



VRIJE
UNIVERSITEIT
BRUSSEL

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

⁵ Isis Van Parijs

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

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

⁸

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

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

Contents

11	1 Theoretical basis	1
12	1.1 Elementary particles and forces	1
13	1.2 Standard Model Lagrangian, connecting fields with particles	2
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	26
23	2.2.1 CMS coordinate system	27
24	2.2.2 Towards the heart of CMS	28
25	2.2.3 Data acquisition	36
26	2.2.4 Phase 1 upgrades	36
27	2.2.5 CMS computing model	38
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	44
33	3.2.3 Generating FCNC top-Z interactions	46
34	3.2.4 Generating SM background events	48
35	3.3 Multivariate analysis techniques: Boosted Decision Trees	50
36	3.4 Statistical methodology	52
37	3.4.1 The absence of signal: limits	53
38	3.4.2 Adding sources of uncertainty	55
39	3.4.3 Asymptotic approximation of the CL method	56
40	3.4.4 Extracting the signal model parameters	57
41	4 Event reconstruction and identification	59

42	4.1 Object Reconstruction	59
43	4.1.1 Charged particle tracks	60
44	4.1.2 Following the Muon's Footsteps	61
45	4.1.3 The path of the Electron	62
46	4.1.4 Primary Vertex Reconstruction	63
47	4.1.5 Calorimeter clusters	63
48	4.2 Putting the pieces together	64
49	4.3 Particle flow identification	65
50	4.3.1 Muons	65
51	4.3.2 Electrons and isolated photons	65
52	4.3.3 Hadrons and non-isolated photons	65
53	4.3.4 Post processing	66
54	4.4 Pile up mitigation and luminosity measurement	66
55	4.5 Physics object reconstruction and identification	67
56	4.5.1 Muons	67
57	4.5.2 Electrons	70
58	4.5.3 Jets	71
59	4.5.4 Jets from b fragmentation	74
60	4.5.5 Missing transverse energy	77
61	4.6 Summary of corrections	78
62	5 Event selection and categorisation	79
63	5.1 Baseline event selection and filters	79
64	5.1.1 Event cleaning	82
65	5.1.2 Estimation of the trigger efficiency	84
66	5.1.3 Corrections	85
67	5.2 Analysis Strategy	90
68	5.3 Data driven NPL background simulation	90
69	5.4 Regions and channels	91
70	5.4.1 WZ control region	93
71	5.4.2 TTCR and STCR	94
72	5.4.3 TTSR and STSR	95
73	6 The search for FCNC involving a top quark and a Z boson	97
74	6.1 Construction of template distributions	97
75	6.2 Systematic uncertainties	97
76	6.3 Limit setting procedure	97
77	6.4 Result and discussion	97
78	7 Denouement of the top-Z FCNC hunt at 13 TeV	99
79	Appendices	99
80	A Trigger scale factors	101■
81	B Dilepton controlplots	113■

CONTENTS

v

82	C Statistical independent regions	115■
83	Bibliography	117■

Theoretical basis

1

85 The Standard Model (SM) [1] is a name given in the 1970s to a theory describing the fundamental
 86 particles and their interactions. This quantum field theory describes the particles and their
 87 interactions as fields and has successfully incorporated three of the four fundamental forces in
 88 the universe. In [Section 1.1](#), the particle content of the SM is summarised, while [Section 1.2](#)
 89 describes the SM Lagrangian and its symmetries. In [Section 1.3](#), the flavour content of the SM
 90 is highlighted, and [Section 1.4](#) focusses on the top quark in the SM. The latest experimental
 91 results related to the properties of the top quark are summarised in [Section ??](#).

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

97 1.1 Elementary particles and forces

98 The interactions in nature can be described by four forces, the strong force, the electromagnetic
 99 (EM) force, the weak force and the gravitational force. These interactions happen via particles
 100 with an integer spin known as bosons. The strong interaction is mediated by eight gluons g ,
 101 while the electromagnetic force is mediated by photons γ , and the weak force by Z and W^\pm
 102 bosons. In [Table 1.1](#), the forces and their characteristics are shown. The gravitational force is
 103 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

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

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

Generation	Particle	Mass	Charge
First	e^-	0.511 MeV	$-q_e$
	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

116

The quarks can also be divided into three generations. Unlike the leptons, they carry colour charge and can interact via the strong interaction. The top quark, discovered in 1995 at the Tevatron [7, 8], is the heaviest SM particle with a mass¹ measured to be $173.34 \pm 0.27(\text{stat}) \pm 0.71(\text{syst})$ GeV [9]. The quarks and their properties are summarised in Table 1.3. In nature, only colour neutral objects can exist. This has as consequence that quarks are bound through gluons into mesons (quark+anti-quark) and baryons (three quarks). These mesons and baryons are mostly short-lived and unstable particles that rapidly decay through W^\pm and Z bosons. The only known stable baryon is the proton, made up of two up quarks and one down quark.

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

1.2 Standard Model Lagrangian, connecting fields with particles

The SM is a quantum field theory and thus describes the dynamics and kinematics of particles 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

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 [6], where q_e represents the elementary charge.

	Generation	Particle	Mass	Charge
First	up u	$2.2^{+0.6}_{-0.4}$ MeV	$\frac{2}{3} q_e$	
	down d	$4.7^{+0.5}_{-0.4}$ MeV	$\frac{-1}{3} q_e$	
Second	charm c	1.28 ± 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	$\frac{-1}{3} q_e$	

132 symmetry, where $SU(2)_L \times U(1)_Y$ describes the electroweak interaction and $SU(3)_C$ the strong
 133 interaction. The indices refer to colour C, the left chiral nature of the $SU(2)_L$ coupling L,
 134 and the weak hypercharge Y. Its Lagrangian is constructed such that contains symmetries
 135 representing physics conservation laws such as conservation of energy, momentum and angular
 136 momentum. The symmetries under local gauge transformations are sustained by demanding
 137 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:

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

The generators T^a satisfy the Lie algebra:

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

138 where ϵ_{abc} is an antisymmetric tensor. The gauge fields of $SU(2)_L$ only couple to left-handed
 139 fermions as required by the observed parity violating nature of the weak force. The $SU(3)_C$
 140 group represents quantum chromodynamics (QCD). It has eight generators corresponding to
 141 eight gluon fields $G_\mu^{1..8}$. 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 singlets:

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

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.

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

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

that are related to the gauge, fermion, Yukawa and scalar sectors. The gauge Lagrangian regroups the gauge fields of all three symmetry groups, and the fermionic part consists of kinetic energy terms for quarks and leptons. The interaction between fermions and the scalar doublet ϕ gives rise to fermion masses and is described by the Yukawa Lagrangian. The scalar part of 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 $SU(2)_L \times U(1)_Y$:

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

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

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

Electroweak symmetry breaking

In $\mathcal{L}_{\text{gauge}}$ and \mathcal{L}_f are no mass terms for fermions present because only singlets under $SU(3)_C \times SU(2)_L \times U(1)_Y$ can acquire a mass with an interaction of the type $m^2 \phi^\dagger \phi$ without breaking the gauge invariance. In order to accommodate mass terms for fermions and gauge fields, 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 Z bosons. 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 Z bosons as:

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

- 160 The Brout-Englert-Higgs field couples universally to fermions with a strength proportional to
161 their masses, and to gauge bosons with a strength proportional to the square of their masses.

162 1.3 Flavours in the SM

Flavour changing charged currents are introduced in 1963 by Nicola Cabibbo [12]. Via interactions with a W boson the flavour of the quarks is changed. At the time of the postulation only up, down, and strange quarks were known and the charged weak current was described as a coupling between the up quark and d_{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) [13–15] to require the existence of a fourth quark with charge $\frac{2}{3}q_e$. This quark, known as the charm quark, couples to s_{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

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

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

Kobayashi and Maskawa generalised the Cabibbo rotation matrix to accommodate a third generation of quarks. The result is a 3×3 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 [16]. 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 [6]

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

165 From Equation 1.15 follows that top quarks predominantly decay via charged weak currents
 166 to bottom quarks, with a probability consistent with unity. In the SM, FCNC can only occur via
 167 higher loop interactions which are highly suppressed. The expected transition probabilities for
 168 a top quark decaying via a FCNC interaction in the SM are given in Table 1.4, where it is clear
 169 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 [17].

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}

170

171 1.4 The top quark in the SM

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

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

yielding a Yukawa coupling of [6]

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

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

179 The CKM matrix element V_{tb} , given in [Equation 1.15](#), is experimentally found to be much
 180 larger than V_{ts} , V_{td} , and close to unity. The top quark decays through electroweak interactions
 181 since the W boson mass is smaller than the top quark mass and the W boson can be on shell. A
 182 consequence of this is that the top quark has a very short lifetime of only $1/\Gamma_t \approx 5 \cdot 10^{-25}$ s [6]
 183 leading to the fact that the formation of bound states involving top quarks are not allowed.
 184 This lifetime is even shorter than the typical hadronisation timescale of $1/\Lambda_{\text{QCD}} \approx 10^{-23}$ s,
 185 prohibiting gluons to radiate from the top quark and keeping its spin coherent. Since the
 186 electroweak interactions have a vector-axial vector (V-A) coupling structure³, the top quark
 187 spin orientation can be derived from the angular distributions of its decay products. This makes
 188 it possible to study the polarisation of top quarks from the angular distributions in various
 189 processes.

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

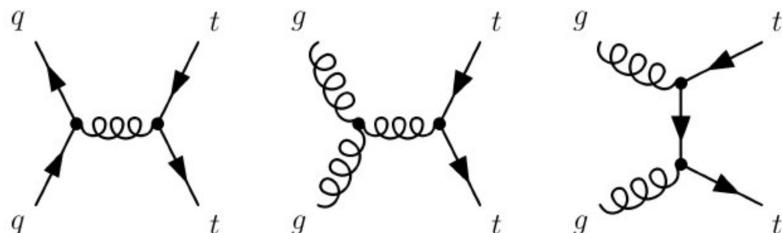


Figure 1.1: Leading order diagrams of the top quark pair production. Gluon fusion (right and middle) are the dominant processes at the LHC, while quark fusion (left) is the dominant one at the Tevatron.

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.

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

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

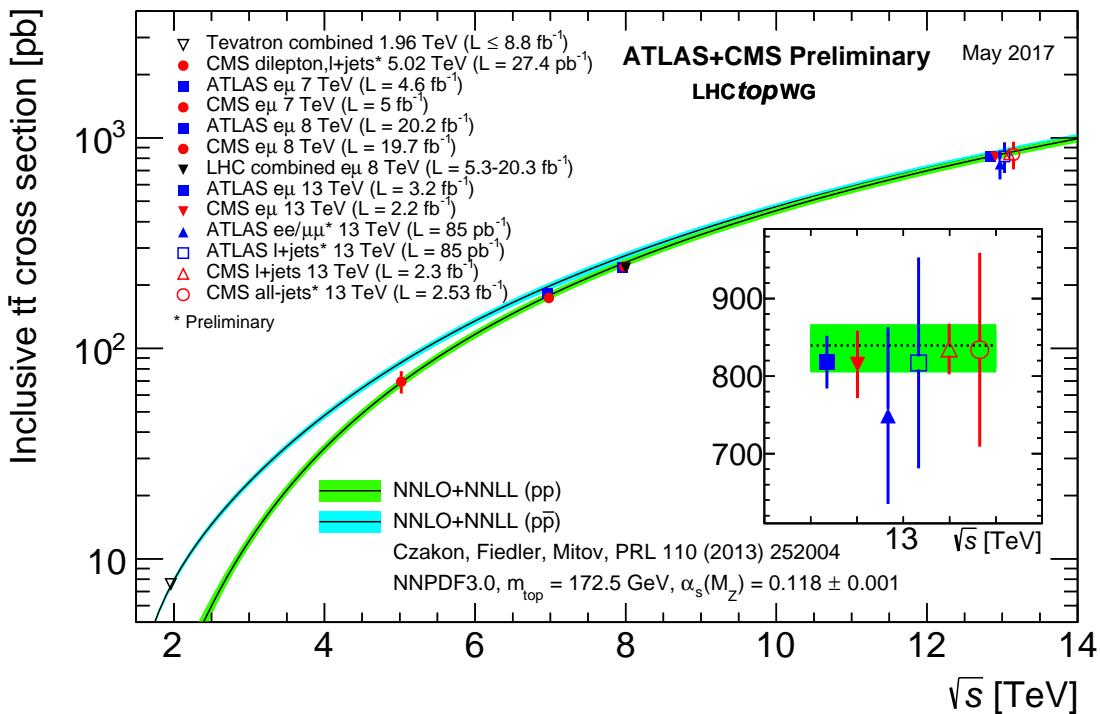


Figure 1.2: Summary of the LHC and the Tevatron measurements of the top quark pair production cross section as function of the centre-of-mass energy compared with the next-to-next-to-leading order QCD calculation. The theory bands are the uncertainties due to renormalization and factorisation scales,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

The singly produced top quarks are produced via the electroweak interaction. These production mechanisms are subdivided at leading order into three main channels based on the virtuality ($Q^2 = -p_\mu p^\mu$) of the exchanged W boson. In Figure 1.3, the corresponding Feynman diagrams are shown. The single top quark production cross sections, given in Table 1.6, are smaller than the top quark pair production cross sections since the electroweak coupling strength is smaller than the strong coupling strength. In addition, for the single top quark production, there is the need of sea quarks (b, \bar{q}) in the initial states for which the parton density functions increase less steeply at low momentum fractions compared to the gluon parton density functions.

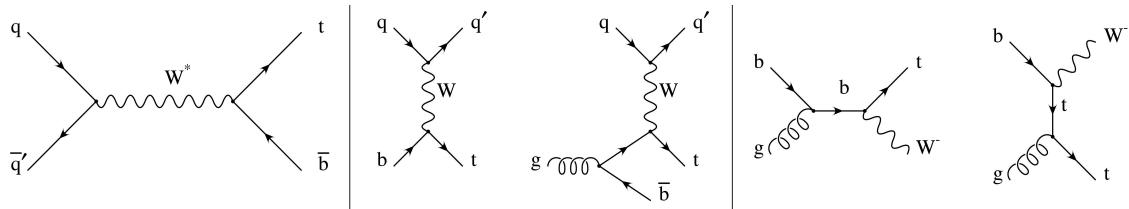


Figure 1.3: Leading order Feynman diagrams of the electroweak production of single top quarks in the s -channel (left), t -channel (middle), and for the tW associated production. Figure taken from [20].

210

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

Table 1.6: Predictions on the single top quark production cross sections at next-to-leading order per centre-of-mass energy [6]. The uncertainties from scale dependence and from parton density functions are combined in quadrature or given separately (scale + PDF). For the t -channel the relative proportions to t and \bar{t} are 65% and 35%. For the s -channel this is respectively 69% and 31%. The tW -channel has an equal proportion of top and antitop quarks. For Tevatron, the top quark mass is assumed to be 173.3 GeV, while the LHC predictions use $m_t = 172.5$ GeV [6, 21].

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

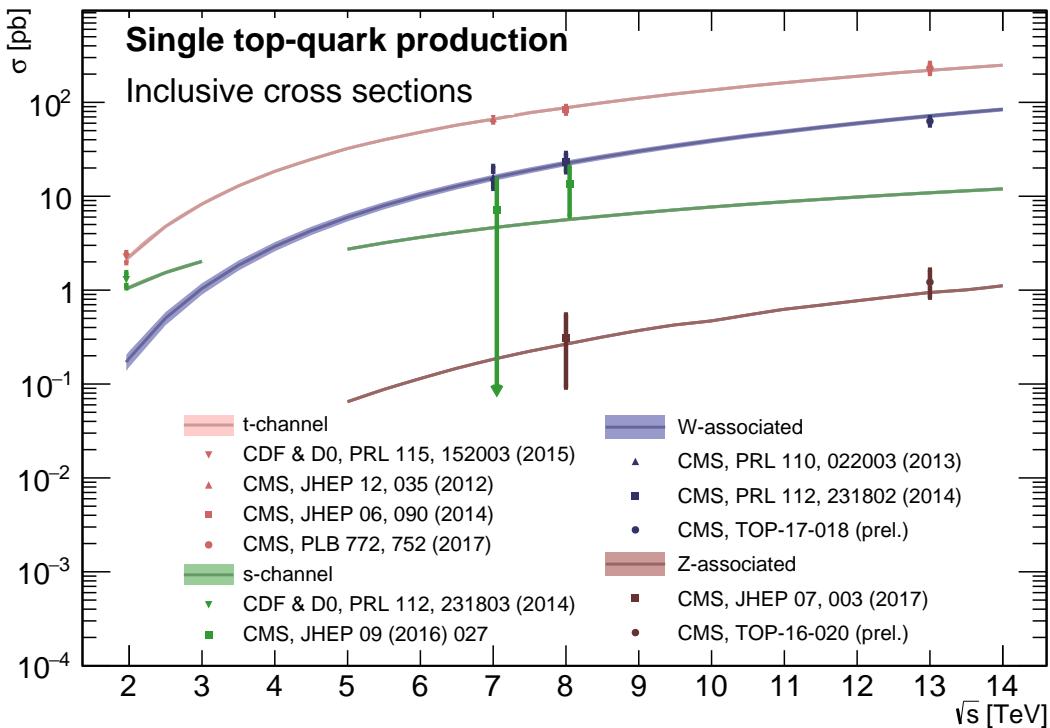


Figure 1.4: Summary of the measurements of the single top quark production cross section as function of the centre-of-mass energy. Figure taken from [22].

226 1.5 Effective field theories

227 Problems can be simplified if one looks at the relevant scale of the process that one want to
 228 investigate, for example the chemical properties of an hydrogen atom can be described without
 229 any knowledge of quark interactions inside the proton. In this case, the proton can be considered
 230 the elementary object (indivisible) due to the fact that the binding energy of the constituents is
 231 much bigger than the energy of the electron in orbit around the proton. Effective field theories
 232 are based on this kind of separation of different energy scales in a system [23]. Effective field
 233 theories can be used for theories where the perturbative expansion cannot be trusted, e.g. QCD
 234 at low energy, or as bottom up approach to look for new physics in a model independent way.
 235 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 [24] 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 will be used further.

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

236 Historically, charged currents were flavour universal and the later discovered parity violation of
 237 the weak interaction led to the V-A structure. After this the $SU(2)_L$ symmetry was postulated
 238 and the existence of neutral currents was predicted. The effective Lagrangian used then (given
 239 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 $\Lambda_F \approx 348 \text{ GeV}$ [23]. From the SM Lagrangian, one could also calculate the muon decay. Considering that the momenta involved are small compared to the W boson mass, the propagator's denominator can be expanded as [1]

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

Looking at the first term, and identifying

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

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

An effective theory is thus a Taylor expansion in the ratio of two scales and the only remnants of the full theory at low energies are the symmetries and the values of the coupling constants. If the expansion parameter is small, one can truncate the series leading to the Lagrangian containing a finite number of free coefficients, making predictions possible. The error on these predictions are 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 [\[25\]](#), 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. [\[25\]](#). 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.

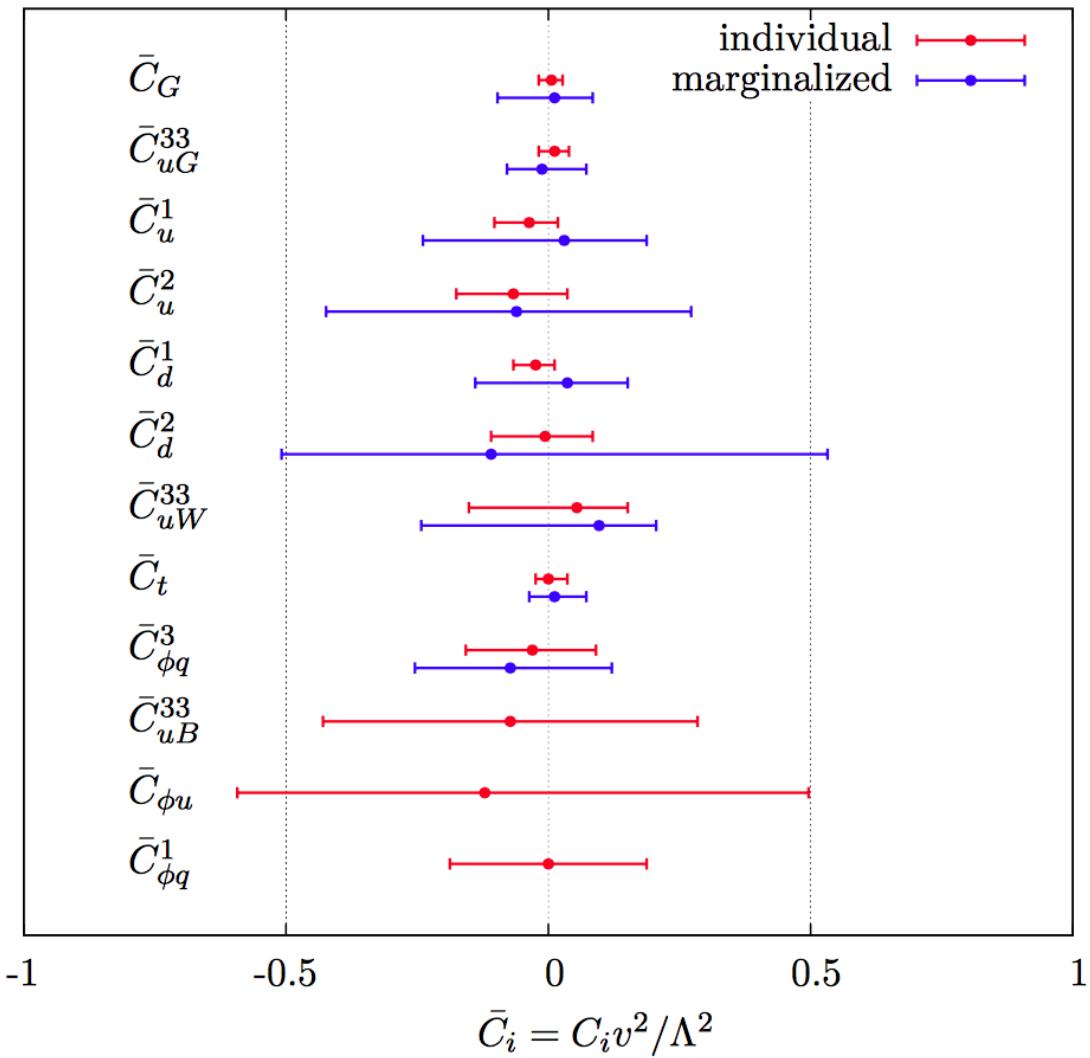


Figure 1.5: Global fit results of top quark effective field theory to experimental data including all constrained operators at dimension six. For the operators, the Warsaw basis of [25] is used. The bounds are set on the Wilson coefficients of various operators contributing to top quark production and decay in two cases (red) all other coefficients set to zero, or (blue) all other coefficients marginalized over. Figure taken from [26].

258 **1.6 Motivation for new physics**

259 Many high energy experiments confirm the success of the SM. In particular the scalar boson,
 260 the cornerstone of the SM, has consecrated the theory. Unfortunately there are also strong
 261 indications that the SM ought to be a lower energy expression of a more global theory. The
 262 existence of physics beyond the SM (BSM) [27] is strongly motivated. These motivations are
 263 based on direct evidence from observation such as the existence of neutrino masses, the existence
 264 of dark matter and dark energy, or the matter-antimatter asymmetry, and also from theoretical
 265 problems such as the hierarchy problem, the coupling unification or the large numbers of free
 266 parameters in the SM.

267 In the SM, the neutrinos are assumed to be massless, while experiments with solar, atmospheric,
 268 reactor and accelerator neutrinos have established that neutrinos can oscillate and change flavour
 269 during flight [4, 5]. These oscillations are only possible when neutrinos have masses. The
 270 flavour neutrinos (ν_e , ν_μ , ν_τ) are then linear expressions of the fields of at least three mass
 271 eigenstate neutrinos ν_1 , ν_2 , and ν_3 .

272 The ordinary or baryonic matter described by the SM describes only 5% of the mass (energy)
 273 content of the universe. Astrophysical evidence indicated that dark matter is contributing
 274 to approximately 27%, and dark energy to 68% of the content of the universe. From the
 275 measurements of the temperature and polarizations anisotropies of the cosmic microwave
 276 background by the Planck experiment [28], the density of cold non baryonic matter is determined.
 277 Cold dark matter is assumed to be only sensitive to the weak and gravitational force, leading
 278 to only one possible SM candidate: the neutrino. However, these are too light to account for
 279 the vast amount of dark matter and other models are needed. Dark energy is assumed to be
 280 responsible for the acceleration in the expansion of the universe [29].

281 At the Big Bang matter and antimatter are assumed to be produced in equal quantities.
 282 However, it is clear that we are surrounded by matter. So where did all the antimatter go?
 283 In 1967, Sakharov identified three mechanisms that are necessary to obtain a global matter
 284 antimatter asymmetry [30]. These mechanisms are those of baryon and lepton number violation,
 285 that at a given moment in time there was a thermal imbalance for the interactions in the universe,
 286 and there is charge C and charge parity CP violation⁴.

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

292 The hierarchy problem [31] is related to the huge difference in energy between the weak
 293 scale and the Planck scale. The vev of the Brout-Englert-Higgs field determines the weak scale
 294 that is approximately 246 GeV. The radiative corrections to the scalar boson squared mass m_H^2 ,
 295 coming from its self couplings and couplings to fermions and gauge bosons, are quadratically

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.

296 proportional to the ultraviolet momentum cut-off Λ_{UV} . This cut-off is at least equal to the energy
 297 to which the SM is valid without the need of new physics. For the SM to be valid up to the
 298 Planck mass, the correction to m_H^2 becomes thirty orders of magnitude larger than m_H^2 . This
 299 implies that an extraordinary cancellation of terms should happen. This is also known as the
 300 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

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

301 As one can see the correction term to m_H^2 is much larger than m_H^2 itself. By introducing BSM
 302 physics models that introduce new scalar particles at the TeV scale that couple to the scalar
 303 boson one can cancel the Λ_{UV}^2 divergence and avoid this fine-tuning.

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

311 1.7 An effective approach beyond the SM: FCNC involving a top 312 quark

313 The closeness of the top quark mass to the electroweak scale led physicist to believe that it
 314 is a sensitive probe for new physics. Studying its properties is therefore an important topic
 315 of the experimental program at the LHC. Several extensions of the SM enhance the FCNC
 316 branching ratios and can be probed at the LHC [17], from which some of them are shown in
 317 Table 1.7. Previous searches have been performed at the Tevatron by the CDF [32] and D0 [33]
 318 collaborations, and at the LHC by the ATLAS [34–38] and CMS [39–43] collaborations.

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

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

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

324 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. Considering that Λ is large, contributions suppressed by powers of Λ greater than two are neglected. Moreover, all four fermion operators are omitted for the rest of this thesis. The Warsaw basis is adopted for the independent effective operators [25], parametrising the new physics effects relevant for the flavour changing neutral current interactions of the top quark as, all flavour indices understood,

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

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

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

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

The hermitian derivative operator is defined as

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

After electroweak symmetry breaking the operators induce [17, 44] 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, and the field strength tensors of the photon A_μ , the gluon field $G_\mu^{1\dots 8}$, and the Z boson Z_μ^0 are defined as

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

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

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

$$\begin{aligned}
 \kappa_{tgq} f_{gq}^L &= \frac{\nu}{g_s \Lambda} [\bar{c}_{uG}]_{i3}^*, & \kappa_{tgq} f_{gq}^R &= \frac{\nu}{g_s \Lambda} [\bar{c}_{uG}]_{3i}, \\
 \kappa_{t\gamma q} f_{\gamma q}^L &= \frac{\nu}{g \Lambda} [\cos\theta_W \bar{c}_{uB} - \sin\theta_W \bar{c}_{uW}]_{i3}^*, & \kappa_{t\gamma q} f_{\gamma q}^R &= \frac{\nu}{g \Lambda} [\sin\theta_W \bar{c}_{uB} - \cos\theta_W \bar{c}_{uW}]_{3i}, \\
 \kappa_{tZq} f_{Zq}^L &= -\frac{2\cos\theta_W \nu}{g \Lambda} [\sin\theta_W \bar{c}_{uB} + \cos\theta_W \bar{c}_{uW}]_{i3}^*, & \kappa_{tZq} f_{Zq}^R &= -\frac{2\cos\theta_W \nu}{g \Lambda} [\cos\theta_W \bar{c}_{uB} + \sin\theta_W \bar{c}_{uW}]_{3i}, \\
 \zeta_{tZq} \tilde{f}_{Zq}^L &= -\frac{2\nu^2}{\Lambda^2} [(\bar{c}_{hq}^{(1)} - \bar{c}_{hq}^{(3)})_{i3} + (\bar{c}_{hq}^{(1)} - \bar{c}_{hq}^{(3)})_{3i}^*], & \zeta_{tZq} \tilde{f}_{Zq}^R &= -\frac{2\nu^2}{\Lambda^2} [(\bar{c}_{hu})_{i3} + (\bar{c}_{hu})_{3i}^*], \\
 \eta_{tHq} \hat{f}_{Hq}^L &= \frac{3\nu^2}{2\Lambda^2} [\bar{c}_{uh}]_{3i}^*, & \eta_{tHq} \hat{f}_{Hq}^R &= \frac{3\nu^2}{2\Lambda^2} [\bar{c}_{uh}]_{i3}.
 \end{aligned} \tag{1.38}$$

332 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

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

333 where $\Gamma_{t \rightarrow qX}$ represents the FCNC decay width⁵ for a coupling strength $\delta_{txq}^2 = 1$, and Γ_t the full
 334 decay width of the top quark. In the SM, supposing a top quark mass of 172.5 GeV, the full
 335 width becomes $\Gamma_t^{\text{SM}} = 1.32$ GeV [45].

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

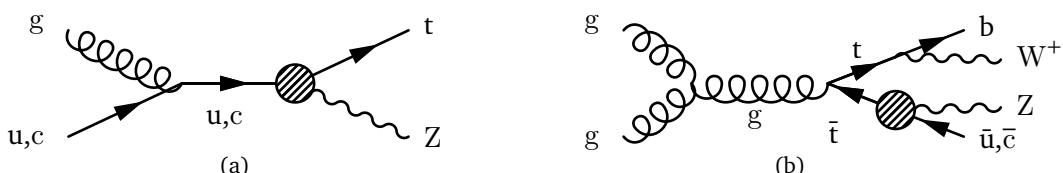


Figure 1.6: Feynman diagrams for the processes with a tZq FCNC interaction, where the FCNC interaction is indicated with the shaded dot. (a) Single top quark production through an FCNC interaction. (b) Top quark pair production with an FCNC induced decay.

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.

341 The observation of top-FCNC interactions has yet to come and experiments have so far only
 342 been able to put upper bounds on the branching ratios. An overview of the best current limits is
 343 given in [Table 1.8](#). In [Figure 1.7](#) a comparison is shown between the current best limits set by
 344 ATLAS and CMS with respect to several BSM model benchmark predictions. From there one can
 345 see that FCNC searches involving a Z or H boson are close to excluding or confirming several
 346 BSM theories. In [Figure 1.9](#), the searches performed by CMS are summarised. For the tZq
 347 vertex, the best limit from CMS comes from Ref. [39] where both single top quark and top quark
 348 pair is studied. The observed (expected) limits 95% CL at 8 TeV for the FCNC tZq interaction
 349 by CMS are $\mathcal{B}(t \rightarrow uZ) < 2.2 \times 10^{-4}$ (2.7×10^{-4}) and $\mathcal{B}(t \rightarrow cZ) < 4.9 \times 10^{-4}$ (12×10^{-4}). In
 350 [Figure 1.8](#), the summary of the 95% confidence level observed limits on the branching ratios
 351 of the top quark decays to a charm or up quark and a neutral boson is given, considering the results from the HERA, the LEP, the Tevatron, and the 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	[38]
$t \rightarrow u\gamma$	single top quark production	1.3×10^{-4}	1.9×10^{-4}	CMS	[41]
$t \rightarrow ug$	single top quark production	4.0×10^{-5}	3.5×10^{-5}	ATLAS	[35]
$t \rightarrow uH$	top quark pair decay	2.4×10^{-3}	1.7×10^{-3}	ATLAS	[37]
$t \rightarrow cZ$	top quark pair decay	2.3×10^{-4}	3.2×10^{-4}	ATLAS	[38]
$t \rightarrow c\gamma$	single top quark production	2.0×10^{-3}	1.7×10^{-3}	CMS	[41]
$t \rightarrow cg$	single top quark production	2.0×10^{-4}	1.8×10^{-4}	ATLAS	[35]
$t \rightarrow cH$	top quark pair decay	2.2×10^{-3}	1.6×10^{-3}	CMS	[37]

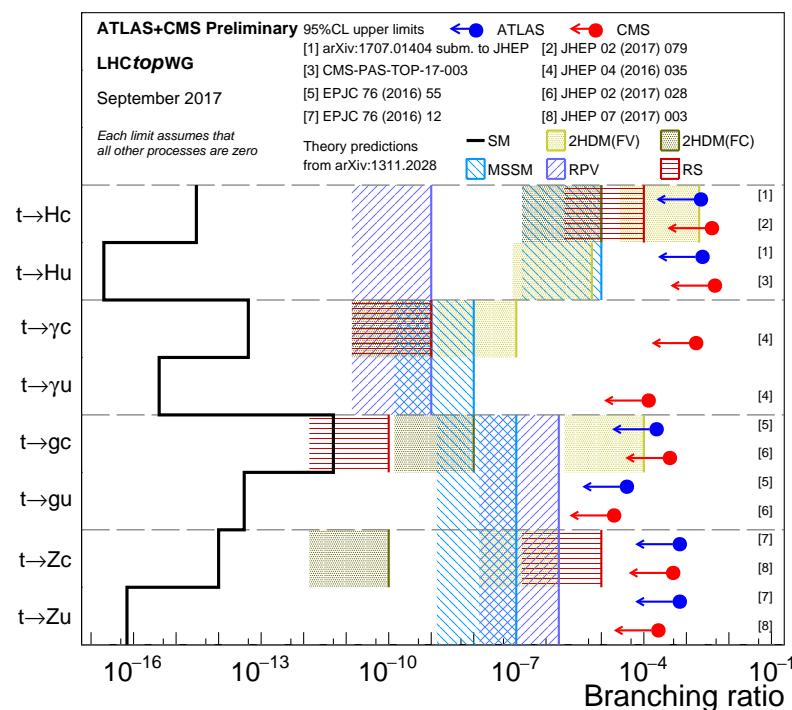


Figure 1.7: Current best limits set by CMS and ATLAS for top-FCNC interactions. Figure taken from [22]. (TO DO Remake with new atlas results)

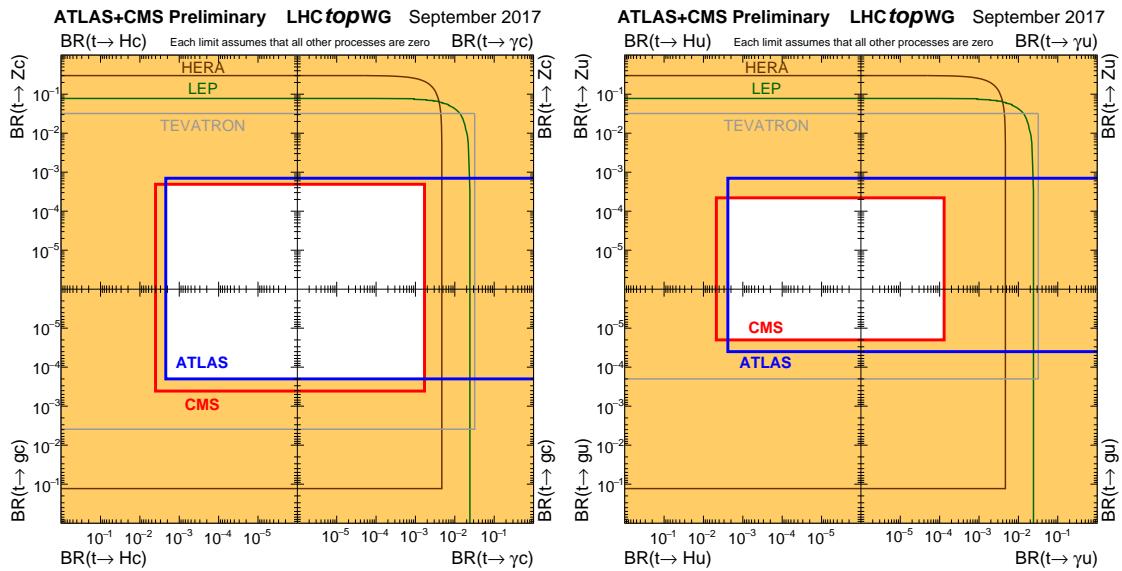
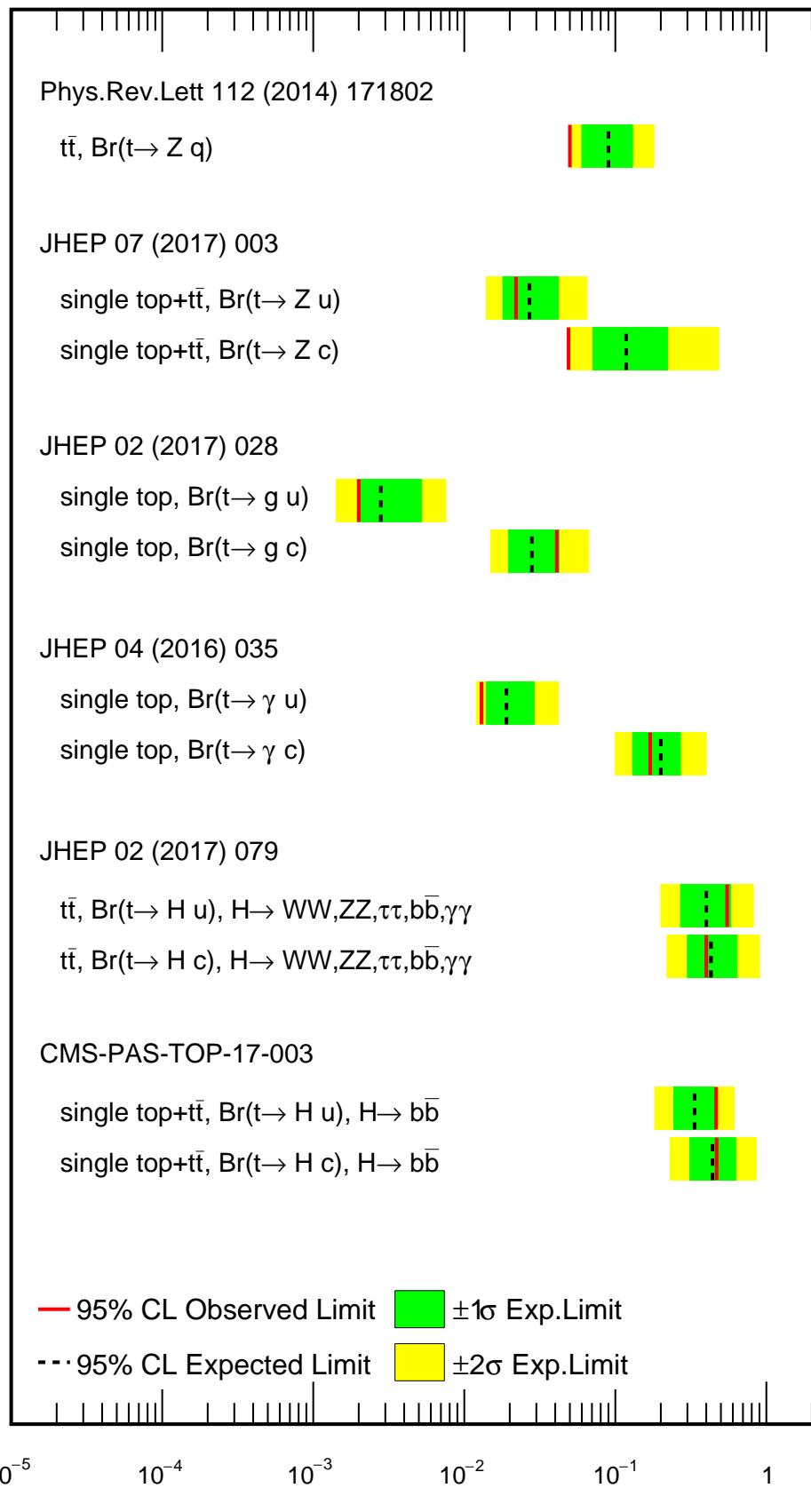


Figure 1.8: Summary of the current 95% confidence level observed limits on the branching ratios of the top quark decays via flavour changing neutral currents to a charm (left) or up (right) quark and a neutral boson. The coloured lines represent the results from HERA (the most stringent limits between 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 [46].

