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

Isis Van Parijs

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

Published in Faculteit Wetenschappen & Bio-ingenieurswetenschappen
Vrije Universiteit Brussel
At 1. June 2017.

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

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

Contents

1 Theoretical basis

The Standard Model (SM) [1] is a name given in 1970s to a theory describing the fundamental particles and their interactions. This quantum field theory describes the particles and their interactions as fields and has successfully incorporated three of the four fundamental forces in the universe. In Section 1.1, the particle content of the SM is summarised, while Section 1.2 describes the SM Lagrangian and its symmetries. In Section 1.3, the flavour content of the SM is highlighted, while Section 1.4 focusses on the SM top quark. The latest experimental results of the top quark are given in Section 1.5.

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

1.1 Getting to the nature of things

The interactions in nature can be described by four forces, the strong force, the electromagnetic (EM) force, the weak force and the gravitational force. These interactions happen via particles with an integer spin known as bosons. The strong interaction is mediated by eight gluons g, while the electromagnetic force is mediated by photons γ, and the weak force by Z and $\mathrm{W}^{\pm}$ bosons. In Table 1.1, the forces and their characteristics are shown. The gravitational force is 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	$10^{e}{-}15$ m	8 gluons
Electromagnetic force	∞	photon
Weak force	10^{-18} m	$\mathrm{W}^{\pm}$, 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 don't 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 compromises the muon μ^- and muon neutrino ν_μ, whereas the third generation consists of the tau τ and tau neutrino ν_τ. The neutrino's are neutral particles, while the other leptons have charge $\pm q_\mathrm{e}$ where q_e represents 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 neutrino's are assumed to be massless, nonetheless it is experimentally established that neutrino do have a tiny non-zero mass. 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 [4], where q_e represents the elementary charge.

Generation	Particle	Mass	Charge
First	e^-	0.511 MeV	$-q_\mathrm{e}$
	ν_e	≈ 0	0
Second	μ^-	106 MeV	$-q_\mathrm{e}$
	ν_μ	≈ 0	0
Third	τ	1 777 MeV	$-q_\mathrm{e}$
	ν_τ	≈ 0	0

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 [**observationtopD0**, **observationtopCDF**], is the heaviest SM particle with a mass close to 173.1 ± 0.6 GeV[1] [4]. The quarks and their properties are summarized 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 particle that rapidly decay through $\mathrm{W}^\pm$ and Z bosons, associated with a fermion. 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 and is discovered in 2012 [5, 6]. It is responsible for the masses of the $\mathrm{W}^\pm$ 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 $\mathscr{L}$. The theory is based on the $\mathrm{SU_C}(3) \times \mathrm{SU_L}(2) \times \mathrm{U_Y}(1)$ gauge

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

Table 1.3: The properties of the quarks in the three generations of the SM [4], 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	$\frac{-1}{3}\ q_e$

symmetry, where $\mathrm{SU_L}(2) \times \mathrm{U_Y}(1)$ describes the electroweak interaction and $\mathrm{SU_C}(3)$ the strong coupling. The indices refer to colour C, the left chiral nature of the $\mathrm{SU_L}(2)$ coupling L, and the weak hypercharge Y. Its Lagrangian is constructed such that contains symmetries representing physics conservation laws such as conservation of energy, momentum and angular momentum. The symmetries under local group transformations are sustained by demanding gauge invariance .

NOTE: should I explain gauge invariance or is a reference enough?

The $\mathrm{U_Y}(1)$ group has one generator Y with an associated gauge field B_μ. The three gauge fields W^1_μ, W^2_μ, and W^3_μ, are associated to $\mathrm{SU_L}(2)$ with three generators that can can be written as half of the Pauli matrices:

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

The generators T^a satisfy the Lie algebra:

$$\left[T^a, T^b\right] = i\epsilon^{abc} T_c \text{ and} \left[T^a, Y\right] = 0, \tag{1.2}$$

where ϵ^{abc} is an antisymmetric tensor. The gauge fields of $\mathrm{SU_L}(2)$ only couple to left-handed fermions as required by the observed parity violating nature of the weak force. The $\mathrm{SU_C}(3)$ group represents quantum chromodynamics (QCD). It has eight generators corresponding to eight gluon fields $\mathrm{G}^{1\ldots8}_\mu$. Unlike $\mathrm{SU_L}(2) \times \mathrm{U_Y}(1)$, $\mathrm{SU_C}(3)$ is not chiral.

Under $\mathrm{SU_C}(3)$ 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:

$$\mathrm{l_{L,i}} = \begin{pmatrix} \mathrm{e^-}_{\mathrm{L,i}} \\ \nu_{\mathrm{L,i}} \end{pmatrix},\ \mathrm{e^-}_{\mathrm{R,i}},\ \mathrm{q_{L,i}} = \begin{pmatrix} \mathrm{u_{L,i}} \\ \mathrm{d_{L,i}} \end{pmatrix},\ \mathrm{u_{R,i}},\ \text{and } \mathrm{d_{R,i}} \tag{1.3}$$

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

$$\mathscr{L}_{\mathrm{SM}} = \mathscr{L}_{\mathrm{gauge}} + \mathscr{L}_{\mathrm{f}} + \mathscr{L}_{\mathrm{Yuk}} + \mathscr{L}_{\phi}, \tag{1.4}$$

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 $\mathrm{U_Y}(1)$ and g for $\mathrm{SU_L}(2)$. The physically observable gauge bosons of this theory are the photon field A_μ, the Z boson field Z^0_μ, and the W field $\mathrm{W}^\pm_\mu$. These are a superposition of the four gauge fields of $\mathrm{SU_L}(2) \times \mathrm{U_Y}(1)$:

$$\mathrm{A}_\mu = \sin\theta_\mathrm{W} \mathrm{W}^1_\mu + \cos\theta_\mathrm{W} \mathrm{B}_\mu,\ \mathrm{Z}^0_\mu = \cos\theta_\mathrm{W} \mathrm{W}^3_\mu - \sin\theta_\mathrm{W} \mathrm{B}_\mu,\ \text{and } \mathrm{W}^\pm_\mu = \sqrt{\frac{1}{2}}\left(\mathrm{W}^1_\mu \mp \mathrm{W}^2_\mu\right), \quad (1.5)$$

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

The coupling constant representing the strength of the QCD interactions is denoted as g_s. In QCD their 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 $\mathscr{L}_\text{gauge}$ and $\mathscr{L}_\text{f}$ are no mass terms for fermions present because only singlets under $\mathrm{SU_C}(3) \times \mathrm{SU_L}(2) \times \mathrm{U_Y}(1)$ 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 $\mathscr{L}_\phi$ is introduced.

The scalar doublet is introduced in the SM as

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

NOTE: check if I need to add constants here

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 ($\not{\mathrm{EW}}$):

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

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 is reduced to three degrees of freedom that couple to the gauge fields and mix with 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:

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

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

1.3 Flavours in the SM

Flavour changing charged currents are introduced in 1963 by Nicola Cabibbo [7]. Via interaction 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 $\mathrm{d_{weak}}$, where $\mathrm{d_{weak}}$ is a linear combination of the down and strange quarks, $\mathrm{d_{weak}} = \cos\theta_c \mathrm{d} + \sin\theta_c \mathrm{s}$. This linear combination is a direct consequence of the chosen rotation

$$\begin{pmatrix} \mathrm{d_{weak}} \\ \mathrm{s_{weak}} \end{pmatrix} = \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix} \begin{pmatrix} \mathrm{d} \\ \mathrm{s} \end{pmatrix} = \mathscr{R} \begin{pmatrix} \mathrm{d} \\ \mathrm{s} \end{pmatrix}, \tag{1.10}$$

where the rotation angle θ_c is known as the Cabibbo angle. This provides a definition for the charged weak current between uand dquarks,

$$J_\mu = \bar{\mu} \gamma_\mu (1 + \gamma_5) \mathrm{d_{weak}}. \tag{1.11}$$

A consequence of Cabibbo's approach is that the $\mathrm{s_{weak}}$ is left uncoupled, leading to Glashow, Iliopoulos and Maiani (GIM) [8–10] to require the existence of a fourth quark with charge $\frac{2}{3} q_e$. This quark, known as the charm quark, couples to $\mathrm{s_{weak}}$ and a new definition of the charged weak current is modified to

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

The neutral weak current is defined as

$$J_3 = \bar{U} \gamma_\mu (1 + \gamma_5) \left[\mathscr{R}, \mathscr{R}^\dagger \right] D, \tag{1.13}$$

and is diagonal in flavour space. This has as consequence that no flavour changing neutral currents occur at tree-level Feynmann diagrams[2] .

NOTE: should I explain feynmann diagrams?

Kobayashi and Maskawa generalised the Cabibbo rotation matrix to accommodate for a third generation of quarks. The result is a 3×3 unitary matrix known as the CKM matrix, responsible

[2]Feynmann diagrams are physical representation of interaction between particles. They are based on Feynmann rules [1].

for the mixing of weak interaction states of down-type quarks:

$$\begin{pmatrix} \mathrm{d_{weak}} \\ \mathrm{s_{weak}} \\ \mathrm{b_{weak}} \end{pmatrix} = \begin{pmatrix} V_{\mathrm{ud}} & V_{\mathrm{us}} & V_{\mathrm{ub}} \\ V_{\mathrm{cd}} & V_{\mathrm{cs}} & V_{\mathrm{cb}} \\ V_{\mathrm{td}} & V_{\mathrm{ts}} & V_{\mathrm{tb}} \end{pmatrix} \begin{pmatrix} \mathrm{d} \\ \mathrm{s} \\ \mathrm{b} \end{pmatrix} = \mathscr{V}_{\mathrm{CKM}} \begin{pmatrix} \mathrm{d} \\ \mathrm{s} \\ \mathrm{b} \end{pmatrix}. \tag{1.14}$$

The unitarity of the matrix ($\mathscr{V}^{\dagger}_{\mathrm{CKM}}\mathscr{V}_{\mathrm{CKM}} = \mathbb{1}$). 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 fives 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 [11]. Each element V_{ij} of $\mathscr{V}_{\mathrm{CKM}}$ represents the transition probability of a quark i going to a quark j, and is experimentally determined to be [4]

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

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

Table 1.4: The predicted branching ratios $\mathscr{B}$ for FCNC interactions involving the top quark in the SM [12]

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

1.4 The top of the SM

Discovered in 1995 by the CDF and D0 collaborations at Tevatron with proton-antiproton data [13, 14], the top quark plays an important role in studying high energy physics. Its Yukawa interaction is given by

$$\mathscr{L}_{\mathrm{top-Yukawa}} = -\frac{\lambda_t v}{\sqrt{2}}\bar{\mathrm{t}}_{\mathrm{L}}\mathrm{t}_{\mathrm{R}} - \frac{\lambda_{\mathrm{t}}}{\sqrt{2}}\mathrm{H}\bar{\mathrm{t}}_{\mathrm{L}}\mathrm{t}_{\mathrm{R}} + \mathrm{h.c.}, \tag{1.16}$$

yielding a Yukawa coupling of

$$\lambda_{\mathrm{t}} = \frac{\sqrt{2}m_{\mathrm{t}}}{v} = 0.991 \pm 0.003, \tag{1.17}$$

with the top mass m_t equal to 172.44 ± 0.49 GeV [4]. This Yukawa coupling is very large compared to the other Yukawa couplings in the SM ($\mathcal{O}(10^{-2})$), leading to the belief that the top quark may have an important role in understanding the mechanism of electroweak symmetry breaking. On top of this, the very short lifetime of the top quark makes it an excellent candidate for property studies. Its high mass, almost 40 times higher than the mass of the closest particle in mass, leads to a large coupling with the Higgs boson and makes the top quark an interesting candidate for the understanding of how particles acquire mass.

The CKM matrix element V_{tb}, given in Equation 1.15, is experimentally found to be much larger than V_{ts}, V_{td}, and close to unity. The top quark decays through electroweak interactions since the W boson mass is smaller than the top mass and the W boson can be on shell. A consequence of this is that the top quark has a very short lifetime of only $1/\Gamma_t \approx 5\,10^{-25}$ s [4] leading to the fact that the formation of bound states involving top quarks are not allowed. This lifetime is even shorter than the typical hadronisation timescale of $1/\Lambda_{QCD} \approx 10^{-23}$ s, prohibiting gluons to radiate from the top quark and keeping its spin coherent. Since the electroweak interactions have a V-A coupling structure , the top quark spin orientation can be derived from the angular distributions of its decay products. This makes it possible to study the polarisation of top quarks from the angular distributions in various processes.

NOTE: Explain V-A

The massiveness of the top quark leads to the fact that a large amount of energy is needed to create one. This is only the case for high energy collisions such as those in the Earth's upper atmosphere as cosmic rays collide with particles in air, or by particle accelerators. The production of top quarks happens in two ways: single via the electroweak interaction or in pairs via the strong interaction. At hadron colliders, the dominant production mechanism is top 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 pair production mechanisms are shown. The production channel of gluon fusion is the main contributor to the top pair cross section the LHC compared to quark fusion at Tevatron. The $gg \rightarrow t\bar{t}$ process contributes 80-90% to the total top pair cross section in the LHC centre-of-mass energy regime of 7-14 TeV [4]. In Table 1.5 the predicted top pair production cross sections are given for the LHC and Tevatron.

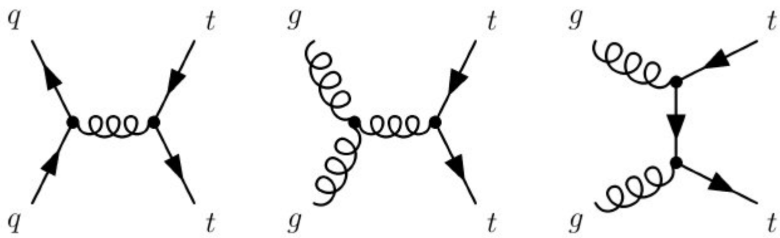

Figure 1.1: Leading order diagrams of the top pair production. Gluon fusion (right and middle)are the dominant processes at the LHC, while quark fusion (left) is the dominant one at Tevatron.

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 1.2, the corresponding Feynman diagrams are shown. The single top quark production cross section, given in Table 1.6, are smaller than

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 [4]. The first uncertainty is from scale dependence, while the second uncertainty originates from parton density functions.

Experiment	Top 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	$\sigma_{t\bar{t}} = 816.0^{+19.4+34.4}_{-28.6-34.4}$

the top pair production cross sections since the electroweak coupling strength is smaller than the strong coupling strength. In addition, for the single top 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.

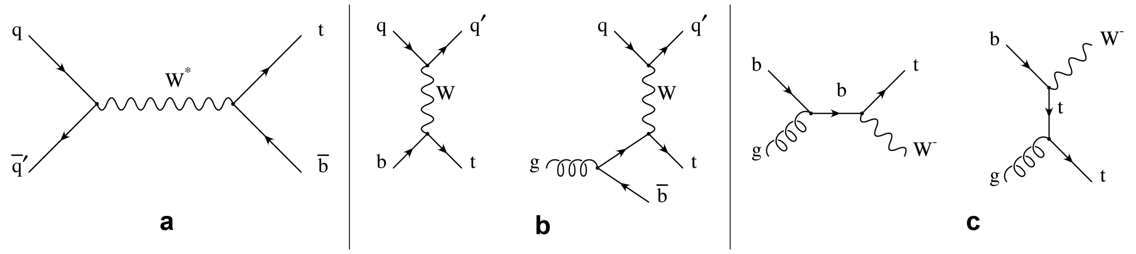

Figure 1.2: Leading order Feynman diagrams of the electroweak production of single top quarks in the s-channel (left), t-channel (middle), and for the tW associated production. Figure taken from [15].

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 $\rightarrow b\bar{b}$). This process is also known as W-gluon fusion production. It has the highest single top quark cross section in proton collisions and the top quark production is roughly twice more than the antitop quarks. 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 Tevatron, but of relevant size at the LHC. The tW-channel is sensitive to new physics affecting the Wtb vertex.

Table 1.6: Predictions on the single top quark production cross sections at next-to-leading order per centre-of-mass energy [4]. 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 $\bar{\mathrm{t}}$ are 65% and 35%. For the s-channel this respectively 69% and 31%. The tW-channel has an equal proportion of top and antitop quarks. For Tevatron, the top mass is assumed to be 173.3 GeV, while for the LHC predictions $m_{\mathrm{t}} = 172.5$ GeV [4, 16].

Experiment	Centre-of-mass energy	Cross section $\sigma_{\mathrm{t}+\bar{\mathrm{t}}}$ (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	$216.99^{+9.04}_{-7.71}$	$10.32^{+0.40}_{-0.36}$	$71.7^{+1.80+3.40}_{-1.80-3.40}$

1.5 Hunting down the SM top quark

In this section a selection of experimental results of measurements on the SM are presented. In Figure 1.3, a summary plot of the CMS cross section measurements can be found. The estimations by the CMS and ATLAS collaborations of the CKM matrix element V_{tb} from single top quark measurement is given in Figure 1.4. The most precise estimation of V_{tb} originates from a combination of t-channel cross section measurements at 7 and 8 TeV by the CMS collaboration resulting in $|f_{\mathrm{L}} V_{\mathrm{tb}}| = 0.998 \pm 0.038$ (exp.) ± 0.016 (theo.). Assuming the $f_{\mathrm{L}} = 1$ and $|V_{\mathrm{tb}}| < 1$, this result yields a limit of $|V_{\mathrm{tb}}| > 0.92$ at 95% confidence level. The most recent top mass measurements are given in Figure 1.5. The CMS combined top mass measurement is $m_{\mathrm{t}} = 172.44 \pm 0.48$ GeV from 7+8 TeV data.

In general the various measurements show a good agreement with the SM predictions and by lack of deviations of 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 1.6. These results are consistent with the SM expectation for which those operators vanish.

NOTE: Should I explain Wilson coeff and Wrasaw basis?

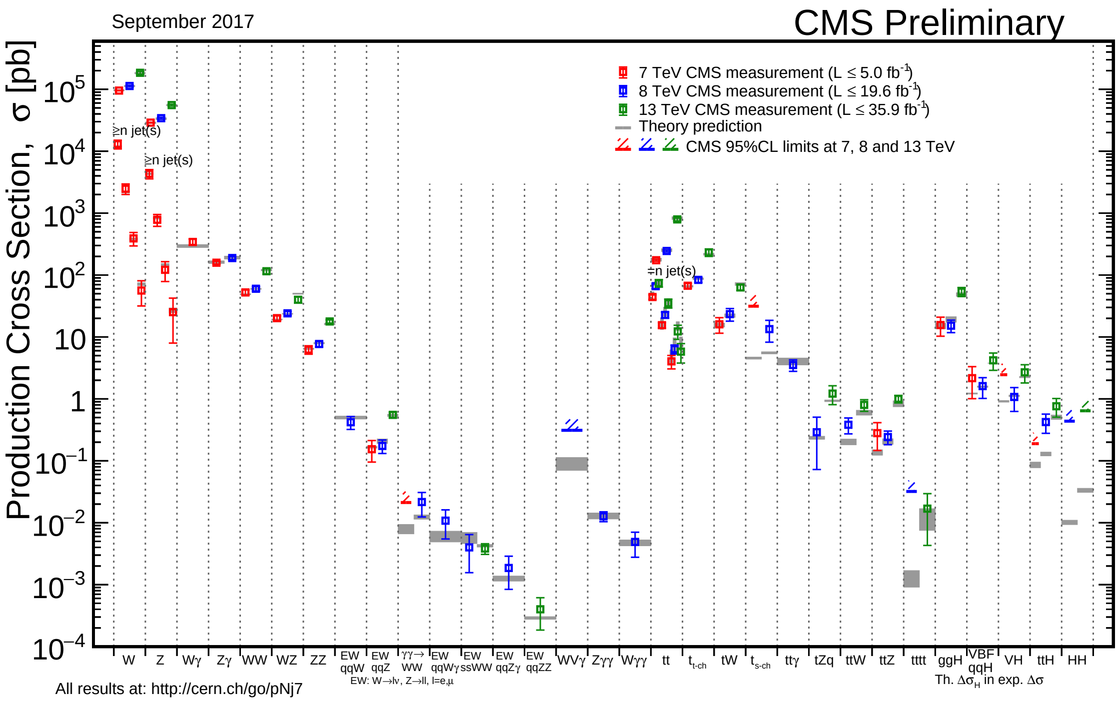

Figure 1.3: Summary of the SM cross section measurements performed by the CMS collaboration. Figure taken from [17]

.

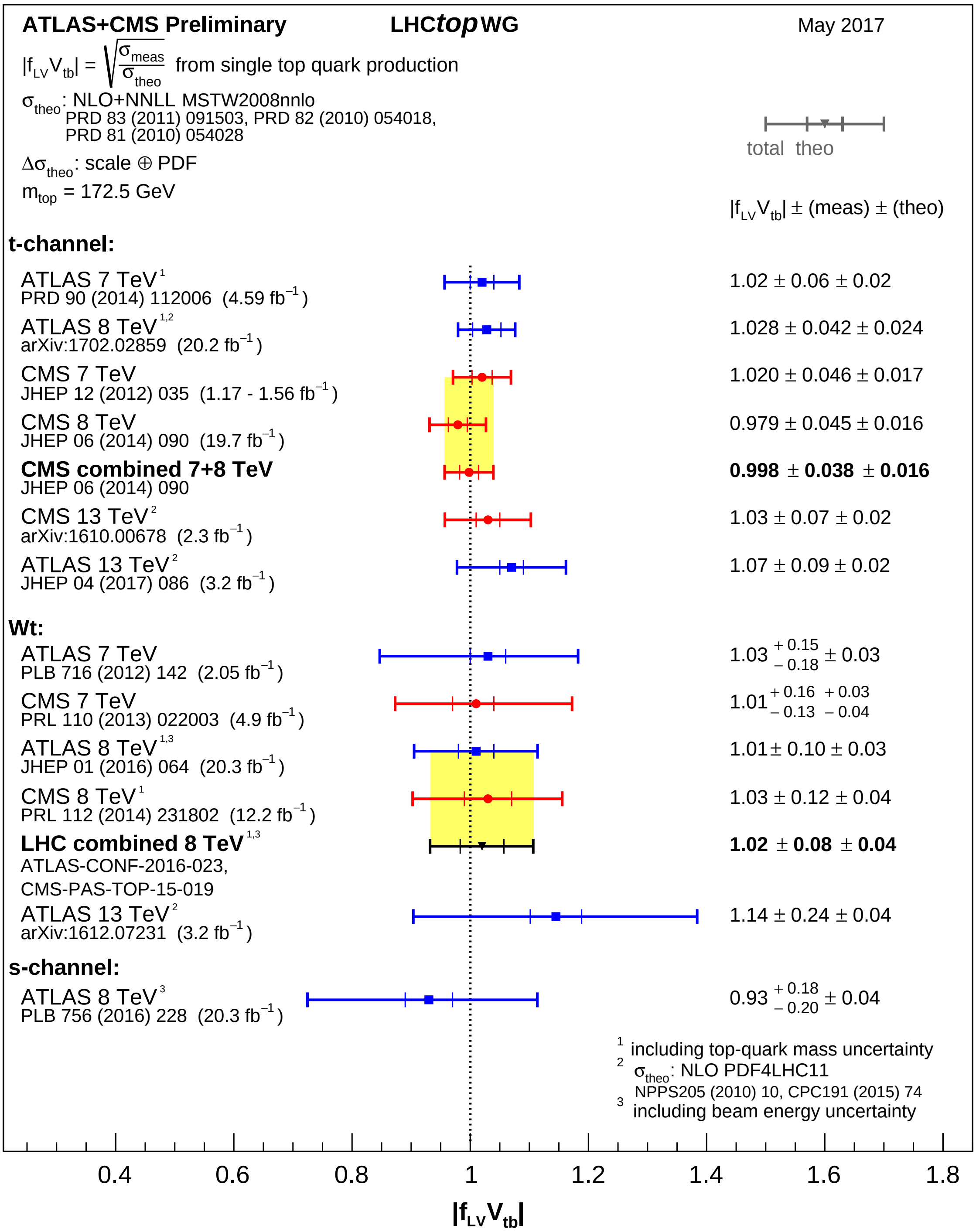

Figure 1.4: Estimations of the SM V_{tb} CKM element from single top cross section measurements. Figure taken from [18].

