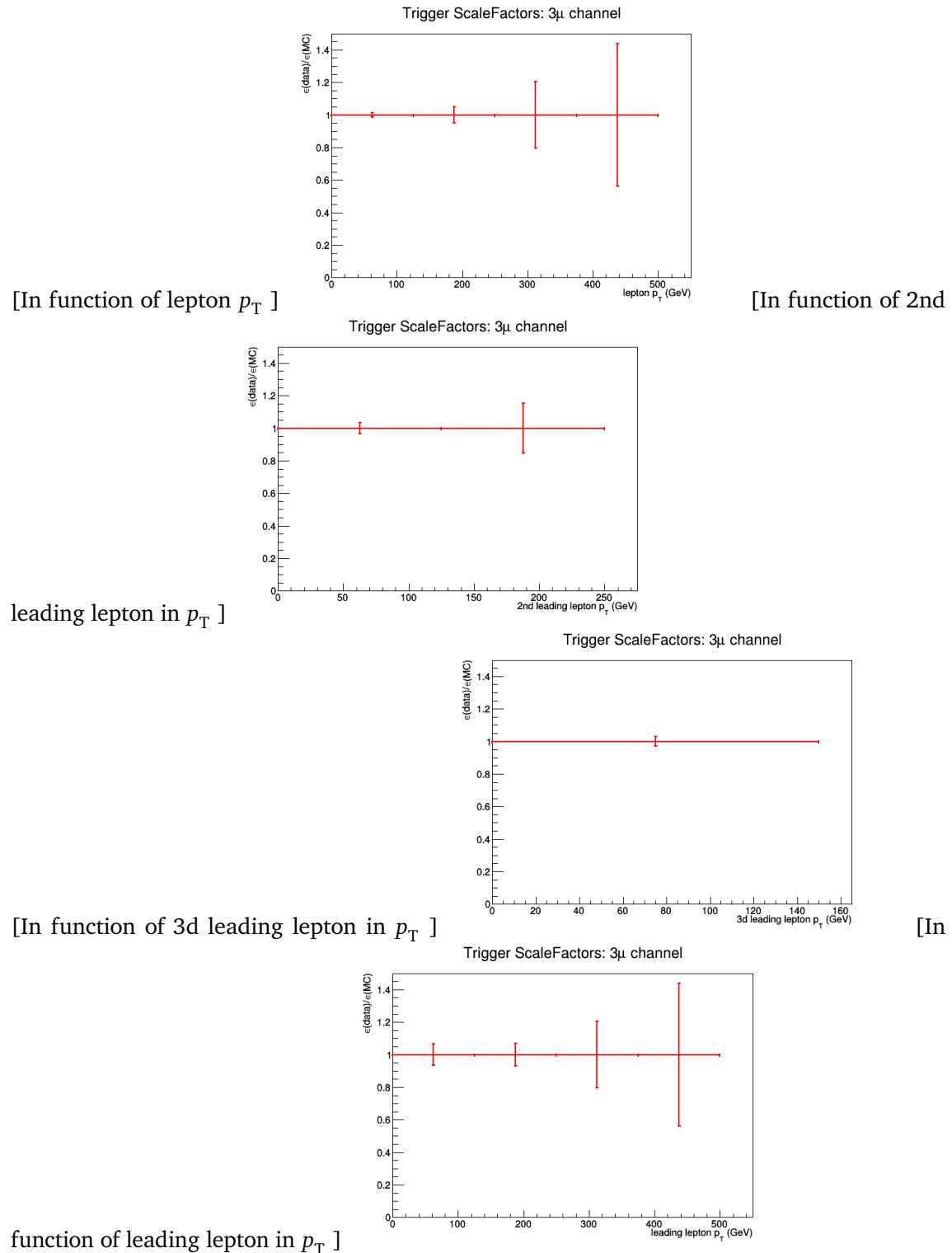


Figure A.9: The trigger scale factors measured as a function of lepton p_T , using the dataset collected by E_T^{miss} triggers and WZ simulation, after a 3 lepton and jets selection, in the Z mass window. All corrections to simulation are applied. 3μ channel.