CMS preliminary September 2017



Experimental set-up

2

354 The main objective of the Large Hadron Collider (LHC) was the search for the Brout-Englert-
 355 Higgs boson. The Large Electron Positron (LEP) [47] and Tevatron [48] experiments had
 356 established that the mass of the scalar boson has to be larger than 114 GeV [49, 50], and smaller
 357 than approximate 1 TeV due to unitarity and perturbativity constraints [51]. On top of this,
 358 the search for new physics such as supersymmetry or the understanding of dark matter were
 359 part of the motivation for building the LHC. Since the start of its operation, the LHC is pushing
 360 the boundaries of the Standard Model, putting the most stringent limits on physics beyond the
 361 Standard Model as well as precision measurements of the parameters of the Standard Model. A
 362 milestone of the LHC is the discovery the scalar boson in 2012 by the two largest experiments
 363 at the LHC [10, 11].

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

370 2.1 The Large Hadron Collider

371 The LHC has started its era of cutting edge science on 10 September 2008 [52] after approval by
 372 the European Organisation of Nuclear Research (CERN) in 1995 [53]. Installed in the previous
 373 LEP tunnels, the LHC consists of a 26.7 km ring, that is installed between 45 and 170 m under
 374 the French-Swiss border amidst Cessy (France) and Meyrin (Switzerland). Built to study rare
 375 physics phenomena at high energies, the LHC can accelerate two type of particles, protons or
 376 ions Pb^{45+} , and provides collisions at four interaction points, where the particle bunches are
 377 crossing. Experiments for studying the collisions are installed on each interaction point.

378 As can be seen in Figure 2.1, the LHC is 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

accelerated to 1.4 GeV and each pack is split up in twelve bunches with 25 ns spacing for Run 2 (50 ns for Run 1). The proton synchrotron (PS) then increases their energy to 25 GeV before the super proton synchrotron (SPS) increases the proton energy up to 450 GeV. Each accelerator ring expands in radius in order to reduce the energy loss of the protons by synchrotron radiation¹. Furthermore, the magnets responsible for the bending of the proton trajectories have to be strong enough to sustain to higher proton energy. Ultimately, the protons are injected into opposite 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) [54]. Before the start of the LHC in 2010, the previous energy record was held by the Tevatron collider at Fermilab, colliding proton 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).

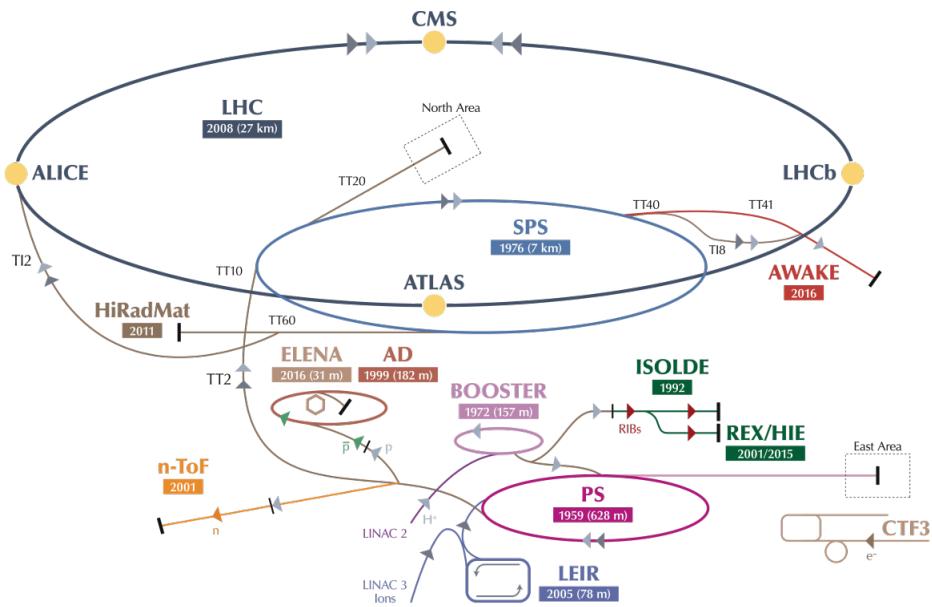


Figure 2.1: Schematic representation of the accelerator complex at CERN [55]. The LHC (dark blue) is the last element in chain of accelerators. Protons are successively accelerated by LINAC 2, the Booster, the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS) before entering the LHC.

Inside the LHC ring [56], the protons are accelerated by the means of radio frequency cavities, while 1232 dipole magnets of approximately 15 m long, each weighing 35 t ensure the deflection of the beams. The two proton beams circulate in opposite direction in separate pipes inside of the magnet. Through the use of a strong electric current in the coils around the beam pipe, magnetic fields are generated and cause the protons to bend in the required orbits. In order to get the coil to become superconducting and able to produce - with the aid of an iron return yoke - a strong magnetic field of 8.3 T, the magnet structure is surrounded by a vessel. This vessel is filled with liquid Helium making it possible to cool down the magnet to 1.9 K. In order to get

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

401 more focussed and stabilised proton beams, additional higher-order multipole and corrector
 402 magnets are placed along the LHC beam line.

403 The LHC is home to seven experiments, each located on an interaction point:

- 404 • A Toroidal LHC ApparatuS (ATLAS) [57] and the Compact Muon Solenoid (CMS) [58]
 405 experiments are the two general purpose detectors at the LHC. They both have a hermetic,
 406 cylindrical structure and were designed to search for new physics phenomena along with
 407 precision measurements of the Standard Model. The existence of two distinct experiments
 408 allows cross-confirmation of any discovery.
- 409 • A Large Ion Collider Experiment (ALICE) [59] and the LHC Beauty (LHCb) [60] exper-
 410 iments are focusing on specific phenomena. ALICE studies strongly interacting matter
 411 at extreme energy densities where a quark-gluon plasma forms in heavy ion collisions
 412 (Pb-Pb or p-Pb). LHCb searches for differences between matter and antimatter with the
 413 focus on b physics..
- 414 • The forward LHC (LHCf) [61] and the TOTal cross section, Elastic scattering and diffraction
 415 dissociation Measurement (TOTEM) [62] experiments are two smaller experiments that
 416 focus on head on collisions. LHCf consists of two parts placed before and after ATLAS
 417 and studies particles created at very small angles. TOTEM is placed in the same cavern as
 418 CMS and measures the total proton-proton cross section and studies elastic and diffractive
 419 scattering.
- 420 • The Monopoles and Exotics Detector At the LHC (MoEDAL) [63] experiment is situated
 421 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 [64] 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 event of interest and L the machine luminosity. This luminosity depends only on the beam parameters and is for a Gaussian beam expressed as

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

422 The number of particles per bunch is expressed by N_b , while n_b is the number of bunches
 423 per beam, f_{rev} the revolution frequency, γ_r the relativistic gamma factor, ϵ_n the normalized
 424 transverse beam emittance - a quality for the confinement of the beam , β^* the beta function at
 425 the collision point - a measurement for the width of the beam, θ_c the angle between two beams
 426 at the interaction point, σ_z the mean length of one bunch, and σ^* the mean height of one bunch.

427 In Equation 2.2, the blue part represents the stream of particles, the red part the brilliance, and
 428 the green part the geometric reduction factor due to the crossing angle at the interaction point.

429 The peak design luminosity for the LHC reached in 2016 is $10^{34} \text{ m}^{-2}\text{s}^{-1}$, which leads to about
 430 1 billion proton interactions per second. In 2016, the LHC was around 10% above this design
 431 luminosity [65]. The luminosity is not a constant in time since it diminishes due to collisions
 432 between the beams, and the interaction of the protons and the particle gas that is trapped in
 433 the centre of the vacuum tubes due to the magnetic field. The diffusion of the beam degrades
 434 the emittance and therefore also the luminosity. For this reason, the mean lifetime of a beam
 435 inside the LHC is around 15 h. The integrated luminosity - the luminosity provided in a certain
 436 time range - recorded by CMS and ATLAS over the year 2016 is given in Figure 2.2. In Run 2,
 the peak luminosity is $13\text{-}17 \cdot 10^{33} \text{ cm}^{-2}\text{s}^{-1}$ compared to $7.7 \cdot 10^{33} \text{ cm}^{-2}\text{s}^{-1}$ in Run 1.

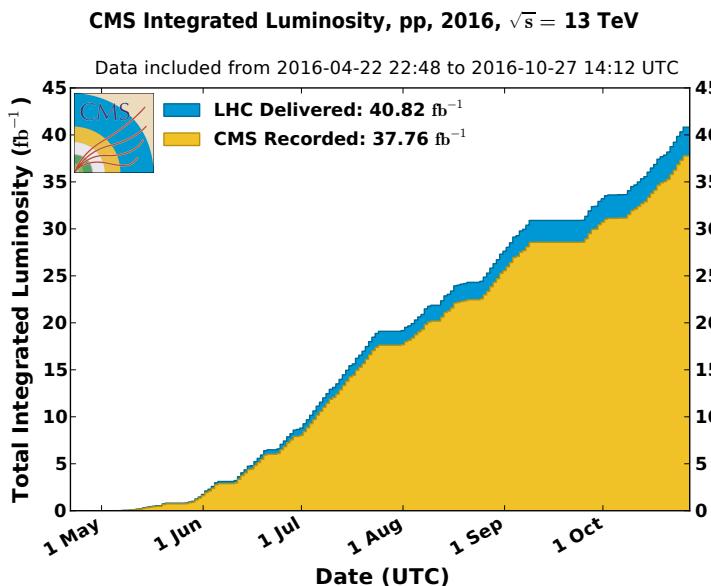


Figure 2.2: Cumulative offline luminosity measured versus day delivered to (blue), and recorded by CMS (orange) during stable beams and for proton collisions at 13 TeV centre-of-mass energy in 2016. The delivered luminosity accounts for the luminosity delivered from the start of stable beams until the LHC requests CMS to turn off the sensitive detectors to allow a beam dump or beam studies [66].

437

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

443 2.2 The Compact Muon Solenoid

444 At one of the collision points of the LHC, the CMS detector[67–69] is placed. Weighing 14 000 t,
 445 this cylindrical detector is about 28.7 m long and 15 m in diameter. It has an onion like structure

of several specialised detectors and contains a superconducting solenoid with a magnetic field of 3.8 T. Living in a hadronic environment, multi-jet processes produced by the strong interaction are a main source of background for rare physics processes. Therefore, good identification, momentum resolution, and charge determination of muons, electrons and photons are one of the main goals of the CMS detector. Additionally, a good charged particle momentum resolution and reconstruction efficiency in the inner tracker provides identification for jets coming from b quarks or tau particles can be identified. Also the electromagnetic resolution for an efficient photon and lepton isolation as well as a good hadronic calorimeter for the missing transverse energy² were kept into account while designing CMS. In Figure 2.3, an overview of the CMS detector is shown.

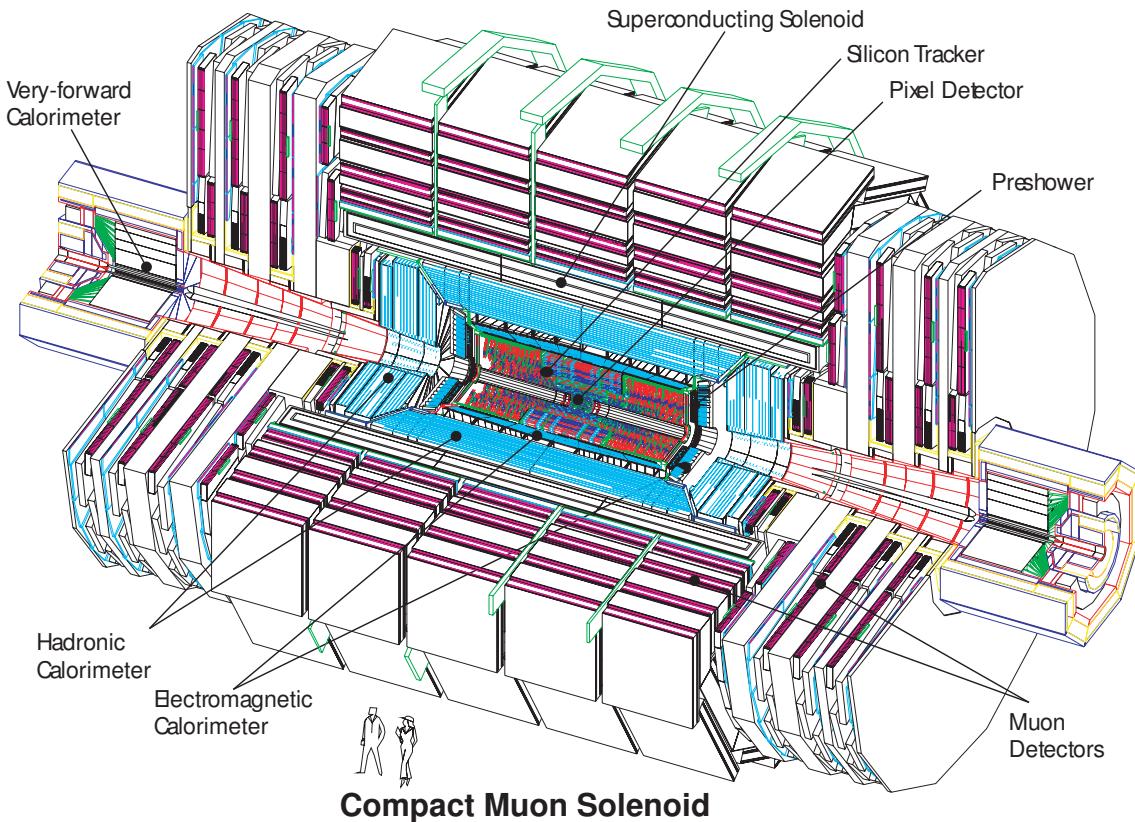


Figure 2.3: Mechanical layout of the CMS detector. Figure taken from [70].

455

456 2.2.1 CMS coordinate system

The coordinate system used by CMS can be found in Figure 2.4. 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

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

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

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

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

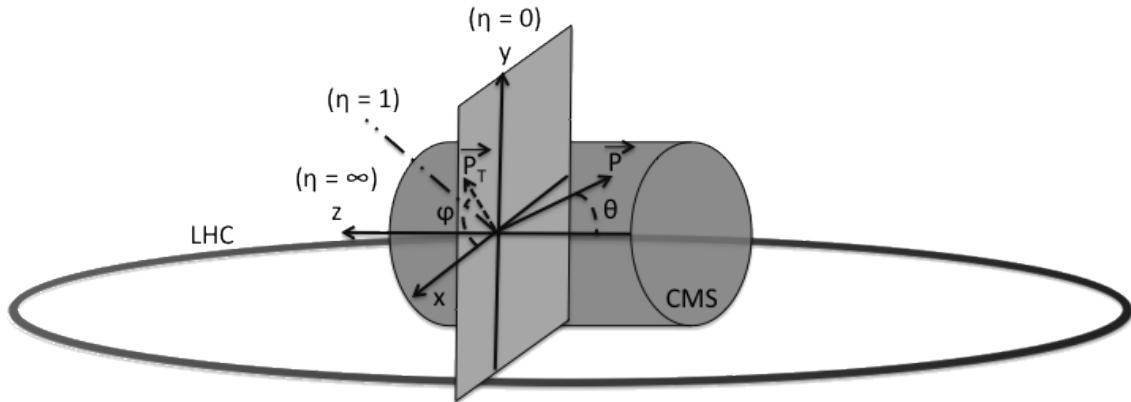


Figure 2.4: Representation of the coordinate system used by CMS. The point of origin is put at the collision point. The x-axis points towards the centre of the LHC ring such that the z-axis lies tangent to the beam axis.

- 458

459 2.2.2 Towards the heart of CMS

- 460 The CMS detector can be divided into two parts. A central barrel is placed around the beam
 461 pipe ($|\eta| < 1.4$), and two plugs (end caps) ensure the hermeticity of the detector. In [Figure 2.3](#)
 462 and [Figure 2.5](#) the onion like structure of the CMS detector is visible. The choice of a solenoid of
 463 12.9 m long and 5.9 m diameter gives the advantage of bending the particle trajectories in the
 464 transverse plane. The hadronic calorimeter ([Section 2.2.2.3](#)), the electromagnetic calorimeter
 465 ([Section 2.2.2.4](#)) and the tracker ([Section 2.2.2.5](#)) are within the solenoid ([Section 2.2.2.2](#)),
 466 while the muon chambers ([Section 2.2.2.1](#)) are placed outside the solenoid. The data used for
 467 the search presented in this thesis is collected after the long shutdown 1. After discussing each
 468 part of CMS in their Run 1 configuration, [Section 2.2.4](#) elaborates on their different upgrades
 469 for the data collected in Run 2.

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.

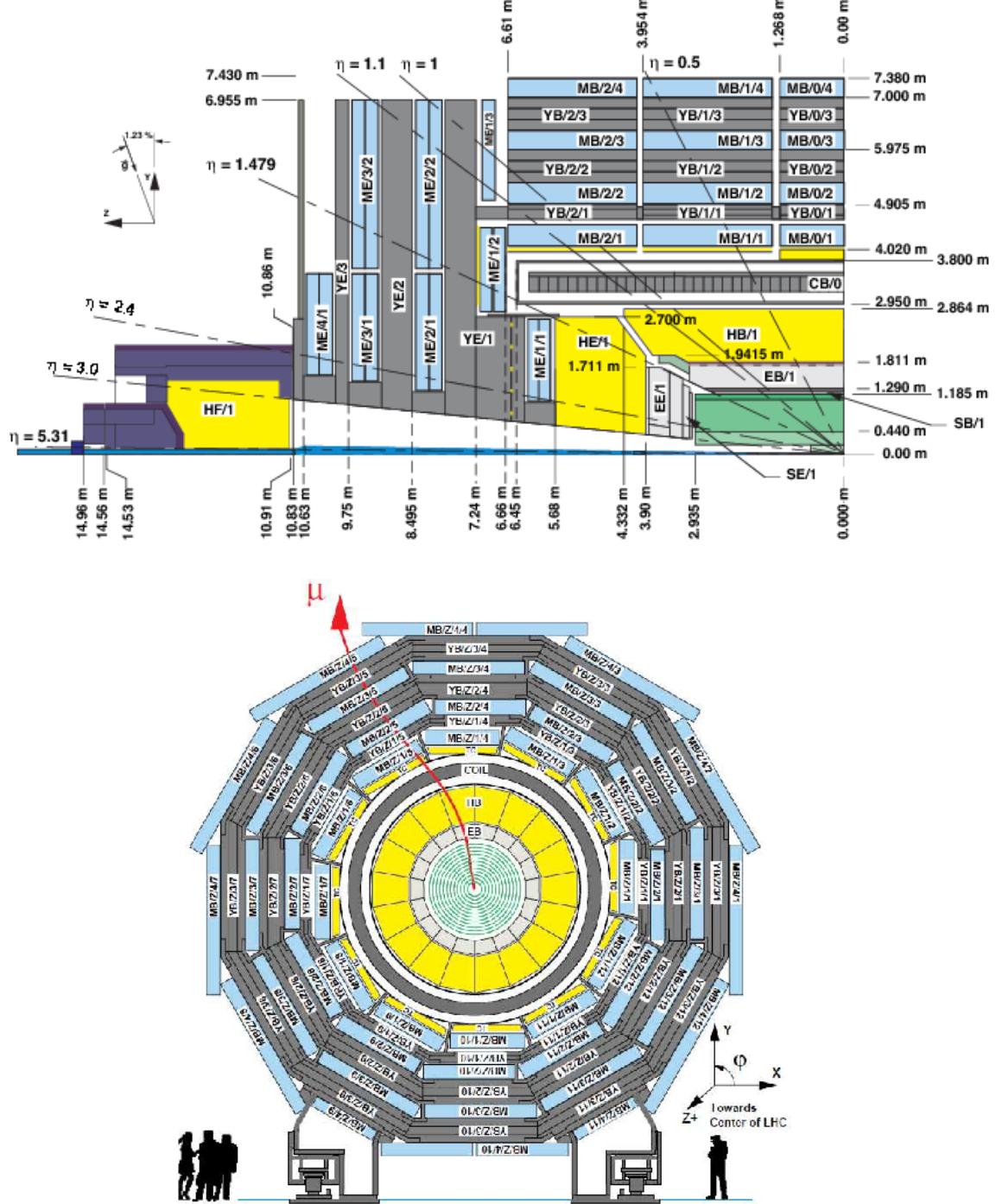


Figure 2.5: Schematic view of the CMS detector in the Run 1 configuration. The longitudinal view of one quarter of the detector is given on top, while the transversal view is shown on the bottom. The muon system barrel elements are denoted as $MBZ/N/S$, where $z = -2 \dots +2$ is the barrel wheel number, $n = 1 \dots 4$ the station number and $S = 1 \dots 12$ the sector number. Similarly, the steel return yokes are denoted as $YBZ/N/S$. The solenoid is denoted as $CB0$, while the hadronic calorimeter is denoted as HE (end cap)/ HB (barrel)/ HF (forward) and the electromagnetic calorimeter as EE (end cap)/ EB (barrel). The green part represents the tracking system (tracker + pixel). Figure taken from [71].

470 **2.2.2.1 Muon system**

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

476 The muon system is divided into three parts, shown in Figure 2.6. The muon rate and neutron
 477 induced backgrounds are small and the magnetic field is very low for the barrel, thus CMS can
 478 use drift tube (DT) chambers. For the end caps however, the muon and background flux is much
 479 higher and there is a need to use cathode strip chambers (CSC) which are able to provide a
 480 faster response, higher granularity and have a better resistance against radiation. In order to
 481 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 2.5 the arrangement is shown.

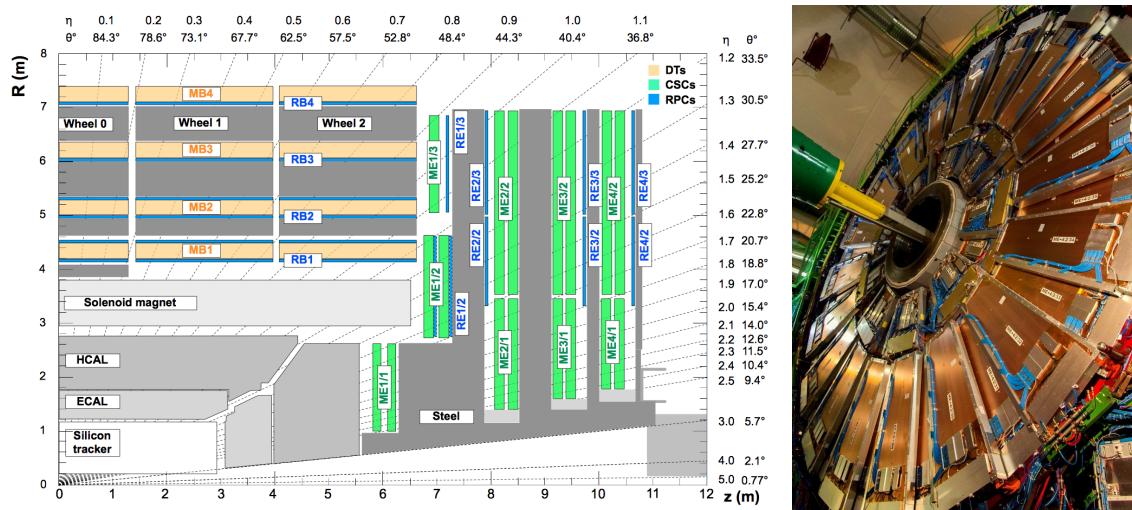


Figure 2.6: (Left) Schematic view of one quarter of the CMS muon system in the Run 1 configuration. The cathode strip chambers (CSC) are shown in green, the drift tubes (DT) are shown in yellow, while the resistive plate chambers (RPC) are shown in blue. Figure taken from [71]. (Right) Cathode strip chambers (ME+4/2 chambers on YE+3). Photo taken from [72].

482

483 Providing a measurement for $|\eta| < 1.2$, the DT chambers in the barrel are on average
 484 $2 \times 2.5 \text{ m}^2$ in size and consist of 12 layers of DT cells⁵ arranged in three groups of four. The
 485 $r\phi$ coordinate is provided by the two outside groups, while the middle group measures the
 486 z coordinate. For the outer muon station, the DT chambers contain only 8 layers of DT cells,
 487 providing a muon position in the $r\phi$ plane. There are four CSC stations in each end cap, providing
 488 muon measurements for $0.9 < |\eta| < 2.4$ (Run 1 configuration). These CSCs are multi-wired
 489 proportional chambers that consist of 6 anode wire planes crossed by 7 copper strips cathode
 490 panels in a gas volume. The r coordinate is provided by the copper strips, while the ϕ coordinate
 491 comes from the anode wires, giving a two dimensional position measurement. There are six

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

492 layers of RPCs in the barrel muon system and one layer into each of the first three stations
 493 of the end cap. They are made from two high resistive plastic plates with an applied voltage
 494 and separated by a gas volume. Read out strips mounted on top of the plastic plates detect the
 495 signal generated by a muon passing through the gas volume. The RPCs provide a fast response
 496 with a time resolution of 1 ns and cover a range of $|\eta| < 1.8$ for the Run 1 configuration.

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

NOTE:
check numbers for run
2

503 2.2.2.2 Solenoid

504 Making use of the knowledge of previous experiments like ALEPH and DELPHI at LEP and H1
 505 at HERA, CMS choose for a large super conducting solenoid with a length of 12.9 m and a
 506 inner bore of 5.9 m [69]. With 2 168 turns, a current of 19.5 kA and a total energy of 2.7 GJ, a
 507 large bending power can be obtained for a modestly-sized solenoid. In order to ensure a good
 508 momentum resolution in the forward regions, a favourable length/radius was necessary. In
 509 [Figure 2.7](#), a photo of the CMS solenoid is shown.

510 The solenoid uses a high-purity aluminium stabilised conductor with indirect cooling from
 511 liquid helium, together with fully epoxy impregnation. A four-layer winding is implemented that
 512 can withstand an outward pressure of 64 atm. The NbTi cable is co-extruded by pure aluminium
 513 that acts as a thermal stabilizer and has an aluminium alloy for mechanical reinforcement. The
 514 return of the magnetic field is done by fives wheels, noted by YB in [Figure 2.5](#).

515 2.2.2.3 Hadronic calorimeter

516 The hadronic calorimeter (HCAL) is dedicated to precisely measure the energy of charged and
 517 neutral hadrons via a succession of absorbers and scintillators. This makes it crucial for physics
 518 analyses with hadronic jets or missing transverse energy. The HCAL barrel extends between
 519 $1.77 < r < 2.95$ m, where r is the radius in the transverse plane with respect to the beam. Due
 520 to space limitations, the HCAL needs to be as small as possible and is made from materials
 521 with short interaction lengths⁶. On top of this, the HCAL should be as hermetic as possible and
 522 extend to large absolute pseudo rapidities such that it can proved a good measurement of the
 523 missing transverse energy.

524 The quality of the energy measurements is dependent on the fraction of the hadronic shower
 525 that can be detected. Therefore, the HCAL barrel (HB) inside the solenoid is reinforced by an
 526 outer hadronic calorimeter between the solenoid and muon detectors (HO, see [Figure 2.8](#)),
 527 using the solenoid as extra absorber. This increases the thickness to 12 interaction lengths.

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.

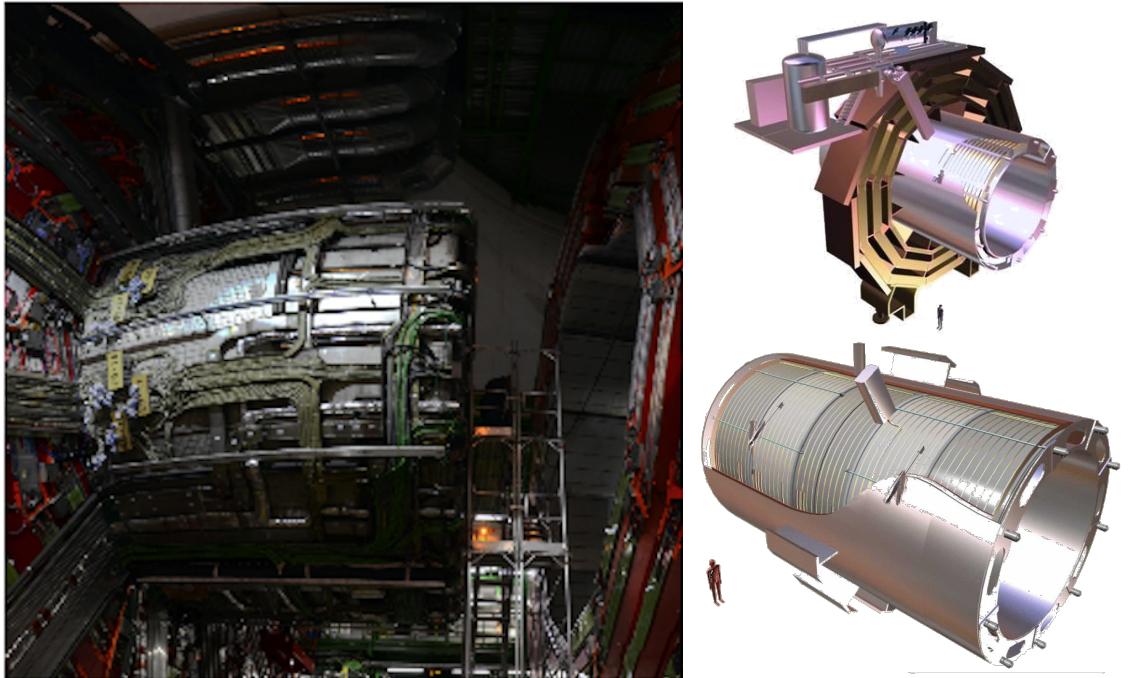


Figure 2.7: (Left) CMS solenoid during the long shutdown in 2013. (Right) An impression of the solenoid magnet taken from [73].

528 The HB and HO provide measurements for $|\eta| < 1.3$, while an end cap on each side (HE,
 529 $1.3 < |\eta| < 3$) and a forward calorimeter (HF, $3.0 < |\eta| < 5.2$) extend the pseudo rapidity
 530 range.

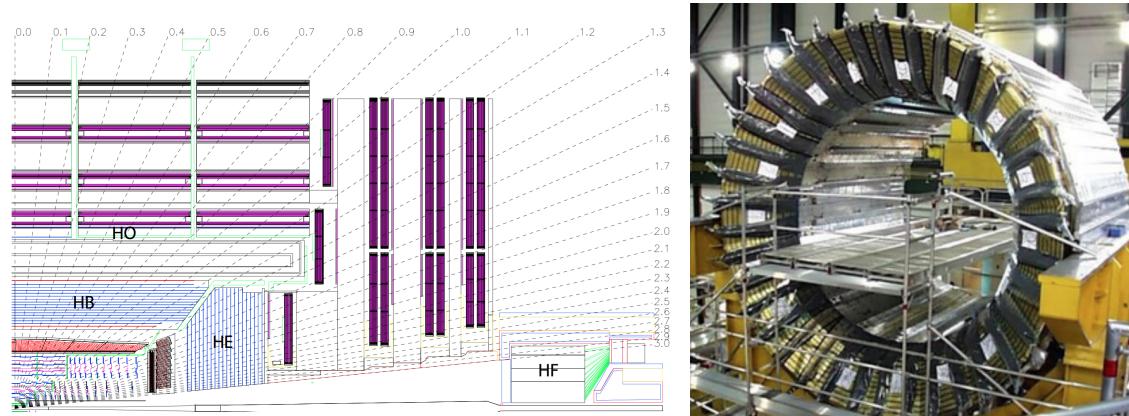


Figure 2.8: (Left) Longitudinal view of the CMS detector showing the locations of the HB, HE, HO, and HF calorimeters. Figure taken from [58]. (Right) CMS barrel calorimeter. Photo taken from [74].

531 The HB is made of 16 absorber plates where most of them are built from brass and others are
 532 made from stainless steel and is about five to ten interaction lengths thick. It is divided in $\eta \times \phi$
 533 towers and contains 2592 read out channels. The HO complements the HB and extends the
 534 reach up to twelve interaction lengths. This subsystem contains 2160 read out channels. The HE

is also composed of brass absorber plates and has a thickness corresponding to approximately ten interaction lengths, with 2592 read out channels.. The HF experiences intense particle fluxes with an expected energy of 760 GeV deposited on average in a proton interaction at a centre-of-mass of 14 TeV, compared to 100 GeV in the rest of the detector. Therefore, these are Cherenkov light detectors made of radiation hard quartz fibers. The main causes of such large energy events are high energy muons, cosmic particles and charged particles from late showering hadrons. During Run 1, it became clear that the glass windows of the photon multiplier tubes (PMTs) had to be replaced which was done during LS1 [75]. The HF represents 1728 read out channels.

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

2.2.2.4 Electromagnetic calorimeter

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

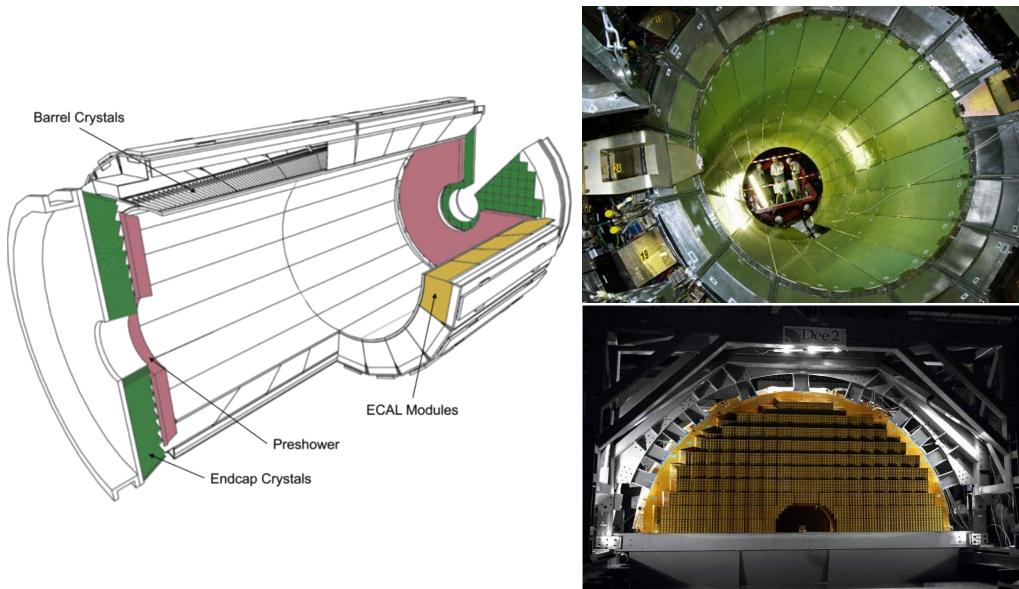


Figure 2.9: (Left) Schematic cross section of the electromagnetic calorimeter taken from [58]. (Right top) The ECAL barrel during construction [76]. (Right bottom) One half of an EE [77].

There are three regions: a central barrel (EB), an endcap region (EE) and a preshower (ES) (Figure 2.9). The EB has an inner radius of 129 cm and corresponds to a pseudo rapidity of $0 < |\eta| < 1.479$. At a distance of 314 cm from the vertex and covering a pseudo rapidity of $1.479 <$

558 $|\eta| < 3.0$, are the EE. They consist of semi-circular aluminium plates from which structural
 559 units of 5×5 crystals (super crystals) are supported. The ES is placed in front of the crystal
 560 calorimeter over the end cap pseudo rapidity range with two planes of silicon strip detectors as
 561 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 [78] for the barrel is 2.10^{-3} rad in ϕ and 10^{-3} in η . For the end caps 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 [58]. The energy resolution is given by

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

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

565 2.2.2.5 Inner tracking system and operations

566 The tracking system (tracker) [79] is the detecting unit closest to the point of interaction.
 567 Responsible for the reconstruction of trajectories from charged particles with $|\eta| < 2.5$ that are
 568 bent by the magnetic field, it provides a measurement of the momentum. The tracker is also
 569 responsible for the determination of the interaction point or vertex. It should be able to provide
 570 high granularity as well as fast read out, and be able to endure high radiation. For this reason,
 571 the CMS collaboration choose silicon detector technology.

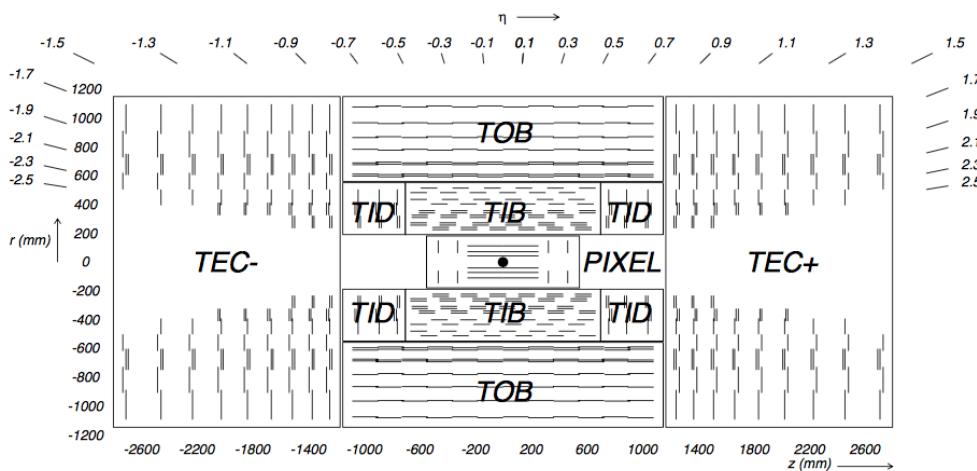


Figure 2.10: Schematic cross section through the CMS tracker. Each line represents a detector module. Double lines indicate back-to-back modules that deliver stereo hits. Figure taken from [58].

572 The tracking system consists of a cylinder of 5.8 m long and 2.5 m in diameter. It is immersed
 573 in a co-axial magnetic field of 3.8 T provided by the solenoid. As shown in Figure 2.10, the

tracker is built up from a large silicon strip tracker with a small silicon pixel tracker inside. The inner pixel region ($4.4 < r < 10.2$ cm), gets the highest flux of particles. Therefore, pixel silicon sensors of $100 \times 150 \mu\text{m}^2$ are used. It consists of three cylindrical barrels that are complemented by two discs of pixel modules at each side. The silicon strip tracker ($20 < r < 116$ cm) has three subdivisions. The Tracker Inner Barrel and Discs (TIB, TID, see Figure 2.12) are composed of four barrel layers accompanied by three discs at each end. The outer part of the tracker - Tracker Outer Barrel (TOB) - consists of 6 barrel layers. In the outer discs, there are nine discs of silicon sensors, referred to as Tracker End Caps (TEC).

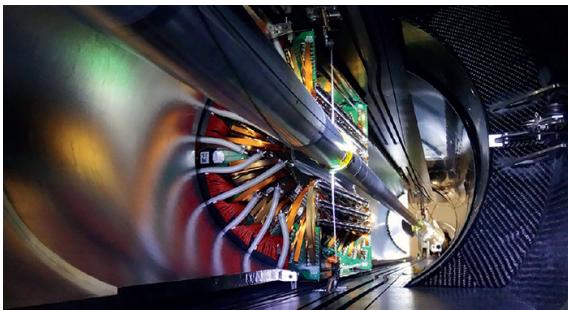


Figure 2.11: The pixel barrel being re-installed after the Long Shutdown in 2015, around the beam pipe at CMS [80].



Figure 2.12: First half of the inner tracker barrel, consisting of three layers of silicon modules [81].

The pixel, shown in Figure 2.11, has 1440 modules that cover an area of about 1 m^2 and have 66 million pixels. It provides a three-dimensional position measurement of the hits arising from the interaction from charged particles with the sensors. In transverse coordinate ($r\phi$), the hit position resolution is about $10 \mu\text{m}$, while $20\text{-}40 \mu\text{m}$ is obtained in the longitudinal coordinate (z). The sensor plane position provides the third coordinate. The TIB/TID, shown in Figure 2.12, delivers up to four $r\phi$ -measurements using a $320 \mu\text{m}$ thick silicon micro-strip sensors. 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 to 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 micor strip 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 sensor 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 rigns of the TIB, TID, and TOB, as wel as rigns 1,2, and 5 of the TECs (blue line in Figure 2.10). 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

605 channels and covers an active area of about 198 m².

606 2.2.3 Data acquisition

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

613 CMS Level-1 Trigger

614 The Level-1 Trigger has to be a flexible, maintainable system, capable of adapting to the
 615 evolving physics programme of CMS [82]. Its output rate is restricted to 100 kHz imposed
 616 by the CMS readout electronics. It is implemented by custom hardware and selects events
 617 containing candidate objects - e.g. ionization deposits consistent with a muon, or energy clusters
 618 corresponding to an electron / photon / tau lepton / missing transverse energy / jet. Collisions
 619 with large momenta can be selected by using scalar sum of the transverse momenta of the jets.