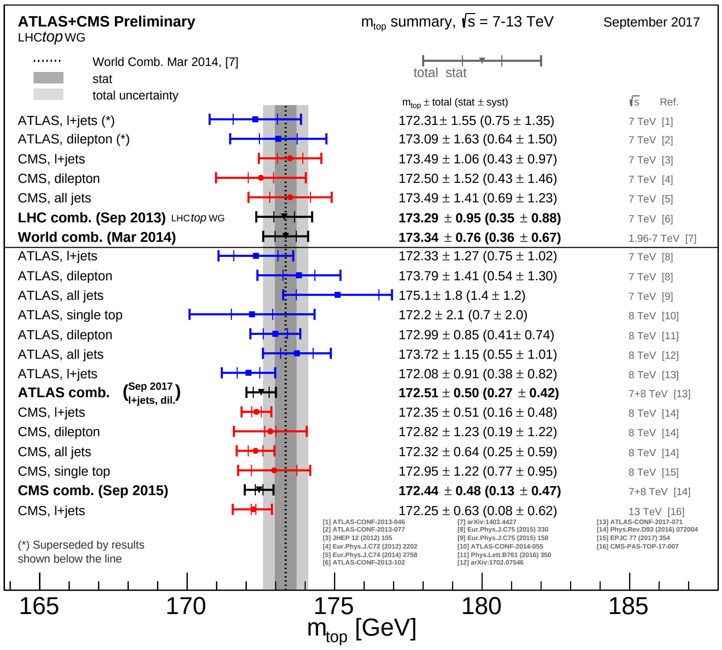

Figure 1.5: Summary of the top mass direct measurements performed by CMS and ATLAS, and compared with the LHC and LHC+Tevatron combinations. The results below the line are produced after the LHC and LHC+Tevatron combinations. Figure taken from [18].

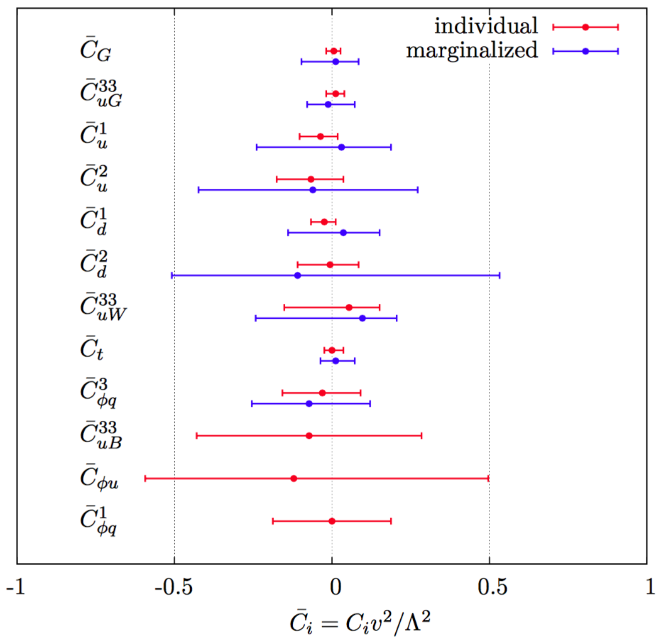

Figure 1.6: Global fit results of top quark effective field theory to experimental data including all constrained operators at dimension six. For the operators, the Warsaw basis of [19] is used. The bounds are set on the Wilson coëfficients of various operators contributing to top quark production and decay in two cases (red) all other coëfficients set to zero, or (blue) all other coëfficient are marginalised over. Figure taken from [20].

1.6 Why to look beyond the SM

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

In the SM, the neutrino is assumed to be massless, whilst experiments with solar, atmospheric, reactor and accelerator neutrinos have established that neutrinos can oscillate and change flavour during flight [21, 22]. These oscillations are only possible when neutrino's have masses. The flavour neutrinos (ν_{e}, ν_{μ}, ν_{τ}) are then linear expressions of the fields of at least three mass eigenstate neutrinos $\nu\ _1$, $\nu\ _2$, and $\nu\ _3$.

The ordinary or baryonic matter described by the SM describes only 5% of the mass (energy) content of the universe. Astrophysical evidence indicated that dark matter is contributing to approximately 27%, and dark energy to 68% of the content of the universe. From the measurements of the temperature and polarizations anisotropies of the cosmic microwave background by the Planck experiment [23], the density of cold non baryonic matter is determined. Cold dark matter is assumed to be only sensitive to the weak and gravitational force, leading to only one possible SM candidate: the neutrino. However, these are too light to account for the vast amount of dark matter and other models are needed. Dark energy is assumed to be responsible for the acceleration in the expansion of the universe [24].

At the Big Bang matter and antimatter is assumed to be produced in equal quantities. However, it is clear that we are surrounded by matter. So where did all the antimatter go? In 1967, Sakharov identified three mechanisms that are necessary to obtain a global matter antimatter asymmetry [25]. These mechanisms are those of baryon and lepton number violation, that at a given moment in time there was a thermal imbalance for the interactions in the universe, and there is charge C and charge parity CP violation[3].

The large numbers of free parameters in the SM are taken as nine fermion masses, three CKM mixing angles and one CP violating phase, one EM coupling constant g', one weak coupling constant g, one strong coupling constant g_s, one QCD vacuum angle, one vacuum expectation value, and one mass of the scalar boson. This large number of free parameters lead to the expectation of a more elegant, general theory beyond the SM.

The hierarchy problem [26] is related to the huge difference in energy between the weak scale and the Planck scale. The vev of the Brout-Englert-Higgs field determines the weak scale that is approximately 246 GeV. The radiative corrections to the scalar boson squared mass m_{H}^2, coming from its self couplings and couplings to fermions and gauge bosons, are quadratically

[3] The rate of a process $i \to f$ can be different from the CP-conjugate process: $\tilde{i} \to \tilde{f}$. 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.

proportional to the ultraviolet momentum cut-off Λ_{UV}. This cut-off is at least equal to the energy to which the SM is valid without the need of new physics. The SM is valid up to the Planck mass making the correction to m_{H}^2 about thirty orders of magnitude larger than m_{H}^2. This implies that an extraordinary cancellation of terms should happen. This is also known as the naturalness problem of the H boson mass.

The correction to the squared mass of the scalar boson coming from a fermion f, coupling to the scalar field ϕ with a coupling λ_{f} is given by

$$\Delta m_{\mathrm{H}}^2 = -\frac{|\lambda_{\mathrm{f}}|^2}{8\pi^2}\Lambda_{\mathrm{UV}}^2, \tag{1.18}$$

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 $-\lambda_{mathrmS}|\phi|^2|\mathrm{S}|^2$ is

$$\Delta m_{\mathrm{H}}^2 = -\frac{|\lambda_{\mathrm{S}}|^2}{16\pi^2}\left(\Lambda_{\mathrm{UV}}^2 - 2m_{\mathrm{S}}^2 \ln\left(\frac{\Lambda_{\mathrm{UV}}}{m_{\mathrm{S}}}\right) + \ldots\right). \tag{1.19}$$

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

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

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

The closeness of the top mass to the electroweak scale led physicist to believe that it is a sensitive probe for new physics. Its property study is therefore an important topic of the experimental program at the LHC. Several extensions of the SM enhance the FCNC branching ratios and can be probed at the LHC [12], from which some of them are shown in Table 1.7. Previous searches have been performed at the Fermilab Tevatron by the CDF [27] and D0 [28] collaborations, and at the LHC by the ATLAS [29–32] and CMS [33–37] collaborations.

The impact of BSM models can written in a model independent way by means of an effective field theory valid up to an energy scale Λ. The leading effects are parametrized by a set of fully gauge symmetric dimension-6 operators that are added to the SM Lagrangian and can be reduced to a minimal set of operators as discussed in [38, 39]. The full Lagrangian, neglecting neutrino physics, in the fully gauge symmetric case is given by

$$\mathscr{L}_{\mathrm{SM+EFT}} = \mathscr{L}_{\mathrm{SM}} + \sum_{\mathrm{i}} \frac{\bar{c}_{\mathrm{i}}}{\Lambda^2}\mathscr{O}_{\mathrm{i}} + \mathscr{O}\left(\frac{1}{\Lambda^3}\right), \tag{1.20}$$

Table 1.7: The predicted branching ratios $\mathscr{B}$ for FCNC interactions involving the top quark in some BSM models [12]: 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
$t \to uZ$	$\leq 1.1\ 10^{-4}$	–	$\leq 2\ 10^{-6}$	$t \to cZ$	$\leq 1.1\ 10^{-4}$	$\leq 10^{-7}$	$\leq 2\ 10^{-6}$
$t \to u\gamma$	$\leq 7.5\ 10^{-9}$	–	$\leq 2\ 10^{-6}$	$t \to c\gamma$	$\leq 7.5\ 10^{-9}$	$\leq 10^{-6}$	$\leq 2\ 10^{-6}$
$t \to ug$	$\leq 1.5\ 10^{-7}$	–	$\leq 8\ 10^{-5}$	$t \to cg$	$\leq 1.5\ 10^{-7}$	$\leq 10^{-4}$	$\leq 8\ 10^{-5}$
$t \to uH$	$\leq 4.1\ 10^{-5}$	$\leq 5.5\ 10^{-6}$	$\leq 10^{-5}$	$t \to cH$	$\leq 4.1\ 10^{-5}$	$\leq 10^{-3}$	$\leq 10^{-5}$

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. After electroweak symmetry breaking the operators induce [12, 40] 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

NOTE: At something about Warsaw basis

$$\mathscr{L}^{t}_{\mathrm{EFT}} = \frac{\sqrt{2}}{2} \sum_{q=u,c} \Bigg[g' \frac{\kappa_{t\gamma q}}{\Lambda} A_{\mu\nu} \bar{t} \sigma^{\mu\nu} \left(f^{L}_{\gamma q} P_L + f^{R}_{\gamma q} P_R \right) q \tag{1.21}$$

$$+ \frac{g}{2\cos\theta_W} \frac{\kappa_{tZq}}{\Lambda} Z_{\mu\nu} \bar{t} \sigma^{\mu\nu} \left(f^{L}_{Zq} P_L + f^{R}_{Zq} P_R \right) q \tag{1.22}$$

$$+ \frac{\sqrt{2} g}{4\cos\theta_W} \zeta_{tZq} \bar{t} \gamma^{\mu} \left(\tilde{f}^{L}_{q} P_L + \tilde{f}^{R}_{q} P_R \right) q Z_{\mu} \tag{1.23}$$

$$+ g_S \frac{\kappa_{gqt}}{\Lambda} Z_{\mu\nu} \bar{t} \sigma^{\mu\nu} \left(f^{L}_{gq} P_L + f^{R}_{gq} P_R \right) q G^{a}_{\mu\nu} \tag{1.24}$$

$$+ \eta_{Hqt} \bar{t} \left(\hat{f}^{L}_{q} P_L + \hat{f}^{R}_{q} P_R \right) q H + \mathrm{h.c.} \Bigg], \tag{1.25}$$

where the the value of the FCNC couplings at scale Λ are represented by $\kappa_{tZq}, \kappa_{gqt}, \kappa_{t\gamma q}, \zeta_{tZq}$, 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 $|f^{L}_{xq}|^2 + |f^{R}_{xq}|^2 = 1$, $|\tilde{f}^{L}_{q}|^2 + |\tilde{f}^{R}_{q}|^2 = 1$, and $|\hat{f}^{L}_{q}|^2 + |\hat{f}^{R}_{q}|^2 = 1$. In the expression for $\mathscr{L}^{t}_{\mathrm{EFT}}$, the unitary gauge is adopted and the scalar field is expanded around its vacuum expectation value with H being the SM scalar boson, and the field strength tensors of the photon A_{μ}, the gluon field $G^{1\ldots 8}_{\mu}$, and the Z boson Z^{0}_{μ} are defined as

$$A_{\mu\nu} = \partial_{\mu} A_{\nu} - \partial_{\nu} A_{\mu},\ Z_{\mu\nu} = \partial_{\mu} Z_{\nu} - \partial_{\nu} Z_{\mu},\ \text{and}\ G_{\mu\nu} = \partial_{\mu} G^{a}_{\nu} - \partial_{\nu} G^{a}_{\mu} + g_S f^{a}_{bc} G^{b}_{\mu} G^{c}_{\nu}. \tag{1.26}$$

Denoting the structure constant of the $SU_C(3)$ group as f^{a}_{bc}. Note that there are two coupling constants arising in $\mathscr{L}^{t}_{\mathrm{EFT}}$, which is a residue of electroweak symmetry breaking. The massive Z boson will appear in both the Z^{0}_{μ} field as well as the covariant derivative , leading to an extra Z-vertex.

NOTE: should I explain cov. der?

1.8 The top-FCNC constrained

Experiments commonly put limits on the branching ratio's which allow an easier interpretation across different EFT models by use of the branching ratio $\mathscr{B}$

$$\mathscr{B}(t \to qX) = \frac{\delta^2_{tXq}\Gamma_{t\to qX}}{\Gamma_t}, \tag{1.27}$$

where $\Gamma_{t\to qX}$ represents the FCNC decay width[4] for a coupling strength $\delta^2_{tXq} = 1$, and Γ_t the full decay width of the top quark. In the SM, supposing a top quark mass of 172.5 GeV, the full width becomes $\Gamma_t^{SM} = 1.32$ GeV [41].

Searches for top-FCNC usually adopt a search strategy depending on the experimental set-up and the FCNC interaction of interest, looking either for FCNC interactions in the production of a single top quark or in its decay for top pair interactions. In Figure 1.7, these two cases are shown for the tZq vertex.

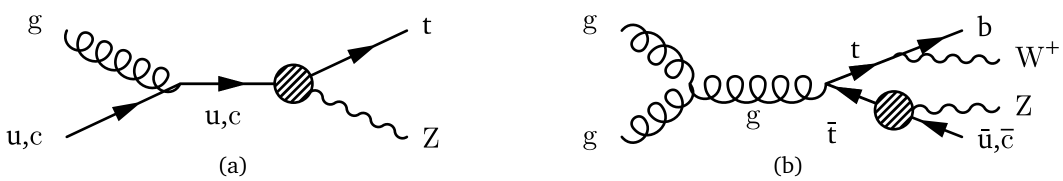

Figure 1.7: Feynman diagrams for the tZq FCNC interaction, where the FCNC interaction is indicated with the shaded dot. (a) Single top production through an FCNC interaction. (b) Top pair production with an FCNC induced decay.

The observation of top-FCNC interactions has yet to come and experiments have so far only been able to put upper bounds on the branching ratios. An overview of the best current limits is given in Table 1.8 . In Figure 1.8 a comparison is shown between the current best limits set by ATLAS and CMS with respect to several BSM model benchmark predictions. From there one can see that FCNC searches involving a Z or H boson are close to excluding or confirming several BSM theories.

NOTE: Check atlas result for tZq from top2017 proceedings when they appear

[4]The decay width of a certain process represents the probability per unit time that a particle will decay. The total decay width, defined as all possible decay widths of a particle, is inversely proportional to its lifetime.

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

Process	Search mode	Observed $\mathscr{B}$	Expected $\mathscr{B}$	Experiment
$t \rightarrow uZ$	top pair decay and single top production	$2.2\ 10^{-4}$	$2.7\ 10^{-4}$	CMS [33]
$t \rightarrow u\gamma$	single top production	$1.3\ 10^{-4}$	$1.9\ 10^{-4}$	CMS [35]
$t \rightarrow ug$	single top production	$4.0\ 10^{-5}$	$3.5\ 10^{-5}$	ATLAS [30]
$t \rightarrow uH$	top pair decay	$2.4\ 10^{-3}$	$1.7\ 10^{-3}$	ATLAS [32]
$t \rightarrow cZ$	top pair decay and single top production	$4.9\ 10^{-4}$	$12\ 10^{-4}$	CMS [33]
$t \rightarrow c\gamma$	single top production	$2.0\ 10^{-3}$	$1.7\ 10^{-3}$	CMS [35]
$t \rightarrow cg$	single top production	$2.0\ 10^{-4}$	$1.8\ 10^{-4}$	ATLAS [30]
$t \rightarrow cH$	top pair decay	$2.2\ 10^{-3}$	$1.6\ 10^{-3}$	CMS [32]

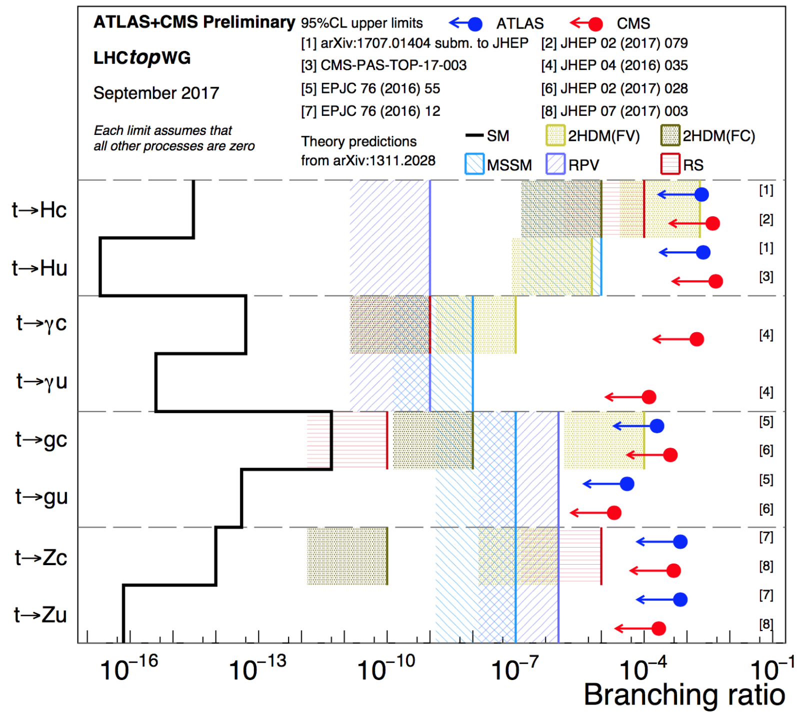

Figure 1.8: Current best limits set by CMS and ATLAS for top-FCNC interactions. Figure taken from [17].

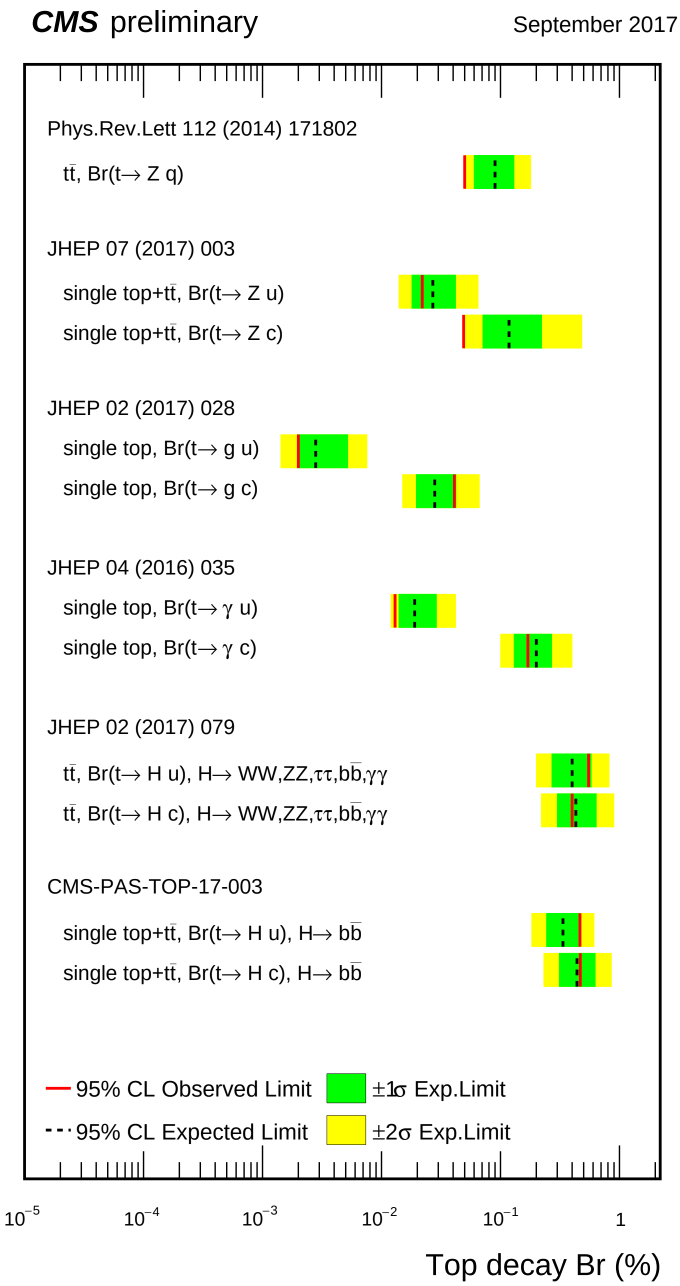

Figure 1.9: Summary of the FCNC branching ratios from CMS searches at 8 TeV. Figure taken from [17].

2 Experimental set-up

The main objective of the Large Hadron Collider (LHC) was the search for the Brout-Englert-Higgs boson. The Large Electron Positron (LEP) [42] and Tevatron [43] experiments had established that the mass of the scalar boson has to be larger than 114 GeV [44, 45], and smaller than approximate 1 TeV due to unitarity and perturbativity constraints [46]. On top of this, the search for new physics such as supersymmetry or the understanding of dark matter were part of the motivation for building the LHC. Since the start of its operation, the LHC is pushing the boundaries of the Standard Model, putting the most stringent limits on physics beyond the Standard Model as well as precision measurements of the parameters of the Standard Model. A milestone of the LHC is the discovery the scalar boson in 2012 by the two largest experiments at the LHC [5, 6].