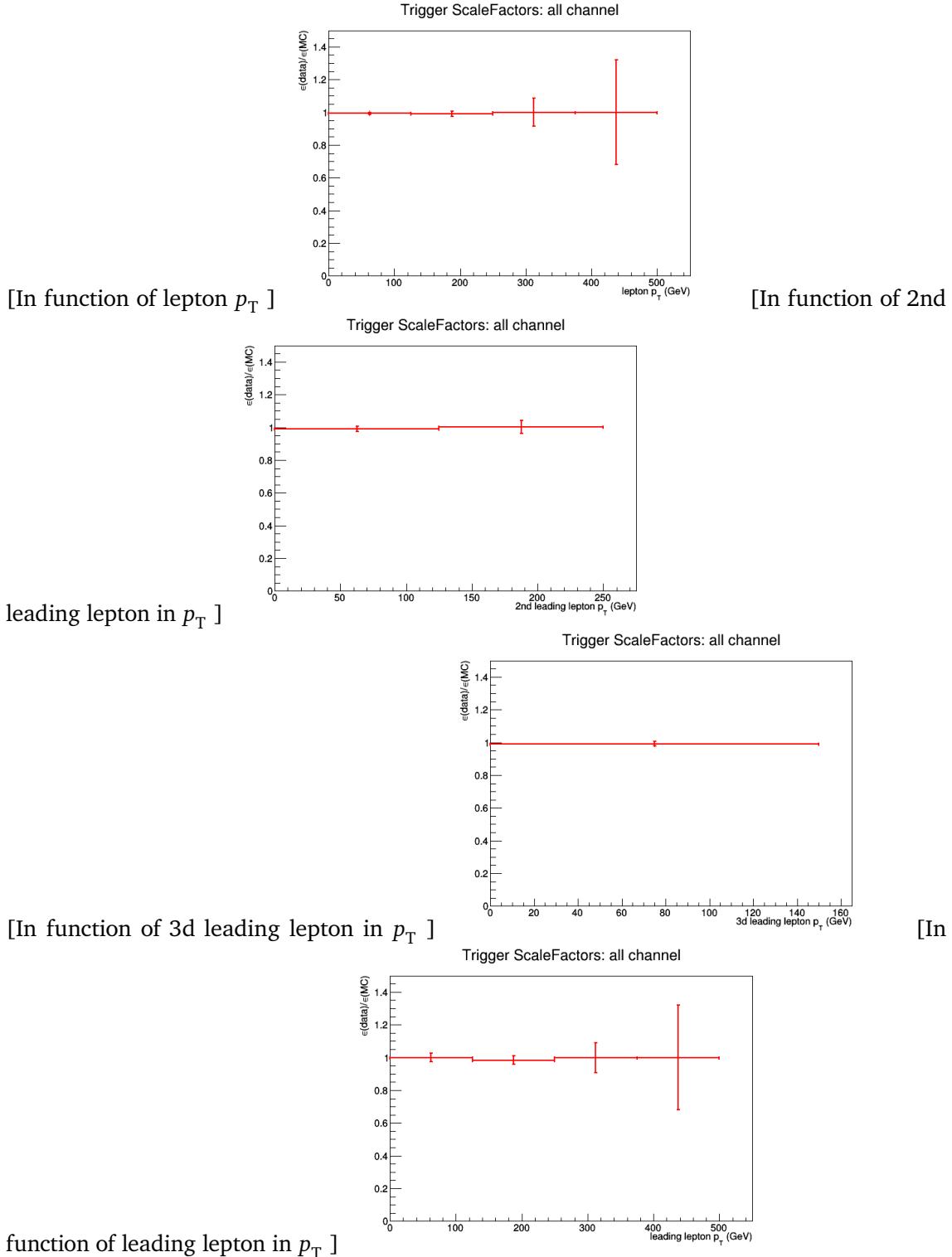


Figure A.10: The trigger scale factors measured as a function of lepton p_T , using the dataset collected by E_T^{miss} triggers and WZ simulation, after a 3 lepton and jets selection, in the Z mass window. All corrections to simulation are applied. All channel.

Dilepton controlplots

B

Statistical independent regions

1715

C

	STS R	TTS R	WZCR	TTCR	STCR	$Tr_{WZCR \rightarrow STSR}$	$Tr_{WZCR \rightarrow TTSR}$	$Tr_{TTCR \rightarrow TTSR}$	$Tr_{STCR \rightarrow STSR}$
FCNC tZu	3.9 ± 0.0	11.2 ± 0.0	7.3 ± 0.0	0.6 ± 0.0	0.3 ± 0.0	0.22 ± 0.00	0.46 ± 0.00	—	—
FCNC tZc	14.1 ± 0.1	30.9 ± 0.1	15.5 ± 0.1	1.7 ± 0.0	1.2 ± 0.0	0.39 ± 0.01	0.76 ± 0.01	—	—
DY	-4.6 ± 21.7	22.3 ± 14.3	136.9 ± 48.5	-14.0 ± 22.2	-15.9 ± 11.3	0.08 ± 0.09	0.10 ± 0.10	—	—
ttbar	13.1 ± 2.5	9.7 ± 2.1	8.5 ± 1.9	22.3 ± 3.1	33.7 ± 3.8	0.54 ± 0.31	0.70 ± 0.31	0.33 ± 0.00	0.33 ± 0.00
DD fake	62.0 ± 7.9	63.0 ± 7.9	204.0 ± 14.3	41.0 ± 6.4	56.0 ± 7.5	—	—	—	—
WZ+jets	60.9 ± 1.6	83.0 ± 1.7	552.7 ± 4.6	11.9 ± 0.6	8.5 ± 0.6	0.10 ± 0.00	0.15 ± 0.00	—	—
tZq	8.0 ± 0.2	16.9 ± 0.2	7.3 ± 0.2	2.2 ± 0.1	1.0 ± 0.1	0.36 ± 0.02	0.67 ± 0.03	—	—
ttZ	3.5 ± 0.3	42.5 ± 0.9	9.6 ± 0.5	6.8 ± 0.4	0.5 ± 0.1	0.14 ± 0.02	0.61 ± 0.05	—	—
ZZ	4.6 ± 0.1	4.8 ± 0.1	46.2 ± 0.3	0.8 ± 0.0	0.6 ± 0.0	0.10 ± 0.00	0.13 ± 0.00	—	—
other	1.2 ± 0.3	5.6 ± 0.3	6.8 ± 0.5	7.5 ± 0.7	2.4 ± 0.5	0.16 ± 0.03	0.30 ± 0.03	—	—

Table C.1: Event yields: All channel. The last three columns represent the transfer factors that have to be applied.