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

627 CMS HLT Trigger

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

633 2.2.4 Phase 1 upgrades

634 Before the start of taking collision data for 13 TeV operations on 3 June 2015, CMS had a long
 635 shutdown (LS1) [83]. During this shutdown, the section of the beryllium beam pipe within CMS
 636 was replaced by a narrower one. This operation required the pixel to be removed and reinserted
 637 into CMS. In Run 2, higher particle fluxes with respect to Run 1 are expected. To avoid long
 638 damage caused by the intense particle flux at the heart of CMS, the tracker is been made ready
 639 to operate at much lower temperature than during Run 1. The electromagnetic calorimeter
 640 preshower system had been damaged during Run 1, therefore the preshower discs were removed,
 641 repaired and reinstalled successfully inside CMS in 2014. To help the discrimination between
 642 interesting low momentum muons coming from collisions and muons caused by backgrounds, a

643 fourth triggering and measurement station for muons was added in each of the end caps. Several
 644 new detectors were installed into CMS for measuring the collision rate within the detector and
 645 monitors beam related backgrounds.

646 During the LS1, the muon system underwent major upgrades [84, 85]. In the fourth station
 647 of each end cap, the outermost rings of CSC and RPC chambers were completed, providing an
 648 angular coverage of $1.2 < |\eta| < 1.8$ for Run 2, increasing the system redundancy, and allowing
 649 tighter cuts on the trigger quality. In order to reduce the environmental noise, outer yoke discs
 650 have been placed on both sides for the end caps. At the innermost rings of the first station,
 651 the CSCs have been upgraded by refurbishing the readout electronics to make use of the full
 652 detector granularity instead of groups of three as was the case for Run 1. In Figure 2.6 (right),
 653 the refurbishing of the CSCs is shown.

654 Since the HF experiences intense particle fluxes, it became clear during Run 1 that the glass
 655 windows of the PMTs need replacing. For the ECAL in Run 1, the energy reconstruction happened
 656 via a weighted sum of the digitized samples [86]. For Run 2 however, the reconstruction had
 657 to be made more resistant for out of time pile up and a multi-fit approach has been set into
 658 place. In this approach, the pulse shape is modelled as a sum of one in-time pulse plus the out
 659 of time pulses [78]. The energy resolution is better than 2% in the central barrel region and
 660 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:

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

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

663 During the LS1, the CMS cooling plant was refurbished [87] and the fluorocarbon cooling
 664 system overhauled. To help to suppress the humidity inside the tracker, new methods for vapour
 665 sealing and insulation were applied. Furthermore, several hundred high-precision sensors are
 666 used to monitor the humidity and temperature. In order to get as dry air as possible, a new
 667 dry-gas plant provides eight times more dry gas (air or nitrogen) than during the first run, and
 668 allows regulation of the flow. As a final addition, the cooling bundles outside the tracker are
 669 equipped with heater wires and temperature sensors in order to maintain safe operations above
 670 the cavern dew point. For the data taking in 2015-2016, the tracker operated at -15°C .

671 In Run 2, with the increase in centre of mass energy and a higher luminosity, a larger number
 672 of simultaneous inelastic collisions per crossing is expected with respect to Run 1. For this, the
 673 CMS Level-1 has been upgraded [88]. All hardware, software, databases and the timing control
 674 system have been replaced for Run 2, where the main changes are that the muon system now
 675 uses the redundancy of the muon detector system earlier to make a high resolution muon trigger.

676 Other upgrades are that the calorimeter system isn't bound any more for streaming data and
677 the global trigger has more Level-1 Trigger algorithms.

678 After the first half of Run 2, the innermost part of detection material in CMS (pixel) was
679 upgraded by adding a fourth layer , enhancing the particle tracking capabilities of CMS. The
680 data used in the framework of this thesis however is from before this upgrade. More information
681 on the Pixel upgrade can be found in Refs. [89, 90].

682 **2.2.5 CMS computing model**

683 The selected data is stored, processed and dispersed via the Worldwide Large Hadron Collider
684 GRID (WLCG) [91, 92]. This has a tiered structure that functions as a single, coherent system.

685 At CERN and the Wigner Research Center for physics, a single Tier-0 is located. The raw data
686 collected by CMS is archived here, and a first reconstruction of the data is done. This data is
687 then already in a file format usable for physics analysis. Furthermore, it is able to reprocess
688 data when new calibrations become available. The Tier-0 site distributes this data to a total of
689 14 Tier-1 centres. They carry out data reprocessing and store real data as well as simulated
690 data. The Tier-1 further distributes the data to over 150 Tier-2 centres. These make the data
691 accessible for physics analysis and are also being used for the production of simulated data. The
692 data is made accessible for physicists around the world.

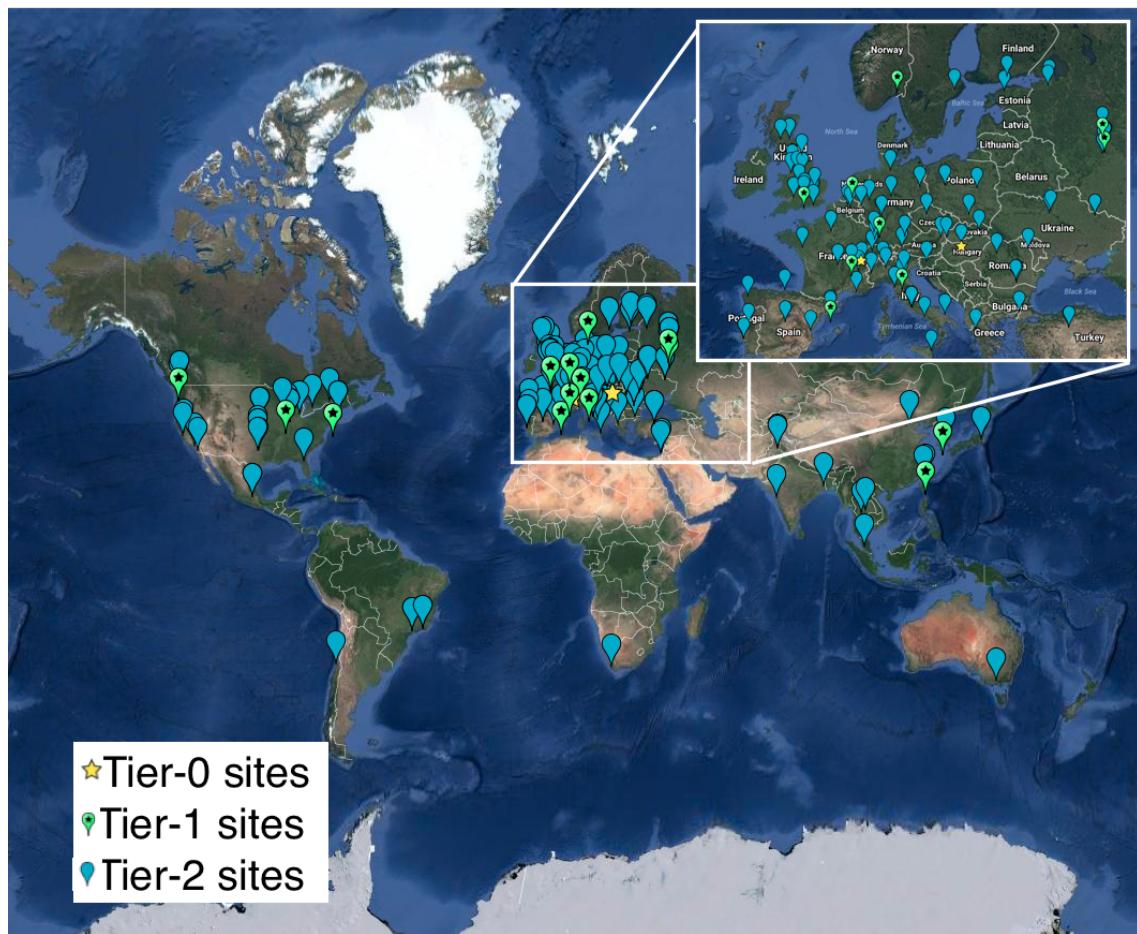


Figure 2.13: Worldwide LHC Computing Grid in 2017 [93].

Analysis techniques

3

694 In order to disentangle the collisions coming from high energy experiments, many tools have
 695 been developed. In [Section 3.1](#), the predictions behind hadron collision at high energies are
 696 presented. These are used to generate events via Monte Carlo event generators, explained in
 697 [Section 3.2](#). Machine learning helps to differentiate between signal- and background like events.
 698 In [Section 3.3](#), the multivariate technique of boosted decision trees is explained. This yields
 699 powerful discriminants for separating signal and background events and provides distributions
 700 that go through template-based maximum likelihood fits. The fitting method used in the search
 701 presented in this thesis is discussed in [Section 3.4](#).

702 3.1 Hadron collisions at high energies

In hadron collisions at sufficiently high momentum transfer, all partons can be approximated as free making it possible to treat 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 $\hat{\sigma}_{ij \rightarrow X}$ [94]

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

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

708 The parton density functions (PDF) [95–97] give the momentum distribution of the proton
 709 amongst its partons at an energy scale μ_F . These function can not be determined from first prin-
 710 ciples and have to obtained from global fits to data. The PDFs are obtained from measurements on

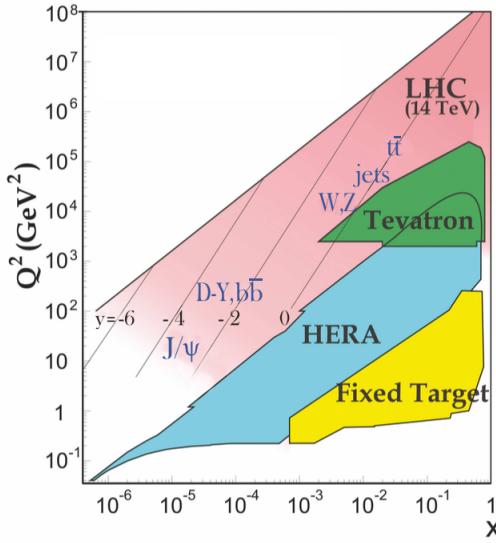


Figure 3.1: Kinematic regions in momentum fraction x and factorisation scale Q^2 probed by fixed-target and collider experiments. Some of the final states accessible at the LHC are indicated in the appropriate regions, where y is the rapidity. In this figure, the incoming partons have $x_{1,2} = (M/14\text{TeV})e^{\pm y}$ 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 [6].

711 deep inelastic scattering using lepton-proton collision by the HERA collider [98], supplemented
 712 with proton-antiproton collisions from Tevatron at Fermi lab [99], and proton collision data
 713 from the ATLAS, CMS and LHCb collaborations at the LHC (Run 1) [100]. These measure-
 714 ments are included in global PDF sets known as the PDF4LHC recommendation [97]. From their
 715 measurement at scale μ_F these PDFs can be extrapolated using the DGLAP equations [101].
 716 The PDFs are used to calculate the cross section of a certain process and are therefore used
 717 as input for the Monte Carlo generators used to make the simulated data samples at the LHC.
 718 In the framework of this thesis, the NLO PDF4LHC15_100 set is used. This set is an envelope
 719 of three sets, CT14, MMHT2014 and NNPDF3.0 [97]. In Figure 3.2 the dependency of the PDFs
 720 on the momentum fraction x is shown for the NNPDF3.0 set on hadronic scale ($\mu_F^2 = (10\text{GeV})^2$)
 721 and LHC scale ($\mu_F^2 = (10^4\text{GeV})^2$). For most values of the momentum fraction, the gluon density
 722 dominates, meaning that it is easier to probe muons than the quarks. For x close to one, the
 723 parton densities of the up and down quarks (the valence quarks of the proton) dominate over
 724 the gluon density. The charm, anti-up, and anti-down quarks have lower densities in general
 725 since those are sea quarks which originate in the proton only through gluon splitting. The
 726 resolution scale Q^2 is typically taken to be the energy scale of the collision. For the top quark pair
 727 production a scale of $Q^2 = (350\text{GeV})^2$ is chosen, meaning that the centre-of-mass energy of the
 728 hard interaction is about twice the top quark mass. The uncertainty on the parton distributions
 729 is evaluated using the Hessian technique [102], where a matrix with a dimension identical to
 730 the number of free parameters needs to be diagonalised. In the case of PDF4LHC15_100 set, this
 731 translates into 100 orthonormal eigenvectors and 200 variations of the PDF parameters in the
 732 plus and minus direction.

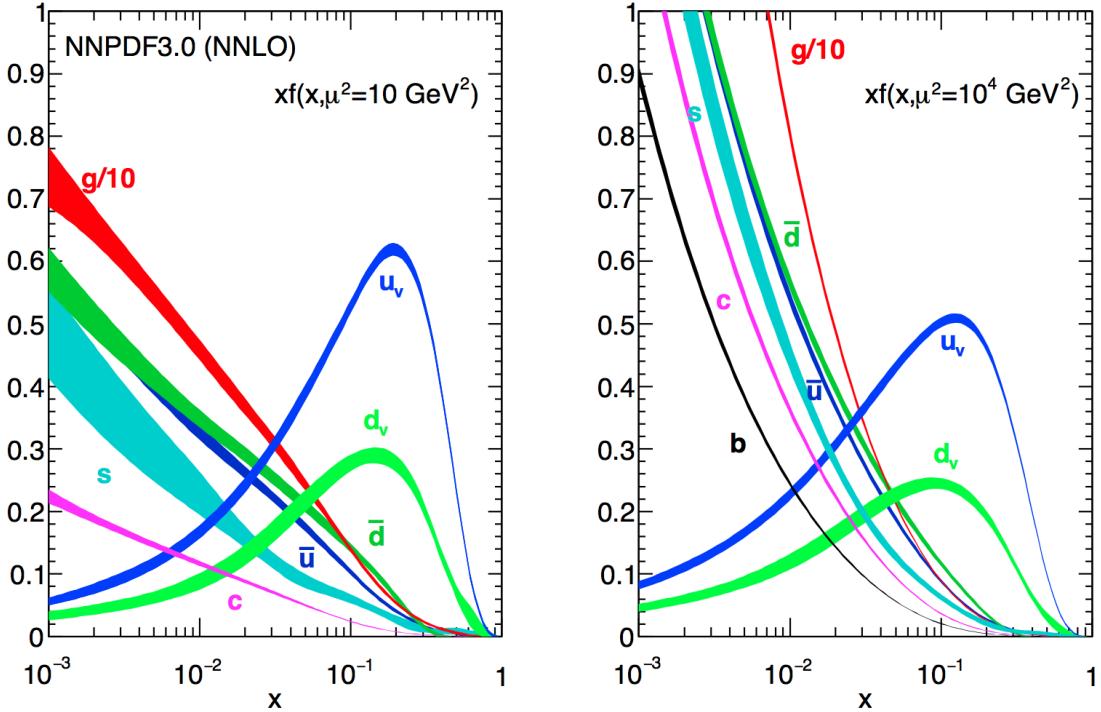


Figure 3.2: The momentum fraction x times the parton distribution functions $f(x)$, where $f = u_v, d_v, \bar{u}, \bar{d}, s, c$, or g as function of the momentum fraction obtained in the NNLO NNPDF3.0 global analysis at factorisation scales $\mu^2 = 10 \text{ GeV}^2$ (left) and $\mu^2 = 10^4 \text{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

At high energies, divergences can appear from quantum fluctuations. For the theory still to be able to describe the experimental regime, a renormalization scale μ_R is used to redefine physical quantities. A consequence of this method is that the coupling constants will run as function of μ_R . Beyond this 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

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

733 with n_f the number of quarks and μ_0 the reference scale on which the coupling is known. The
 734 current world average of the strong coupling constant at the Z boson mass is $\alpha_s(\mu_R = m_Z) =$
 735 0.1181 ± 0.0011 [6]. From Equation 3.3 one can see easily that the coupling strength decreases
 736 with increasing renormalization scale, this known as asymptotic freedom. Additionally, following
 737 the behaviour of $\alpha_s(\mu_R^2)$, a limit $\Lambda_{\text{QCD}} \approx 200 \text{ MeV}$ is found for which α_s becomes larger than
 738 one. Under this limit, the perturbative calculations of observables can no longer be done.

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

Cross sections can be written in terms of interacting vertices contributing to the matrix element (ME) originating from elements of a perturbative series [103], allowing them to be expanded as a power series of the coupling constant α

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

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

744 3.2 Event generation

745 In order to compare reconstructed data with theoretical predictions, collision events are gener-
 746 ated and passed through a simulation of the CMS detector and an emulation of its readout. For
 747 the detector simulation, a so-called Full Simulation package [104, 105] based on the Geant4
 748 toolkit [106] is employed. It allows a detailed simulation of the interactions of the particles
 749 with the detector material.

750 3.2.1 Fundamentals of simulating a proton collision

751 The procedure of to generate $\text{pp} \rightarrow \text{X}$ events can be subdivided into sequential steps [107–109],
 752 as shown in Figure 3.3.

753 The interaction of two incoming protons is often soft and elastic leading to events that are not
 754 interesting in the framework of this thesis. More intriguing are the hard interaction between two
 755 partons from the incoming protons. The matrix elements of a hard scattering process of interest
 756 is the starting point of the generation of events. Monte Carlo techniques are used to sample the
 757 corresponding cross section integral and the resulting sample of events reflect the probability
 758 distribution of a process over its final state phase space. After obtaining the sample of events of
 759 the hard interaction, a parton shower (PS) program is used to simulate the hadronisation of
 760 final state partons into hadrons which then decay further. Additionally, radiation of soft gluons
 761 or quarks from initial or final state partons is simulated. These are respectively referred to as
 762 initial state radiation (ISR) or final state radiation (FSR). Contributions from soft secondary
 763 interactions, the so-called underlying event (UE), and colour reconnection effects are also taken
 764 into account. A brief overview of the employed programs used for the event generation of the
 765 signal and main background processes used in the search presented in the thesis are given in
 766 Section 3.2.2.

NOTE: 764
Should I
add more 765
details? 766

767 3.2.2 Programs for event generation

768 The FEYNRULES package [110] allows the calculation of the Feynman rules in momentum space
 769 for any quantum field theory model. By use of a Lagrangian, the set of Feynman rules associated
 770 with this Lagrangian are calculated. Via the Universal FeynRules Output (UFO) [111] the
 771 results are then passed to matrix element generators.

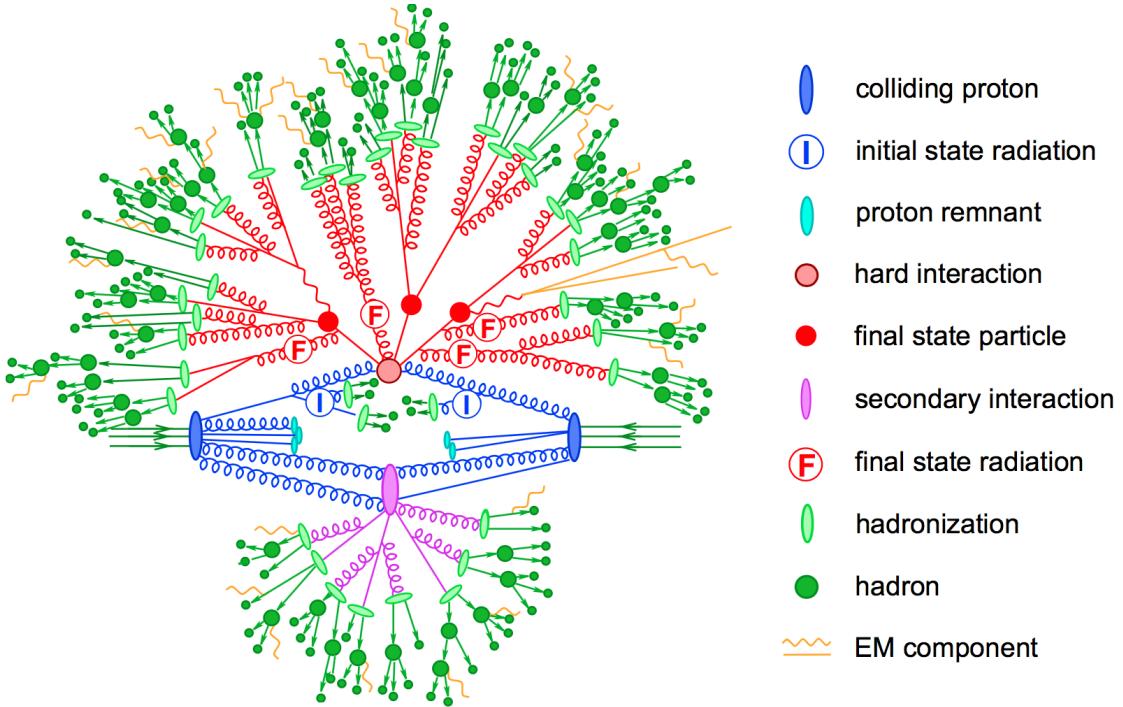


Figure 3.3: Sketch of a hadron collision as simulated by a Monte-Carlo event generator. The red blob in the centre represents the hard collision, surrounded by a tree-like structure representing Bremsstrahlung as simulated by parton showers. The purple blob indicates a secondary hard scattering event. Parton-to-hadron transitions are represented by light green blobs, dark green blobs indicate hadron decays, while yellow lines signal soft photon radiation. Figure taken from [109].

772 The MadGraph program [112] is used to interpret the physics model and calculate the cor-
 773 responding Feynman diagrams and matrix elements. After this, MadEvent [113] is used to
 774 calculate the corresponding partons. These generated parton configurations are then merged
 775 with Pythia [114–116] parton showers using the MLM merging scheme [117].

776 The MadGraph5_aMC@NLO program [118] combines the LO MadGraph [112] and the aMC@NLO
 777 program into a common framework. This combination supports the generation of samples
 778 at LO or NLO together with a dedicated matching to parton showers using the MLM [117]
 779 or FXFX [119] schemes respectively. The FXFX scheme produces a certain fraction of events
 780 with negative weights originating from the subtraction of amplitudes that contain additional
 781 emissions from the NLO matrix element to prevent double-counting.

782 The POWHEG box (versions 1,2) [120–125] contains predefined implementations of various
 783 processes at NLO. It applies the POWHEG method for ME- to PS- matching, where the hardest
 784 radiation generated from the ME has priority over subsequent PS emission to remove the overlap
 785 with the PS simulation.

786 The JHU generator (version 7.02) [126–129] is used to generate the parton level information
 787 including full spin and polarization correlations. It is commonly used for studying the spin and
 788 parity properties of new resonances such as $ab \rightarrow X \rightarrow VV$, where $V = Z, W, \gamma$.

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 that the resonant particle is on-shell. This makes the production and decay amplitude factorize, allowing to perform the simulation of the production and decay of heavy resonances like top quarks or Higgs bosons to be performed in separate steps. The MadSpin program [130] extends this approach and accounts for off-shell effects through a partial reweighting of the events. Additionally, spin correlation effects between production and decay products are taken into account.

The Pythia program (versions 6,8) [114–116] generates events of various processes at LO. However more commonly it is only used for its PS simulation and is then interfaced with other LO and NLO event generators to perform subsequent parton showering, hadronisation, and simulation of the underlying event. In this thesis the underlying event tunes [131] are the CUETP8M2T4, CUETP8M1 and CUETP8M2.

The detector response is simulated via the Geant4 [106] program. This program tracks the particles through the detector material via a detailed description of the detector and generates several hits throughout several sensitive layers. In addition, the response of the detector electronics to these hits are simulated.

3.2.3 Generating FCNC top-Z interactions

The FCNC processes are generated by interfacing the Lagrangian in Equation 1.36 with 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 $f_{Xq}^R = 1$. The signal rates are estimated by use of the MadGraph5_aMC@NLO program for estimating the partial widths. The anomalous couplings are left free to float for this estimation, and only one coupling allowed to be non-vanishing at a time. The results are presented in Table 3.1.

The anomalous single top cross sections are calculated by convolution of the hard scattering matrix elements with the LO order set of NN2.3LO [Ball:2012cx] partons densities. The NLO effects are modelled by multiplying each LO cross section by a global k -factor. The LO single top production cross section and the global k -factors for the top-Z production are 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.

The top pair cross sections are derived from the SM $t\bar{t}$ cross section, calculated with MadGraph5_aMC@NLO at NLO ($\sigma_{t\bar{t}} = 6.741 \cdot 10^2 \text{ pb}$), and considering the decay $t\bar{t} \rightarrow (bW^\pm)(X_{qt})$. 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, 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.

The generation at leading order of the single top FCNC process $tZ + 0,1 \text{ jet}$ including a merging technique can not be done since the $tZ+1 \text{ jet}$ also contains contribution of top pair where one

NOTE: W₈₀₉
RH and not
LH?

NOTE: 819
these par-
820 tial widths
are at LO,
how does
this relate
821 to NLO that
is used? Or
is there no
difference?

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_{gqt}/Λ	$t g u$	$3.665220 \cdot 10^5$	$(\kappa_{tg u}/\Lambda)^2$
	$t g c$	$3.664620 \cdot 10^5$	$(\kappa_{tg c}/\Lambda)^2$
$\kappa_{t\gamma q}/\Lambda$	$t\gamma u$	$1.989066 \cdot 10^4$	$(\kappa_{t\gamma u}/\Lambda)^2$
	$t\gamma c$	$1.988904 \cdot 10^4$	$(\kappa_{t\gamma c}/\Lambda)^2$
κ_{tZq}/Λ	$tZ u$	$1.637005 \cdot 10^4$	$(\kappa_{tZ u}/\Lambda)^2$
	$tZ c$	$1.636554 \cdot 10^4$	$(\kappa_{tZ c}/\Lambda)^2$
ζ_{tZq}	$tZ u$	$1.685134 \cdot 10^{-1}$	$(\zeta_{tZ u})^2$
	$tZ c$	$1.684904 \cdot 10^{-1}$	$(\zeta_{tZ c})^2$
η_{tHq}	$tH u$	$1.904399 \cdot 10^{-1}$	$(\eta_{tH u})^2$
	$tH c$	$1.904065 \cdot 10^{-1}$	$(\eta_{tH c})^2$

Table 3.2: Leading order single top production cross section for $pp \rightarrow tZ$ or $\bar{t}Z$, where the new physics scale is given in GeV. The NLO k -factors [132] are given in the last column.

Anomalous coupling	Cross section (pb)	NLO k -factor
$\kappa_{tg u}/\Lambda$	$3.272 \cdot 10^7$	$(\kappa_{tg u}/\Lambda)^2$
$\kappa_{tg c}/\Lambda$	$3.021 \cdot 10^6$	$(\kappa_{tg c}/\Lambda)^2$
$\kappa_{t\gamma u}/\Lambda$	$2.260 \cdot 10^5$	$(\kappa_{t\gamma u}/\Lambda)^2$
$\kappa_{t\gamma c}/\Lambda$	$2.654 \cdot 10^4$	$(\kappa_{t\gamma c}/\Lambda)^2$
$\kappa_{tZ u}/\Lambda$	$1.728 \cdot 10^6$	$(\kappa_{tZ u}/\Lambda)^2$
$\kappa_{tZ c}/\Lambda$	$2.040 \cdot 10^5$	$(\kappa_{tZ c}/\Lambda)^2$
$\zeta_{tZ u}$	7.484	$(\zeta_{tZ u})^2$
$\zeta_{tZ c}$	1.038	$(\zeta_{tZ c})^2$

Table 3.3: Next to leading order top pair cross section for the top-Z FCNC interactions with with a full leptonic decay.

Anomalous coupling	Process	Cross section (pb)
κ_{tZu}/Λ	$t\bar{t} \rightarrow (b\ell^+\nu)(\bar{u}\ell^+\ell^-)$	$2.727008 \cdot 10^5 \left(\kappa_{tZu}/\Lambda\right)^2$
	$t\bar{t} \rightarrow (\bar{b}\ell^-\bar{\nu})(u\ell^+\ell^-)$	$2.727008 \cdot 10^5 \left(\kappa_{tZu}/\Lambda\right)^2$
κ_{tZc}/Λ	$t\bar{t} \rightarrow (b\ell^+\nu)(\bar{c}\ell^+\ell^-)$	$2.726257 \cdot 10^5 \left(\kappa_{tZc}/\Lambda\right)^2$
	$t\bar{t} \rightarrow (\bar{b}\ell^-\bar{\nu})(c\ell^+\ell^-)$	$2.726257 \cdot 10^5 \left(\kappa_{tZc}/\Lambda\right)^2$
ζ_{tZu}	$t\bar{t} \rightarrow (b\ell^+\nu)(\bar{u}\ell^+\ell^-)$	$2.827184 \left(\zeta_{tZu}\right)^2$
	$t\bar{t} \rightarrow (\bar{b}\ell^-\bar{\nu})(u\ell^+\ell^-)$	$2.827184 \left(\zeta_{tZu}\right)^2$
ζ_{tZc}	$t\bar{t} \rightarrow (b\ell^+\nu)(\bar{c}\ell^+\ell^-)$	$2.806801 \left(\zeta_{tZc}\right)^2$
	$t\bar{t} \rightarrow (\bar{b}\ell^-\bar{\nu})(c\ell^+\ell^-)$	$2.806801 \left(\zeta_{tZc}\right)^2$

823 quark is decaying in tZ. Therefore, single top and top pair processes must be made independently
 824 though this violates gauge invariance. In the official release of `MadGraph5_aMC@NLO` this is not
 825 allowed and the generator experts bypassed this by creating the single top process without
 826 the extra hard jet. Since cross section for tZ + 1 jet is relatively small, these contributions
 827 are neglected and the single top signal is generated without any extra jets, while top pair is
 828 generated up to two extra jets.

829 3.2.4 Generating SM background events

830 The SM tZq events were generated using the `MadGraph5_aMC@NLO` generator, interfaced with
 831 `Pythia` version 8.2 [116] for parton showering and hadronisation. The WZ+jets, tZ,
 832 and tW samples are produced using the `MadGraph5_aMC@NLO`(version 5.222) [118], which
 833 includes up to one hadronic jet at next to leading order (NLO) QCD accuracy. Other minor
 834 background (e.g. WW, ZZ, tWZ and tH) are simulated using different generators such as
 835 `MadGraph` [112], `MadSpin` [130] and `JHU` [126–129]. All events are interfaced to `Pythia` for
 836 parton shower and hadronisation.

NOTE: Add source

The complete list of SM samples is given in Table 3.4 , along with their cross sections. The cross sections without a reference are coming from the generator with which the sample has been made, for some of them the uncertainties are provided by the Generator Group . For each MC sample, the integrated luminosity that the sample represents is estimated as the number of simulated events divided by the cross section of the generated process. For processes generated with `MadGraph5_aMC@NLO`, the effective number of simulated events is used, taking into account positive and negative event weights. The correction factor for those events is defined as

$$C = \frac{\text{Nb. of pos. weights} + \text{Nb. of neg. weights}}{\text{Nb. of pos. weights} - \text{Nb. of neg. weights}} \times \text{mc baseweight} \quad (3.6)$$

Table 3.4: SM MC samples used in this analysis with their corresponding cross section and MadGraph5_aMC@NLO correction C when applicable. The generators used for each sample are indicated.

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

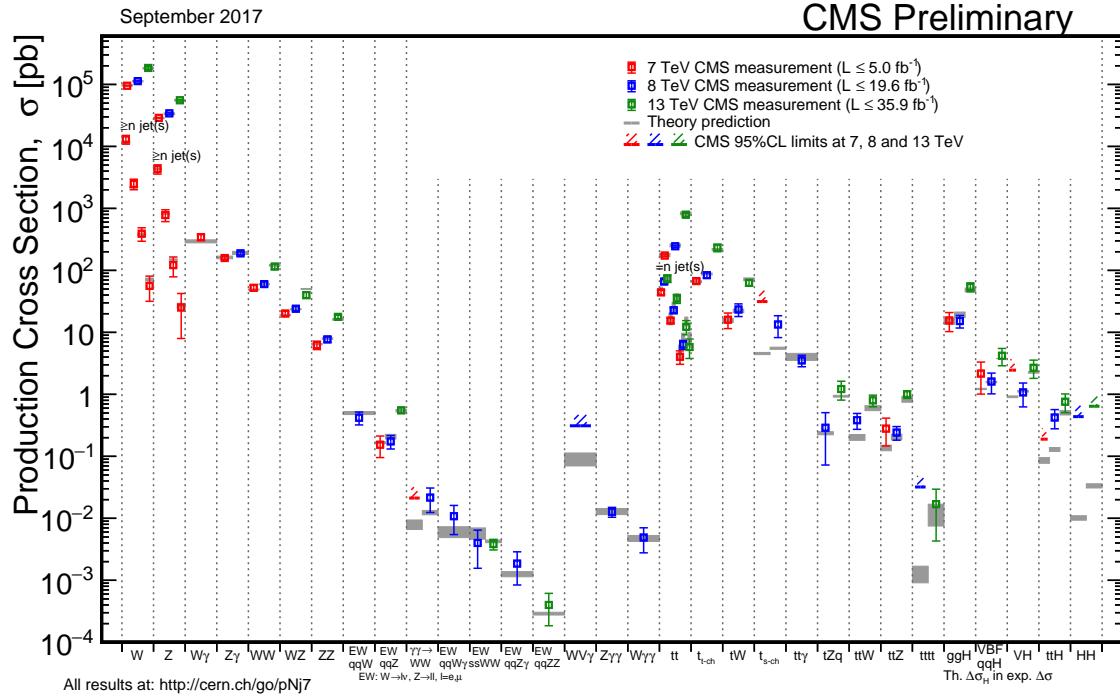


Figure 3.4: Summary of the SM cross section measurements performed by the CMS collaboration. Figure taken from [22]

837 3.3 Multivariate analysis techniques: Boosted Decision Trees

838 The need of processing large quantities of data and discriminating between events with largely
 839 similar experimental signatures makes multivariate statistical analysis (MVA) a largely used
 840 method in the physics community. Multivariate classification methods based on machine
 841 learning techniques are a fundamental ingredient to most analyses. The advantage of using
 842 a MVA classifier is that it can achieve a better discrimination power with respect to a simple
 843 cut and count analysis with poorly discriminating variables. These variables are referred to
 844 as weak variables and have similar distributions for signal and background samples. A risk of
 845 using MVA classifiers is overtraining. This happens when there are too many model parameters
 846 of an algorithm adjusted to too few data points. This leads to an increase in the classification
 847 performance over the objectively achievable one.

848 There are many software tools that exist for MVA. In this thesis the Tool for Multivariate
 849 Analysis (TMVA) [133] is used. This software is an open source project included into
 850 ROOT [134]. All multivariate techniques in TMVA belong to supervised learning algorithms. By
 851 training on events for which the outcome is known, a mapping function is determined that
 852 describes a classification or an approximation of the underlying behaviour defining the target
 853 value (regression).

854 In this thesis boosted decision trees (BDT) are employed for the classification of events as
 855 implemented in the TMVA framework [133]. This multivariate techniques is based on a set of

856 decision trees where each yields a binary output depending on the fact that an event is signal- or
 857 background-like. The advantage of such a multivariate technique is that several discriminating
 858 variables can be combined into a powerful one-dimensional discriminant D.

In Figure 3.5 a schematic view of de decision tree is shown. The starting point is the root node. Then a consecutive set of a total of i questions (nodes) regarding discriminating variables x_i are asked with only two possible answers per question (binary splits). The decision tree is constructed by training on a dataset for which the outcome is already provided, such as simulation dataset with signal and background processes (supervised learning). For each node a criterion $x_i > C_i$ is found by maximizing the separation gain between nodes

$$\text{separationgain} \approx \text{gain}(\text{parent}) - \text{gain}(\text{daughter, Signal}) - \text{gain}(\text{daughter, Background}), \quad (3.7)$$

with the gain computed using the Gini index

$$\text{gain}(\text{cell}) \approx p(1-p), \quad (3.8)$$

859 where p denotes the purity of a selection $x > C$. This is repeated until the maximum of nodes is
 860 reached and at the end of the sequence, the leaf nodes are labelled either signal S or background
 B, depending on the majority of events that end up on those nodes.

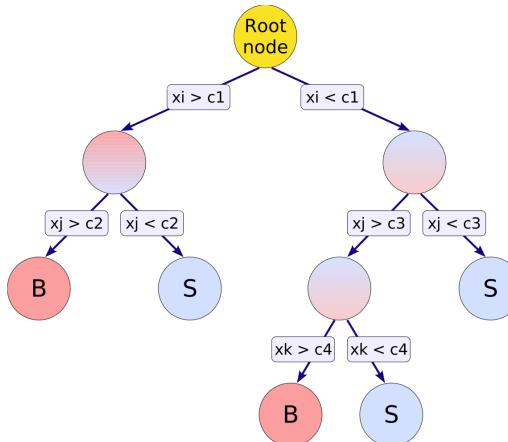


Figure 3.5: Schematic view of a decision tree. Figure taken from [133].

861

Different trees can be combined into a forest where the final output is determined by the majority vote of all trees, forming the sum of so-called weak learners into one strong learner. From one training collection, trees are derived by reweighting events, and combined into a single classifier as the weighted average of each individual decision tree. A method for making such forests is boosting a tree. In this method, misclassified events are weighted higher so that future learner concentrate on these events. This has as advantage that the response of the decision trees are stabilised against fluctuations in the training sample which enhances the performance. Additionally, the trees can be kept very shallow, in this thesis the maximal number of nodes is set the three, which improves the robustness against overtraining. Examples of such boosting algorithms are Adaptive Boosting (AdaBoost) and Gradient Boosting [135]. In

AdaBoost, each weight of the misclassified events are enhanced while reducing the weight of correctly classified events after each training such that future events learn those better

$$\alpha_{n+1} = \left(\frac{1 - \epsilon_n}{\epsilon_n} \right)^\beta, \quad (3.9)$$

where ϵ_n denotes the misclassification error of the current tree n and β is a learning rate. The weight w_i at node i is then equal to $w_i = \ln \alpha_i$. The final weight is the sum of all classifiers weighted by their errors. The learning rate is typically chosen to be $\beta \leq 0.5$ to allow more boosting steps. Gradient boosting has a similar approach and combines a gradient descent with boosting. Instead of fitting the base-learner to the reweighted data as in AdaBoost, it is fitted to the negative gradient vector of the loss function evaluated at the previous node. Misclassified events will result in a majority vote with large gradients of the loss function. Also for the Gradient boost, the learning rate is typically slow, this also known as shrinkage. In this thesis Gradient boost is used with a shrinkage of 0.2-0.3.

Additionally, the Gradient boost is used in combination with bagging, so-called stochastic gradient boosting. Bagging is a resampling technique draws a subset of events from the training data where the same event is allowed to be randomly picked several times from the parent sample. The tree is then trained on this subset and this is repeated many times. It is based on the assumption that sampling from a dataset that follows a distribution is the same as sampling from the distribution itself [136]. If one draws an event out of the parent sample, it is more likely to draw an event out of the phase space that has a high probability density, as the original dataset will have more events in the regions. Since the selected event is kept in the original sample, the parent sample stays unchanged so that randomly extracted samples have the same parent distribution, albeit statistically fluctuated. Bagging smears over the statistical fluctuations in the training data, making it suitable for stabilising the response of the classifier and increasing the performance by eliminating overtraining. In stochastic gradient boosting the bagging resampling procedure uses random sub-samples of the training events for growing the trees.

The discriminating power of a BDT is assessed by analysing the receiver operating characteristic (ROC) curve. This curves show the background rejection over the signal efficiency of the remaining sample. By looking at the area under the curve with respect to random guessing (AUC), the best classifier can be identified. This follows the Neyman-Pearson lemma that the best ROC curve is given by the likelihood ratio $\mathcal{L}(x|Signal)/\mathcal{L}(x|Background)$ [136]. No discrimination power will result in an AUC of 0%, while 50% means fully separated event classes. In Figure 3.6 an example of ROC curve is shown.

3.4 Statistical methodology

The search performed in the framework of this thesis requires the simultaneous analysis of data from different decay channels. The statistical methodology used for this search is developed by the ATLAS and CMS collaborations in the context of the LHC Higgs Combination group. The description of the methodology can be found in Refs. [138–141].

The Higgs Combined Tool [142] is a RooStats [143] framework which runs different statistical

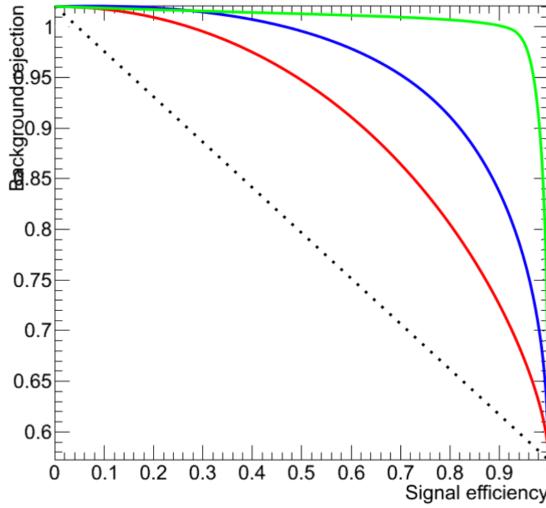


Figure 3.6: Example of ROC curves. In this example, the green method is better than the red one, which is better than the blue one. The dashed line represents a case where there is no separation. Figure taken from [137].