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

2.1 The Large Hadron Collider

The LHC has started its era of cutting edge science on 10 September 2008 [47] after approval by the European Organisation of Nuclear Research (CERN) in 1995 [48]. Installed in the previous LEP tunnels, the LHC consists of a 26.7 km ring, that is installed between 45 and 170 m under the French-Swiss border amidst Cessy (France) and Meyrin (Switzerland). Built to study rare physics phenomena at high energies, the LHC can accelerate two type of particles, protons or ions Pb^{45+}, and provides collisions at four interaction points, where the particle bunches are crossing. Experiments for studying the collisions are installed on each interaction point.

As can be seen in Figure 2.1, the LHC is last element in a chain that creates, injects and accelerates protons. The starting point is the ionisation of hydrogen, creating protons that are injected in a linear accelerator (LINAC 2). Here, the protons obtain an energy of 50 MeV. They 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 radiation[1] 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) [49]. 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).

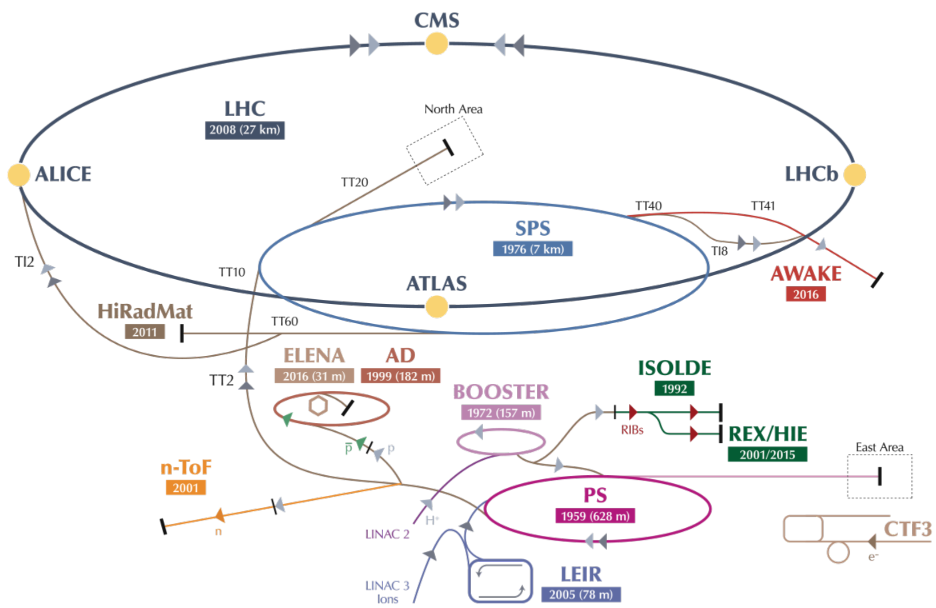

Figure 2.1: Schematic representation of the accelerator complex at CERN [50]. The LHC (dark blue) is the last element in chain of accelerators. Protons are successively accelerated by LINAC 2, the Booster, the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS) before entering the LHC.

Inside the LHC ring [51], 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

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

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

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

- A Toroidal LHC ApparatuS (ATLAS) [52] and the Compact Muon Solenoid (CMS) [53] experiments are the two general purpose detectors at the LHC. They both have a hermetic, cylindrical structure and were designed to search for new physics phenomena along with precision measurements of the Standard Model. The existence of two distinct experiments allows cross-confirmation of any discovery.
- A Large Ion Collider Experiment (ALICE) [54] and the LHC Beauty (LHCb) [55] experiments are focusing on specific phenomena. ALICE studies strongly interacting matter at extreme energy densities where a quark-gluon plasma forms in heavy ion collisions (Pb-Pb or p-Pb). LHCb searches for differences between matter and antimatter with the focus on b physics..
- The forward LHC (LHCf) [56] and the TOTal cross section, Elastic scattering and diffraction dissociation Measurement (TOTEM) [57] experiments are two smaller experiments that focus on head on collisions. LHCf consists of two parts placed before and after ATLAS and studies particles created at very small angles. TOTEM is placed in the same cavern as CMS and measures the total proton-proton cross section and studies elastic and diffractive scattering.
- The Monopoles and Exotics Detector At the LHC (MoEDAL) [58] experiment is situated near LHCb and tries to find magnetic monopoles.

For the enhancement of the exploration of rare events and thus enhancing the number of collisions, high beam energies as well as high beam intensities are required. The luminosity [59] 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

$$N_{\mathrm{event}} = L\sigma_{\mathrm{event}}, \tag{2.1}$$

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

$$L = \frac{1}{4\pi} N_{\mathrm{b}} n_{\mathrm{b}} f_{rev} \frac{N_{\mathrm{b}}}{\epsilon_n} \left(1 + \left(\frac{\theta_c \sigma_z}{2\sigma^*}\right)^2\right)^{\frac{-1}{2}} \frac{\gamma_{\mathrm{r}}}{\beta^*}. \tag{2.2}$$

The number of particles per bunch is expressed by N_{b}, while n_{b} is the number of bunches per beam, f_{rev} the revolution frequency, γ_{r} the relativistic gamma factor, ϵ_n the normalized transverse beam emittance - a quality for the confinement of the beam , β^* the beta function at the collision point - a measurement for the width of the beam, θ_{c} the angle between two beams at the interaction point, σ_{z} the mean length of one bunch, and σ^* the mean height of one bunch.

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

The peak design luminosity for the LHC reached in 2016 is $10^{34}\,\mathrm{m}^{-2}\mathrm{s}^{-1}$, which leads to about 1 billion proton interactions per second. In 2016, the LHC was around 10% above this design luminosity [60]. The luminosity is not a constant in time since it diminishes due to collisions between the beams, and the interaction of the protons and the particle gas that is trapped in the centre of the vacuum tubes due to the magnetic field. The diffusion of the beam degrades the emittance and therefore also the luminosity. For this reason, the mean lifetime of a beam inside the LHC is around 15 h. The integrated luminosity - the luminosity provided in a certain time range - recorded by CMS and ATLAS over the year 2016 is given in Figure 2.2. In Run 2, the peak luminosity is 13-17 $10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ compared to 7.7 $10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ in Run 1.

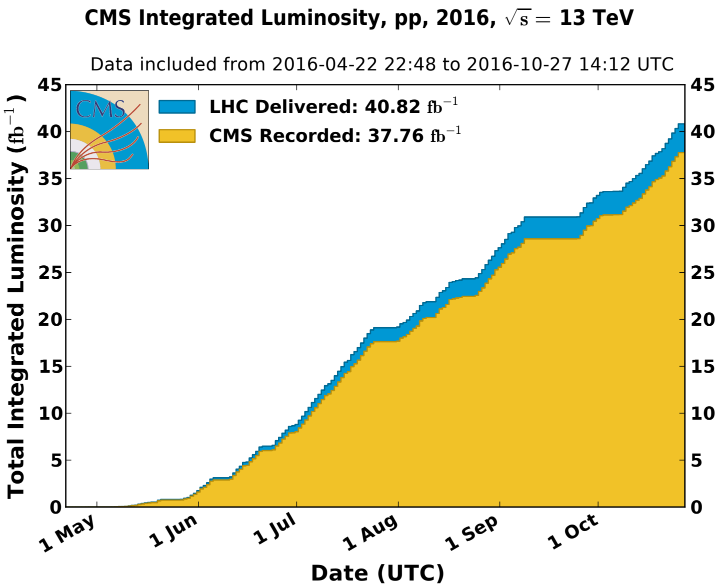

Figure 2.2: Cumulative offline luminosity measured versus day delivered to (blue), and recorded by CMS (orange) during stable beams and for proton collisions at 13 TeV centre-of-mass energy in 2016. The delivered luminosity accounts for the luminosity delivered from the start of stable beams until the LHC requests CMS to turn off the sensitive detectors to allow a beam dump or beam studies [61].

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

2.2 The Compact Muon Solenoid

At one of the collision points of the LHC, the CMS detector[62–64] is placed. Weighing 14 000 t, 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 energy[2] were kept into account while designing CMS. In Figure 2.3, an overview of the CMS detector is shown.

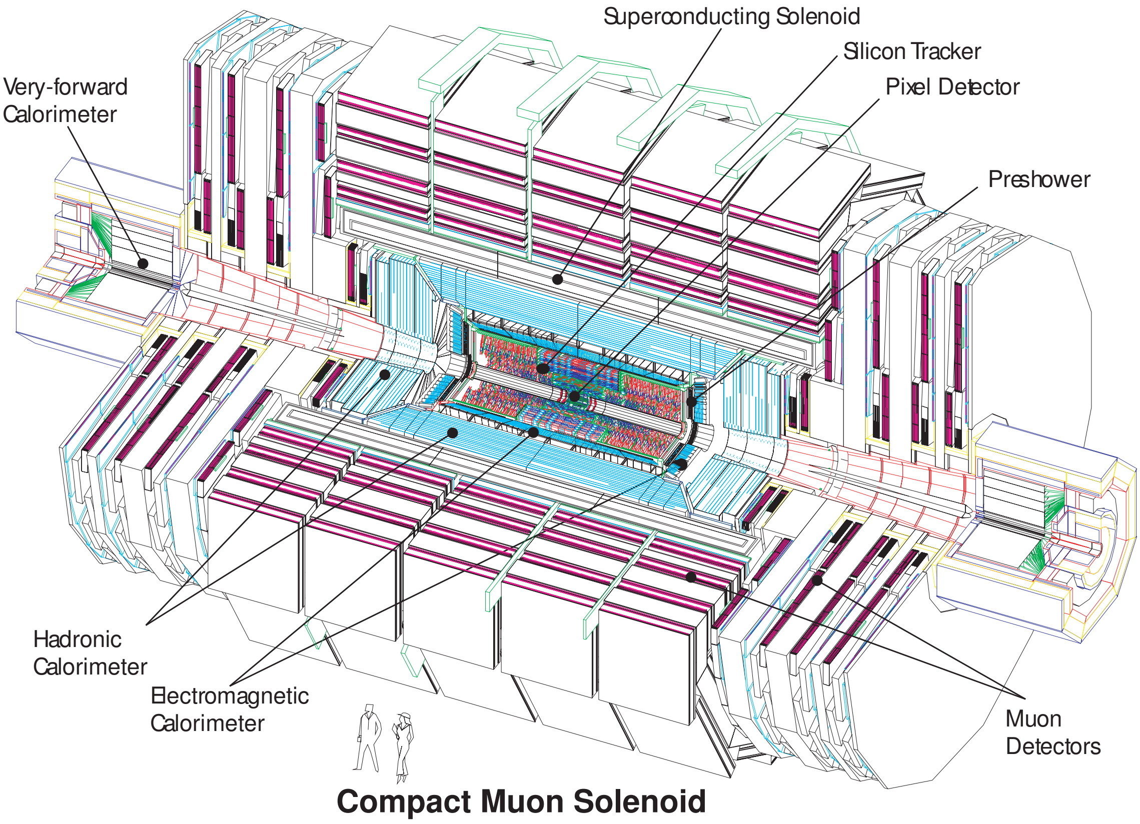

Figure 2.3: Mechanical layout of the CMS detector. Figure taken from [65].

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

[2] The 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 angle[3] $\phi \in [-\pi, \pi]$, the pseudo-rapidity[4] η :

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

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

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

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

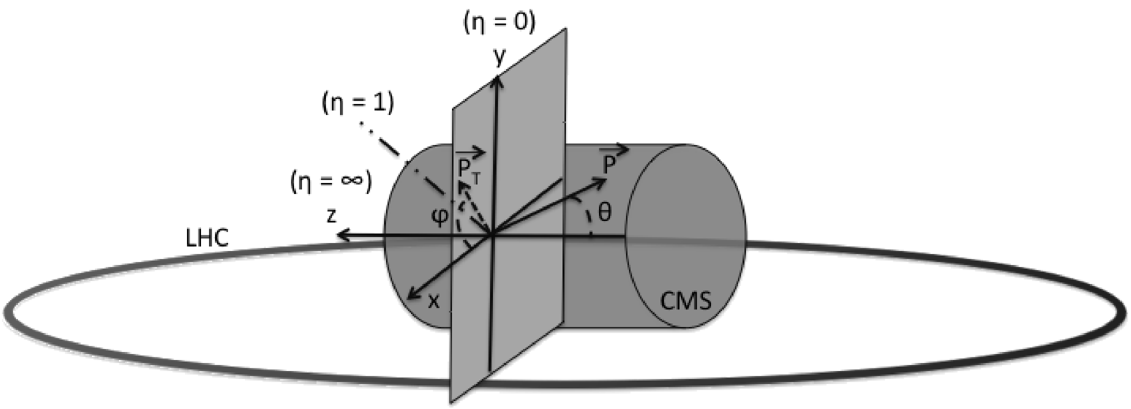

Figure 2.4: Representation of the coordinate system used by CMS. The point of origin is put at the collision point. The x-axis points towards the centre of the LHC ring such that the z-axis lies tangent to the beam axis.

2.2.2 Towards the heart of CMS

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

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

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

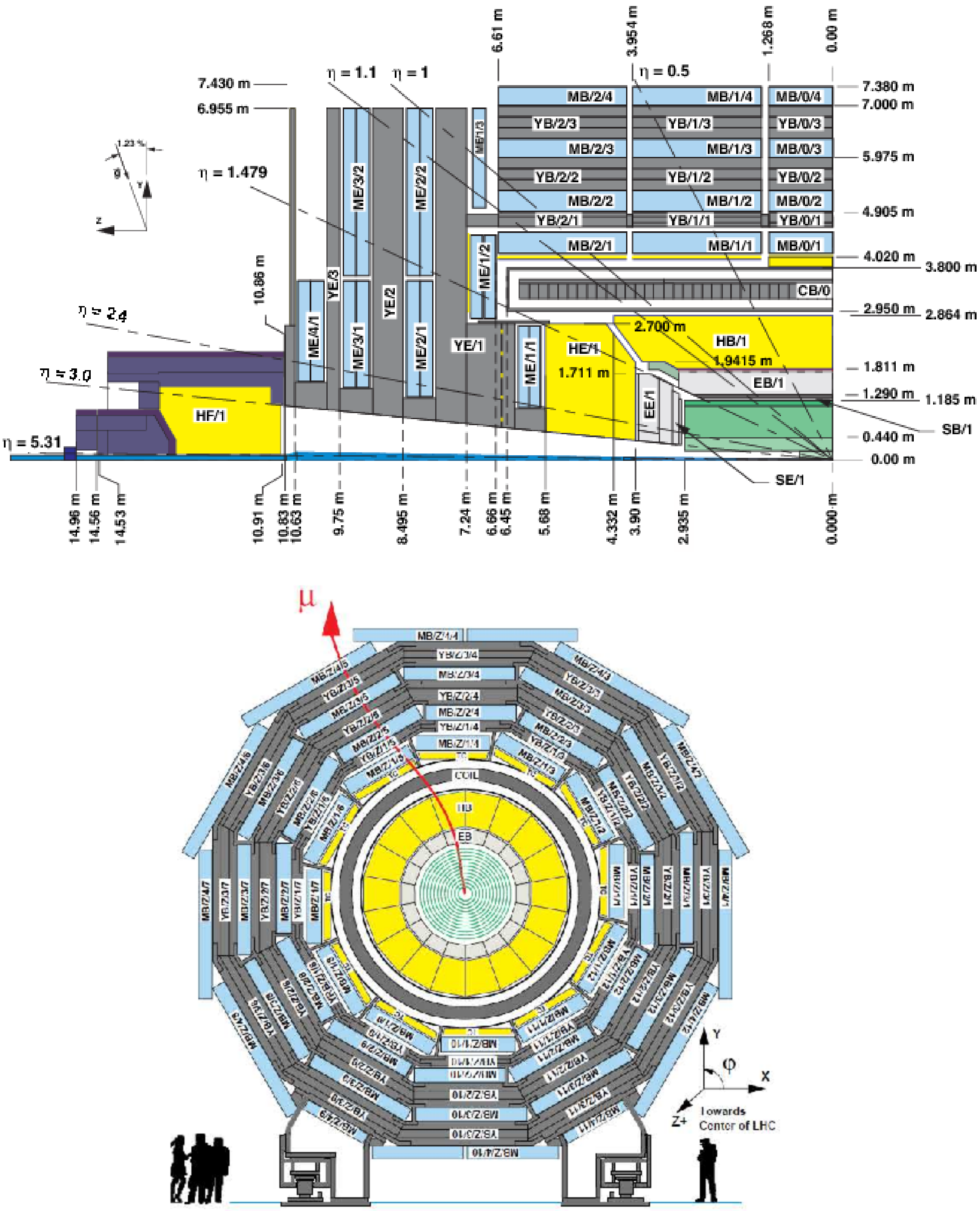

Figure 2.5: Schematic view of the CMS detector in the Run 1 configuration. The longitudinal view of one quarter of the detector is given on top, while the transversal view is shown on the bottom. The muon system barrel elements are denoted as MBZ/N/S, where z= −2...+2 is the barrel wheel number, n= 1...4 the station number and S= 1...12 the sector number. Similary, 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 electromagnet calorimeter as EE (end cap)/EB (barrel). The green part represents the tracking system (tracker + pixel). Figure taken from [66].

2.2.2.1 Muon system

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

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

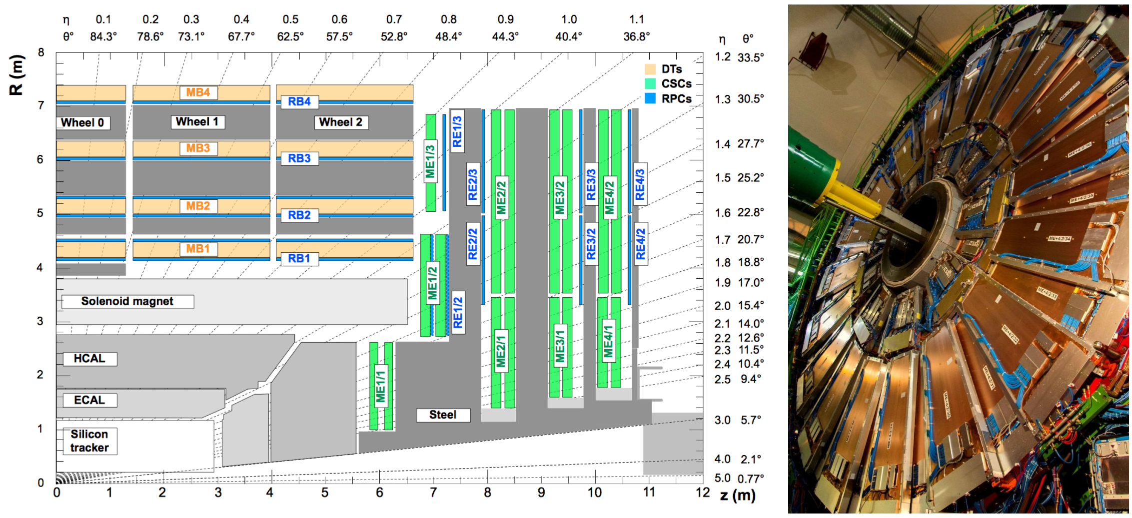

Figure 2.6: (Left) Schematic view of one quarter of the CMS muon system in the Run 1 configuration. The cathode strip chambers (CSC) are shown in green, the drift tubes (DT) are shown in yellow, while the resistive plate chambers (RPC) are shown in blue. Figure taken from [66]. (Right) Cathode strip chambers (ME+4/2 chambers on YE+3). Photo taken from [67].

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

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

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

The muon system provides triggering on muons, identifying muons and improves the momentum measurement and charge determination of high p_T muons. On top of the muon system, the muon energy is deposited in the electromagnetic calorimeter, the hadronic calorimeter, and outer calorimeter. The high magnetic field enables an efficient first level trigger and allows a 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. There is an efficient muon measurement up to $|\eta| < 2.4$.

NOTE: check numbers for run 2

2.2.2.2 Solenoid

Making use of the knowledge of previous experiments like ALEPH and DELPHI at LEP and H1 at HERA, CMS choose for a large super conducting solenoid with a a length of 12.9 m and a inner bore of 5.9 m [64]. With 2 168 turns, a current of 19.5 kA and a total energy of 2.7 GJ, a large bending power can be obtained for a modestly-sized solenoid. In order to ensure a good momentum resolution in the forward regions, a favourable length/radius was necessary. In Figure 2.7, a photo of the CMS solenoid is shown.

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

2.2.2.3 Hadronic calorimeter

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

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

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

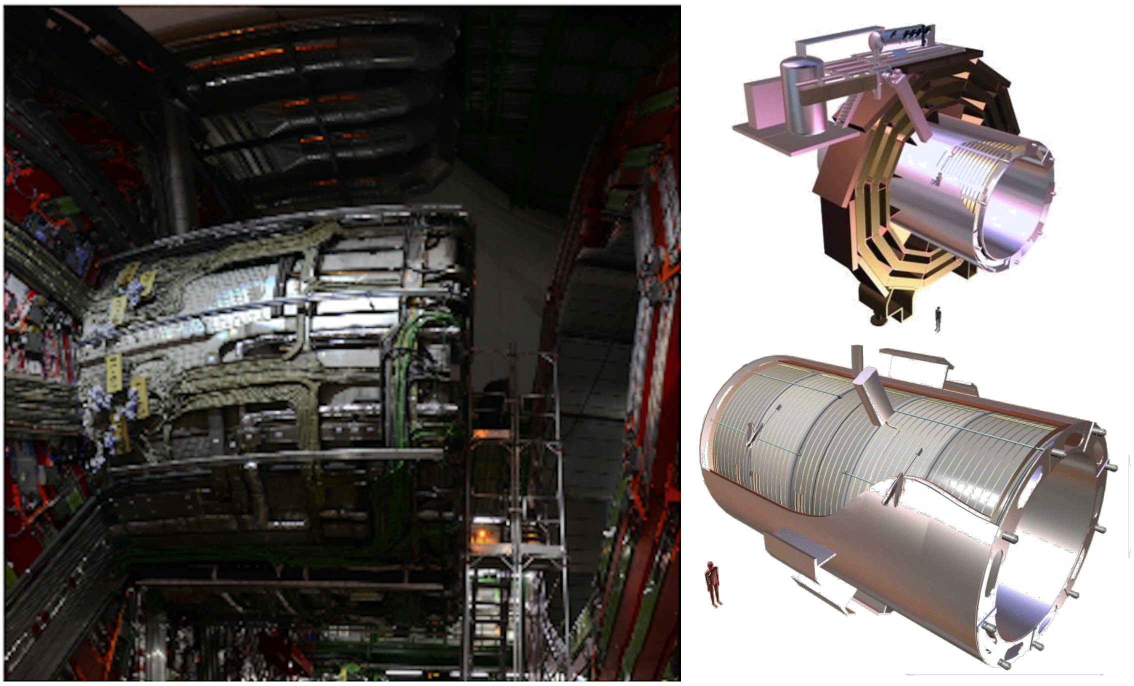

Figure 2.7: (Left) CMS solenoid during the long shutdown in 2013. (Right) An impression of the solenoid magnet taken from [68].