898 methods. In this section, only the statistical tools necessary for the performed search are
899 described. The results presented in this thesis are obtained using the asymptotic formulae [144].

900 In general the event yields of signal and background processes are denoted as s and b
901 respectively. These represent event counts in multiple bins or unbinned probability density
902 functions. By use of simulation, predictions on both signal and background yields are made.
903 These predictions are subject to multiple uncertainties that are accounted for by introducing
904 nuisance parameters θ such that $s = s(\theta)$ and $b = b(\theta)$. In the following, the actual observed
905 events are denoted as data or observation.

906 3.4.1 The absence of signal: limits

907 The absence of a signal is characterised in high energy physics by the Bayesian and modified
908 classical frequentist statistical approaches. They allow to quantify the level of incompatibility of
909 data with a signal hypothesis in terms of confidence levels (CL). The convention is to require a
910 95% CL for excluding a signal.

911 An analysis targeting a certain signal production mechanism can either set approximate
912 model-independent limits on signal cross sections times branching ratio ($\sigma \times \mathcal{B}$) or on the
913 signal cross section times branching ratio times detector acceptance ($\sigma \times \mathcal{B} \times \mathcal{A}$). In order to
914 test various theories, the latter is not useful unless the acceptance \mathcal{A} is provided. However, many
915 analysis are not able to present result in a form of limits on $\sigma \times \mathcal{B} (\times \mathcal{A})$, therefore an alternative
916 is adopted to set limits in the signal strength modifier μ . The signal strength modifier is defined
917 to equally change all the cross sections of all production mechanisms of the signal by the same
918 scale.

In this thesis, the modified frequentist approach for confidence levels is used [145, 146]. The classical frequentist uses a test statistic q_μ based on the profile likelihood ratio to determine

how signal- or background-like the data is. However, it does not allow nuisance parameters and is modified to incorporate these. First a likelihood $\mathcal{L}(\text{data}|\mu, \theta)$ is constructed as

$$\mathcal{L}(\text{data}|\mu, \theta) = \text{Poisson}(\text{data}|\mu s(\theta) + b(\theta)) p(\tilde{\theta}|\theta). \quad (3.10)$$

The probability density function (pdf) $p(\tilde{\theta}|\theta)$ describes all sources of uncertainty and is described in [Section 3.4.2](#). The data in [Equation 3.10](#) 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

$$\text{Poisson}(\text{data}|\mu s(\theta) + b(\theta)) = \prod_i \frac{(\mu s_i(\theta) + b_i(\theta))^{n_i}}{n_i!} e^{-\mu s_i(\theta) - b_i(\theta)}. \quad (3.11)$$

At the LHC, the test statistic is defined as

$$q_\mu = -2 \ln \frac{\mathcal{L}(\text{data}|\mu, \hat{\theta}_\mu)}{\mathcal{L}(\text{data}|\hat{\mu}, \hat{\theta}_\mu)}, \quad (3.12)$$

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 to guarantee a one sided confidence interval. This has as consequence that upward fluctuations of the data ($\hat{\mu} > \mu$) are not considered against the signal hypothesis of data with a signal with strength μ .

The criterion for excluding the signal at $1 - \alpha$ confidence level is the ratio of the probabilities 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 ($s + b$) and background only (b) hypothesis is defined as

$$\text{CL} = \frac{P(q_\mu \geq q_\mu^{\text{obs}} | \mu s + b)}{P(q_\mu \geq q_\mu^{\text{obs}} | b)} \leq \alpha. \quad (3.13)$$

These probabilities are defined as

$$p_\mu = P(q_\mu \geq q_\mu^{\text{obs}} | \mu s + b) = \int_{q_\mu^{\text{obs}}}^{\infty} f(q_\mu | \mu, \hat{\theta}_\mu^{\text{obs}}) dq_\mu, \\ 1 - p_b = P(q_\mu \geq q_\mu^{\text{obs}} | b) = \int_{q_{\mu=0}^{\text{obs}}}^{\infty} f(q_\mu | \mu = 0, \hat{\theta}_{\mu=0}^{\text{obs}}) dq_\mu, \quad (3.14)$$

where p_μ and p_b are called 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 pdfs of the signal plus background and background only hypothesis

928 constructed from toy Monte Carlo pseudo data. These pdfs are shown in [Figure 3.7](#) and are
 929 generated with nuisance parameters fixed to $\hat{\theta}_{\mu=0}^{\text{obs}}$ and $\hat{\theta}_{\mu}^{\text{obs}}$. These values of the nuisance
 930 parameters for the background only $\hat{\theta}_{\mu=0}^{\text{obs}}$ and signal plus background $\hat{\theta}_{\mu}^{\text{obs}}$ hypothesis that best
 931 describe the data are found by maximising the likelihood from [Equation 3.10](#). The 95% CL
 level upper limit on μ is achieved by adjusting μ until CL = 0.05

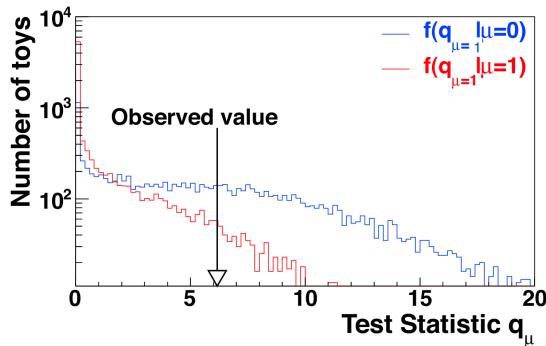


Figure 3.7: Test statistic distributions for pseudo data generated for the signal plus background ($\mu = 1$) and background only ($\mu = 0$) hypothesis. Figure taken from [\[141\]](#).

932
 933 The expected median upper limit and the $\pm 1\sigma$ and $\pm 2\sigma$ bands for a hypothesis is generated
 934 by a large set of pseudo data and calculate the CLs and the value of μ at 95% CL for each of
 935 them. A cumulative probability distribution can be build by starting the integration from the
 936 side corresponding to low event yields. The median expected value is where the cumulative
 937 distribution function crosses the 50% quantile. The $\pm 1\sigma$ (68%) and $\pm 2\sigma$ (95%) bands are
 938 defined by the crossings of the 16% and 84%, and 2.5% and 97.5% quantiles.

939 3.4.2 Adding sources of uncertainty

940 In this thesis, all sources of uncertainties are assumed to be either 100% correlated or uncor-
 941 related. Partially correlated uncertainties are broken down to subcomponents that fit those
 942 requirements, allowing to include all constraints in the likelihoods in a clean factorised form.

A systematic uncertainty pdf $p(\theta|\tilde{\theta})$ 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 $p(\theta|\tilde{\theta})$ are re-interpret as posteriors of real or imaginary measurements $\tilde{\theta}$

$$p(\theta|\tilde{\theta}) \sim p(\tilde{\theta}|\theta) \pi_\theta(\theta), \quad (3.15)$$

943 where $\pi_\theta(\theta)$ is the hyper prior for the (imaginary) measurements. For the pdfs used by the
 944 Higgs Combine Tool (normal, log normal, gamma distribution), hyper priors can remain flat.
 945 This allows to use the pdf $p(\tilde{\theta}|\theta)$ to constrain the likelihood of the main measurement in a
 946 frequentist calculation. Additionally this allows to build a sampling distribution of the test
 947 statistic [\[141\]](#).

The statistical uncertainties on the Monte Carlo prediction in each bin are obtained following the Barlow-Beeston-light approach [\[147\]](#). In this approach a single Gaussian constrained

nuisance parameter is assigned to scale the sum of the process yields in each bin, constrained by the total uncertainty. This method has as advantage that it minimises the number of parameters required in the maximum likelihood fit. Considering n_{tot} events in a bin with background process i in the bin

$$n_{\text{tot}} = \sum_{i \in \text{bkg}} n_i, \quad (3.16)$$

the total uncertainty e_{tot} is given by

$$e_{\text{tot}} = \sqrt{\sum_{i \in \text{bkg}} e_i^2}, \quad (3.17)$$

948 with e_i the uncertainty on background i and is given by the sum of squares of weights used to
 949 fill the bins. The Gaussian constrained parameter x has then a nominal value of zero and scales
 950 the yield as $n_{\text{tot}} + x e_{\text{tot}}$.

951 Choices of systematic uncertainty density functions

For uncertainties that are unconstrained by a priori measurements that do not involve the data going into the statistical analysis, flat priors are used. When there are a priori measurements available such as those from control regions, one can use either a Gaussian pdf, a log-normal pdf, or a gamma distribution. The Gaussian pdf is suited for describing uncertainties on parameters with both positive and negative values. This prior is however not suitable for positively defined observables such as cross sections, cut efficiencies, luminosity, etc. and is not used in this thesis. An alternative option is the log normal pdf which is used in the rest of this thesis

$$\rho(\theta) = \frac{1}{\sqrt{2\pi} \ln(\kappa)} \exp\left(-\frac{(\ln(\theta/\tilde{\theta}))^2}{2(\ln \kappa)^2}\right) \frac{1}{\theta}. \quad (3.18)$$

The parameter κ characterises the width of the log normal pdf. For example $\kappa = 1.10$ implies that the observable can be larger or smaller by a factor 1.10, both deviation having a chance of 16%. The gamma distribution is used for describing statistical uncertainties associated with a number of Monte Carlo events in simulation or a number of observed events in a data control sample. In this thesis, the gamma distribution is only used for the latter. The event rate in the signal region n is related to the number of events in the control region N as $n = \alpha N$. Ignoring the uncertainties on α , the predicted rate follows

$$\rho(n) = \frac{1}{\alpha} \frac{n/\alpha)^N}{N!} \exp(-n/\alpha). \quad (3.19)$$

952 The mapping between the posteriors $\rho(\theta|\tilde{\theta})$ and the auxiliary measurement pdfs $p(\tilde{\theta}|\theta)$ are
 953 given in [141].

954 3.4.3 Asymptotic approximation of the CL method

955 In order to significantly reduce computing time, the Asymptotic CL method is used. This method
 956 avoids an ensemble of toy Monte Carlo samples and instead replaces it by one representative
 957 dataset, called Asimov dataset. This dataset is constructed such that all observed quantities are
 958 set equal to their MLE values ($\hat{\theta}_{\text{Asimov}} = \theta_0$). More information about this procedure can be
 959 found in Refs. [139].

960 **3.4.4 Extracting the signal model parameters**

From a scan of the profile likelihood ratio,

$$q(a) = -2 \ln \frac{\mathcal{L}(\text{obs} | s(a) + b, \hat{\theta}_a)}{\mathcal{L}(\text{obs} | s(\hat{a}) + b, \hat{\theta})}, \quad (3.20)$$

the signal model parameters are evaluated. The likelihood is maximised by the parameters \hat{a} and $\hat{\theta}$. The likelihood

$$\mathcal{L}_{\max} = \mathcal{L}(\text{obs} | s(\hat{a}) + b, \hat{\theta}) \quad (3.21)$$

961 is called the best-fit set.

962 The 68% and 95% CL on a given parameter of interest a_i is then evaluated from $q(a_i) = 1$ or
 963 $q(a_i) = 3.84$ respectively, where all other unconstrained model parameters are treated in the
 964 same way as the nuisance parameters [140].

Event reconstruction and identification

4

966 After the detector simulation described in [Section 3.2](#), the simulated data has the exact same
 967 format as the real collision data recorded at the CMS experiment. Therefore the same software
 968 can be used for the reconstruction of both simulation and real data. In [Section 4.1](#), the object
 969 reconstruction is explained. After reconstructing the objects, they are connected ([Section 4.2](#)) to
 970 physics objects need to be identified ([Section 4.3](#)) and corrected for pile up ([Section 4.4](#)). The
 971 objects used for physics analysis have extra requirements as shown in [Section 4.5](#). A summary
 972 of all the corrections applied to data and simulation is given in [Section 4.6](#).

973 4.1 Object Reconstruction

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

985 The traditional hadron colliders reconstruction is as follows. The reconstruction of isolated
 986 photons and electrons is primarily done by the ECAL, while the identification of muons is based
 987 on the muon detectors. Hadrons and photons form jets which are measured by the calorimeters
 988 without any contribution from the tracker or muon detectors. Jets can be tagged using the
 989 tracker as coming from hadronic τ decays or b hadronisation based on the properties of the
 990 relevant charged particle tracks. The missing transverse energy \vec{p}_T is defined as the vectorial
 991 sum of the undetectable particle transverse momenta, and can be reconstructed without any
 992 information from the tracker. The particle flow (PF) [148] reconstruction correlates the tracks
 993 and clusters from all detector layers with the identification of each final state particle, and

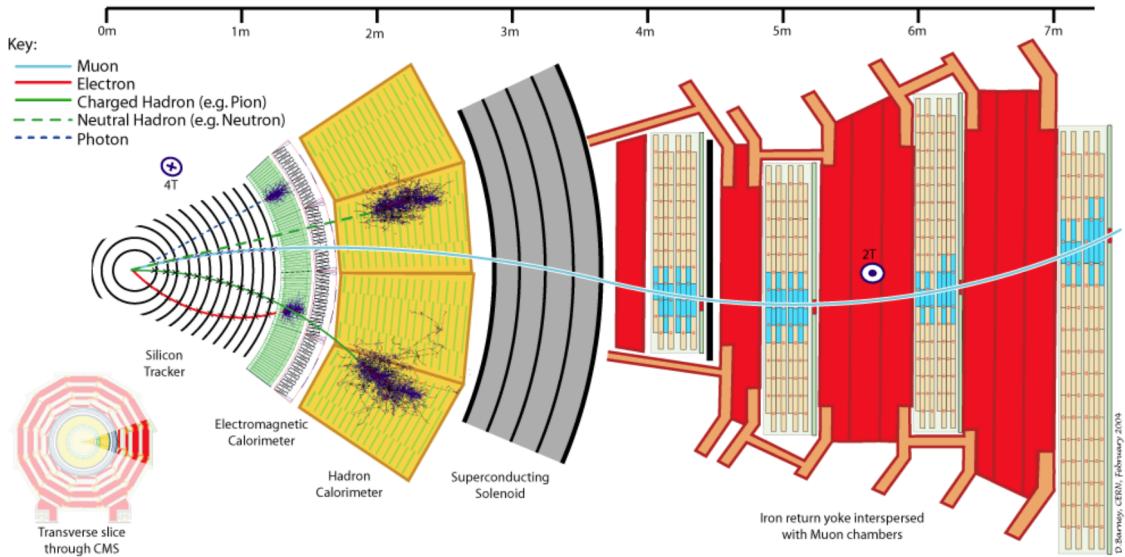


Figure 4.1: Cross-section of the CMS detector with all parts of the detector labelled. This sketch shows the specific particle interactions from a beam interaction reign to the muon detector. The muon and charged pion are positively charged, the electron is negatively charged. Figure taken from [148].

994 combines the corresponding measurements to reconstruct the properties. In the particle flow
 995 algorithm, the muon is identified by a track in the inner tracker connected to a track in the
 996 muon detector as described in Section 4.1.2. The electrons are identified by a track and an
 997 ECAL cluster, not connected to an HCAL cluster as described in Section 4.1.3. The ECAL and
 998 HCAL clusters without a track link identify the photons and neutral hadrons, while the addition
 999 of the tracker determines the energy and direction of a charged hadron.

1000 Coarse-grained detectors can cause signals of different particles to merge and reduce the
 1001 ability of identifying and reconstructing the particles. Therefore, particle flow identification
 1002 requires sufficiently segmented subdetectors such that a global event description is possible.
 1003 The CMS detector is built to meet to requirements of the particle flow reconstruction. It has
 1004 an efficient and pure muon identification system, a hermetic HCAL with coarse segmentation,
 1005 a higher segmented ECAL, a fine-grained tracker and a large magnetic field to separate the
 1006 calorimeter deposits of charged and neutral particles in jets.

1007 4.1.1 Charged particle tracks

1008 An iterative tracking algorithm is responsible for the reconstruction of the tracks made by
 1009 charged particles in the inner tracking system. Each iteration consists of four steps [69]: the
 1010 track-seed generation, the pattern recognition algorithm, removal of track-hit ambiguities and
 1011 a final track fit.

1012 The seed generation is the first step. It consists of finding reconstructed hits that are usable
 1013 for seeding the subsequent track-finding algorithm. They are identified from a group of at
 1014 least three reconstructed hits in the tracker, or from a pair of hits while requiring the origin
 1015 of the track segment to be compatible with the nominal beam-collision point. Since the pixel

1016 has a higher granularity compared to the strip tracker, its seed generation efficiency is higher.
 1017 The overall efficiency exceeds 99%. The second step of each iteration, the pattern recognition
 1018 algorithm, uses the seeds as a starting point for a Kalman filter method [149, 150]. This
 1019 algorithm extrapolates the seed trajectory towards the next tracker layer taking into account
 1020 the magnetic field and multiple scattering effects. The track parameters are updated when a
 1021 compatible hit in the next layer is found. This procedure continues until the outermost layer is
 1022 reached. Since the Kalman filter method can result in multiple tracks associated to the same
 1023 seed, or different tracks sharing the same hits, a removal of ambiguities is necessary. This
 1024 ambiguity resolving is done by removing tracks that are sharing too many hits from the list
 1025 of track candidates. The tracks with the highest number of hits or with the lowest χ^2 in the
 1026 track fit is kept. The updated track parameters are then refitted using the Kalman filter method,
 1027 where all hits found in the pattern recognition step are taken into account. The fit is done twice
 1028 - once outwards from the beam line towards the calorimeters, and inwards from the outermost
 1029 track hit to the beam line -, improving the estimation of the track parameters.

1030 All hits that are unambiguously associated to the final track are removed from the list of
 1031 available hits. In order to associate the remaining hits, the procedure is repeated with looser
 1032 track reconstruction criteria. The use of the iterative track reconstruction procedure has a
 1033 high track finding efficiency, where the fake track reconstruction rate is negligible. For muons,
 1034 this results in a global track reconstruction efficiency exceeding 98%, and 75-98% for charged
 1035 hadrons.

1036 4.1.2 Following the Muon's Footsteps

1037 The muon reconstruction [151] has three subdivisions: local reconstruction, regional reconstruc-
 1038 tion and global reconstruction. The local reconstruction is performed on individual detector
 1039 elements such as strip and pixel hits in the inner tracking system, and muon hits and/or seg-
 1040 ments in the muon chambers. Independent tracks are reconstructed in the inner tracker - called
 1041 tracker tracks - and in the muon system, called standalone muon tracks. Based on these tracks,
 1042 two reconstructions are considered.

1043 The outside-in approach is referred to as Global Muon reconstruction. For each standalone
 1044 muon track, a inner tracker track is found by comparing the parameters of the two tracks prop-
 1045 agated onto a common surface. Combining the hits from the tracker track and the standalone
 1046 track, gives a fit via the Kalman filter technique [149, 150] for a global muon track.

1047 The second approach is an inside-out reconstruction, creating tracker muons. All candidate
 1048 tracker tracks with a $p_T > 0.5$ GeV and total momentum $p > 2.5$ GeV are extrapolated to the
 1049 muon system taking into account the magnetic field, the average expected energy losses, and
 1050 multiple Coulomb scattering in the detector material. The extrapolated track and the muon
 1051 segments are considered matched when the difference in the position in the x coordinates is
 1052 smaller than 3 cm, or when the ratio of this distance to its uncertainty is smaller than four. When
 1053 at least one muon segment - DT or CSC hits - matches the extrapolated track, the corresponding
 1054 tracker track is indicated as a tracker muon.

1055 For low transverse momenta ($p_T \lesssim 5$ GeV), the tracker muon reconstruction is more efficient
 1056 than the global muon approach. This is due to the fact that tracker muons only require a

single muon segment in muon system, while the global muon approach requires typically segments in at least two muon stations. These tracker muons are used for identifying muons from the hadronisation of b or c quarks. The global muon approach typically improves the tracker reconstruction for $p_T \gtrsim 200$ GeV. These are labelled isolated when in a cone of $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} = 0.3$ around the muon, the sum of the transverse momenta of additional tracker tracks and energy deposits in the calorimeter is less than 10% of the muon's transverse momentum.

4.1.3 The path of the Electron

The electrons in CMS radiate more than 70% of their energy in the inner track through bremsstrahlung before reaching the ECAL. This has as consequence that the electron tracks are increasingly curved in the magnetic field as a function of its flight distance. Standard tracking algorithms are based on Kalman filtering which assume that the energy loss is Gaussian distributed, and are therefore not suitable to fit the electron tracks. A different filtering algorithm, the Gaussian sum filter (GSF) [152] is used in the electron track reconstruction instead.

In CMS, the electrons are reconstructed in two ways. The older ECAL based tracking is developed to identify high energetic isolated electrons. This tracking algorithm starts from ECAL clusters with a transverse energy above 4 GeV and extrapolates from these cluster the position of the hits in the tracker. In order to account for bremsstrahlung, neighbouring clusters in η and ϕ are grouped together into a supercluster from which then the direction is determined to find the position of the particles in the tracker. This has as consequence that for electrons or positrons in jets, energy deposits of surrounding particles will be entering the supercluster leading to a wrong position of the electron/positron in the tracker. Another disadvantage of the ECAL based tracking is that for low p_T electrons, the trajectories will be very curved and the supercluster will not contain all of the energy deposit, leading to a higher misconstruction rate.

The faults of the ECAL based tracking are lifted by adding a tracker based algorithm. This algorithm uses all the tracks with a p_T higher than 2 GeV found with iterative tracking as seeds. Iterative tracking uses the Kalman Filter algorithm several times with an average track reconstruction efficiency but high purity. In contrary with a global combinatorial fit, the iterative tracking accepts tracks with a small transverse momentum that are not leaving any energy in the ECAL, and tracks from particles that only interact with the inner tracker layers. When the electron or positron radiated a small amount of energy, the corresponding track can be reconstructed across the whole tracker and safely propagated to the ECAL surface. When there is a larger amount of energy radiated however, the pattern recognition might fail to accommodate for the change in the electron momentum leading to a track reconstructed with a small number of hits. The solution for this is a preselection based on the χ^2 and number of hits and the selected tracks are fitted again with Gaussian-Sum-Filter which can accommodate substantial energy losses across the trajectory.

The electron seeds from the ECAL- and tracker-based procedures are merged into a unique collection and are then refitted by using the summed Gaussian distributions as uncertainty per hit in the track fit.

1097 The electron efficiency is measured in 8 TeV proton collision data to be better than 93%
 1098 for electrons with an ECAL supercluster energy of $E_T > 20$ GeV [153]. For electrons with an
 1099 $E_T > 25$ GeV in 13 TeV proton collision data, the efficiency is about 96%[154].

1100 4.1.4 Primary Vertex Reconstruction

1101 The primary vertex (PV) reconstruction should be able to measure the location of all proton
 1102 interaction vertices in each event: the signal vertex an all vertices from pile up events. It
 1103 consists of a vertex finding and a vertex fitting algorithm and happens in three steps. Tracks are
 1104 selected to be consistent with being produced promptly in the primary interaction by imposing
 1105 requirements on the track parameters [79]. By grouping reconstructed tracks according to the
 1106 z coordinate of their closest approach to the beam line, vertices for all interaction in the same
 1107 beam crossing are found, at CMS this is done by a deterministic annealing algorithm [155] based
 1108 on a statistical mechanics model. On top of this, a vertex fitting algorithm like the Adaptive
 1109 Vertex fitter [156], is performed. This creates the three-dimensional primary-vertex position.
 1110 With this fit, the contribution from long-lived hadron decays is reduced by down weighting the
 1111 tracks with a larger distance to the vertex. The primary vertex corresponding to the highest
 1112 sum of squared track transverse momenta is noted as the point of the main interaction. The
 1113 resolution on the primary vertex is about 14 μm in $r\phi$ and about 19 μm in the z direction for
 1114 primary vertices with the sum of the track $p_T > 100$ GeV for 2016 data taking.

1115 4.1.5 Calorimeter clusters

1116 The cluster algorithm in the calorimeter detects and measures the energy and direction of stable
 1117 neutral particles such as photons and neutral hadron. Additionally, it used to separate neutral
 1118 particles from charged hadron energy deposits, and reconstructs and identifies electrons and
 1119 their bremsstrahlung photons. Furthermore, the cluster algorithm contributes to the energy
 1120 measurements of charged hadrons that don't have accurate tracks parameters, e.g. for low
 1121 quality and high transverse momentum tracks. The clustering is performed separately in each
 1122 subdetector: ECAL barrel and endcaps, HCAL barrel and end caps, and the two preshower
 1123 layers. The HF has no clustering algorithm since the electromagnetic or hadronic components
 1124 give rise to an HF EM or HF HAD cluster.

1125 The clustering algorithm consist of different steps. First seeds are identified when cells have
 1126 an energy larger than the seeding threshold and larger than their neighbouring cells. Then
 1127 topological clusters are made by accumulating cells that share at least a corner with a cell
 1128 already in the cluster and an energy above a cell threshold set to twice the noise level. The third
 1129 step is a expectation maximization algorithm that reconstructs the cluster [148]. This algorithm
 1130 assumes that the energy deposits are Gaussian distributed and is an iterative algorithm with
 1131 two steps at each iteration. The first step calculates the expected fraction of the energy, while
 1132 the second step performs a maximum likelihood fit. The positions and energies of the Gaussian
 1133 functions are then taken as cluster parameters.

1134 The calorimeter clusters are used for reconstructing photons and neutral hadrons. The clusters
 1135 that are not in the vicinity of the extrapolated charged tracks are easily identified as neutral
 1136 hadrons or photons. For the energy deposits that overlap with charged hadrons however, the
 1137 neutral particle energy deposit can only be detected as an excess over the charged particle

deposit. For this reason, a good calibration of the electromagnetic and hadronic calorimeter is vital.

The ECAL calibration is performed before the hadron cluster calibration or particle identification. For run 1, the ECAL response to electrons and photons as well as the cell-to-cell relative calibration is determined with test beam data, radio active sources, and cosmic ray measurements. For run 2, the collision data collected at 7 and 8 TeV was used to refine the calibration. The effect of the thresholds in the clustering algorithm are estimated from simulated single photons with energies varying from 0.25 to 100 GeV. The photons used for the calibration should not have a conversion prior to their entrance to ensure the calibration of single clusters. In all ECAL regions and for all energies, the calibrated photon energies agree with the true photon energies within 1%.

In contrary to the photons, the hadrons deposit in general energy in both ECAL and HCAL. Since the calorimeter response in the HCAL depends on the fraction of shower energy deposited in the ECAL, the ECAL and HCAL cluster energies are recalibrated together to get an estimate of the true hadron energy. Since the calibration is done for hadrons, single neutral hadrons such as K_L^0 are used for determining the calibration constants. The hadrons interaction with the tracker material are rejected for the calibration purposes. This calibration is checked with isolated charged hadron selected from early data recorded at $\sqrt{s} = 0.9, 2.2$ and 7 TeV.

4.2 Putting the pieces together

A link algorithm connects the several PF elements from the various CMS subdetectors. It tests any pair of elements in an event and is restricted to considering nearest neighbours in the $\eta\phi$ -plane. The quality of the link is determined via the distance between the two elements and PF blocks of elements are formed from elements with a direct link or indirect link through common elements.

The link between a central tracker track and a calorimeter clusters is made by extrapolating the tracker track to the two layers of the preshower, the ECAL, and the HCAL. If this extrapolated position is within the cluster area, the two are linked. When there are several ECAL or HCAL clusters for the same track, the link with the smallest distance is kept. A dedicated cluster algorithm accounts for the energy of the photons emitted through bremsstrahlung at for photons that have converted to an electron-positron pair.

The ECAL to HCAL cluster and ECAL to preshower cluster links are established when the cluster position in the more granular calorimeter, ECAL or preshower, is in accordance with the cluster envelope of the less granular calorimeter, HCAL or ECAL. When there are multiple HCAL clusters linked to the same ECAL cluster, the link with the smallest distance is kept. This is also true for multiple ECAL clusters with the same preshower clusters. The ECAL supercluster is linked with the ECAL cluster when they share at least one ECAL cell.

Nuclear interactions in the tracker can lead to kinks in hadron trajectories as well as the production of secondary particles. This leads to charged particle tracks linked together via a common displaced vertex. The displaced vertices considered should have at least three tracks, with at most one incoming track, and the invariant mass of the outgoing tracks should exceed 0.2 GeV. The link between a track and the muon detectors is done via local, regional, and global reconstruction as explained in [Section 4.1.2](#).

1180 4.3 Particle flow identification

1181 The identification and reconstruction follows a particular order in each PF block. After each
 1182 identification and reconstruction the corresponding PF elements (tracks and clusters) are
 1183 removed from the PF block. The muons are the first to be identified and reconstructed. These
 1184 are reconstructed if their momenta are compatible with corresponding track only momenta. Then
 1185 the electron and its corresponding brehmstrahung photons, are identified and reconstructed by
 1186 using of the GSF tracking. At the same time, the energetic and isolated photons are identified as
 1187 well. The remaining elements in the PF block are subjected to a cross identification of charged
 1188 hadrons, neutral hadrons, and photons that arise from parton fragmentation, hadronisation,
 1189 and decays in jets. The charged hadron candidate is made from the remaining candidates that
 1190 have a charged particle track associated with them. Then the charged particle energy fraction
 1191 is subtracted from the calibrated energy of the linked calorimeter clusters and the remaining
 1192 energy is assigned to the neutral energy. Depending on the excess of neutral energy in the ECAL
 1193 and HCAL clusters, a photon or a neutral hadron is assigned respectively. The pseudorapidity
 1194 range of the inner tracker limits the information on the particles charge to $|\eta| < 2.4$. Outside
 1195 this range a simplified identification is done for hadronic and electromagnetic candidates only.

1196 4.3.1 Muons

1197 A set of selection requirements based on the global and tracker muon properties is responsible
 1198 for muon identification. The muons are considered isolated when the additional inner tracks
 1199 and calorimeter energy deposits within a distance to the muon direction in the $\eta\phi$ -plane is
 1200 smaller than 0.3. The muons coming from charged hadron decays or heavy flavour decays need
 1201 more stringent criteria. This due to the fact that charged hadrons can be misidentified as muons
 1202 because of e.g. punch-through, or muons can be seen as charged hadrons, and will absorb the
 1203 energy deposits of nearby particles.

1204 4.3.2 Electrons and isolated photons

1205 The electrons and photons are reconstructed together as discussed before. An electron candidate
 1206 seeded from a GSF track is considered an electron when the linked ECAL cluster is not linked
 1207 to three or more additional tracks. The photon seeds are ECAL superclusters with transverse
 1208 energies above 10 GeV that have no links with a GSF track. After associating photons from
 1209 brehmstrahung with the associated electrons, the remaining energy is associated to the photons
 1210 and the photon direction is taken to be that of the supercluster. The electron direction is
 1211 chosen to be that of the GSF track and its energy is a combination of the ECAL energy with
 1212 the momentum of the GSF track. Photons are retained if they are isolated, while electrons
 1213 should satisfy additional criteria based on a multivariate analysis for isolated and non-isolated
 1214 electrons.

1215 4.3.3 Hadrons and non-isolated photons

1216 After muon, electron and isolated photon identification, the remaining particles are hadrons
 1217 from jet fragmentation and hadronisation. These can show up as charged hadrons (e.g. π^\pm ,
 1218 K^\pm , or protons), neutral hadrons (e.g. K_L^0 or neutrons), non isolated photons (e.g. from π^0
 1219 decays), and additional muons from early decays of charged hadrons.

1220 The photons and neutral hadrons are assigned to calorimeter clusters without any link to
 1221 tracks. When the calorimeter clusters between the ECAL and HCAL are linked, the clusters are
 1222 assumed to arise from the same hadron shower. If their is not such a link, HCAL clusters are
 1223 assigned to neutral hadrons, while the ECAL clusters are assigned to photons based on the fact
 1224 that neutral hadrons leave only 3% of their energy in the ECAL. The HCAL clusters linked with
 1225 tracks, that are not linked with other HCAL clusters, are assigned to charged hadrons. These
 1226 tracks are then linked with remaining ECAL clusters.

1227 Hadron interactions can result in the creation of extra particles originating from a secondary
 1228 vertex. These extra particles are identified by having a common secondary vertex and replaced
 1229 in the PF list as one single primary charged hadron.

1230 4.3.4 Post processing

1231 After identification and reconstruction of all particles as described above. An artificial large
 1232 missing transverse momentum \vec{p}_T can be reconstructed. The cause of the \vec{p}_T is mostly
 1233 misidentified or misreconstructed high- p_T muons originating from cosmic rays, misconstruction
 1234 of the muon's momentum, or punch-through charged hadrons. A post processing step is applied
 1235 to solve this \vec{p}_T . Events with genuine large \vec{p}_T due to the presence of neutrino's are unaffected
 1236 by this post processing.

1237 4.4 Pile up mitigation and luminosity measurement

1238 The particle flow algorithm is design without taking pile up into account. For the 8 TeV dataset,
 1239 an average of about 21 pile up interactions happen per bunch cross section. For the dataset
 1240 taken at 13 TeV, the number of pile up interactions increases to about 27 interactions per
 1241 bunch crossing. These interactions are spread around the beam axis around the centre of the
 1242 CMS coordinate system and follow a normal distribution with a standard deviation of about
 1243 5 cm [148]. The number of pile up interactions is estimated from the number of interaction
 1244 vertices reconstructed from charged particle tracks, or from the instantaneous luminosity of the
 1245 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 [157]. The instantaneous luminosity from recorded process rate R is then determined as

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

1246 where $\sigma_{fid} = \sigma \times A$ corresponds to the fiducial cross section recorded in the luminometer
 1247 acceptance A which is determined using van der Meer scans [158]. The overall uncertainty on
 1248 the luminosity measurement is estimated to be 2.5%.

1249 The luminosity is used to infer to number of pile up interactions in data, which can be used
 1250 to corrected the predefined pile up interactions in simulation. The average number of pile up
 1251 interactions in data is estimated per luminosity section by multiplying the luminosity times

the total inelastic cross section. Then an event weight can be derived from the ratio of the distributions of pile up interactions in data and simulation. For 13 TeV collisions, the inelastic cross section is measured to be 71.3 ± 3.5 mb [158]. However a better agreement in data and simulation for the pile up sensitive variables, such as the number of primary vertices, is found with a lower cross section of 69 mb.

The pile up vertices are separated from the primary vertex by requiring that the primary vertex is the vertex with the highest quadratic sum of the transverse momenta of the corresponding tracks. The charged hadrons coming from these pile up vertices are identified via their tracks and are removed from the list of reconstructed particles to be used for physics analysis. This method is the so-called pile up charged hadron subtraction and denoted as CHS [159]. For the reconstructed particles outside the tracker acceptance as well as photons and neutral hadrons, the CHS method doesn't work. Therefore, the transverse density from pile up interactions is estimated using jet clustering techniques and their effect is subtracted from the particles transverse momenta. Additionally, the pile up contribution can be estimated locally as described for the muons and electrons described in [Section 4.5.1](#) and [Section 4.5.2](#).

4.5 Physics object reconstruction and identification

The particle flow objects are used for building physics objects that are used for analysis. These objects are jets, muons, electrons, photons, taus and missing transverse momentum \vec{p}_T . They are used to compute other quantities such as particle isolation and have extra requirements that are analysis dependent. In the following section, only the physics objects used throughout this thesis are discussed.

4.5.1 Muons

The muon candidates used for analysis in this thesis correspond to the tight working point (WP). The tight working point yields the most genuine muons and rejects falsely reconstructed ones. Detailed reports on the performance can be found in [160].

In order to reject objects wrongly reconstructed as muons from hadron showers that reach the muon system (punch-throughs), the global muon fit is required to include at least one valid hit in the muon chambers and for which at least two muon segments in two muon stations is present. Additionally, the muon tracks should have a global fit yielding a goodness-of-fit of $\chi^2/\text{ndof} < 10$. The decay of muons in flight is suppressed by requiring at least one pixel hit in the muon track. Furthermore, a minimum of five hits in the tracker is required. Cosmic muons and muons originating from pile up 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 [Figure 4.2](#), the muon efficiencies for data and simulation is presented. These efficiencies are estimated from tag-and-probe methods that select $Z \rightarrow \mu^- \mu^+$ and tag one muon that passes the identification criteria. The other muon is used as probe and one measures how many times it passes the identification criteria to get the efficiency. Overall, the efficiency is about 95-100%, with two drops due to the crack between the wheels of the DT system. The differences between

1292 data and simulation are corrected by applying p_T - and η -dependent scale factors ($\epsilon_{\text{data}}/\epsilon_{\text{MC}}$)
 1293 to simulated events. In Table 4.1, the muon requirements for the muons used throughout this
 thesis are summarised.

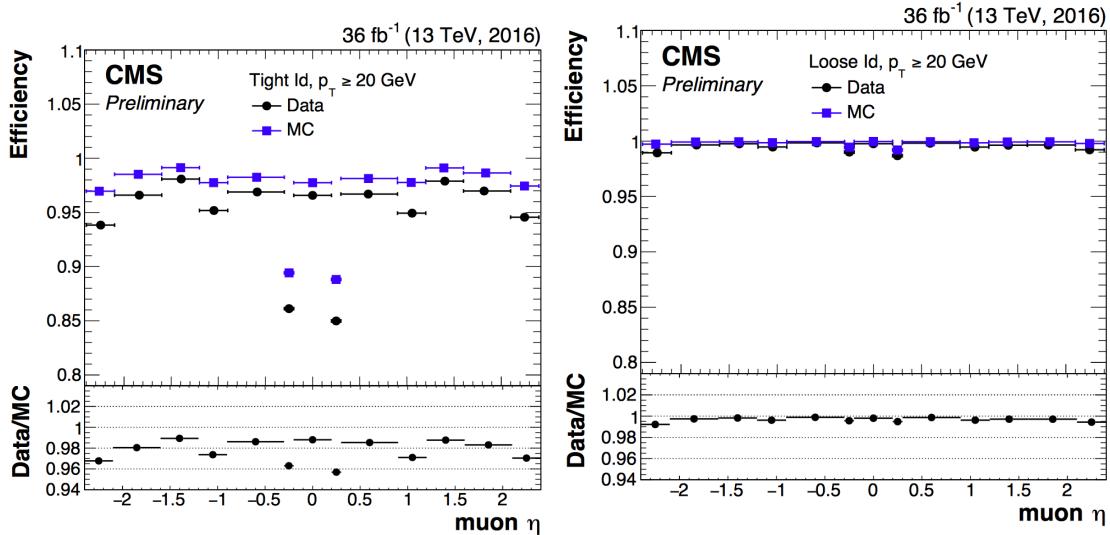


Figure 4.2: Comparison of the muon tight ID (left) and loose ID (right) efficiencies in data and simulation as a function of the pseudorapidity of the muon using the full 2016 dataset. Figure taken from [160].

1294

In addition to the identification criteria, the muons are required to be spatially isolated from EM and hadronic activity. The 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 pile up interactions. The relative isolation for muons \mathcal{I}_μ is defined as [148]:

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

1295 where a cone of $\Delta R = 0.4$ is adopted and the pile up mitigation is based on the $\Delta\beta$ correction.
 1296 The $\Delta\beta$ correction estimates the pile up energy as half of the contribution coming from charged
 1297 hadrons. For tight ID muons, this relative isolation should $\mathcal{I}_\mu < 0.15$, while for loose muons
 1298 this should be $\mathcal{I}_\mu < 0.25$. In Figure 4.3, the isolation efficiencies as a function of the pseudo
 1299 rapidities using the tag and probe method are shown for the tight muon ID. The efficiencies are
 1300 85-100% with a decline for low- p_T muons since they are most likely coming from hadronic
 1301 or heavy flavour decays. The differences between data and simulation are accounted for by
 1302 applying η - and p_T -dependent scale factors on the simulation.

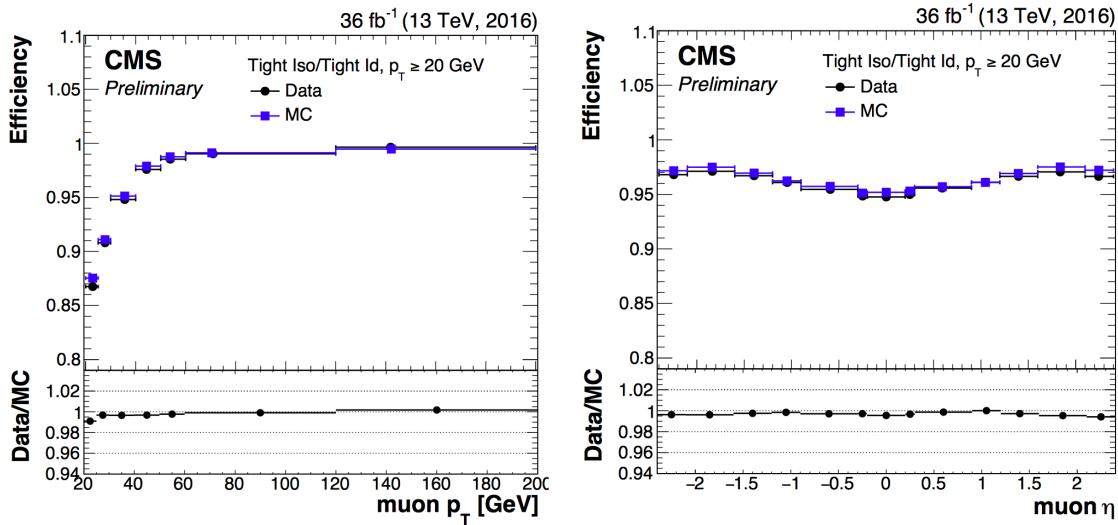


Figure 4.3: Comparison of the muon tight isolation requirement with the muon tight ID efficiencies in data and simulation as a function of the transverse momentum (left) or pseudorapidity (right) of the muon using the full 2016 dataset. Figure taken from [160].

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

Property	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	$d_{xy} < 2$ mm
Longitudinal distance wrt. the PV	N/A	$d_z < 5$ mm
Nb. of pixel hits	N/A	> 0
Nb. of tracker layers with hits	N/A	> 5
Relative Isolation	<0.25	<0.15

1303 **4.5.2 Electrons**

1304 The electrons candidates used correspond to the tight and veto working points. The study of
 1305 the electron reconstruction and identification performance can be found in [154].

1306 Starting from an electron PF candidate with a GSF track that is outside the barrel-endcap
 1307 transition region ($1.4443 < |\eta| < 1.5660$), several requirements are set. The electrons from
 1308 photon conversions are dismissed by requiring the electron track to have not have more than one
 1309 (two or three) missing hit in the innermost layer for the tight (veto) working point. Additionally,
 1310 a photon conversion veto is applied by testing if a pair of electron tracks is originating from a
 1311 common displaced vertex. For the 8 TeV datasets more refined cuts are placed on the electron
 1312 object using a multivariate analysis. For the 13 TeV dataset this is replaced with more refined
 1313 cuts on the shower shape variables such as the difference in η or ϕ between the energy weighted
 1314 supercluster position in the ECAL and the track direction in at the innermost tracker position
 1315 ($\Delta\eta_{in}$, $\Delta\phi_{in}$), and the ECAL crystal based shower covariance in the η direction ($\sigma_{\eta\eta}$). These
 1316 cuts also include energy related variables such as the absolute difference between the inverse
 1317 electron energy measured in the ECAL and the inverse momentum measured in the tracker
 1318 ($|1/E - 1/p|$), and the ratio of the energy measured in the HCAL and ECAL (H/E). Unlike the
 1319 muon case, the identification criteria 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 pile up mitigation is applied. For this correction, the expected pile up energy inside the isolation cone is estimated from the median density energy per area of pile up contamination (ρ), computed event by event, and the effective area ($A_{eff.}$) [148]. This effective area is estimated from simulation and denotes the expected amount of neutral energy from pile up 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

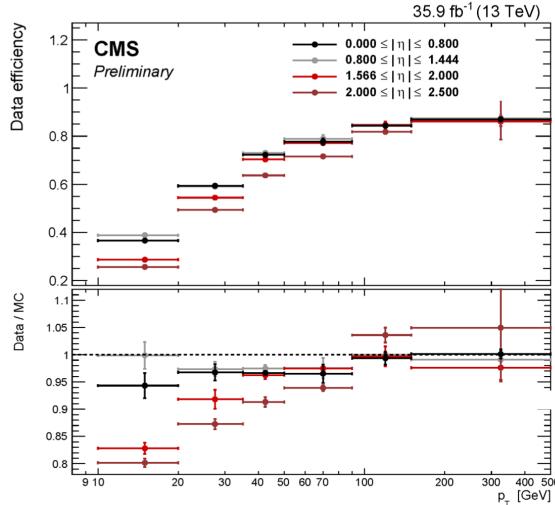
$$\mathcal{I}_e = \frac{\sum p_T(CH) + \max(0., \sum E_T(NH), \sum E_T(\gamma) - \rho \times A_{eff.})}{p_T(e)}. \quad (4.3)$$

1320

1321 The efficiency of electron identification is estimated from $Z \rightarrow e^- e^+$ events via the tag-and-
 1322 probe method and is shown in Figure 4.4 for the tight working point. The efficiencies reach
 1323 95 – 100%. The difference between data and simulation are corrected by dedicated p_T - and η
 1324 dependent scale factors as well.

Table 4.2: The effective areas A_{eff} used for the electron relative isolation [161].

η region	A_{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 [154].

1325 4.5.3 Jets

Jets are reconstructed from all reconstructed particles without the charged hadrons associated to pile up vertices (PF+CHS jets). The clustering is done via the anti- k_T algorithm [162] with a radius parameter for the cone size of the resulting jet of $R = 0.5$ for 8 TeV data and $R = 0.4$ for the more boosted 13 TeV dataset. The initial step of the anti- k_T algorithm considers all candidates as protojets and starts to calculate the distances for protojets i and j as

$$d_{ij} = \min \left(\frac{1}{p_{T,i}^2}, \frac{1}{p_{T,j}^2} \right),$$

$$d_i = \frac{1}{p_{T,i}^2}. \quad (4.4)$$

Table 4.3: 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$	
	Veto electron	Tight electron	Veto electron	Tight electron
$\sigma_{\eta\eta}$	< 0.0115	< 0.00998	< 0.037	< 0.0292
$ \Delta\eta_{in} $	< 0.00749	< 0.00308	< 0.00895	< 0.00605
$ \Delta\phi_{in} $	< 0.228	< 0.0816	< 0.213	< 0.0394
H/E	< 0.356	< 0.0414	< 0.211	< 0.0641
relative isolation	< 0.175	< 0.0588	< 0.159	< 0.0571
$ 1/E - 1/p $	< 0.299 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

1326 For each iteration the two distances are calculated. When $d_{ij} < d_i$, the two protojets are merged
 1327 and their four momentum is summed. If d_i is the smallest distance, the protojet is renamed
 1328 as final jet and ignored in the subsequent steps. More information about the jet algorithm
 1329 performance can be found in [163].

1330 The jets used for the analysis discussed in this thesis uses the loose identification working point
 1331 given. The criteria on the constituents of the jets are given in Table 4.4. These requirements
 1332 find their origin on the assumption that a proper jet originating from the hadronisation of a
 1333 quark or gluon consists of multiple PF particles and types. The jet should consist of more than
 1334 one constituent and the neutral hadron fraction and neutral EM energy fractions should be less
 1335 than 99%. For the jets within the tracker acceptance ($|\eta| < 2.4$), at least one constituent has to
 1336 be a charged hadron resulting in a charged hadron energy fraction above 0%. Additionally the
 1337 charged EM energy fraction should be less than 99%. On top of these requirements, objects
 1338 that are labelled as jets and found in vicinity of any isolated lepton, $\Delta R < 0.3$, are removed
 1339 from the jet collection in that event.

1340 The energy of the reconstructed jets deviate from the energies of the corresponding jets
 1341 clustered from the hadronisation products of true partons from simulations due to non linear
 1342 subdetetector response and efficiencies. The jet energy corrections (JEC) calibrate the jets
 1343 in order to have the correct energy scale and resolution. Jet energy scale corrections (JES)
 1344 are determined as a function of pseudorapidity and the transverse momentum from data and
 1345 simulated events by combining several channels and methods. This is extensively described
 1346 in [164]. These corrections account for the effects of pile up, the uniformity of the detector
 1347 response, and residual data-simulation jet energy scale differences. Furthermore, the jet energy
 1348 resolution (JER) is measured in data and simulation as function of pile up, jet size and jet flavour.
 1349 A detailed understanding of both the energy scale and the transverse momentum resolution

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

1350 of the jets is crucial for many physics analysis, and these are commonly the main source of
 1351 systematic uncertainty. The performance of the jet energy corrections for the 13 TeV dataset
 1352 can be found in [165].

1353 The JEC are factorised and subsequently correct for the offset energy due to pile up, the
 1354 detector response to hadrons, and residual differences between data and simulation as a function
 1355 of 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

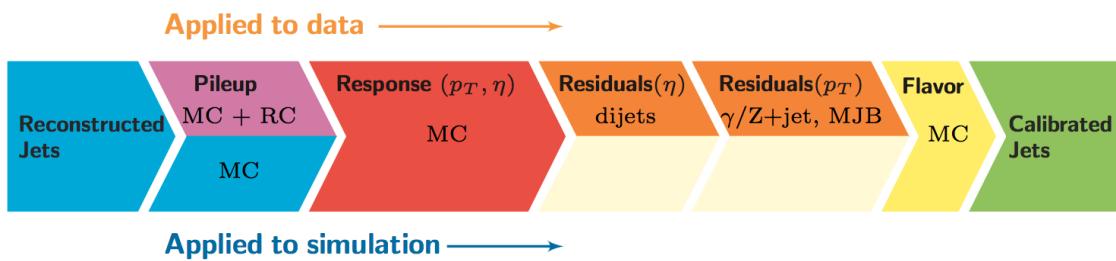


Figure 4.5: The sequence of the JEC for data and simulations. The corrections marked with MC are derived from simulation studies, while RC stands for random cone, and MJB for the analysis of multi-jet events. Figure taken from [164].

1356 additional pile up activity. It is based on the jet area method, which uses the effective area
 1357 of the jets multiplied by the average density in the event to calculate the offset energy to
 1358 be subtracted of the jets. The correction factors are derived by comparing the jet response
 1359 with and without pile up events overlaid. The residual differences between data and detector
 1360 simulation are determined using the random cone method (RC). For this method, many jets
 1361 are reconstructed in each event by clustering particles through placing random cones. This
 1362 provides a mapping of the $\eta\phi$ -space and the average p_T of those jets gives the average energy
 1363 offset due to pile up [164]. The next level of corrections have as goal to have an uniform
 1364 energy response independent of the transverse momentum or pseudorapidity of the jet. These
 1365 corrections are determined from simulated events by matching the reconstructed to true particle

1367 jets and comparing there momenta. The residual corrections between data and simulation are
 1368 determined by comparing the transverse momentum balance in various types of events (multi-
 1369 jet, Z + jets, and γ + jets), using a reference jet in the barrel region. The jet flavour corrections
 1370 are optional and not used for this thesis. More information on the jet flavour corrections can be
 1371 found in [164]. For jets with a transverse momentum above 30 GeV, the uncertainties from the
 1372 various corrections are 3-5% for the 13 TeV dataset [165].

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

$$c_{\text{matched}} = 1 + \frac{p_T^{\text{reco.}} - p_T^{\text{true}}}{p_T^{\text{reco.}}} (s_{\text{JER}} - 1), \\ c_{\text{unmatched}} = 1 + N(0, \sigma_{\text{JER}}) \sqrt{\max(s_{\text{JER}}^2 - 1, 0)}, \quad (4.5)$$

1373 where $N(0, \sigma_{\text{JER}})$ denotes a sample value from a normal distribution centred at zero with
 1374 as standard deviation the relative resolution in simulation σ_{JER} , and s_{JER} the η -dependent
 1375 resolution scale factors. These scale factors are derived from data from di-jet or γ + jets events
 1376 and analysing the p_T balance. The resolution scale factors (data/simulation) are found to be
 1377 1.1-1.2 except for the transition regions around $|\eta| = 3$ and $|\eta| = 1.4$ [165] 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

1378

1379 4.5.4 Jets from b fragmentation

1380 Jets originating from the hadronisation of bottom quarks can be discriminated from jets from glu-
 1381 ons and light-flavour quarks as well as charm quark fragmentation through the use of b-tagging.
 1382 There are a multitude of algorithms developed within CMS to perform b-tagging [166, 167] on

1383 jets that fall within the pseudorapidity acceptance of the trackers. These algorithms exploit
 1384 the properties of the b quark to identify the jets formed by its fragmentation. These hadrons
 1385 have relative large masses, long lifetimes and daughter particles with hard momentum spectra.
 1386 Additionally, their semi-leptonic decays can be exploited as well. To use b-jet identification
 1387 in an analysis, one needs to know its efficiency and misidentification probability. In general
 1388 these are function of the pseudorapidity and transverse momentum of the considered jet. Their
 1389 performances are directly measured from data by use of b-jet enriched jet samples (multi-jet or
 1390 top-quark decays).

1391 This thesis uses b-jets identified by the Combined Secondary Vertex version 2 (CSVv2)
 1392 algorithm [166]. This algorithm combines secondary vertices together with track based lifetime
 1393 information by use of a multivariate technique. The secondary vertex is reconstructed from
 1394 displaced tracks within a jet, as illustrated in Figure 4.6. The final state b quark is encapsulated
 1395 in a B meson (e.g. B^\pm , B_0 , B_S) after the hadronisation. This B meson has relatively long
 1396 lifetime and can travel a measurable distance from the primary vertex before decaying¹. After
 1397 reconstruction, the secondary vertices are required to be in accordance with the B meson
 1398 hypothesis bases on the amount of shared tracks with the primary vertex, the invariant vertex
 mass to reject kaon decays, and the direction of the tracks compared to the jet axis.

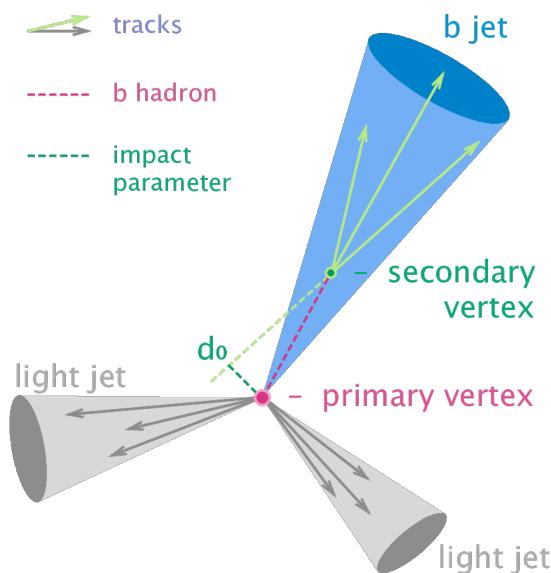


Figure 4.6: Sketch showing the common principle of the identification of b-jets. Figure taken from [168]

1399

1400 The b-tagging algorithm performances are evaluated taking into account two cases: discrim-
 1401 ination of b-tagged jets originating from charm quarks, and discrimination of b-tagged jets
 1402 against jets coming from gluons or light (u, d, s) quarks. In Figure 4.7, the misidentification
 1403 probabilities for different b-tagging algorithms within CMS are shown. Based on the misiden-
 1404 tification probabilities for a certain threshold on the CSVv2 discriminator, different working

¹For example, B^\pm mesons have a lifetime of about 1.6 ps [6] and travel 4-9 mm before decaying if their momenta is 40-100 GeV.

1405 points (WP) are defined. These are shown in [Table 4.6](#). The analysis presented in this thesis
 1406 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?

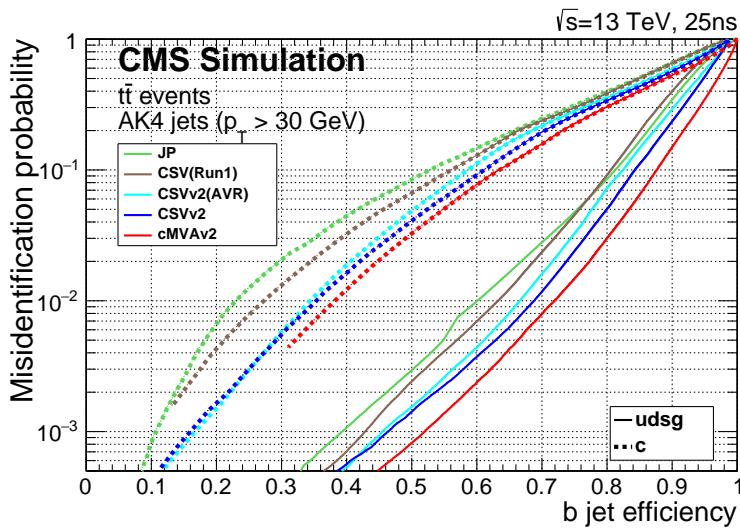


Figure 4.7: Misidentification probabilities of various b-tagging algorithm in simulation. Figure taken from [\[167\]](#).

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	≈ 81%	≈ 10%
Medium (M)	> 0.8484	≈ 66%	≈ 1%
Tight (T)	> 0.9535	≈ 46%	≈ 0.1%

1407

The efficiency of tagging a jet as coming from a bottom quark in simulation typically deviates somewhat from data. Efficiency scale factors $\epsilon_b^{\text{data}}/\epsilon_b^{\text{MC}}$ are derived from data to account for those differences. These scale factors are η -, p_T -, and flavour dependent, where the flavour of the jet is determined from matched generated hadrons. For cut based analyses these scale factors are applied to the b-tagging efficiencies and mistag rates according to the chosen working point [\[167\]](#). For shape based analysis however, such as the one presented in this thesis, the scale factor are applied on the distribution of the b-tagging discriminator. This is the so-called IterativeFit method [\[169\]](#). It uses a tag and probe method to measure the scale factors for both b, c, and light flavoured jets simultaneously. The scale factors to account for the differences in simulation and data for the probe jet are determined iteratively to account for the impact of the b-, c-flavour, and light flavour scale factors on eachother. In a fist step, no scale factors are applied. Then the scale factor is measured by applying the scale factors of the previous iteration to simulation until the scale factors become stable. Throughout the procedure, the scale factor

for charm jets are set unity with an uncertainty that is twice the one of the b scale factor. The scale factors obtained in η -, p_T -, and CSVv2 discriminant values are determined with the bin content N of the considered (η, p_T ,discriminant) bin as

$$\begin{aligned} \text{SF}_b &= \frac{N_b^{\text{data}} - N_b^{\text{MC}}}{N_b^{\text{MC}}}, \\ \text{SF}_{g,u,d,s} &= \frac{N_{g,u,d,s}^{\text{data}} - N_{g,u,d,s}^{\text{MC}}}{N_{g,u,d,s}^{\text{MC}}}, \\ \text{SF}_c &= 1. \end{aligned} \quad (4.6)$$

1408 The uncertainties related to the IterativeFit method cover possible shape discrepancies between
 1409 data and simulation. The uncertainty coming from the purity of the sample is subdivided into
 1410 two uncorrelated uncertainties based on the purity of the light flavoured (lf) and heavy flavoured
 1411 (hf) jet contributions in the sample. Furthermore, the jet energy scale results in jets migrating
 1412 from one p_T bin to an other, having an influence on bin dependent scale factors. The statistical
 1413 fluctuations of the limited amount of entries in each bin are also accounted for and have an
 1414 influence on the scale factor uncertainties. These have four uncorrelated sources: two for heavy
 1415 flavour and two for light flavour jets. Since the uncertainty on the scale factors for the jets
 1416 originating from a charm quark (cf) is determined from the uncertainty on the b scale factors
 1417 resulting in two independent uncertainties [169].

1418 4.5.5 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 to balance the vectorial sum of the transverse momenta of all particles:

$$\begin{aligned} E_T &= |\vec{p}_T|, \\ \vec{p}_T &= - \sum_{i=1}^{N_{\text{particles}}} \vec{p}_{T,i}. \end{aligned} \quad (4.7)$$

1419 The z-component can not be calculated from the momentum imbalance since the boost along
 1420 the z-axis, determined by the momentum fraction, can not be reconstructed.

The missing transverse energy is influenced by the minimum thresholds in calorimeters, the inefficiencies in the tracker, and the non-linear response of the calorimeter to hadronic particles. The bias is reduced by correcting the transverse momentum of the jets too particle jet p_T via the JEC and propagating it to the missing transverse momentum taking into account the energy

$$\begin{aligned} \vec{p}_T^{\text{corr}} &= - \sum_{i=1}^{N_{\text{jets}}} \vec{p}_{T,i}^{\text{corr.}} - \sum_{i=1}^{N_{\text{unclustered}}} \vec{p}_{T,i}^{\text{raw}}, \\ \vec{p}_T^{\text{corr}} &= \vec{p}_T^{\text{raw}} - \sum_{i=1}^{N_{\text{jets}}} (\vec{p}_{T,i}^{\text{JEC}} - \vec{p}_{T,i}^{\text{PU-only}}). \end{aligned} \quad (4.8)$$

1421 The $\vec{p}_{T,i}^{\text{PU-only}}$ denotes the transverse momentum of the jet, where only the pile up related
 1422 corrections are applied. The performance of the missing transverse energy reconstruction can
 1423 be found in [170].

1424 4.6 Summary of corrections

1425 Throughout the chapter several corrections are introduced to improve the agreement between
 1426 data and simulation. These corrections are sources of systematic uncertainties for the anal-
 1427 ysis presented in this thesis. Therefore a summary of the corrections and their associated
 1428 uncertainties is provided.

1429 **Lepton scale factors** The systematic uncertainty on the lepton scale factors consist of three
 1430 sources: identification, isolation and tracking. The applied scale factors are varied
 1431 independently within one standard deviation of their measured uncertainties to account
 1432 for their systematic impact on the measurements.

1433 **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.
 1434 Furthermore, residual corrections and smearing is applied to match the overall energy
 1435 scale and resolution for simulation and data. These corrections are also propagated to
 1436 the missing transverse energy. The systematic uncertainties due to these scale factors are
 1437 estimated by varying them within their uncertainties and repeating the measurements
 1438 with recalibrated jets and missing transverse energy.

1440 **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.

1455 **Pile up** Varying the minimum bias cross section, used to calculate the pile up distribution by $\pm 4.6\%$, results in a systematic shift in the pile up distribution. The uncertainty is estimated by recalculating the pile up weights to the distributions associated to the minimum bias cross sections.

1459 **Luminosity** The luminosity is measured with a global uncertainty of 2.5%, affecting the ex-
 1460 pected number of events.

Event selection and categorisation

5

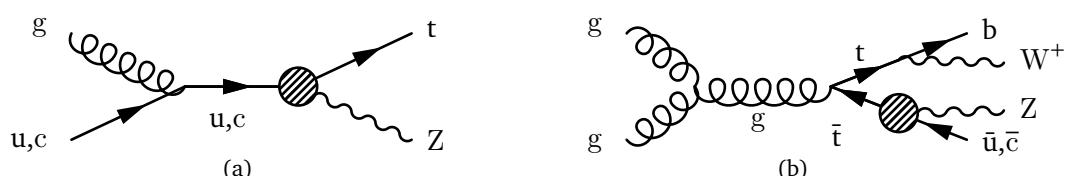
1461

1462 A basic event selection is made for selecting signal like events. The necessary event requirement
1463 are discussed in [Section 5.2](#).

1464 The analysis uses signal and background regions to constrain the huge SM background
1465 compared to the expected signal. [Section 5.4](#) discusses each region that is entering the analysis.
1466 On top of the use of background estimation from control regions, backgrounds that have prompt
1467 leptons contaminated by real leptons either from decays of tau leptons or from hadronized
1468 mesons or baryons (collectively commonly referred as “non-prompt leptons”) as well as by
1469 hadrons or jets misidentified as leptons¹ are evaluated with a data-driven method discussed in
1470 [Section 5.3](#).

1471 5.1 Baseline event selection and filters

1472 The CMS collaboration recorded in the course of 2016 proton at 13 TeV collisions data with a
1473 total recorded integrated luminosity of $35.9 \pm 2.5\% \text{ fb}^{-1}$. The baseline event selection has a
1474 goal to substantially reject SM background events, whilst maintaining a high signal efficiency.
1475 In this analysis a search is performed in a final state made up of a Z boson and a top quark,
1476 associated or not with a jet, in the leptonic decay of the Z boson and the top quark. The leading
order Feynman diagrams can be seen in [Figure 5.1](#).



NOTE:
adapt to
leptonic de-
cay

Figure 5.1: Feynman diagrams for the tZq FCNC interaction with a fully leptonic decay, where the FCNC interaction is indicated with the shaded dot. (a) Single top production through an FCNC interaction. (b) Top pair production with an FCNC induced decay.

1477

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

1478 The signal considers both the single top FCNC (tZ in the final state) and the top pair FCNC
 1479 (tZq in the final state) events. Their final state signatures consist of three leptons, considering
 1480 electrons or muons in our analysis, and a jet originating from a b quark. For FCNC tZq , there is
 1481 an additional up or charm jet. Leptons from tau decays are not vetoed and are entering the
 1482 analysis. Four different lepton channels based on lepton flavour are considered (eee , $e\mu\mu$,
 1483 $\mu\mu\mu$).

1484 The CMS trigger system, described in [Section 2.2.3](#), filters out the main of the collision events
 1485 from uninteresting processes and dedicated trigger paths are defined to single out the events
 1486 with our required detector signature. The trigger paths are chosen based on on-line triggering
 1487 objects with at least one muon (M), at least one electron (E), at least two muons (MM), at
 1488 least two electrons (EE), at least one muon and an electron (ME), at least three muons (MMM),
 1489 at least three electrons (EEE), at least two muons and one electron (MME), or at least two
 1490 electrons and one muon (EEM). For the MC simulation a simple *or* of all triggers is taken and
 1491 the event is taken if it passes one of the trigger paths. For data however, double counting of the
 1492 same event has to be taken into account and a procedure to avoid double counting has been
 1493 put into place. It consists of vetoing in a given dataset the events that are already selected in
 another, as given in [Table 5.1](#).

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

Dataset	Trigger Logic
$e\mu$	$EM \parallel EEM \parallel MME$
$\mu\mu$	$(MM \parallel MMM) \&& !(EM \parallel EEM \parallel MME)$
ee	$(EE \parallel EEE) \&& !(MM \parallel MMM) \&& !(EM \parallel EEM \parallel MME)$
single μ	$M \&& !(EE \parallel EEE) \&& !(MM \parallel MMM) \&& !(EM \parallel EEM \parallel MME)$
single e	$E \&& !M \&& !(EE \parallel EEE) \&& !(MM \parallel MMM) \&& !(EM \parallel EEM \parallel MME)$

1494

1495 This trigger selection strategy is to allow the maximum statistics on the signal region since it
 1496 does not discard events from any dataset. For the single lepton triggers, at least one electron
 1497 (muon) with a transverse momentum p_T higher than 32 GeV (24 GeV) is required. The dilepton
 1498 triggers require an electron (muon) with $p_T > 23$ GeV and a muon (electron) with $p_T > 8$ GeV,
 1499 or an electron (muon) with $p_T > 23$ GeV (17 GeV) and an electron (muon) with $p_T > 12$ GeV
 1500 (8 GeV). Events collected by the trilepton triggers require an electron (muon) with $p_T > 16$ GeV
 1501 (12 GeV), a second electron (muon) of $p_T > 12$ GeV (10 GeV), and a third electron (muon)
 1502 with $p_T > 8$ GeV (5 GeV). The mixed trilepton trigger events require two electrons (muons)
 1503 with $p_T > 12$ GeV (9 GeV) and a third muon (electron) with $p_T > 8$ GeV (9 GeV). The HLT
 1504 trigger paths used in data and simulation are summarised in [Table 5.2](#).

1505 In order to ensure a full trigger efficiency, the offline p_T thresholds are set higher than the on-
 1506 line trigger thresholds. Selected electrons (muons) are required to have $p_T > 35$ (30) GeV and
 1507 $|\eta| < 2.1(2.4)$. A quantity for evaluating lepton isolation is calculated as the summed energy of

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

1508 all particles (charged hadrons, neutral hadrons, photons) in a cone of radius $\Delta R < 0.3$ (0.4)
 1509 around the electron (muon), excluding the electron (muon) itself, and divided by the lepton p_T
 1510 . An electron candidate is selected if this isolation quantity is below 0.059 in the barrel region
 1511 and below 0.057 in the end caps. A muon candidate is selected if this isolation quantity is
 1512 below 0.15. Other lepton selection criteria are applied in analysis based on the values of various
 1513 quantities determined during the reconstruction, such as number of missing track hits or the
 1514 electromagnetic shower created by the particles. These are given in [Section 4.5.1 \(Table 4.1\)](#)
 1515 and [Section 4.5.2 \(Table 4.3\)](#). The trigger efficiency estimation is described in [Section 5.1.2](#).

1516 The samples are pre-selected off-line to ensure that all reconstructed particles considered
 1517 for the analysis are corresponding to a proton interaction and that signals from beam halo
 1518 particles as well as detector noise is removed. For this reason, several filters are used. These are
 1519 described in [Section 5.1.1](#).

1520 On top of leptons, jets and missing transverse energy are expected from the signal signature.
 1521 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. More information about the jets used for this analysis can be found in [Section 4.5.3](#). 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. A selected jet may still overlap with the selected leptons leading to a double-counting. 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 set 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 4.5.4](#).

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 4.5.5](#). 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. The events and their corresponding object collections reconstructed using the reconstruction criteria described in [Section 4.1](#), are used as input for the analysis. Further requirements on the momenta and the pseudo rapidities are made to fulfil the trigger requirements and reconstruction algorithms.

5.1.1 Event cleaning

Some events arising from instrumental noise and beam backgrounds might end up in the data [[170](#), [171](#)]. Spurious deposits may appear in the ECAL from non collision origins such as beam halo particles, or from particles striking the sensors in the ECAL photo detectors. Conjointly, dead ECAL cells can cause artificial missing transverse energy. The HCAL can cause spurious energy from particle interactions with the light guides and the photomultiplier tubes of the HF, as well as noisy hybrid photo diodes. In CMS, different algorithms, so-called filters, are developed to identify and suppress these events.

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

1563 The machine induced particles, via beam-gas / beam-pipe/... interactions, that are flying with
 1564 the beam affect the physics analysis. They leave a calorimeter deposit along a line at constant
 1565 ϕ in the calorimeter, and interactions in the CSCs will often line up with this deposit. This
 1566 can be seen in [Figure 5.2](#). Therefore, events containing such beam halo particles are removed
 1567 from the selection with the CSC Beam Halo Filter. This filter uses information related to the
 1568 geometric quantities, energy deposits, and timing signatures. For 2016, the filter rejects 85% in
 1569 a halo-enriched sample, whereas the mistag probability determined from simulation if found to
 be less than 0.01%.

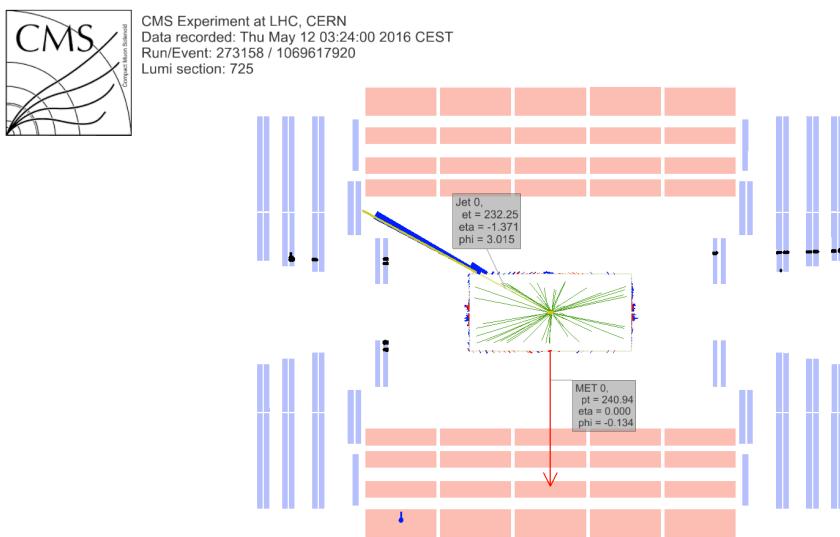


Figure 5.2: Event display of a beam halo event with collinear hits in the CSC (black), missing transverse energy of 250 GeV and a jet of 232 GeV. The hadronic deposit is spread in η , but narrow in ϕ . Figure taken from [\[170\]](#).

1570

1571 Furthermore, there is anomalous high missing transverse energy coming from low quality
 1572 muons that lead to high- p_T tracks, but are considered not good by the particle flow algorithm.
 1573 These low quality tracks will be mislabelled as charged hadrons and will therefore be used in
 1574 the calculation of the missing transverse energy. By investigation the purity of the reconstructed
 1575 tracks and the relative transverse momentum error of the muons, these events can be filtered
 1576 out.

1577 Supplementary to previous filters, only events with where the first primary vertex is a well
 1578 reconstructed primary vertex are selected. The reconstructed primary vertex should have at
 1579 least five degrees of freedom, the longitudinal distance from the beam spot is maximally 24 cm
 1580 ($d_z < 24$ cm), and the transversal distance from the beam spot is maximally 2 cm ($d_{xy} < 2$ cm).

1581 **5.1.2 Estimation of the trigger efficiency**

1582 The trigger efficiency in data is estimated using a data sample collected using unprescaled E_T^{miss}
 1583 triggers. These allow events with a missing transverse energy higher than 110 GeV(120 GeV)
 1584 and that the scalar sum of the transverse momenta of the reconstructed PF jets $H_T^{\text{trig.}}$ is at least
 1585 300 GeV (120 GeV), or a calorimeter/PF E_T^{miss} higher than 200 GeV/300 GeV. For an HB-HE
 1586 cleaned event the PF missing transverse energy threshold is lowered to 170 GeV. These trigger
 1587 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

The studied simulation sample is the main background, 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

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

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

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

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

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

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	$\mu\mu\mu$	eee	$ee\mu$	$e\mu\mu$
	1.0000	1.0000	0.9541	1.0006	1.0004%

1589

1590 The trigger efficiencies are also measured in function of the p_T of the leptons for which the
 1591 distributions can be found in [Appendix A](#). The resulting scale factors in can be found in Figure
 1592 [Figure 5.3](#). The trigger efficiencies are measured to be nearly 100% for both simulation and
 1593 data. The results are dominated by statistics and assigning a large uncertainty to the trigger
 1594 efficiency based on the dataset collected by E_T^{miss} triggers would be over conservative. A one
 1595 percent uncertainty on the trigger selection for the $ee\mu$ and $\mu\mu\mu$ final states, and 5% for the
 1596 eee and $e\mu\mu$ final states is assigned instead. No scale factors will be applied on simulation as
 1597 they are close to unity. Control plots are made in the dilepton region to validate all corrections
 1598 applied to simulation and can be found in [Appendix B](#).

NOTE:
Make sure
this is the
case

1599 5.1.3 Corrections

1600 Mismatches between data and simulation are corrected for via the use of scale frcators. These
 1601 are elaborately discussed in [Section 4.5](#). In this section a short overview of the applied corrections
 1602 is given and their effect on a dilepton selection is shown.

1603 Pile up reweighting

1604 In data, the number of interactions per bunch crossing (pile up) is calculated with a minimum
 1605 bias cross section of 69.2 mb. The number of simulated pile up events is then reweighed to
 1606 match the expected number of pile up events in data. Pile up reweighting manifests itself as an
 1607 altered shape of the number of reconstructed primary vertices as can be seen in [Figure 5.4](#).

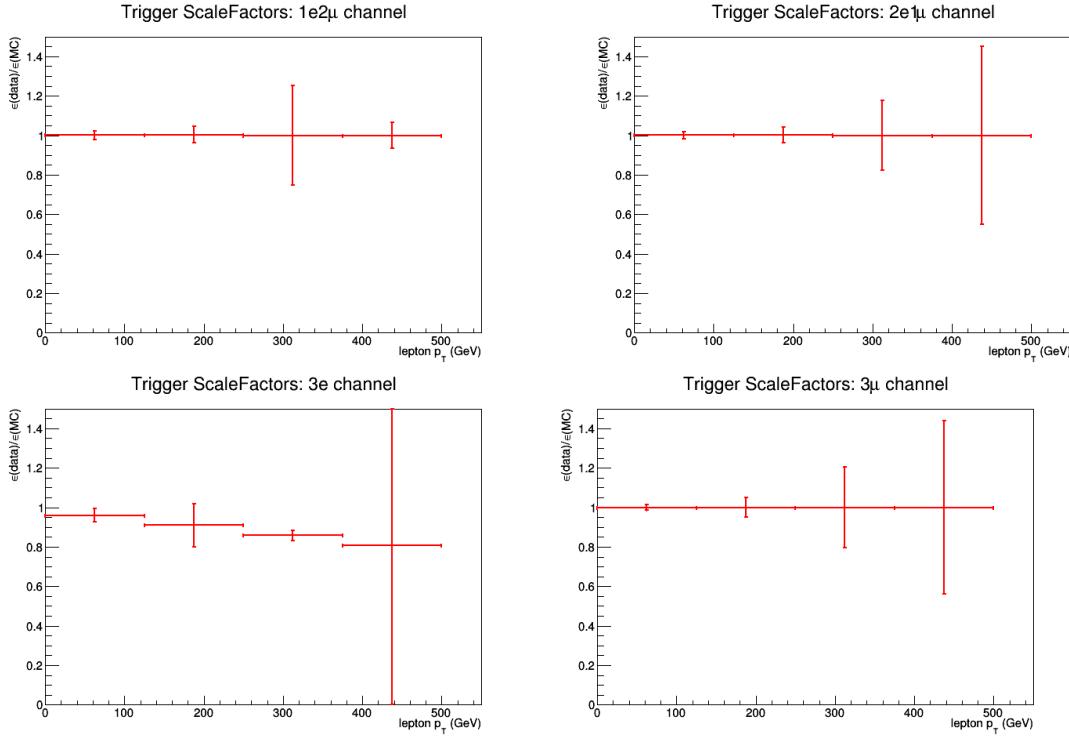


Figure 5.3: The trigger scale factors measured as a function of lepton p_T , using the dataset collected by E_T^{miss} triggers and WZ simulation, after a 3 lepton and jets selection, in the Z mass window. All corrections to simulation are applied. Left, upper: eeμ channel. Right, upper: eeμ channel. Left, lower: eee channel. Right, lower: μμμ channel

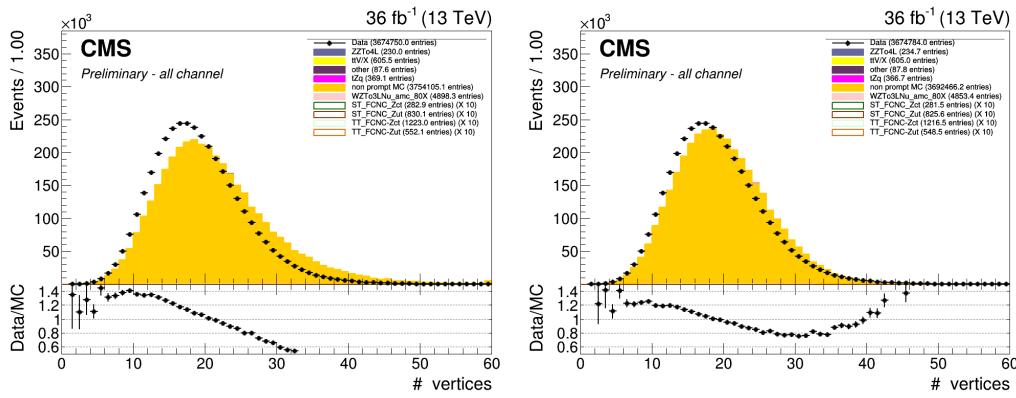


Figure 5.4: The number of primary vertices before (left) and after (right) pile up reweighting. After a 2 lepton plus jets selection, in the Z mass window.

1608 Note that [Figure 5.4](#) indicates that even after pile up reweighting, the primary vertex multiplicity
 1609 is not well described by simulation. This is a known effect, and using a minimum bias
 1610 cross section with a slightly lower value is found to better describe the data. However, the b
 1611 tagging scale factors are only provided for the nominal inelastic cross section, and thus this
 1612 value is used.

1613 **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 4.5.1](#) and [Section 4.5.2](#)). Multiplying these scale factors for each lepton provides an overall efficiency:

$$SF_{\text{global}}^{\mu} = \prod_i^{\# \mu} SF_{\text{ID}}^{\mu}(p_T, \eta) SF_{\text{Iso.}}^{\mu}(p_T, \eta) SF_{\text{Trig.}}^{\mu}(p_T, \eta) SF_{\text{track}}^{\mu}(p_T, \eta), \quad (5.5)$$

$$SF_{\text{global}}^e = \prod_i^{\# e} SF_{\text{ID}}^e(p_T, \eta) SF_{\text{Iso.}}^e(p_T, \eta) SF_{\text{Trig.}}^e(p_T, \eta) SF_{\text{track}}^e(p_T, \eta). \quad (5.6)$$

1614 The effect of the scale factors can be found in [Figure 5.6](#) for the electron scaling and [Figure 5.5](#)
 1615 for the muons. The trigger efficiencies are estimated in [Section 5.1.2](#).