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

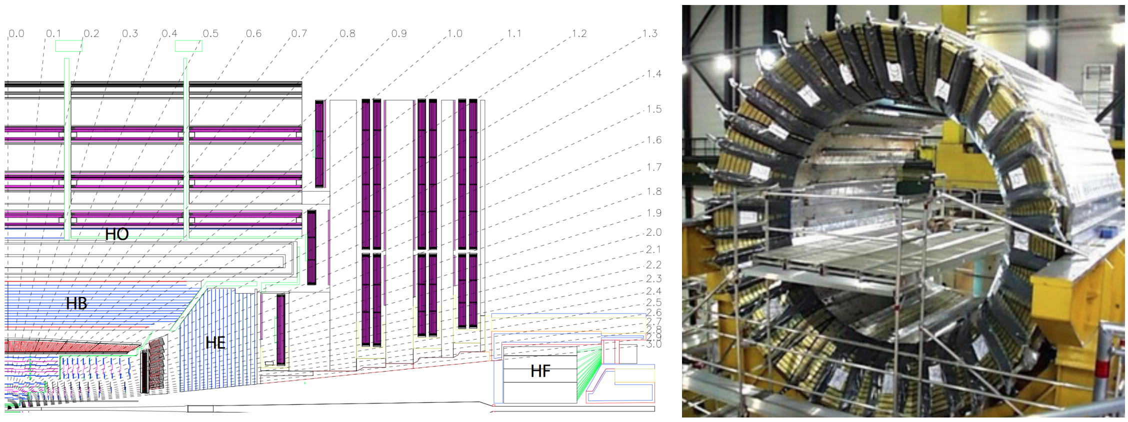

Figure 2.8: (Left) Longitudinal view of the CMS detector showing the locations of the HB, HE, HO, and HF calorimeters. Figure taken from [53]. (Right) CMS barrel calorimeter. Photo taken from [69].

The HB is made of 16 absorber plates where most of them are built from brass and others are made from stainless steal and is about five to ten interaction lengths thick. It is divided in $\eta \times \phi$ towers and contains 2592 read out channels. The HO complements the HB and extends the 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 [70]. The HF represents 1728 read out channels.

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

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.

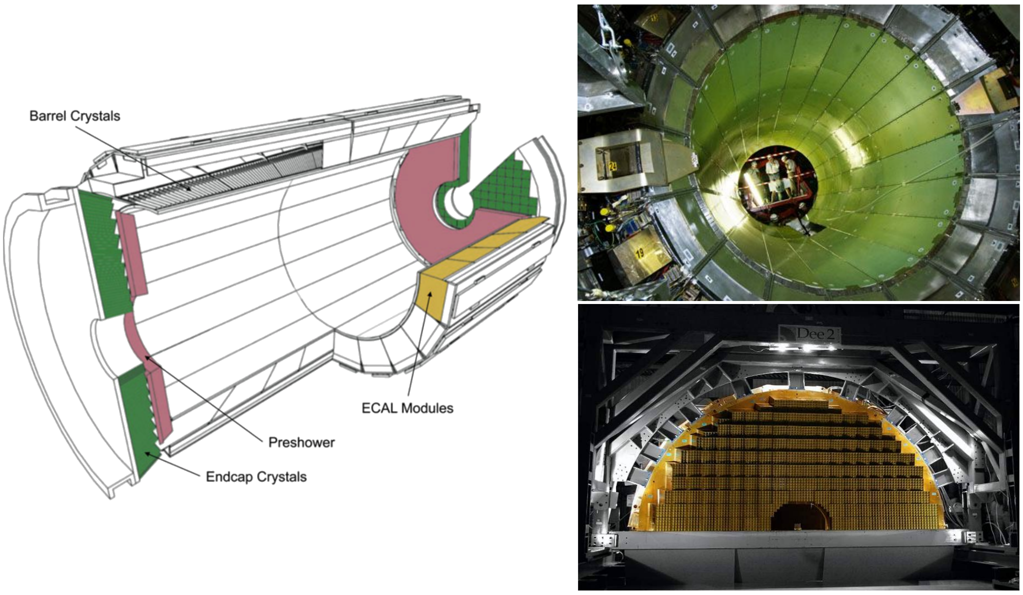

Figure 2.9: (Left) Schematic cross section of the electromagnetic calorimeter taken from [53]. (Right top) The ECAL barrel during construction [71]. (Right bottom) One half of an EE [72].

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 $1.479 <$

$|\eta| < 3.0$, are the EE. They consist of semi-circular aluminium plates from which structural units of 5×5 crystals (super crystals) are supported. The ES is placed in front of the crystal calorimeter over the end cap pseudo rapidity range with two planes of silicon strip detectors as active elements.

The electromagnetic shower will typically involve more than one channel. More than 90% of the energy of a 35 GeV electron or photon is contained in a 5×5 matrix of crystals. Therefore, a clustering algorithm is performed in order to associate the energy deposits to the particles impinging the calorimeter. The achieved precision [73] 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 [53]. The energy resolution is given by

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

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

2.2.2.5 Inner tracking system and operations

The tracking system (tracker) [74] is the detecting unit closest to the point of interaction. Responsible for the reconstruction of trajectories from charged particles with $|\eta| < 2.5$ that are bent by the magnetic field, it provides a measurement of the momentum. The tracker is also responsible for the determination of the interaction point or vertex. It should be able to provide high granularity as well as fast read out, and be able to endure high radiation. For this reason, the CMS collaboration choose silicon detector technology.

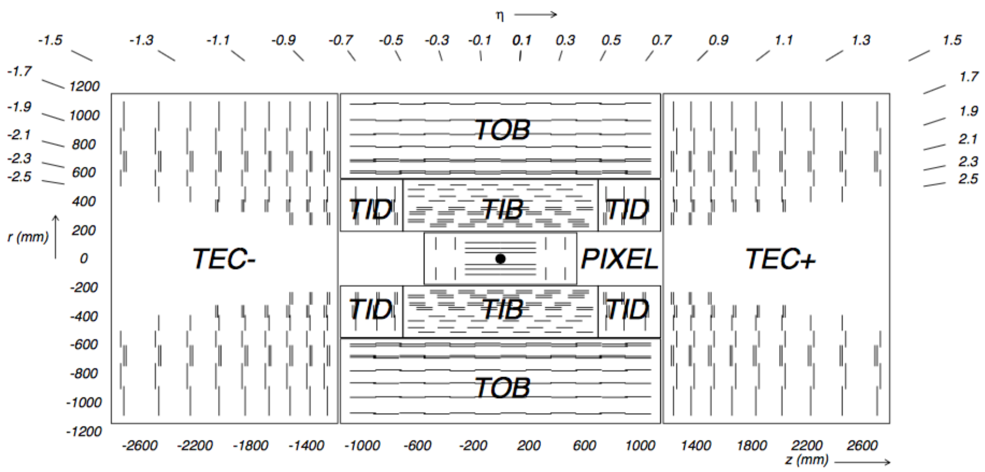

Figure 2.10: Schematic cross section through the CMS tracker. Each line represents a detector module. Double lines indicate back-to-back modules that deliver stereo hits. Figure taken from [53].

The tracking system consists of a cylinder of 5.8 m long and 2.5 m in diameter. It is immersed 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 100×150 μm^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 [75].

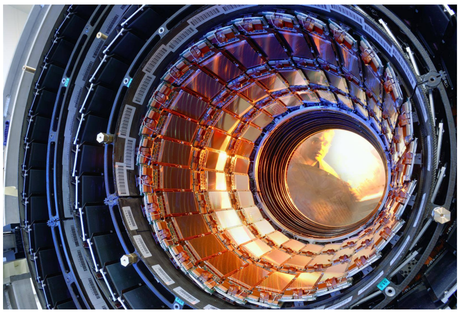

Figure 2.12: First half of the inner tracker barrel, consisting of three layers of silicon modules [76].

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 μm, while 20-40 μm 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 μm 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 80μm, while the remaining to have a strip pitch of 120 μm. This leads to a respective single point resolution of 23 μm and 35 μm. For the TID, the pitch varies between 100 μmand 141 μm. The TOB provides six $r\phi$-measurements with a single point resolutions of 53 μm in the first four layers, and 35 μm in the last two layers. It consists of 500 μmthick micor strip sensors with strip pitches of 183 μm (first 4 layers) or 122 μm (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 μmto 184 μmaverage 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 μm in the TIB and 530 μm in the TOB, and is varying with pitch in the TID and TEC. To allow overlay and avoid gaps in acceptance, each module is shifted slightly in r or z with respect to its neighbouring modules within a layer. With this detector lay out, at least nine points per charged particle trajectory can be measured in an $|\eta|$ range up to 2.4, where at least four of them being two dimensional. The CMS silicon tracker provides 9.3 million readout channels and covers an active area of about 198 m^2.

2.2.3 Data acquisition

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

CMS Level-1 Trigger

The Level-1 Trigger has to be a flexible, maintainable system, capable of adapting to the evolving physics programme of CMS [77]. Its output rate is restricted to 100 kHz imposed by the CMS readout electronics. It is implemented by custom hardware and selects events containing candidate objects - e.g. ionization deposits consistent with a muon, or energy clusters corresponding to an electron / photon / tau lepton / missing transverse energy / jet. Collisions with large momenta can be selected by using scalar sum of the transverse momenta of the jets.

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

CMS HLT Trigger

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

2.2.4 Phase 1 upgrades

Before the start of taking collision data for 13 TeV operations on 3 June 2015, CMS had a long shutdown (LS1) [78]. During this shutdown, the section of the beryllium beam pipe within CMS was replaced by a narrower one. This operation required the pixel to be removed and reinserted into CMS. In Run 2, higher particle fluxes with respect to Run 1 are expected. To avoid long damage caused by the intense particle flux at the heart of CMS, the tracker is been made ready to operate at much lower temperature than during Run 1. The electromagnetic calorimeter preshower system had been damaged during Run 1, therefore the preshower discs were removed, repaired and reinstalled successfully inside CMS in 2014. To help the discrimination between interesting low momentum muons coming from collisions and muons caused by backgrounds, a fourth triggering and measurement station for muons was added in each of the end caps. Several new detectors were installed into CMS for measuring the collision rate within the detector and monitors beam related backgrounds.

During the LS1, the muon system underwent major upgrades [79, 80]. In the fourth station of each end cap, the outermost rings of CSC and RPC chambers were completed, providing an angular coverage of $1.2 < |\eta| < 1.8$ for Run 2, increasing the system redundancy, and allowing tighter cuts on the trigger quality. In order to reduce the environmental noise, outer yoke discs have been placed on both sides for the end caps. At the innermost rings of the first station, the CSCs have been upgraded by refurbishing the readout electronics to make use of the full detector granularity instead of groups of three as was the case for Run 1. In Figure 2.6 (right), the refurbishing of the CSCs is shown.

Since the HF experiences intense particle fluxes, it became clear during Run 1 that the glass windows of the PMTs need replacing. For the ECAL in Run 1, the energy reconstruction happened via a weighted sum of the digitized samples [81]. For Run 2 however, the reconstruction had to be made more resistant for out of time pile up and a multi-fit approach has been set into place. In this approach, the pulse shape is modelled as a sum of one in-time pulse plus the out of time pulses [73]. The energy resolution is better than 2% in the central barrel region and 2-5 % elsewhere.

During the first data taking period of the LHC (2010 to 2013), the tracker operated at +4°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:

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

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

During the LS1, the CMS cooling plant was refurbished [82] and the fluorocarbon cooling system overhauled. To help to suppress the humidity inside the tracker, new methods for vapour sealing and insulation were applied. Furthermore, several hundred high-precision sensors are used to monitor the humidity and temperature. In order to get as dry air as possible, a new dry-gas plant provides eight times more dry gas (air or nitrogen) than during the first run, and allows regulation of the flow. As a final addition, the cooling bundles outside the tracker are equipped with heater wires and temperature sensors in order to maintain safe operations above the cavern dew point. For the data taking in 2015-2016, the tracker operated at −15°C.

In Run 2, with the increase in centre of mass energy and a higher luminosity, a larger number of simultaneous inelastic collisions per crossing is expected with respect to Run 1. For this, the CMS Level-1 has been upgraded [83]. All hardware, software, databases and the timing control system have been replaced for Run 2, where the main changes are that the muon system now uses the redundancy of the muon detector system earlier to make a high resolution muon trigger.

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

After the first half of Run 2, the innermost part of detection material in CMS (pixel) was upgraded by adding a fourth layer , enhancing the particle tracking capabilities of CMS. The data used in the framework of this thesis however is from before this upgrade. More information on the Pixel upgrade can be found in Refs. [84, 85].

2.2.5 CMS computing model

The selected data is stored, processed and dispersed via the Worldwide Large Hadron Collider GRID (WLCG) [86, 87]. This has a tiered structure that functions as a single, coherent system.

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

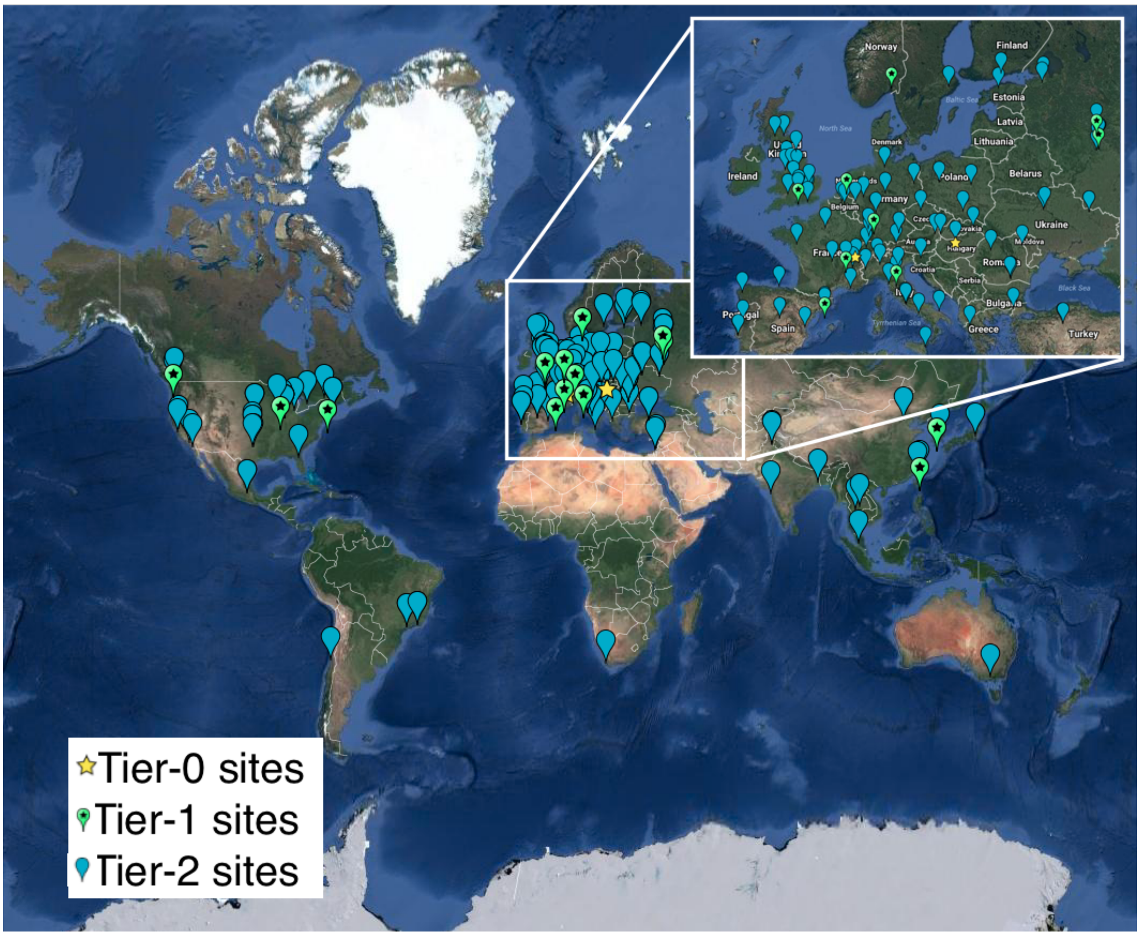

Figure 2.13: Worldwide LHC Computing Grid in 2017 [88].

Analysis techniques 3

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

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

$$\vec{p}_{\text{parton}} = x\vec{p}_{\text{hadron}}, \tag{3.1}$$

where x is referred to as the Björken scaling variable. The interaction $\mathrm{p_A p_B} \rightarrow \mathrm{X}$ can then be factorised in terms of partonic cross sections $\hat{\sigma}_{\mathrm{ij}\rightarrow\mathrm{X}}$ [89]

$$\sigma_{\mathrm{p_A p_B}\rightarrow\mathrm{X}} = \sum_{\mathrm{ij}} \iint dx_1 dx_2\, f_{\mathrm{i}}^{\mathrm{A}}(x_1,Q^2) f_{\mathrm{j}}^{\mathrm{B}}(x_2,Q^2) d\hat{\sigma}_{\mathrm{ij}\rightarrow\mathrm{X}}, \tag{3.2}$$

where i and j are the partons resolved from protons A and B, $f_{\mathrm{i}}(x_{\mathrm{i}},Q^2)$ the parton density functions (PDF), and Q^2 the factorisation scale more commonly denoted as μ_{F}. The factorisation scale is the scale at which the hadronic interaction can be expressed as a product of the partonic cross section and the process independent PDF. In Figure 3.1, the kinematic regions in x and μ_{F} are shown for fixed target and collider experiments.

The parton density functions (PDF) [90–92] give the momentum distribution of the proton amongst its partons at an energy scale μ_{F}. These function can not be determined from first principles and have to obtained from global fits to data. The PDFs are obtained from measurements on

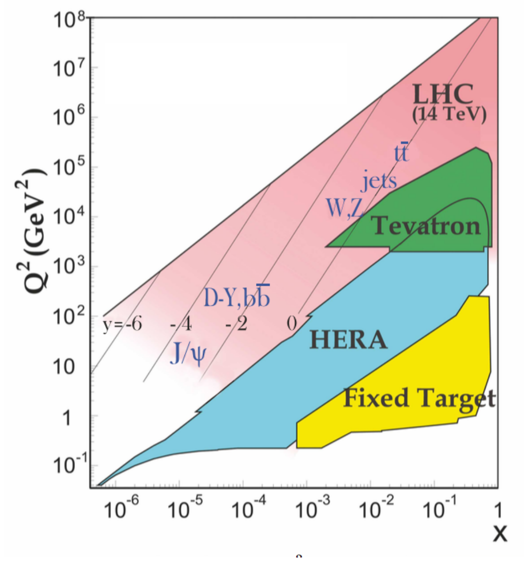

Figure 3.1: Kinematic regions in momentum fraction x and factorisation scale Q^2 probed by fixed-target and collider experiments. Some of the final states accessible at the LHC are indicated in the appropriate regions, where y is the rapidity. In this figure, the incoming partons have $x_{1,2} = (M/14\text{TeV})e^{\pm y}$ 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 [4].

deep inelastic scattering using lepton-proton collision by the HERA collider [93], supplemented with proton-antiproton collisions from Tevatron at Fermi lab [94], and proton collision data from the ATLAS, CMS and LHCb collaborations at the LHC (Run 1) [95]. These measurements are included in global PDF sets known as the PDF4LHC recommendation [92]. From their measurement at scale μ_F these PDFs can be extrapolated using the DGLAP equations [96]. The PDFs are used to calculate the cross section of a certain process and are therefore used as input for the Monte Carlo generators used to make the simulated data samples at the LHC. In the framework of this thesis, the NLO PDF4LHC15_100 set is used. This set is an envelope of three sets, CT14, MMHT2014 and NNPDF3.0 [92]. In Figure 3.2 the dependency of the PDFs on the momentum fraction x is shown for the NNPDF3.0 set on hadronic scale ($\mu_F^2 = (10\text{GeV})^2$ and LHC scale ($\mu_F^2 = (10^4\text{GeV})^2$. For most values of the momentum fraction, the gluon density dominates, meaning that it is easier to probe muons than the quarks. For x close to one, the parton densities of the up and down quarks (the valence quarks of the proton) dominate over the gluon density. The charm, anti-up, and anti-down quarks have lower densities in general since those are sea quarks which originate in the proton only through gluon splitting. The resolution scale Q^2 is typically taken to be the energy scale of the collision. For the top quark pair production a scale of $Q^2 = (350\,\text{GeV})^2$ is chosen, meaning that the centre-of-mass energy of the hard interaction is about twice the top quark mass. The uncertainty on the parton distributions is evaluated using the Hessian technique [97], where a matrix with a dimension identical to the number of free parameters needs to be diagonalised. In the case of PDF4LHC15_100 set, this translates into 100 orthonormal eigenvectors and 200 variations of the PDF parameters in the plus and minus direction.

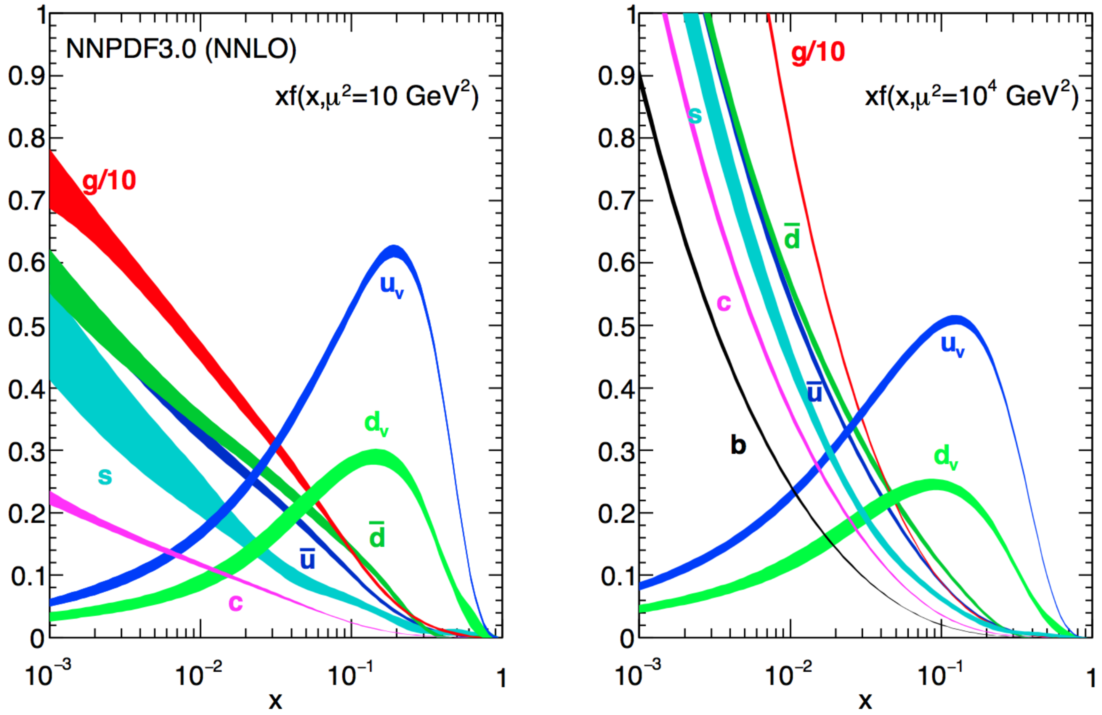

Figure 3.2: The momentum fraction x times the parton distribution functions $f(x)$, where $f = \mathrm{u_v}, \mathrm{d_v}, \bar{\mathrm{u}}, \bar{\mathrm{d}}, \mathrm{s}, \mathrm{c}$, or g as function of the momentum fraction obtained in the NNLO `NNPDF3.0` global analysis at factorisation scales $\mu^2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

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 constant[1] α_S is found to be

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

with n_f the number of quarks and μ_0 the reference scale on which the coupling is known. The current world average of the strong coupling constant at the Z boson mass is $\alpha_S(\mu_R = m_Z) = 0.1181 \pm 0.0011$ [4]. From Equation 3.3 one can see easily that the coupling strength decreases with increasing renormalization scale, this known as asymptotic freedom. Additionally, following the behaviour of $\alpha_S(\mu_R^2)$, a limit $\Lambda_{QCD} \approx 200\,\mathrm{MeV}$ is found for which α_S becomes larger than one. Under this limit, the perturbative calculations of observables can no longer be done.

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

Cross sections can be written in terms of interacting vertices contributing to the matrix element (ME) originating from elements of a perturbative series [98], allowing them to be expanded as a power series of the coupling constant α

$$\sigma = \sigma_{\mathrm{LO}} \left(1 + \left(\frac{\alpha}{2\pi} \right) \sigma_1 + \left(\frac{\alpha}{2\pi} \right)^2 \sigma_2 + \ldots \right). \tag{3.4}$$

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

3.2 Event generation

In order to compare reconstructed data with theoretical predictions, collision events are generated and passed through a simulation of the CMS detector and an emulation of its readout. For the detector simulation, a so-called Full Simulation package [99, 100] based on the `Geant4` toolkit [101] is employed. It allows a detailed simulation of the interactions of the particles with the detector material.

3.2.1 Fundamentals of simulating a proton collision

The procedure of to generate $\mathrm{pp} \to \mathrm{X}$ events can be subdivided into sequential steps [102–104], as shown in Figure 3.3.

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

NOTE: Should I add more details?

3.2.2 Programs for event generation

The `FEYNRULES` package [105] allows the calculation of the Feynman rules in momentum space for any quantum field theory model. By use of a Lagrangian, the set of Feynman rules associated with this Lagrangian are calculated. Via the `Universal Feynrules Output` (UFO) [106] the results are then passed to matrix element generators.

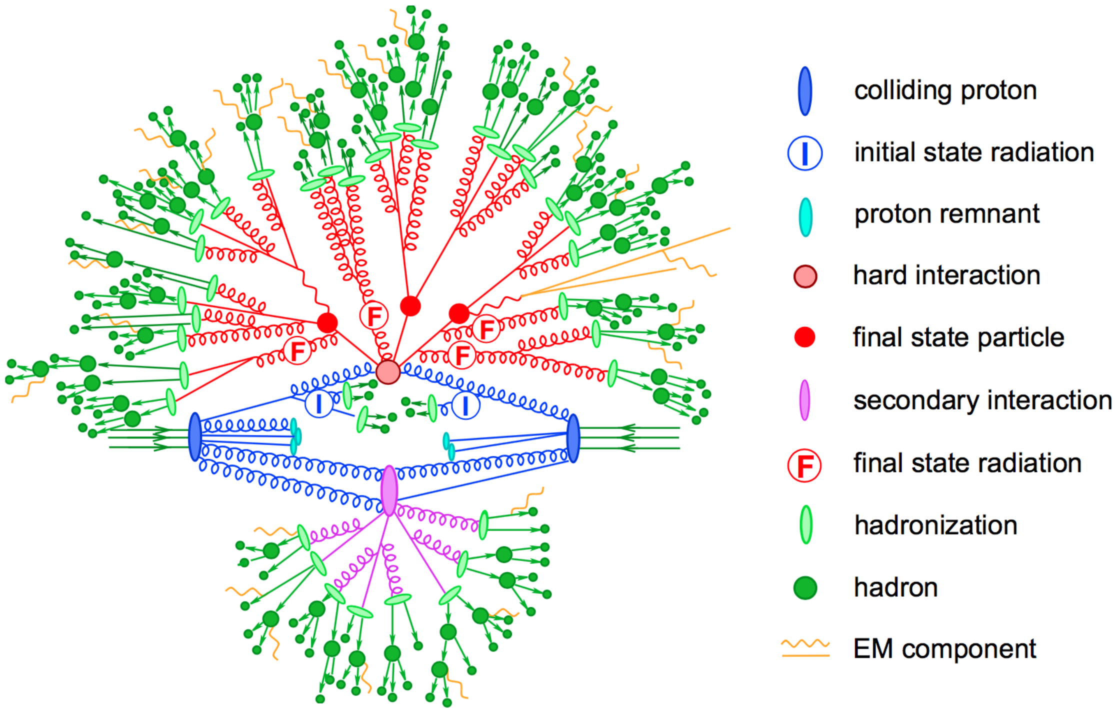

Figure 3.3: Sketch of a hadron collision as simulated by a Monte-Carlo event generator. The red blob in the centre represents the hard collision, surrounded by a tree-like structure representing Bremsstrahlung as simulated by parton showers. The purple blob indicates a secondary hard scattering event. Parton-to-hadron transitions are represented by light green blobs, dark green blobs indicate hadron decays, while yellow lines signal soft photon radiation. Figure taken from [104].

The `MadGraph` program [107] is used to interpret the physics model and calculate the corresponding Feynman diagrams and matrix elements. After this, `MadEvent` [108] is used to calculate the corresponding partons. These generated parton configurations are then merged with `Pythia` [109–111] parton showers using the MLM merging scheme [112].

The `MadGraph5_aMC@NLO` program [113] combines the LO `MadGraph` [107] and the `aMC@NLO` program into a common framework. This combination supports the generation of samples at LO or NLO together with a dedicated matching to parton showers using the MLM [112] or FXFX [114] schemes respectively. The FXFX scheme produces a certain fraction of events with negative weights originating from the subtraction of amplitudes that contain additional emissions from the NLO matrix element to prevent double-counting.

The `POWHEG` box (versions 1,2) [115–120] contains predefined implementations of various processes at NLO. It applies the `POWHEG` method for ME- to PS- matching, where the hardest radiation generated from the ME has priority over subsequent PS emission to remove the overlap with the PS simulation.

The `JHU` generator (version 7.02) [121–124] is used to generate the parton level information including full spin and polarization correlations. It is commonly used for studying the spin an parity properties of new resonances such as $ab \to X \to VV$, 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 [125] 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) [109–111] 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 [126] are the CUETP8M2T4, CUETP8M1 and CUETP8M2.

The detector response is simulated via the `Geant4` [101] 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 1.25 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 $f^{L}_{Xq} = 0$ and $f^{R}_{Xq} = 1$. 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.

NOTE: Why RH and not LH?

The anomalous single top cross sections are calculated by convolution of the hard scattering matrix elements with the NLO order set of `CTEQ6` partons densities [127]. 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 ($\sigma_{t\bar{t}} = 6.741\ 10^{2}$pb), and considering the decay $t\bar{t} \to (bW^{\pm})(Xqt)$. The branching ratio $\mathcal{B}(t \to bW^{\pm})$ is assumed to be equal to one and the FCNC branching ratio is calculated as

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

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

NOTE: these partial widths are at LO, how does this relate 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	tg u	$3.665220\ 10^5$	$\left(\kappa_{tgu}/\Lambda\right)^2$
	tg c	$3.664620\ 10^5$	$\left(\kappa_{tgc}/\Lambda\right)^2$
$\kappa_{t\gamma q}/\Lambda$	tγ u	$1.989066\ 10^4$	$\left(\kappa_{t\gamma u}/\Lambda\right)^2$
	tγ c	$1.988904\ 10^4$	$\left(\kappa_{t\gamma c}/\Lambda\right)^2$
κ_{tZq}/Λ	tZ u	$1.637005\ 10^4$	$\left(\kappa_{tZu}/\Lambda\right)^2$
	tZ c	$1.636554\ 10^4$	$\left(\kappa_{tZc}/\Lambda\right)^2$
ζ_{tZq}	tZ u	$1.685134\ 10^{-1}$	$\left(\zeta_{tZu}\right)^2$
	tZ c	$1.684904\ 10^{-1}$	$\left(\zeta_{tZc}\right)^2$
η_{tHq}	tH u	$1.904399\ 10^{-1}$	$\left(\eta_{tHu}\right)^2$
	tH c	$1.904065\ 10^{-1}$	$\left(\eta_{tHc}\right)^2$

Table 3.2: Leading order single top production cross section for $\mathrm{pp} \to$ tZ or $\bar{\mathrm{t}}$Z, where the new physics scale is given in GeV. The NLO k—factors [128] are given in the last column.

Anomalous coupling	Cross section (pb)		NLO k—factor
κ_{tgu}/Λ	$3.272\ 10^7$	$\left(\kappa_{tgu}/\Lambda\right)^2$	1.00
κ_{tgc}/Λ	$3.021\ 10^6$	$\left(\kappa_{tgc}/\Lambda\right)^2$	1.00
$\kappa_{t\gamma u}/\Lambda$	$2.260\ 10^5$	$\left(\kappa_{t\gamma u}/\Lambda\right)^2$	1.00
$\kappa_{t\gamma c}/\Lambda$	$2.654\ 10^4$	$\left(\kappa_{t\gamma c}/\Lambda\right)^2$	1.00
κ_{tZu}/Λ	$1.728\ 10^6$	$\left(\kappa_{tZu}/\Lambda\right)^2$	1.40
κ_{tZc}/Λ	$2.040\ 10^5$	$\left(\kappa_{tZc}/\Lambda\right)^2$	1.40
ζ_{tZu}	7.484	$\left(\zeta_{tZu}\right)^2$	1.40
ζ_{tZc}	1.038	$\left(\zeta_{tZc}\right)^2$	1.40

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	$t\bar{t} \to (b\ell^+\nu)(\bar{u}\ell^+\ell^-)$	$2.727008\ 10^5$	$\left(\kappa_{tZu}/\Lambda\right)^2$
	$t\bar{t} \to (\bar{b}\ell^-\bar{\nu})(u\ell^+\ell^-)$	$2.727008\ 10^5$	$\left(\kappa_{tZu}/\Lambda\right)^2$
κ_{tZc}/Λ	$t\bar{t} \to (b\ell^+\nu)(\bar{c}\ell^+\ell^-)$	2.72625710^5	$\left(\kappa_{tZc}/\Lambda\right)^2$
	$t\bar{t} \to (\bar{b}\ell^-\bar{\nu})(c\ell^+\ell^-)$	$2.726257\ 10^5$	$\left(\kappa_{tZc}/\Lambda\right)^2$
ζ_{tZu}	$t\bar{t} \to (b\ell^+\nu)(\bar{u}\ell^+\ell^-)$	2.827184	$\left(\zeta_{tZu}\right)^2$
	$t\bar{t} \to (\bar{b}\ell^-\bar{\nu})(u\ell^+\ell^-)$	2.827184	$\left(\zeta_{tZu}\right)^2$
ζ_{tZc}	$t\bar{t} \to (b\ell^+\nu)(\bar{c}\ell^+\ell^-)$	2.806801	$\left(\zeta_{tZc}\right)^2$
	$t\bar{t} \to (\bar{b}\ell^-\bar{\nu})(c\ell^+\ell^-)$	2.806801	$\left(\zeta_{tZc}\right)^2$

3.2.4 Generating SM background events

The SM tZq events were generated using the MadGraph5_aMC@NLO generator, interfaced with Pythia version 8.2 [111] for parton showering and hadronisation. The WZ+jets, $t\bar{t}Z$, tZq, and $t\bar{t}W$ samples are produced using the MadGraph5_aMC@NLO(version 5.222) [113], which includes up to one hadronic jet at next to leading order (NLO) QCD accuracy. Other minor background (e.g. WW, ZZ, tWZ and $t\bar{t}H$) are simulated using different generators such as MadGraph [107],MadSpin [125] and JHU [121–124]. All events are interfaced to Pythia for parton shower and hadronisation.

NOTE: Add source

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

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

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
$WZ \to 3\ell\nu$	MadGraph5_aMC@NLO+Pythia	5.26	1.61
tZq with $Z \to \ell^+\ell^-$	MadGraph5_aMC@NLO+Pythia	0.0758	3.77
tqH with $H \to ZZ \to \ell^+\ell^-\ell^+\ell^-$	JHU+Pythia	$8.80\ 10^{-6}$	-
$t\bar{t}W$+jets with $W \to \ell\nu$	MadGraph5_aMC@NLO+MadSpin+Pythia	0.2043 ± 0.0020	1.94
$t\bar{t}Z \to 2\ell + 2\nu +$ 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	-
$WW \to 2\ell 2\nu$	POWHEG+Pythia	12.178	-
$ZZ \to 4\ell$	POWHEG+Pythia	0.3366	-
WZZ	MadGraph5_aMC@NLO+Pythia	0.05565	1.14
ZZZ	MadGraph5_aMC@NLO+Pythia	0.01398	1.17
single top tWZ, with $Z_\mu \to \ell^+\ell^-$	MadGraph+Pythia	0.001123	-
single top t-channel $\bar{t}$	POWHEG+MadSpin+Pythia	$44.33\ ^{+1.76}_{-1.49}$	-
single top t-channel t	POWHEG+MadSpin+Pythia	$26.38\ ^{+1.32}_{-1.18}$	-
single top $\bar{t}W$	POWHEG+Pythia	35.85 ± 0.90 (scale) ± 1.70 (PDF)	-
single top tW	POWHEG+Pythia	35.85 ± 0.90 (scale) ± 1.70 (PDF)	-
$t\bar{t}$	POWHEG+Pythia	$831.76\ ^{+19.77+35.06}_{-29.20-35.06}$	-
$Z/\gamma^* +$jets, with $m_{\ell\ell} >$ 50GeV	MadGraph5_aMC@NLO+Pythia	$3 \times (\ 1921.8 \pm 0.6 \pm 33.2\ )$	1.49
$Z/\gamma^* +$jets, with 10 GeV $< m_{\ell\ell} <$ 50 GeV	MadGraph+Pythia	18610	-

3.3 Multivariate analysis techniques: Boosted Decision Trees

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

There are many software tools that exist for MVA. In this thesis the `Tool for Multivariate Analysis` (TMVA) [129] is used. This software is an open source project included into `ROOT` [130]. All multivariate techniques in TMVA belong to supervised learning algorithms. By training on events for which the outcome is known, a mapping function is determined that describes a classification or an approximation of the underlying behaviour defining the target value (regression).

In this thesis boosted decision trees (BDT) are employed for the classification of events as implemented in the `TMVA` framework [129]. This multivariate techniques is based on a set of decision trees where each yields a binary output depending on the fact that an event is signal- or background-like. The advantage of such a multivariate technique is that several discriminating variables can be combined into a powerful one-dimensional discriminant D.

In Figure 3.4 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 $x_{\mathrm{i}} > C_{\mathrm{i}}$ is found by maximizing the separation gain between nodes

$$\mathrm{separation gain} \approx \mathrm{gain(parent)} - \mathrm{gain(daughter, Signal)} - \mathrm{gain(daughter, Background)}, \tag{3.7}$$

with the gain computed using the Gini index

$$\mathrm{gain(cell)} \approx p(1-p), \tag{3.8}$$

where p denotes the purity of a selection $x > C$. This is repeated until the maximum of nodes is 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.

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

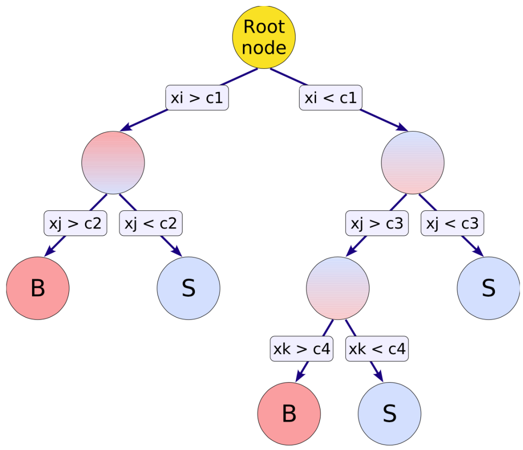

Figure 3.4: Schematic view of a decision tree. Figure taken from [129].

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 [131]. 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

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

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 $\beta \leqslant 0.5$ 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 is 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 [132]. 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 $\mathcal{L}(x|\text{Signal})/\mathcal{L}(x|\text{Background})$ [132]. No discrimination power will result in an AUC of 0%, while 50% means fully separated event classes. In Figure 3.5 an example of ROC curve is shown.

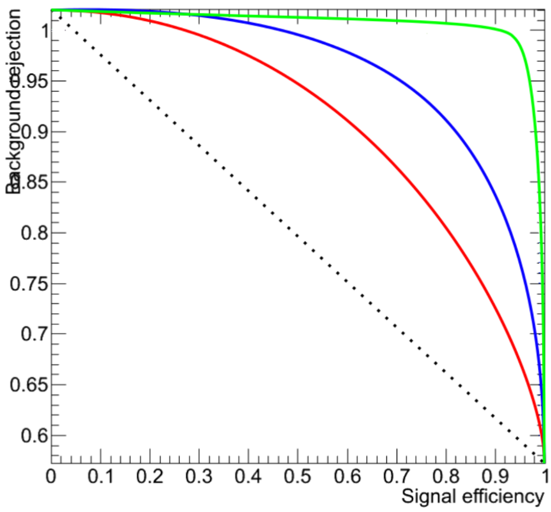

Figure 3.5: Example of ROC curves. In this example, the green method is better than the red one, which is better than the blue one. The dashed line represents a case where there is no separation. Figure taken from [133].

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. [134–137].
The `Higgs Combined Tool` [138] is a `RooStats` [139] framework which runs different statistical methods. In this section, only the statistical tools necessary for the performed search are described. The results presented in this thesis are obtained using the asymptotic formulae [140].

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

3.4.1 The absence of signal: limits

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

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

In this thesis, the modified frequentist approach for confidence levels is used [141, 142]. 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 $\mathfrak{L}(\text{data}|\,\mu, \theta)$ is constructed as

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

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

$$\text{Poisson}(\text{data}|\,\mu s(\theta) + b(\theta)) = \prod_\text{i} \frac{(\mu s_\text{i}(\theta) + b_\text{i}(\theta))^{n_\text{i}}}{n_\text{i}!} e^{-\mu s_\text{i}(\theta) - b_\text{i}(\theta)}. \tag{3.11}$$

At the LHC, the test statistic is defined as

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

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

$$\mathrm{CL} = \frac{\mathrm{P}\left(q_\mu \geq q_\mu^{obs} | \, \mu s + b\right)}{\mathrm{P}\left(q_\mu \geq q_\mu^{obs} | \, b\right)} \leq \alpha. \tag{3.13}$$

These probabilities are defined as

$$\begin{aligned} p_\mu &= \mathrm{P}\left(q_\mu \geq q_\mu^{obs} | \, \mu s + b\right) = \int_{q_\mu^{obs}}^{\infty} f(q_\mu | \, \mu, \hat{\theta}_\mu^{obs}) \, dq_\mu, \\ 1 - p_b &= \mathrm{P}\left(q_\mu \geq q_\mu^{obs} | \, b\right) = \int_{q_{\mu=0}^{obs}}^{\infty} f(q_\mu | \, \mu = 0, \hat{\theta}_{\mu=0}^{obs}) \, dq_\mu, \end{aligned} \tag{3.14}$$

where p_μ and p_b are called the p-values associated to the two hypothesis, and $f(q_\mu | \mu, \hat{\theta}_\mu^{obs})$ and $f(q_\mu | \, \mu = 0, \hat{\theta}_{\mu=0}^{obs})$ are the pdfs of the signal plus background and background only hypothesis constructed from toy Monte Carlo pseudo data. These pdfs are shown in Figure 3.6 and are generated with nuisance parameters fixed to $\hat{\theta}_{\mu=0}^{obs}$ and $\hat{\theta}_\mu^{obs}$. These values of the nuisance parameters for the background only $\hat{\theta}_{\mu=0}^{obs}$ and signal plus background $\hat{\theta}_\mu^{obs}$ hypothesis that best 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$= 0.05$

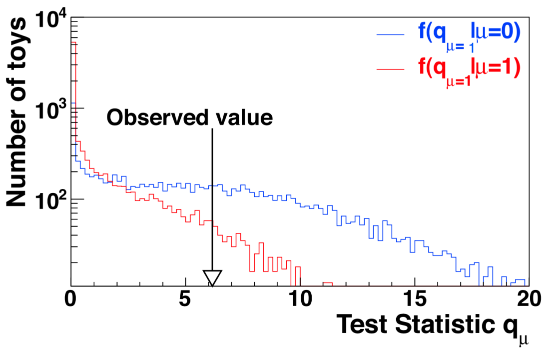

Figure 3.6: Test statistic distributions for pseudo data generated for the signal plus background ($\mu = 1$) and background only ($\mu = 0$) hypothesis. Figure taken from [137].

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

3.4.2 Adding sources of uncertainty

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

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

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

where $\pi_{\theta}(\theta)$ is the hyper prior for the (imaginary) measurements. For the pdfs used by the `Higgs Combine Tool` (normal, log normal, gamma distribution), hyper priors can remain flat. This allows to use the pdf $p(\tilde{\theta}|\theta)$ to constrain the likelihood of the main measurement in a frequentist calculation. Additionally this allows to build a sampling distribution of the test statistic [137].

The statistical uncertainties on the Monte Carlo prediction in each bin are obtained following the Barlow-Beeston-light approach [143]. 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

$$n_{\mathrm{tot}} = \sum_{\mathrm{i} \in \mathrm{bkg}} n_{\mathrm{i}}, \tag{3.16}$$

the total uncertainty e_{tot} is given by

$$e_{\mathrm{tot}} = \sqrt{\sum_{\mathrm{i} \in \mathrm{bkg}} e_{\mathrm{i}}^2}, \tag{3.17}$$

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

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

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

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

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

The mapping between the posteriors $\rho(\theta|\tilde{\theta})$ and the auxiliary measurement pdfs $p(\tilde{\theta}|\theta)$ are given in [137].