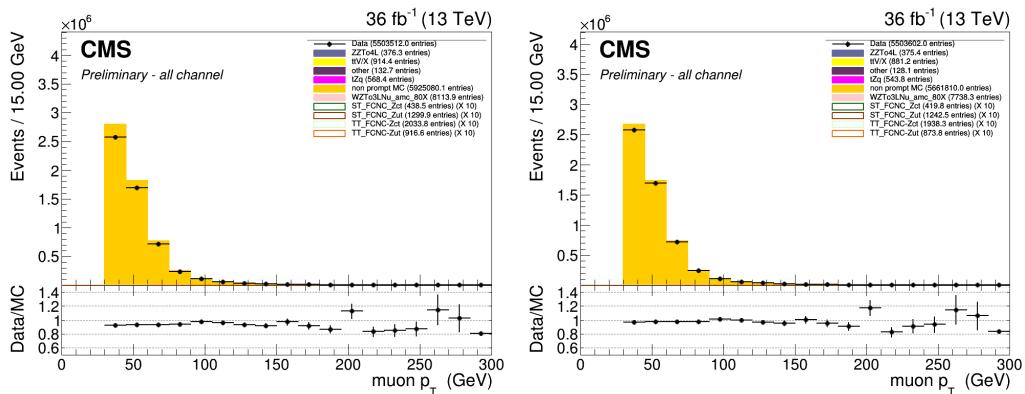


Figure 5.5: The p_T of the muons before (left) and after (right) muon scale factors. After a 2 lepton plus jets selection, in the Z mass window. Both after the Rochester correction.

1616 Additionally, corrections are made for the energy resolution of the leptons. For the electrons,
 1617 energy smearing and regression is applied [172]. The energy regression uses the detector
 1618 information to correct the electron energy in order to have the best energy resolution by
 1619 correcting local energy containment, material effects, etc.. The energy scale and smearing

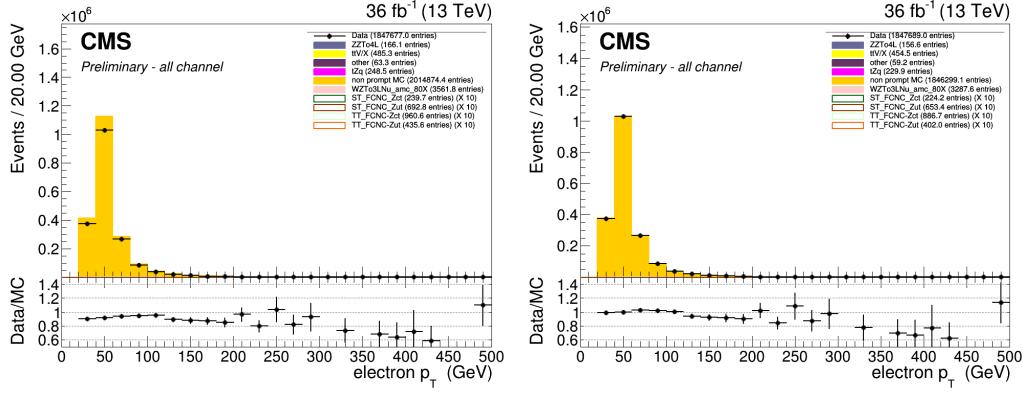


Figure 5.6: The p_T of the electrons before (left) and after (right) electron scale factors. After a 2 lepton plus jets selection, in the Z mass window. Both after energy scale corrections and smearing.

1620 is done in order to bring the data energy scale to simulation level. It smears the simulation
 1621 energies to have identical energy resolution in simulation and data. For the muons, the p_T is
 1622 corrected using the Rochester method [173, 174]. This correction removes the bias of the muon
 1623 p_T from any detector misalignment or any possible error of the magnetic field. The effect of
 1624 the Rochester correction can be found in [Figure 5.7](#).

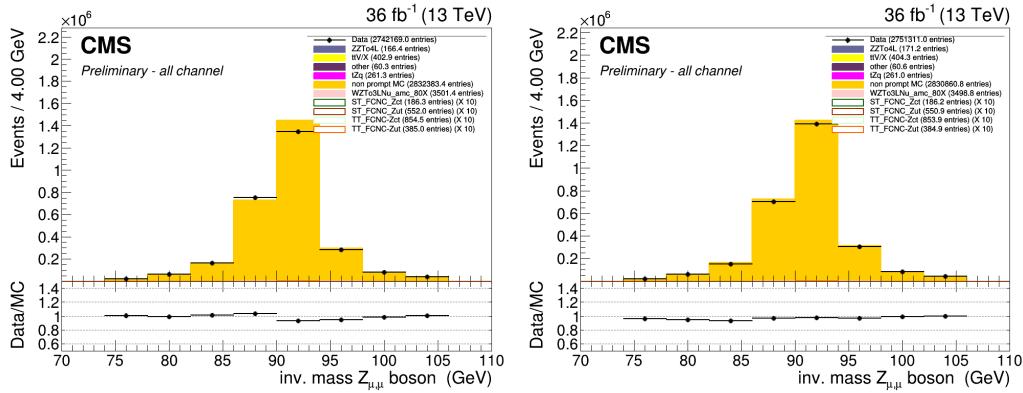


Figure 5.7: The mass of the Z boson consisting of the muons before (left) and after (right) the rochester correction. After a 2 lepton plus jets selection, in the Z mass window.

1625 **CSVv2 shape correction**

1626 In order to make the shape of the CSVv2 b-tagging discriminant in simulation agree with data,
 1627 jet-by-jet based scale factors are applied. These scale factors are a function of the p_T , η and
 1628 CSVv2 value of the jet as discussed in [Section 4.5.4](#). The effect of these scale factors can be
 1629 found in [Figure 5.8](#).

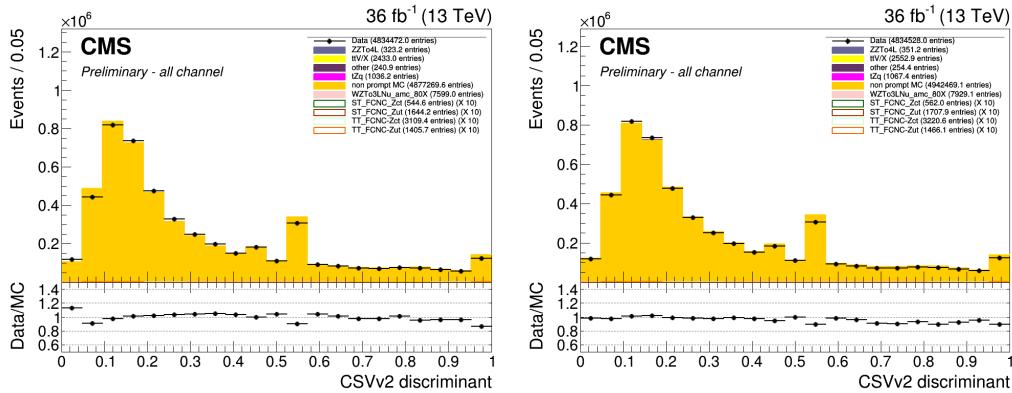


Figure 5.8: The CSVv2 discriminant of the jets before (left) and after (right) b-tag scale factors. After a 2 lepton plus jets selection, in the Z mass window.

1630 **Jet energy**

1631 The jet energy in data and simulation is corrected by the measured energy response of the
 1632 detector. This provides p_T - η dependent scale factors and are directly taken from the frontier
 1633 condition database as discussed in [Section 4.5.3](#). The effect of the jet energy corrections can be
 1634 found in [Figure 5.9](#) and [Figure 5.10](#).

NOTE: Fix
jer plot

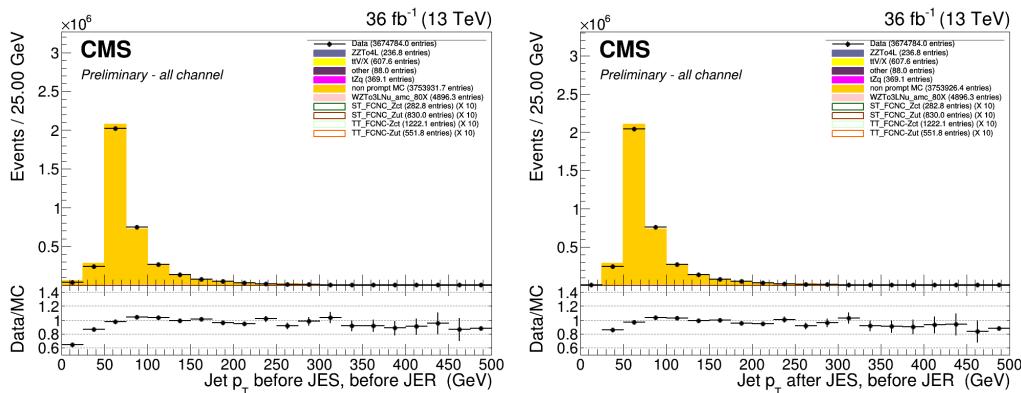


Figure 5.9: The p_T of the jets before (left) and after (right) jet energy scale corrections. After a 2 lepton plus jets selection, in the Z mass window.

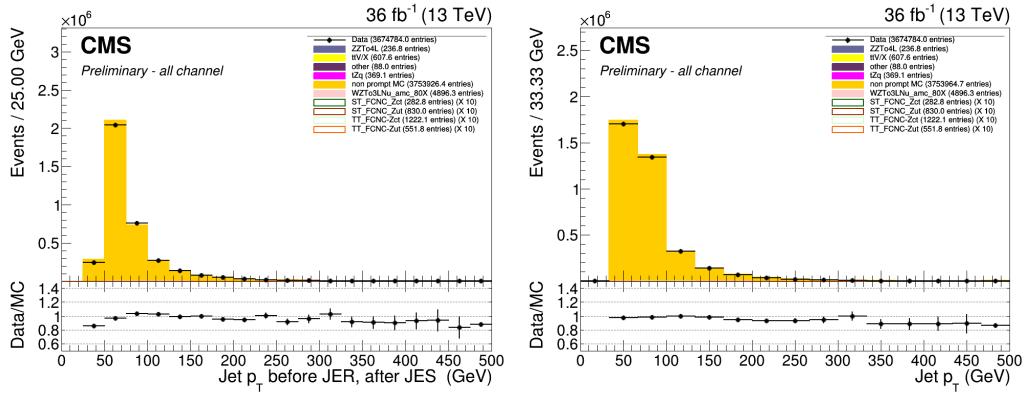


Figure 5.10: The p_T of the jets before (left) and after (right) jet energy resolution smearing. After a 2 lepton plus jets selection, in the Z mass window.

1635 Missing transverse energy

1636 The energy scale and resolution corrections applied to the jets are propagated back to the
1637 missing transverse energy (smeared Type I correction). This rebalances the transverse net
1638 momentum of the event and improves the missing transverse energy resolution itself.

1639 5.2 Analysis Strategy

1640 The analysis strategy uses five statistically independent regions to extract limits using a likelihood
1641 fit of various observables. Two signal regions, the tZ (STS_R) and tZ_q (TTS_R) signal region, are
1642 constructed using the jet multiplicity, focussed on each signal signature (see Tab. 5.7). In order
1643 to constrain the rate of WZ+jet events as well as that of NPL backgrounds three control regions
1644 are defined. The WZ control region (WZCR) focusses on NPLs originating from Z/γ^* + jets and
1645 simultaneously constrains the WZ+jets background rate. The NPL backgrounds coming from
1646 $t\bar{t}$, are constrained by two control regions, TT_{CR} and ST_{CR}, one for each signal region (TTS_R
1647 and ST_R). In the ST_R and TTS_R multivariate discriminants based on Boosted Decision Trees
1648 (BDT) are used to respectively discriminate FCNC tZ and FCNC tZ_q from backgrounds. In the
1649 WZCR a discriminating variable between the two backgrounds, WZ+jets and NPLs, is used.
1650 In TT_{CR} and ST_{CR} the dominating process is $t\bar{t}$ +jets, and its rate is estimated by subtracting
1651 all other background predictions from data. A simultaneous global fit is performed taking
1652 into account each region (STS_R, TTS_R, WZCR, TT_{CR} and ST_{CR}) for the four different leptonic
1653 channels.

1654 5.3 Data driven NPL background simulation

1655 The MC samples are used to model the backgrounds as well as for training the boosted decision
1656 trees for signal to background separation. One of the most important background consist of
1657 events with not prompt leptons. These are mostly instrumental background and are therefore
1658 very difficult to model. The NPL background is estimated from data for both its shape and its
1659 normalisation.

1660 The NPL sources are

- 1661 • hadronic objects wrongly reconstructed as leptons,
- 1662 • real leptons coming from the semi leptonic decay of a b or c hadron,
- 1663 • real leptons coming from the conversion of photons,

1664 that pass the identification and isolation requirements. The dominant source of these NPLs
 1665 depend on the flavour of the lepton and therefore the events with a noy prompt muon (NP μ)
 1666 are treated differently than those with a not prompt electron (NPE). For muons, the dominant
 1667 source is the semi leptonic decay of heavy flavour hadrons. For electrons, the dominant sources
 1668 are hadrons and photon conversions.

1669 The backgrounds causing NPL contributions are mostly $Z/\gamma^* + \text{jets}$ (Drell–Yan) and $t\bar{t} + \text{jets}$
 1670 dilepton processes, and in a smaller amount WW and tWZ. All of these backgrounds contain
 1671 two real leptons and one NPLDue to the fact that the probability for a lepton to be a NPL is
 1672 small, backgrounds containing two or more NPL s are neglected. The assumption is made that
 1673 for DY the two leptons compatible with a Z boson decay are the real leptons, and the additional
 1674 lepton is coming from a NPL source, while for $t\bar{t} + \text{jets}$ the NPL is associated to the Z boson.

1675 The NPL sample is constructed from data by requiring exactly three leptons, from which two
 1676 are considered real, isolated leptons and the third is a NPL. This NPL is created by loosening its
 1677 identification and inverting its isolation criteria. The full requirements on the not prompt leptons
 1678 are given in [Table 5.5](#) and [Table 5.6](#). For NPEs, a large fraction is coming from misidentified
 1679 photons. These are removed by applying a tighter cut on the $1/E - 1/p$ variable, and by limiting
 1680 the isolation values to be smaller than one.

1681 The NPL samples are defined in a given control region and are used to describe their contribu-
 1682 tion in the other regions.

1683 5.4 Regions and channels

1684 The regions are defined as in Table [5.7](#) after a common selection of

- 1685 • exactly 3 leptons containing one opposite sign, same flavour pair that are assigned to the
 1686 Z boson,
- 1687 • at least 1 jet and at the most 3 jets,
- 1688 • the transverse mass of the W boson to be maximal 300 GeV,

The cut on the transverse mass of the W boson is done to remove events that are passing the events cleaning elading to anomalous large missing transverse energy. 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 5.5: Non prompt electron requirements used in this analysis. The requirements are set in the barrel ($|\eta_{supercluster}| \leq 1.479$) and the end caps ($|\eta_{supercluster}| > 1.479$).

	$ \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 5.6: Non prompt muon requirements used in the analysis.

	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	$> 30 \text{ GeV}$

Table 5.7: 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

1689 Additional leptons are vetoed in order to reduce the contamination of backgrounds with four
 1690 or more leptons in the final state, e.g. ZZ, $t\bar{t}Z$, and $t\bar{t}H$. The most important backgrounds are
 1691 the ones that contain three prompt leptons in the final state. These are mainly WZ+jets, $t\bar{t}Z$
 1692 and SM tZq. For these backgrounds, the three lepton topology is identical to the FCNC signal:
 1693 two opposite sign leptons of the same flavour decaying from the Z boson, and a third additional,
 1694 high p_T lepton coming from the W boson decay.

1695 For the FCNC tZ final state, one b jet coming from the SM top decay is expected. For the
 1696 FCNC tZq, an additional light jet is expected. In the $t\bar{t}Z$ final state, two b jets are present in the
 1697 final state. However, due to inefficiencies of the b-tagging algorithm, one of the two b jets may
 1698 be identified as a light quark jet, giving the same final state as the FCNC tZq final state. For
 1699 the WZ+jets final states, one of the b jets produced by gluon splitting can be b-tagged or light
 1700 flavour jets coming from the WZ+jets production can be mis-tagged as b jets. The SM tZq final
 1701 state expects the same signal as FCNC tZq.

1702 The NPL events give a significant background contribution. This background is coming mainly
 1703 from $Z/\gamma^* + \text{jets}$ and $t\bar{t} + \text{jets}$ processes (in a less significant way, also WW and tWZ contributes),
 1704 which have very high cross sections and causes a large number of NPL background events
 1705 compared to signal.

1706 In order to reduce the large uncertainties in backgrounds, five independent regions are used
 1707 as defined in Table 5.7.

1708 5.4.1 WZ control region

1709 In this region, a fit is performed on the transverse mass of the W boson, in order to estimate
 1710 the NPL yield 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 one b-tagged jet, or at least one b-tagged jet. For this the probability of tagging at least one jet with the CSVv2 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

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

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

NOTE: [1715](#)
make ap-
pendix [1716](#)

1718 5.4.2 TTCSR and STCSR

The TTCSR and STCSR have the same selection criteria as TTSR and STSR but are outside the Z mass window (sidebands):

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

where $M(Z_{\text{reco}})$ is the reconstructed mass of the Z boson in the event, and $M(Z)$ the mass of the Z boson. These regions are dominated by $t\bar{t}$ +jets (see [Appendix C](#)) and are used to estimate the NPL coming from $t\bar{t}$ +jets in the STSR and TTSR. Since there aren't enough events entering these regions, no shapes are used in the fit. The distribution of the mass of the Z boson is flat for $t\bar{t}$ +jets events, as shown in Fig. [Figure 5.11](#), 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

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

1719 The resulting transfer factors are given in [Appendix C](#). The expected yield in the signal region
1720 estimated from the TTCSR (STCSR) is in agreement with the yield calculated from simulated
1721 events.

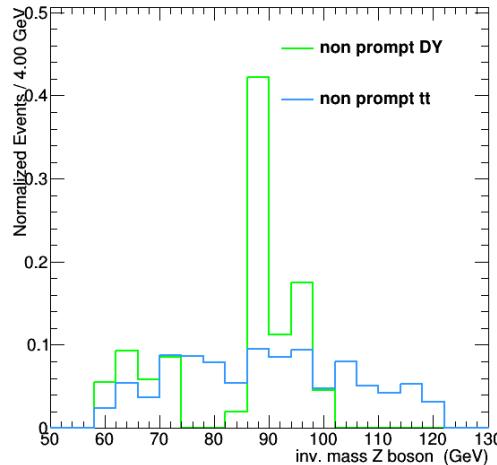


Figure 5.11: The normalized distribution for $Z/\gamma^* + \text{jets}$ and $t\bar{t} + \text{jets}$ events before dividing the events in to regions, after $|m_Z^{\text{reco}} - m_Z| < 30$ GeV. All leptonic channels combined.

1722 5.4.3 TTSR and STSR

1723 The TTSR is defined to target top quark pair FCNC (tZq), while the STSR focusses on single top
 1724 quark FCNC (tZ). They have NPL contributions coming from $Z/\gamma^* + \text{jets}$ and $t\bar{t} + \text{jets}$ events. In
 1725 this region, the data driven NPL template is split into two templates based on the presence of
 1726 the NPL in the Z boson:

- 1727 • NPL associated with W boson is assigned to $Z/\gamma^* + \text{jets}$ and estimated in the WZCR.
- 1728 • NPL associated with Z boson is assigned to $t\bar{t} + \text{jets}$ and estimated in the TTCR and STCR.

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

1729

6

1730 **6.1 Construction of template distributions**

1731 **6.2 Systematic uncertainties**

1732 **6.3 Limit setting procedure**

1733 **6.4 Result and discussion**

Denouement of the top-Z FCNC hunt at 13 TeV

1734

7

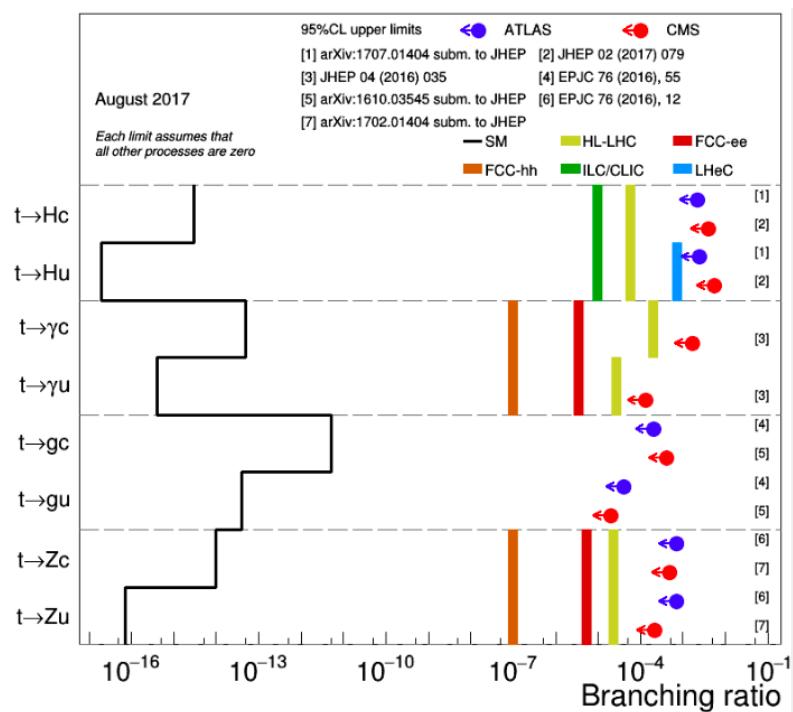


Figure 7.1:

1735

Trigger scale factors



1736 The trigger scale factors measured as a function of lepton p_T , using the dataset collected by
 1737 E_T^{miss} triggers and WZ simulation, after a 3 lepton and jets selection, in the Z mass window. All
 1738 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 $\mu\mu\mu$ 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.

$\mu\mu\mu$ 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 eee 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.

eee 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%
STSR	4/4 = 100.00%	70.71%	176/176 = 100.00%	10.66%
TTSR	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 ee 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.

ee 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%
STSR	1/1 = 100.00%	141.42%	255/255 = 100%	8.86%
TTSR	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 e $\mu\mu$ 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.

e $\mu\mu$ 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%
STSR	1/1 = 100.00%	141.42%	423/423 = 100.00%	6.88%
TTSR	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%

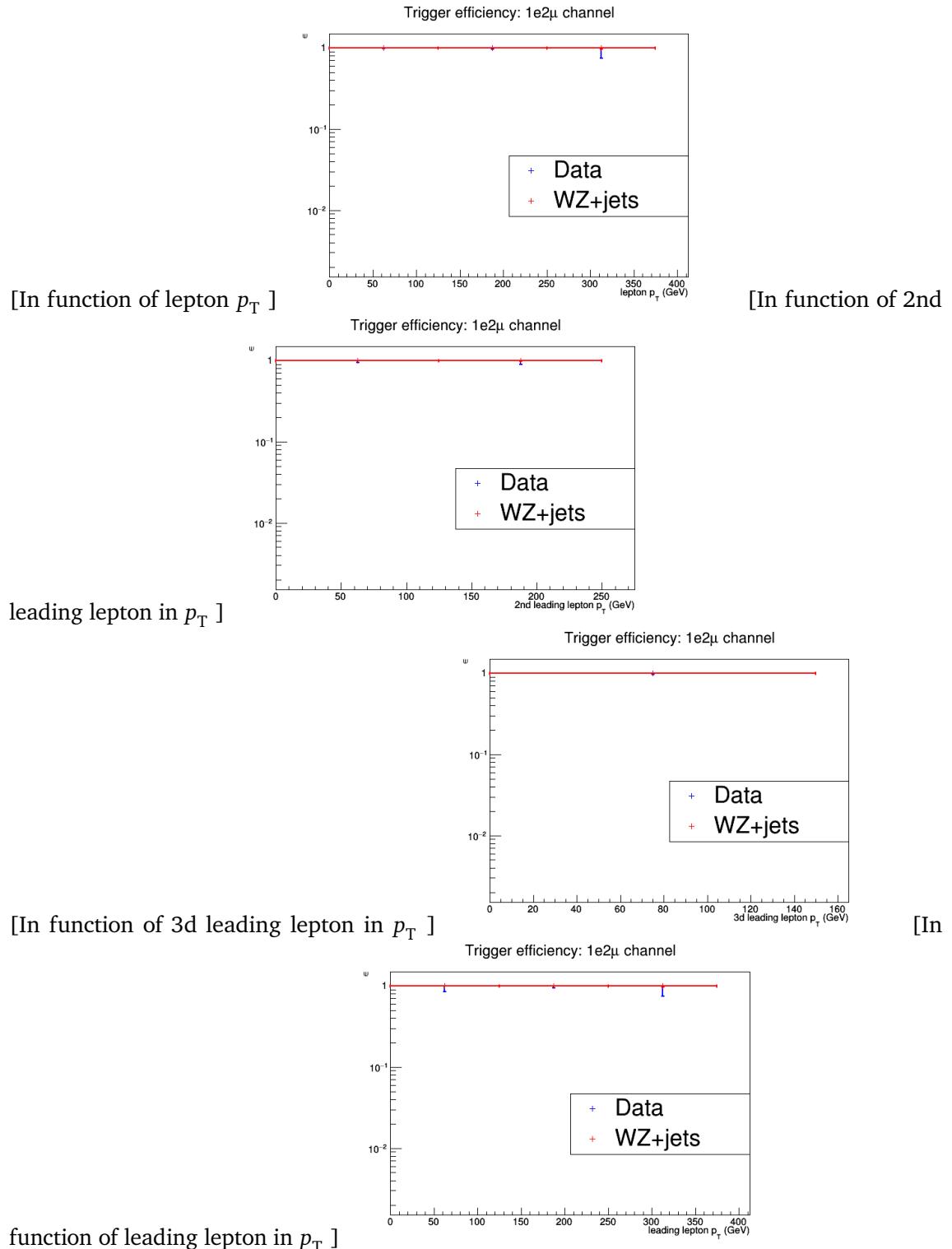


Figure A.1: The trigger efficiencies measured as a function of lepton p_T , using the dataset collected by E_T^{miss} triggers and WZ simulation, after a 3 lepton and jets selection, in the Z mass window. All corrections to simulation are applied. 1e2μ channel.

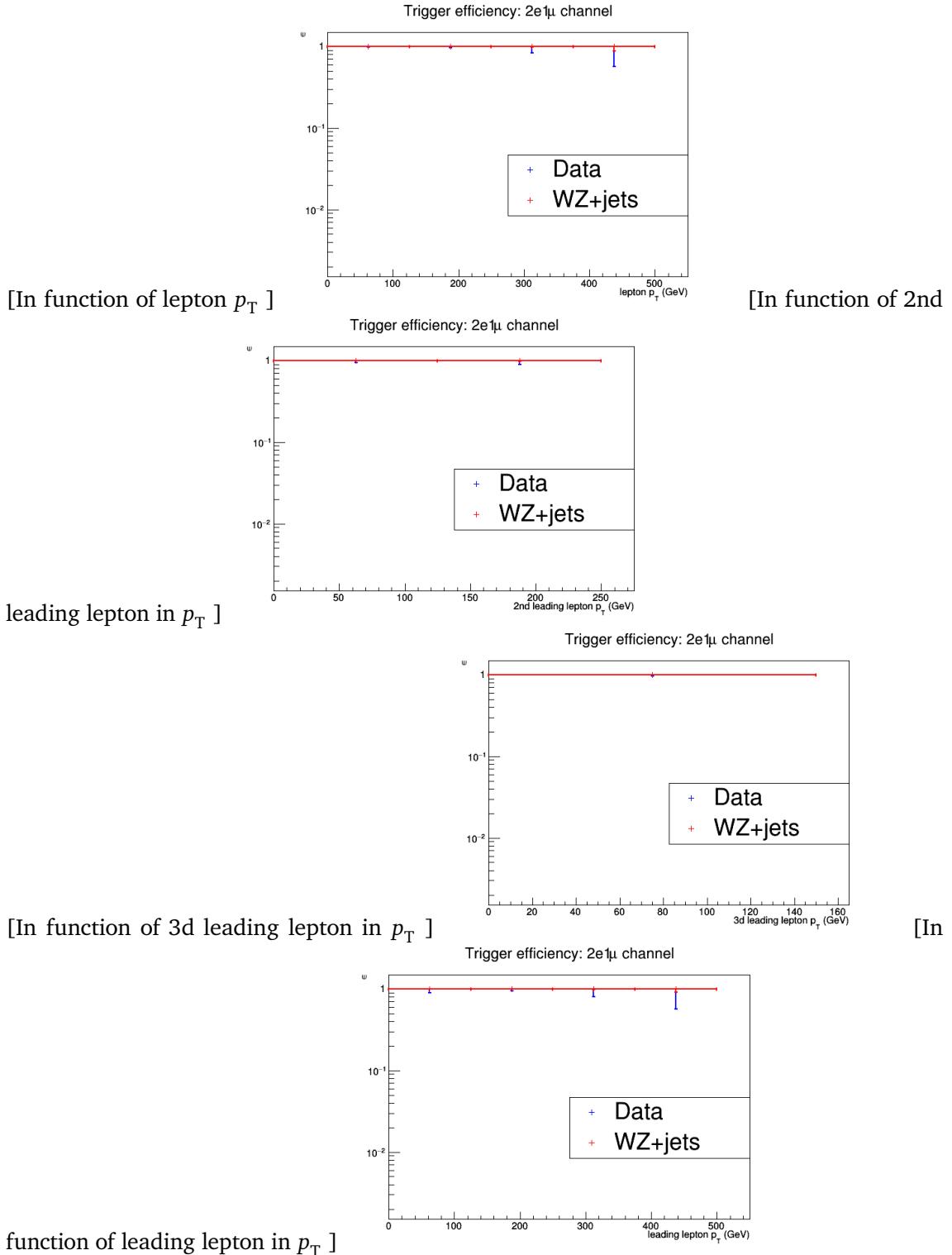


Figure A.2: The trigger efficiencies measured as a function of lepton p_T , using the dataset collected by E_T^{miss} triggers and WZ simulation, after a 3 lepton and jets selection, in the Z mass window. All corrections to simulation are applied. 2e1μ channel.

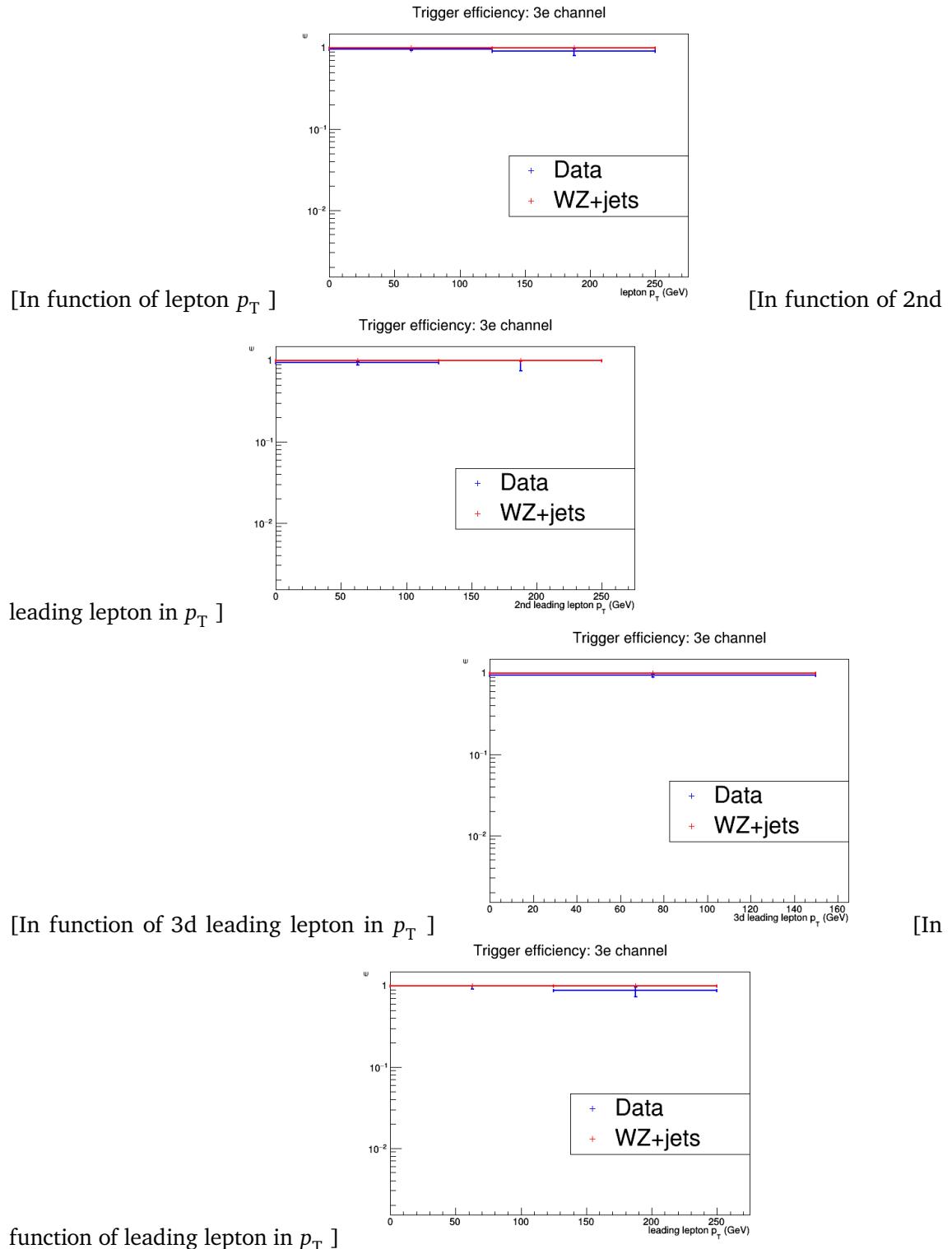


Figure A.3: The trigger efficiencies measured as a function of lepton p_T , using the dataset collected by E_T^{miss} triggers and WZ simulation, after a 3 lepton and jets selection, in the Z mass window. All corrections to simulation are applied. 3e channel.

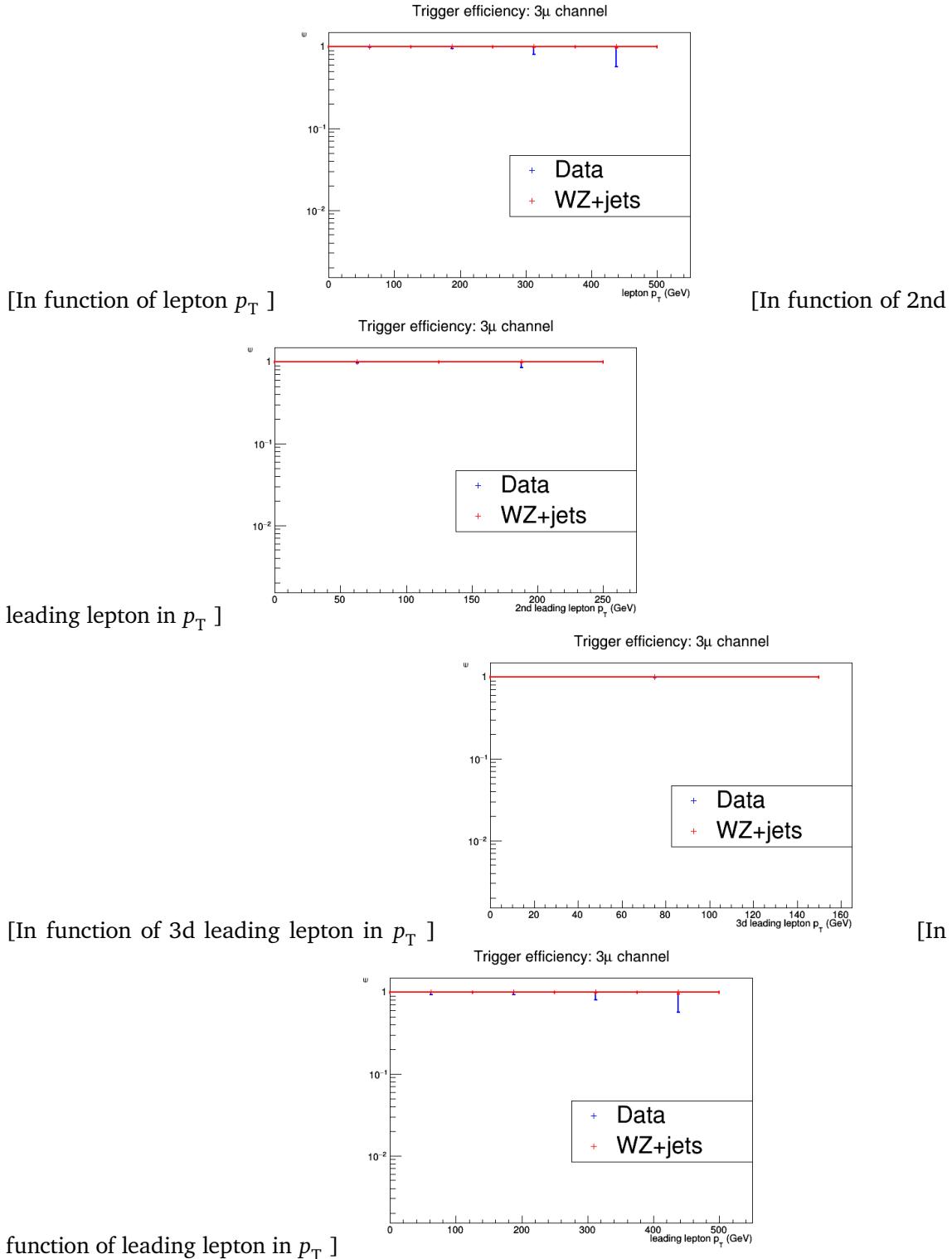


Figure A.4: The trigger efficiencies measured as a function of lepton p_T , using the dataset collected by E_T^{miss} triggers and WZ simulation, after a 3 lepton and jets selection, in the Z mass window. All corrections to simulation are applied. 3μ channel.

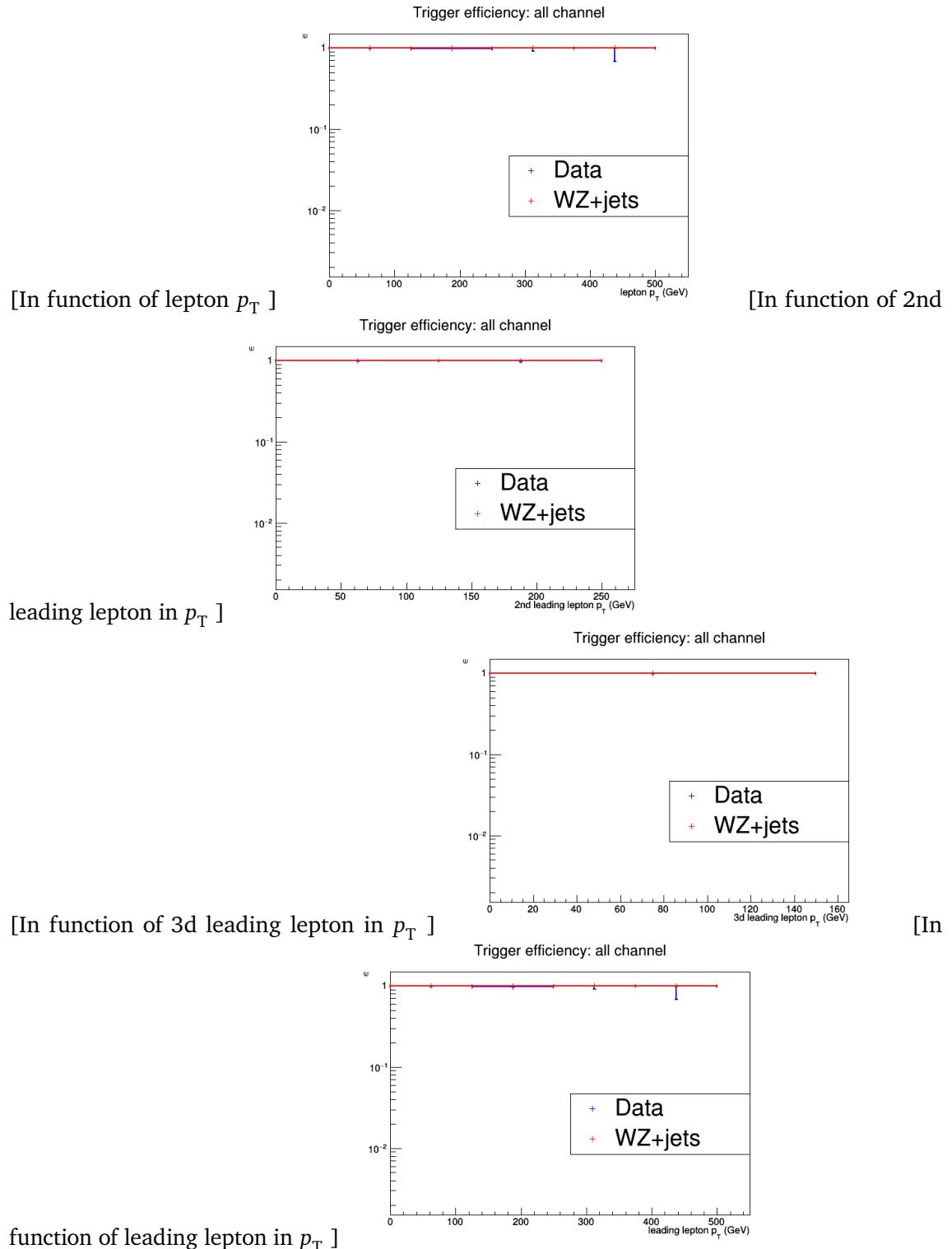


Figure A.5: The trigger efficiencies measured as a function of lepton p_T , using the dataset collected by E_T^{miss} triggers and WZ simulation, after a 3 lepton and jets selection, in the Z mass window. All corrections to simulation are applied. All channel.