3.4.3 Asymptotic approximation of the CL method

In order to significantly reduce computing time, the Asymptotic CL method is used. This method avoids an ensemble of toy Monte Carlo samples and instead replaces it by one representative dataset, called Asimov dataset. This dataset is constructed such that all observed quantities are set equal to their MLE values ($\hat{\theta}_{\text{Asimov}} = \theta_0$). More information about this procedure can be found in Refs. [135].

3.4.4 Extracting the signal model parameters

From a scan of the profile likelihood ratio,

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

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

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

is called the best-fit set.

The 68% and 95% CL on a given parameter of interest a_{i} is then evaluated from $q(a_{\text{i}}) = 1$ or $q(a_{\text{i}}) = 3.84$ respectively, where all other unconstrained model parameters are treated in the same way as the nuisance parameters [136].

Event reconstruction and identification 4

After the detector simulation described in Section 3.2, the simulated data has the exact same format as the real collision data recorded at the CMS experiment. Therefore the same software can be used for the reconstruction of both simulation and real data. In Section 4.1, the object reconstruction for physics analysis is shown. After reconstructing the objects, the objects are connected (Section 4.2) to physics objects need to be identified (Section 4.3) and corrected for pile up (Section 4.4). The objects used for phyiscs analysis have extra requirements as shown in Section 4.5. A summary of all the corrections applied to data and simulation is given in Section 4.7.

4.1 Object Reconstruction

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

The traditional hadron colliders reconstruction is as follows. The reconstruction of isolated photons and electrons is primarily done by the ECAL, while the identification of muons is based on the muon detectors. Hadrons and photons form jets which are measured by the calorimeters without any contribution from the tracker or muon detectors. Jets can be tagged using the tracker as coming from hadronic τ decays or b hadronisation based on the properties of the properties the relevant charged particle tracks. The missing transverse energy is defined as the vectorial sum of the undetectable particle transverse momenta, and can be reconstructed without any information from the tracker. The particle flow (PF) [144] reconstruction correlates

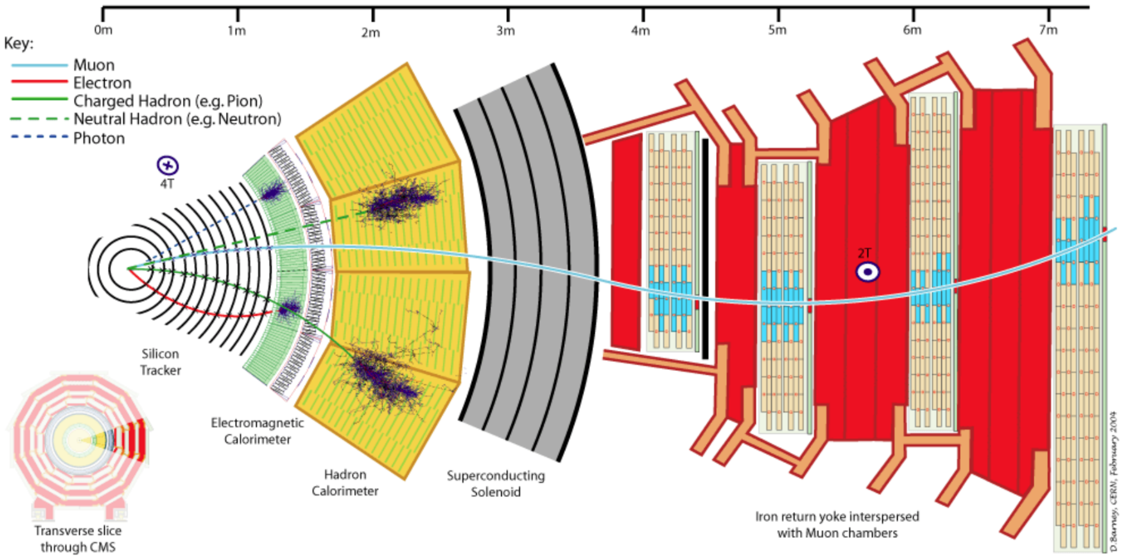

Figure 4.1: Cross-section of the CMS detector with all parts of the detector labelled. This sketch shows the specific particle interactions from a beam interaction reign to the muon detector. The muon and charged pion are positively charged, the electron is negatively charged. Figure taken from [144].

the tracks and clusters from all detector layers with the identification of each final state particle, and combining the corresponding measurements to reconstruct the properties. Here, the muon is identified by a track in the inner tracker connected to a track in the muon detector as described in Section 4.1.2. The electrons are identified by a track and ECAL cluster, and not connected to an HCAL cluster as described in Section 4.1.3. The ECAL and HCAL clusters without a track link identify the photons and neutral hadrons, while the addition of the tracker determines the energy and direction of a charged hadron.

Coarse-grained detectors can cause signals of different particles to merge and reduce the ability of identifying and reconstructing the particles. Therefore, particle flow identification requires sufficiently segmented subdetectors such that a global event description is possible. From a list of identified particles that are reconstructed from a combined fit of all relevant measurements, the physics objects are determined. The CMS detector is built to meet to requirements of the particle flow reconstruction. It has an efficient and pure muon identification system, a hermetic HCAL with coarse segmentation, a higher segmented ECAL, a fine-grained tracker and a large magnetic field to separate the calorimeter deposits of charged and neutral particles in jets.

4.1.1 Charged particle tracks

An iterative tracking algorithm is responsible for the reconstruction of the tracks made by charged particles in the inner tracking system. Each iteration consists of four steps [64]: the track-seed generation, the pattern recognition algorithm, removal of track-hit ambiguities and a final track fit.

The seed generation is the first step. It consists of finding reconstructed hits that are usable for seeding the subsequent track-finding algorithm. They are identified from a group of at least three reconstructed hits in the tracker, or from a pair of hits while requiring the origin of the track segment to be compatible with the nominal beam-collision point. Since the pixel has a higher granularity compared to the strip tracker, its seed generation efficiency is higher. The overall efficiency exceeds 99%. The second step of each iteration, the pattern recognition algorithm, uses the seeds as a starting point for a Kalman filter method [145, 146]. This algorithm extrapolates the seed trajectory towards the next tracker layer taking into account the magnetic field and multiple scattering effects. The track parameters are updated when a compatible hit in the next layer is found. This procedure continues until the outermost layer is reached. Since the Kalman filter method can result in multiple tracks associated to the same seed, or different tracks sharing the same hits, a removal of ambiguities is necessary. This ambiguity resolving is done by removing tracks that are sharing too many hits from the list of track candidates. The tracks with the highest number of hits or with the lowest χ^2 in the track fit is kept. The updated track parameters are then refitted using the Kalman filter method, where all hits found in the pattern recognition step are taken into account. The fit is done twice - once outwards from the beam line towards the calorimeters, and inwards from the outermost track hit to the beam line -, improving the estimation of the track parameters.

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

4.1.2 Following the Muon's Footsteps

The muon reconstruction [147] has three subdivisions: local reconstruction, regional reconstruction and global reconstruction. The local reconstruction is performed on individual detector elements such as strip and pixel hits in the inner tracking system, and muon hits and/or segments in the muon chambers. Independent tracks are reconstructed in the inner tracker - called tracker tracks - and in the muon system, called standalone muon tracks. Based on these tracks, two reconstructions are considered.

The outside-in approach is referred to as Global Muon reconstruction. For each standalone muon track, a inner tracker track is found by comparing the parameters of the two tracks propagated onto a common surface. Combining the hits from the tracker track and the standalone track, gives a fit via the Kalman filter technique [145, 146] for a global muon track.

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

For low transverse momenta ($p_{\mathrm{T}} \lesssim 5$ GeV), the tracker muon reconstruction is more efficient 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 bor c quarks. The global muon approach typically improves the tracker reconstruction for $p_{\mathrm{T}} \gtrsim 200$ 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) is used in the electron track reconstruction instead.

NOTE: Add source

In CMS, the electrons are reconstructed in two ways. The older ECAL based tracking is developed to identify high energy, 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 GeVfound 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 enrgy 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 enery 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.

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

NOTE: Add source

NOTE: Add source

4.1.4 Primary Vertex Reconstruction

The primary vertex (PV) reconstruction should be able to measure the location of all proton interaction vertices in each event: the signal vertex an all vertices from pile up events. It consists of a vertex finding and a vertex fitting algorithm and happens in three steps. Tracks are selected to be consistent with being produced promptly in the primary interaction by imposing requirements on the track parameters [74]. By grouping reconstructed tracks according to the z coordinate of their closest approach to the beam line, vertices for all interaction in the same beam crossing are found, at CMS this is done by a deterministic annealing algorithm [148] . On top of this, a vertex fitting algorithm like the Adaptive Vertex fitter [149], is performed. This creates the three-dimensional primary-vertex position. With this fit, the contribution from long-lived hadron decays is reduced by down weighting the tracks with a larger distance to the vertex. The primary vertex corresponding to the highest sum of squared track transverse momenta is noted as the point of the main interaction. The resolution on the primary vertex is about 14 μm in $r\phi$ and about 19 μm in the z direction for primary vertices with the sum of the track $p_T > 100$ GeV for 2016 data taking.

NOTE: Check text with PFlow paper!

4.1.5 Calorimeter clusters

The cluster algorithm in the calorimeter

1. detects and measures the energy and direction of stable neutral particles such as photons and neutral hadron,
2. separates neutral particles from charged hadron energy deposits, /item reconstructs and identifies electrons and their bremsstrahlung photons,
3. contributes to the energy measurements of charged hadrons that don't have accurate tracks parameters, e.g. for low quality and high transverse momentum tracks.

The clustering is performed separately in each subdetetector: ECAL barrel and endcaps, HCAL barrel and end caps, and the two preshower layers. The HF has no clustering algorithm since the electromagnetic or hadronic components give rise to an HF EM or HF HAD cluster.

The clustering algorithm consist of different steps. First seeds are identified when cells have an energy larger than the seeding threshold and lager than their neighbouring cells. Then topological clusters are made by accumulating cells that share at least a corner with a cell already in the cluster and an energy above a cell threshold set to twice the noise level. The third step is a expectation maximization algorithm that reconstructs the cluster [144]. This algorithm assumes that the energy deposits are Gaussian distributed and is an iterative algorithm with two steps at each iteration. A first step calculated the expected fraction if the energy in a certain step, while the second step performs a maximum likelihood fit. The positions and energies of the Gaussian functions are then taken as cluster parameters.

The calorimeter clusters are used for reconstructing photons and neutral hadrons. The clusters that are not in the vicinity of the extrapolated charged tracks are easily identified as neutral hadrons or photons. For the energy deposits that overlap with charged hadrons however, the 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[1]. 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 responce in the HCAL depends on the fraction of shower energy deposited in the ECAL, the ECAL and HCAl cluster energyes are recalibrated together to get an estimate of the true hadron energy. Since now the calibration is done for hadrons, single neutral hadrons such as K_L^0 are used for determining the calibration constants. The hadrons interactiong 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 ar 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

[1] Specifically electron and photon energy corrections are performed after the identification step.

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.

4.3 Particle flow identification

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

4.3.1 Muons

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

4.3.2 Electrons and isolated photons

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

4.3.3 Hadrons and non-isolated photons

After muon, electron and isolated photon identification, the remaining particles are hadrons from jet fragmentation and hadronisation. These can show up ad charged hadrons (e.g. $\pi\pm$, $K^{\pm}$, or protons), netutral hadrons (e.g. K^0_L or neutrons), non isolated photons (e.g. from π^0 decays), and additional muons from early decays of charged hadrons.

The photons and neutral hadrons are assigned to calorimeter clusters without any link to tracks. When the calorimeter clusters between the ECAL and HCAL are linked, the clusters are assumed to aris from the same hadron shower. if their is not such a link, HCAL clusters are assigned to neutral hadrons, while the ECAL clusters are assigned to photons based on the fact that neutral hadrons leave only 3% of their energy in the ECAL.Then the HCAL clusters linked with tracks, that are not linked with other HCAL clusters, are assigned to charged hadrons. These tracks can be linked with remaining ECAL clusters.

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

4.3.4 Post processing

After identification and reconstruction of all particles as described above. An artificial large missing transverse momentum $\vec{\not{p}}_{\mathrm{T}}$ can be reconstructed. The cause of the $\vec{\not{p}}_{\mathrm{T}}$ is mostly misidentified or misreconstructed high-p_{T} muons originating from cosmic rays, misconstruction of the muon's momentum, or punch-throuvh charged hadrons. A post processing step is applied to solve this $\vec{\not{p}}_{\mathrm{T}}$. Events with genuine large $\vec{\not{p}}_{\mathrm{T}}$ due to the presence of neutrino's are unaffected by this post processing.

4.4 Pile up mitigation

The particle flow algorithm is design without taking pile up into account. For the 8 TeV dataset, an avarage of about 21 pile up interactions of pile up interactions happened per bunch cross section. For the dataset taken at 13 TeV, the number of pile up interactions increased to about 27 interactions per bunch crossing. THese interactions are spread around the beam axis around the centre of the CMS coordinate system and follow a normal distribution with a standard deviation of about 5 cm [144]. The number of pile up interactions is iestimated from th enumber of interaction vertices reconstructed from charged particle tracks , or from the instantaneous luminosity of the given bunch crossing with dedicated detectors and the inelastic proton-proton crossing . This minimum bias cross section is measured to be 69.2 mb ± 4.6% [150]. The distribution of the number of pile up interaction is reweighted to take care of discrepancies. The uncertainty on the minimum bias cross section measurement results in a systematic shift in the pileup distribution.

NOTE: Add source

NOTE: Add source

The pile up vertices are seperated 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 [151]. 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.

NOTE: Add source

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 $\vec{\not{p}}_{\mathrm{T}}$. 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 and loose working points. The tight working point yields the most genuine muons and rejects falsely reconstructed ones. While the loose working point yields as many reconstructed muons as possible. Detailed reports on the performance can be found in .

NOTE: Add source

In order to reject objects wrongly reconstructed as muons from hadron showers that reach he 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 $\chi^2/\mathrm{ndof} < 10$. 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 $d_{\mathrm{x,y}} < 2$ mm and $d_{\mathrm{z}} < 5$ mm. Also muons according to the loose muon ID 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 $Z \to \mu^-\mu^+$ 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 data and simulation are corrected by applying p_{T} - and η-dependent scale factors ($\epsilon_{\mathrm{data}/\epsilon_{\mathrm{MC}}}$) to simulated events. In Table 4.1, the muon requirements for the muons used throughout this theses are summarised.

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

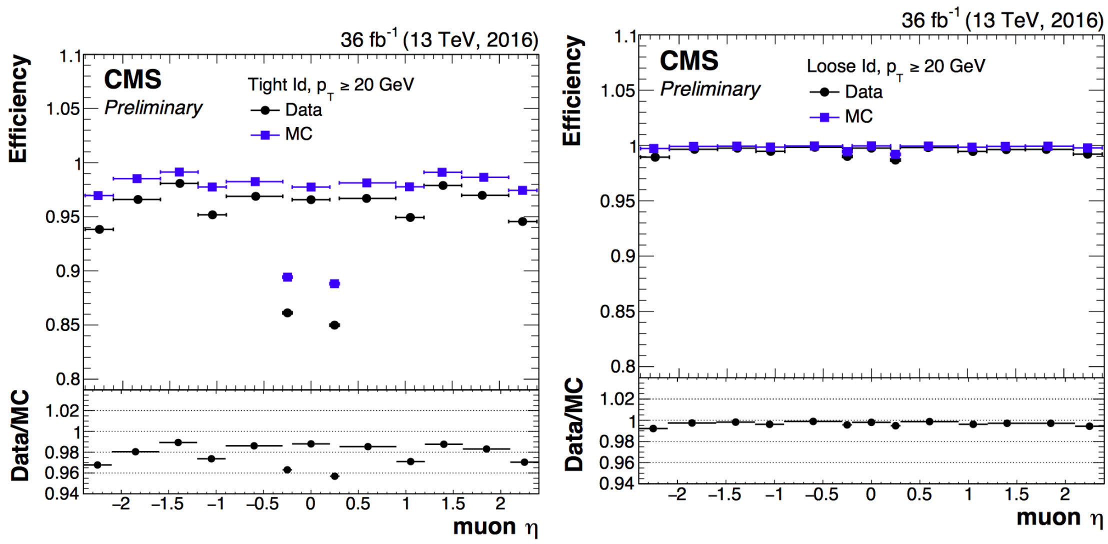

Figure 4.2: Comparison of the muon tight ID (left) and loose ID (right) efficiencies in data and simulation as a function of the pseudorapidity of the muon using the full 2016 dataset. Figure taken from [152].

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 enrgy coming from pile up interactions. The relative isolation for muons $\mathscr{I}$ is defined as

NOTE: Add source

$$\mathscr{I}_\mu = \frac{\sum p_T(\mathrm{CH}) + \max\left(0., \sum E_T(\mathrm{NH}), \sum E_T(\gamma) - 0.5 \times \sum E_T(\mathrm{CH})\right)}{p_T(\mu)}, \tag{4.1}$$

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

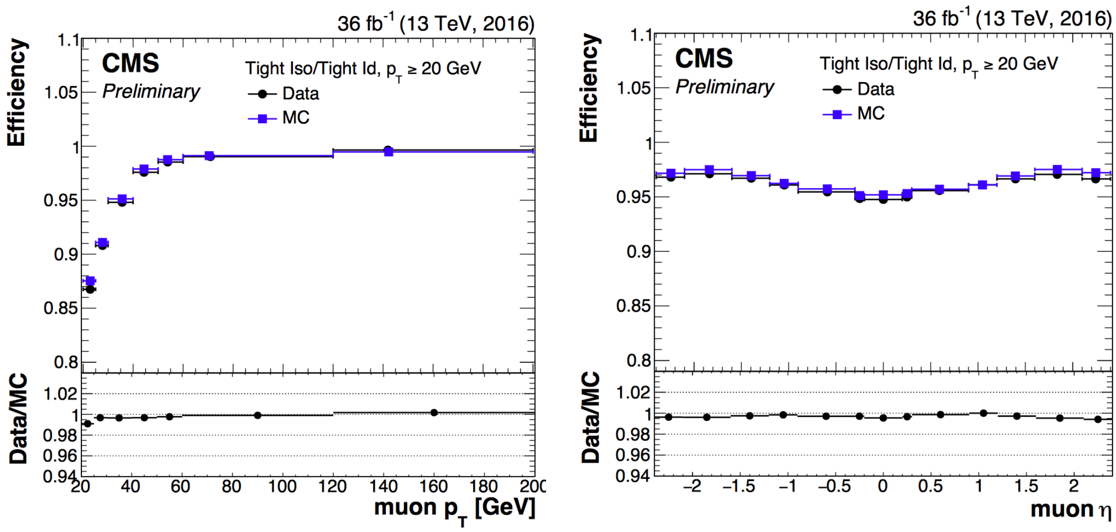

Figure 4.3: Comparison of the muon tight isolation requirement with the muon tight ID efficiencies in data and simulation as a function of the transvers emomentum (left) or pseudorapidity (right) of the muon using the full 2016 dataset. Figure taken from [152].

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

4.5.2 Electrons

The electrons candidates used correspond to the tight and veto working points. The study of the electron reconstruction and identification performance can be found in [153].

Starting from an electron PF candidate with a GSF track that is outside the barrel-endcap transition region ($1.4443 < |\eta| < 1.5660$), several requirements are set. The electrons from photon conversions are dismissed by requiring the electron track to have not have more than one (two or three) missing hit in the innermost layer for the tight (veto) working point. Additionally, a photon conversion veto is applied by testing if a pair of electron tracks is originating from a common displaced vertex. For the 8 TeV datasets more refined cuts are placed on the electron object using a multivariate analysis. For the 13 TeV dataset this is replaced with more refined cuts on the shower shape variables such as the difference in η or ϕ between the energy weighted supercluster position in the ECAL and the track direction in at the innermost tracker position ($\Delta\eta_{in}$, $\Delta\phi_{in}$), and the ECAL crystal based shower covariance in the η direction ($\sigma_{\eta\eta}$). These cuts also include energy related variables such as the absolute difference between the inverse electron energy measured in the ECAL and the inverse momentum measured in the tracker ($|1/E - 1/p|$), and the ratio of the energy measured in the HCAL and ECAL (H/E). Unlike the 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}) . 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 superclustes. Table 4.2 shows the values used for 13 TeV data. The relative electron isolation $\mathscr{I}_{\mathrm{e}}$ is calculated as

NOTE: Add source

$$\mathscr{I}_{\mathrm{e}} = \frac{\sum p_{\mathrm{T}}(\mathrm{CH}) + \max\left(0., \sum E_{\mathrm{T}}(\mathrm{NH}), \sum E_{\mathrm{T}}(\gamma) - \rho \times A_{\mathrm{eff}}\right)}{p_{\mathrm{T}}(\mathrm{e})}. \tag{4.2}$$

Table 4.2: The effective areas A_{eff} used for the electron relative isolation [154].

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

The efficiency of electron identification is estimated from $\mathrm{Z} \to \mathrm{e}^-\mathrm{e}^+$ events via the tag-and-probe method and is shown in Figure 4.4 for the tight working point. The efficiencies reach

$\approx 95-100\%$. The difference between data and simulation are corrected by dedicated p_T- and η dependent scale factors as well.

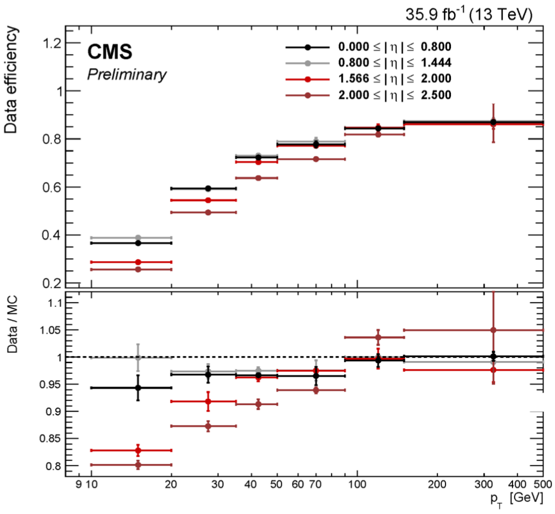

Figure 4.4: Electron identification efficiency as function of the electron transverse momentum from the full 2016 dataset. Figure taken from [153].

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	$\lvert\eta_{supercluster}\rvert \leq 1.479$		$\lvert\eta_{supercluster}\rvert > 1.479$	
	Veto electron	Tight electron	Veto electron	Tight electron
$\sigma_{\eta\eta}$	< 0.0115	< 0.00998	< 0.037	< 0.0292
$\lvert\Delta\eta_{in}\rvert$	< 0.00749	< 0.00308	< 0.00895	< 0.00605
$\lvert\Delta\phi_{in}\rvert$	< 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
$\lvert 1/E-1/p\rvert$	< 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

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$-\mathrm{k_T}$ algorithm [155] 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$-\mathrm{k_T}$ algorithm considers all candidates as protojets and starts to calculate the distances for protojets i and j as

$$
\begin{aligned}
d_{\mathrm{ij}} &= \min\left(\frac{1}{p_{\mathrm{T,i}}^2}, \frac{1}{p_{\mathrm{T,j}}^2}\right), \\
d_{\mathrm{i}} &= \frac{1}{p_{\mathrm{T,i}}^2}.
\end{aligned}
\tag{4.3}
$$

For each iteration the two distances are calculated. When $d_{\mathrm{ij}} < d_{\mathrm{i}}$, the two protojets are merged and their four momentum is summed. If d_{i} is the smallest distance, the protojet is renamed as final jet and ignored in the subsequent steps. More information about the jet algorithm performance can be found in [156].