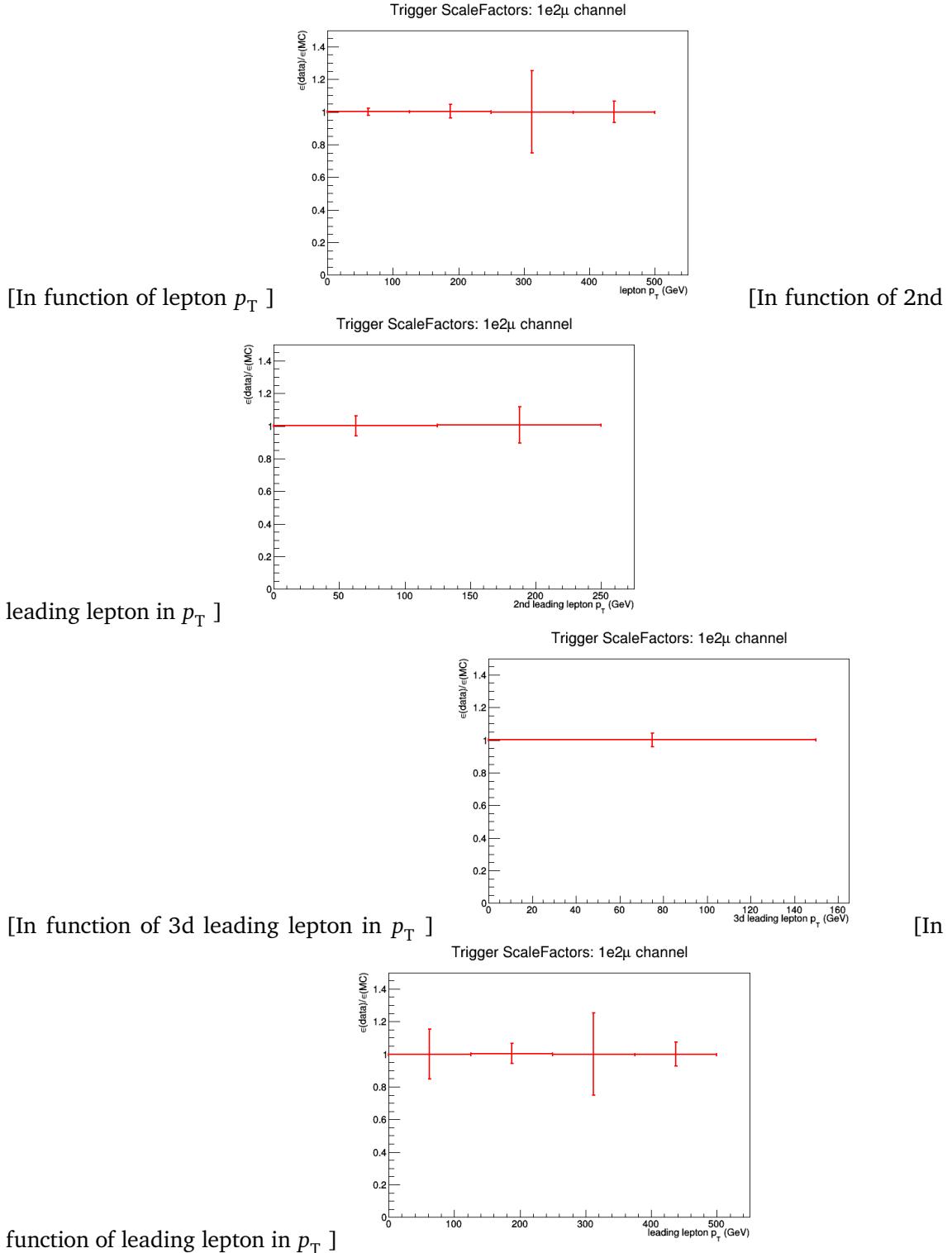


Figure A.6: The trigger scale factors measured as a function of lepton p_T , using the dataset collected by E_T^{miss} triggers and WZ simulation, after a 3 lepton and jets selection, in the Z mass window. All corrections to simulation are applied. 1e2μ channel.

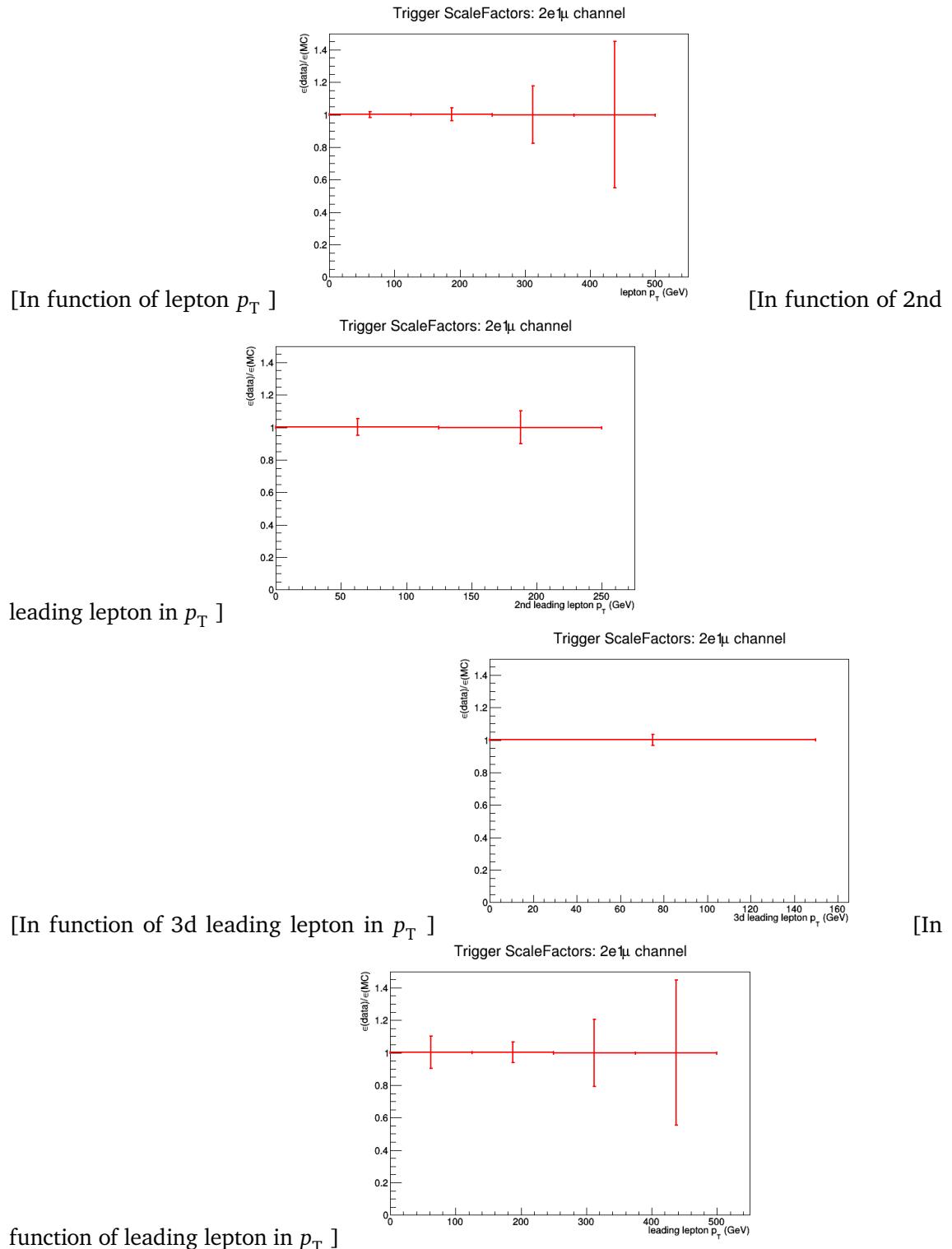


Figure A.7: The trigger scale factors measured as a function of lepton p_T , using the dataset collected by E_T^{miss} triggers and WZ simulation, after a 3 lepton and jets selection, in the Z mass window. All corrections to simulation are applied. 2e1μ channel.

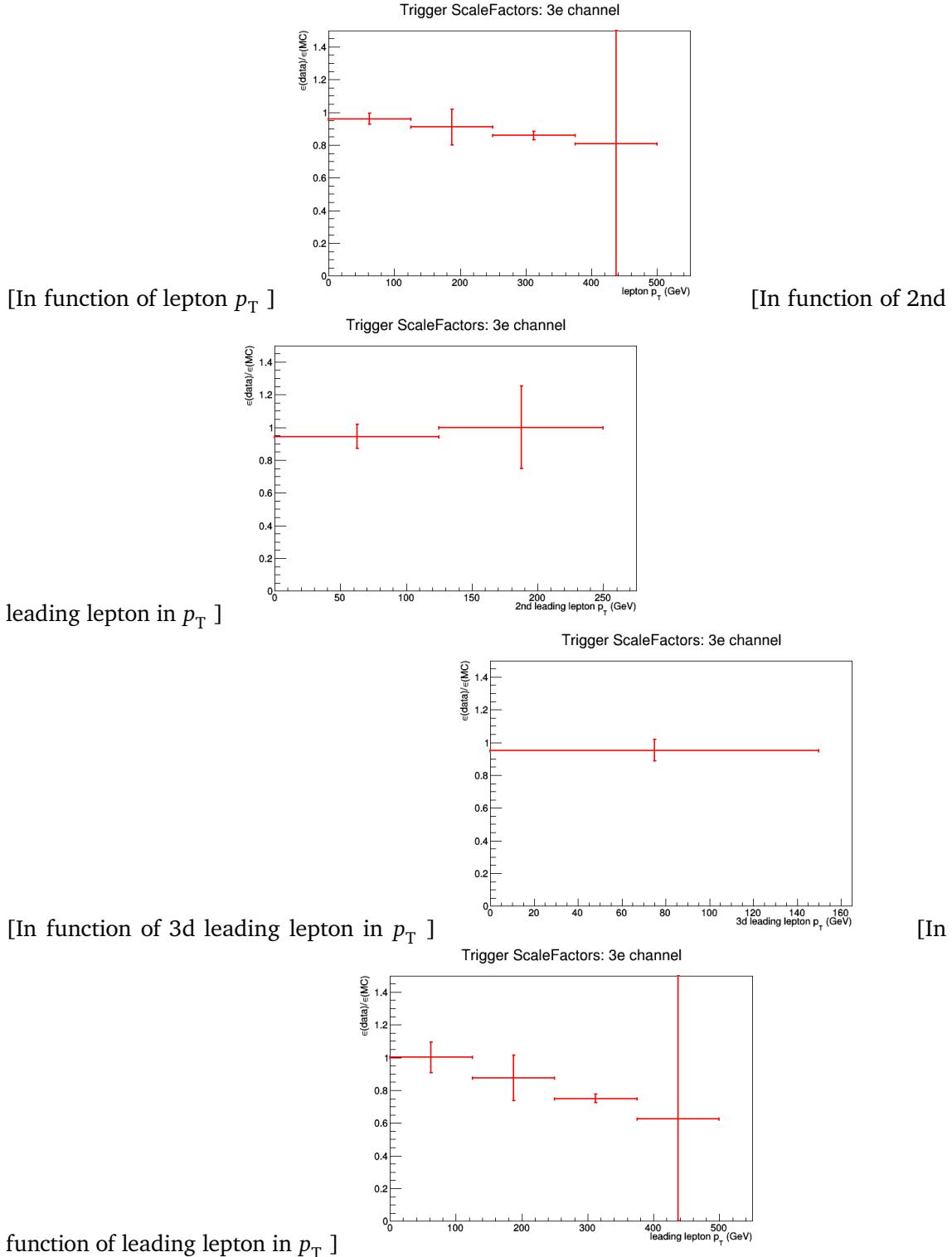


Figure A.8: The trigger scale factors measured as a function of lepton p_T , using the dataset collected by E_T^{miss} triggers and WZ simulation, after a 3 lepton and jets selection, in the Z mass window. All corrections to simulation are applied. 3e channel.

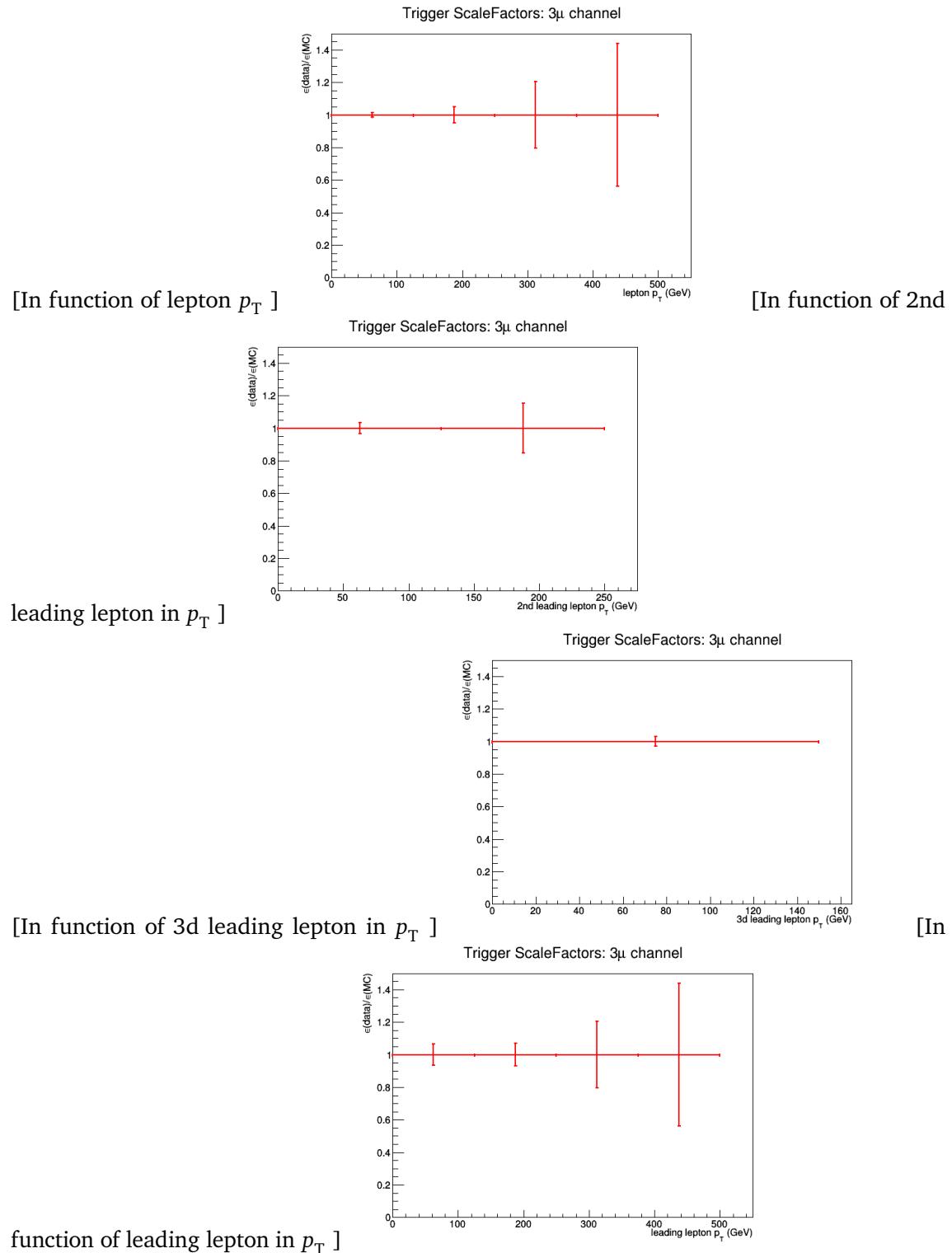


Figure A.9: The trigger scale factors measured as a function of lepton p_T , using the dataset collected by E_T^{miss} triggers and WZ simulation, after a 3 lepton and jets selection, in the Z mass window. All corrections to simulation are applied. 3μ channel.

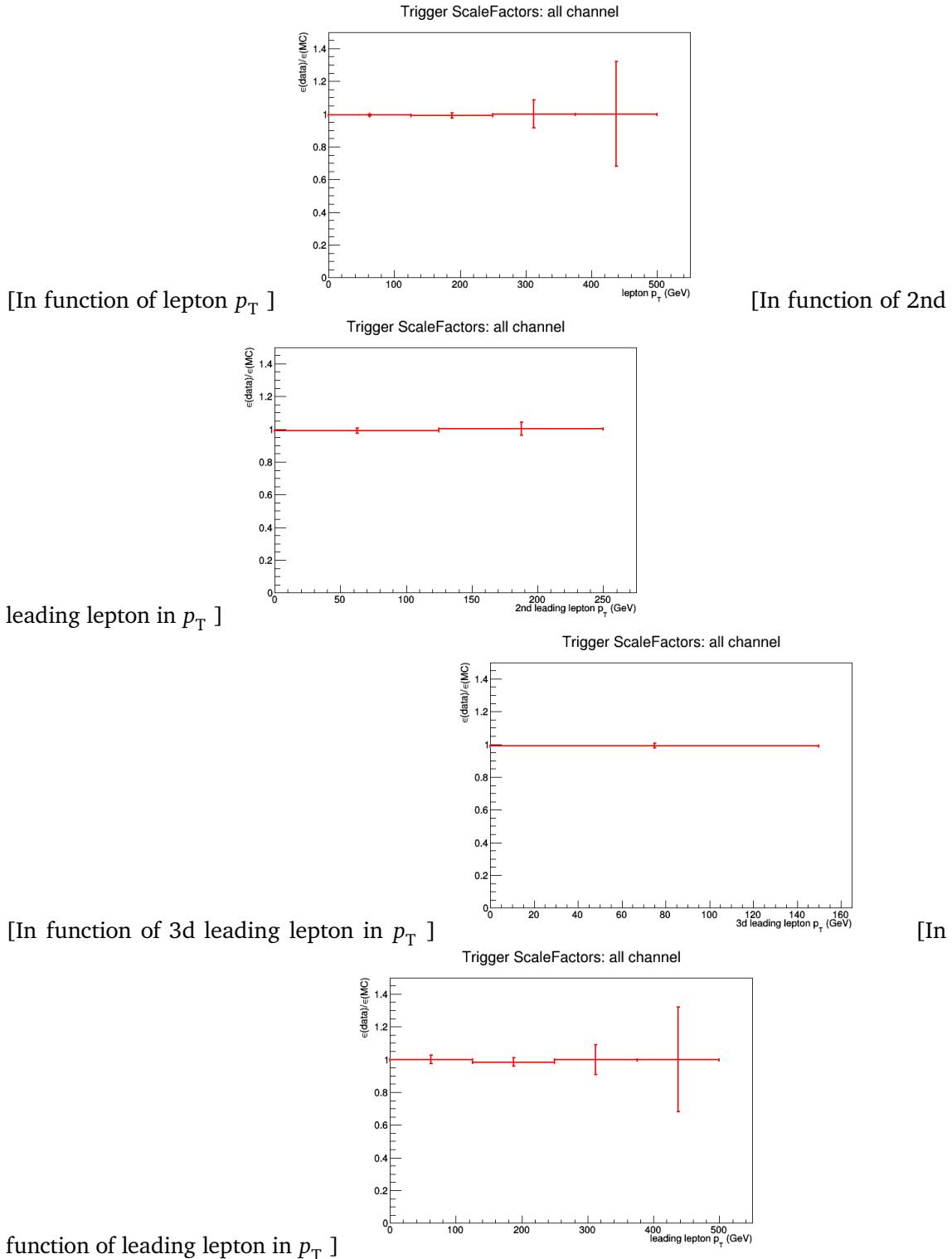


Figure A.10: The trigger scale factors measured as a function of lepton p_T , using the dataset collected by E_T^{miss} triggers and WZ simulation, after a 3 lepton and jets selection, in the Z mass window. All corrections to simulation are applied. All channel.

Dilepton controlplots

B

Statistical independent regions

1740

C

Bibliography

- [1] MICHAEL E PESKIN and DANIEL V SCHROEDER: **An introduction to quantum field theory; 1995 ed.** Includes exercises. Boulder, CO: Westview, 1995. URL: <https://cds.cern.ch/record/257493> (see pp. 1, 5, 11, 16).
- [2] C. P. BURGESS: **Introduction to Effective Field Theory.** In: *Ann. Rev. Nucl. Part. Sci.*, **57**: (2007), pp. 329–362. DOI: [10.1146/annurev.nucl.56.080805.140508](https://doi.org/10.1146/annurev.nucl.56.080805.140508). arXiv: [hep-th/0701053 \[hep-th\]](https://arxiv.org/abs/hep-th/0701053) (see p. 1).
- [3] NADIA ROBOTTI: **The discovery of the electron: I.** In: *European Journal of Physics*, **18**:3 (1997), p. 133. URL: <http://stacks.iop.org/0143-0807/18/i=3/a=002> (see p. 2).
- [4] Y. FUKUDA et al.: **Evidence for oscillation of atmospheric neutrinos.** In: *Phys. Rev. Lett.*, **81**: (1998), pp. 1562–1567. DOI: [10.1103/PhysRevLett.81.1562](https://doi.org/10.1103/PhysRevLett.81.1562). arXiv: [hep-ex/9807003 \[hep-ex\]](https://arxiv.org/abs/hep-ex/9807003) (see pp. 2, 14).
- [5] Y. ABE et al.: **Indication of Reactor $\bar{\nu}_e$ Disappearance in the Double Chooz Experiment.** In: *Phys. Rev. Lett.*, **108**: (13 Mar. 2012), p. 131801. DOI: [10.1103/PhysRevLett.108.131801](https://doi.org/10.1103/PhysRevLett.108.131801) (see pp. 2, 14).
- [6] C. PATRIGNANI et al.: **Review of Particle Physics.** In: *Chin. Phys.*, **C40**:10 (2016), p. 100001. DOI: [10.1088/1674-1137/40/10/100001](https://doi.org/10.1088/1674-1137/40/10/100001) (see pp. 2–3, 6–8, 10, 42–43, 75).
- [7] S. ABACHI et al.: **Observation of the top quark.** In: *Phys. Rev. Lett.*, **74**: (1995), pp. 2632–2637. DOI: [10.1103/PhysRevLett.74.2632](https://doi.org/10.1103/PhysRevLett.74.2632). arXiv: [hep-ex/9503003 \[hep-ex\]](https://arxiv.org/abs/hep-ex/9503003) (see p. 2).
- [8] F. ABE et al.: **Observation of top quark production in $\bar{p}p$ collisions.** In: *Phys. Rev. Lett.*, **74**: (1995), pp. 2626–2631. DOI: [10.1103/PhysRevLett.74.2626](https://doi.org/10.1103/PhysRevLett.74.2626). arXiv: [hep-ex/9503002 \[hep-ex\]](https://arxiv.org/abs/hep-ex/9503002) (see p. 2).
- [9] First combination of Tevatron and LHC measurements of the top-quark mass. In: (2014). arXiv: [1403.4427 \[hep-ex\]](https://arxiv.org/abs/1403.4427) (see p. 2).
- [10] SERGUEI CHATRCHYAN et al.: **Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC.** In: *Phys. Lett.*, **B716**: (2012), pp. 30–61. DOI: [10.1016/j.physletb.2012.08.021](https://doi.org/10.1016/j.physletb.2012.08.021). arXiv: [1207.7235 \[hep-ex\]](https://arxiv.org/abs/1207.7235) (see pp. 2, 23).
- [11] GEORGES AAD et al.: **Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC.** In: *Phys. Lett.*, **B716**: (2012), pp. 1–29. DOI: [10.1016/j.physletb.2012.08.020](https://doi.org/10.1016/j.physletb.2012.08.020). arXiv: [1207.7214 \[hep-ex\]](https://arxiv.org/abs/1207.7214) (see pp. 2, 23).

- 1773 [12] NICOLA CABIBBO: **Unitary Symmetry and Leptonic Decays**. In: *Phys. Rev. Lett.*, **10**:
 1774 (12 June 1963), pp. 531–533. doi: [10.1103/PhysRevLett.10.531](https://doi.org/10.1103/PhysRevLett.10.531) (see p. 5).
- 1775 [13] S. L. GLASHOW, J. ILIOPoulos, and L. MAIANI: **Weak Interactions with Lepton-Hadron
 1776 Symmetry**. In: *Phys. Rev. D*, **2**: (7 Oct. 1970), pp. 1285–1292. doi: [10.1103/PhysRevD.2.1285](https://doi.org/10.1103/PhysRevD.2.1285) (see p. 5).
- 1778 [14] B.J. BJØRKEN and S.L. GLASHOW: **Elementary particles and SU(4)**. In: *Physics Letters*,
 1779 **11**:3 (1964), pp. 255–257. doi: [https://doi.org/10.1016/0031-9163\(64\)90433-0](https://doi.org/10.1016/0031-9163(64)90433-0) (see
 1780 p. 5).
- 1781 [15] LUCIANO MAIANI: **The GIM Mechanism: origin, predictions and recent uses**. In:
 1782 *Proceedings, 48th Rencontres de Moriond on Electroweak Interactions and Unified Theories:
 1783 La Thuile, Italy, March 2-9, 2013*. 2013, pp. 3–16. arXiv: [1303.6154 \[hep-ph\]](https://arxiv.org/abs/1303.6154). URL:
 1784 <https://inspirehep.net/record/1225307/files/arXiv:1303.6154.pdf> (see p. 5).
- 1785 [16] PATRICK KOPPENBURG and SEBASTIEN DESCOTES-GENON: **The CKM Parameters**. In:
 1786 (2017). arXiv: [1702.08834 \[hep-ex\]](https://arxiv.org/abs/1702.08834) (see p. 6).
- 1787 [17] J. A. AGUILAR-SAAVEDRA: **Top flavor-changing neutral interactions: Theoretical
 1788 expectations and experimental detection**. In: *Acta Phys. Polon.*, **B35**: (2004),
 1789 pp. 2695–2710. arXiv: [hep-ph/0409342 \[hep-ph\]](https://arxiv.org/abs/hep-ph/0409342) (see pp. 6, 15–17).
- 1790 [18] S. ABACHI et al.: **Observation of the top quark**. In: *Phys. Rev. Lett.*, **74**: (1995),
 1791 pp. 2632–2637. doi: [10.1103/PhysRevLett.74.2632](https://doi.org/10.1103/PhysRevLett.74.2632). arXiv: [hep-ex/9503003 \[hep-ex\]](https://arxiv.org/abs/hep-ex/9503003)
 1792 (see p. 6).
- 1793 [19] F. ABE et al.: **Observation of top quark production in $\bar{p}p$ collisions**. In: *Phys. Rev.
 1794 Lett.*, **74**: (1995), pp. 2626–2631. doi: [10.1103/PhysRevLett.74.2626](https://doi.org/10.1103/PhysRevLett.74.2626). arXiv: [hep-ex/9503002 \[hep-ex\]](https://arxiv.org/abs/hep-ex/9503002) (see p. 6).
- 1796 [20] ANDREA GIAMMANCO and JEANNINE WAGNER-KUHR: **Measurement of the t-channel
 1797 single Top-quark production rates in pp collisions at 7 TeV**. 2011. URL: <http://cms.web.cern.ch/news/measurement-t-channel-single-top-quark-production-rates-pp-collisions-7-tev> (see p. 9).
- 1800 [21] LHCTOP WORKING GROUP: **ATLAS-CMS recommended predictions for single top
 1801 cross sections using the Hathor v2.1 program**. 2017. URL: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SingleTopRefXsec#Predictions_at_7_8_13_and_14_TeV
 1803 (see p. 10).
- 1804 [22] CMS COLLABORATION: **Summaries of CMS cross section measurements**. 2017. URL:
 1805 <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsCombined> (see
 1806 pp. 10, 20, 22, 50).
- 1807 [23] CÉLINE DEGRANDE: **Effective field theories in the Standard Model and beyond**. Thesis.
 1808 Université Catholique de Louvain. URL: <https://cp3.irmp.ucl.ac.be/upload/theses/phd/degrande.pdf> (see p. 11).
- 1810 [24] ENRICO FERMI: **Tentativo di una Teoria Dei Raggi β** . In: *Il Nuovo Cimento* (1924–
 1811 1942), **11**:1 (Sept. 20, 2008), p. 1. doi: [10.1007/BF02959820](https://doi.org/10.1007/BF02959820) (see p. 11).
- 1812 [25] B. GRZADKOWSKI, M. ISKRZYNSKI, M. MISIAK, and J. ROSIEK: **Dimension-Six Terms
 1813 in the Standard Model Lagrangian**. In: *JHEP*, **1010**: (2010), p. 085. doi: [10.1007/JHEP10\(2010\)085](https://doi.org/10.1007/JHEP10(2010)085). arXiv: [1008.4884 \[hep-ph\]](https://arxiv.org/abs/1008.4884) (see pp. 12–13, 16).

- 1815 [26] ANDY BUCKLEY, CHRISTOPH ENGLERT, JAMES FERRANDO, et al.: **Constraining top**
 1816 **quark effective theory in the LHC Run II era.** In: *JHEP*, **04**: (2016), p. 015. doi:
 1817 [10.1007/JHEP04\(2016\)015](https://doi.org/10.1007/JHEP04(2016)015). arXiv: [1512.03360 \[hep-ph\]](https://arxiv.org/abs/1512.03360) (see p. 13).
- 1818 [27] YORIKIYO NAGASHIMA: **Beyond the standard model of elementary particle physics.**
 1819 Weinheim, Germany: Wiley-VCH Verlag, 2014. url: [http://www-spires.fnal.gov/
 1820 spires/find/books/www?cl=QC793.2.N34::2014](http://www-spires.fnal.gov/spires/find/books/www?cl=QC793.2.N34::2014) (see p. 14).
- 1821 [28] P. A. R. ADE et al.: **Planck 2015 results. XIII. Cosmological parameters.** In: *Astron.*
 1822 *Astrophys.*, **594**: (2016), A13. doi: [10.1051/0004-6361/201525830](https://doi.org/10.1051/0004-6361/201525830). arXiv: [1502.01589
 1823 \[astro-ph.CO\]](https://arxiv.org/abs/1502.01589) (see p. 14).
- 1824 [29] P. J. E. PEEBLES and BHARAT RATRA: **The Cosmological constant and dark energy.**
 1825 In: *Rev. Mod. Phys.*, **75**: (2003), pp. 559–606. doi: [10.1103/RevModPhys.75.559](https://doi.org/10.1103/RevModPhys.75.559). arXiv:
 1826 [astro-ph/0207347 \[astro-ph\]](https://arxiv.org/abs/astro-ph/0207347) (see p. 14).
- 1827 [30] A. D. SAKHAROV: **Violation of CP Invariance, c Asymmetry, and Baryon Asym-**
 1828 **metry of the Universe.** In: *Pisma Zh. Eksp. Teor. Fiz.*, **5**: (1967). [*Usp. Fiz.*
 1829 *Nauk* **161**, 61 (1991)], pp. 32–35. doi: [10.1070/PU1991v034n05ABEH002497](https://doi.org/10.1070/PU1991v034n05ABEH002497) (see p. 14).
- 1830 [31] GUSTAVO BURDMAN: **New solutions to the hierarchy problem.** In: *Braz. J. Phys.*, **37**:
 1831 (2007), pp. 506–513. doi: [10.1590/S0103-97332007000400006](https://doi.org/10.1590/S0103-97332007000400006). arXiv: [hep-ph/0703194
 1832 \[hep-ph\]](https://arxiv.org/abs/hep-ph/0703194) (see p. 14).
- 1833 [32] T. ET AL. AALTONEN: **Search for the Flavor-Changing Neutral-Current Decay $t \rightarrow Zq$**
 1834 **in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV.** In: *Phys. Rev. Lett.*, **101**: (19 Nov. 2008),
 1835 p. 192002. doi: [10.1103/PhysRevLett.101.192002](https://doi.org/10.1103/PhysRevLett.101.192002) (see p. 15).
- 1836 [33] VICTOR MUKHAMEDOVICH ABAZOV et al.: **Search for flavour changing neutral cur-**
 1837 **rents via quark-gluon couplings in single top quark production using 2.3 fb^{-1} of**
 1838 **$p\bar{p}$ collisions.** In: *Phys. Lett.*, **B693**: (2010), pp. 81–87. doi: [10.1016/j.physletb.2010.08.011](https://doi.org/10.1016/j.physletb.2010.08.011). arXiv: [1006.3575 \[hep-ex\]](https://arxiv.org/abs/1006.3575) (see p. 15).
- 1840 [34] GEORGES AAD et al.: **Search for flavour-changing neutral current top-quark decays**
 1841 **to qZ in pp collision data collected with the ATLAS detector at $\sqrt{s} = 8$ TeV.** In: *Eur. Phys. J.*, **C76**:1 (2016), p. 12. doi: [10.1140/epjc/s10052-015-3851-5](https://doi.org/10.1140/epjc/s10052-015-3851-5). arXiv:
 1842 [1508.05796 \[hep-ex\]](https://arxiv.org/abs/1508.05796) (see p. 15).
- 1844 [35] GEORGES AAD et al.: **Search for single top-quark production via flavour-changing**
 1845 **neutral currents at 8 TeV with the ATLAS detector.** In: *Eur. Phys. J.*, **C76**:2 (2016),
 1846 p. 55. doi: [10.1140/epjc/s10052-016-3876-4](https://doi.org/10.1140/epjc/s10052-016-3876-4). arXiv: [1509.00294 \[hep-ex\]](https://arxiv.org/abs/1509.00294) (see pp. 15,
 1847 19).
- 1848 [36] GEORGES AAD et al.: **Search for flavour-changing neutral current top quark decays**
 1849 **$t \rightarrow Hq$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector.** In: *JHEP*, **12**:
 1850 (2015), p. 061. doi: [10.1007/JHEP12\(2015\)061](https://doi.org/10.1007/JHEP12(2015)061). arXiv: [1509.06047 \[hep-ex\]](https://arxiv.org/abs/1509.06047) (see
 1851 p. 15).
- 1852 [37] MORAD AABOUD et al.: **Search for top quark decays $t \rightarrow qH$, with $H \rightarrow \gamma\gamma$, in**
 1853 **$\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector.** In: (2017). arXiv: [1707.01404
 1854 \[hep-ex\]](https://arxiv.org/abs/1707.01404) (see pp. 15, 19).

- [38] **Search for flavour-changing neutral current top quark decays $t \rightarrow qZ$ in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector.** Tech. rep. ATLAS-CONF-2017-070. Geneva: CERN, Sept. 2017. URL: <https://cds.cern.ch/record/2285808> (see pp. 15, 19).
- [39] ALBERT M SIRUNYAN et al.: **Search for associated production of a Z boson with a single top quark and for tZ flavour-changing interactions in pp collisions at $\sqrt{s} = 8$ TeV.** In: (2017). arXiv: [1702.01404 \[hep-ex\]](https://arxiv.org/abs/1702.01404) (see pp. 15, 19).
- [40] SERGUEI CHATRCHYAN et al.: **Search for Flavor-Changing Neutral Currents in Top-Quark Decays $t \rightarrow Zq$ in pp Collisions at $\sqrt{s} = 8$ TeV.** In: *Phys. Rev. Lett.*, **112**:17 (2014), p. 171802. doi: [10.1103/PhysRevLett.112.171802](https://doi.org/10.1103/PhysRevLett.112.171802). arXiv: [1312.4194 \[hep-ex\]](https://arxiv.org/abs/1312.4194) (see p. 15).
- [41] VARDAN KHACHATRYAN et al.: **Search for anomalous single top quark production in association with a photon in pp collisions at $\sqrt{s} = 8$ TeV.** In: *JHEP*, **04**: (2016), p. 035. doi: [10.1007/JHEP04\(2016\)035](https://doi.org/10.1007/JHEP04(2016)035). arXiv: [1511.03951 \[hep-ex\]](https://arxiv.org/abs/1511.03951) (see pp. 15, 19).
- [42] VARDAN KHACHATRYAN et al.: **Search for top quark decays via Higgs-boson-mediated flavor-changing neutral currents in pp collisions at $\sqrt{s} = 8$ TeV.** In: *JHEP*, **02**: (2017), p. 079. doi: [10.1007/JHEP02\(2017\)079](https://doi.org/10.1007/JHEP02(2017)079). arXiv: [1610.04857 \[hep-ex\]](https://arxiv.org/abs/1610.04857) (see p. 15).
- [43] **Search for the flavor-changing interactions of the top quark with the Higgs boson in $H \rightarrow b\bar{b}$ channel at $\sqrt{s} = 13$ TeV.** Tech. rep. CMS-PAS-TOP-17-003. Geneva: CERN, 2017. URL: <https://cds.cern.ch/record/2284743> (see p. 15).
- [44] M. BENEKE et al.: **Top quark physics.** In: *1999 CERN Workshop on standard model physics (and more) at the LHC, CERN, Geneva, Switzerland, 25-26 May: Proceedings*. 2000, pp. 419–529. arXiv: [hep-ph/0003033 \[hep-ph\]](https://arxiv.org/abs/hep-ph/0003033). URL: <http://weblib.cern.ch/abstract?CERN-TH-2000-100> (see p. 17).
- [45] JUN GAO, CHONG SHENG LI, and HUA XING ZHU: **Top Quark Decay at Next-to-Next-to Leading Order in QCD.** In: *Phys. Rev. Lett.*, **110**:4 (2013), p. 042001. doi: [10.1103/PhysRevLett.110.042001](https://doi.org/10.1103/PhysRevLett.110.042001). arXiv: [1210.2808 \[hep-ph\]](https://arxiv.org/abs/1210.2808) (see p. 18).
- [46] LHCTOP WORKING GROUP: **LHCTopWG Summary plots.** 2017. URL: <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCTopWGSummaryPlots> (see p. 21).
- [47] STEPHEN MYERS: **The LEP Collider, from design to approval and commissioning.** John Adams' Lecture. Delivered at CERN, 26 Nov 1990. Geneva: CERN, 1991. URL: [http://cds.cern.ch/record/226776](https://cds.cern.ch/record/226776) (see p. 23).
- [48] STEPHEN HOLMES, RONALD S MOORE, and VLADIMIR SHILTSEV: **Overview of the Tevatron collider complex: goals, operations and performance.** In: *Journal of Instrumentation*, **6**:08 (2011), T08001. URL: <http://stacks.iop.org/1748-0221/6/i=08/a=T08001> (see p. 23).
- [49] R. BARATE et al.: **Search for the standard model Higgs boson at LEP.** In: *Phys. Lett.*, **B565**: (2003), pp. 61–75. doi: [10.1016/S0370-2693\(03\)00614-2](https://doi.org/10.1016/S0370-2693(03)00614-2). arXiv: [hep-ex/0306033 \[hep-ex\]](https://arxiv.org/abs/hep-ex/0306033) (see p. 23).
- [50] KENNETH HERNER: **Higgs Boson Studies at the Tevatron.** In: *Nucl. Part. Phys. Proc.*, **273-275**: (2016), pp. 852–856. doi: [10.1016/j.nuclphysbps.2015.09.131](https://doi.org/10.1016/j.nuclphysbps.2015.09.131) (see p. 23).

- 1897 [51] ABDELHAK DJOUADI: **The Anatomy of electro-weak symmetry breaking. I: The**
 1898 **Higgs boson in the standard model.** In: *Phys. Rept.*, **457**: (2008), pp. 1–216. DOI:
 1899 [10.1016/j.physrep.2007.10.004](https://doi.org/10.1016/j.physrep.2007.10.004). arXiv: [hep-ph/0503172 \[hep-ph\]](https://arxiv.org/abs/hep-ph/0503172) (see p. 23).
- 1900 [52] LYNDON EVANS and PHILIP BRYANT: **LHC Machine.** In: *Journal of Instrumentation*,
 1901 **3**:08 (2008), S08001. URL: <http://stacks.iop.org/1748-0221/3/i=08/a=S08001> (see
 1902 p. 23).
- 1903 [53] THOMAS SVEN PETTERSSON and P LEFÈVRE: **The Large Hadron Collider: conceptual**
 1904 **design.** Tech. rep. CERN-AC-95-05-LHC. Oct. 1995, p. 20 and 22. URL: <https://cds.cern.ch/record/291782> (see p. 23).
- 1906 [54] JORG WENNINGER and EZIO TODESCO: **Large Hadron Collider momentum calibration**
 1907 **and accuracy.** Tech. rep. CERN-ACC-2017-0007. Geneva: CERN, Feb. 2017. URL:
 1908 <https://cds.cern.ch/record/2254678> (see p. 24).
- 1909 [55] CINZIA DE MELIS: **The CERN accelerator complex. Complexe des accélérateurs du**
 1910 **CERN.** In: (July 2016). General Photo. URL: <https://cds.cern.ch/record/2197559>
 1911 (see p. 24).
- 1912 [56] OLIVER SIM BRÜNING, PAUL COLLIER, P LEBRUN, et al.: **LHC Design Report.** CERN
 1913 Yellow Reports: Monographs. Geneva: CERN, 2004. URL: <https://cds.cern.ch/record/782076> (see p. 24).
- 1915 [57] G. AAD et al.: **The ATLAS Experiment at the CERN Large Hadron Collider.** In: *JINST*,
 1916 **3**: (2008), S08003. DOI: [10.1088/1748-0221/3/08/S08003](https://doi.org/10.1088/1748-0221/3/08/S08003) (see p. 25).
- 1917 [58] S. CHATRHYAN et al.: **The CMS Experiment at the CERN LHC.** In: *JINST*, **3**: (2008),
 1918 S08004. DOI: [10.1088/1748-0221/3/08/S08004](https://doi.org/10.1088/1748-0221/3/08/S08004) (see pp. 25, 32–34).
- 1919 [59] K. AAMODT et al.: **The ALICE experiment at the CERN LHC.** In: *JINST*, **3**: (2008),
 1920 S08002. DOI: [10.1088/1748-0221/3/08/S08002](https://doi.org/10.1088/1748-0221/3/08/S08002) (see p. 25).
- 1921 [60] A. AUGUSTO ALVES JR. et al.: **The LHCb Detector at the LHC.** In: *JINST*, **3**: (2008),
 1922 S08005. DOI: [10.1088/1748-0221/3/08/S08005](https://doi.org/10.1088/1748-0221/3/08/S08005) (see p. 25).
- 1923 [61] M. BONGI et al.: **Astroparticle physics at LHC: The LHCf experiment ready for data**
 1924 **taking.** In: *Nucl. Instrum. Meth.*, **A612**: (2010), pp. 451–454. DOI: [10.1016/j.nima.2009.08.039](https://doi.org/10.1016/j.nima.2009.08.039) (see p. 25).
- 1926 [62] G. ANELLI et al.: **The TOTEM experiment at the CERN Large Hadron Collider.** In: *JINST*, **3**: (2008),
 1927 S08007. DOI: [10.1088/1748-0221/3/08/S08007](https://doi.org/10.1088/1748-0221/3/08/S08007) (see p. 25).
- 1928 [63] B. ACHARYA et al.: **The Physics Programme Of The MoEDAL Experiment At The LHC.**
 1929 In: *Int. J. Mod. Phys.*, **A29**: (2014), p. 1430050. DOI: [10.1142/S0217751X14300506](https://doi.org/10.1142/S0217751X14300506).
 1930 arXiv: [1405.7662 \[hep-ph\]](https://arxiv.org/abs/1405.7662) (see p. 25).
- 1931 [64] BY JAMES GILLIES: **Luminosity? Why don't we just say collision rate?** In: (Mar.
 1932 2011). URL: [http://cds.cern.ch/record/1997001](https://cds.cern.ch/record/1997001) (see p. 25).
- 1933 [65] BY HARRIET JARLETT and HARRIET KIM JARLETT: **LHC pushes limits of performance.**
 1934 In: (Aug. 2016). URL: [http://cds.cern.ch/record/2212301](https://cds.cern.ch/record/2212301) (see p. 26).
- 1935 [66] CMS COLLABORATION. 2017. URL: https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults#Online_Luminosity_AN2 (see p. 26).

- 1937 [67] **Technical proposal.** LHC Tech. Proposal. Cover title : CMS, the Compact Muon
 1938 Solenoid : technical proposal. Geneva: CERN, 1994. URL: <https://cds.cern.ch/record/290969> (see p. 26).
- 1940 [68] G. L. BAYATIAN et al.: **CMS physics: Technical design report.** In: (2006) (see p. 26).
- 1941 [69] G L BAYATIAN, S CHATRCHYAN, G HMAYAKYAN, et al.: **CMS Physics: Technical Design**
 1942 **Report Volume 1: Detector Performance and Software.** Technical Design Report
 1943 CMS. There is an error on cover due to a technical problem for some items. Geneva:
 1944 CERN, 2006. URL: <https://cds.cern.ch/record/922757> (see pp. 26, 31, 60).
- 1945 [70] CMS COLLABORATION: **Detector Drawings.** CMS Collection. Mar. 2012. URL: <https://cds.cern.ch/record/1433717> (see p. 27).
- 1947 [71] S CHATRCHYAN, V KHACHATRYAN, A M SIRUNYAN, et al.: **Performance of the CMS Drift**
 1948 **Tube Chambers with Cosmic Rays.** In: *J. Instrum.*, 5:arXiv:0911.4855. CMS-CFT-
 1949 09-012 (Nov. 2009), T03015 . 47 p. URL: <http://cds.cern.ch/record/1223944> (see
 1950 pp. 29–30).
- 1951 [72] TOM DODINGTON: **News from the CMS experimental site: 22 November 2013.** 2013.
 1952 URL: <http://cms.web.cern.ch/news/news-point-5-22-november-2013> (see p. 30).
- 1953 [73] INSTITUTE OF RESEARCH INTO THE FUNDAMENTAL LAWS THE UNIVERSE: **The CMS**
 1954 **detector superconducting solenoid.** 2006. URL: http://irfu.cea.fr/en/Phoebe/Vie_des_labos/Ast/ast_visu.php?id_ast=839 (see p. 32).
- 1956 [74] FERGUS WILSON: **Experimental Particle Physics.** 2012. URL: <http://slideplayer.com/slide/794631/> (see p. 32).
- 1958 [75] EMRAH TIRAS, BURAK BILKI, and YASAR ONEL: **Commissioning of CMS Forward**
 1959 **Hadron Calorimeters with Upgraded Multi-anode PMTs and μTCA Readout.** In:
 1960 (2016). arXiv: 1611.05232 [physics.ins-det] (see p. 33).
- 1961 [76] LUCAS TAYLOR: **Experimental Particle Physics.** 2011. URL: <http://cms.web.cern.ch/news/electromagnetic-calorimeter> (see p. 33).
- 1963 [77] **Proceedings, 34th International Conference on High Energy Physics (ICHEP 2008)** ■
 1964 URL: <http://www.slac.stanford.edu/econf/C080730> (see p. 33).
- 1965 [78] L. BRIANZA: **Precision crystal calorimetry in LHC Run II with the CMS ECAL.** In:
 1966 *Journal of Instrumentation*, 12:01 (2017), p. C01069. URL: <http://stacks.iop.org/1748-0221/12/i=01/a=C01069> (see pp. 34, 37).
- 1968 [79] SERGUEI CHATRCHYAN, VARDAN KHACHATRYAN, ALBERT M SIRUNYAN, et al.: **Description**
 1969 **and performance of track and primary-vertex reconstruction with the CMS**
 1970 **tracker.** In: *J. Instrum.*, 9:arXiv:1405.6569. CERN-PH-EP-2014-070. CMS-TRK-11-001
 1971 (May 2014). Comments: Replaced with published version. Added journal reference
 1972 and DOI, P10009. 80 p. URL: <http://cds.cern.ch/record/1704291> (see pp. 34, 63).
- 1973 [80] BY CHRISTINE SUTTON: **Chronicles of CMS: the saga of LS1.** In: (May 2015). URL:
 1974 <http://cds.cern.ch/record/2024986> (see p. 35).
- 1975 [81] **A beautiful barrel for CMS.** In: (Mar. 2014). URL: <http://cds.cern.ch/record/1998635> (see p. 35).

- 1977 [82] VARDAN KHACHATRYAN et al.: **The CMS trigger system**. In: *JINST*, **12**:01 (2017),
1978 P01020. doi: [10.1088/1748-0221/12/01/P01020](https://doi.org/10.1088/1748-0221/12/01/P01020). arXiv: [1609.02366 \[physics.ins-det\]](https://arxiv.org/abs/1609.02366) (see p. 36).
- 1980 [83] BY CORINNE PRALAVORIO and CORINNE PRALAVORIO: **Major work to ready the LHC
1981 experiments for Run 2**. In: (May 2015). URL: <http://cds.cern.ch/record/2024977> (see p. 36).
- 1983 [84] LUIGI GUIDUCCI: **CMS muon system towards LHC Run 2 and beyond**. Tech. rep.
1984 CMS-CR-2014-333. Geneva: CERN, Oct. 2014. URL: <https://cds.cern.ch/record/1966038> (see p. 37).
- 1986 [85] CARLO BATTILANA: **The CMS muon system status and upgrades for LHC run-2 and
1987 performance of muon reconstruction with 13 TeV data**. Tech. rep. CMS-CR-2016-
1988 437. Geneva: CERN, Dec. 2016. URL: <http://cds.cern.ch/record/2239185> (see
1989 p. 37).
- 1990 [86] SERGUEI CHATRCHYAN et al.: **Energy Calibration and Resolution of the CMS Elec-
1991 tronagnetic Calorimeter in pp Collisions at $\sqrt{s} = 7$ TeV**. In: *JINST*, **8**: (2013).
1992 [JINST8,9009(2013)], P09009. doi: [10.1088/1748-0221/8/09/P09009](https://doi.org/10.1088/1748-0221/8/09/P09009). arXiv: [1306.2016 \[hep-ex\]](https://arxiv.org/abs/1306.2016) (see p. 37).
- 1994 [87] **Cool running for CMS tracker**. In: (Mar. 2014). URL: <http://cds.cern.ch/record/1998606> (see p. 37).
- 1996 [88] L. CADAMURO: **The CMS Level-1 trigger system for LHC Run II**. In: *Journal of
1997 Instrumentation*, **12**:03 (2017), p. C03021. URL: <http://stacks.iop.org/1748-0221/12/i=03/a=C03021> (see p. 37).
- 1999 [89] DAVID AARON MATZNER DOMINGUEZ, D. ABBANEO, K. ARNDT, et al.: **CMS Technical
2000 Design Report for the Pixel Detector Upgrade**. In: (2012) (see p. 38).
- 2001 [90] HANNO CHRISTOPHER PERREY: **Plans and Status of the Phase I Upgrade of the
2002 CMS Pixel Tracker**. Tech. rep. CMS-CR-2014-005. Geneva: CERN, Jan. 2014. URL:
2003 <http://cds.cern.ch/record/1644757> (see p. 38).
- 2004 [91] CLAUDIO GRANDI, DAVID STICKLAND, LUCAS TAYLOR, ACHILLE PETRILLI, and ALAIN
2005 HERVÉ: **CMS Computing Model: The "CMS Computing Model RTAG"**. Tech. rep.
2006 CMS-NOTE-2004-031. CERN-LHCC-2004-035. LHCC-G-083. Geneva: CERN, Dec.
2007 2004. URL: <http://cds.cern.ch/record/814248> (see p. 38).
- 2008 [92] CHRISTOPH ECK, J KNOBLOCH, LESLIE ROBERTSON, et al.: **LHC computing Grid: Tech-
2009 nical Design Report. Version 1.06 (20 Jun 2005)**. Technical Design Report LCG.
2010 Geneva: CERN, 2005. URL: <https://cds.cern.ch/record/840543> (see p. 38).
- 2011 [93] WORLDWIDE LHC COMPUTING GRID: **WorldWide LHC Computing Gird - 2017**. 2017.
2012 URL: <http://wlcg-public.web.cern.ch> (see p. 39).
- 2013 [94] JOHN C. COLLINS, DAVISON E. SOPER, and GEORGE F. STERMAN: **Factorization of Hard
2014 Processes in QCD**. In: *Adv. Ser. Direct. High Energy Phys.*, **5**: (1989), pp. 1–91. doi:
2015 [10.1142/9789814503266_0001](https://doi.org/10.1142/9789814503266_0001). arXiv: [hep-ph/0409313 \[hep-ph\]](https://arxiv.org/abs/hep-ph/0409313) (see p. 41).

- 2016 [95] RINGAILE PLACAKYTE: **Parton Distribution Functions**. In: *Proceedings, 31st International Conference on Physics in collisions (PIC 2011): Vancouver, Canada, August 28-September 1, 2011*. 2011. arXiv: 1111.5452 [hep-ph]. URL: <https://inspirehep.net/record/954990/files/arXiv:1111.5452.pdf> (see p. 41).
- 2020 [96] RICHARD D. BALL, VALERIO BERTONE, STEFANO CARRAZZA, et al.: **Parton distributions for the LHC run II**. In: *Journal of High Energy Physics*, **2015**:4 (2015), p. 40. doi: 10.1007/JHEP04(2015)040 (see p. 41).
- 2023 [97] JON BUTTERWORTH et al.: **PDF4LHC recommendations for LHC Run II**. In: *J. Phys.*, **G43**: (2016), p. 023001. doi: 10.1088/0954-3899/43/2/023001. arXiv: 1510.03865 [hep-ph] (see pp. 41–42).
- 2026 [98] H. ABRAMOWICZ and A. CALDWELL: **HERA collider physics**. In: *Rev. Mod. Phys.*, **71**: (1999), pp. 1275–1410. doi: 10.1103/RevModPhys.71.1275. arXiv: hep-ex/9903037 [hep-ex] (see p. 42).
- 2029 [99] STEPHEN HOLMES, RONALD S. MOORE, and VLADIMIR SHILTSEV: **Overview of the Tevatron Collider Complex: Goals, Operations and Performance**. In: *JINST*, **6**: (2011), T08001. doi: 10.1088/1748-0221/6/08/T08001. arXiv: 1106.0909 [physics.acc-ph] (see p. 42).
- 2033 [100] JUAN ROJO et al.: **The PDF4LHC report on PDFs and LHC data: Results from Run I and preparation for Run II**. In: *J. Phys.*, **G42**: (2015), p. 103103. doi: 10.1088/0954-3899/42/10/103103. arXiv: 1507.00556 [hep-ph] (see p. 42).
- 2036 [101] ALAN D. MARTIN: **Proton structure, Partons, QCD, DGLAP and beyond**. In: *Acta Phys. Polon.*, **B39**: (2008), pp. 2025–2062. arXiv: 0802.0161 [hep-ph] (see p. 42).
- 2038 [102] J. PUMPLIN, D. STUMP, R. BROCK, et al.: **Uncertainties of predictions from parton distribution functions. 2. The Hessian method**. In: *Phys. Rev.*, **D65**: (2001), p. 014013. doi: 10.1103/PhysRevD.65.014013. arXiv: hep-ph/0101032 [hep-ph] (see p. 42).
- 2041 [103] FRANZ MANDL and GRAHAM G SHAW: **Quantum field theory; 2nd ed.** New York, NY: Wiley, 2010. URL: <https://cds.cern.ch/record/1236742> (see p. 44).
- 2043 [104] S BANERJEE: **CMS Simulation Software**. In: *Journal of Physics: Conference Series*, **396**:2 (2012), p. 022003. URL: <http://stacks.iop.org/1742-6596/396/i=2/a=022003> (see p. 44).
- 2046 [105] M HILDRETH, V N IVANCHENKO, D J LANGE, and M J KORTELAINEN: **CMS Full Simulation for Run-2**. In: *Journal of Physics: Conference Series*, **664**:7 (2015), p. 072022. URL: <http://stacks.iop.org/1742-6596/664/i=7/a=072022> (see p. 44).
- 2049 [106] S. AGOSTINELLI, J. ALLISON, K. AMAKO, et al.: **Geant4-a simulation toolkit**. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **506**:3 (2003), pp. 250–303. doi: [https://doi.org/10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8) (see pp. 44, 46).
- 2053 [107] MICHAEL H. SEYMOUR and MARILYN MARX: **Monte Carlo Event Generators**. In: *Proceedings, 69th Scottish Universities Summer School in Physics : LHC Phenomenology (SUSSP69): St Andrews, Scotland, August 19-September 1, 2012*. 2013, pp. 287–319. doi: 10.1007/978-3-319-05362-2_8. arXiv: 1304.6677 [hep-ph] (see p. 44).

- 2057 [108] TORBJORN SJÖSTRAND: **Monte Carlo Tools**. In: *Proceedings, 65th Scottish Universities Summer School in Physics: LHC Physics (SUSSP65): St. Andrews, UK, August 16-29, 2009*. 2009, pp. 309–339. doi: [10.1201/b11865-14](https://doi.org/10.1201/b11865-14). arXiv: [0911.5286](https://arxiv.org/abs/0911.5286) [hep-ph] (see p. 44).
- 2061 [109] STEFAN HÖCHE: **Introduction to parton-shower event generators**. In: *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders (TASI 2014): Boulder, Colorado, June 2-27, 2014*. 2015, pp. 235–295. doi: [10.1142/9789814678766_0005](https://doi.org/10.1142/9789814678766_0005). arXiv: [1411.4085](https://arxiv.org/abs/1411.4085) [hep-ph] (see pp. 44–45).
- 2066 [110] ADAM ALLOUL, NEIL D. CHRISTENSEN, CELINE DEGRANDE, CLAUDE DUHR, and BENJAMIN FUKS: **FeynRules 2.0 - A complete toolbox for tree-level phenomenology**. In: *Comput. Phys. Commun.*, **185**: (2014), pp. 2250–2300. doi: [10.1016/j.cpc.2014.04.012](https://doi.org/10.1016/j.cpc.2014.04.012). arXiv: [1310.1921](https://arxiv.org/abs/1310.1921) [hep-ph] (see p. 44).
- 2070 [111] CELINE DEGRANDE, CLAUDE DUHR, BENJAMIN FUKS, et al.: **UFO - The Universal FeynRules Output**. In: *Comput. Phys. Commun.*, **183**: (2012), pp. 1201–1214. doi: [10.1016/j.cpc.2012.01.022](https://doi.org/10.1016/j.cpc.2012.01.022). arXiv: [1108.2040](https://arxiv.org/abs/1108.2040) [hep-ph] (see p. 44).
- 2073 [112] JOHAN ALWALL, MICHEL HERQUET, FABIO MALTONI, OLIVIER MATTELAER, and TIM STELZER: **MadGraph 5 : Going Beyond**. In: *JHEP*, **06**: (2011), p. 128. doi: [10.1007/JHEP06\(2011\)128](https://doi.org/10.1007/JHEP06(2011)128). arXiv: [1106.0522](https://arxiv.org/abs/1106.0522) [hep-ph] (see pp. 45, 48).
- 2076 [113] MICHELANGELO L. MANGANO, MAURO MORETTI, FULVIO PICCININI, and MICHELE TREC-CANI: **Matching matrix elements and shower evolution for top-quark production in hadronic collisions**. In: *JHEP*, **01**: (2007), p. 013. doi: [10.1088/1126-6708/2007/01/013](https://doi.org/10.1088/1126-6708/2007/01/013). arXiv: [hep-ph/0611129](https://arxiv.org/abs/hep-ph/0611129) [hep-ph] (see p. 45).
- 2080 [114] TORBJORN SJOSTRAND, STEFAN ASK, JESPER R. CHRISTIANSEN, et al.: **An introduction to {PYTHIA} 8.2**. In: *Computer Physics Communications*, **191**: (2015), pp. 159–177. doi: <http://dx.doi.org/10.1016/j.cpc.2015.01.024> (see pp. 45–46).
- 2083 [115] TORBJORN SJOSTRAND, STEPHEN MRENNA, and PETER Z. SKANDS: **PYTHIA 6.4 Physics and Manual**. In: *JHEP*, **0605**: (2006), p. 026. doi: [10.1088/1126-6708/2006/05/026](https://doi.org/10.1088/1126-6708/2006/05/026). arXiv: [hep-ph/0603175](https://arxiv.org/abs/hep-ph/0603175) [hep-ph] (see pp. 45–46).
- 2086 [116] TORBJORN SJOSTRAND, STEFAN ASK, JESPER R. CHRISTIANSEN, et al.: **An Introduction to PYTHIA 8.2**. In: *Comput. Phys. Commun.*, **191**: (2015), pp. 159–177. doi: [10.1016/j.cpc.2015.01.024](https://doi.org/10.1016/j.cpc.2015.01.024). arXiv: [1410.3012](https://arxiv.org/abs/1410.3012) [hep-ph] (see pp. 45–46, 48).
- 2089 [117] JOHAN ALWALL et al.: **Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions**. In: *Eur. Phys. J.*, **C53**: (2008), pp. 473–500. doi: [10.1140/epjc/s10052-007-0490-5](https://doi.org/10.1140/epjc/s10052-007-0490-5). arXiv: [0706.2569](https://arxiv.org/abs/0706.2569) [hep-ph] (see p. 45).
- 2093 [118] J. ALWALL, R. FREDERIX, S. FRIXIONE, et al.: **The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations**. In: *JHEP*, **07**: (2014), p. 079. doi: [10.1007/JHEP07\(2014\)079](https://doi.org/10.1007/JHEP07(2014)079). arXiv: [1405.0301](https://arxiv.org/abs/1405.0301) [hep-ph] (see pp. 45, 48).
- 2097 [119] RIKKERT FREDERIX and STEFANO FRIXIONE: **Merging meets matching in MC@NLO**. In: *JHEP*, **12**: (2012), p. 061. doi: [10.1007/JHEP12\(2012\)061](https://doi.org/10.1007/JHEP12(2012)061). arXiv: [1209.6215](https://arxiv.org/abs/1209.6215) [hep-ph] (see p. 45).

- 2100 [120] SIMONE ALIOLI, PAOLO NASON, CARLO OLEARI, and EMANUELE RE: A general frame-
 2101 **work for implementing NLO calculations in shower Monte Carlo programs: the**
 2102 **POWHEG BOX.** In: *Journal of High Energy Physics*, **2010**:6 (2010), p. 43. doi: [10.1007/JHEP06\(2010\)043](https://doi.org/10.1007/JHEP06(2010)043) (see p. 45).
- 2104 [121] SIMONE ALIOLI, PAOLO NASON, CARLO OLEARI, and EMANUELE RE: NLO single-top
 2105 **production matched with shower in POWHEG: s - and t -channel contributions.**
 2106 In: *Journal of High Energy Physics*, **2009**:09 (2009), p. 111. url: <http://stacks.iop.org/1126-6708/2009/i=09/a=111> (see p. 45).
- 2108 [122] STEFANO FRIXIONE, PAOLO NASON, and CARLO OLEARI: Matching NLO QCD com-
 2109 **putations with parton shower simulations: the POWHEG method.** In: *Journal of*
 2110 *High Energy Physics*, **2007**:11 (2007), p. 070. url: <http://stacks.iop.org/1126-6708/2007/i=11/a=070> (see p. 45).
- 2112 [123] SIMONE ALIOLI, PAOLO NASON, CARLO OLEARI, and EMANUELE RE: A general frame-
 2113 **work for implementing NLO calculations in shower Monte Carlo programs: the**
 2114 **POWHEG BOX.** In: *JHEP*, **06**: (2010), p. 043. doi: [10.1007/JHEP06\(2010\)043](https://doi.org/10.1007/JHEP06(2010)043). arXiv:
 2115 [1002.2581 \[hep-ph\]](https://arxiv.org/abs/1002.2581) (see p. 45).
- 2116 [124] STEFANO FRIXIONE, PAOLO NASON, and CARLO OLEARI: Matching NLO QCD com-
 2117 **putations with Parton Shower simulations: the POWHEG method.** In: *JHEP*, **11**:
 2118 (2007), p. 070. doi: [10.1088/1126-6708/2007/11/070](https://doi.org/10.1088/1126-6708/2007/11/070). arXiv: [0709.2092 \[hep-ph\]](https://arxiv.org/abs/0709.2092) (see
 2119 p. 45).
- 2120 [125] PAOLO NASON: **A New method for combining NLO QCD with shower Monte Carlo**
 2121 **algorithms.** In: *JHEP*, **11**: (2004), p. 040. doi: [10.1088/1126-6708/2004/11/040](https://doi.org/10.1088/1126-6708/2004/11/040).
 2122 arXiv: [hep-ph/0409146 \[hep-ph\]](https://arxiv.org/abs/hep-ph/0409146) (see p. 45).
- 2123 [126] ANDREI V. GRITSAN, RAOUL RÖNTSCH, MARKUS SCHULZE, and MENG XIAO: **Constraining**
 2124 **anomalous Higgs boson couplings to the heavy flavor fermions using matrix**
 2125 **element techniques.** In: *Phys. Rev.*, **D94**:5 (2016), p. 055023. doi: [10.1103/PhysRevD.94.055023](https://doi.org/10.1103/PhysRevD.94.055023). arXiv: [1606.03107 \[hep-ph\]](https://arxiv.org/abs/1606.03107) (see pp. 45, 48).
- 2127 [127] IAN ANDERSON et al.: **Constraining anomalous HVV interactions at proton and**
 2128 **lepton colliders.** In: *Phys. Rev.*, **D89**:3 (2014), p. 035007. doi: [10.1103/PhysRevD.89.035007](https://doi.org/10.1103/PhysRevD.89.035007). arXiv: [1309.4819 \[hep-ph\]](https://arxiv.org/abs/1309.4819) (see pp. 45, 48).
- 2130 [128] SARA BOLOGNESI, YANYAN GAO, ANDREI V. GRITSAN, et al.: **On the spin and parity of**
 2131 **a single-produced resonance at the LHC.** In: *Phys. Rev.*, **D86**: (2012), p. 095031.
 2132 doi: [10.1103/PhysRevD.86.095031](https://doi.org/10.1103/PhysRevD.86.095031). arXiv: [1208.4018 \[hep-ph\]](https://arxiv.org/abs/1208.4018) (see pp. 45, 48).
- 2133 [129] YANYAN GAO, ANDREI V. GRITSAN, ZIJIN GUO, et al.: **Spin determination of single-**
 2134 **produced resonances at hadron colliders.** In: *Phys. Rev.*, **D81**: (2010), p. 075022.
 2135 doi: [10.1103/PhysRevD.81.075022](https://doi.org/10.1103/PhysRevD.81.075022). arXiv: [1001.3396 \[hep-ph\]](https://arxiv.org/abs/1001.3396) (see pp. 45, 48).
- 2136 [130] PIERRE ARTOISENET, RIKKERT FREDERIX, OLIVIER MATTELAER, and ROBBERT RIETK-
 2137 **ERK: Automatic spin-entangled decays of heavy resonances in Monte Carlo simula-**
 2138 **tions.** In: *JHEP*, **03**: (2013), p. 015. doi: [10.1007/JHEP03\(2013\)015](https://doi.org/10.1007/JHEP03(2013)015). arXiv: [1212.3460 \[hep-ph\]](https://arxiv.org/abs/1212.3460) (see pp. 46, 48).
- 2140 [131] V. KHACHATRYAN and ETAL: **Event generator tunes obtained from underlying event**
 2141 **and multiparton scattering measurements.** In: *The European Physical Journal C*,
 2142 **76**:3 (Mar. 17, 2016), p. 155. doi: [10.1140/epjc/s10052-016-3988-x](https://doi.org/10.1140/epjc/s10052-016-3988-x) (see p. 46).

- 2143 [132] YUE ZHANG, BO HUA LI, CHONG SHENG LI, JUN GAO, and HUA XING ZHU: **Next-to-leading order QCD corrections to the top quark associated with γ production via model-independent flavor-changing neutral-current couplings at hadron colliders.** In: *Phys. Rev.*, D83: (2011), p. 094003. DOI: [10.1103/PhysRevD.83.094003](https://doi.org/10.1103/PhysRevD.83.094003). arXiv: [1101.5346 \[hep-ph\]](https://arxiv.org/abs/1101.5346) (see p. 47).
- 2148 [133] A. HOECKER, P. SPECKMAYER, J. STELZER, et al.: **TMVA - Toolkit for Multivariate Data Analysis.** In: *ArXiv Physics e-prints*, (Mar. 2007). eprint: [physics/0703039](https://arxiv.org/abs/physics/0703039) (see pp. 50–51).
- 2151 [134] R. BRUN and F. RADEMAKERS: **ROOT: An object oriented data analysis framework.** In: *Nucl. Instrum. Meth.*, A389: (1997), pp. 81–86. DOI: [10.1016/S0168-9002\(97\)00048-X](https://doi.org/10.1016/S0168-9002(97)00048-X) (see p. 50).
- 2154 [135] A. MAYR, H. BINDER, O. GEFFELER, and M. SCHMID: **The Evolution of Boosting Algorithms - From Machine Learning to Statistical Modelling.** In: *ArXiv e-prints*, (Mar. 2014). arXiv: [1403.1452 \[stat.ME\]](https://arxiv.org/abs/1403.1452) (see p. 51).
- 2157 [136] OLAF BEHNKE, KEVIN KRONINGER, GREGORY SCHOTT, and THOMAS SCHORNER-SADENIUS: **Data Analysis in High Energy Physics: A Practical Guide to Statistical Methods.** 1st. Wiley-VCH, 2013 (see p. 52).
- 2160 [137] CHRISTIAN BÖSER, SIMON FINK, and STEFFEN RÖCKER: **Introduction to Boosted Decision Trees: A multivariate approach to classification problems.** Presented at the 'KSETA Doktoranden' workshop, Lauterbad. July 2014. URL: <https://indico.scc.kit.edu/indico/event/48/session/4/contribution/35/material/slides/0.pdf> (see p. 53).
- 2165 [138] SERGUEI CHATRHYAN et al.: **Combined results of searches for the standard model Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV.** In: *Phys. Lett.*, B710: (2012), pp. 26–48. DOI: [10.1016/j.physletb.2012.02.064](https://doi.org/10.1016/j.physletb.2012.02.064). arXiv: [1202.1488 \[hep-ex\]](https://arxiv.org/abs/1202.1488) (see p. 52).
- 2168 [139] GLEN COWAN, KYLE CRANMER, EILAM GROSS, and OFER VITELLS: **Asymptotic formulae for likelihood-based tests of new physics.** In: *Eur. Phys. J.*, C71: (2011). [Erratum: *Eur. Phys. J.*C73,2501(2013)], p. 1554. DOI: [10.1140/epjc/s10052-011-1554-0](https://doi.org/10.1140/epjc/s10052-011-1554-0), [10.1140/epjc/s10052-013-2501-z](https://doi.org/10.1140/epjc/s10052-013-2501-z). arXiv: [1007.1727 \[physics.data-an\]](https://arxiv.org/abs/1007.1727) (see pp. 52, 56).
- 2173 [140] **Observation of a new boson with a mass near 125 GeV.** Tech. rep. CMS-PAS-HIG-12-020. Geneva: CERN, 2012. URL: <https://cds.cern.ch/record/1460438> (see pp. 52, 57).
- 2176 [141] **Procedure for the LHC Higgs boson search combination in Summer 2011.** Tech. rep. CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11. Geneva: CERN, Aug. 2011. URL: <http://cds.cern.ch/record/1379837> (see pp. 52, 55–56).
- 2179 [142] HIGGS WORKING GROUP: **Documentation of the RooStats based statistics tools for Higgs PAG.** 2017. URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideHiggsAnalysisCombinedLimit> (see p. 52).
- 2182 [143] LORENZO MONETA, KEVIN BELASCO, KYLE S. CRANMER, et al.: **The RooStats Project.** In: *PoS, ACAT2010*: (2010), p. 057. arXiv: [1009.1003 \[physics.data-an\]](https://arxiv.org/abs/1009.1003) (see p. 52).

- 2184 [144] GLEN COWAN, KYLE CRANMER, EILAM GROSS, and OFER VITELLS: **Asymptotic formulae**
 2185 **for likelihood-based tests of new physics.** In: *The European Physical Journal C*, 71:2
 2186 (Feb. 2011). DOI: [10.1140/epjc/s10052-011-1554-0](https://doi.org/10.1140/epjc/s10052-011-1554-0). arXiv: [1007.1727 \[hep-ph\]](https://arxiv.org/abs/1007.1727) (see
 2187 p. 53).
- 2188 [145] THOMAS JUNK: **Confidence level computation for combining searches with small**
 2189 **statistics.** In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators,*
 2190 *Spectrometers, Detectors and Associated Equipment*, 434:2 (1999), pp. 435–443.
 2191 DOI: [https://doi.org/10.1016/S0168-9002\(99\)00498-2](https://doi.org/10.1016/S0168-9002(99)00498-2) (see p. 53).
- 2192 [146] A L READ: **Presentation of search results: the CL s technique.** In: *Journal of Physics G: Nuclear and Particle Physics*, 28:10 (2002), p. 2693. URL: <http://stacks.iop.org/0954-3899/28/i=10/a=313> (see p. 53).
- 2195 [147] J. S. CONWAY: **Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra.** In: *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN, Geneva, Switzerland 17-20 January 2011.* 2011, pp. 115–120. DOI: [10.5170/CERN-2011-006.115](https://doi.org/10.5170/CERN-2011-006.115). arXiv: [1103.0354 \[physics.data-an\]](https://arxiv.org/abs/1103.0354) (see p. 55).
- 2200 [148] A. M. SIRUNYAN et al.: **Particle-flow reconstruction and global event description with the CMS detector.** In: *JINST*, 12: (2017), P10003. DOI: [10.1088/1748-0221/12/10/P10003](https://doi.org/10.1088/1748-0221/12/10/P10003). arXiv: [1706.04965 \[physics.ins-det\]](https://arxiv.org/abs/1706.04965) (see pp. 59–60, 63, 66, 68, 70).
- 2203 [149] R. FRÜHWIRTH: **Application of Kalman filtering to track and vertex fitting.** In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 262:2 (1987), pp. 444–450. DOI: [http://dx.doi.org/10.1016/0168-9002\(87\)90887-4](https://doi.org/10.1016/0168-9002(87)90887-4) (see p. 61).
- 2207 [150] PIERRE BILLOIR: **Progressive track recognition with a Kalman like fitting procedure.** In: *Comput. Phys. Commun.*, 57: (1989), pp. 390–394. DOI: [10.1016/0010-4655\(89\)90249-X](https://doi.org/10.1016/0010-4655(89)90249-X) (see p. 61).
- 2210 [151] SERGUEI CHATRHYAN et al.: **Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV.** In: *JINST*, 7: (2012), P10002. DOI: [10.1088/1748-0221/7/10/P10002](https://doi.org/10.1088/1748-0221/7/10/P10002). arXiv: [1206.4071 \[physics.ins-det\]](https://arxiv.org/abs/1206.4071) (see p. 61).
- 2213 [152] W ADAM, R FRÜHWIRTH, A STRANDLIE, and T TODOROV: **Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC.** In: *Journal of Physics G: Nuclear and Particle Physics*, 31:9 (2005), N9. URL: <http://stacks.iop.org/0954-3899/31/i=9/a=N01> (see p. 62).
- 2217 [153] **Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV.** In: *Journal of Instrumentation*, 10:06 (2015), P06005. URL: <http://stacks.iop.org/1748-0221/10/i=06/a=P06005> (see p. 63).
- 2220 [154] **Electron and photon performance in CMS with the full 2016 data sample.** In: (Mar. 2017). URL: <http://cds.cern.ch/record/2255497> (see pp. 63, 70–71).
- 2222 [155] K. ROSE: **Deterministic annealing for clustering, compression, classification, regression, and related optimization problems.** In: *Proceedings of the IEEE*, 86:11 (Nov. 1998), pp. 2210–2239. DOI: [10.1109/5.726788](https://doi.org/10.1109/5.726788) (see p. 63).

- 2225 [156] WOLFGANG WALTENBERGER: **Adaptive Vertex Reconstruction**. Tech. rep. CMS-NOTE-
 2226 2008-033. Geneva: CERN, July 2008. URL: <https://cds.cern.ch/record/1166320> (see
 2227 p. 63).
- 2228 [157] ANDREAS KORNMAYER: **The CMS Pixel Luminosity Telescope**. Tech. rep. CMS-CR-
 2229 2015-121. Geneva: CERN, June 2015. URL: <https://cds.cern.ch/record/2039978> (see
 2230 p. 66).
- 2231 [158] **CMS Luminosity Measurements for the 2016 Data Taking Period**. Tech. rep. CMS-
 2232 PAS-LUM-17-001. Geneva: CERN, 2017. URL: <http://cds.cern.ch/record/2257069>
 2233 (see pp. 66–67).
- 2234 [159] **Pileup Removal Algorithms**. Tech. rep. CMS-PAS-JME-14-001. Geneva: CERN, 2014.
 2235 URL: <http://cds.cern.ch/record/1751454> (see p. 67).
- 2236 [160] **Muon Identification and Isolation efficiency on full 2016 dataset**. In: (Mar. 2017).
 2237 URL: <http://cds.cern.ch/record/2257968> (see pp. 67–69).
- 2238 [161] **Effective areas used for Summer16 samples**. 2017. URL: https://indico.cern.ch/event/482673/contributions/2187022/%20attachments/1282446/1905912/talk_electron_ID_spring16.pdf (see p. 71).
- 2241 [162] MATTEO CACCIARI, GAVIN P. SALAM, and GREGORY SOYEZ: **The Anti-k(t) jet clustering
 2242 algorithm**. In: *JHEP*, **04**: (2008), p. 063. DOI: <10.1088/1126-6708/2008/04/063>. arXiv:
 2243 [0802.1189 \[hep-ph\]](0802.1189) (see p. 71).
- 2244 [163] **Jet algorithms performance in 13 TeV data**. Tech. rep. CMS-PAS-JME-16-003.
 2245 Geneva: CERN, 2017. URL: <http://cds.cern.ch/record/2256875> (see p. 72).
- 2246 [164] V. KHACHATRYAN et al.: **Jet energy scale and resolution in the CMS experiment in
 2247 pp collisions at 8 TeV**. In: *Journal of Instrumentation*, **12**:02 (2017), P02014. URL:
 2248 <http://stacks.iop.org/1748-0221/12/i=02/a=P02014> (see pp. 72–74).
- 2249 [165] **Jet energy scale and resolution performances with 13TeV data**. In: (June 2016).
 2250 URL: <http://cds.cern.ch/record/2160347> (see pp. 73–74).
- 2251 [166] THE CMS COLLABORATION: **Identification of b-quark jets with the CMS experiment**.
 2252 In: *Journal of Instrumentation*, **8**:04 (2013), P04013. URL: <http://stacks.iop.org/1748-0221/8/i=04/a=P04013> (see pp. 74–75).
- 2254 [167] **Identification of b quark jets at the CMS Experiment in the LHC Run 2**. Tech. rep.
 2255 CMS-PAS-BTV-15-001. Geneva: CERN, 2016. URL: <http://cds.cern.ch/record/2138504> (see pp. 74, 76).
- 2257 [168] NAZAR BARTOSIK: **Diagram showing the common principle of identification of jets
 2258 initiated by b-hadron decays**. 2016. URL: http://bartosik.pp.ua/hep_sketches/btagging (see p. 75).
- 2260 [169] **Identification of b quark jets at the CMS Experiment in the LHC Run 2**. Tech. rep.
 2261 CMS-PAS-BTV-16-001. unpublished. Geneva: CERN, 2017 (see pp. 76–77).
- 2262 [170] **Performance of missing energy reconstruction in 13 TeV pp collision data using
 2263 the CMS detector**. Tech. rep. CMS-PAS-JME-16-004. Geneva: CERN, 2016. URL:
 2264 <http://cds.cern.ch/record/2205284> (see pp. 78, 82–83).

- 2265 [171] JORDAN DAMGOV: **Missing Et Optional Filters Run 2.** 2017. URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETOptionalFiltersRun2> (see p. 82).
- 2266
- 2267 [172] **EGM corrections for Moriond17.** URL: https://indico.cern.ch/event/613162/contributions/2472046/attachments/1410308/2156822/EGMSmearerAndRegression_PPD_022017.pdf (see p. 87).
- 2268
- 2269
- 2270 [173] **Misalignment and Muon Scale corrections.** URL: https://www-cdf.fnal.gov/~jyhan/cms_momscl/rochcor_cmsnote.pdf (see p. 88).
- 2271
- 2272 [174] **Rochestor Correction.** Accessed: 2017-05-20. URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/RochcorMuon> (see p. 88).
- 2273