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

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

The energy of the reconstructed jets deviate from the energies of the corresponding jets clustered from the hadronisation products of true partons from simulations due to non linear subdetetector response and efficiencies. The jet energy corrections (JEC) calibrate the jets in order to have the correct energy scale. Jet energy scale corrections (JES) are determined as a function of pseudorapidity ant the transverse momentum from data and simulated events by combining several channels and methods. This is extensively described in [157]. These corrections account for the effects of pile up, the uniformity of the detector response, and residual data-simulation jet energy scale differences. Furthermore, the jet energy resolution (JER) is measured in data and simulation as function of pile up, jet size and jet flavour. A detailed understanding of both the energy scale and the transverse momentum resolution of the jets is crucial for many physics analysis, and these are commonly the main source of systematic uncertainty. The performance of the jet energy corrections for the 13 TeV dataset can be found in [158].

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

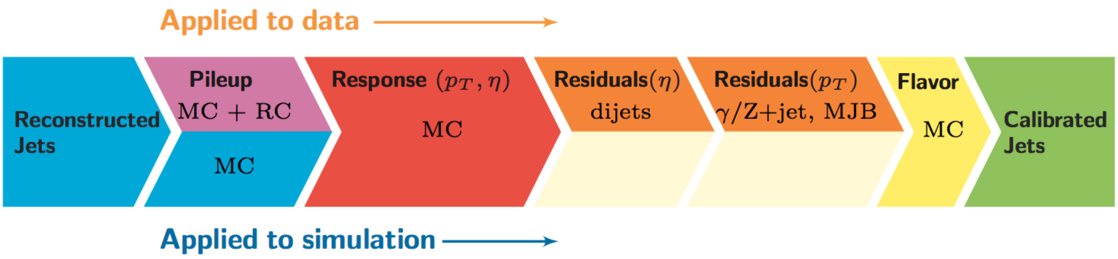

Figure 4.5: The sequence of the JEC for data and simulations The corrections marked with MC are derived from simulation studies, while RC stands for random cone, and MJB for the analysis of multi-jet events. Figure taken from [157].

additional pile up activity. It is based on the jet area method, which uses the effective area of the jets multiplied by the average density in the event to calculated the offset energy to be subtracted of the jets. The correction factors are derived by comparing the jet response with and without pile up events overlaid. The residual differences between data and detector simulation are determined using the random cone method (RC). For this method, many jets are reconstructed in each event by clustering particles through placing random cones. This provides a mapping of the $\eta\phi$-space and the average p_{T} of those jets gives the average energy offset due to pile up [157]. The next level of corrections have as goal to have an uniform energy response independent of the transverse momentum or pseudorapidity of the jet. These corrections are determined from simulated events by matching the reconstructed to true particle jets and comparing there momenta. The residual corrections between data and simulation are determined by comparing the transverse momentum balance in various types of events (multi-jet, $\mathrm{Z}+$jets, and $\gamma+$jets), using a reference jet in the barrel region. The jet flavour corrections are optional and not used for this thesis More information on the jet flavour corrections can be found in [157]. For jets with a transverse momentum above 30 GeV, the uncertainties from the various corrections are 3-5% for the 13 TeV dataset [158].

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

$$
\begin{aligned}
c_{\text{matched}} &= 1 + \frac{p_{\mathrm{T}}^{\text{reco.}} - p_{\mathrm{T}}^{\text{true}}}{p_{\mathrm{T}}^{\text{reco.}}}\left(s_{\text{JER}} - 1\right), \\
c_{\text{unmatched}} &= 1 + \mathrm{N}(0, \sigma_{\text{JER}})\sqrt{\max\left(s_{\text{JER}}^2 - 1, 0\right)},
\end{aligned}
\tag{4.4}
$$

where $\mathrm{N}(0, \sigma_{\text{JER}})$ denotes a sample value from a normal distribution centred at zero with as standard deviation the relative resolution in simulation σ_{JER}, and s_{JER} the η-dependent resolution scale factors. These scale factors are derived from data from di-jet or $\gamma+$jets events and analysing the p_{T} balance. The resolution scale factors (data/simulation) are found to be 1.1-1.2 except for the transition regions around $|\eta| = 3$ and $|\eta| = 1.4$ [158].

4.5.4 Jets from b fragmentation

Jets originating from the hadronisation of bottom quarks can be discriminated from jets from gluons and light-flavour quarks as well as charm quark fragmentation through the use of b-tagging. There are a multitude of algorithms developed within CMS to perform b-tagging [159, 160] on jets that fall within the pseudorapidity acceptance of the trackers. These algorithms exploit the properties of the b quark to identify the jets formed by its fragmentation. These hadrons have relative large masses, long lifetimes and daughter particles with hard momentum spectra. Additionally, their semi-leptonic decays can be exploited as well. To use b-jet identification in an analysis, one needs to know its efficiency and misidentification probability. In general these are function of the pseudorapidity and transverse momentum of the considered jet. Their performances are directly measured from data by use of b-jet enriched jet samples (multi-jet or top-quark decays).

This thesis uses b-jets identified by the Combined Secondary Vertex version 2 (CSVv2) algorithm [159]. This algorithm combines secondary vertices together with track based lifetime information by use of a multivariate technique. The secondary vertex is reconstructed from diplaced tracks within a jet, as illustrated in Figure 4.6. The final state b quark is encapsulated in a B meson (e;g. $\mathrm{B}^{\pm}$, B_0, B_S) after the hadronisation. This B meson has relatively long lifetime and can travel a measurable distance from the primary vertex before decaying[2]. After reconstruction, the seondary vertices are required to be in accordance with the B meson hypothesis bases an 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.

The b-tagging algorithm performances are evaluated taking into account two cases: discrimination of b-tagged jets originating from charm quarks, and discrimination of b-tagged jets against jets coming from gluons or light (u, d, s) quarks.In Figure 4.7, the misidentification probabilities for different b-tagging algorithms within CMS are shown. Based on the misidentification probabilities for a certain threshold on the CSVv2 discriminator, different working points (WP) are defined. These are shown in Table 4.5. The analysis presented in this thesis uses the loose working point which has an average efficiency of 81% and a misidentification rate of 10%.

NOTE: Find reason why I am not using cMVA, omdat CVSv2 op ttbar events is gemeten en cMVA op multijet?

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

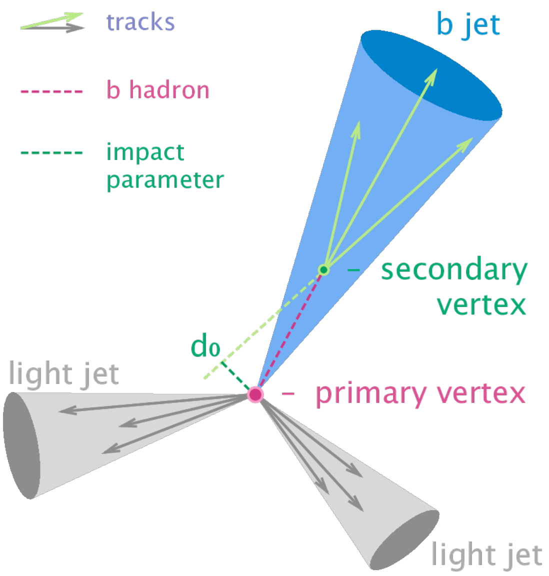

Figure 4.6: Sketch showing the common principle of the identification of b-jets. Figure taken from [161]

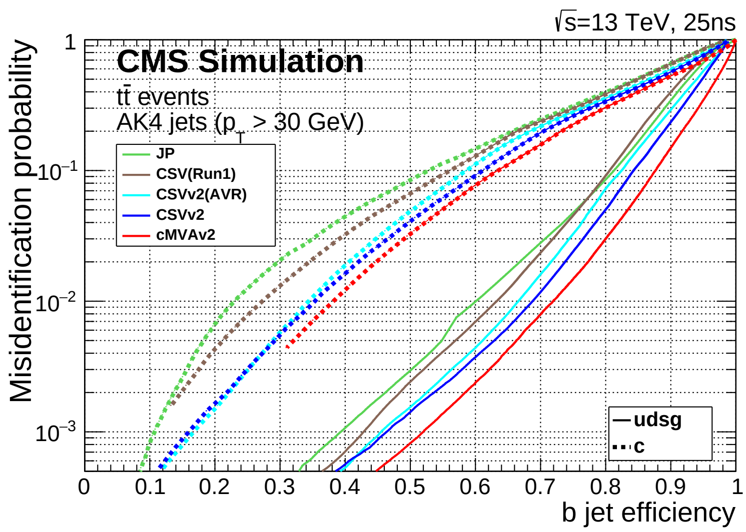

Figure 4.7: Misidentification probabilities of various b-tagging algorithm in simulation. Figure taken from [160].

Table 4.5: Working points used for tagging jets as coming from b quarks for the CSVv2 discriminant.

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

The efficiency of tagging a jet as coming from a bottom quark in simulation typically deviates somewhat from data. Efficiency scale factors $\epsilon_{\mathrm{b}}^{\mathrm{data}}/\epsilon_{\mathrm{b}}^{\mathrm{MC}}$ 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 [160]. 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 [162]. 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 af 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

$$
\begin{aligned}
\mathrm{SF}_{\mathrm{b}} &= \frac{N_{\mathrm{b}}^{\mathrm{data}} - N_{\mathrm{b}}^{\mathrm{MC}}}{N_{\mathrm{b}}^{\mathrm{MC}}}, \\
\mathrm{SF}_{\mathrm{g,u,d,s}} &= \frac{N_{\mathrm{g,u,d,s}}^{\mathrm{data}} - N_{\mathrm{g,u,d,s}}^{\mathrm{MC}}}{N_{\mathrm{g,u,d,s}}^{\mathrm{MC}}}, \\
\mathrm{SF}_{\mathrm{c}} &= 1.
\end{aligned}
\tag{4.5}
$$

The uncertainties related to the IterativeFit method cover possible shape discrepancies between data and simulation. The purity of the sample on which the scale factors are determined has an influence on the scale factors. 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. Furthermore, the jet energy scale results in jets migrating from one p_{T} bin to an other, having an influence on bin dependent scale factors. The statistical fluctuations of the limited amount of entries in each bin are also accounted for and have an influence on the scale factor uncertainties. These have four uncorrelated sources: two for heavy flavour and two for light flavour jets. Since the uncertainty on the scale factors for the jets originating from a charm quark (cf) is determined from the uncertainty on the b scale factors resulting in two independent uncertainties [162].

4.5.5 Missing transverse energy

The missing transverse momentum $\vec{\not p}_{\mathrm{T}}$ and energy $E_{\mathrm{T}}^{\mathrm{miss}}$ resulting from particles that do not interact with the detecor material, are calculated to balance the vectorial sum of the transverse momenta of all particles:

$$
\begin{aligned}
E_{\mathrm{T}} &= \left|\vec{\not p}_{\mathrm{T}}\right|, \\
\vec{\not p}_{\mathrm{T}} &= - \sum_{\mathrm{i=1}}^{\mathrm{N_{particles}}} \vec{p}_{\mathrm{T,i}}.
\end{aligned}
\tag{4.6}
$$

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

The missing transverse energy is influenced by the minimum tresholds in calorimeters, the inefficiencies in the tarcker, and the nonlinear response of the calorimeter to hadronic particles. Ths bias is reduced by correctiong the trannsverse momentum of the jets too partcle jet p_{T} via the JEC and propagating it to the missing transverse momentum taking into account the energy

$$\begin{aligned}
\vec{\not{p}}_{\mathrm{T}}^{\mathrm{corr}} &= -\sum_{\mathrm{i}=1}^{\mathrm{N_{jets}}} \vec{p}_{\mathrm{T,i}}^{\mathrm{corr.}} - \sum_{\mathrm{i}=1}^{\mathrm{N_{unlustered}}} \vec{p}_{\mathrm{T,i}}^{\mathrm{raw}}, \\
\vec{\not{p}}_{\mathrm{T}}^{\mathrm{corr}} &= \vec{\not{p}}_{\mathrm{T}}^{\mathrm{raw}} - \sum_{\mathrm{i}=1}^{\mathrm{N_{jets}}} \left(\vec{p}_{\mathrm{T,i}}^{\mathrm{JEC}} - \vec{p}_{\mathrm{T,i}}^{\mathrm{PU-only}} \right).
\end{aligned} \tag{4.7}$$

The $\vec{p}_{\mathrm{T,i}}^{\mathrm{PU-only}}$ denotes the transverse momentum of the jet, where only the pile up related corrections are applied. The performance of the missing transverse energy reconstruction can be found in [163].

4.6 Luminosity

The luminsity 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 [164]. The instantaneous luminosity from recorded process rate R is then determined as

$$\mathrm{L}dt = \frac{R dt}{\sigma_{\mathrm{fid}}}, \tag{4.8}$$

where $\sigma_{\mathrm{fid}} = \sigma \times A$ corresponds to the fiducial cross section recorded in the luminometer acceptance A which is determined using van der Meer scans [150]. The overall uncertainty on the luminosity measurement is estimated to be 2.5%.

The luminosity is used to infer to number of pile up interactions in data, which can be used to corrected the predefined pile up interactions in simulation. The average number of pile up 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 [150]. 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.

4.7 Summary of corrections

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

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

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

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.

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.

Luminosity The luminosity is measured with a global uncertainty of 2.5%, affecting the expected number of events.

5 Event selection and categorisation

A basic event selection is made for selecting signal like events. The necessary event requirement are discussed in Section 5.2.

The analysis uses signal and background regions to constrain the huge SM background compared to the expected signal. Section 5.4 discusses each region that is entering the analysis. On top of the use of background estimation from control regions, backgrounds that have prompt leptons contaminated by real leptons either from decays of tau leptons or from hadronized mesons or baryons (collectively commonly referred as “non-prompt leptons") as well as by hadrons or jets misidentified as leptons[1] are evaluated with a data-driven method discussed in Section 5.5.

5.1 Baseline event selection and filters

Trigger and filters

5.2 Event selection

5.3 Effect of the corrections in dilepton events

5.4 Regions and channels

5.5 Data driven background simulation

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

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

6.1 Construction of template distributions

6.2 Systematic uncertainties

6.3 Limit setting procedure

6.4 Result and discussion

Denouement of the top-Z FCNC hunt at 13 TeV 7

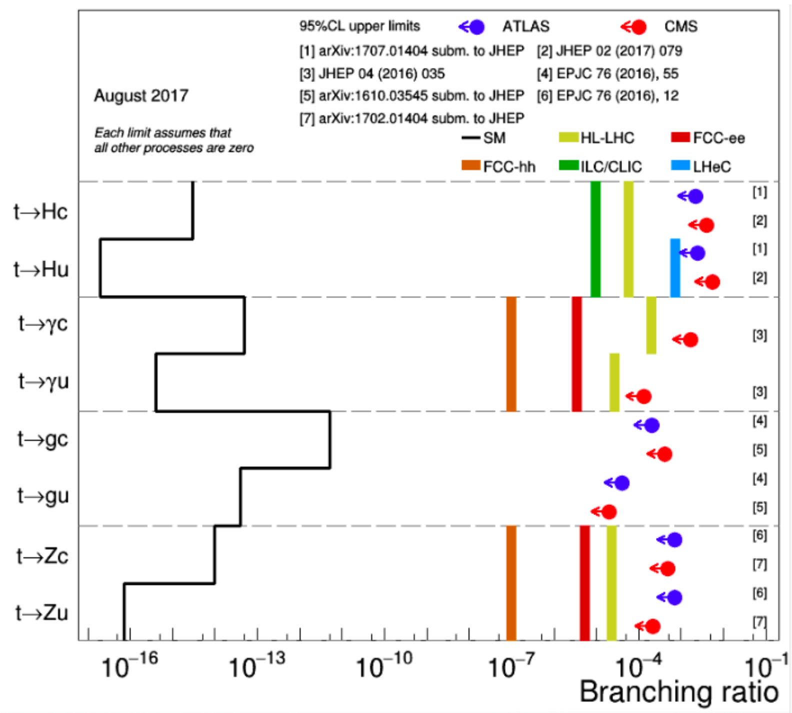

Figure 7.1:

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).

[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. arXiv: hep-th/0701053 [hep-th] (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] C. Patrignani et al.: **Review of Particle Physics**. In: *Chin. Phys.*, **C40**:10 (2016), p. 100001. doi: 10.1088/1674-1137/40/10/100001 (see pp. 2–3, 6–9, 38–39, 68).

[5] 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. arXiv: 1207.7235 [hep-ex] (see pp. 2, 19).

[6] 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. arXiv: 1207.7214 [hep-ex] (see pp. 2, 19).

[7] Nicola Cabibbo: **Unitary Symmetry and Leptonic Decays**. In: *Phys. Rev. Lett.*, **10**: (12 June 1963), pp. 531–533. doi: 10.1103/PhysRevLett.10.531 (see p. 5).

[8] S. L. Glashow, J. Iliopoulos, and L. Maiani: **Weak Interactions with Lepton-Hadron Symmetry**. In: *Phys. Rev. D*, **2**: (7 Oct. 1970), pp. 1285–1292. doi: 10.1103/PhysRevD.2.1285 (see p. 5).

[9] B.J. Bjørken and S.L. Glashow: **Elementary particles and SU(4)**. In: *Physics Letters*, **11**:3 (1964), pp. 255–257. doi: https://doi.org/10.1016/0031-9163(64)90433-0 (see p. 5).

[10] Luciano Maiani: **The GIM Mechanism: origin, predictions and recent uses**. In: *Proceedings, 48th Rencontres de Moriond on Electroweak Interactions and Unified Theories: La Thuile, Italy, March 2-9, 2013*. 2013, pp. 3–16. arXiv: 1303.6154 [hep-ph]. url: https://inspirehep.net/record/1225307/files/arXiv:1303.6154.pdf (see p. 5).

[11] Patrick Koppenburg and Sebastien Descotes-Genon: **The CKM Parameters**. In: (2017). arXiv: 1702.08834 [hep-ex] (see p. 6).

[12] J. A. Aguilar-Saavedra: **Top flavor-changing neutral interactions: Theoretical expectations and experimental detection**. In: *Acta Phys. Polon.*, **B35**: (2004), pp. 2695–2710. arXiv: hep-ph/0409342 [hep-ph] (see pp. 6, 14–15).

[13] S. Abachi et al.: **Observation of the top quark**. In: *Phys. Rev. Lett.*, **74**: (1995), pp. 2632–2637. doi: 10.1103/PhysRevLett.74.2632. arXiv: hep-ex/9503003 [hep-ex] (see p. 6).

[14] 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. arXiv: hep-ex/9503002 [hep-ex] (see p. 6).

[15] Andrea Giammanco and Jeannine Wagner-Kuhr: **Measurement of the t-channel 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. 8).

[16] LHCTop working group: **ATLAS-CMS recommended predictions for single top 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 (see p. 9).

[17] CMS collaboration: **Summaries of CMS cross section measurements**. 2017. url: https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsCombined (see pp. 10, 17–18).

[18] LHCTop working group: **LHCTopWG Summary plots**. 2017. url: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCTopWGSummaryPlots (see pp. 11–12).

[19] B. Grzadkowski, M. Iskrzynski, M. Misiak, and J. Rosiek: **Dimension-Six Terms in the Standard Model Lagrangian**. In: *JHEP*, **10**: (2010), p. 085. doi: 10.1007/JHEP10(2010)085. arXiv: 1008.4884 [hep-ph] (see p. 12).

[20] Andy Buckley, Christoph Englert, James Ferrando, et al.: **Constraining top quark effective theory in the LHC Run II era**. In: *JHEP*, **04**: (2016), p. 015. doi: 10.1007/JHEP04(2016)015. arXiv: 1512.03360 [hep-ph] (see p. 12).

[21] 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. arXiv: hep-ex/9807003 [hep-ex] (see p. 13).

[22] Y. Abe et al.: **Indication of Reactor $\overline{\nu}_e$ Disappearance in the Double Chooz Experiment**. In: *Phys. Rev. Lett.*, **108**: (13 Mar. 2012), p. 131801. doi: 10.1103/PhysRevLett.108.131801 (see p. 13).

[23] P. A. R. Ade et al.: **Planck 2015 results. XIII. Cosmological parameters**. In: *Astron. Astrophys.*, **594**: (2016), A13. doi: 10.1051/0004-6361/201525830. arXiv: 1502.01589 [astro-ph.CO] (see p. 13).

[24] P. J. E. Peebles and Bharat Ratra: **The Cosmological constant and dark energy**. In: *Rev. Mod. Phys.*, **75**: (2003), pp. 559–606. doi: 10.1103/RevModPhys.75.559. arXiv: astro-ph/0207347 [astro-ph] (see p. 13).

[25] A. D. SAKHAROV: **Violation of CP Invariance, c Asymmetry, and Baryon Asymmetry of the Universe**. In: *Pisma Zh. Eksp. Teor. Fiz.*, **5**: (1967). [Usp. Fiz. Nauk161,61(1991)], pp. 32–35. DOI: 10.1070/PU1991v034n05ABEH002497 (see p. 13).

[26] GUSTAVO BURDMAN: **New solutions to the hierarchy problem**. In: *Braz. J. Phys.*, **37**: (2007), pp. 506–513. DOI: 10.1590/S0103-97332007000400006. arXiv: hep-ph/0703194 [hep-ph] (see p. 13).

[27] T. ET AL. AALTONEN: **Search for the Flavor-Changing Neutral-Current Decay $t \to Zq$ in $p\overline{p}$ Collisions at $\sqrt{s} = 1.96$ TeV**. In: *Phys. Rev. Lett.*, **101**: (19 Nov. 2008), p. 192002. DOI: 10.1103/PhysRevLett.101.192002 (see p. 14).

[28] VICTOR MUKHAMEDOVICH ABAZOV et al.: **Search for flavour changing neutral currents via quark-gluon couplings in single top quark production using 2.3 fb^{-1} of $p\bar{p}$ collisions**. In: *Phys. Lett.*, **B693**: (2010), pp. 81–87. DOI: 10.1016/j.physletb.2010.08.011. arXiv: 1006.3575 [hep-ex] (see p. 14).

[29] GEORGES AAD et al.: **Search for flavour-changing neutral current top-quark decays 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. arXiv: 1508.05796 [hep-ex] (see p. 14).

[30] GEORGES AAD et al.: **Search for single top-quark production via flavour-changing neutral currents at 8 TeV with the ATLAS detector**. In: *Eur. Phys. J.*, **C76**:2 (2016), p. 55. DOI: 10.1140/epjc/s10052-016-3876-4. arXiv: 1509.00294 [hep-ex] (see pp. 14, 17).

[31] GEORGES AAD et al.: **Search for flavour-changing neutral current top quark decays $t \to Hq$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector**. In: *JHEP*, **12**: (2015), p. 061. DOI: 10.1007/JHEP12(2015)061. arXiv: 1509.06047 [hep-ex] (see p. 14).

[32] MORAD AABOUD et al.: **Search for top quark decays $t \to qH$, with $H \to \gamma\gamma$, in $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector**. In: (2017). arXiv: 1707.01404 [hep-ex] (see pp. 14, 17).

[33] 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] (see pp. 14, 17).

[34] SERGUEI CHATRCHYAN et al.: **Search for Flavor-Changing Neutral Currents in Top-Quark Decays $t \to Zq$ in pp Collisions at $\sqrt{s} =$ 8TeV**. In: *Phys. Rev. Lett.*, **112**:17 (2014), p. 171802. DOI: 10.1103/PhysRevLett.112.171802. arXiv: 1312.4194 [hep-ex] (see p. 14).

[35] 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. arXiv: 1511.03951 [hep-ex] (see pp. 14, 17).

[36] 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. arXiv: 1610.04857 [hep-ex] (see p. 14).

[37] **Search for the flavor-changing interactions of the top quark with the Higgs boson in $H \to b\bar{b}$ 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. 14).

[38] J. A. Aguilar-Saavedra: **A Minimal set of top anomalous couplings**. In: *Nucl. Phys.*, **B812**: (2009), pp. 181–204. DOI: 10.1016/j.nuclphysb.2008.12.012. arXiv: 0811.3842 [hep-ph] (see p. 14).

[39] J. A. Aguilar-Saavedra: **A Minimal set of top-Higgs anomalous couplings**. In: *Nucl. Phys.*, **B821**: (2009), pp. 215–227. DOI: 10.1016/j.nuclphysb.2009.06.022. arXiv: 0904.2387 [hep-ph] (see p. 14).

[40] 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]. URL: http://weblib.cern.ch/abstract?CERN-TH-2000-100 (see p. 15).

[41] 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. arXiv: 1210.2808 [hep-ph] (see p. 16).

[42] 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 (see p. 19).

[43] 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. 19).

[44] 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. arXiv: hep-ex/0306033 [hep-ex] (see p. 19).

[45] 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 (see p. 19).

[46] Abdelhak Djouadi: **The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model**. In: *Phys. Rept.*, **457**: (2008), pp. 1–216. DOI: 10.1016/j.physrep.2007.10.004. arXiv: hep-ph/0503172 [hep-ph] (see p. 19).

[47] Lyndon Evans and Philip Bryant: **LHC Machine**. In: *Journal of Instrumentation*, **3**:08 (2008), S08001. URL: http://stacks.iop.org/1748-0221/3/i=08/a=S08001 (see p. 19).

[48] Thomas Sven Pettersson and P Lefèvre: **The Large Hadron Collider: conceptual design**. Tech. rep. CERN-AC-95-05-LHC. Oct. 1995, p. 20 and 22. URL: https://cds.cern.ch/record/291782 (see p. 19).

[49] Jorg Wenninger and Ezio Todesco: **Large Hadron Collider momentum calibration and accuracy**. Tech. rep. CERN-ACC-2017-0007. Geneva: *CERN*, Feb. 2017. URL: https://cds.cern.ch/record/2254678 (see p. 20).

[50] Cinzia De Melis: **The CERN accelerator complex. Complexe des accélérateurs du CERN**. In: (July 2016). General Photo. URL: https://cds.cern.ch/record/2197559 (see p. 20).

[51] Oliver Sim Brüning, Paul Collier, P Lebrun, et al.: **LHC Design Report**. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. URL: https://cds.cern.ch/record/782076 (see p. 20).

[52] G. Aad et al.: **The ATLAS Experiment at the CERN Large Hadron Collider**. In: *JINST*, **3**: (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003 (see p. 21).

[53] S. Chatrchyan et al.: **The CMS Experiment at the CERN LHC**. In: *JINST*, **3**: (2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004 (see pp. 21, 28–30).

[54] K. Aamodt et al.: **The ALICE experiment at the CERN LHC**. In: *JINST*, **3**: (2008), S08002. DOI: 10.1088/1748-0221/3/08/S08002 (see p. 21).

[55] A. Augusto Alves Jr. et al.: **The LHCb Detector at the LHC**. In: *JINST*, **3**: (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005 (see p. 21).

[56] M. Bongi et al.: **Astroparticle physics at LHC: The LHCf experiment ready for data taking**. In: *Nucl. Instrum. Meth.*, **A612**: (2010), pp. 451–454. DOI: 10.1016/j.nima.2009.08.039 (see p. 21).

[57] G. Anelli et al.: **The TOTEM experiment at the CERN Large Hadron Collider**. In: *JINST*, **3**: (2008), S08007. DOI: 10.1088/1748-0221/3/08/S08007 (see p. 21).

[58] B. Acharya et al.: **The Physics Programme Of The MoEDAL Experiment At The LHC**. In: *Int. J. Mod. Phys.*, **A29**: (2014), p. 1430050. DOI: 10.1142/S0217751X14300506. arXiv: 1405.7662 [hep-ph] (see p. 21).

[59] by James Gillies: **Luminosity? Why don't we just say collision rate?** In: (Mar. 2011). URL: http://cds.cern.ch/record/1997001 (see p. 21).

[60] by Harriet Jarlett and Harriet Kim Jarlett: **LHC pushes limits of performance**. In: (Aug. 2016). URL: http://cds.cern.ch/record/2212301 (see p. 22).

[61] CMS Collaboration. 2017. URL: https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults#Online_Luminosity_AN2 (see p. 22).

[62] **Technical proposal**. LHC Tech. Proposal. Cover title : CMS, the Compact Muon Solenoid : technical proposal. Geneva: CERN, 1994. URL: https://cds.cern.ch/record/290969 (see p. 22).

[63] G. L. Bayatian et al.: **CMS physics: Technical design report**. In: (2006) (see p. 22).

[64] G L Bayatian, S Chatrchyan, G Hmayakyan, et al.: **CMS Physics: Technical Design Report Volume 1: Detector Performance and Software**. Technical Design Report CMS. There is an error on cover due to a technical problem for some items. Geneva: CERN, 2006. URL: https://cds.cern.ch/record/922757 (see pp. 22, 27, 54).

[65] CMS Collaboration: **Detector Drawings**. CMS Collection. Mar. 2012. URL: https://cds.cern.ch/record/1433717 (see p. 23).

[66] S Chatrchyan, V Khachatryan, A M Sirunyan, et al.: **Performance of the CMS Drift Tube Chambers with Cosmic Rays**. In: *J. Instrum.*, **5**:arXiv:0911.4855. CMS-CFT-09-012 (Nov. 2009), T03015 . 47 p. URL: http://cds.cern.ch/record/1223944 (see pp. 25–26).

[67] Tom Dodington: **News from the CMS experimental site: 22 November 2013.** 2013. URL: http://cms.web.cern.ch/news/news-point-5-22-november-2013 (see p. 26).

[68] Institute of Research into the Fundamental Laws the Universe: **The CMS detector superconducting solenoid**. 2006. URL: http://irfu.cea.fr/en/Phocea/Vie_des_labos/Ast/ast_visu.php?id_ast=839 (see p. 28).

[69] Fergus Wilson: **Experimental Particle Physics**. 2012. URL: http://slideplayer.com/slide/794631/ (see p. 28).

[70] Emrah Tiras, Burak Bilki, and Yasar Onel: **Commissioning of CMS Forward Hadron Calorimeters with Upgraded Multi-anode PMTs and μTCA Readout**. In: (2016). arXiv: 1611.05232 [physics.ins-det] (see p. 29).

[71] Lucas Taylor: **Experimental Particle Physics**. 2011. URL: http://cms.web.cern.ch/news/electromagnetic-calorimeter (see p. 29).

[72] **Proceedings, 34th International Conference on High Energy Physics (ICHEP 2008)** URL: http://www.slac.stanford.edu/econf/C080730 (see p. 29).

[73] L. Brianza: **Precision crystal calorimetry in LHC Run II with the CMS ECAL**. In: *Journal of Instrumentation*, **12**:01 (2017), p. C01069. URL: http://stacks.iop.org/1748-0221/12/i=01/a=C01069 (see pp. 30, 33).

[74] Serguei Chatrchyan, Vardan Khachatryan, Albert M Sirunyan, et al.: **Description and performance of track and primary-vertex reconstruction with the CMS tracker**. In: *J. Instrum.*, **9**:arXiv:1405.6569. CERN-PH-EP-2014-070. CMS-TRK-11-001 (May 2014). Comments: Replaced with published version. Added journal reference and DOI, P10009. 80 p. URL: http://cds.cern.ch/record/1704291 (see pp. 30, 57).

[75] by Christine Sutton: **Chronicles of CMS: the saga of LS1**. In: (May 2015). URL: http://cds.cern.ch/record/2024986 (see p. 31).

[76] **A beautiful barrel for CMS**. In: (Mar. 2014). URL: http://cds.cern.ch/record/1998635 (see p. 31).

[77] Vardan Khachatryan et al.: **The CMS trigger system**. In: *JINST*, **12**:01 (2017), P01020. DOI: 10.1088/1748-0221/12/01/P01020. arXiv: 1609.02366 [physics.ins-det] (see p. 32).

[78] by Corinne Pralavorio and Corinne Pralavorio: **Major work to ready the LHC experiments for Run 2**. In: (May 2015). URL: http://cds.cern.ch/record/2024977 (see p. 32).

[79] Luigi Guiducci: **CMS muon system towards LHC Run 2 and beyond**. Tech. rep. CMS-CR-2014-333. Geneva: *CERN*, Oct. 2014. URL: https://cds.cern.ch/record/1966038 (see p. 33).

[80] Carlo Battilana: **The CMS muon system status and upgrades for LHC run-2 and performance of muon reconstruction with 13 TeV data**. Tech. rep. CMS-CR-2016-437. Geneva: *CERN*, Dec. 2016. URL: http://cds.cern.ch/record/2239185 (see p. 33).

[81] Serguei Chatrchyan et al.: **Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in pp Collisions at $\sqrt{s} = 7$ TeV**. In: *JINST*, **8**: (2013). [JINST8,9009(2013)], P09009. doi: 10.1088/1748-0221/8/09/P09009. arXiv: 1306.2016 [hep-ex] (see p. 33).

[82] **Cool running for CMS tracker**. In: (Mar. 2014). url: http://cds.cern.ch/record/1998606 (see p. 33).

[83] L. Cadamuro: **The CMS Level-1 trigger system for LHC Run II**. In: *Journal of Instrumentation*, **12**:03 (2017), p. C03021. url: http://stacks.iop.org/1748-0221/12/i=03/a=C03021 (see p. 33).

[84] David Aaron Matzner Dominguez, D. Abbaneo, K. Arndt, et al.: **CMS Technical Design Report for the Pixel Detector Upgrade**. In: (2012) (see p. 34).

[85] Hanno Christopher Perrey: **Plans and Status of the Phase I Upgrade of the CMS Pixel Tracker**. Tech. rep. CMS-CR-2014-005. Geneva: *CERN*, Jan. 2014. url: http://cds.cern.ch/record/1644757 (see p. 34).

[86] Claudio Grandi, David Stickland, Lucas Taylor, Achille Petrilli, and Alain Hervé: **CMS Computing Model: The "CMS Computing Model RTAG"**. Tech. rep. CMS-NOTE-2004-031. CERN-LHCC-2004-035. LHCC-G-083. Geneva: *CERN*, Dec. 2004. url: http://cds.cern.ch/record/814248 (see p. 34).

[87] Christoph Eck, J Knobloch, Leslie Robertson, et al.: **LHC computing Grid: Technical Design Report. Version 1.06 (20 Jun 2005)**. Technical Design Report LCG. Geneva: CERN, 2005. url: https://cds.cern.ch/record/840543 (see p. 34).

[88] Worldwide LHC Computing Grid: **WorldWide LHC Computing Gird - 2017**. 2017. url: http://wlcg-public.web.cern.ch (see p. 35).

[89] John C. Collins, Davison E. Soper, and George F. Sterman: **Factorization of Hard Processes in QCD**. In: *Adv. Ser. Direct. High Energy Phys.*, **5**: (1989), pp. 1–91. doi: 10.1142/9789814503266_0001. arXiv: hep-ph/0409313 [hep-ph] (see p. 37).

[90] 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. 37).

[91] 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. 37).

[92] 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. 37–38).

[93] 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. 38).

[94] 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. 38).

[95] 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. 38).

[96] 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. 38).

[97] 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. 38).

[98] 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. 40).

[99] 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. 40).

[100] 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. 40).

[101] 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 (see pp. 40, 42).

[102] 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. 40).

[103] 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. arXiv: 0911.5286 [hep-ph] (see p. 40).

[104] 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. arXiv: 1411.4085 [hep-ph] (see pp. 40–41).

[105] 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. arXiv: 1310.1921 [hep-ph] (see p. 40).

[106] 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. arXiv: 1108.2040 [hep-ph] (see p. 40).

[107] 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. arXiv: 1106.0522 [hep-ph] (see pp. 41, 44).

[108] Michelangelo L. Mangano, Mauro Moretti, Fulvio Piccinini, and Michele Treccani: **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. arXiv: hep-ph/0611129 [hep-ph] (see p. 41).

[109] 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. 41–42).

[110] 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. arXiv: hep-ph/0603175 [hep-ph] (see pp. 41–42).

[111] 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. arXiv: 1410.3012 [hep-ph] (see pp. 41–42, 44).

[112] 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. arXiv: 0706.2569 [hep-ph] (see p. 41).

[113] 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. arXiv: 1405.0301 [hep-ph] (see pp. 41, 44).

[114] Rikkert Frederix and Stefano Frixione: **Merging meets matching in MC@NLO**. In: *JHEP*, **12**: (2012), p. 061. doi: 10.1007/JHEP12(2012)061. arXiv: 1209.6215 [hep-ph] (see p. 41).

[115] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re: **A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX**. In: *Journal of High Energy Physics*, **2010**:6 (2010), p. 43. doi: 10.1007/JHEP06(2010)043 (see p. 41).

[116] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re: **NLO single-top production matched with shower in POWHEG: s - and t -channel contributions**. In: *Journal of High Energy Physics*, **2009**:09 (2009), p. 111. url: http://stacks.iop.org/1126-6708/2009/i=09/a=111 (see p. 41).

[117] Stefano Frixione, Paolo Nason, and Carlo Oleari: **Matching NLO QCD computations with parton shower simulations: the POWHEG method**. In: *Journal of High Energy Physics*, **2007**:11 (2007), p. 070. url: http://stacks.iop.org/1126-6708/2007/i=11/a=070 (see p. 41).

[118] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re: **A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX**. In: *JHEP*, **06**: (2010), p. 043. doi: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph] (see p. 41).

[119] Stefano Frixione, Paolo Nason, and Carlo Oleari: **Matching NLO QCD computations with Parton Shower simulations: the POWHEG method**. In: *JHEP*, **11**: (2007), p. 070. doi: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph] (see p. 41).

[120] Paolo Nason: **A New method for combining NLO QCD with shower Monte Carlo algorithms**. In: *JHEP*, **11**: (2004), p. 040. doi: 10.1088/1126-6708/2004/11/040. arXiv: hep-ph/0409146 [hep-ph] (see p. 41).

[121] Andrei V. Gritsan, Raoul Röntsch, Markus Schulze, and Meng Xiao: **Constraining anomalous Higgs boson couplings to the heavy flavor fermions using matrix element techniques**. In: *Phys. Rev.*, **D94**:5 (2016), p. 055023. doi: 10.1103/PhysRevD.94.055023. arXiv: 1606.03107 [hep-ph] (see pp. 41, 44).

[122] Ian Anderson et al.: **Constraining anomalous HVV interactions at proton and lepton colliders**. In: *Phys. Rev.*, **D89**:3 (2014), p. 035007. doi: 10.1103/PhysRevD.89.035007. arXiv: 1309.4819 [hep-ph] (see pp. 41, 44).

[123] Sara Bolognesi, Yanyan Gao, Andrei V. Gritsan, et al.: **On the spin and parity of a single-produced resonance at the LHC**. In: *Phys. Rev.*, **D86**: (2012), p. 095031. doi: 10.1103/PhysRevD.86.095031. arXiv: 1208.4018 [hep-ph] (see pp. 41, 44).

[124] Yanyan Gao, Andrei V. Gritsan, Zijin Guo, et al.: **Spin determination of single-produced resonances at hadron colliders**. In: *Phys. Rev.*, **D81**: (2010), p. 075022. doi: 10.1103/PhysRevD.81.075022. arXiv: 1001.3396 [hep-ph] (see pp. 41, 44).

[125] Pierre Artoisenet, Rikkert Frederix, Olivier Mattelaer, and Robbert Rietkerk: **Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations**. In: *JHEP*, **03**: (2013), p. 015. doi: 10.1007/JHEP03(2013)015. arXiv: 1212.3460 [hep-ph] (see pp. 42, 44).

[126] V. Khachatryan and etal: **Event generator tunes obtained from underlying event and multiparton scattering measurements**. In: *The European Physical Journal C*, **76**:3 (Mar. 17, 2016), p. 155. doi: 10.1140/epjc/s10052-016-3988-x (see p. 42).

[127] J. Pumplin, D. R. Stump, J. Huston, et al.: **New generation of parton distributions with uncertainties from global QCD analysis**. In: *JHEP*, **07**: (2002), p. 012. doi: 10.1088/1126-6708/2002/07/012. arXiv: hep-ph/0201195 [hep-ph] (see p. 42).

[128] 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. arXiv: 1101.5346 [hep-ph] (see p. 43).

[129] A. Hoecker, P. Speckmayer, J. Stelzer, et al.: **TMVA - Toolkit for Multivariate Data Analysis**. In: *ArXiv Physics e-prints*, (Mar. 2007). eprint: physics/0703039 (see pp. 46–47).

[130] 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 (see p. 46).

[131] A. Mayr, H. Binder, O. Gefeller, 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] (see p. 47).

[132] 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 pp. 47–48).

[133] 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. 48).

[134] Serguei Chatrchyan 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. arXiv: 1202.1488 [hep-ex] (see p. 48).

[135] 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, 10.1140/epjc/s10052-013-2501-z. arXiv: 1007.1727 [physics.data-an] (see pp. 48, 52).

[136] **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. 48, 52).

[137] **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. 48, 50–52).

[138] 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. 48).

[139] Lorenzo Moneta, Kevin Belasco, Kyle S. Cranmer, et al.: **The RooStats Project**. In: *PoS*, **ACAT2010**: (2010), p. 057. arXiv: 1009.1003 [physics.data-an] (see p. 48).

[140] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells: **Asymptotic formulae for likelihood-based tests of new physics**. In: *The European Physical Journal C*, **71**:2 (Feb. 2011). DOI: 10.1140/epjc/s10052-011-1554-0. arXiv: 1007.1727 [hep-ph] (see p. 48).

[141] Thomas Junk: **Confidence level computation for combining searches with small statistics**. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **434**:2 (1999), pp. 435–443. DOI: https://doi.org/10.1016/S0168-9002(99)00498-2 (see p. 49).

[142] 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. 49).

[143] 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. arXiv: 1103.0354 [physics.data-an] (see p. 51).

[144] 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. arXiv: 1706.04965 [physics.ins-det] (see pp. 53–54, 57, 60).

[145] 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 (see p. 55).

[146] 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 (see p. 55).

[147] Serguei Chatrchyan 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. arXiv: 1206.4071 [physics.ins-det] (see p. 55).

[148] 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 (see p. 57).

[149] Wolfgang Waltenberger: **Adaptive Vertex Reconstruction**. Tech. rep. CMS-NOTE-2008-033. Geneva: *CERN*, July 2008. URL: https://cds.cern.ch/record/1166320 (see p. 57).

[150] **CMS Luminosity Measurements for the 2016 Data Taking Period**. Tech. rep. CMS-PAS-LUM-17-001. Geneva: *CERN*, 2017. URL: http://cds.cern.ch/record/2257069 (see pp. 60, 71).

[151] **Pileup Removal Algorithms**. Tech. rep. CMS-PAS-JME-14-001. Geneva: *CERN*, 2014. URL: http://cds.cern.ch/record/1751454 (see p. 61).

[152] **Muon Identification and Isolation efficiency on full 2016 dataset**. In: (Mar. 2017). URL: http://cds.cern.ch/record/2257968 (see pp. 62–63).

[153] **Electron and photon performance in CMS with the full 2016 data sample.** In: (Mar. 2017). URL: http://cds.cern.ch/record/2255497 (see pp. 64–65).

[154] **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. 64).

[155] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez: **The Anti-k(t) jet clustering algorithm**. In: *JHEP*, **04**: (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph] (see p. 66).

[156] **Jet algorithms performance in 13 TeV data**. Tech. rep. CMS-PAS-JME-16-003. Geneva: *CERN*, 2017. URL: http://cds.cern.ch/record/2256875 (see p. 66).

[157] V. Khachatryan et al.: **Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV**. In: *Journal of Instrumentation*, **12**:02 (2017), P02014. URL: http://stacks.iop.org/1748-0221/12/i=02/a=P02014 (see pp. 66–67).

[158] **Jet energy scale and resolution performances with 13TeV data**. In: (June 2016). URL: http://cds.cern.ch/record/2160347 (see pp. 67–68).

[159] The CMS collaboration: **Identification of b-quark jets with the CMS experiment**. In: *Journal of Instrumentation*, **8**:04 (2013), P04013. URL: http://stacks.iop.org/1748-0221/8/i=04/a=P04013 (see p. 68).

[160] **Identification of b quark jets at the CMS Experiment in the LHC Run 2**. Tech. rep. CMS-PAS-BTV-15-001. Geneva: *CERN*, 2016. URL: http://cds.cern.ch/record/2138504 (see pp. 68–70).

[161] Nazar Bartosik: **Diagram showing the common principle of identification of jets initiated by b-hadron decays**. 2016. URL: http://bartosik.pp.ua/hep_sketches/btagging (see p. 69).

[162] **Identification of b quark jets at the CMS Experiment in the LHC Run 2**. Tech. rep. CMS-PAS-BTV-16-001. unpublished. Geneva: *CERN*, 2017 (see p. 70).

[163] **Performance of missing energy reconstruction in 13 TeV pp collision data using the CMS detector**. Tech. rep. CMS-PAS-JME-16-004. Geneva: *CERN*, 2016. URL: http://cds.cern.ch/record/2205284 (see p. 71).

[164] Andreas Kornmayer: **The CMS Pixel Luminosity Telescope**. Tech. rep. CMS-CR-2015-121. Geneva: *CERN*, June 2015. URL: https://cds.cern.ch/record/2039978 (see p. 71).